

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4608740号  
(P4608740)

(45) 発行日 平成23年1月12日(2011.1.12)

(24) 登録日 平成22年10月22日(2010.10.22)

|                             |                 |
|-----------------------------|-----------------|
| (51) Int.Cl.                | F I             |
| <b>G06F 17/30 (2006.01)</b> | G06F 17/30 370Z |
| <b>G06F 3/048 (2006.01)</b> | G06F 17/30 170A |
| <b>G06F 13/00 (2006.01)</b> | G06F 17/30 210A |
|                             | G06F 3/048 651A |
|                             | G06F 3/048 653A |
| 請求項の数 8 (全 25 頁) 最終頁に続く     |                 |

|              |                               |           |                            |
|--------------|-------------------------------|-----------|----------------------------|
| (21) 出願番号    | 特願2000-187152 (P2000-187152)  | (73) 特許権者 | 000002185                  |
| (22) 出願日     | 平成12年6月22日(2000.6.22)         |           | ソニー株式会社                    |
| (65) 公開番号    | 特開2001-312515 (P2001-312515A) |           | 東京都港区港南1丁目7番1号             |
| (43) 公開日     | 平成13年11月9日(2001.11.9)         | (74) 代理人  | 100082131                  |
| 審査請求日        | 平成19年2月20日(2007.2.20)         |           | 弁理士 稲本 義雄                  |
| (31) 優先権主張番号 | 特願2000-42303 (P2000-42303)    | (72) 発明者  | 斎藤 真理                      |
| (32) 優先日     | 平成12年2月21日(2000.2.21)         |           | 東京都品川区北品川6丁目7番35号 ソニー株式会社内 |
| (33) 優先権主張国  | 日本国(JP)                       | (72) 発明者  | 山本 則行                      |
| (31) 優先権主張番号 | 特願2000-42305 (P2000-42305)    |           | 東京都品川区北品川6丁目7番35号 ソニー株式会社内 |
| (32) 優先日     | 平成12年2月21日(2000.2.21)         | (72) 発明者  | 法林 浩之                      |
| (33) 優先権主張国  | 日本国(JP)                       |           | 東京都品川区北品川6丁目7番35号 ソニー株式会社内 |
| 最終頁に続く       |                               |           |                            |

(54) 【発明の名称】 情報処理装置および方法、並びにプログラム格納媒体

(57) 【特許請求の範囲】

【請求項1】

1以上の既存のテキストデータから、少なくとも日時情報を含む属性情報を抽出し、抽出した前記属性情報に含まれる各項目の情報の共通性に基づいて前記既存のテキストデータをグループ化するテキストデータグループ化手段と、

グループ化された前記既存のテキストデータに含まれる各キーワードに対し、前記キーワードの出現頻度または複数の前記既存のテキストデータに亘る分布情報の少なくとも一方に基づいて重要度を算出する算出手段と、

グループ化された前記既存のテキストデータに含まれるキーワードの中から、算出された前記重要度に基づいて検索用キーワードを選択する選択手段と、

選択された前記検索用キーワードに関連する関連情報を取得する取得手段と、

グループ化された前記既存のテキストデータを前記属性情報の前記日時情報に従ってソートし、前記既存のテキストデータそれぞれにおける各キーワードの前記重要度を保持するとともに、選択された前記検索用キーワードに取得された前記関連情報を対応付けたデータベースを構築する構築手段と、

新たなテキストデータの出現をイベントとして検知する検知手段と、

検知された前記イベントに対応する前記新たなテキストデータからキーワードを検出する検出手段と、

前記新たなテキストデータから検出された前記キーワードのうち、前記データベースに保持されている前記重要度が所定の条件を満たすものを検索し、検索したキーワードに対

応する前記関連情報を前記データベースから読み出す検索手段と、  
読み出された前記関連情報の表示を制御する表示制御手段と  
を含み、

前記検出手段は、前記新たなテキストデータから検出された前記キーワードのうち、前記データベースに保持されている前記重要度が、現時点以前の期間 X に所定の閾値 A 以下であり、且つ、現時点以前の前記期間 X よりも長い期間 Y に異なる 2 以上のテキストデータにおいて所定の閾値 B 以上であるものを検索する

情報処理装置。

【請求項 2】

前記検知手段は、電子メールの送信、受信、または編集による前記新たなテキストデータの出現を前記イベントとして検知する

10

請求項 1 に記載の情報処理装置。

【請求項 3】

前記取得手段は、前記検索用キーワードが含まれる Web ページのタイトルおよび URL を前記関連情報として取得する

請求項 1 に記載の情報処理装置。

【請求項 4】

前記取得手段は、所定のタイミングにおいて、選択した前記検索用キーワードに関連する前記関連情報を取得する

請求項 3 に記載の情報処理装置。

20

【請求項 5】

前記構築手段は、所定の更新条件が満たされた場合、構築した前記データベースを更新する

請求項 1 に記載の情報処理装置。

【請求項 6】

前記所定の更新条件は、ユーザが設定可能である

請求項 5 に記載の情報処理装置。

【請求項 7】

グループ化手段と、

算出手段と、

30

選択手段と、

取得手段と、

構築手段と、

検知手段と、

検出手段と、

検索手段と、

表示制御手段とを備える情報処理装置の情報処理方法において、

前記グループ化手段による、1 以上の既存のテキストデータから、少なくとも日時情報を含む属性情報を抽出し、抽出した前記属性情報に含まれる各項目の情報の共通性に基づいて前記既存のテキストデータをグループ化するテキストデータグループ化ステップと、

40

前記算出手段による、グループ化された前記既存のテキストデータに含まれる各キーワードに対し、前記キーワードの出現頻度または複数の前記既存のテキストデータに亘る分布情報の少なくとも一方に基づいて重要度を算出する算出ステップと、

前記選択手段による、グループ化された前記既存のテキストデータに含まれるキーワードの中から、算出された前記重要度に基づいて検索用キーワードを選択する選択ステップと、

前記取得手段による、選択された前記検索用キーワードに関連する関連情報を取得する取得ステップと、

前記構築手段による、グループ化された前記既存のテキストデータを前記属性情報の前記日時情報に従ってソートし、前記既存のテキストデータそれぞれにおける各キーワード

50

の前記重要度を保持するとともに、選択された前記検索用キーワードに取得された前記関連情報を対応付けたデータベースを構築する構築ステップと、

前記検知手段による、新たなテキストデータの出現をイベントとして検知する検知ステップと、

前記検出手段による、検知された前記イベントに対応する前記新たなテキストデータからキーワードを検出する検出ステップと、

前記検索手段による、前記新たなテキストデータから検出された前記キーワードのうち、前記データベースに保持されている前記重要度が所定の条件を満たすものを検索し、検索したキーワードに対応する前記関連情報を前記データベースから読み出す検索ステップと、

前記表示制御手段による、読み出された前記関連情報の表示を制御する表示制御ステップと

を含み、

前記検出手段による、前記検索ステップは、前記新たなテキストデータから検出された前記キーワードのうち、前記データベースに保持されている前記重要度が、現時点以前の期間 X に所定の閾値 A 以下であり、且つ、現時点以前の前記期間 X よりも長い期間 Y に異なる 2 以上のテキストデータにおいて所定の閾値 B 以上であるものを検索する

情報処理方法。

#### 【請求項 8】

コンピュータを、

1 以上の既存のテキストデータから、少なくとも日時情報を含む属性情報を抽出し、抽出した前記属性情報に含まれる各項目の情報の共通性に基づいて前記既存のテキストデータをグループ化するテキストデータグループ化手段と、

グループ化された前記既存のテキストデータに含まれる各キーワードに対し、前記キーワードの出現頻度または複数の前記既存のテキストデータに亘る分布情報の少なくとも一方に基づいて重要度を算出する算出手段と、

グループ化された前記既存のテキストデータに含まれるキーワードの中から、算出された前記重要度に基づいて検索用キーワードを選択する選択手段と、

選択された前記検索用キーワードに関連する関連情報を取得する取得手段と、

グループ化された前記既存のテキストデータを前記属性情報の前記日時情報に従ってソートし、前記既存のテキストデータそれぞれにおける各キーワードの前記重要度を保持するとともに、選択された前記検索用キーワードに取得された前記関連情報を対応付けたデータベースを構築する構築手段と、

新たなテキストデータの出現をイベントとして検知する検知手段と、

検知された前記イベントに対応する前記新たなテキストデータからキーワードを検出する検出手段と、

前記新たなテキストデータから検出された前記キーワードのうち、前記データベースに保持されている前記重要度が所定の条件を満たすものを検索し、検索したキーワードに対応する前記関連情報を前記データベースから読み出す検索手段と、

読み出された前記関連情報の表示を制御する表示制御手段と

して動作させるためのプログラムが格納されており、

前記検出手段は、前記新たなテキストデータから検出された前記キーワードのうち、前記データベースに保持されている前記重要度が、現時点以前の期間 X に所定の閾値 A 以下であり、且つ、現時点以前の前記期間 X よりも長い期間 Y に異なる 2 以上のテキストデータにおいて所定の閾値 B 以上であるものを検索する

プログラム格納媒体。

#### 【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、情報処理装置および方法、並びにプログラム格納媒体に関し、例えば、電子メ

10

20

30

40

50

ールが送受信されたときに、当該電子メールの文面に関連する情報をデータベースから抽出して表示させるようにした情報処理装置および方法、並びにプログラム格納媒体に関する。

【0002】

【従来の技術】

従来、パーソナルコンピュータのデスクトップ（表示画面）に、いわゆるデスクトップマスケットと呼ばれるキャラクタを表示させるアプリケーションプログラムが存在する。

【0003】

デスクトップマスケットは、例えば、電子メールの着信等をユーザに通知する機能やデスクトップ上を移動する機能などを有している。

10

【0004】

【発明が解決しようとする課題】

ところで、例えばユーザが電子メールとして送信する文書等を入力している時や、受信した文書を開覧している時などに、対象としている文書に関連する情報（以下、関連情報と記述する）を、デスクトップマスケットが提示するようになれば、ユーザにとって利便性が向上するとともに、デスクトップマスケットに対して一層愛着を感じるようになる。しかしながら、そのようなデスクトップマスケットは実現されていない課題があった。

