

(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2004-280190
(P2004-280190A)

(43) 公開日 平成16年10月7日(2004.10.7)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
G06F 17/27	G06F 17/27	5B091
G06F 17/28	G06F 17/28	

審査請求 未請求 請求項の数 4 O L (全 10 頁)

(21) 出願番号	特願2003-67188 (P2003-67188)	(71) 出願人	000004226 日本電信電話株式会社 東京都千代田区大手町二丁目3番1号
(22) 出願日	平成15年3月12日 (2003.3.12)	(74) 代理人	100077481 弁理士 谷 義一
		(74) 代理人	100088915 弁理士 阿部 和夫
		(72) 発明者	佐野 直美 東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
		(72) 発明者	浅倉 剛 東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内

最終頁に続く

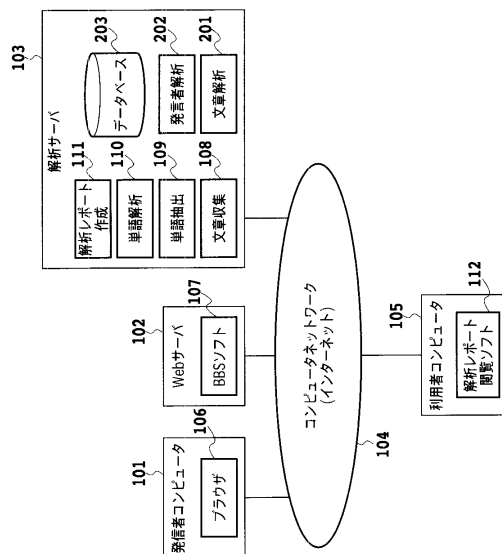
(54) 【発明の名称】 口コミ情報解析方法

(57) 【要約】

【課題】多数の人の口コミ情報を収集して、信頼性の高い解析レポートを提供することができる口コミ情報解析方法を提供する。

【解決手段】解析サーバ103は投稿された発言メッセージを複数のWebサーバ102から収集し、収集した発言の中から同一の発言者の発言を取り出し、取り出した発言に関して解析を行う。

【選択図】 図2



【特許請求の範囲】**【請求項 1】**

Webサーバが提供する発言を解析サーバにより収集し、該解析サーバにより収集した発言を解析する口コミ情報解析方法において、
前記Webサーバは複数あり、
前記解析サーバは発言の中に含まれる同一の発言者の発言を、収集した発言の中から取り出し、
当該取り出した発言に関して解析を行うことを特徴とする口コミ情報解析方法。

【請求項 2】

請求項 1 に記載の 口コミ情報解析方法において、前記発言者に関する解析の中には全発言数、日時ごとの発言数および発言期間のいずれかを含むことを特徴とする 口コミ情報解析方法。 10

【請求項 3】

請求項 1 または 2 に記載の 口コミ情報解析方法において、前記解析サーバにより収集した発言の意味内容を分析し、意味内容の種類ごとの発言数を取得する文章解析を前記解析サーバによりさらに実行することを特徴とする 口コミ情報解析方法。

【請求項 4】

請求項 1 から 3 のいずれかに記載の 口コミ情報解析方法において、前記Webサーバが提供する発言は発言者のコンピュータから投稿された発言であることを特徴とする 口コミ情報解析方法。 20

【発明の詳細な説明】**【0001】****【発明の属する技術分野】**

本発明は、掲示板に掲載された投稿メッセージを 口コミ情報として収集し、収集した 口コミ情報を解析する 口コミ情報解析方法に関する。

【0002】**【従来の技術】**

従来この種のシステムを代表的なシステム構成例を図 1 に示す。

【0003】

図 1 において、101 は発言者のコンピュータであり、パソコンなど通信機能を有するコンピュータが使用される。発言者のコンピュータ 101 にはWebページ（ホームページとも呼ばれる）を表示させるためのブラウザ106が搭載されている。102 はWebサーバであり、発言者のコンピュータ 101 に提供するWebページおよびBBSソフト（掲示板用ソフト）107を搭載している。BBSソフト107は、そのWebページを使用して発言者のコンピュータ 101 からの発言者の発言メッセージを受付け、複数の発言者のコンピュータ 101 から受付けた発言のメッセージを表示目的のために提供するソフトウェアのことである。 30

