

(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2005-182804

(P2005-182804A)

(43) 公開日 平成17年7月7日(2005.7.7)

(51) Int.C1.⁷

F 1

テーマコード(参考)

G06F 17/30

G06F 17/30

340B

5B075

G06F 13/00

G06F 17/30

380D

5B085

G06F 15/00

G06F 13/00

540B

G06F 15/00

310A

審査請求 未請求 請求項の数 66 O L (全 29 頁)

(21) 出願番号 特願2004-362224 (P2004-362224)
 (22) 出願日 平成16年12月15日 (2004.12.15)
 (31) 優先権主張番号 10/739185
 (32) 優先日 平成15年12月19日 (2003.12.19)
 (33) 優先権主張国 米国(US)

(71) 出願人 000005496
 富士ゼロックス株式会社
 東京都港区赤坂二丁目17番22号
 (74) 代理人 100079049
 弁理士 中島 淳
 (74) 代理人 100084995
 弁理士 加藤 和詳
 (72) 発明者 ダニエルーアレキサンダー ビルス
 アメリカ合衆国 94131 カリフォル
 ニア州 サンフランシスコ シーザー チ
 ャベス ストリート 3943
 (72) 発明者 デイビッド エム. ヒルバート
 アメリカ合衆国 95126 カリフォル
 ニア州 サンノゼ マートル ストリート
 877

最終頁に続く

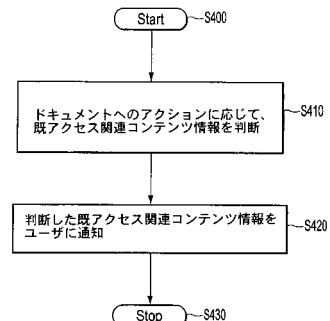
(54) 【発明の名称】コンテンツ情報を処理する方法、システム、及びプログラム

(57) 【要約】

【課題】コンテンツ情報を処理する方法、システム及びプログラムを提供する。

【解決手段】本発明のコンテンツ情報を処理する方法は、情報を含むドキュメントになされるアクションに応じて、及び、該情報のコンテクストに基づいて、関連する既アクセスのコンテンツ情報を判断し、判断された関連する前記コンテンツ情報をユーザに示す。

【選択図】 図4



【特許請求の範囲】**【請求項 1】**

情報を含むドキュメントになされるアクションに応じて、及び、該情報のコンテキストに基づいて、関連する既アクセスのコンテンツ情報を判断し、

判断された関連する前記コンテンツ情報をユーザに示す、
コンテンツ情報を処理する方法。

【請求項 2】

ユーザに示すことは、少なくとも通知特性を活性化することを含み、
該通知特性は、判断された関連する前記コンテンツ情報の可用性を示す、
請求項 1 に記載のコンテンツ情報を処理する方法。

10

【請求項 3】

ドキュメントになされる前記アクションは、少なくとも、ドキュメントをリトリープするアクション、ドキュメントを開くアクション、ドキュメントを保存するアクション、ドキュメントをファイルするアクション、ドキュメントを編集するアクション、ドキュメントを削除するアクション、ドキュメントを転送するアクション、及び、ドキュメントをブックマークするアクション、の一以上を含む、請求項 1 に記載のコンテンツ情報を処理する方法。

【請求項 4】

既アクセスのコンテンツ情報のグループに含まれるように、前記ドキュメントの情報を処理することをさらに含む、請求項 1 に記載のコンテンツ情報を処理する方法。

20

【請求項 5】

情報を処理することは、前記ドキュメントから該情報を抽出することを含む、請求項 4 に記載のコンテンツ情報を処理する方法。

【請求項 6】

情報を処理することは、媒体記憶装置に抽出された該情報を出力することをさらに含む、請求項 5 に記載のコンテンツ情報を処理する方法。

【請求項 7】

前記媒体記憶装置に抽出した情報を出力することは、該媒体記憶装置に抽出した該情報を周期的に出力することを含む、請求項 6 に記載のコンテンツ情報を処理する方法。

30

【請求項 8】

前記情報はテキスト情報であり、
情報を処理することは、抽出された情報をテキスト表現方式に変換すること、をさらに含む、
請求項 6 に記載のコンテンツ情報を処理する方法。

【請求項 9】

情報を処理することは、変換された情報と既アクセス・コンテンツとの類似性を定量化する、少なくとも一つの類似性スコアを判断すること、をさらに含む、請求項 8 に記載のコンテンツ情報を処理する方法。

【請求項 10】

情報を処理することは、既アクセスのコンテンツ情報及び変換された情報のグループから、類似性閾値を上回る類似性スコアを有するコンテンツ情報の集合をさらに判断すること、をさらに含む、請求項 9 に記載のコンテンツ情報を処理する方法。

40

【請求項 11】

さらに判断することは、逆インデックス技法を使用して実行される、請求項 10 に記載のコンテンツ情報を処理する方法。

【請求項 12】

さらに判断することは、変換された情報に基づいて、自動的に生成されたテキスト照会を使用してインデックスを照会すること、をさらに含む、請求項 11 に記載のコンテンツ情報を処理する方法。

【請求項 13】

50

情報を処理することは、既アクセスのコンテンツ情報を含む逆インデックスに変換された情報を追加すること、をさらに含む、請求項 1 1 に記載のコンテンツ情報を処理する方法。

【請求項 1 4】

情報を処理することは、少なくとも類似性スコアに基づいて、コンテンツ情報の順位を判断すること、をさらに含む、請求項 1 0 に記載のコンテンツ情報を処理する方法。

【請求項 1 5】

関連するコンテンツ情報は、少なくともコンピュータ・アプリケーション・ドキュメント、テキスト・メッセージ、eメール・メッセージ、カレンダ・エントリまたはウェブ・ページの一部である、請求項 1 に記載のコンテンツ情報を処理する方法。

10

【請求項 1 6】

関連するコンテンツ情報は、少なくとも、ページ全体、一頁の中に含まれる個々のテキスト文字、ワード、句、テキスト行、文、段落、テキストの段、テキストのブロック、テキスト記事、複数頁ドキュメント、單一頁ドキュメントの集合体、及び複数頁ドキュメントの集合体、を含むリストから選択される、請求項 1 に記載のコンテンツ情報を処理する方法。

【請求項 1 7】

関連する既アクセス・コンテンツは、既アクセス・コンテンツのグループにコンテンツを含める、コンピュータ・アプリケーションのタイプに関係なく判断される、請求項 1 に記載のコンテンツ情報を処理する方法。

20

【請求項 1 8】

示すことは、関連するコンテンツ情報のダイジェスト・ドキュメントを生成すること、を含む、請求項 1 に記載のコンテンツ情報を処理する方法。

【請求項 1 9】

関連するコンテンツ情報の焦点要約を生成すること、をさらに含む、請求項 1 8 に記載のコンテンツ情報を処理する方法。

【請求項 2 0】

情報を含むドキュメントになされるアクションに応じて、及び、該情報のコンテクストに基づいて、関連する既アクセスのコンテンツ情報を判断し、

判断された関連する前記コンテンツ情報をユーザに示す、

30

ことを、プロセッサにより実行された場合に、該プロセッサに実施させる、コンテンツ情報を処理するプログラム。

【請求項 2 1】

ユーザに示すことは、少なくとも通知特性を活性化することを含み、
該通知特性は、判断された関連する前記コンテンツ情報の可用性を示す、
請求項 2 0 に記載のコンテンツ情報を処理するプログラム。

【請求項 2 2】

ドキュメントになされる前記アクションは、少なくとも、ドキュメントをリトリープするアクション、ドキュメントを開くアクション、ドキュメントを保存するアクション、ドキュメントをファイルするアクション、ドキュメントを編集するアクション、ドキュメントを削除するアクション、ドキュメントを転送するアクション、及び、ドキュメントをブックマークするアクション、の一以上を含む、請求項 2 0 に記載のコンテンツ情報を処理するプログラム。

40

【請求項 2 3】

既アクセスのコンテンツ情報のグループに含まれるように、前記ドキュメントの情報を処理することをさらに含む、請求項 2 0 に記載のコンテンツ情報を処理するプログラム。

【請求項 2 4】

情報を処理することは、前記ドキュメントから該情報を抽出することを含む、請求項 2 3 に記載のコンテンツ情報を処理するプログラム。

【請求項 2 5】

50

情報を処理することは、媒体記憶装置に抽出された該情報を出力することをさらに含む、請求項 24 に記載のコンテンツ情報を処理するプログラム。

【請求項 26】

前記媒体記憶装置に抽出した情報を出力することは、該媒体記憶装置に抽出した該情報を周期的に出力することを含む、請求項 25 に記載のコンテンツ情報を処理するプログラム。

【請求項 27】

前記情報はテキスト情報であり、

情報を処理することは、抽出された情報をテキスト表現方式に変換すること、をさらに含む、

10

請求項 25 に記載のコンテンツ情報を処理するプログラム。

【請求項 28】

情報を処理することは、変換された情報と既アクセス・コンテンツとの類似性を定量化する、少なくとも一つの類似性スコアを判断すること、をさらに含む、請求項 27 に記載のコンテンツ情報を処理するプログラム。

【請求項 29】

情報を処理することは、既アクセスのコンテンツ情報及び変換された情報のグループから、類似性閾値を上回る類似性スコアを有するコンテンツ情報の集合をさらに判断すること、をさらに含む、請求項 28 に記載のコンテンツ情報を処理するプログラム。

【請求項 30】

さらに判断することは、逆インデックス技法を使用して実行される、請求項 29 に記載のコンテンツ情報を処理するプログラム。

【請求項 31】

さらに判断することは、変換された情報に基づいて、自動的に生成されたテキスト照会を使用してインデックスを照会すること、をさらに含む、請求項 30 に記載のコンテンツ情報を処理するプログラム。

20

【請求項 32】

情報を処理することは、既アクセスのコンテンツ情報を含む逆インデックスに変換された情報を追加すること、をさらに含む、請求項 30 に記載のコンテンツ情報を処理するプログラム。

30

【請求項 33】

情報を処理することは、少なくとも類似性スコアに基づいて、コンテンツ情報の順位を判断すること、をさらに含む、請求項 29 に記載のコンテンツ情報を処理するプログラム。

