

(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2006-236149

(P2006-236149A)

(43) 公開日 平成18年9月7日(2006.9.7)

(51) Int. Cl.

G06F 13/00 (2006.01)

F I

G06F 13/00 650B

テーマコード (参考)

審査請求 未請求 請求項の数 7 O L (全 15 頁)

(21) 出願番号 特願2005-52041 (P2005-52041)
 (22) 出願日 平成17年2月25日 (2005.2.25)

(71) 出願人 304024430
 国立大学法人北陸先端科学技術大学院大学
 石川県能美市旭台一丁目1番地
 (74) 代理人 100105809
 弁理士 木森 有平
 (74) 代理人 100126398
 弁理士 浅野 典子
 (72) 発明者 西本 一志
 石川県能美市旭台1-1 国立大学法人北
 陸先端科学技術大学院大学内
 (72) 発明者 小倉 加奈代
 石川県能美市旭台1-1 国立大学法人北
 陸先端科学技術大学院大学内

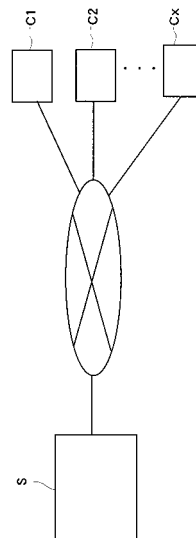
(54) 【発明の名称】 コミュニケーション装置

(57) 【要約】

【課題】 音声によるコミュニケーションを支援するものであり、音声データ間の相互関係を把握可能とすることによりマルチスレッド対話が実現可能なコミュニケーション装置を提供する。

【解決手段】 ユーザのクライアント端末に対して受信した音声データを提供することにより音声によるコミュニケーションを支援するコミュニケーション装置において、受信した音声データに関するログデータに音声データ間の関連を示す情報を付加する手段を備える。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

ユーザのクライアント端末に対して受信した音声データを提供することにより音声によるコミュニケーションを支援するコミュニケーション装置において、受信した音声データに関するログデータに音声データ間の関連を示す情報を付加する手段を備えることを特徴とするコミュニケーション装置。

【請求項 2】

前記音声データ間の関連を示す情報には、発言相手を特定する情報が含まれていることを特徴とする請求項 1 記載のコミュニケーション装置。

【請求項 3】

前記音声データ間の関連を示す情報には、先行する音声データを特定する情報が含まれていることを特徴とする請求項 1 又は請求項 2 に記載のコミュニケーション装置。

10

【請求項 4】

前記音声データ間の関連を示す情報を付加する手段は、発言相手が指定されると、その発言相手を特定する情報をログデータに付加することを特徴とする請求項 1 記載のコミュニケーション装置。

【請求項 5】

前記音声データ間の関連を示す情報を付加する手段は、先行する音声データの一覧から音声データが選択されると、その選択された音声データを特定する情報をログデータに付加することを特徴とする請求項 1 記載のコミュニケーション装置。

20

【請求項 6】

信号を発信する発信手段と、当該発信手段から発信された信号を受信する受信手段とを備え、

発言者のクライアント端末から音声データを受信し、当該発信手段からの信号を受信した受信手段に接続される発言相手のクライアント端末から当該発言相手の識別子を受信し、当該音声データのログデータに当該発言相手の識別子を付与することを特徴とする請求項 1 記載乃至請求項 5 のいずれか 1 項に記載のコミュニケーション装置。

【請求項 7】

テキストデータが入力可能であり、入力されたテキストデータは、前記ログデータに付加されることを特徴とする請求項 1 乃至請求項 6 のいずれか 1 項に記載のコミュニケーション装置。

30

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、複数ユーザのクライアント端末に対して受信した音声データを提供することにより音声によるコミュニケーションを支援するコミュニケーション装置に関する。

【背景技術】**【0002】**

インターネット環境が普及した現在、電子メールや電子掲示板、チャットのような非対面型のテキストをベースとしたコミュニケーションが日常的に利用されている。中でもテキストチャットは逐次やりとりをするという点で対面対話に通じる点がある。また一方で、「マルチスレッド対話」が容易に可能であるという対面対話にはない利点を持っている。ここで、マルチスレッド対話とは、会話空間で同時並行的に複数の話題についての対話が進行し、しかもある一人の参加者が同時に複数の話題に参加しているような対話である。対面対話でも、多人数が集まれば、同じ場で複数話題が同時に展開されることはあるが、これは話題に基づく集団の単なる分割にすぎないため、マルチスレッド対話とは見なさない。

40

【0003】

日常の対面対話では、一定時間は、対話参加者全員が単一の話話を共有し、しかもその話題についての発話を行う必要があり、話者は一人だけであることが求められる（同期性

50

の制約)。そのため、対面対話は非効率的に進行していると言える。また、この結果、今交わされている話題とは別のことを思いついても、すぐに発話することができず、思いついた内容そのものを忘れてしまうことがしばしば起こる。

【0004】

また、対話参加者の位置関係の影響の有無という点でもテキストチャットと対面対話では大きな違いがある。対面対話の場合、特に会議や会食のような多人数で1つの場を共有する状況では、最初についた座席の位置関係によって会話を交わしやすい人、そうでない人の区分が決定する(近接性の制約)。自分の座席から離れている人と話をするためには、座席を移動するか、大声で相手に呼びかける必要がある。前者の場合は、会話そのものを中断させ、後者の場合は、呼びかけた相手の会話を遮ることになり、どちらの場合も、会話進行の大きな妨げになる。

10

【0005】

これに対して、テキストチャット対話では、このような制約が無い場合、思いついたことや過去の任意の発話に関連する内容をいつでも発言順序を気にせず発言可能となる。つまり、マルチスレッド対話を行うことが可能となるため、対話が効率化される。また、テキストチャット対話では、発言履歴が会話空間であり、発言履歴に表示される参加者個々のログイン名が参加者の分身であると言える。ここでは単に誰が参加しているのかという存在情報のみが意味を持ち、参加者の位置関係という概念そのものが存在しない。そのため、対面対話とは異なり参加者の位置関係を気にせずに発言することが可能である。ゆえに、たとえば企画会議などにおいてアイデアをもれなく収集するような際に非常に有効である。