【0004】

103 は解析サーバであり、Webサーバ102に蓄積されている発言者のメッセージ（テキスト形態）を収集し、収集した発言のメッセージ（以下、発言と略記することがある）の解析を行う。 40

【0005】

解析サーバ103には、以下のプログラム、すなわち、発言を収集するためのプログラム108、収集した発言から形態素解析等により単語を抽出するためのプログラム109および抽出した複数の単語について多変量解析や統計的な解析を行うプログラム110および、その解析結果のレポートを作成するプログラム111が搭載されている。

【0006】

104 は複数の発言者のコンピュータ 101、Webサーバ 102 および解析サーバ 103 が接続されるコンピュータネットワーク、たとえば、インターネットである。

【0007】

105は上記レポートを利用する利用者のコンピュータである。解析サーバ103により作成されたレポートはCDROMに書き込まれて、このCDROMが利用者のコンピュータ105で閲覧される。利用者のコンピュータ105には閲覧のためのソフトが搭載されている。CDROMはコンピュータネットワーク104に接続されていない利用者のコンピュータ105に供されるが、発言者のコンピュータ101からも解析レポートを入手することができる。

【0008】

このシステム動作を図3を参照して説明する。

【0009】

WEBサーバ102は、BBSソフト107により発言者のコンピュータから寄せられる発言(メッセージ)を内部のハードディスク記憶装置等に蓄積する。 10

【0010】

利用者のコンピュータ105の解析レポート閲覧ソフト112を使用して、公衆電話回線(図1には不図示)を介して解析サーバ103に口コミ収集サービスを依頼する。解析サーバ103の文章収集プログラムがこの依頼を受付け、Webサーバ102に蓄積の発言(メッセージ)を要求する。

【0011】

Webサーバ102から送られ、解析サーバ103の文章収集プログラム108により受信された発言は単語抽出プログラム108に送られる。単語抽出プログラム109は形態素解析等により発言を構成する複数の単語を抽出する。 20

【0012】

その抽出結果は単語解析プログラム110に送られて、単語解析プログラム110により単語に関する統計解析が行われる。その統計解析結果が解析レポート作成プログラム111に送られて、レポートが作成される。

【0013】

作成レポートはCDROMに書き込まれ、このCDROMが郵送等で利用者に引渡される。利用者は利用者のコンピュータ105で解析レポート閲覧ソフト112を使用してCDROMに書き込まれたレポートを閲覧する。

【0014】

このようなシステムで提供される口コミ情報分析サービスとしては、野村総合研究所が提供するサービスが知られている(たとえば、非特許文献1)。 30

【0015】

このサービスではある特定の1つの掲示板に投稿されたメッセージについて以下の解析サービスを提供している。

(1)クチコミの文章の中で使用されている単語について、出現頻度の高い順にランキングを表示する。単語ランキングの表には、単語と品詞(名詞、形容詞、動詞)、出現頻度と全体に対する割合(%表示)が示される。ドリルダウン(関連語の詳細分析)も可能。

(2)選択されたグループのクチコミ情報の中から、数多く存在する話題を「主な話題」として抽出する。

主な話題は、次の2つの単語の組み合わせで表現されるので、 40

「 は です」(名詞・形容詞)

「 が する」(名詞・動詞)

2種類の話題が抽出される。

(3)単語のランキングに表示されるスコアは、判別の強さを表し、グループの特徴を表す度合いを示すので、複数のグループ間で特徴的な単語を比較することも可能となる。この機能を用いて、顧客の属性別グループや対象商品グループなど、各グループでの話題の違いを把握することが可能となる。

(4)選択されたグループとそのグループ内の単語間の関係を、「コレスポネンス分析」により2次元で表現する。マップ上には、各要素の相対的な位置が散布図で表示される。マッピングにより、各グループと関連性の高い単語やグループ間の類似度を、視覚的に 50

把握することができる。

(5) 選択されたグループのテキスト属性について分析する。テキスト属性として、メーカー名、性別、年代のほか、感情(喜怒哀楽)、内容分類(質問、使用后レポート、うわさ等)などといった情報がある。この機能を用いることで、属性間のクロス分析を行なうことが可能となる。

(6) 時系列モニタ

選択されたグループやその配下グループに分類されたテキストの時系列推移を表示する。テキスト件数のほか、出現頻度の上位単語の件数、およびスコアの3種類での分析が可能。