【請求項 34】

関連するコンテンツ情報は、少なくともコンピュータ・アプリケーション・ドキュメント、テキスト・メッセージ、eメール・メッセージ、カレンダ・エントリまたはウェブ・ページの一部である、請求項 20 に記載のコンテンツ情報を処理するプログラム。

【請求項 35】

関連するコンテンツ情報は、少なくとも、ページ全体、一頁の中に含まれる個々のテキスト文字、ワード、句、テキスト行、文、段落、テキストの段、テキストのブロック、テキスト記事、複数頁ドキュメント、單一頁ドキュメントの集合体、及び複数頁ドキュメントの集合体、を含むリストから選択される、請求項 20 に記載のコンテンツ情報を処理するプログラム。

40

【請求項 36】

関連する既アクセス・コンテンツは、既アクセス・コンテンツのグループにコンテンツを含める、コンピュータ・アプリケーションのタイプに関係なく判断される、請求項 20 に記載のコンテンツ情報を処理するプログラム。

【請求項 37】

示すことは、関連するコンテンツ情報のダイジェスト・ドキュメントを生成すること、を含む、請求項 20 に記載のコンテンツ情報を処理するプログラム。

50

【請求項 3 8】

関連するコンテンツ情報の焦点要約を生成すること、をさらに含む、請求項 2 0 に記載のコンテンツ情報を処理するプログラム。

【請求項 3 9】

メモリと、
制御装置と、
を備え、
該制御装置は、

情報を含むドキュメントになされるアクションに応じて、及び、該情報のコンテキストに基づいて、関連する既アクセスのコンテンツ情報を判断し、

10 判断された関連する前記コンテンツ情報をユーザに示す、
コンテンツ情報を処理するシステム。

【請求項 4 0】

ユーザに示すことは、少なくとも通知特性を活性化することを含み、
該通知特性は、判断された関連する前記コンテンツ情報の可用性を示す、
請求項 3 9 に記載のコンテンツ情報を処理するシステム。

【請求項 4 1】

ドキュメントになされる前記アクションは、少なくとも、ドキュメントをリトリープするアクション、ドキュメントを開くアクション、ドキュメントを保存するアクション、ドキュメントをファイルするアクション、ドキュメントを編集するアクション、ドキュメントを削除するアクション、ドキュメントを転送するアクション、及び、ドキュメントをブックマークするアクション、の一以上を含む、請求項 3 9 に記載のコンテンツ情報を処理するシステム。

【請求項 4 2】

クライアント・モジュール及びサーバ・モジュールをさらに備える、請求項 3 9 に記載のコンテンツ情報を処理するシステム。

【請求項 4 3】

少なくとも前記クライアント・モジュールは、情報を含む前記ドキュメントにアクセスするために使用されるアプリケーションに埋め込まれている、請求項 4 2 に記載のコンテンツ情報を処理するシステム。

【請求項 4 4】

既アクセスのコンテンツ情報のグループに含まれるように、前記ドキュメントの情報を処理することをさらに含む、請求項 3 9 に記載のコンテンツ情報を処理するシステム。

【請求項 4 5】

情報を処理することは、前記ドキュメントから該情報を抽出することを含む、請求項 4 4 に記載のコンテンツ情報を処理するシステム。

【請求項 4 6】

情報を処理することは、媒体記憶装置に抽出された該情報を出力することをさらに含む、請求項 4 5 に記載のコンテンツ情報を処理するシステム。

【請求項 4 7】

前記媒体記憶装置に抽出した情報を出力することは、該媒体記憶装置に抽出した該情報を周期的に出力することを含む、請求項 4 6 に記載のコンテンツ情報を処理するシステム。

【請求項 4 8】

前記情報はテキスト情報であり、

情報を処理することは、抽出された情報をテキスト表現方式に変換すること、をさらに含む、

請求項 4 6 に記載のコンテンツ情報を処理するシステム。

【請求項 4 9】

情報を処理することは、変換された情報と既アクセス・コンテンツとの類似性を定量化する、少なくとも一つの類似性スコアを判断すること、をさらに含む、請求項 4 8 に記載の

10

20

30

40

50

コンテンツ情報を処理するシステム。

【請求項 5 0】

情報を処理することは、既アクセスのコンテンツ情報及び変換された情報のグループから、類似性閾値を上回る類似性スコアを有するコンテンツ情報の集合をさらに判断すること、をさらに含む、請求項 4 9 に記載のコンテンツ情報を処理するシステム。

【請求項 5 1】

さらに判断することは、逆インデックス技法を使用して実行される、請求項 5 0 に記載のコンテンツ情報を処理するシステム。

【請求項 5 2】

さらに判断することは、変換された情報に基づいて、自動的に生成されたテキスト照会を使用してインデックスを照会すること、をさらに含む、請求項 5 1 に記載のコンテンツ情報を処理するシステム。 10

【請求項 5 3】

情報を処理することは、既アクセスのコンテンツ情報を含む逆インデックスに変換された情報を追加すること、をさらに含む、請求項 5 1 に記載のコンテンツ情報を処理するシステム。

【請求項 5 4】

情報を処理することは、少なくとも類似性スコアに基づいて、コンテンツ情報の順位を判断すること、をさらに含む、請求項 5 0 に記載のコンテンツ情報を処理するシステム。

【請求項 5 5】

関連するコンテンツ情報は、少なくともコンピュータ・アプリケーション・ドキュメント、テキスト・メッセージ、eメール・メッセージ、カレンダ・エントリまたはウェブ・ページの一部である、請求項 3 9 に記載のコンテンツ情報を処理するシステム。 20

【請求項 5 6】

関連するコンテンツ情報は、少なくとも、ページ全体、一頁の中に含まれる個々のテキスト文字、ワード、句、テキスト行、文、段落、テキストの段、テキストのブロック、テキスト記事、複数頁ドキュメント、單一頁ドキュメントの集合体、及び複数頁ドキュメントの集合体、を含むリストから選択される、請求項 3 9 に記載のコンテンツ情報を処理するシステム。

【請求項 5 7】

関連する既アクセス・コンテンツは、既アクセス・コンテンツのグループにコンテンツを含める、コンピュータ・アプリケーションのタイプに関係なく判断される、請求項 3 9 に記載のコンテンツ情報を処理するシステム。 30

【請求項 5 8】

示すことは、関連するコンテンツ情報のダイジェスト・ドキュメントを生成すること、を含む、請求項 3 9 に記載のコンテンツ情報を処理するシステム。

【請求項 5 9】

関連するコンテンツ情報の焦点要約を生成すること、をさらに含む、請求項 5 8 に記載のコンテンツ情報を処理するシステム。

【請求項 6 0】

ー以上の既アクセス・ドキュメントからテキストを抽出し、
抽出した該テキストをサーバ・モジュールに先行送信し、
前記サーバ・モジュールから受信した応答に基づいて、該サーバ・モジュールにより判斷された関連するコンテンツ情報をユーザに先行通知する、
ことを、クライアント・モジュールのプロセッサにより実行された場合に、該プロセッサに実施させる、コンテンツ情報を処理するプログラム。 40

【請求項 6 1】

判断された関連する前記コンテンツ情報に電子的接続を提供する、ことをさらに含む、請求項 6 0 に記載のコンテンツ情報を処理するプログラム。

【請求項 6 2】

50

前記クライアント・モジュールは、ホストコンピュータ・アプリケーションに埋め込まれている、請求項 6 0 に記載のコンテンツ情報を処理するプログラム。

【請求項 6 3】

前記クライアント・モジュール及び前記サーバ・モジュールが別々の情報処理装置で提供される、請求項 6 0 に記載のコンテンツ情報を処理するプログラム。

【請求項 6 4】

クライアント・モジュールにより出力された現在表示されているテキストをデータ記憶媒体に記憶し、

現在表示されている前記テキストのインデックスを生成し、

ユーザの現在のコンテクストに密接に関連する既アクセス・コンテンツを判断するため 10 に、前記インデックスを検索する、

ことを、サーバ・モジュールのプロセッサにより実行された場合に、該プロセッサに実施させる、コンテンツ情報を処理するプログラム。

【請求項 6 5】

ユーザによって定義されたダイジェスト特定に従って、既アクセス・コンテンツのダイジェストを生成すること、をさらに含む、請求項 6 4 に記載のコンテンツ情報を処理するプログラム。

【請求項 6 6】

判断されたコンテンツの焦点要約を生成すること、をさらに含む請求項 6 4 に記載のコンテンツ情報を処理するプログラム。 20

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本発明は既にアクセスした情報を処理することに関する。詳細には、本発明はコンテンツ情報を処理する方法、システム、及びプログラムに関する。

【背景技術】

【0 0 0 2】

例えば、ウェブ・ページ、e メール・メッセージまたはドキュメントなどの既にアクセスした情報を検出することは、コンピュータを使用して行われている最も頻繁なアクションの一つである。例えば、最近の研究では、5 回のウェブ・ページ閲覧の内、4 回が既にアクセスした、例えば、既に見たウェブ・ページに対してあることが示唆されている。 30

【0 0 0 3】

特に、再閲覧サポートのために設計された数多くのツールがある。例えば、最近アクセスしたファイルの履歴をよく維持しているオペレーティング・システム及びコンピュータ・アプリケーションがある。別の例では、ウェブ・ブラウザ及びいくつかのファイル・マネージャは、ブックマークを生成、編成するための機能だけではなく、「戻る」ボタン、最近アクセスしたファイル、及び URL の履歴も介して再閲覧をサポートしている。さらに、ウェブ・ページにまつわる別の例は、検索ツールが再閲覧ツールとして頻繁に使用されることである。

【0 0 0 4】

多くの場合、既に見たコンテンツを検出することは難しく、時間がかかる。コンテンツを再閲覧することがこのように広く行われている一般的な活動であるという事実にも関わらず、既存の検索ツールは、既にアクセスした情報を再閲覧することではなく情報の検出に重点を置いている。 40

【0 0 0 5】

多くの場合、既に見たコンテンツに戻る理由は新しいドキュメントを生成するためである。例えば、最も関連性のあるリソースと数日間分の価値があるウェブに基づいた調査の主な発見事項を収集し、要約するダイジェストの生成が、他者と該発見事項を共有すること、レポートを生成すること、あるいはさらに後で参照するために密接に関連したリソースの集合を保存することなどの複数の理由から実行される場合がある。 50