20

【0006】

しかし、テキスト対話では文字入力のわずらわしさのために、思うような発言を十分に行うことが難しい。このため、音声対話によりコミュニケーションを支援する装置、特に、音声対話によるマルチスレッド対話を実現可能な装置が求められる。

【0007】

音声データによりコミュニケーションを支援する装置として、下記非特許文献1には、空間的位置関係にかかわらず、音声認識を用いて同一スレッドに属する話者の発話音声により明確に聞こえる等の自動的音響効果を付加した同時的会話環境である“The Mad Hatter’s Cocktail Party”が開示されている。このシステムでは、音声認識を用いた同一スレッドの自動判定を行っているため、認識精度の問題が生じてしまい確実性が低くなる。

30

【0008】

【非特許文献1】P.M.Aoki, M. Romaine, M.H. Szymanski, J.D. Thornton, D. Wilson, and A. Woodruff: The Mad Hatter's Cocktail Party: A Social Mobile Audio Space Supporting Multiple Conversations, Proc. ACM SIGCHI Conf. on Human Factors in Computing Systems, pp.425-432, 2003.

【発明の開示】

【発明が解決しようとする課題】

【0009】

ところが、音声を用いた対話(対面対話、電話、音声チャット)では、今のところマルチスレッド対話は実現不可能ないし非常に困難である。その最大の理由は、一般に人は複数の発言を同時に聞いてそれらを記憶し、理解することができないことにある。

40

【0010】

対面でマルチスレッド対話を行なうためには、我々は、一度に複数の発言を聞き分け、記憶し、理解することが必要とされる。しかし、これは人間の認知能力の範囲内では困難である。したがって、マルチスレッド対話が行なわれることはめったにない。一方テキストチャットでは、誰が、いつ、何を発言したかという発言にかかわるデータが保存され、それが発言履歴として表示され、参加者は自由に履歴を閲覧できる。これが人間の短期記憶を補う役割を果たしているために同期性の制約が解消され、また発言履歴が対話の場として機能するために近接性の制約も解消される。この結果、テキストチャットではマルチ

50

スレッド対話が可能となっていると思われる。つまりマルチスレッド対話の実現には、発言履歴が不可欠であると考えられる。

【0011】

音声チャットシステムでは、個々の発言が音声データとして計算機上に記録され、いつ誰が発言したかを示すログリストも表示されるために、何度でも必要に応じて聞きなおすことができる。つまり音声データやログの保存によって人の短期記憶が補助されているため、対面対話や電話よりはマルチスレッド対話を実現しやすい音声コミュニケーション・メディアとなっている。

【0012】

しかし、ログ上では個々の発言（音声データ）の関連性が見えず、音を再生して聞いてみないとどの発言（音声データ）とどの発言（音声データ）とがつながっているのかわからない。このため、既存の音声チャットシステムでは依然としてマルチスレッド対話を実現することが難しい。

10

【0013】

本発明は、このような実情に鑑みてなされたものであり、音声によるコミュニケーションを支援するものであり、音声データ間の相互関係を把握可能とすることによりマルチスレッド対話が可能となるコミュニケーション装置を提供することを目的とする。

【課題を解決するための手段】

【0014】

発明者等は、従来のテキストチャットシステムを用いた場合に、マルチスレッド状況を維持するために、どのような手段がとられているか、その傾向を調査した。その結果を下記に示す。

20

【0015】

チャット対話収録で得たデータについて予備分析を行なった結果、発言履歴の参照しにくさがある中で話の流れを追い、できる限りスムーズに対話を進めるために貢献していると考えられる3種類の表現が存在することがわかった。

1) 誰に向けた発言であるかを明記する表現（固有名詞を含む）（図6では“>人”）

例）まだまだ今年はこれからですよ。> Bさん

2) どの話に関連した発言をしているかを明記する表現（図6では“>単語”）

例）意見わかれるとこみたいです> マロンクリーム

30

3) どの発言に対して発言をしているかを明記する表現（この場合はコピー&ペーストを行っていると推測される場合）（図6では“>コピペ”）

例）A：栗と生クリームって合わないと思いませんか？> A L L > とても合うと思うよ。マロンクリームってめっちゃうまいやん！

（「栗と生クリームって合わないと思いませんか？> A L L」がコピー&ペーストされた部分。）

【0016】

これら3つの表現が対話中に出現した割合のグラフを図6に示す。さらに、発言間の意味的つながりの判定作業を行なった上で、3つの表現を発言間のインターバル（発言間距離）ごとに分類した出現頻度の結果を示すグラフを図7に示す。なお、分析対象としたデータは、2人対話が3対話311発言分、3人対話が5対話559発言分、計870発言分の発言履歴データであり、全てチャット経験者を被験者としたデータである。

40

【0017】

図6を見ると、3つの表現をあわせると分析対象発言の約1/4の割合で、話をスムーズに進めるためにいずれかの表現を用いていることがわかる。また、図7を見ると、隣接する発言同士が異なる話題である場合（発言間距離2以上の場合）にどの表現も出現頻度が増加している。このことから、より複雑な状況になればなるほど、話の流れを追いやすくするために、誰に対する発言なのか、どの話、どの発言に関する発言なのかなど、発言間の関連を明記して対応していると推測できる。

【0018】

50

以上の結果から、発明者等は、本発明のコミュニケーション装置を開発した。本発明のコミュニケーション装置は、ユーザのクライアント端末に対して受信した音声データを提供することにより音声によるコミュニケーションを支援するコミュニケーション装置において、受信した音声データに関するログデータに音声データ間の関連を示す情報を付加する手段を備えることを特徴とする。この本発明によれば、ログデータを参照することにより、音声データの相互関係を把握することができる。