【0016】

このような機械的な解析システムを使用しない場合、ユーザは周知の検索エンジンを使用して特定の単語を含むホームページを検索する。検索結果の示すサイトを閲覧して、その掲載記事を目視確認する。

【0017】

【非特許文献1】

<http://www.true teller.net/kuchikomikdc/index.html>

【0018】

【発明が解決しようとする課題】

従来の機械的な解析システムを使用する口コミ情報解析は特定の単一の掲示板に投稿された複数のメッセージからそのメッセージに記載された単語を抽出して、単語数や単語の種類内容に基づいて口コミ情報の分析を行う。このため、複数のWEBサーバから口コミ情報を収集していないので、解析結果が多数の人の口コミ情報を反映していない場合が生じる。

【0019】

一方、検索エンジンを使用したホームページ(サイト)検索では、多数の検索結果が得られた場合、ユーザがその全てを閲覧することには膨大な時間を要する。

【0020】

そこで、本発明の目的は、上述の点に鑑みて多数の人の口コミ情報を収集して、信頼性の高い解析レポートを提供することができる口コミ情報解析方法を提供することにある。

【0021】

【課題を解決するための手段】

このような目的を達成するために、請求項1の発明は、Webサーバが提供する発言を解析サーバにより収集し、該解析サーバにより収集した発言を解析する口コミ情報解析方法において、前記Webサーバは複数あり、前記解析サーバは発言の中に含まれる同一の発言者の発言を、収集した発言の中から取り出し、当該取り出した発言に関して解析を行うことを特徴とする。

【0022】

請求項2の発明は、請求項1に記載の口コミ情報解析方法において、前記発言者に関する解析の中には全発言数、日時ごとの発言数および発言期間のいずれかを含むことを特徴とする。

【0023】

請求項3の発明は、請求項1または2に記載の口コミ情報解析方法において、前記解析サーバにより収集した発言の意味内容を分析し、意味内容の種類ごとの発言数を取得する文章解析を前記解析サーバによりさらに実行することを特徴とする。

【0024】

請求項4の発明は、請求項1から3のいずれかに記載の口コミ情報解析方法において、前記Webサーバが提供する発言は発言者のコンピュータから投稿された発言であることを特徴とする。

【0025】

10

20

30

40

50

【発明の実施の形態】

以下、図面を参照して本発明の実施形態を詳細に説明する。

【0026】

図2は本発明実施形態のシステム構成を示す。なお、図1の従来例と同様の箇所には同一の符号を付しているため、重複的な説明を避け、従来例との相違点を説明する。

【0027】

図2において、Webサーバ102には複数設置される。解析サーバ103には、文章解析プログラム201、発言者解析プログラム202が搭載され、内部のハードディスク記憶装置にはデータベース203が構築されている。

【0028】

文章解析プログラム201は収集した発言(メッセージ)の意味内容を解析するプログラムである。意味内容の解析は複数の単語とその複数の単語が示す意味内容を記載した対応表を参照することで、文章の意味内容の解析が行われる。文章解析プログラム201はさらに意味内容の種類ごとの発言数を取得する機能を有している。

【0029】

文章解析の意味内容については適宜定めればよく、たとえば、クレーム(苦情)や依頼、製品の評判等、種々の意味内容に関する文章解析を行うとよい。

【0030】

発言者解析プログラム202は同一の発言者(同一の名前の発言者)の発言を抽出した発言の中で検索して、取り出し、同一の発言者の発言を解析するプログラムである。具体的な解析の内容としては全発言数、日時ごとの発言数、発言を行った期間、発言を投稿したWebサーバごとの発言数、発言者の地位や職種に関する情報である。

【0031】

全発言数は上記検索で取得した発言の個数を計数することにより得られる。日時ごとの発言数については、発言メッセージの中に記載されている投稿日時をソーティング(並び換え)することにより日時順の発言を取得する。さらに、ソーティング後の発言の中の最も古い日時と最も新しい日時を取得することにより発言期間を取得することができる。発言者の地位や職種はたとえば、発言者のメッセージが「私は消防士です。」という「私」および「消防士」というように、発言メッセージの中に含まれている地位/職種を表す単語を検出する。

【0032】

このためには、予め解析サーバ103に地位/職種を表す単語を記載したテーブルを用意すればよく、単語抽出プログラム109により抽出した1つの文の中の単語と上記テーブルとを照合することにより発言者の地位や職種を判別することが可能となる。