【 0 0 0 6 】

一般的に使用される再閲覧手法には多くの欠点がある。例えば、ウェブ・ブラウザとファイル・マネージャで使用される戻るボタンは短期間のドキュメント・レビューをサポートするだけである。ナビゲーションは通常純粹に逐次的である。つまり、都合良く到達するのは最近閲覧したページとドキュメントだけである。加えて、あるページに戻り、別のパスを辿ると、通常、最初のパスから閲覧したページが削除され、戻るボタンを介して最近閲覧したすべてのページに達するのは不可能になる。

【 0 0 0 7 】

ファイル・アクセス及びURL履歴機能は戻るボタンより多くの柔軟性を与えるが、アクセスしたファイルとURLの長いリストをスキミングすることは効率的でも、便利でもない。さらに、ユーザはページ・タイトル、URLまたはファイル名を探している情報と関連付けることができなければならず、これは、ユーザが既に見た情報の出所を知らない場合、あるいはページ・タイトルが参考にならない場合には特に困難である。加えて、ファイル・アクセスまたはURL履歴は、通常アプリケーション単位で維持されている。例えば、ウェブ・ページにアクセスし、eメールを読み、ドキュメントを開くと、通常3つの別々の履歴が生じる。

10

【 0 0 0 8 】

ブックマークを使用することの主要な欠点は、ユーザはページに含まれている情報に将来ニーズが出てくる可能性があるかどうかを前もって評価しなければならないという点である。非常に多くのページをブックマークすることは、ブックマークの数と、それらを効率的に活用するために必要とされる編成と、がこのような大規模では難しくなるという理由からよい解決策ではない場合が多い。一般的にはブックマークの活用は、ユーザがそれらを生成、維持する際に費やす作業量に直接的に関係している。

20

【 0 0 0 9 】

多くの場合URL、ファイル名及びディレクトリを記憶することが困難であるために、既に見たコンテンツに戻るためには頻繁に検索ツールが再利用される。しかしながら、通常、検索ツール使用時に使用可能な特定の再閲覧サポートはない。例えば、既にアクセスしたウェブ・ページをリトリープすることは、多くの場合、所望されるリンクが発見されるまで複数回照会を繰り返すことを必要とする。さらに、既にアクセスしたウェブ・ページがまだ存在しているという保証はない。

30

【 0 0 1 0 】

閲覧したコンテンツを活用し、自動的にダイジェスト、ログまたは要約を生成することは今日入手可能なツールによって十分にサポートされていない。関連するリソースを収集し、ダイジェストまたはレポートを生成するプロセスは、将来のアクセスのために関連性のあるコンテンツをマーキングすること、ページを再閲覧すること、最も関連性のあるセクションを検出すこと、及び／又は最重要点を要約すること、次にすべての検出した情報を一箇所に収集するドキュメントをアセンブルすることから構成される大部分は手作業による活動である。これは非常に煩わしいプロセスである。既存のシステムまたはツールで既にアクセスしたコンテンツに基づいたものはない。それらは主に新しい情報を検出することに力を注いでいる。

40

【 0 0 1 1 】

例えば、非特許文献1～3は、本願発明に関連する技術について記載している。

【非特許文献1】バレットら(Barrett et al.)、「ウェブ・パーソナル化法(How to Personalize the Web)」、計算機学会(ACM)、シカゴ、1997年、頁75～82

【非特許文献2】バドジグら(Budzig et al.)、「コンテクストにおける情報アクセス(Information Access In Context)」、知識ベース・システム(Knowledge Based Systems)、エルセヴィア・サイエンス(Elsevier Science)、2001年、14(1-2)、頁37～53

【非特許文献3】コックバーンら(Cockburn et al.)、「ウェブ・ページ再閲覧に関する改良：分析、設計、及び、評価(Improving WEB Page Revisitation : Analysis, Design and Evaluation)」、IT & ソサエティ(IT & Society)、2003年、3(1)、頁159～

50

183

【発明の開示】

【発明が解決しようとする課題】

【0012】

本発明の課題は、コンテンツの再閲覧を容易にするコンテンツ情報を処理する方法、システム、及びプログラムを提供することである。

【課題を解決するための手段】

【0013】

本発明は、ユーザが既にアクセスしたコンテンツを再閲覧するのを補助するために自動コンテンツ・インデックス生成技法及び自動コンテンツ・リトリープ技法を使用する方法及びシステムを提供する。 10

【0014】

本発明は、先行型照会生成機能と推奨機能に自動化されたコンテンツ・インデックス生成を統合し、既に見たコンテンツへの自動コンテキスト・アクセスを可能にする方法及びシステムを別に提供する。

【0015】

本発明は、既にアクセスしたコンテンツに基づいて焦点（ポイントを絞った）ダイジェスト・ドキュメントを生成するために自動インデックス生成技法とリトリープ技法を使用する方法及びシステムを別に提供する。

【0016】

本発明は、既にアクセスしたコンテンツに基づいてダイジェストの中にドキュメントのコンテキストに特殊な要約を生成するために自動インデックス生成技法とリトリープ技法を使用する方法及びシステムを別に提供する。 20

【0017】

本発明は既に見た、あるいはアクセスしたリソースの完全に自動化されたリトリープ及び要約に対処し、コンテキストに合焦したダイジェストの自動生成を可能にする方法及びシステムを別に提供する。

【0018】

本発明は、例えば、ウェブ・ブラウザ、eメール・クライアントまたはワープロなどの一般的なアプリケーションを一般的に使用している間に、ユーザがインタラクトするすべてのコンテンツの全文インデックスを自動的に生成する方法及びシステムを別に提供する。 30

【0019】

テキスト・ドキュメントに関する多様な例示的な実施形態においては、本発明によるシステム及び方法は現在表示されているテキストを、送信されてきたテキストを全文インデックスに追加するサーバに先行して送信する。テキスト・ドキュメントに関する多様な例示的な実施形態においては、本発明にかかるシステム及び方法が、例えば、現在表示されているウェブ・ページ、受信及び／又は表示、編集等されているeメール・メッセージなどの、ユーザの現在のコンテキストに密接に関係する、既にアクセスしたコンテンツを先行して決定する、または検出するために、生成されたインデックスを使用する。

【0020】

本発明は、リトリープされたドキュメントの集まりがコンテンツ再閲覧とダイジェスト生成の両方のために使用される方法及びシステムを別に提供する。 40

【0021】

本発明の第1の態様のコンテンツ情報を処理する方法は、情報を含むドキュメントになされるアクションに応じて、及び、該情報のコンテキストに基づいて、関連する既アクセスのコンテンツ情報を判断し、判断された関連する前記コンテンツ情報をユーザに示す。

【0022】

本発明の第2の態様は、第1の態様のコンテンツ情報を処理する方法であって、ユーザに示すことは、少なくとも通知特性を活性化することを含み、該通知特性は、判断された関連する前記コンテンツ情報の可用性を示す。 50

【 0 0 2 3 】

本発明の第3の態様は、第1の態様のコンテンツ情報を処理する方法であって、ドキュメントになされる前記アクションは、少なくとも、ドキュメントをリトリープするアクション、ドキュメントを開くアクション、ドキュメントを保存するアクション、ドキュメントをファイルするアクション、ドキュメントを編集するアクション、ドキュメントを削除するアクション、ドキュメントを転送するアクション、及び、ドキュメントをブックマークするアクション、の一以上を含む。

【 0 0 2 4 】

本発明の第4の態様は、第1の態様のコンテンツ情報を処理する方法であって、既アクセスのコンテンツ情報のグループに含まれるように、前記ドキュメントの情報を処理することをさらに含む。

10

【 0 0 2 5 】

本発明の第5の態様は、第4の態様のコンテンツ情報を処理する方法であって、情報を処理することは、前記ドキュメントから該情報を抽出することを含む。

【 0 0 2 6 】

本発明の第6の態様は、第5の態様のコンテンツ情報を処理する方法であって、情報を処理することは、媒体記憶装置に抽出された該情報を出力することをさらに含む。

20

【 0 0 2 7 】

本発明の第7の態様は、第6の態様のコンテンツ情報を処理する方法であって、前記媒体記憶装置に抽出した情報を出力することは、該媒体記憶装置に抽出した該情報を周期的に出力することを含む。

20

【 0 0 2 8 】

本発明の第8の態様は、第6の態様のコンテンツ情報を処理する方法であって、前記情報はテキスト情報であり、情報を処理することは、抽出された情報をテキスト表現方式に変換すること、をさらに含む。

【 0 0 2 9 】

本発明の第9の態様は、第8の態様のコンテンツ情報を処理する方法であって、情報を処理することは、変換された情報と既アクセス・コンテンツとの類似性を定量化する、少なくとも一つの類似性スコアを判断すること、をさらに含む。

30

【 0 0 3 0 】

本発明の第10の態様は、第9の態様のコンテンツ情報を処理する方法であって、情報を処理することは、既アクセスのコンテンツ情報及び変換された情報のグループから、類似性閾値を上回る類似性スコアを有するコンテンツ情報を集合をさらに判断すること、をさらに含む。

【 0 0 3 1 】

本発明の第11の態様は、第10の態様のコンテンツ情報を処理する方法であって、さらに判断することは、逆インデックス技法を使用して実行される。

40

【 0 0 3 2 】

本発明の第12の態様は、第11の態様のコンテンツ情報を処理する方法であって、さらに判断することは、変換された情報に基づいて、自動的に生成されたテキスト照会を使用してインデックスを照会すること、をさらに含む。

【 0 0 3 3 】

本発明の第13の態様は、第11の態様のコンテンツ情報を処理する方法であって、情報を処理することは、既アクセスのコンテンツ情報を含む逆インデックスに変換された情報を追加すること、をさらに含む。

【 0 0 3 4 】

本発明の第14の態様は、第10の態様のコンテンツ情報を処理する方法であって、情報を処理することは、少なくとも類似性スコアに基づいて、コンテンツ情報を順位を判断すること、をさらに含む。

50

【 0 0 3 5 】

本発明の第15の態様は、第1の態様のコンテンツ情報を処理する方法であって、関連するコンテンツ情報は、少なくともコンピュータ・アプリケーション・ドキュメント、テキスト・メッセージ、eメール・メッセージ、カレンダ・エントリまたはウェブ・ページの一部である。