【0019】

前記音声データ間の関連を示す情報には、発言相手を特定する情報が含まれていることを特徴とする。この発明によれば、ログデータを参照することにより、音声データごとに発言者や発言相手を知ることができ、音声データの相互関係を把握することができる。

10

【0020】

前記音声データ間の関連を示す情報には、先行する音声データを特定する情報が含まれていることを特徴とする。この発明によれば、ログデータを参照することにより、その音声データがいずれの音声データに対する発言であるかを把握することができる。

【0021】

前記音声データ間の関連を示す情報を付加する手段は、発言相手が指定されると、その発言相手を特定する情報をログデータに付加することを特徴とする。この発明によれば、発言相手となるユーザを指定するだけで、ログデータに発言相手の情報を付与することができる。

【0022】

前記音声データ間の関連を示す情報を付加する手段は、先行する音声データの一覧から音声データが選択されると、その選択された音声データを特定する情報をログデータに付加することを特徴とする。この発明によれば、先行する音声データの一覧から音声データを選択するだけで、そのログデータに対応する音声データが先行するどの音声データを受けてなされたもののかの情報を付加することができる。

20

【0023】

信号を発信する発信手段と、当該発信手段から発信された信号を受信する受信手段とを備え、発言者のクライアント端末から音声データを受信し、当該発信手段からの信号を受信した受信手段に接続される発言相手のクライアント端末から当該発言相手の識別子を受信し、当該音声データのログデータに当該発言相手の識別子を付与することを特徴とする。この発明によれば、発言者がクライアント端末から音声データをコミュニケーション装置に送信し、自己の発信手段を用いて発言相手の受信手段に信号を送信すると、その信号を受け取った受信手段に接続される発言相手のクライアント端末から発言相手の識別子がコミュニケーション装置に送信される。コミュニケーション装置は、発言者のクライアント端末から受信した音声データのログデータに、発言相手のクライアント端末から受信した発言相手の識別子を付加する。発言者は、発言相手を指し示すように、送信手段を用いて発言相手の受信手段に信号を送る動作をするだけで、音声データのログデータに発言相手の情報を付加することができる。

30

【0024】

詳しくは、信号を発信する発信手段と、当該発信手段から発信された信号を受信する受信手段とを備え、発言者のクライアント端末から音声データと当該発言者の識別子を受信し、発信手段から発信された識別子を受信した受信手段が接続される発言相手のクライアント端末から、当該受信した識別子と発言相手の識別子とを受信すると、当該発言者の識別子と受信した識別子とが一致する場合は、当該受信した音声データに関するログデータに当該発言相手の識別子を付加することを特徴とする。

40

【0025】

また、本発明のコミュニケーション装置は、テキストデータが入力可能であり、入力されたテキストデータは、前記ログデータに付加されることを特徴とする。この発明によれば、音声データのみならず、テキストデータによってもコミュニケーションが図られる。たとえば、具体的かつ詳細な内容は音声で入力し、その内容のタイトルやキーワードなど

50

だけを文字として入力するようなことが可能となり、個々の音声データ間の関係を意味的に明示することが可能となる。また、音声では入力できない「顔文字」のようなものも音声に付加して入力可能となる。

【発明の効果】

【0026】

本発明のコミュニケーション装置によれば、コミュニケーションに参加するユーザは、ログデータを参照することにより、音声データの相互関係を把握することができるため、音声によるマルチスレッド型の対話が可能となる。とくに、音声データ間の関連を示す情報として、発言者や発言相手を特定する情報が含まれている場合は、個々の音声データごとに、誰の発言であるか、誰に対しての発言であるか、を把握することができる。また、音声データ間の関連を示す情報として、先行する音声データを特定する情報が含まれる場合は、個々の音声データが、先行するいずれの音声データを受けた発言であるかを把握することができる。また、これらの両方が含まれる場合は、音声データ間の相互関係を更に詳しく把握することができる。

10

【0027】

ユーザは、ユーザー一覧から発言相手を選択するだけで、その音声データの発言相手を指定でき、また、先行する音声データの一覧から音声データを選択するだけで、その音声データがいずれの音声データを受けて発言されたもののかの情報を付加することができるため、音声データ間の関連付けを簡単に行うことができ、マルチスレッド型の対話が滞りなく円滑に進行される。

20

【0028】

さらに、信号を発信する発信手段と、その信号を受信する受信手段とを用いることにより、発言者が発言相手を指し示すような感覚で、音声データに関するログデータに発言相手を特定する情報を付加することができる。このシステムを対面型の対話に導入すれば、マルチスレッド型の対話が可能となるだけでなく、全く新しくしかも効率的な対話スタイルを確立することが可能となる。これは会議などにおいて非常に有効であると期待できる。また、発信手段を発言相手側に向ける動作により発言相手を見ることとなり、ログデータを見続けるような状況が回避され、対話がより自然に進行する。

【0029】

さらに、ログデータに、入力されたテキストデータが付加されると、ログデータを参照することによって、音声データ間の意味的なつながりを把握することが可能となる。これによってマルチスレッド対話をより効率的に行うことができる。また、従来のテキスト・音声混合チャットシステムでは、音声の再生手段を持たない参加者は音声データの内容を全く理解できなかった。本システムによれば、付加されたタイトルやキーワードなどのテキストデータで、おおまかながら音声データの内容を理解可能となる。

30

【発明を実施するための最良の形態】

【0030】

(第1の実施の形態)

以下、本発明に係るコミュニケーション装置Sについて図面を参照しながら説明する。図1は、本実施の形態のコミュニケーション装置Sを説明する説明図である。コミュニケーション装置Sは、コンピュータシステムであるサーバにより実現され、インターネットなどのネットワークを介して同じくコンピュータシステムである複数のクライアント端末C1, , , Cxと接続可能となっている。

40

【0031】

コミュニケーション装置Sは、例えば、音声チャットシステムのように、ユーザのクライアント端末C1, , , Cxから音声データを受信し、その受信した音声データを提供することにより、音声によるコミュニケーションを支援するシステムである。