【0033】

文章解析プログラム201および発言者解析プログラム202の解析結果が解析レポート作成プログラムによりレポートが作成される。

【0034】

203は文章収集プログラム108により複数のWebサーバから収集した発言を記憶し、上述の解析結果を記憶しておくデータベースである。

【0035】

この実施形態では利用者のコンピュータ105はコンピュータネットワーク104に接続されて、解析サーバ103で作成されたレポートをオンラインで受け取る。なお、従来例のようにCDROMの形態でオフラインでレポートを受け取るようにしてもよいことは勿論である。

【0036】

このようなシステムの動作手順を図4に示す。

【0037】

発言者コンピュータ101のブラウザ106を使用して発言者は掲示板に発言を行うと、その発言がWebサーバ102のBBSソフト107によりWEBサーバ102内の記憶

10

20

30

40

50

装置に蓄積される。利用者のコンピュータ105の解析レポート閲覧ソフト112により利用者から口コミ情報依頼が解析サーバ103に送られると、文章収集プログラム108はクローラーを使用して予め設定された複数のWebサーバ102から蓄積の発言を収集する。文章収集プログラム108は収集した発言を文章解析プログラム201に引渡す。

【0038】

文章解析プログラム203は発言者解析プログラム202と共同して発信日時・発言者・発言サイト、その他予め定められているテーマに関してそれぞれ共通の発言を取りまとめて、データベース203に蓄積する。なお、文章解析プログラム201では口コミ解析(従来と同様の口コミ解析も行われる)。

【0039】

解析結果1~3はデータベース203に蓄積された後、レポート作成プログラム111により使用されて、レポートが作成される。

【0040】

作成されたレポートは、利用者のコンピュータ105の解析レポート閲覧ソフト112により利用者に閲覧に供される。

【0041】

発言の収集形態として次の形態を適宜定めればよい。

(a) 解析サーバ103は予め定めたテーマについて、依頼がなくても一定周期で発言を収集し、データベース203に収集した発言を蓄積する。蓄積した発言に関する解析結果を利用者に提供する。

(b) 依頼者は、収集したいテーマ、たとえば、発言者名、地位/職種等を指定する。指定されたテーマに関する発言を文章収集プログラム108により収集して利用者に発言の解析結果を提供する。

(c) その他、依頼者、利用者に好適なテーマを予め受付け、定期的に解析結果を提供する。

などが考えられる。

【0042】

(具体的な実施形態)

具体的な実施形態の一例を図5に示す。

図5は、利用者がクレームの探索を行う場合の処理例を示す。利用者がレポート閲覧ソフト112の設定により、利用者のコンピュータ105の表示画面上に図5(A)のようなクレームの苦情発言の解析結果を表示させる。解析結果は図5(A)に示すようにサイトごとおよび週ごとの発言者数が3次元表示される。利用者が解析結果の種類を切り替えると、図5(B)に示すように発言者ごとおよび週ごとの発言数が3次元表示される。この画面の右側の条件設定画面で絞り込み条件、たとえば、クレームの発言の意味内容について絞り込みを行うように利用者が条件設定すると、その条件設定が利用者のコンピュータ105から解析サーバ103に伝えられる。解析サーバ103の中の条件設定に対応する解析プログラム、この例ではクレームの発言について発言者解析プログラム202が起動する。このためには、設定条件と、起動する解析プログラムの名前の対応表(あるいは関数)を定めておき、利用者のコンピュータ105から渡された条件に対応する解析プログラム名を取得する。

【0043】

発言者解析プログラム202により解析されたクレームに関する発言の解析結果が利用者のコンピュータ105に送られて、図5(C)に示すように利用者のコンピュータ105に表示される。

【0044】

図5(C)では、発言のポジティブ(肯定的)とネガティブ(否定的)との度合いが発言数と発言率についてグラフで表される例を示している。この画面を見ることにより利用者はポジティブな発言が多いかネガティブな発言が多いかを知ることができる。

【0045】

10

20

30

40

50

利用者が条件設定画面で、表示画面上の特定のエリアについての発言数（記事数）と全体の発言数（記事数）との比率を求めるように条件設定を行うと、解析サーバ103の対応プログラムにより比率が計算されて、その計算結果とグラフ表示を行うためのデータが解析サーバ103から利用者のコンピュータ105に送られる。利用者のコンピュータ105の解析レポート閲覧ソフト112により利用者のコンピュータ105の表示画面には図5（D）に示すように条件設定で求められた解析結果がグラフの形態で表示される。