【0005】

また、関連情報をユーザに提示するためには、関連情報のデータベースを予め構築しておく必要があるが、データベースを構築するためには、データの入力等の手間が掛かる課題があった。

20

【0006】

本発明はこのような状況に鑑みてなされたものであり、処理済みのテキストファイルを用いて、自動的にデータベースを構築できるようにするとともに、送受信などの対象とした文書に関連する関連情報をユーザに提示できるようにすることを目的とする。

【0007】

【課題を解決するための手段】

本発明の情報処理装置は、1以上の既存のテキストデータから、少なくとも日時情報を含む属性情報を抽出し、抽出した属性情報に含まれる各項目の情報の共通性に基づいて既存のテキストデータをグループ化するテキストデータグループ化手段と、グループ化された既存のテキストデータに含まれる各キーワードに対し、キーワードの出現頻度または複数の既存のテキストデータに亘る分布情報の少なくとも一方に基づいて重要度を算出する算出手段と、グループ化された既存のテキストデータに含まれるキーワードの中から、算出された重要度に基づいて検索用キーワードを選択する選択手段と、選択された検索用キーワードに関連する関連情報を取得する取得手段と、グループ化された既存のテキストデータを属性情報の日時情報に従ってソートし、既存のテキストデータそれぞれにおける各キーワードの重要度を保持するとともに、選択された検索用キーワードに取得された関連情報を対応付けたデータベースを構築する構築手段と、新たなテキストデータの出現をイベントとして検知する検知手段と、検知されたイベントに対応する新たなテキストデータからキーワードを検出する検出手段と、新たなテキストデータから検出されたキーワードのうち、データベースに保持されている重要度が所定の条件を満たすものを検索し、検索したキーワードに対応する関連情報をデータベースから読み出す検索手段と、読み出された関連情報の表示を制御する表示制御手段とを含み、検出手段は、新たなテキストデータから検出されたキーワードのうち、データベースに保持されている重要度が、現時点以前の期間Xに所定の閾値A以下であり、且つ、現時点以前の期間Xよりも長い期間Yに異なる2以上のテキストデータにおいて所定の閾値B以上であるものを検索する。

30

40

【0008】

前記検知手段は、電子メールの送信、受信、または編集による新たなテキストデータの出現をイベントとして検知するようにすることができる。

【0009】

50

前記取得手段は、検索用キーワードが含まれるWebページのタイトルおよびURLを関連情報として取得するようにすることができる。

【0010】

前記取得手段は、所定のタイミングにおいて、選択した検索用キーワードに関連する関連情報を取得するようにすることができる。

【0011】

前記構築手段は、所定の更新条件が満たされた場合、構築したデータベースを更新するようにすることができる。

【0012】

前記所定の更新条件は、ユーザが設定可能であるようにすることができる。

10

【0013】

本発明の情報処理方法は、グループ化手段による、1以上の既存のテキストデータから、少なくとも日時情報を含む属性情報を抽出し、抽出した属性情報に含まれる各項目の情報の共通性に基づいて既存のテキストデータをグループ化するテキストデータグループ化ステップと、算出手段による、グループ化された既存のテキストデータに含まれる各キーワードに対し、キーワードの出現頻度または複数の既存のテキストデータに亘る分布情報の少なくとも一方に基づいて重要度を算出する算出ステップと、選択手段による、グループ化された既存のテキストデータに含まれるキーワードの中から、算出された重要度に基づいて検索用キーワードを選択する選択ステップと、取得手段による、選択された検索用キーワードに関連する関連情報を取得する取得ステップと、構築手段による、グループ化された既存のテキストデータを属性情報の日時情報に従ってソートし、既存のテキストデータそれぞれにおける各キーワードの重要度を保持するとともに、選択された検索用キーワードに取得された関連情報を対応付けたデータベースを構築する構築ステップと、検知手段による、新たなテキストデータの出現をイベントとして検知する検知ステップと、検出手段による、検知されたイベントに対応する新たなテキストデータからキーワードを検出する検出ステップと、検索手段による、新たなテキストデータから検出されたキーワードのうち、データベースに保持されている重要度が所定の条件を満たすものを検索し、検索したキーワードに対応する関連情報をデータベースから読み出す検索ステップと、表示制御手段による、読み出された関連情報の表示を制御する表示制御ステップとを含み、検出手段による、検索ステップは、新たなテキストデータから検出されたキーワードのうち、データベースに保持されている重要度が、現時点以前の期間Xに所定の閾値A以下であり、且つ、現時点以前の期間Xよりも長い期間Yに異なる2以上のテキストデータにおいて所定の閾値B以上であるものを検索する。

20

30

【0014】

本発明のプログラム格納媒体は、コンピュータを、1以上の既存のテキストデータから、少なくとも日時情報を含む属性情報を抽出し、抽出した属性情報に含まれる各項目の情報の共通性に基づいて既存のテキストデータをグループ化するテキストデータグループ化手段と、グループ化された既存のテキストデータに含まれる各キーワードに対し、キーワードの出現頻度または複数の既存のテキストデータに亘る分布情報の少なくとも一方に基づいて重要度を算出する算出手段と、グループ化された既存のテキストデータに含まれるキーワードの中から、算出された重要度に基づいて検索用キーワードを選択する選択手段と、選択された検索用キーワードに関連する関連情報を取得する取得手段と、グループ化された既存のテキストデータを属性情報の日時情報に従ってソートし、既存のテキストデータそれぞれにおける各キーワードの重要度を保持するとともに、選択された検索用キーワードに取得された関連情報を対応付けたデータベースを構築する構築手段と、新たなテキストデータの出現をイベントとして検知する検知手段と、検知されたイベントに対応する新たなテキストデータからキーワードを検出する検出手段と、新たなテキストデータから検出されたキーワードのうち、データベースに保持されている重要度が所定の条件を満たすものを検索し、検索したキーワードに対応する関連情報をデータベースから読み出す検索手段と、読み出された関連情報の表示を制御する表示制御手段として動作させるため

40

50

のプログラムが格納されており、検出手段は、新たなテキストデータから検出されたキーワードのうち、データベースに保持されている重要度が、現時点以前の期間 X に所定の閾値 A 以下であり、且つ、現時点以前の期間 X よりも長い期間 Y に異なる 2 以上のテキストデータにおいて所定の閾値 B 以上であるものを検索する。

【 0 0 1 5 】

本発明においては、1 以上の既存のテキストデータが属性情報に含まれる各項目の情報の共通性に基づいてグループ化され、グループ化された既存のテキストデータに含まれる各キーワードに対し、キーワードの出現頻度または複数の既存のテキストデータに亘る分布情報の少なくとも一方に基づいて重要度が算出され、グループ化された既存のテキストデータに含まれるキーワードの中から、算出された重要度に基づいて検索用キーワードが選択される。また、選択された検索用キーワードに関連する関連情報が取得され、グループ化された既存のテキストデータが属性情報の日時情報に従ってソートされ、既存のテキストデータそれぞれにおける各キーワードの重要度が保持されるとともに、選択された検索用キーワードに取得された関連情報を対応付けたデータベースが構築される。さらに、新たなテキストデータの出現がイベントとして検知され、検知されたイベントに対応する新たなテキストデータからキーワードが検出され、新たなテキストデータから検出されたキーワードのうち、データベースに保持されている重要度が所定の条件を満たすものが検索され、検索されたキーワードに対応する関連情報がデータベースから読み出され、その表示が制御される。なお、検索の処理では、新たなテキストデータから検出されたキーワードのうち、データベースに保持されている重要度が、現時点以前の期間 X に所定の閾値 A 以下であり、且つ、現時点以前の期間 X よりも長い期間 Y に異なる 2 以上のテキストデータにおいて所定の閾値 B 以上であるものが検索される。

【 0 0 2 5 】

【発明の実施の形態】

図 1 は、本発明を適用したデスクトップマスコット（以下、エージェントと記述する）をデスクトップ上に表示するアプリケーションプログラム（以下、エージェントプログラムと記述する）の機能ブロックを示している。エージェントプログラムは、例えば、パーソナルコンピュータ（詳細は図 2 を参照して後述する）にインストールされて実行されるものである。

【 0 0 2 6 】

エージェントプログラムの機能ブロックは、処理の対象となる文書の関連情報（後述）を蓄積してデータベースを構築する蓄積部 1、処理の対象となる文書に対応する関連情報をユーザに提示する提示部 7、および、エージェントの表示等を制御するエージェント制御部 11 から構成される。

【 0 0 2 7 】

蓄積部 1 の文書取得部 2 は、電子メールを送受信するアプリケーションプログラム（以下、メーラ(mailer)と記述する）によって送受信した文書やワードプロセッサプログラムによって編集された文書などのうち、自己が未処理の文書を取得して文書属性処理部 3 および文書内容処理部 4 に供給する。なお、以下においては、主に、メーラにより電子メールとして送受信した文書を例として説明する。

【 0 0 2 8 】

文書属性処理部 3 は、文書取得部 2 から供給される文書の属性情報を抽出し、属性情報に基づき、文書をグループ化して、文書内容処理部 4 および文書特徴データベース作成部 5 に供給する。電子メールである場合、文書の属性情報としてヘッダに記述されている情報（対象となっているメールを特定するメッセージ ID、参照しているメールのメッセージ ID (References, In-Reply-To)、宛先(To,Cc,Bcc)、送信元(From)、日付(Date)）が抽出される。

【 0 0 2 9 】

文書内容処理部 4 は、文書属性処理部 3 でグループ化された文書の文面を抽出し、形態素解析を施してキーワード（単語）に分割し、キーワードの出現頻度、複数の文書に亘る分