【0036】

本発明の第16の態様は、第1の態様のコンテンツ情報を処理する方法であって、関連するコンテンツ情報は、少なくとも、ページ全体、一頁の中に含まれる個々のテキスト文字、ワード、句、テキスト行、文、段落、テキストの段、テキストのブロック、テキスト記事、複数頁ドキュメント、單一頁ドキュメントの集合体、及び複数頁ドキュメントの集合体、を含むリストから選択される。

10

【0037】

本発明の第17の態様は、第1の態様のコンテンツ情報を処理する方法であって、関連する既アクセス・コンテンツは、既アクセス・コンテンツのグループにコンテンツを含める、コンピュータ・アプリケーションのタイプに關係なく判断される。

【0038】

本発明の第18の態様は、第1の態様のコンテンツ情報を処理する方法であって、示すことは、関連するコンテンツ情報のダイジェスト・ドキュメントを生成すること、を含む。

【0039】

本発明の第19の態様は、第18の態様のコンテンツ情報を処理する方法であって、関連するコンテンツ情報の焦点要約を生成すること、をさらに含む。

20

【0040】

本発明の第20の態様は、コンテンツ情報を処理するプログラムであって、情報を含むドキュメントになされるアクションに応じて、及び、該情報のコンテクストに基づいて、関連する既アクセスのコンテンツ情報を判断し、判断された関連する前記コンテンツ情報をユーザに示す、ことを、プロセッサにより実行された場合に、該プロセッサに実施させる、コンテンツ情報を処理するプログラムである。

【0041】

本発明の第21の態様は、第20の態様のコンテンツ情報を処理するプログラムであって、ユーザに示すことは、少なくとも通知特性を活性化することを含み、該通知特性は、判断された関連する前記コンテンツ情報の可用性を示す。

30

【0042】

本発明の第22の態様は、第20の態様のコンテンツ情報を処理するプログラムであって、ドキュメントになされる前記アクションは、少なくとも、ドキュメントをリトリープするアクション、ドキュメントを開くアクション、ドキュメントを保存するアクション、ドキュメントをファイルするアクション、ドキュメントを編集するアクション、ドキュメントを削除するアクション、ドキュメントを転送するアクション、及び、ドキュメントをブックマークするアクション、の一以上を含む。

【0043】

本発明の第23の態様は、第20の態様のコンテンツ情報を処理するプログラムであって、既アクセスのコンテンツ情報のグループに含まれるように、前記ドキュメントの情報を処理することをさらに含む。

40

【0044】

本発明の第24の態様は、第23の態様のコンテンツ情報を処理するプログラムであって、情報を処理することは、前記ドキュメントから該情報を抽出することを含む。

【0045】

本発明の第25の態様は、第24の態様のコンテンツ情報を処理するプログラムであって、情報を処理することは、媒体記憶装置に抽出された該情報を出力することをさらに含む。

【0046】

本発明の第26の態様は、第25の態様のコンテンツ情報を処理するプログラムであって

50

、前記媒体記憶装置に抽出した情報を出力することは、該媒体記憶装置に抽出した該情報を周期的に出力することを含む。

【0047】

本発明の第27の態様は、第25の態様のコンテンツ情報を処理するプログラムであって、前記情報はテキスト情報であり、情報を処理することは、抽出された情報をテキスト表現方式に変換すること、をさらに含む。

【0048】

本発明の第28の態様は、第27の態様のコンテンツ情報を処理するプログラムであって、情報を処理することは、変換された情報と既アクセス・コンテンツとの類似性を定量化する、少なくとも一つの類似性スコアを判断すること、をさらに含む。 10

【0049】

本発明の第29の態様は、第28の態様のコンテンツ情報を処理するプログラムであって、情報を処理することは、既アクセスのコンテンツ情報及び変換された情報のグループから、類似性閾値を上回る類似性スコアを有するコンテンツ情報を集合をさらに判断すること、をさらに含む。

【0050】

本発明の第30の態様は、第29の態様のコンテンツ情報を処理するプログラムであって、さらに判断することは、逆インデックス技法を使用して実行される。

【0051】

本発明の第31の態様は、第30の態様のコンテンツ情報を処理するプログラムであって、さらに判断することは、変換された情報に基づいて、自動的に生成されたテキスト照会を使用してインデックスを照会すること、をさらに含む。 20

【0052】

本発明の第32の態様は、第30の態様のコンテンツ情報を処理するプログラムであって、情報を処理することは、既アクセスのコンテンツ情報を含む逆インデックスに変換された情報を追加すること、をさらに含む。

【0053】

本発明の第33の態様は、第29の態様のコンテンツ情報を処理するプログラムであって、情報を処理することは、少なくとも類似性スコアに基づいて、コンテンツ情報の順位を判断すること、をさらに含む。 30

【0054】

本発明の第34の態様は、第20の態様のコンテンツ情報を処理するプログラムであって、関連するコンテンツ情報は、少なくともコンピュータ・アプリケーション・ドキュメント、テキスト・メッセージ、eメール・メッセージ、カレンダ・エントリまたはウェブ・ページの一部である。

【0055】

本発明の第35の態様は、第20の態様のコンテンツ情報を処理するプログラムであって、関連するコンテンツ情報は、少なくとも、ページ全体、一頁の中に含まれる個々のテキスト文字、ワード、句、テキスト行、文、段落、テキストの段、テキストのブロック、テキスト記事、複数頁ドキュメント、單一頁ドキュメントの集合体、及び複数頁ドキュメントの集合体、を含むリストから選択される。 40

【0056】

本発明の第36の態様は、第20の態様のコンテンツ情報を処理するプログラムであって、関連する既アクセス・コンテンツは、既アクセス・コンテンツのグループにコンテンツを含める、コンピュータ・アプリケーションのタイプに関係なく判断される。

【0057】

本発明の第37の態様は、第20の態様のコンテンツ情報を処理するプログラムであって、示すことは、関連するコンテンツ情報のダイジェスト・ドキュメントを生成すること、を含む。

【0058】

10

20

30

40

50

本発明の第38の態様は、第20の態様のコンテンツ情報を処理するプログラムであって、関連するコンテンツ情報の焦点要約を生成すること、をさらに含む。

【0059】

本発明の第39の態様は、コンテンツ情報を処理するシステムであって、メモリと、制御装置と、を備え、該制御装置は、情報を含むドキュメントになされるアクションに応じて、及び、該情報のコンテクストに基づいて、関連する既アクセスのコンテンツ情報を判断し、判断された関連する前記コンテンツ情報をユーザに示す。

【0060】

本発明の第40の態様は、第39の態様のコンテンツ情報を処理するシステムであって、ユーザに示すことは、少なくとも通知特性を活性化することを含み、該通知特性は、判断された関連する前記コンテンツ情報の可用性を示す。 10

【0061】

本発明の第41の態様は、第39の態様のコンテンツ情報を処理するシステムであって、ドキュメントになされる前記アクションは、少なくとも、ドキュメントをリトリーブするアクション、ドキュメントを開くアクション、ドキュメントを保存するアクション、ドキュメントをファイルするアクション、ドキュメントを編集するアクション、ドキュメントを削除するアクション、ドキュメントを転送するアクション、及び、ドキュメントをブックマークするアクション、の一以上を含む。

【0062】

本発明の第42の態様は、第39の態様のコンテンツ情報を処理するシステムであって、クライアント・モジュール及びサーバ・モジュールをさらに備える。 20

【0063】

本発明の第43の態様は、第42の態様のコンテンツ情報を処理するシステムであって、少なくとも前記クライアント・モジュールは、情報を含む前記ドキュメントにアクセスするために使用されるアプリケーションに埋め込まれている。

【0064】

本発明の第44の態様は、第39の態様のコンテンツ情報を処理するシステムであって、既アクセスのコンテンツ情報のグループに含まれるように、前記ドキュメントの情報を処理することをさらに含む。

【0065】

本発明の第45の態様は、第44の態様のコンテンツ情報を処理するシステムであって、情報を処理することは、前記ドキュメントから該情報を抽出することを含む。 30

【0066】

本発明の第46の態様は、第45の態様のコンテンツ情報を処理するシステムであって、情報を処理することは、媒体記憶装置に抽出された該情報を出力することをさらに含む。

【0067】

本発明の第47の態様は、第46の態様のコンテンツ情報を処理するシステムであって、前記媒体記憶装置に抽出した情報を出力することは、該媒体記憶装置に抽出した該情報を周期的に出力することを含む。

【0068】

本発明の第48の態様は、第46の態様のコンテンツ情報を処理するシステムであって、前記情報はテキスト情報であり、情報を処理することは、抽出された情報をテキスト表現方式に変換すること、をさらに含む。 40

【0069】

本発明の第49の態様は、第48の態様のコンテンツ情報を処理するシステムであって、情報を処理することは、変換された情報と既アクセス・コンテンツとの類似性を定量化する、少なくとも一つの類似性スコアを判断すること、をさらに含む。

【0070】

本発明の第50の態様は、第49の態様のコンテンツ情報を処理するシステムであって、情報を処理することは、既アクセスのコンテンツ情報及び変換された情報のグループから 50

、類似性閾値を上回る類似性スコアを有するコンテンツ情報を集合をさらに判断すること、をさらに含む。

【0071】

本発明の第51の態様は、第50の態様のコンテンツ情報を処理するシステムであって、さらに判断することは、逆インデックス技法を使用して実行される。

【0072】

本発明の第52の態様は、第51の態様のコンテンツ情報を処理するシステムであって、さらに判断することは、変換された情報に基づいて、自動的に生成されたテキスト照会を使用してインデックスを照会すること、をさらに含む。

【0073】

本発明の第53の態様は、第51の態様のコンテンツ情報を処理するシステムであって、情報を処理することは、既アクセスのコンテンツ情報を含む逆インデックスに変換された情報を追加すること、をさらに含む。

【0074】

本発明の第54の態様は、第50の態様のコンテンツ情報を処理するシステムであって、情報を処理することは、少なくとも類似性スコアに基づいて、コンテンツ情報の順位を判断すること、をさらに含む。

【0075】

本発明の第55の態様は、第39の態様のコンテンツ情報を処理するシステムであって、関連するコンテンツ情報は、少なくともコンピュータ・アプリケーション・ドキュメント、テキスト・メッセージ、eメール・メッセージ、カレンダ・エントリまたはウェブ・ページの一部である。