【0046】

このようにして、利用者は予め用意されている条件設定を行って、所望の解析結果を解析サーバ103から取得して利用者のコンピュータ105の表示画面で見ることができる。図5の例では、図5（E）はメディアごとの記事の分布の表示例を示し、図5（F）は特定メディアにおける発声単語の出現回数のグラフの表示例を示している。

10

【0047】

上述の実施形態の他に次の形態を実施できる。

1) Webサーバと解析サーバとは上記実施形態においては別体としているが、同一システムでその機能を実現してもよい。

2) サーバ、コンピュータは汎用の周知の情報処理装置を使用すればよい。

【0048】

【発明の効果】

以上、説明したように、請求項1の発明によれば、Webサーバが提供する発言を解析サーバにより収集し、解析サーバにより収集した発言を解析する際に、解析サーバは発言の中に含まれる同一の発言者の発言を、収集した発言の中から取り出し、当該取り出した発言に関して解析を行うので、従来にはない発言者についての種々の発言の解析を行うことができる。収集の対象のWebサーバも複数なので、多数の発言を収集可能であり、発言の解析結果の信頼性も高くなる。さらに、従来から周知あるいは公知の解析手法と組み合わせることにより、同一の発言者に関する発言について内容を絞り込んで解析することも可能となる。

20

【0049】

請求項2の発明によれば、発言者に関する解析の中には全発言数、日時ごとの発言数および発言期間のいずれかを含むので、発言の影響度や口コミの時間的な変化を知ることができる。

30

【0050】

請求項3の発明によれば、前記解析サーバにより収集した発言の意味内容を分析し、意味内容の種類ごとの発言数を取得する文章解析を解析サーバによりさらに実行することにより製品の苦情等の度合いなどを知ることができる。