10

20

30

40

50

布状況を求め、文書のグループ毎に、キーワードの重み（文書の主旨に関係する程度を示す値）を演算する。重みの演算には、例えば、tf・idf法(Salton,G.:Automatic Text Processing:The Transformation,Analysis, and Retrieval of Information by Computer,Addison-Wesley,1989)等を用いる。ただし、広範囲に亘って分布している単語、すなわち、例えば、ほぼ全ての文書に含まれている「よろしく」、「お願いします」等の単語は、キーワードには成り得ないので除外する。

#### 【0030】

文書特徴データベース作成部5は、文書属性処理部3でグループ化された文書毎に、各文書の属性情報と、文書内に含まれる全てのキーワードの重みをデータベース化してハードディスクドライブなどよりなる記憶部29（図2）に記録する。文書特徴データベース作成部5はまた、キーワードの重みに基づき、検索用のキーワード（重要語）を選択して記録する。文書特徴データベース作成部5はさらに、検索用のキーワードを関連情報検索部6に供給し、それに対応して関連情報検索部6から返される関連情報を、検索用のキーワードに対応付けて記録する。

10

#### 【0031】

関連情報検索部6は、文書特徴データベース作成部5が選択して記録した検索用のキーワードの関連情報を検索し、検索結果のインデックスを文書特徴データベース作成部5に返す。キーワードの関連情報を検索するには、例えば、インターネット上の検索エンジンに用いる方法がある。この場合、WebページのURL(Uniform Resource Locator)やWebページのタイトルが、関連情報として文書特徴データベース作成部5に返される。

20

#### 【0032】

提示部7のイベント抽出部8は、メーラが電子メールの送信や受信を終えたことや、編集中の文書のテキストデータ量が所定の閾値を超えたことを検知して、データベース問い合わせ部9に通知する。以下、メーラが電子メールの送信や受信を終えたことや、編集中の文書のテキストデータ量が所定の閾値を超えたことをイベント発生と記述する。

#### 【0033】

データベース問い合わせ部9は、イベント抽出部8からのイベント発生の通知に対応して、イベント発生に対応する文書（例えば、受信した電子メール）を取得し、文書属性処理部3および文書内容処理部4の処理と同様に、当該文書から属性情報とキーワードを取得する。データベース問い合わせ部9はさらに、取得した属性情報とキーワードを検索条件として文書特徴データベース作成部5が作成したデータベースを検索し、検索結果のうちのキーワードの重みが所定の条件（詳細については図6を参照して後述する）を満たすものを選択して、対応する関連情報を、イベント抽出部8を介して関連情報提示部10に供給する。

30

#### 【0034】

関連情報提示部10は、イベント抽出部8を介してデータベース問い合わせ部9から供給される関連情報をデスクトップ上に表示させる。

#### 【0035】

なお、蓄積部1の処理によるデータベースの作成、更新は、所定のタイミング（図9および図10を参照して後述する）において実行される。一方、提示部7の処理は、イベント抽出部8がイベント発生を検知する毎に実行される。

40

#### 【0036】

図2は、エージェントプログラムがインストールされて実行されるパーソナルコンピュータの構成例を示している。

#### 【0037】

パーソナルコンピュータは、CPU(Central Processing Unit)21を内蔵している。CPU21にはバス24を介して、入出力インタフェース25が接続されている。入出力インタフェース25には、キーボード、マウスなどの入力デバイスよりなる入力部26、処理結果としての例えば音声信号を出力する出力部27、処理結果としての画像を表示するディスプレイなどよりなる表示部28、プログラムや各種データを格納するハードディスクドラ

50

イブなどよりなる記憶部 29、LAN(Local Area Network)やインターネットを介してデータを通信するモデムなどよりなる通信部 30、および、磁気ディスク 32乃至半導体メモリ 35などのプログラム格納媒体に対してデータを読み書きするドライブ 31が接続されている。バス 24には、ROM(Read Only Memory) 22およびRAM(Random Access Memory) 23が接続されている。

【0038】

本発明のエージェントプログラムは、磁気ディスク 32乃至半導体メモリ 35に格納された状態でパーソナルコンピュータに供給され、ドライブ 31によって読み出されて、記憶部 29に内蔵されるハードディスクドライブにインストールされている。記憶部 29にインストールされているエージェントプログラムは、入力部 26に入力されるユーザからのコマンドに対応するCPU 21の指令によって、記憶部 29からRAM 23にロードされて実行される。なお、パーソナルコンピュータの起動時において自動的にエージェントプログラムが実行されるように設定しておくようにしてもよい。

10

【0039】

なお、記憶部 29に内蔵されるハードディスクドライブには、エージェントプログラムの他、WWWブラウザ、メーラ、およびワードプロセッサなどのアプリケーションプログラムもインストールされており、エージェントプログラムと同様に、入力部 26に入力されるユーザからの起動コマンドに対応するCPU 21の指令によって、記憶部 29からRAM 23にロードされて実行される。

【0040】

次に、エージェントプログラムのデータベース作成処理について、図3のフローチャートを参照して説明する。このデータベース作成処理は、エージェントプログラムが実行する処理のうちの一つであり、エージェントプログラムが実行された場合において、データベースが作成されていないときに開始される。

20

【0041】

ステップ S1において、文書取得部 2は、未処理の文書(例えば、前回のデータベース作成処理が終了された以降に送受信された電子メール)を、記録部 29に内蔵されるハードディスクドライブから取得して文書属性処理部 3および文書内容処理部 4に供給する。

【0042】

ステップ S2において、文書属性処理部 3は、文書取得部 2から供給された電子メールの文書から属性情報(メッセージID等のヘッダ情報)を抽出し、属性情報に基づき、文書をグループ化して、文書内容処理部 4および文書特徴データベース作成部 5に供給する。

30

【0043】

ステップ S3において、文書内容処理部 4は、文書属性処理部 3でグループ化された文書の本文のテキストデータを抽出し、形態素解析を施し、キーワードとして抽出する。ステップ S4において、文書内容処理部 4は、キーワードの出現頻度、複数の文書に亘る分布状況を求め、文書のグループ毎にキーワードとなる単語の重みを、例えばtf・idf法を用いて演算する。

【0044】

ステップ S5において、文書特徴データベース作成部 5は、文書属性処理部 3でグループ化された文書毎に、各文書の属性情報と文書内に含まれる全てのキーワードの重みをデータベース化する。具体的には、図4に示すように、グループ化された文書を時系列順にソートした後、グループ化された文書に含まれる全キーワード 1乃至 nの重みも時系列順にソートして記憶部 29に記録する。ここで、図4においては、例えば、重み A1は、文書 Aにおけるキーワード 1の重みの値を示し、重み B2は、文書 Bにおけるキーワード 2の重みの値を示している。さらに例えば、文書 Bにキーワード 1が含まれていない場合、重み B1は 0となる。

40

【0045】

ステップ S6において、文書特徴データベース作成部 5は、重みが所定の閾値以上であるキーワードを検索用キーワード(重要語)に選択し、または、重みが大きい順に指定した

50



個数のキーワードを選択し、関連情報検索部 6 に供給する。

関連情報検索部 6 は、文書特徴データベース作成部 5 からの検索用キーワードを検索条件として、インターネット上の検索エンジンにアクセスして検索結果を取得し、得られた Web ページの URL やタイトルを関連情報として文書特徴データベース作成部 5 に出力する。

【 0 0 4 6 】

ステップ S 7 において、文書特徴データベース作成部 5 は、関連情報検索部 6 からの関連情報を検索用キーワードに対応付けて記憶部 2 9 に記録することにより、データベースを作成する。なお、ステップ S 6 の処理とステップ S 7 の処理は、ステップ S 5 までの一連の処理に継続して実行する場合と、一連の処理に継続せず、所定のタイミングにおいて実行する場合がある。

10

【 0 0 4 7 】

以上のように、データベース作成処理が所定の頻度で実行されることにより、送受信した電子メールの文書に対応した関連情報がデータベース内に蓄積されることになる。なお、データベース作成処理を任意のタイミングで開始させることも可能である。さらに、このようにして作成されたデータベースは、所定の条件が満たされたときに更新される（更新のタイミングについては、図 9 および図 1 0 を参照して後述する）。

【 0 0 4 8 】

ところで、当該データベース作成処理をユーザが強制的に中断することができるようにするために、中断要求があった場合、中断された時点で処理済みの文書を記録し、再開要求があった場合、未処理の文書から処理を再開するようにする。

20

【 0 0 4 9 】

次に、エージェントプログラムの関連情報提示処理について、図 5 のフローチャートを参照して説明する。この関連情報提示処理は、上述したデータベース作成処理とは異なり、エージェントプログラムが実行されている間、繰り返して実行される。

【 0 0 5 0 】

ステップ S 1 1 において、エージェントプログラムの終了が指示されたか否かが判定され、エージェントプログラムの終了が指示されていないと判定された場合、ステップ S 1 2 に進む。ステップ S 1 2 において、イベント抽出部 8 は、イベント発生（電子メールの送受信の終了等）を監視する。イベント発生が検知されない場合、ステップ S 1 1 に戻り、それ以降の処理が繰り返される。ステップ S 1 2 において、イベント発生が検知された場合、ステップ S 1 3 に進む。

30

【 0 0 5 1 】

ステップ S 1 3 において、イベント抽出部 8 は、イベント発生をデータベース問い合わせ部 9 に通知する。データベース問い合わせ部 9 は、イベント抽出部 8 からのイベント発生の通知に対応して、イベント発生に対応する文書（送受信された電子メール）を取得し、当該文書から属性情報を抽出する。ステップ S 1 4 において、データベース問い合わせ部 9 は、当該文書からキーワードを抽出する。

【 0 0 5 2 】

ステップ S 1 5 において、データベース問い合わせ部 9 は、ステップ S 1 3 の処理で抽出した属性情報と、ステップ S 1 4 の処理で抽出したキーワードを検索条件として、文書特徴データベース作成部 5 が作成したデータベースを検索し、合致するものを抽出する。さらに、データベース問い合わせ部 9 は、検索結果として抽出したもののうち、キーワードの重みの時系列推移に着目して、以下に示す条件 1 および条件 2 を満たすキーワードを選択する。