【0076】

本発明の第56の態様は、第39の態様のコンテンツ情報を処理するシステムであって、関連するコンテンツ情報は、少なくとも、ページ全体、一頁の中に含まれる個々のテキスト文字、ワード、句、テキスト行、文、段落、テキストの段、テキストのブロック、テキスト記事、複数頁ドキュメント、單一頁ドキュメントの集合体、及び複数頁ドキュメントの集合体、を含むリストから選択される。

【0077】

本発明の第57の態様は、第39の態様のコンテンツ情報を処理するシステムであって、関連する既アクセス・コンテンツは、既アクセス・コンテンツのグループにコンテンツを含める、コンピュータ・アプリケーションのタイプに関係なく判断される。

【0078】

本発明の第58の態様は、第39の態様のコンテンツ情報を処理するシステムであって、示すことは、関連するコンテンツ情報のダイジェスト・ドキュメントを生成すること、を含む。

【0079】

本発明の第59の態様は、第58の態様のコンテンツ情報を処理するシステムであって、関連するコンテンツ情報の焦点要約を生成すること、をさらに含む。

【0080】

本発明の第60の態様は、コンテンツ情報を処理するプログラムであって、一以上の既アクセス・ドキュメントからテキストを抽出し、抽出した該テキストをサーバ・モジュールに先行送信し、前記サーバ・モジュールから受信した応答に基づいて、該サーバ・モジュールにより判断された関連するコンテンツ情報をユーザに先行通知する、ことを、クライアント・モジュールのプロセッサにより実行された場合に、該プロセッサに実施させる、コンテンツ情報を処理するプログラムである。

【0081】

本発明の第61の態様は、第60の態様のコンテンツ情報を処理するプログラムであって、判断された関連する前記コンテンツ情報に電子的接続を提供する。

【0082】

10

20

30

40

50

本発明の第 62 の態様は、第 60 の態様のコンテンツ情報を処理するプログラムであって、前記クライアント・モジュールは、ホストコンピュータ・アプリケーションに埋め込まれている。

【 0 0 8 3 】

本発明の第 63 の態様は、第 60 の態様のコンテンツ情報を処理するプログラムであって、前記クライアント・モジュール及び前記サーバ・モジュールが別々の情報処理装置で提供される。

【 0 0 8 4 】

本発明の第 64 の態様はコンテンツ情報を処理するプログラムであって、クライアント・モジュールにより出力された現在表示されているテキストをデータ記憶媒体に記憶し、現在表示されている前記テキストのインデックスを生成し、ユーザの現在のコンテキストに密接に関連する既アクセス・コンテンツを判断するために、前記インデックスを検索する、ことを、サーバ・モジュールのプロセッサにより実行された場合に、該プロセッサに実施させる、コンテンツ情報を処理するプログラムである。10

【 0 0 8 5 】

本発明の第 65 の態様は、第 64 の態様のコンテンツ情報を処理するプログラムであって、ユーザによって定義されたダイジェスト特定に従って、既アクセス・コンテンツのダイジェストを生成すること、をさらに含む。

【 0 0 8 6 】

本発明の第 66 の態様は、第 64 の態様のコンテンツ情報を処理するプログラムであって、判断されたコンテンツの焦点要約を生成すること、をさらに含む。20

【 発明の効果 】

【 0 0 8 7 】

本発明のコンテンツ情報を処理する方法は、情報を含むドキュメントになされるアクションに応じて、及び、該情報のコンテキストに基づいて、関連する既アクセスのコンテンツ情報を判断し、判断された関連する前記コンテンツ情報をユーザに示すようにしているので、コンテンツの再閲覧を容易にすることができる。

【 発明を実施するための最良の形態 】

【 0 0 8 8 】

本発明のシステム及び方法の多様な例示的な実施形態は、図を参照して以下に詳細に説明される。30

【 0 0 8 9 】

図 1 は、本発明に従って既にアクセスしたコンテンツを再閲覧し、既にアクセスしたコンテンツに基づき焦点ダイジェスト・ドキュメントを生成するための方法及びシステムの実現の一つの例示的な実施形態のハイ・レベル概略表現である。

【 0 0 9 0 】

本発明によるシステム及び方法の多様な例示的な実施形態においては、コンテンツはテキスト、数字、記号、マーキング、メタ・データ等を含む。さらに本発明によるシステム及び方法の多様な例示的な実施形態においては、コンテンツは例えば、コンピュータ・アプリケーション・ドキュメント、テキスト・メッセージ、e メール・メッセージ、カレンダ・エントリ、ウェブ・ページ等のドキュメントの一部である。さらに本発明によるシステム及び方法の多様な例示的な実施形態においては、コンテンツは、ページ全体、一頁の中に含まれる個々のテキスト文字、ワード、句、テキスト行、文、段落、テキストの段、テキストのブロック、テキスト記事、複数頁ドキュメント、單一頁ドキュメントの集合体、複数頁ドキュメントの集合体等の中に含まれるか、あるいはそれ自体がページ全体、一頁の中に含まれる個々のテキスト文字、ワード、句、テキスト行、文、段落、テキストの段、テキストのブロック、テキスト記事、複数頁ドキュメント、單一頁ドキュメントの集合体、複数頁ドキュメントの集合体等である。40

【 0 0 9 1 】

本発明によるシステム及び方法の多様な例示的な実施形態においては、システム 10 は、50

一人または複数人のユーザが既にアクセスしたコンテンツを再閲覧するのを補助するためにインデックス生成技法とリトリープ技法を利用する。システム10はさらに既にアクセスしたコンテンツに基づいて焦点ダイジェスト・ドキュメントを生成するためにもインデックス生成技法とリトリープ技法を利用する。本発明によるシステム及び方法の多様な例示的な実施形態においては、これらの機能は自動的に実行される。

【0092】

図1に図示されるように、多様な例示的な実施形態においては、システム10は少なくとも2つのシステム・モジュールを利用する。これらの例示的な実施形態においては、第1のシステム・モジュールはクライアント・モジュール20である。クライアント・モジュール20は再閲覧及びダイジェスト生成のためのユーザ・インターフェースを提供する。第2のシステム・モジュールは、コンテンツ情報及び/又はコンテンツ・ドキュメントを記憶し、インデックス生成し、インデックス60に対して、及びインデックス60からリトリープするサーバ・モジュール30である。多様な例示的な実施形態においては、インデックス60は全文インデックスである。

【0093】

多様な例示的な実施形態においては、クライアント20は、例えば、ウェブ・ブラウザ、eメール・クライアント、プレゼンテーション・ソフトウェア、またはワープロ等のあらゆる一般的に使用されているアプリケーション50の中に埋め込まれている。例えば、図1に示されるように、多様な例示的な実施形態においては、グラフィック・ユーザ・インターフェース(GUI)環境において、クライアント・モジュール20は、ホスト・アプリケーション50のユーザ・インターフェース内に組み込まれるツールバー40として実現されている。代替の例示的な実施形態においては、クライアント・モジュール20は他の既知の、あるいは今後開発される方法または技法を使用して実現される。

【0094】

多様な例示的な実施形態においては、クライアント・モジュール20及びサーバ・モジュール30は、スタンドアロン單一ユーザ・セットアップが所望されるときなど、例えば、單一デスクトップ・コンピュータなどの同じホスト情報処理装置にインストールされる。代わりに、多様な例示的な実施形態においては、クライアント・モジュール20とサーバ・モジュール30は、例えば、システム10が同じサーバ・システムから複数のユーザをサポートすることになる場合など別々のホスト情報処理装置にインストールされる。

【0095】

多様な例示的な実施形態においては、複数のユーザが、ネットワーク環境を使用して、例えば、モバイル機器、複数のデスクトップ、パーソナル・デジタル・アシスタント・デバイス、移動計算装置及び通信装置などの複数の異なるシステムからサーバ・モジュール30に接続する。複数のユーザが複数の装置を介して同じサーバ・モジュール30にアクセスできるようにすることは、それによりユーザは使用される処理装置、アプリケーション及び/又は通信リンクの種類に関係なく同じコンテンツ履歴にアクセスできるため好適である。

【0096】

多様な例示的な実施形態においては、クライアント・モジュール20は、一般的に使用されている生産性アプリケーション50を使用して生成された一つまたは複数のドキュメント52からテキスト21を抽出すること、該抽出したテキストをサーバ・モジュール30に先行して送信すること22、検出された密接に関連する既にアクセスしたコンテンツの存在をユーザに先行して通知すること23、検出された密接に関連する既にアクセスしたコンテンツに電子接続を提供すること24、検出され、アクセス及び/又はリトリープされたユーザのコンテンツの明示的履歴25を提供すること、サーバ・モジュール30により生成されるダイジェストを特定するために使用されるダイジェスト生成構成要素を含むメニュー26を提供すること、及び類似する機能を含むが、それらに限られない、リストから選択される多様な機能を実行する。

【0097】

10

20

30

30

40

50

多様な例示的な実施形態においては、サーバ・モジュール30はインデックス60に現在表示されているテキストを記憶し、インデックスを生成する。多様な例示的な実施形態においては、サーバ・モジュール30は、ユーザの現在のコンテクストに密接に関係する既にアクセスしたコンテンツを決定するためにインデックス60のリトリーブ、例えば、照会32を実行する。多様な例示的な実施形態においては、サーバ・モジュール30はユーザの特定に従ってドキュメント70のダイジェスト33を生成する。

【0098】

図2は、本発明によるシステム及び方法に関するためのネットワーク環境200の一つの例示的な実施形態を概略して示す。図2に示されているように、コンテンツ再閲覧及びダイジェスト生成システムは、複数のユーザ5を有する環境で使用されるように構成されている。マルチ・ユーザ環境の多様な例示的な実施形態においては、ネットワーク環境200は、それぞれの単独のユーザ5がクライアント・モジュール20を含む情報処理装置210を有するように配置される。多様な例示的な実施形態においては、サーバ・モジュール30はサーバ・デバイスなどのスタンドアロン・情報処理装置220に含まれる。多様な代替の例示的な実施形態(不図示)においては、サーバ・モジュール30はユーザ5の情報処理装置210の一つに常駐する。多様な例示的な実施形態においては、ユーザ5のそれぞれは、一つまたは複数の通信リンク230を使用してネットワーク205上でコンテンツ再閲覧及びダイジェスト生成システムのサーバ・モジュール30に接続される。