【0032】

クライアント端末C1, , , Cxは、音声によるコミュニケーション参加のために、ネットワークを介してコミュニケーション装置Sにアクセス可能となっている。クライアン

50

ト端末 C 1 , , , C x は、音声を録音する録音機能を備え、録音された音声は、音声データとしてネットワークを介してコミュニケーション装置 S にアップロード可能となっている。

【 0 0 3 3 】

コミュニケーション装置 S は、クライアント端末 C 1 , , , C x からの要求に応じて、各クライアント端末 C 1 , , , C x にユーザインタフェース画面を表示する。図 2 は、コミュニケーション装置 S がクライアント端末 C 1 , , , C x に提供するユーザインタフェース画面を示す図である。

【 0 0 3 4 】

クライアント端末 C 1 , , , C x において、参加者の識別子であるハンドル名を入力する欄 1 にハンドル名が入力され、ログインのボタン 2 が押下されると、コミュニケーション装置 S は、その信号を受信し、ユーザインタフェース画面のユーザー一覧の欄 3 にログイン状態にあるユーザの識別子（ここではハンドル名）と、履歴一覧の欄 4 に蓄積記憶されているログデータを表示する。

10

【 0 0 3 5 】

コミュニケーション装置 S は、ユーザインタフェース画面と同時に提供するプログラムにより、クライアント端末 C 1 , , , C x に下記の機能を実現させる。

【 0 0 3 6 】

（録音機能）

コミュニケーション装置 S は、クライアント端末 C 1 , , , C x に、音声を録音して音声データを生成させ、その音声データをコミュニケーション装置 S にアップロードさせる録音機能を実現させる。実現させる録音機能としては三種類有り、（ 1 ）通常録音機能、（ 2 ）先行音声データ指定録音機能、（ 3 ）発言相手指定録音機能、を備える。

20

【 0 0 3 7 】

（通常録音機能）

クライアント端末 C 1 , , , C x は、ユーザインタフェース画面の録音関連ボタン群 5 のうち「録音」ボタン 6 の押下を検知すると、図 3 に示されるような「発言完了」ボタンを、ユーザインタフェース画面をすべて覆い隠す形で画面上に表示し、録音を開始する。クライアント端末 C 1 , , , C x は「発言完了」ボタンの押下を検知すると録音を終了し、生成された音声データ、及び、その音声データの発言者の識別子（ここではハンドル名）をコミュニケーション装置 S にアップロードする。ここで、発言者とは、音声データの元となる発言をした者であり、音声データを生成してアップロードする者のことである。

30

【 0 0 3 8 】

図 4 は、ユーザインタフェース画面の履歴一覧の欄 4 のログデータを抽出して示す図である。コミュニケーション装置 S は、音声データを特定する識別子 a 1 , 発言者を特定する識別子（ハンドル名） a 2 , 発言時刻となるアップロード時刻 a 3 を、受信した音声データに関する基本ログデータとし、受信した音声データと基本ログデータとを関連付けて蓄積記憶する。つぎに、コミュニケーション装置 S は、各クライアント端末 C 1 , , , C x に対してその基本ログデータを提供し、再表示させたユーザインタフェース画面の履歴一覧の欄 4 に、生成された基本ログデータを追加表示させる。

40

【 0 0 3 9 】

（先行音声データ指定録音機能 1）

また、クライアント端末 C 1 , , , C x は、生成した音声データと、先行する音声データとを対応付けてアップロードする機能を備える。クライアント端末 C 1 , , , C x は、上記通常録音機能により音声の録音を完了した後に、ユーザインタフェース画面の履歴一覧の欄 4 から、先行する音声データが選択されるとともに、「先行発言指定」ボタン 7 が押下されると、選択された先行する音声データを特定する識別子 a 1 を、音声データや発言者の識別子（ハンドル名）とともにコミュニケーション装置 S にアップロードする。なお、履歴一覧は音声データの履歴を示すものであり、先行する音声データの一覧の役割も同時に果たす。

50

【 0 0 4 0 】

(先行音声データ指定録音機能 2)

また、別の先行音声データ指定録音機能として、履歴一覧の欄 4 から先行する音声データが選択された後に、「発言への返信録音」ボタン 8 が押下されると、「発言完了」ボタンを表示し、発言の録音を開始する。「発言完了」ボタンが押下されると録音を終了し、生成された音声データと、発言者の識別子(ハンドル名)と、選択された音声データを特定する識別子 1 a とを、コミュニケーション装置 S にアップロードする。

【 0 0 4 1 】

コミュニケーション装置 S は、上記二通りのいずれかの先行音声データ指定録音機能により生成された音声データ、発言者の識別子(ハンドル名)、及び、選択された先行する音声データを特定する識別子 a 1 を受信すると、生成された音声データと基本ログデータに、選択された先行する音声データを特定する識別子 a 1 を付加し、互いに関連付けて蓄積記憶する。その後、各クライアント端末 C 1 , , , C x に対して、基本ログデータ、及び、選択された先行する音声データの識別子 a 1 を提供する。クライアント端末 C 1 , , , C x の履歴一覧には、基本ログデータが表示されるとともに、基本ログデータの文末に、例えば「>> [2]」のように、先行する音声データを特定する識別子 a 4 が付与される。

10

【 0 0 4 2 】

(発言相手指定録音機能 1)

また、クライアント端末 C 1 , , , C x は、発言相手を指定して録音を行う発言相手指定録音機能を備える。この機能としては 3 パターン存在する。第 1 のパターンとして、上記通常録音機能で発言を録音し、ユーザー一覧の欄 3 から発言相手が選択されるとともに「発言相手指定」ボタン 9 の押下を検知すると、確認のために「発言相手指定」ボタン 9 を選択した発言相手の識別子(ハンドル名)に変更する。その変更されたボタン 9 の押下を検知すると、生成された音声データと発言相手を特定する識別子(ハンドル名)とをコミュニケーション装置にアップロードする。例えば、通常録音完了後に、ユーザー一覧の欄 3 から「Bob」が選択されたと仮定する。そうすると、「発言相手指定」ボタン 9 を「>Bob」という表示のボタンに変更し、そのボタン 9 の押下を検知すると、「Bob」を指定した発言が完了することになる。