40

【 0 0 5 3 】

キーワードの選択条件について、図 6 を参照して説明する。図 6 は、データベースに蓄積されているキーワードの重みの時系列推移の例を示している。条件 1 は、「キーワードの重みが、現時点以前の所定の期間 X（例えば、2 週間）、所定の閾値 A 以下であること」である。条件 2 は、「現時点以前の所定の期間 Y（例えば、5 週間）において、異なる 2 以上の文書で、キーワードの重みが所定の閾値 B 以上であること」である。なお、条件 3

50

「条件2における異なる2以上の文書のうち、最も古い文書と最も新しい文書が所定の期間Z以上離れていること」を追加すればさらによい。

【0054】

このような条件を用いることにより、ユーザの興味が高いと思われるキーワードを選択することが可能となる。特に、条件1によって、現時点に近い文書に含まれるキーワードは除外されるので、ユーザが現時点で意識していると思われる関連情報（新しすぎる情報）を改めて選択することを避けることができ、且つ、かなり以前の文書に含まれるキーワードも除外されるので、ユーザが現時点で思い出すことができないと思われる関連情報（古すぎる情報）を改めて選択することを避けることもできる。

【0055】

データベース問い合わせ部9は、このようにして選択したキーワードに対応する関連情報を、イベント抽出部8を介して関連情報提示部10に供給する。

【0056】

図5に戻り、ステップS16において、エージェント制御部11は、ステップS15の処理で選択したキーワードが含まれている文書の属性情報を、選択した理由として表示させるとともに、対応する関連情報を表示するか否かをユーザに問う入力ウィンドウ61（図17）をデスクトップに表示させる。

【0057】

なお、入力ウィンドウ61を表示させず、直接、デスクトップ上に関連情報を表示させるようにしてもよい。

【0058】

ステップS17において、ステップS16の処理で表示された入力ウィンドウ61に呼応して、ユーザが「見る」ボタンを選択したか否かが判定される。ユーザが「見る」ボタンを選択したと判定された場合、ステップS18に進む。

【0059】

ステップS18において、関連情報提示部10は、イベント抽出部8を介してデータベース問い合わせ部9から供給された関連情報をデスクトップに表示させる。

【0060】

なお、関連情報として表示される情報は、キーワードが付与された所定のデータベースに蓄積された情報であれば、Webページのタイトルでなくてもかまわない。例えば、所定のデータベースに蓄積されている情報のインデックスを表示するようにして、ユーザのアクセス指令に対応して、当該インデックスの更に詳細な情報を表示させるようにしてもよい。

【0061】

ステップS19において、ステップS18の処理により関連情報として表示されたWebページのタイトルに対して、ユーザがアクセスを指令したと判定された場合、ステップS20に進む。ステップS20において、WWWブラウザが起動され、当該Webページへのアクセスが行われる。

【0062】

ステップS19において、ステップS18の処理により関連情報として表示されたWebページのタイトルに対して、ユーザが記録を指令したと判定された場合、ステップS21に進む。ステップS21において、当該WebページのタイトルおよびURLはスクラップ帳ウィンドウ54（図20）に記録される。なお、ステップS19において、ステップS18の処理により関連情報として表示されたWebページのタイトルに対して、ユーザから何の指令もなされずに所定の時間が経過したと判定された場合、ステップS20またはステップS21の処理はスキップされて、ステップS11に戻る。

【0063】

なお、ステップS17において、ユーザが「見る」ボタンを選択しないと判定された場合、ステップS18乃至ステップS21の処理はスキップされて、ステップS11に戻る。

【0064】

10

20

30

40

50

ここで、関連情報提示処理に関して、イベント発生に対応する電子メールを効率よく取得する手法について説明する。まず、大多数のメーラがメールの保持形式に関し、次のような4つの特徴を有していることに着目する。第1に、メーラにおける1つのフォルダは、パーソナルコンピュータにおける1つのメールボックスファイルに対応していること。第2に、新たに受信したメールは、特定のフォルダに格納されるようになっており、パーソナルコンピュータでは当該フォルダに対応するファイルの末尾に追加され、このとき、1つのファイルには一般に複数のメールが含まれるので、メールとメールとの境界に、特定の文字列パターン(メーラによって異なる)からなる行が挿入されていること。第3に、送信したメールの記録も、同様の形式で特定のファイルに保存されること。第4に、一般に、送受信したメールが含まれるファイルはサイズが大きいこと(数百KB乃至1MB)。

10

**【0065】**

以上の第1乃至第4の特徴を考慮して、次に手順によってイベント発生に対応するメールを取得する。始めに、メールボックスファイルの更新日時を検出し、新たなメールが追加されたか否かを判断する。次に、新たにメールが追加されたメールボックスファイルを末尾から先頭方向に1行ずつ操作して、メールとメールとの境界を示す特定の文字列を検出する。境界を示す文字列を検出した場合、その位置からメールボックスファイルの末尾までデータを抽出する。

**【0066】**

このような手法を用いることにより、イベント発生に対応する電子メールを効率的に取得することが可能となる。

20

**【0067】**

次に、関連情報提示処理に関して、同一のメールに対して何度も関連情報を提示しないようにする手法について説明する。まず、関連情報を提示したメールのメッセージIDを記録するためのデータ構造を設定する。そして、イベントが発生した場合、当該イベントに対応するメールのメッセージIDを取得して、設定したデータ構造と比較する。当該データ構造の中に同じメッセージIDがある場合、当該メールに対しては既に関連情報を提示しているので、関連情報を提示しない。当該データ構造の中に同じメッセージIDがない場合、当該メールに対しては関連情報を提示していないので関連情報を提示し、メッセージIDを当該データ構造に記録する。

**【0068】**

このような手法を用いることにより、同一のメールに対して何度も関連情報を提示するような事態の発生を抑止することが可能となる。

30

**【0069】**

次に、関連情報提示処理のうち、ステップS17乃至ステップS21の処理に関連するエージェントの動作および台詞等について、図7および図8のフローチャートに従って説明する。なお、以下に説明する各ステップの処理のうち、エージェントの動作および台詞等の処理については、主に、機能ブロックのうちのエージェント制御部11が実行する。

**【0070】**

例えば、エージェントプログラムが起動されている状態においてメーラが起動された場合、ステップS31において、図11に示すように、メーラのウィンドウ(以下、メーラウィンドウと記述する)51の表示と重畳しない位置に、エージェント52が登場する(表示される)。

40

**【0071】**

なお、エージェント52の登場は、例えば、図12(A)乃至図12(D)に示す画像が順次表示されることによって、エージェントが前転しながらデスクトップ上に出現する動画が表現される。エージェント51の登場とともに、エージェント52の台詞として吹き出し53と、保存されている関連情報が一覧表示されているスクラップ帳ウィンドウ54(後述)が表示される。吹き出し53の中には、例えば、図13に示すように、登場の挨拶「おはよう、SAITOさん!」と、自己紹介「ぼく、alfだよ。」の台詞が表示される。

**【0072】**

50

また、吹き出し53の表示と同期して、吹き出し53に表示された台詞と同じ意味を持つ他の言語（例えば、英語の場合、“Good morning, SAITO”、“I'm Alf”）の音声信号が音声合成部（不図示）によって合成されて出力される。なお、吹き出し53に表示された言語（いまの場合、日本語）と音声信号の言語（いまの場合、英語）を同じ言語に統一してもよい。なお、以降に表示される吹き出し53にも対応する音声信号が同期して出力される。

【0073】

その後、エージェント52の表示は、ステップS32において、例えば、図14に示すように、待機中の様子（手を後に組み、つま先を上下させる）を示す動画に推移する。

【0074】

ステップS33において、メーラが終了されたか否かが判定される。メーラが終了されていないと判定された場合、ステップS34に進む。ステップS34において（上述した図5のステップS12に対応する）、メーラに対してユーザから何らかのコマンド（メールの送受信、メールの編集、あるいは関連情報の編集等）が入力されたか否かが判定される。何らかのコマンドが入力されたと判定された場合、ステップS35に進む。メーラはコマンドに対応する処理を開始する。

【0075】

ステップS35において、入力されたコマンドが判定される。メールの送信、受信、または編集のコマンドが入力されたと判定された場合、ステップS36に進む。

【0076】

ステップS36において、エージェント52の表示は、待機中の様子から、図15に示すように、作業中の様子（手足を激しく移動する）を示す動画に推移する。この間、図5のステップS13乃至S15の処理（ユーザに推薦する関連情報を選択する処理）が実行される。

【0077】

ステップS37において、コマンドに対応して開始されたメーラの処理（例えば、メール送信など）が継続中であるか否かが判定され、メーラの作業中の処理が終了するまで待機される。待機の間、エージェント52の表示は作業中の状態である。メーラの処理が継続中ではない、すなわち、コマンドに対応して開始されたメーラの処理が終了したと判定された場合、ステップS38に進む。

【0078】

ステップS38において、再度、メーラが終了されたか否かが判定される。メーラが終了されていないと判定された場合、ステップS39に進む。

【0079】

ステップS39（図5のステップS16に対応する）において、エージェント52の吹き出し53には、先程までのメーラの処理がメール送信であった場合、例えば、「今、Aさんにメール送ったけど、某月某日にAさんと（タイトル）について話してたよね。その中でできた（キーワード）について、関連するページを見つけたよ。見てみる？」と表示される。

【0080】

また例えば、先程までのメーラの処理がメール受信であった場合、台詞の前半が「今、Aさんからメールがきたけど」となり、先程までのメーラの処理がメールの編集であった場合、台詞の前半が「今、Aさんにメール書いてるけど」となる。