【0099】

多様な例示的な実施形態においては、ネットワーク205は、例えば、ローカル・エリア・ネットワーク、ワイド・エリア・ネットワーク、ストレージ・エリア・ネットワーク、イントラネット、エクストラネット、インターネット、または他の種類の分散型ネットワークを含むが、それらに限定されない。多様な例示的な実施形態においては、ネットワーク205は有線部分及び/又は無線部分を含む。多様な例示的な実施形態においては、リンク230は、公衆電話交換網上の接続、直接ケーブル接続、ワイド・エリア・ネットワーク、ローカル・エリア・ネットワークまたはストレージ・エリア・ネットワーク上の接続、イントラネットまたはエクストラネット上の接続、インターネット上の接続、あるいは他の分散処理ネットワークまたはシステム上の接続を含む、ネットワーク205に対して、例えば、クライアント・モジュール20とサーバ・モジュール30などのコンテンツ再閲覧及びダイジェスト生成システムの多様な構成要素を接続するためのあるいは既知のあるいは今後開発される装置またはシステムである。一般的に、リンク230は例えば、クライアント・モジュール20及びサーバ・モジュール30などのコンテンツ再閲覧及びダイジェスト生成システムの多様な構成要素をネットワーク205に接続するために使用できる既知のまたは今後開発される接続システムまたは構造である場合がある。

【0100】

図3は、本発明に従って既にアクセスしたコンテンツを再閲覧し、焦点ダイジェスト・ドキュメントを生成するためのシステム300の一つの例示的な実施形態の機能ブロック図である。システム300は外部データ接続370を含む。多様な例示的な実施形態では、データ接続370は、例えば、公衆電話交換網上の接続、直接ケーブル接続、ワイド・エリア・ネットワーク、ローカル・エリア・ネットワークまたはストレージ・エリア・ネットワーク上の接続、イントラネットまたはエクストラネット上の接続、インターネット上の接続、あるいは他の分散処理ネットワークまたはシステムを含む、ネットワーク250などの外部にシステム300を接続するための既知のまたは今後開発される装置またはシステムである。一般的には、データ接続370は、例えば、ネットワーク205などの外部に対してシステム300を接続するために使用できる既知のまたは今後開発される接続システムまたは構造である。

【0101】

図3に示されるように、システム300はユーザに対して情報を表示するために使用可能な一または複数の表示装置340、及び一人または複数のユーザがシステム300にデー

10

20

30

40

50

タを入力できるようにするために使用できる一または複数のユーザ入力装置 350 を含む。該一または複数の表示装置 340 及び該一または複数の入力装置 350 は、それぞれ一または複数の通信リンク 341 及び 351 を介して、入出力インターフェース 330 を通じてシステム 300 に接続される。該一または複数の通信リンク 341 及び 351 は、通常、前記データ接続 370 と同様である。

【0102】

多様な例示的な実施形態では、システム 300 は、制御装置 320、メモリ 310、テキスト抽出回路またはルーチン 305、テキスト先行伝送回路またはルーチン 315、既閲覧コンテンツ・アクセス回路またはルーチン 325、明示履歴アクセス回路またはルーチン 335、ダイジェスト特定回路またはルーチン 345、コンテンツ持続及びインデックス生成回路またはルーチン 355、照会生成回路またはルーチン 365、再閲覧コンテンツ推奨回路またはルーチン 375、並びにダイジェスト生成回路またはルーチン 385 の一つまたは複数を含み、そのすべては一つまたは複数のデータバス及び／又は制御バス及び／又はアプリケーション・プログラミング・インターフェース 360 上で相互接続されている。

【0103】

多様な例示的な実施形態においては、制御装置 320 はシステム 300 の他の構成要素の動作を制御する。多様な例示的な実施形態においては、制御装置 320 は、必要に応じてシステム 300 の多様な構成要素間でのデータの流れも制御する。多様な例示的な実施形態においては、メモリ 310 はシステム 300 に入る、またはシステム 300 から出る情報を記憶する。多様な例示的な実施形態では、メモリ 310 は、システム 300 の機能を実現するあらゆる必要なプログラム及び／又はデータを記憶し、及び／又は多様な処理の段階で既にアクセスしたドキュメント・コンテンツ情報のインデックスなどのデータを記憶する。

【0104】

様々な例示的な実施形態において、メモリ 310 は、可変な揮発性又は不揮発性メモリ、若しくは、不变、すなわち、固定メモリの任意な適切な組合せであってよい。様々な例示的な実施形態において、揮発性又は不揮発性の可変なメモリは、一又は複数のスタティック若しくはダイナミック RAM、フロッピィ・ディスク、及びフロッピィ・ディスク・ドライブ、書き可能若しくは書き替可能光ディスク及び光ディスク・ドライブ、ハード・ディスク・ドライブ、フラッシュ・メモリなどであってよい。同様に、様々な実施形態において、不变、すなわち、固定メモリは、一又は複数の ROM、PROM、EPROM、EEPROM、CD-ROM 若しくは DVD-ROM ディスクなどの光 ROM、及び、光 ROM のディスク・ドライブなどであってよい。

【0105】

多様な例示的な実施形態においては、クライアント・モジュールは、一つまたは複数のアクセスしたドキュメントからテキストを抽出すること、該抽出したテキストをサーバ・モジュールに先行して送信すること、検出された密接に関係する既にアクセスしたコンテンツの存在をユーザに先行して通知すること、検出された密接に関連する既にアクセスしたコンテンツに電子接続を提供すること、検出され、アクセス及び／又はリトリープされたユーザのコンテンツの明示的な履歴を提供すること、サーバ・モジュールによって生成されるダイジェストを特定するために使用されるダイジェスト生成構成要素を提供すること、及び他の類似した機能を含むが、それらに限定されない、リストから選択される多様な機能を実行する。

【0106】

多様な例示的な実施形態においては、ユーザがドキュメントを開くなどのユーザ・アクションに応じて関連付けられた既にアクセスしたコンテンツ情報を判断するとき、クライアント・モジュール内のテキスト抽出回路またはルーチン 305 が、表示されているテキストを自動的に抽出する。多様な例示的な実施形態においては、クライアント・モジュールはホスト・アプリケーション内で実行する。これらの例示的な実施形態においては、クラ

10

20

30

40

50

イアント・モジュールは現在表示されているドキュメントにアクセスできるため、追加の処理のために表示されているテキストを容易に抽出する。

【0107】

多様な例示的な実施形態においては、テキスト抽出ステップに続き、クライアント・モジュール内のテキスト先行伝送回路またはルーチン315が、抽出されたテキストをサーバ・モジュールに、すなわち、サーバに先行して送信する。多様な例示的な実施形態においては、この伝送は、ユーザが、例えば、新しいドキュメントを開く、既存のドキュメントを開く、eメール・メッセージを見る、新しいURLにナビゲーションする等のドキュメントに対するアクションを実行すると必ず発生する。多様な追加の例示的な実施形態においては、ユーザが新しいドキュメントまたはeメール・メッセージを生成している間に、抽出されたテキストの周期的な伝送が発生する。追加の例示的な実施形態においては、抽出されたテキストの周期的な伝送は、ユーザが既存のドキュメント、eメール・メッセージまたは他のアプリケーション・ドキュメントを編集している間に発生する。

【0108】

多様な例示的な実施形態においては、テキスト伝送の目的は二重である。第1に、それによりサーバ・モジュールは現在表示されているテキストにインデックスを付けることができる。第2に、それによりサーバ・モジュールは、ユーザの現在のコンテクストに密接に関係している既にアクセスしたコンテンツをリトリープできる。多様な例示的な実施形態においては、コンテクストは、現在表示されているページ、前のn枚の表示頁、または例えば、時間、ロケーション、カレンダから抽出された約束等の他のコンテクスト情報により定義される。

【0109】

多様な例示的な実施形態においては、サーバ・モジュールへのコンテンツ伝送の後、サーバ・モジュールは、サーバが、ユーザの現在のコンテクストに密接に関係している既にアクセスしたコンテンツを含むかどうかを判断する。このようなコンテンツが存在する場合は、サーバ・モジュールは、情報処理装置生成クライアントに、密接に関係する既にアクセスしたコンテンツについての情報を送り返す。多様な例示的な実施形態においては、送信された情報は、一致するテキスト・セグメントだけではなく、一致するリソースのファイル名またはURL、頁またはドキュメントのタイトル、アクセス日付も含む。サーバはすべての伝送の全文を記憶しているため、コンテンツの元のロケーションが変更してもサーバが既にアクセスしたコンテンツをリトリープできることが理解されなければならない。多様な例示的な実施形態において、任意の種類の密接に関係する既にアクセスしたコンテンツは処理され、クライアント・モジュールに、従ってユーザ情報処理装置に送り返される。

【0110】

多様な例示的な実施形態においては、既に見たコンテンツに対するアクセスまたは既に見たコンテンツの表示は、既閲覧コンテンツ・アクセス回路またはルーチン325のアクセス/表示により提供される。多様な例示的な実施形態においては、クライアント・モジュールがユーザに検出された関係するコンテンツの存在を知らせるか、または通知する。多様な例示的な実施形態においては、ユーザは次に受信された情報を見ることを要求する。多様な例示的な実施形態においては、ユーザは一致するドキュメント、eメール・メッセージまたはウェブ・ページを開く。

【0111】

多様な例示的な実施形態においては、既に見たコンテンツに対するアクセスは、クライアント・ツールバーの「履歴」ボタンを使用して実現される。多様な例示的な実施形態においては、履歴ボタンは、クライアント・モジュール（すなわち、ユーザ情報処理装置）が一致するリソースについての情報を受け取るとその外観を変更する。これらの例示的な実施形態においては、ユーザが履歴ボタンをクリックすると、一致するリソースに対するアクセスを提供するメニューが表示される。

【0112】

10

20

30

40

50

多様な例示的な実施形態においては、ユーザは、サーバ・モジュールが例えば、密接に関連する既に見たコンテンツなどの一致するリソースをリトリフォードするために実行する自動サーチを修正する。例えば、多様な例示的な実施形態においては、ユーザは特定の日付範囲を特定するか、あるいは遠隔的に関係する結果を含むために類似性閾値を変更する。