20

【 0 0 4 3 】

(発言相手指定録音機能 2)

また、第 2 のパターンとして、クライアント端末 C 1 , , , C x は、ユーザー一覧の欄 3 から発言相手となるユーザが選択された後に、「相手指定録音」ボタン 10 の押下を検知すると、「発言完了」ボタンを表示画面に表示し、録音を開始する。「発言完了」ボタンの押下を検知すると、録音を終了し、生成された音声データと、発言者の識別子(ハンドル名)と、選択された発言相手の識別子(ハンドル名)をコミュニケーション装置 S にアップロードする。

30

【 0 0 4 4 】

(発言相手指定録音機能 3)

また、第 3 のパターンでは、履歴一覧から選んだ音声データの発言者を発言相手として指定する。クライアント端末 C 1 , , , C x は、履歴一覧の欄 4 から音声データが選択されるとともに、「発言者への返答録音」ボタン 11 の押下を検知すると、「発言完了」ボタンを表示画面に表示し、録音を開始する。「発言完了」ボタンの押下を検知すると、録音を終了し、生成された音声データと、発言者の識別子(ハンドル名)と、選択された音声データの発言者の識別子(ハンドル名)とをコミュニケーション装置 S にアップロードする。選択された音声データの発言者の識別子は、発言相手の識別子として取り扱われる。

40

【 0 0 4 5 】

コミュニケーション装置 S は、上記機能により受信した音声データと、基本ログデータと、選択された発言相手の識別子(発言相手指定録音機能 3 においては「選択された音声

50

データの発言者の識別子」とを蓄積記憶し、各クライアント端末C1, , , Cxに対してこれらの情報を提供する。クライアント端末C1, , , Cxの履歴一覧の欄4には、基本ログデータが表示されるとともに、発言相手として選択したユーザの識別子(ハンドル名)a5が付加される。たとえば、クライアント端末C1, , , Cxに表示される履歴一覧の欄4には、基本ログデータの文末に「>Bob」のよう発言相手の識別子(ハンドル名)が付与される。なお、自己が他のユーザから発言相手として指定を受けた場合には、行頭に「>>>You:」と表示される。

【0046】

また、以上のそれぞれの機能は組み合わせても動作可能である。たとえば、基本ログデータに、「>Bob>>[4]」のように、発言相手の指定と先行音声データの指定の両方を付与する機能も備える。さらに、発言相手指定録音機能は、「>Susie>Andy」のように複数の発言相手の指定を行なう機能も備える。

10

【0047】

(再生機能)

コミュニケーション装置Sは、クライアント端末C1, , , Cxに、受信した音声データを再生する再生機能を実現させる。クライアント端末C1, , , Cxは、履歴一覧の欄4からログデータが選択されるとともに、「これを聞く」ボタン12が押下されると、その選択されたログデータに対応する音声データをコミュニケーション装置Sに要求する。コミュニケーション装置Sは、クライアント端末C1, , , Cxから、音声データの送信要求を受け取ると、要求を受けた音声データをクライアント端末C1, , , Cxに提供する。クライアント端末C1, , , Cxは、受信した音声データを再生する。

20

【0048】

また、クライアント端末C1, , , C3xは、履歴一覧の欄4でのダブルクリックを検知した場合にも、そのダブルクリックされたログデータに対応する音声データをコミュニケーション装置Sに要求する。コミュニケーション装置Sは、要求された音声データをクライアント端末C1, , , Cxに提供し、それを受信したクライアント端末C1, , , Cxは音声データを再生する。これによれば、ダブルクリックのみで音声データを再生できるため操作性が良い。

【0049】

また、再生機能の付加機能として、「次を聞く」ボタン13の押下を検知すると、ログデータの配列順に、直前に再生した音声データの次の音声データをコミュニケーション装置Sに要求する。コミュニケーション装置Sは、要求された音声データをクライアント端末C1, , , Cxに提供し、それを受信したクライアント端末C1, , , Cxは音声データを再生する。この機能を使うと、履歴一覧のログデータの配列順(時系列順)で、音声データを聞くことができる。

30

【0050】

また、再生機能の付加機能として、「自分宛を聞く」ボタン14の押下を検知すると、それを操作したユーザの識別子(ハンドル名)が発言相手として付与されているログデータに対応する音声データのみをコミュニケーション装置Sに要求する。コミュニケーション装置Sはその音声データのみをクライアント端末C1, , , Cxに提供し、クライアント端末C1, , , Cxはそれを再生する。自分宛が複数ある場合は、履歴一覧の欄4で選択されている音声データよりも後の音声データ(後にアップロードされた音声データ)で、かつその選択されている音声データにもっとも時間的に近い音声データを再生する。時系列に順次連続して再生するようにしても良い。この機能を使うと、履歴一覧のなかから自分宛の音声データのみを聞くことができる。

40

【0051】

また、再生機能の付加機能として、「先行発言を聞く」ボタン15の押下を検知すると、履歴一覧の欄4から指定されているログデータに付加されている先行する音声データをコミュニケーション装置Sに要求する。コミュニケーション装置Sはその音声データのみをクライアント端末C1, , , Cxに提供し、クライアント端末C1, , , Cxはそれを

50

再生する。この機能を使うと、関連する先行の音声データを遡って聞くことができる。

【0052】

(第2の実施の形態)

図5(a)(b)は、本実施の形態のコミュニケーション装置Sを説明する説明図である。本実施の形態のコミュニケーション装置Sは、主に、会議などの対面方式の対話の場面において使用され、対面対話を行ないながらマルチスレッド対話を実現するものである。コミュニケーション装置Sは、上記実施の形態のコミュニケーション装置Sに、発言者が発言相手を指し示すような動作を行なうことにより、発言相手を指定することができる機能を備える。