【0081】

なお、「某月某日にAさんと（タイトル）について話してたよね。」の部分が当該関連情報が選択された理由に相当するが、関連情報の選択理由の表示を、ステップS39において実行せずに、後述するステップS42の処理（関連情報の表示）の後に表示するようにしてもよい。また、関連情報の選択理由の表示をユーザの指示により任意のタイミングで実行するようにしてもよい。

【0082】

10

20

30

40

50

吹き出し53に隣接する位置には、図16に示すように、入力ウィンドウ61が表示される。入力ウィンドウ61には、図17に示すように、関連情報の表示を指示するときに選択する「見る」ボタン、関連情報を表示させない時に選択する「見ない」ボタン、関連情報が選択された背景（選択理由）の再表示を指示するときに選択する「背景をもう一度教えて」ボタンが表示される。

**【0083】**

入力ウィンドウ61が表示された状態で、ステップS40において、エージェント52の表示は、図14に示した待機中の様子を示す動画に推移し、ステップS41（図5のステップS17に対応する）において、入力ウィンドウ61の中の「見る」ボタン、「見ない」ボタン、または「背景をもう一度教えて」ボタンのいずれがユーザにより選択されたかが判定される。入力ウィンドウ61の「見る」ボタンが選択されたと判定された場合、ステップS42に進む。

10

**【0084】**

ステップS42（図5のステップS18に対応する）において、図18および図19に示すように、関連情報として推薦URL71が表示され、エージェント52の表示は、表示された推薦URL71を指し示す動画に推移する。吹き出し53には、台詞「どう？」が表示される。推薦URL71には、通常、推薦されるWebページのタイトルが表示され、推薦URL71の上にマウスカーソルが置かれたときだけURLも重畳して表示される。推薦URL71は、マウスカーソルでドラッグすることにより移動可能である。

**【0085】**

20

ステップS43（図5のステップS19に対応する）において、表示した推薦URL71に対するユーザのコマンドが検出される。複数表示される個々の推薦URL71に対するユーザのコマンドとしては、記録、アクセス、消去等がある。

**【0086】**

推薦URL71に対する記録コマンドは、例えば、記録する推薦URL71をスクラップ帳ウィンドウ54までドラッグアンドドロップする方法や、マウスで右クリックし、表示されるメニューの中から記録を選択する方法などが考えられる。アクセスコマンドや消去コマンドについても同様に、WWWブラウザのアイコンやゴミ箱のアイコンにドラッグアンドドロップする方法や、マウスで右クリックし、表示されるメニューの中から選択する方法、あるいはクリッカブルにする方法などが考えられる。

30

**【0087】**

ステップS43で、推薦URL71に対する記録コマンドが検出された場合、ステップS44（図5のステップS21に対応する）において、エージェント52の表示は、図18に示すように頷く動作に推移する。スクラップ帳ウィンドウ54の中に、記録が指示された推薦URL71に対応するWebページのタイトルが追加される。

**【0088】**

また、ステップS43で、推薦URL71に対するアクセスコマンドが検出された場合、ステップS44（図5のステップS20に対応する）において、エージェント52の表示は、図21（A）および図21（B）に示すように、笑顔となって喜ぶ様子に推移する。吹き出し53には、「わーい」と表示され、対応する音声信号が出力される。

40

**【0089】**

また、ステップS43で、推薦URL71に対する消去コマンドが検出された場合、ステップS44において、エージェント52の表示は、図20（A）および図22（B）に示すように、泣き顔となって悲しみ失望した様子に推移する。吹き出し53には、「だめかぁ」と表示され、対応する音声信号が出力される。

**【0090】**

ステップS44の処理が終了した後、ステップS32に戻り、それ以降の処理が繰り返される。

**【0091】**

なお、ステップS41において、入力ウィンドウ61の「見ない」ボタンが選択されたと

50

判定された場合、ステップS 3 2に戻り、それ以降の処理が繰り返される。また、ステップS 4 1において、入力ウィンドウ6 1の「背景をもう一度教えて」ボタンが選択されたと判定された場合、ステップS 3 9に戻り、ステップS 3 9乃至S 4 1の処理が繰り返される。

**【0092】**

ステップS 3 8において、メーラが終了されたと判定された場合、ステップS 4 5に進む。ステップS 4 5において、吹き出し5 3には、「え、そんなぁ」と表示され、対応する音声信号が出力される。ステップS 4 6において、エージェント5 2の表示は消失される(図2 7を参照して後述する)。

**【0093】**

ステップS 3 5において、関連情報の編集を指示するコマンドが入力されたと判定された場合、ステップS 4 7に進む。ステップS 4 7において、関連情報編集用ウィンドウ(不図示)が表示され、エージェント5 2の表示は、待機中の様子から、図1 9と同様に、関連情報編集用ウィンドウを指し示す様子に推移する。その後、ユーザが関連情報編集用ウィンドウに対して編集のための入力を開始すると、ステップS 4 8において、エージェント5 2の表示は、関連情報編集用ウィンドウを指し示す様子から、図1 5に示した作業中の様子を示す動画に推移する。

**【0094】**

ステップS 4 9において、関連情報編集処理が継続中であるか否かが判定され、関連情報編集処理が終了するまで待機する。この待機の間、エージェント5 2の表示は作業中の状態である。関連情報編集処理が継続中ではない、すなわち、関連情報編集処理が終了したと判定された場合、ステップS 5 0に進む。

**【0095】**

ステップS 5 0において、エージェント5 2の表示は、図2 0と同様に、頷く様子に推移する。吹き出し5 3には、「変更したよ」と表示され、対応する音声信号が出力される。この後、ステップS 3 2に戻り、それ以降の処理が繰り返される。

**【0096】**

ステップS 3 4において、メーラに対してユーザから何らかのコマンドが入力されない状態が所定の時間以上継続した場合、ステップS 5 1に進む。ステップS 5 1において、エージェント5 2の表示は、所定の時間が経過する毎に、移動の状態、遊びの状態、睡眠の状態に順次推移する。この待機中の処理の詳細について、図8のフローチャートを参照して説明する。

**【0097】**

ステップS 6 1において、エージェント5 2の表示は、図1 4に示した待機中の状態から、図2 3または図2 4に示した画像を用いて表現される移動の状態に推移する。

**【0098】**

エージェント5 2の移動は、表示されているウィンドウと重畳しないようにデスクトップ上を横方向あるいは縦方向に行われる。なお、アクティブであるウィンドウ(いまの場合、メーラウィンドウ5 1)を検出して、その周囲を横方向あるいは縦方向に行うようにしてもよい。エージェント5 2がデスクトップ上を横方向(例えば、右方向)に移動するときには、例えば図2 3(A)乃至図2 3(D)に示す画像が順次用いられることにより、いわゆるワープしたかのような動画表現が実現される。具体的には、エージェント5 2の表示は、移動開始位置において、図2 3(A)に示すように、体の向きが移動する方向とされ、その後、指向している方向にジャンプすると、図2 3(B)に示すように、頭部から順に消滅して行く。そして、移動終了位置において、図2 3(C)に示すように、脚部から順に表示されて、最終的には、図2 3(D)に示すように全身が表示される。

**【0099】**

エージェント5 2がデスクトップ上を上下方向に移動するときには、例えば図2 4(A)乃至図2 4(G)に示す画像が順次用いられる。すなわち、エージェント5 2は、移動開始位置において、図2 4(A)に示すように、自身の尻尾(先端がコンセントプラグの形

10

20

30

40

50

状をしている)を手で握り、図24(B)に示すように、尻尾の先端を頭上付近に差し込む。この後、エージェント52の表示は、図24(C)、図24(D)に順次示すように、体の下部から徐々にロープに変身し、図24(E)に示すように、1本のロープになってその状態で移動終了位置まで移動する。移動終了位置においては、図24(F)、図24(G)に順次示すように、頭部から順に復元されて、最終的に全身が表示される。

【0100】

このように、エージェント52の移動を、いわゆるワープによって表現したり、1本のロープに変身させて表現したりすることにより、移動中を表現するために使われるリソース(メモリなど)の量を軽減させることが可能となる。

【0101】

図8に戻り、ステップS62において、イベント(メールの送受信、メールの編集、あるいは関連情報の編集等を指示するコマンドの入力)が発生したか否かが判定される。イベントが発生していないと判定された場合、ステップS63に進む。

【0102】

ステップS63において、エージェント52の表示が移動の状態に推移した後、所定の時間が経過したか否かが判定され、所定の時間が経過したと判定されるまで、ステップS62およびステップS63の処理が繰り返される。ステップS63において、所定の時間が経過したと判定された場合、ステップS64に進む。

【0103】

ステップS64において、エージェント52の表示は、移動の状態から、例えば図25に示す画像で表現される遊びの状態に推移する。図25(A)は、エージェント52が蛇と戯れている状態を示しており、図25(B)は、エージェント52が尻尾の先端を上方に差し込み、そこを支点としてぶら下がり揺れている状態を示している。

【0104】

ステップS65において、イベントが発生したか否かが判定される。イベントが発生していないと判定された場合、ステップS66に進む。ステップS66において、エージェント52の表示が遊びの状態に推移した後、所定の時間が経過したか否かが判定され、所定の時間が経過したと判定されるまで、ステップS65およびステップS66の処理が繰り返される。ステップS66において、所定の時間が経過したと判定された場合、ステップS67に進む。

【0105】

ステップS67において、エージェント52の表示は、遊びの状態から、例えば図26に示す画像で表現される睡眠の状態に推移する。ステップS68において、イベントが発生したか否かが判定され、イベントが発生したと判定されるまで待機される。ステップS68において、イベントが発生したと判定された場合、図7のステップS35にリターンし、それ以降の処理が実行される。

【0106】

なお、ステップS62、またはステップS65において、イベントが発生したと判定された場合も同様に、図7のステップS35にリターンし、それ以降の処理が実行される。

【0107】

また、図8のフローチャートには図示していないが、待機中の処理が実行されている最中において、メーラが終了されたと判定された場合は、ステップS46にリターンする。同様に、図7のステップS33において、メーラが終了されたと判定された場合はステップS46に進む。