【0113】

多様な例示的な実施形態においては、クライアント・モジュールによって実行される別の機能は、明示履歴アクセス回路またはルーチン335を通してユーザのコンテンツ履歴に対する明示的なアクセスのためのインターフェースを提供することである。多様な例示的な実施形態においては、ユーザは既にアクセスしたコンテンツの履歴を獲得するためにシステムの先行照会機能を使用する。代わりに、多様な例示的な実施形態においては、ユーザは既にアクセスしたコンテンツをリトリードするためにマニュアル照会を利用する。他の多様な例示的な実施形態においては、クライアント・モジュールが、例えば、ウェブ・ページ、eメール、ドキュメント等などの最近アクセスしたリソースのリストに明示的なアクセスを提供する。

【0114】

既にアクセスしたコンテンツを再閲覧することに加え、多様な例示的な実施形態においては、クライアント・モジュールは、サーバ・モジュールによって生成されるダイジェストを特定するために使用されるダイジェスト特定回路またはルーチン345を含む。ユーザは多くの異なる活動及びトピックに関係するコンテンツにアクセスする傾向があるため、多様な例示的な実施形態においては、このプロセスの一つの態様はダイジェストのトピックまたは焦点を特定することである。

【0115】

多様な例示的な実施形態においては、再閲覧アプローチと同様に、トピックまたは焦点は、例えば、現在表示されているウェブ・ページ、調査報告またはeメール・メッセージなどの、ユーザの現在のコンテキストにより提供される。代わりに、多様な例示的な実施形態においては、ユーザは代表的なテキストを含むURLまたはファイルのリストを特定するか、あるいはダイジェストに焦点を合わせるためにテキスト照会を入力する。加えて、多様な例示的な実施形態においては、要約の所望される長さ、含まれるドキュメントの最大数、及び/又はドキュメントの日付範囲もしくはファイル・タイプについての情報が、トピックまたは焦点に含まれる。

【0116】

多様な例示的な実施形態においては、特定がサーバ・モジュールに送信されると、サーバ・モジュールは、ダイジェストを生成するために、収集、要約及びドキュメントの構築を開始する。多様な例示的な実施形態においては、次に、サーバ・モジュールは、新規に生成されたドキュメントが表示される準備ができるとクライアント・モジュールに通知する。

【0117】

自動コンテキスト関連のコンテンツ・リトリフォード及び/又は情報ダイジェスト生成に対処する他の機能は本発明の範囲内にある。当業者にとって周知であるこれらの機能は、多様な例示的な実施形態でクライアント・モジュールを使用して実現される。

【0118】

多様な例示的な実施形態においては、サーバ・モジュールは、クライアント・モジュールから受け取る要求に基づいて、コンテンツを持続し(persistent)、インデックスを付け、リトリフォードする。加えて、多様な例示的な実施形態においては、サーバ・モジュールはユーザの特定に従ってダイジェスト・ドキュメントを生成する。

【0119】

多様な例示的な実施形態においては、サーバ・モジュールは既に送信したドキュメントの全文を含むデータベースを維持するために、コンテンツ持続及びインデックス生成回路、またはルーチン355を使用する。多様な例示的な実施形態においては、これはさらに、例えば、パス及びURL情報、ファイル・タイプ、アクセス日付、アクセス頻度等などの

10

20

30

40

50

メタ・データの一つまたは複数、などの情報を含む。加えて、多様な例示的な実施形態においては、サーバ・モジュールはすべての受け取られたコンテンツの全文インデックスを付加的に構築する。多様な例示的な実施形態においては、既にインデックスを付けたコンテンツをいつ削除するのかの基準が特定される。

【0120】

多様な例示的な実施形態においては、照会生成回路またはルーチン365は、サーバ・モジュールが任意の長さのテキスト・フラグメントを一つの加重照会に変換できるようとするアルゴリズムを使用する。多様な例示的な実施形態においては、サーバ・モジュールは該加重照会を使用して、再閲覧サポートまたはダイジェスト生成のために関係するコンテンツをリトリープする。

10

【0121】

多様な例示的な実施形態においては、サーバ・モジュールは再閲覧サポートとダイジェスト生成サポートの両方のために既にインデックスを付けたリソースをリトリープする。多様な例示的な実施形態においては、コンテンツがクライアントから送信されるとき、サーバ・モジュールは照会を生成し、それを全文テキストに対して実行し、関連性スコアをn個の最良の一一致に割り当てる。多様な例示的な実施形態においては、特定された閾値tを上回る関連性スコアが付いた一致が、サーバ・モジュール内の再閲覧のためのコンテンツ推奨回路またはルーチン375によって処理される。次に、多様な例示的な実施形態においては、特定された閾値tを上回る関連性スコアの付いた一致は、ユーザに提示されるためにクライアント・モジュールに送り返される。

20

【0122】

多様な例示的な実施形態においては、サーバ・モジュール内のダイジェスト生成回路またはルーチン385がダイジェスト生成要求を受け取ると、ダイジェスト生成回路またはルーチン385はユーザによって特定されるダイジェスト焦点に関するドキュメントをリトリープする。多様な例示的な実施形態においては、ユーザが、要約が含まれることを要求すると、システムは一致ドキュメントを要約する。ダイジェスト特定に応じて、多様な例示的な実施形態においては、システムは、例えば、URL、タイトル、アクセス日付などの、例えば、一致するドキュメントを説明する情報を含むウェブ・ページなどのドキュメントを生成する。さらに、多様な例示的な実施形態においては、システムは、オプションの要約、及び例えば、一致するドキュメントからのすべての画像を含む、ユーザが要求した他の情報も提供する。一致するドキュメントからの画像が、ユーザのタスクによってはテキストよりさらに有効である場合があることが理解されなければならない。

30

【0123】

多様な例示的な実施形態においては、コンテンツ情報は、例えば、テキスト、数字、記号、マーキング、メタ・データ等を含む。さらに、多様な例示的な実施形態においては、コンテンツは、例えば、コンピュータ・アプリケーション・ドキュメント、テキスト・メッセージ、eメール・メッセージ、カレンダ・エントリ、ウェブ・ページ等のドキュメントの一部である。さらに、多様な例示的な実施形態においては、本発明のシステム及び方法が適用されるコンテンツは、例えば、ページ全体、一頁に含まれる個々のテキスト文字、ワード、句、テキスト行、文、段落、テキストの段、テキストのブロック、テキスト記事、複数頁ドキュメント、單一頁ドキュメントの集合体、複数頁ドキュメントの集合体等の中に含まれるか、あるいはそれら全体である。

40

【0124】

図4は、本発明に従って既にアクセスしたコンテンツ情報を再閲覧するための方法の一つの例示的な実施形態を概略するフローチャートである。図4に示されるように、本方法はステップS400で開始し、既にアクセスした関連コンテンツ情報が、情報を含むドキュメントに対して実行されるアクションに応えて判断されるステップS410に続行する。多様な例示的な実施形態においては、該アクションは、ドキュメントをリトリープするアクション、ドキュメントを開くアクション、ドキュメントを保存するアクション、ドキュメントをファイルするアクション、ドキュメントを編集するアクション、ドキュメントを

50

削除するアクション、ドキュメントを転送するアクション、及びドキュメントをブックマークするアクションの一つまたは複数を含む。多様な例示的な実施形態においては、例えば、ドキュメントにアクセスする、ドキュメントをレビューする、及び／又はドキュメントを編集するときに、ユーザがドキュメントに対して実行する他のドキュメント・アクションは、現在既知であるものと今後開発されるものを含み、本発明の範囲内にある。

【0125】

多様な例示的な実施形態において、既にアクセスした関連コンテンツ情報は、アクセスされているドキュメントに含まれる情報に基づいて決定される。多様な例示的な実施形態においては、これは自動的に実行される。通常、決定された関連コンテンツ情報が、媒体記憶装置デバイスに記憶されている、既にアクセスしたコンテンツ情報ドキュメントのグループのサブ・パートであることが理解される必要がある。しかしながら、必ずしもこれが当てはまらないことも理解される必要がある。

10

【0126】

次にステップS420では、ユーザは判断された既にアクセスした任意の関連コンテンツ情報を通知される。多様な例示的な実施形態においては、通知は、ユーザに対して、判断された関連コンテンツ情報の可用性（アクセス可能か否か）を示す通知特性を少なくとも活性化する（オン状態にする）ことを含む。多様な例示的な実施形態においては、これは自動的に実行される。次に、動作は、本方法の動作が停止するステップS430に続行する。

20

【0127】

図5は、ステップS410の一つの例示的な実施形態をさらに詳細に概略するフローチャートである。多様な例示的な実施形態においては、これはアクセスされているドキュメント内に含まれる情報のコンテクストに基づいて関連コンテンツ情報を決定するために使用される。図5に示されるように、ステップS410は、例えば、現在表示されているテキストまたはアクセスされているドキュメントのようなコンテンツがさらなる処理のために抽出されるサブ・ステップS4110で始まる。次に、ステップS4120で、抽出されたコンテンツがさらなる処理のためにサーバ・モジュールに送信される。

【0128】

ステップS4130では、送信された抽出コンテンツが処理されるか、あるいはコンテンツ表現方式に変換される。テキストに関する多様な例示的な実施形態においては、サブ・ステップS4130が、テキスト・ドキュメントの類似性を評価するだけではなく、個々のワードの加重をサポートするテキスト表現に基づいた多様なアルゴリズムまたは技法を使用することによっても実行される。テキストに関する他の例示的な実施形態は、ステップS4130で他のアルゴリズムまたは技法を使用する。テキストに関する多様な例示的な実施形態においては、ベクトル空間モデルがテキスト表現方式として利用される。ベクトル空間モデルは情報リトリープ・システムにおいて一般的に使用されるテキスト表現パラダイムである。しかしながら、多様な例示的な実施形態においては、現在既知であるか、今後開発される他のモデルがこの機能を実行するために使用される。多様な例示的な実施形態では、コンテンツはテキスト以外のコンテンツである。

30

【0129】

ベクトル空間モデルにおいては、ドキュメントは、各ベクトル寸法がシステムの全体的な語彙の一つの用語に一致する用語重みのベクトルとして表現され、各用語は用語とドキュメントとの間の関連を定量化する。用語 - 重みは、頻繁にtf - idf用語加重方式、つまり用語 - 頻度 / 逆ドキュメント頻度に基づいている。この方式では、用語重みは用語がドキュメントの中に現れる回数(tf)、及びドキュメント集合体全体を通して該用語が現れる回数(df)に基づいて求められる。