【0053】

コミュニケーション装置Sは、信号を発信する発信手段b1と、発信手段から発信された信号を受信する受信手段b2とを備える。発信手段b1は、先端に赤外線発光ダイオード(LED)が内蔵されており、スイッチを押下することにより、赤外線光で識別子が発信される。発信手段b1は、すべてのユーザが一つずつ持ち、発信される識別子は発信手段b1ごとに異なるように、各々に固有の識別子が設定されている。各クライアント端末C1, , , CxにはRS-232Cインタフェースを介して受信手段b2が一つずつ接続されている。

【0054】

以下、発言者のクライアント端末C1と、発言相手のクライアント端末C2を例に説明する(図5(b))。発言者は、クライアント端末C1から音声録音機能を用いて音声データと発言者の識別子をコミュニケーション装置Sにアップロードし、図5(b)に示すように、発信手段b1を発言相手の受信手段b2に向けてスイッチを押下する。発信手段b1が赤外線光で発言者の識別子Xを発信すると、受信手段b2が赤外線受光部でその識別子Xを受信し、受信手段b2が接続されている発言相手のクライアント端末C2にその識別子Xを伝える。クライアント端末C2は、受信した識別子Xと、そのクライアント端末C2を使用する発言相手の識別子Yとを、コミュニケーション装置Sに通知する。

【0055】

コミュニケーション装置Sは、発言者のクライアント端末C1から、音声データと発言者の識別情報Aとを受信し、また、発言相手のクライアント端末C2から、発言者の識別子Xと発言相手の識別子Yとを受信する。発言者のクライアント端末C1から受信した発言者の識別子Xと、発言相手のクライアント端末C2から受信した発言者の識別子Xが一致する場合は、受信した音声データのログデータに発言相手の識別子Yを付加する。発言者は、発信手段b1により受信手段b2に信号を送るだけで、発言相手を指し示すような簡単な動作で音声データのログデータに発言相手の情報を付加することができる。受信した発言者の識別子Xを音声データのログデータに付加してもよい。これにより、発言者が発言相手を指し示すような簡単な動作で、音声データのログデータに発言者の情報を付加することができる。発言相手の識別子Yと発言者の識別子Xとを両方付加すると、より効果的である。

【0056】

(第3の実施の形態)

本実施の形態のコミュニケーション装置Sは、音声データに加えて、テキストデータを入力可能とするものである。コミュニケーション装置Sは、クライアント端末C1, , , Cxからテキストデータと音声データを特定する識別子a1を受信すると、そのテキストデータを音声データのログデータに関連付けて記憶する。そのテキストデータは音声データのログデータとともに、クライアント端末C1, , , Cxに提供され、各クライアント端末C1, , , Cxの履歴一覧の欄4にログデータとともにテキストデータが表示される。これにより、個々の音声データ間の関係を意味的に明示することが可能となる。また、音声では入力できない「顔文字」のようなものも、音声に付加して入力可能となる。

【0057】

(有効性に関する実験)

10

20

30

40

50

以下に、第1の実施の形態を例として、本発明の有効性を確認すべく、実験をおこなった。7人の大学院生から成るグループ2組計14人に対し、以下3つの条件での実験を行った。被験者は全員、何らかの形でテキストチャットを利用した経験はあるが、ボイスチャットの経験はない。また、システムの慣れによる影響を抑えるため、個々のグループの使用順序は異なる。

【0058】

・Base：一般的なインターフェースをもつボイスチャットシステムを非対面状況で使用。実際には、本発明のコミュニケーション装置（以下、ChaTELという）が提供する機能のうち、発言履歴と、「これを聞く」および「録音」ボタンのみを使用可能としたものを使用した。

- ・非対面ChaTEL：ChaTELをそのまま非対面状況で使用。
- ・対面のChaTEL：ChaTELをそのまま対面状況で使用。

【0059】

実験では、最初に7人の被験者からなるグループを3人と4人の2つのサブグループに分け、それぞれのサブグループに異なる話題を与え、これらの話題について2つのサブグループが同時並行的に約20分間話をするように教示した。与えた話題は、「行ってみたい場所」、「昔よくした遊びについて」など、比較的自由なテーマである。なお、これらの与えた話題については、各サブグループでひと通り完結するまで話を続けることを求めたが、それ以外の話をする事や、別のサブグループの話題に参加することについては禁止していない。また非対面条件では、全員が完全に離れた場所で行い、対面条件では、全員が円状に配置するようなセッティングを行なった。

【0060】

本システムで付与した、発言相手指定および先行発言指定機能により、マルチスレッド状況に対応しやすくなると考えられるため、対話構造そのものに影響が出る事予想される。そこで、まず、非対面状況で、発言相手指定および先行発言指定機能を持たないBaseシステムを使用した場合とChaTELを使用した場合の対話構造比較を行なった。ここで、対話構造を比較するために、実験で取得したデータの個々の発言がどの発言と意味的につながりがあるかを同定し、木構造の概念を用いて、始端数（スレッドの開始点の数）、パス数（1スレッド内の個々の発言を結びつける経路数）、終端数（スレッドの末端の数）を算出した。また、各スレッドの始端発言のIDと、最後の終端発言のIDの差をスレッド長とする。これは、それぞれのスレッドが幅広く展開されるのか、深く展開されるのかを判断するための材料のひとつとして使用できる。結果を表1に示す。

【0061】

【表1】

| | A-Base | A-ChaTEL | B-Base | B-ChaTEL |
|-------|--------|----------|--------|----------|
| 全発言数 | 154 発言 | 171 発言 | 161 発言 | 160 発言 |
| 始端数 | 13 | 10 | 10 | 6 |
| パス数 | 142 | 159 | 155 | 154 |
| 終端数 | 49 | 53 | 46 | 50 |
| スレッド長 | 23.1 | 49.6 | 27.9 | 67.8 |

【0062】

表1より、A、Bどちらの被験者群もBase、本システムどちらについてもスレッド長を除いた項目に大きな違いはないことがわかる。スレッド長については、A、Bどちらの被験者群も、Baseよりも本システムを利用したほうが大きくなる傾向にあることがわかる。これは、発言の分岐が起こっても、本システムを利用したほうが1スレッドが長く継続さ