【0051】

請求項4の発明によれば、発言者のコンピュータから投稿された発言の影響度を知ることができる。

【図面の簡単な説明】

【図1】従来例のシステム構成を示すブロック図である。

【図2】本発明実施形態のシステム構成を示すブロック図である。

40

【図3】従来の口コミ解析サービスの処理手順を示す説明図である。

【図4】本発明実施形態の口コミ解析サービスの処理手順を示す説明図である。

【図5】（A）～（F）は利用者のコンピュータの表示例を示す説明図である。

【符号の説明】

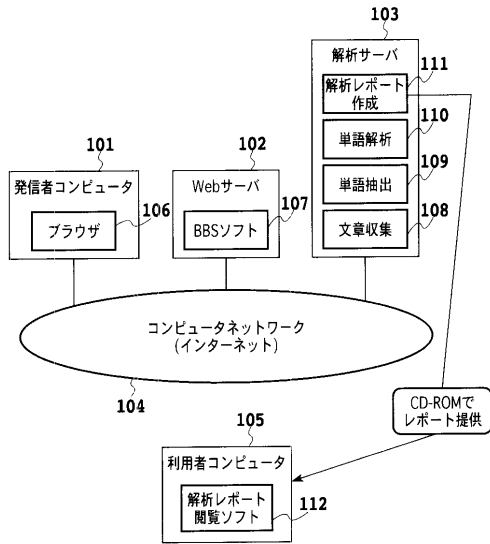
101、105 コンピュータ

102 Webサーバ

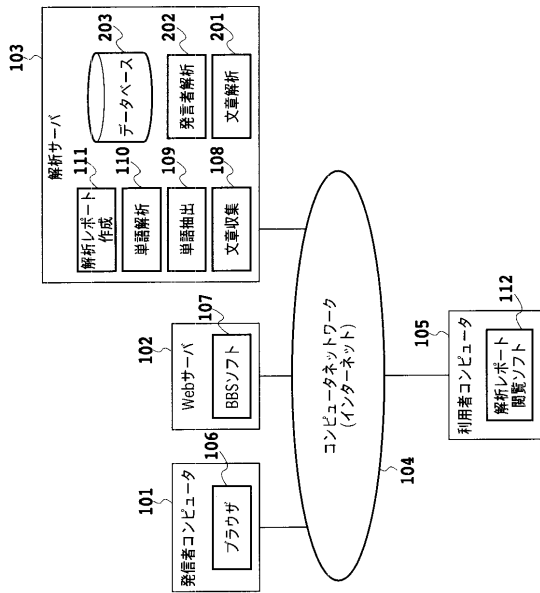
103 解析サーバ

104 コンピュータネットワーク

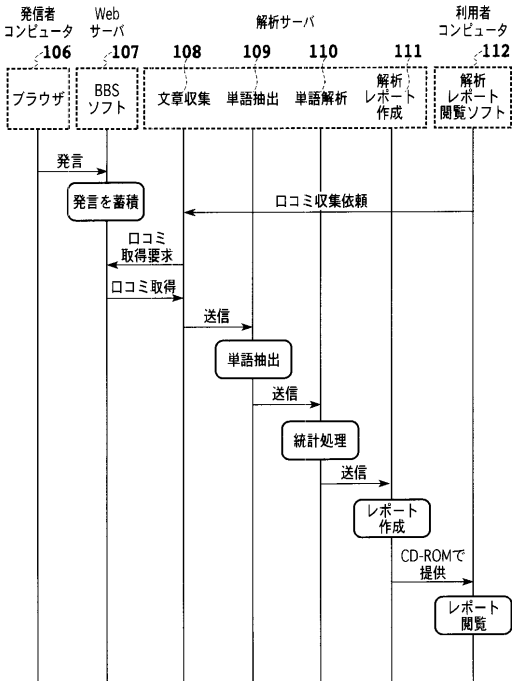
【図1】



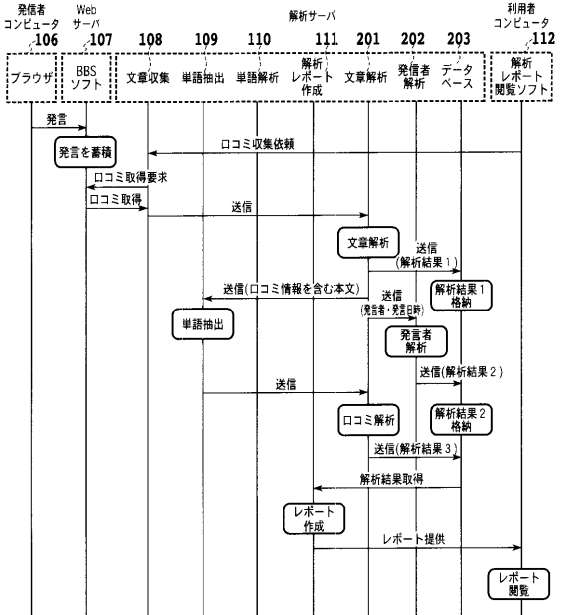
【図2】



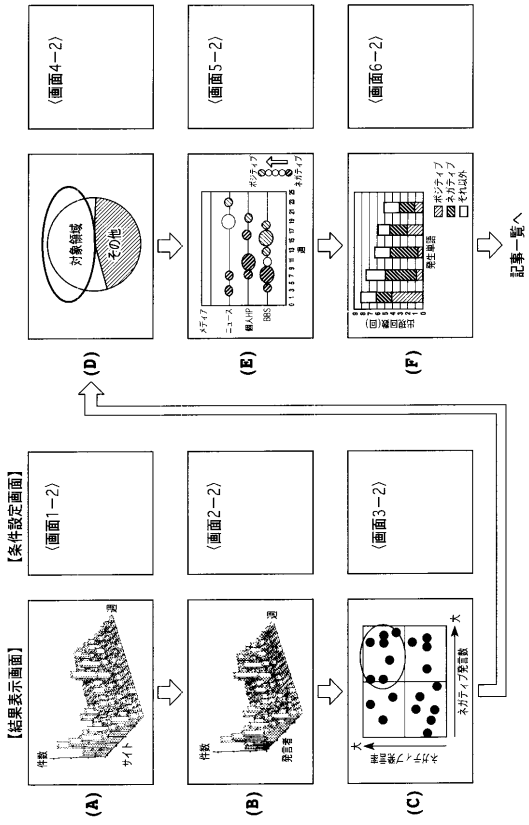
【図3】



【図4】



【図 5】



フロントページの続き

(72)発明者 齊藤 典明

東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内

Fターム(参考) 5B091 AA15 BA15 CA14