【0108】

ステップS46において、エージェント52の表示は、例えば図27(A)および図27(B)に示す画像を用いた表現により、手を振りながら背を向けて遠方に立ち去り、その姿が徐々に小さくなってやがて消失する。

【0109】

なお、エージェント52の消去とともに、吹き出し53、スクラップ帳ウィンドウ54、

10

20

30

40

50

および推薦URL 7 1 等の表示も消去される。

【0110】

以上のように、本発明によれば、電子メール等の文書からキーワードを抽出し、関連情報を推薦する一連の処理に対応して、エージェントが動作するので、エージェント 5 2 に対して信頼性が感じられるようになる。

【0111】

ところで、上述したエージェントの動作および吹き出しの中の台詞の表示、並びに、表示された台詞に対応する音声信号の出力については、本発明のエージェントプログラムだけでなく、他のアプリケーション、例えば、ゲームやワードプロセッサのヘルプ画面等に適用することが可能である。さらに、テレビジョン受像機、ビデオカメラ、カーナビゲーション等のディスプレイに表示されるキャラクタに適用することも可能である。

10

【0112】

また、同一のパーソナルコンピュータを複数のユーザが操作する場合、エージェントの種類を複数用意して、ユーザ毎に表示されるエージェントの種類を違えるようにしてもよい。さらに、同一のユーザが複数のパーソナルコンピュータ上で当該エージェントプログラムを利用する場合、異なるパソコン上においても同じエージェントが表示されるようにしてもよい。

【0113】

次に、蓄積部 1 によるデータベースの作成、更新のタイミングについて説明する。データベースは、図 3 のフローチャートを参照して上述したように作成されるが、次のような第 1 乃至第 3 の状況になった場合、データベースの更新が必要である。

20

【0114】

すなわち、第 1 の状況として、データベースが作成または更新されてから所定の期間が経過している場合、データベース内の関連情報が古くなってしまっているので更新が必要である。第 2 の状況として、データベースに蓄積されている関連情報のうちの所定の割合が提示済みとなった場合、データベース内の関連情報を繰り返し提示してしまったり、提示する関連情報が不足してしまうので更新が必要である。第 3 の状況として、特徴抽出に用いた文書が、例えば電子メールである場合、電子メールの送受信を繰り返していると、その文書の内容が変化するので更新が必要である。

【0115】

なお、データベースの更新が必要である状況になった場合、ユーザに対して更新を指示するように促すこともできるし、ユーザに対する促しを実行せずに、自動的にデータベースを実行するように設定することも可能である。また、ユーザが指定する任意のタイミングで更新することも可能である。

30

【0116】

前記第 1 乃至第 3 の状況を考慮したデータベース更新処理について、図 9 のフローチャートを参照して説明する。このデータベース更新処理は、エージェントプログラムが実行する処理のうちの一つであり、エージェントプログラムの起動とともに開始され、エージェントプログラムが終了されるまで繰り返し実行される。

【0117】

ステップ S 8 1 において、既にデータベースを作成済みであるか否かが判定される。データベースを作成済みであると判定された場合、ステップ S 8 2 に進み、データベースを作成済みではないと判定された場合、ステップ S 8 6 (図 3 のデータベース作成処理に相当する)。

40

【0118】

ステップ S 8 2 において、作成済みのデータベースは更新が必要であるか否かを判断され、更新が必要であると判断されるまで待機する。この判断基準は、例えば、図 10 に示すようなユーザインタフェースの画面を用いて予めユーザが設定するものとする。図 10 の例では、4 つの条件が示されており、左端の がユーザによってチェックされた場合、対応する条件が有効となる。なお、1 番目の条件では回数が設定可能とされており、3 番目

50



の条件では日数が設定可能とされている。

【0119】

ステップS82において、更新が必要であると判断された場合、ステップS83に進む。ステップS83において、データベースを自動的に更新するように設定されているか否かが判定される。自動的に更新するように設定されていないと判定された場合、ステップS84に進む。反対に、ステップS83において、自動的に更新するように設定されていると判定された場合、ステップS84はスキップされる。

【0120】

ステップS84において、データベースの更新が必要である旨がユーザに通知され、さらに、当該通知に対応して、ユーザから更新の指示がなされたか否かが判定される。ユーザから更新の指示がなされた場合、ステップS85に進む。反対に、ステップS84において、ユーザから更新の指示がなされないと判定された場合、ステップS82に戻り、それ以降の処理が繰り返される。

【0121】

ステップS85において、既に作成されているデータベースが更新される。具体的には、新たに追加された文書や内容が変更された文書が選別され、選別された文書から特徴抽出が行われて分析結果が蓄積され、分析結果が用いられて関連情報が検索される。

【0122】

より具体的には、文書取得部2乃至文書内容処理部4によって、電子メールのメールボックスファイル（例えば、mbx等の拡張子が付加されていることが多い）を検出し、その更新日時を取得して、以前に取得した更新日時と比較して、異なる日付であれば、ファイルが更新されていると判断され、追加または変更された部分が抽出される。この場合、メールのグループ化、ヘッダの解析、形態素解析、特徴ベクトル算出等、一連のファイル内の分析が行われ、得られる検索語が関連情報検索部6に供給される。

【0123】

ただし、メールグループが変化せず、分析の結果、更新以前の検索語と更新後の検索語が同じであれば、重み等の計算値だけを変更し、関連情報検索部6による関連情報の検索を実行しないようにしてもよい。また、関連情報をインターネット等のネットワーク上で検索する場合、ネットワークに接続されているか否かが検出され、ネットワークに接続されていない場合、関連情報の検索を行わず、それ以後、ネットワークに接続されたときに関連情報を検索するか否かをユーザに問うようにしてもよい。

【0124】

「同じ関連情報を何度も推薦（提示）しないようにするために、あるメールグループの関連情報を、所定の回数以上推薦したら更新が必要と判断する」との条件に関連して、取得したメールと類似性の高いメールグループを選択する際に、同じメールグループから何度も推薦を行わないように、次のような処理を行う。

【0125】

メールグループ自体に推薦の優先度の順位を付与し（例えば、メールグループ内での特徴語の重みの最大値を当該メールグループの優先度の値とし、優先度の値を降順に並べたものを優先度の順位として付与する）、一度推薦を行ったメールグループを優先順位の最後尾に並び替えるようにする。このようにすることによって、類似度の範囲内にあるメールグループでも、同じメールグループから推薦する頻度が減少する。また、優先順位の変更だけなので、関連情報を大量に検索して準備しておけば、なるべく同じメールグループからの推薦が減り、且つ、情報自体も不足することなく用いることができる。

【0126】

これに関連して、特徴抽出に用いる文書の量に応じて、類似文書を抽出する際の範囲を変化させることができるようにする。具体的には、特徴抽出する文書量または文書のデータサイズに応じて何段階かの類似度の範囲を設定する。例えば、文書量が10ファイル以内である場合は類似度を0.01以上、11ファイル以上50ファイル未満の場合は類似度を0.03以上、51ファイル以上である場合は類似度を0.05以上とする。または、

10

20

30

40

50

文書の容量が500Kb以内である場合は類似度を0.01以上、500Kb以上である場合は類似度を0.02以上とする。

【0127】

そして、事前に設定された範囲の文書のうち、優先度の高い文書から検索された関連情報を提示するようにする。このようにすると、文書量の減少により、データベースの内容が更新されると、類似度の範囲が変化し、類似度の範囲が狭すぎて関連情報が不足したり、反対に、類似度の範囲が広すぎてユーザにとってあまり関連性が明確でない関連情報が提示されるような事態の発生を抑止することが可能となる。

【0128】

以上説明したように、データベース更新処理においては、追加された文書や変更された文書だけを処理の対象とするので、データベース作成処理を再度実行する場合に比べて、処理時間が短縮される。

【0129】

なお、本発明のエージェントプログラムは、上述したようにメーラによって送受信される文書やワードプロセッサプログラムで編集される文書の他、例えば、チャット、電子ニュース、電子掲示板等の文書や音声信号をテキスト化した文書など、属性情報としてタイムスタンプが付与されている文書に対応して動作するようにさせることができる。

【0130】

次に、本発明のプログラム格納媒体について説明する。上述した一連の処理は、ハードウェアにより実行させることもできるが、通常、ソフトウェアにより実行させる。一連の処理をソフトウェアにより実行させる場合には、そのソフトウェアを構成するプログラムが、専用のハードウェアに組み込まれているコンピュータ、または、各種のプログラムをインストールすることで、各種の機能を実行することが可能な、例えば汎用のパーソナルコンピュータなどに、プログラム格納媒体からインストールされる。

【0131】

コンピュータにインストールされ、コンピュータによって実行可能な状態とされるプログラムを格納するプログラム格納媒体は、図2に示すように、コンピュータとは別に、ユーザにプログラムを提供するために配布される、プログラムが記録されている磁気ディスク32(フロッピディスクを含む)、光ディスク33(CD-ROM(Compact Disc-Read Only Memory)、DVD(Digital Versatile Disc)を含む)、光磁気ディスク34(MD(Mini Disc)を含む)、もしくは半導体メモリ35などよりなるパッケージメディア、または、プログラムが一時的もしくは永続的に格納されるROM22や記憶部29を構成するハードディスクなどにより構成される。プログラム格納媒体へのプログラムの格納は、必要に応じてルータ、モデムなどのインタフェースを介して、ローカルエリアネットワーク、インターネット、デジタル衛星放送といった、有線または無線の通信媒体を利用して行われる。

【0132】

なお、本明細書において、プログラム格納媒体に格納されるプログラムを記述するステップは、記載された順序に従って時系列的に行われる処理はもちろん、必ずしも時系列的に処理されなくとも、並列的あるいは個別に実行される処理をも含むものである。