れるということであり、スレッド数そのものがほぼ同じであれば、スレッドが長く続けば続くほどマルチスレッド状況が生じやすくなっているということの意味する。よって、本システムを利用した場合のほうが、マルチスレッドを長く継続できるということになる。

【0063】

また、発言相手指定および先行発言指定機能を持たないBaseシステムと本システムとでは、個々の参加メンバーがシステム利用中に同時参加しているスレッド数にどの程度差があるのかを算出した。なお、各参加者の同時参加スレッド数は、以下のようにして求めた。まず、各スレッドについて、個々の参加者による最初の発言と最後の発言を求め、その両発言の間はその参加者はそのスレッドに参加しているものとみなすこととした。その上で、個々の発言が行なわれた時に、個々の参加メンバーが、その時点で存在しているスレッドのうちいくつかに参加していたかということを出して求めた。結果を表2に示す。

10

【0064】

【表2】

| | A・Base | A・Chat | B・Base | B・Chat |
|-------------|--------|--------|--------|--------|
| 全発言数 | 154 発言 | 171 発言 | 161 発言 | 160 発言 |
| 平均同時参加スレッド数 | 0.75 | 1.34 | 0.82 | 1.43 |

20

【0065】

表2より、A、Bどちらの被験者群もBaseよりも本システムを利用した場合のほうが平均同時参加スレッド数が多くなっていることがわかる。また、Baseの平均同時参加スレッド数が1.0以下であるということは、自分が確実に発言できる機会がくるまでは、聞くことに専念していることを、また本システムの平均同時参加スレッド数が1.0以上であるということは、常にいずれかのスレッドで発言を行いながら、別のスレッドでも発言を行おうとしていることを示唆している。

【0066】

前節より、発言相手指定および先行発言指定機能をもつ本システムが、マルチスレッド状況の発生に有効に機能していることがわかった。しかし、これだけでは、対面状況でマルチスレッド状況に有効であるかという点では不十分である。そこで、本システムを非対面状況で用いた場合と対面状況で用いた場合との比較実験を行なった。まず、この場合も、前節同様の方法で、両システムを用いた際の対話構造の比較を行なった。結果を表3に示す。

30

【0067】

【表3】

| | A・非対面 | A・対面 | B・非対面 | B・対面 |
|-------|--------|--------|--------|--------|
| 全発言数 | 171 発言 | 162 発言 | 160 発言 | 158 発言 |
| 始端数 | 10 | 9 | 6 | 5 |
| パス数 | 159 | 155 | 154 | 149 |
| 終端数 | 53 | 51 | 50 | 55 |
| スレッド長 | 49.6 | 52.7 | 67.8 | 62.2 |

40

表3より、どちらの被験者群も、非対面、対面状況での対話構造に大きな差は見られなかった。

50

【0068】

さらに、非対面状況、対面状況での差異を見るため、対面状況での実験時に収録したビデオ収録データを検討した。その中で、非対面状況では、通常の対面対話とほぼ同じ音量で発言をしていたにも関わらず、対面状況では、小声で発言するということが音声データ、ビデオの両方から確認できた。さらに対面状況で参加者メンバーが目の前にいるにも関わらず、大半の参加者が発言履歴が表示されているモニターに注視している時間が長いということが観察された。これらから、対面状況での本システムの利用は、物理的には同じ場所を共有しつつも、心理的には独立した自分の空間を保持したまま対話していることが推測される。一方で、他の参加メンバーの笑い声に反応したり場の雰囲気共有している場面も見受けられた。

10

【0069】

本発明のコミュニケーション装置によれば、近接性、同期性の制約なしに対面でのマルチスレッド対話を可能とするために、音声による効率的な発言入力を可能とし、発言履歴と相手指定/対応発言指定機能の提供によって、対面状況でも音声によるマルチスレッド対話を可能とした。さらに、開発したシステムがマルチスレッド対話の進行に有効となることを、実験で得た対話データを中心に分析を行ない、検証した。

【0070】

非対面条件で、通常のインターフェースを持つボイスチャットシステムと本研究で開発したシステムとの比較では、前者よりも後者のほうが、1スレッドあたりのスレッド長が長くなることが確認された。また、本研究で開発したシステムを用いたほうが、各参加者が同時に関与するスレッド数が増えることが確認された。これらのことから、相手指定や先行発言指定を行なうことで、同じ話題を長く維持しながら、複数の話題に同時に参加できることが示され、本システムを用いるとマルチスレッド状況が生じやすくなるということが示唆された。

20

【0071】

対面状況時のビデオ記録から、笑いを共有するなどの場の雰囲気を共有している場面はあるにせよ、発言履歴に注視しがちになってしまい、参加者同士が表情を確認する等の対面対話の利点が生かされていないことが示唆された。心理的には互いに独立したままの対話空間を、実空間に接地することにより、対面状況対話空間を共有しやすくする工夫が必要である。そこで、対面状況で、相手の表情を確認できるくらいの視線の移動がごく自然な動作で可能となるよう、相手指定を、発言者が相手を「指差しする」動作によって行なう機能を追加した。これにより、実世界での身体的に相手を指差しするという自然な行為による発言の対象者指定が可能となり、その行為の再にも相手の方を見ることで、場の雰囲気そのものもより自然なものとなった。