【0133】

【発明の効果】

以上のように、本発明によれば、自動的にデータベースを構築し、且つ、送受信などの対象とした文書に関連する関連情報をユーザに提示することが可能となる。

【図面の簡単な説明】

【図1】本発明を適用したエージェントプログラムの機能ブロックの一構成例を示す図である。

【図2】エージェントプログラムをインストールして実行させるパーソナルコンピュータの構成例を示すブロック図である。

【図3】エージェントプログラムのデータベース作成処理を説明するフローチャートである。

10

20

30

40

50

【図4】図3のステップS5の処理を説明するための図である。

【図5】エージェントプログラムの関連情報提示処理を説明するフローチャートである。

【図6】図5のステップS15の処理を説明するための図である。

【図7】エージェントの動作等を説明するフローチャートである。

【図8】図7のステップS51の待機中の処理の詳細を説明するフローチャートである。

【図9】エージェントプログラムのデータベース更新処理を説明するフローチャートである。

【図10】データベースを更新させる条件を入力するユーザインタフェースの表示例を示す図である。

【図11】デスクトップ上に表示されたエージェント52の表示例を示す図である。

10

【図12】エージェント52が登場するときの表示例を示す図である。

【図13】エージェント52の台詞である吹き出し53の表示例を示す図である。

【図14】エージェント52が待機中であるときの表示例を示す図である。

【図15】エージェント52が作業中であるときの表示例を示す図である。

【図16】デスクトップ上に表示された入力ウィンドウ61の表示例を示す図である。

【図17】入力ウィンドウ61の表示例を示す図である。

【図18】デスクトップ上に表示された推薦URL71の表示例を示す図である。

【図19】エージェント52が指示中であるときの表示例を示す図である。

【図20】デスクトップ上に表示されたスクラップ帳ウィンドウ54の表示例を示す図である。

20

【図21】エージェント52が喜びの状態であるときの表示例を示す図である。

【図22】エージェント52が悲しみの状態であるときの表示例を示す図である。

【図23】エージェント52が横方向に移動するときの表示例を示す図である。

【図24】エージェント52が縦方向に移動するときの表示例を示す図である。

【図25】エージェント52が遊びの状態であるときの表示例を示す図である。

【図26】エージェント52が睡眠の状態であるときの表示例を示す図である。

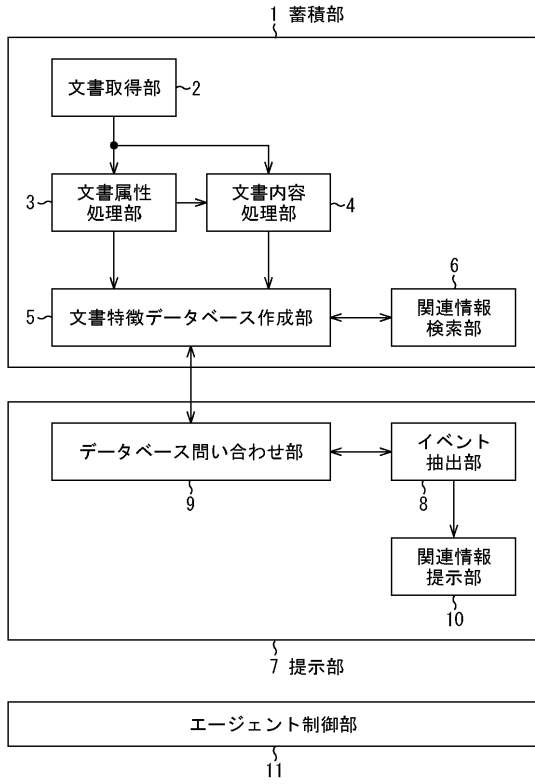
【図27】エージェント52が立ち去るときの表示例を示す図である。

【符号の説明】

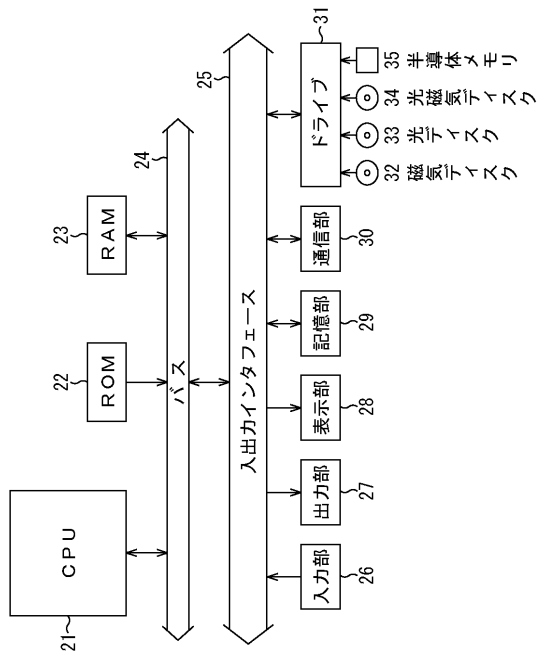
1 蓄積部, 2 文書取得部, 3 文書属性処理部, 4 文書内容処理部, 5 文書特徴データベース作成部, 6 関連情報検索部, 7 提示部, 8 イベント抽出部, 9 データベース問い合わせ部, 10 関連情報提示部, 11 エージェント制御部, 32 磁気ディスク, 33 光ディスク, 34 光磁気ディスク, 35 半導体メモリ, 52 エージェント, 53 吹き出し, 54 スクラップ帳ウィンドウ, 61 入力ウィンドウ, 71 推薦URL