30

【0072】

本発明は、その趣旨を逸脱しない範囲で適宜変更が可能であることはいうまでもない。

【図面の簡単な説明】

【0073】

【図1】本実施の形態のコミュニケーション装置を説明する説明図である。

40

【図2】コミュニケーション装置がクライアント端末に提供するユーザインタフェース画面を示す図。

【図3】「発言完了」ボタンを示す図。

【図4】ユーザインタフェース画面の履歴一覧の欄のログデータを抽出して示す図。

【図5】本実施の形態のコミュニケーション装置を説明する説明図。

【図6】3つの表現が対話中に出現した割合のグラフ。

【図7】3つの表現を発言間のインターバル(発言間距離)ごとに分類した出現頻度の結果を示すグラフ。

【符号の説明】

【0074】

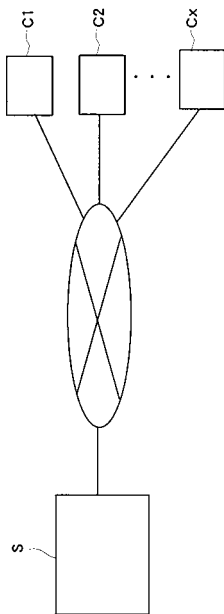
50

- S コミュニケーション装置
- C 1, ..., C x クライアント端末
- 1 ハンドル名を入力する欄
- 2 ログインのボタン
- 3 ユーザー一覧の欄
- 4 履歴一覧の欄
- 5 録音関連ボタン群
- 6 「録音」ボタン
- 7 「先行発言指定」ボタン
- 8 「発言への返信録音」ボタン
- 9 「発言相手指定」ボタン
- 10 「相手指定録音」ボタン
- 11 「発言者への返信録音」ボタン
- 12 「これを聞く」ボタン
- 13 「次を聞く」ボタン
- 14 「自分宛を聞く」ボタン
- 15 「先行発言を聞く」ボタン
- a 1 音声データを特定する識別子
- a 2 発言者を特定する識別子
- a 3 発言時刻となるアップロード時刻
- a 4 先行する音声データを特定する識別子
- b 1 発信手段
- b 2 受信手段

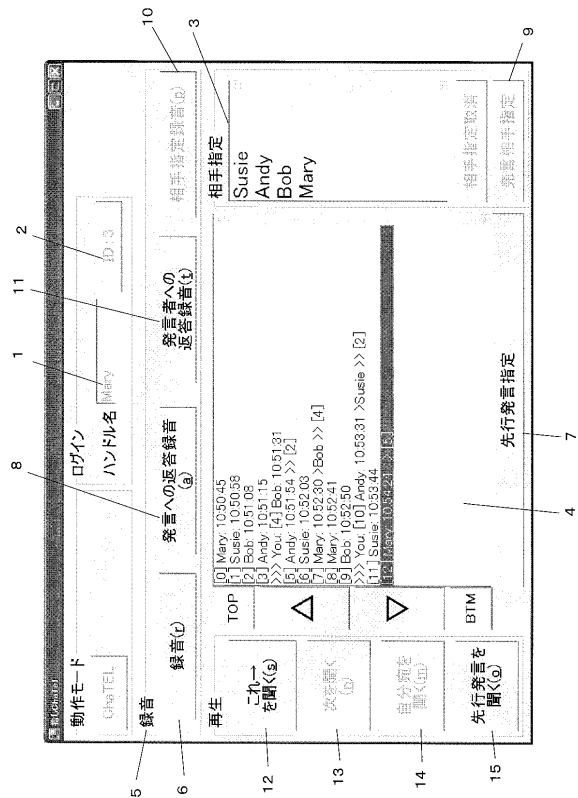
10

20

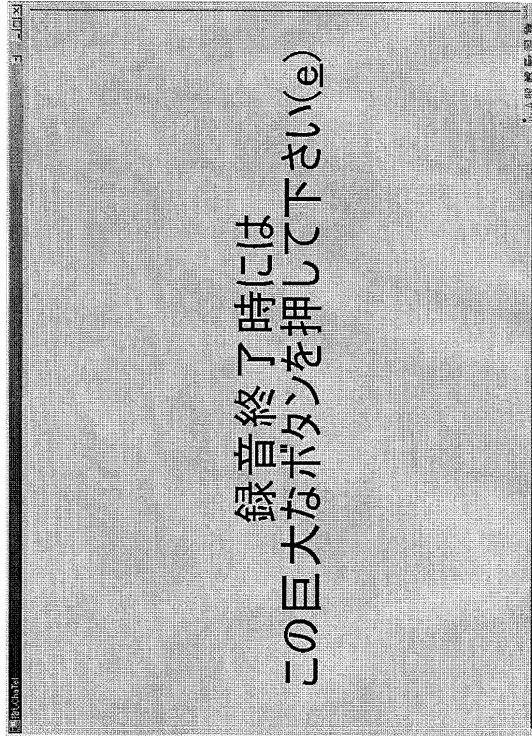
【図1】



【図2】



【 図 3 】



【 図 4 】

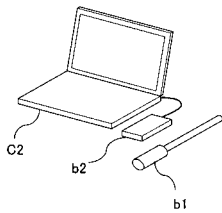
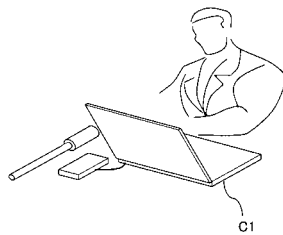
```

a1 { [0] Mary:23:11:49
      [1] Susie:23:12:02
      [2] Andy:23:12:36
      >>> You :[4]Bob:23:13:02 a4
      [5] Andy:23:13:23 >>[2]
      [6] Susie:23:13:44
      [7] Mary:23:13:03 >Bob >>[4]
      [8] Mary:23:14:18
      [9] Bob:23:14:47
      [10] Andy:23:15:06 >Susie >Andy
      [11] Susie:20:15:30
      [12] Mary:23:16:14 >>[12]
    }
a2
a3
a4
a5

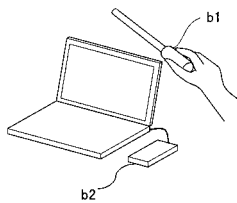
```

【 図 5 】

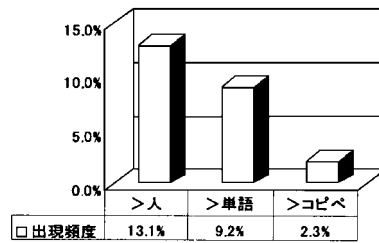
(a)



(b)



【 図 6 】



【 図 7 】