30

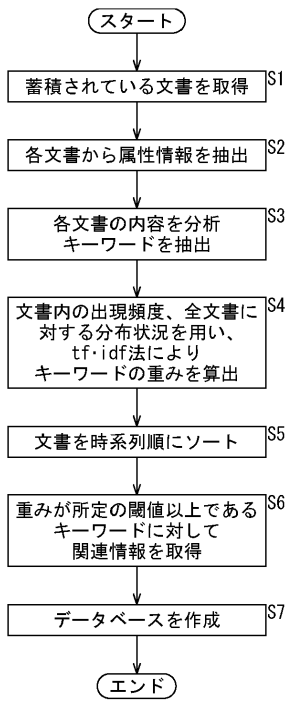
【図1】



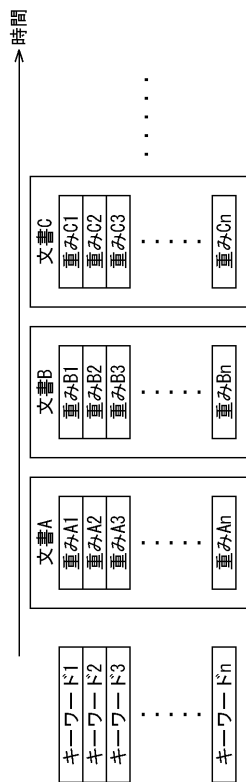
【図2】



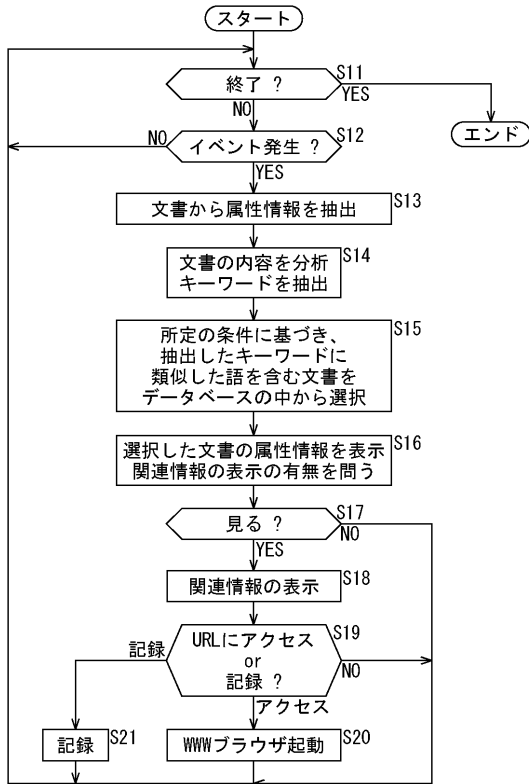
【図3】



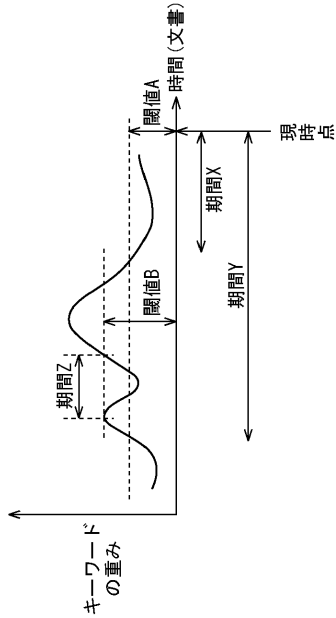
【図4】



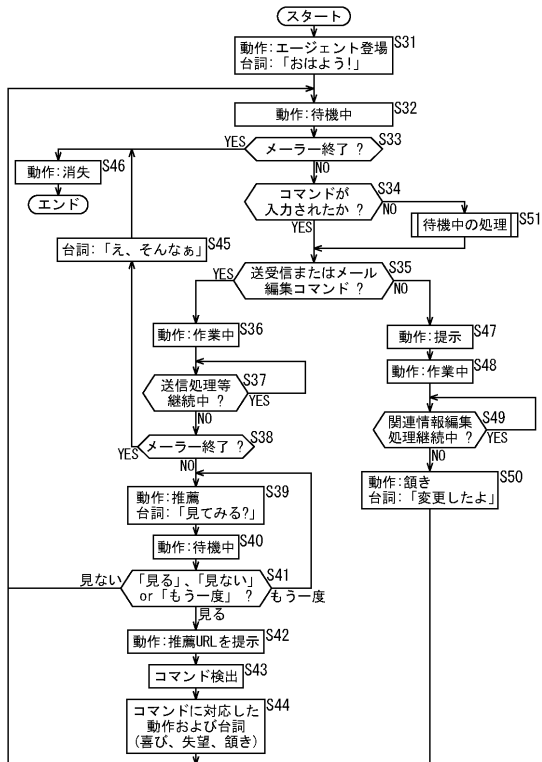
【図5】



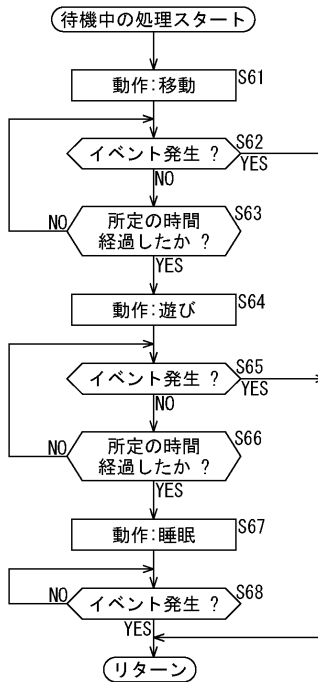
【図6】



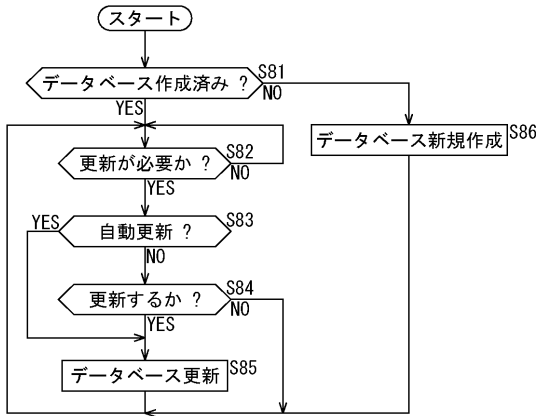
【図7】



【図8】



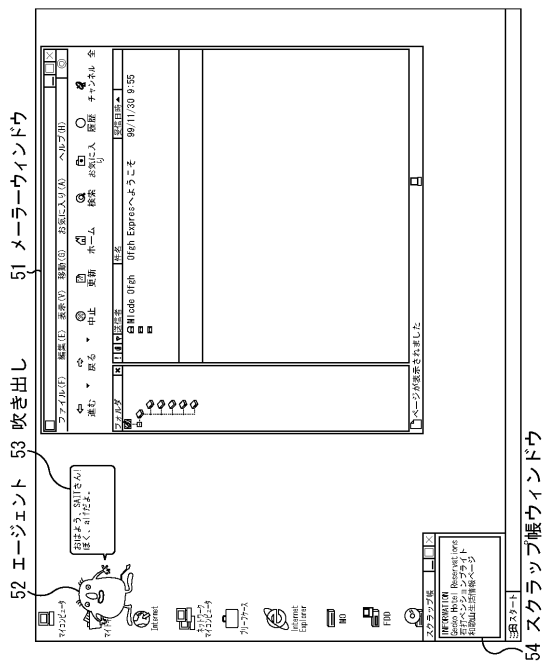
【図 9】



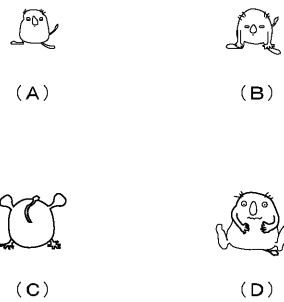
【図 10】

URLが  回行われたメールグループは更新必要と判断する  
 推奨URLを使い切ったメールグループは更新必要と判断する  
 最終更新から  日経過したメールグループは更新必要と判断する  
 新着メールの分析により、検索語が変更されたメールグループは更新必要と判断する

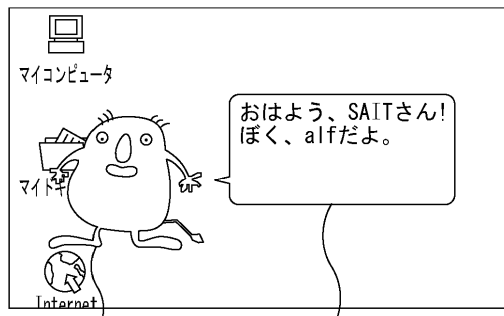
【図 11】



【図 12】



【図 13】



52 エージェント

53 吹き出し

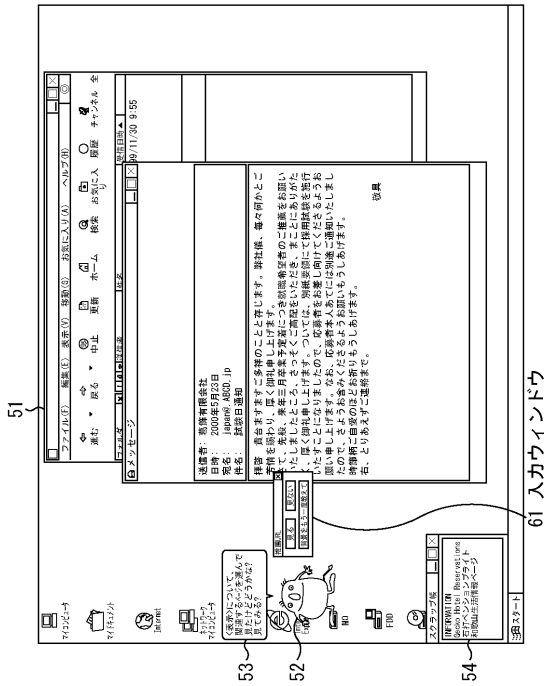
【図 14】



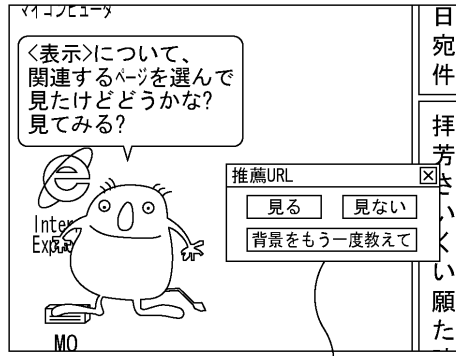
【 図 15 】



【 図 16 】

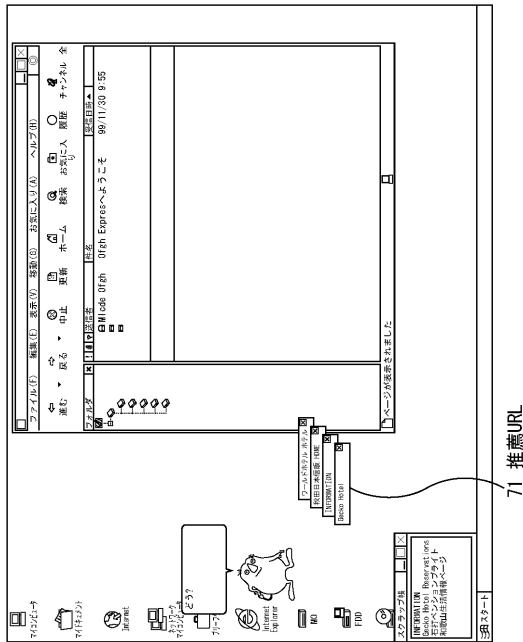


【 図 17 】

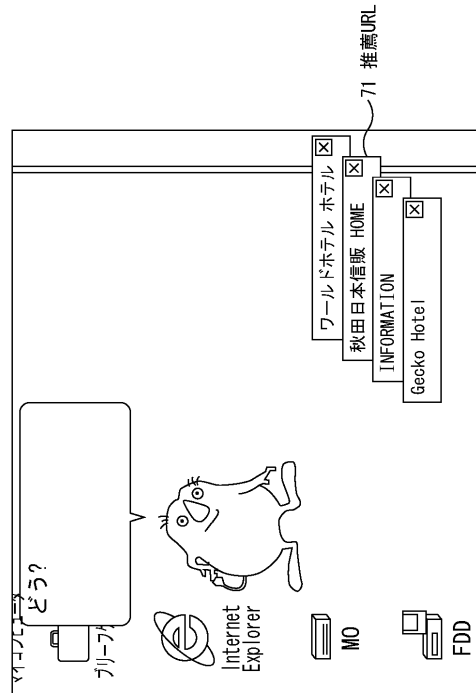


61 入力ウィンドウ

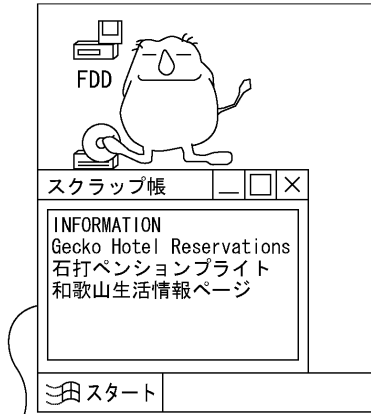
【 図 18 】



【 図 19 】



【 図 2 0 】



54 スクラップ帳ウィンドウ

【 図 2 1 】



【 図 2 2 】



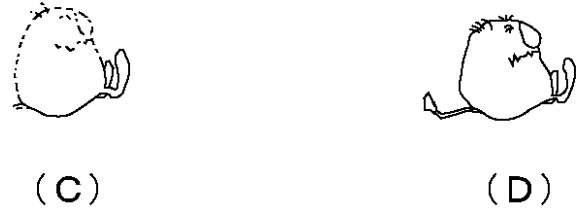
【 図 2 4 】



【 図 2 5 】



【 図 2 3 】



【 図 2 6 】



【 図 2 7 】





---

フロントページの続き

(51)Int.Cl. F I  
G 0 6 F 13/00 6 1 0 F  
G 0 6 F 13/00 6 2 0

(72)発明者 大村 和典  
東京都品川区北品川6丁目7番35号 ソニー株式会社内

審査官 長 由紀子

(56)参考文献 特開平11-065965(JP,A)  
特開平11-346267(JP,A)  
特開平10-093620(JP,A)  
特開平10-107833(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 17/30  
G06F 3/048  
G06F 13/00