40

【0130】

多様な例示的な実施形態においては、df値を求めるために、本方法はすべての既にインデックスが付けられたコンテンツを使用する。代わりに、多様な例示的な実施形態においては、本方法は、ユーザドキュメントの集合をすべてのユーザと対比してユーザによって

50

アクセスされるドキュメントに制限する。多様な他の例示的な実施形態においては、方法はリトリープ日付範囲に基づいてドキュメントを制限する。

【0131】

テキスト・ドキュメントを用語ベクトルに変換するために、多様な例示的な実施形態においては、本発明によるシステム及び方法は、例えば、htmlタグなどのすべてのフォーマット情報を削除し、次に結果として生じるテキストを個々の用語に分割する。

【0132】

サブ・ステップS4140では、多様な例示的な実施形態において、既に見たコンテンツを再閲覧提案の集合の中に含めるのか、あるいはダイジェストの中に含めるのかを判断するために、本発明によるシステムがドキュメントの類似性を定量化する。多様な例示的な実施形態においては、ドキュメントが用語ベクトルに変換されると、別の用語ベクトルに対するその類似性はベクトルを正規化してから、点乗積を取ることによって決定される。結果として生じる類似性スコアは、一般的にはコサイン類似性基準として知られている。多様な例示的な実施形態においては、コサイン類似性基準が使用される。

【0133】

コンピュータ処理の効率の理由から、例えば、現在表示されているドキュメントなどの一つのドキュメントと、サーバ内に常駐する他の既に見たすべてのドキュメントとの類似性を計算することは望ましくない場合がある。このようにして、多様な例示的な実施形態においては、すべての類似性値が使用されるわけではない。類似性評価の目標は、n個の最も類似したドキュメントを検出することであり、その結果それらは再閲覧提案として使用できるか、あるいはダイジェストに含むことができる。このようにして、多様な例示的な実施形態においては、n個の最も類似したドキュメントの集合が逆インデックスを使用して効率的に近似される。この近似は、オリジナルのドキュメントを、次にすべての既に見たドキュメントの逆インデックスに対して実行される短い照会に変換することによって達成される。これらの多様な例示的な実施形態においては、戻されたドキュメントは最終結果集合として処理されるか、あるいはさらに正確な結果のために正確な類似性を決定すべくオリジナル・ドキュメントと比較される。多様な例示的な実施形態においては、照会は、オリジナルのドキュメントをそのtf-idfベクトル表現に変換し、すべての用語をそれらのそれぞれの用語重みにより並べ替えてから、照会を上位n個の用語に制限することによって自動的に生成される。多様な例示的な実施形態においては、基調を成す検索エンジンが加重照会をサポートする場合、用語重みは個々の照会用語の重みとして使用される。

【0134】

サブ・ステップS4150では、サーバにより受け取られるドキュメントが逆インデックスに追加される。これは、既に見たドキュメントの効率的なリトリープを容易にし、リアル・タイム再閲覧提案を実現するために好適である。次に動作は、照会がアクセスされたドキュメントの特性に基づき生成されるサブ・ステップS4160に続行する。次に動作は、本方法の動作が図4のステップS420に戻るステップS4170に続行する。

【0135】

図6は、本発明に従って既にアクセスしたコンテンツに基づき焦点ダイジェスト・ドキュメントを生成するための方法の一つの例示的な実施形態を概略するフローチャートである。図6に示されるように、本方法はステップS600で始まり、ユーザが情報を含むドキュメントに対してアクションを実行するステップS610に続行する。多様な例示的な実施形態においては、該アクションは、ドキュメントをリトリープするアクション、ドキュメントを開くアクション、ドキュメントを保存するアクション、ドキュメントをファイルするアクション、ドキュメントを編集するアクション、ドキュメントを削除するアクション、ドキュメントを転送するアクション、及びドキュメントをブックマークするアクションの一つまたは複数を含む。多様な例示的な実施形態においては、例えば、ドキュメントにアクセスする、ドキュメントをレビューする、及び/又はドキュメントを編集するときにユーザがドキュメントで実行する他のドキュメント・アクションは、現在既知であるも

10

20

30

40

50

のと今後開発されるものを含み、本発明の範囲内にある。

【0136】

次にステップS620では、既にアクセスした関連コンテンツのダイジェスト・ドキュメントが生成される。多様な例示的な実施形態においては、これは現在表示されているテキスト特性に基づく。多様な例示的な実施形態においては、ダイジェスト・ドキュメント生成は、ステップS410に関連して前述された、既にアクセスしたコンテンツ情報を再閲覧するために利用される技法に類似した技法を使用して実行される。多様な例示的な実施形態においては、例えば、ウェブ・ページ、eメール・メッセージまたは別の種類のドキュメントなどの、ユーザのコンテクストに類似しているすべての既に見たコンテンツが、例えば、サブ・ステップS4160及びその過去のサブ・ステップに関連して前述された手法を使用してリトリープされる。次に、多様な例示的な実施形態においては、ウェブ・ページなどの新しいドキュメントは、ドキュメント・タイトル、オリジナル・ドキュメントまたはキャッシュに入れたテキストに対する参照、オプションの要約、及び、例えば、アクセス日付と画像などのダイジェストに含まれるようにユーザによって特定される他の情報を含んでコンパイルされる。10

【0137】

多様な例示的な実施形態においては、ステップS610は省略されてよく、本方法は、ステップS600から直接、既にアクセスした関連コンテンツのダイジェスト・ドキュメントが、ユーザによって定義されるダイジェスト特定に合わせて生成されるステップS620に続行してよい。20

【0138】

ステップS630では、ユーザは特定のリソースの包含または除外をインタラクティブに誘導するか、あるいは要約・パラメータを修正することにより生成されるドキュメントを繰り返して改善する。多様な例示的な実施形態においては、ステップS630が除外される。20

【0139】

多様な例示的な実施形態においては、ダイジェストのためのトピックまたは焦点の特定は、自動的に提供される。多様な例示的な実施形態においては、ダイジェストのためのトピックまたは焦点は、現在表示されているウェブ・ページ、調査報告またはeメール・メッセージなどの、ユーザの現在のコンテクストにより提供される。代わりに、多様な例示的な実施形態においては、ユーザは代表的なテキストを含むURLまたはファイルのリストを特定するか、ダイジェストに焦点を当てるためにテキスト照会を入力する。さらに、多様な例示的な実施形態においては、要約の所望の長さ、含まれるドキュメントの最大数、及び／又はドキュメントの日付範囲もしくはファイル・タイプについての情報が含まれる。30

【0140】

多様な例示的な実施形態においては、ダイジェスト生成構成要素は、現在既知の、または今後開発される技法によるドキュメントの自動的な要約にも対処する。

【0141】

テキスト・コンテンツに関する多様な例示的な実施形態においては、自動要約は、それらの文の特性に基づき一つまたは複数のドキュメントから文を選択することにより開始する。多様な例示的な実施形態においては、文は直接的に要約に含まれる、及び／又は分析される、及び／又は再公式化される。多様な例示的な実施形態においては、要約は、その長さを調整することによって、または文の特性により多くのまたはより少ない重みを与えることによって様々な目的に合わせられる。また、多様な例示的な実施形態においては、一般的であるよりも、ある特定の主題または照会に向けられる要約も生成される。40

【0142】

ステップS640では、要約を含む生成されたダイジェストが次にユーザに提供される。次に動作は、本方法の動作が停止するステップS650に続行する。

【0143】

10

20

30

40

50

本発明は、前記に概略された例示的な実施形態に関連して記載されている。多様な代替例、修正例、変形例及び／又は改良例は、既知であるのか、あるいは現在は予見されないかに関係なく本発明の精神及び範囲内にある。したがって、前記に述べられるような本発明の例示的な実施形態は、限定的ではなく例示的であることが意図される。多様な変更は、本発明の精神及び範囲から逸脱することなく加えられてよい。したがって、本発明はすべての既知の、または今後開発される代替例、修正例、変形例及び／又は改良例を包含することが意図される。

【図面の簡単な説明】

【0144】

【図1】本発明に従って既にアクセスしたコンテンツを再閲覧し、焦点ダイジェスト・ドキュメントを生成する方法及びシステムの一つの例示的な実施形態のハイ・レベル概略表現である。10

【図2】本発明による方法及びシステムと関連して使用するためのネットワーク環境の一つの例示的な実施形態を示す。

【図3】本発明に従って既にアクセスしたコンテンツを再閲覧し、焦点ダイジェスト・ドキュメントを生成するためのシステムの一つの例示的な実施形態の機能ブロック図である。20

【図4】本発明に従って既にアクセスしたコンテンツ情報を再閲覧するための方法の一つの例示的な実施形態を概略するフローチャートである。

【図5】本発明による方法及びシステムと関連して使用するためのステップS410の一つの例示的な実施形態をさらに詳細に概略するフローチャートである。20

【図6】本発明に従って既にアクセスしたコンテンツに基づき焦点ダイジェスト・ドキュメントを生成するための方法の一つの例示的な実施形態を概略するフローチャートである。30

【符号の説明】

【0145】

20 クライアント・モジュール

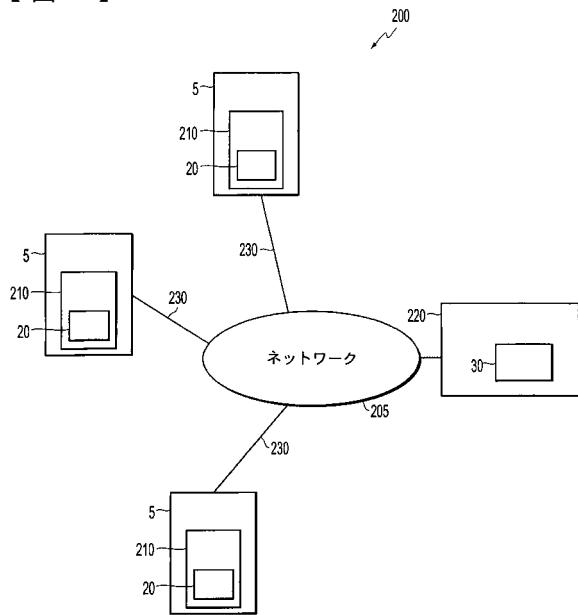
30 サーバ・モジュール

300 コンテンツ再閲覧及びダイジェスト生成システム

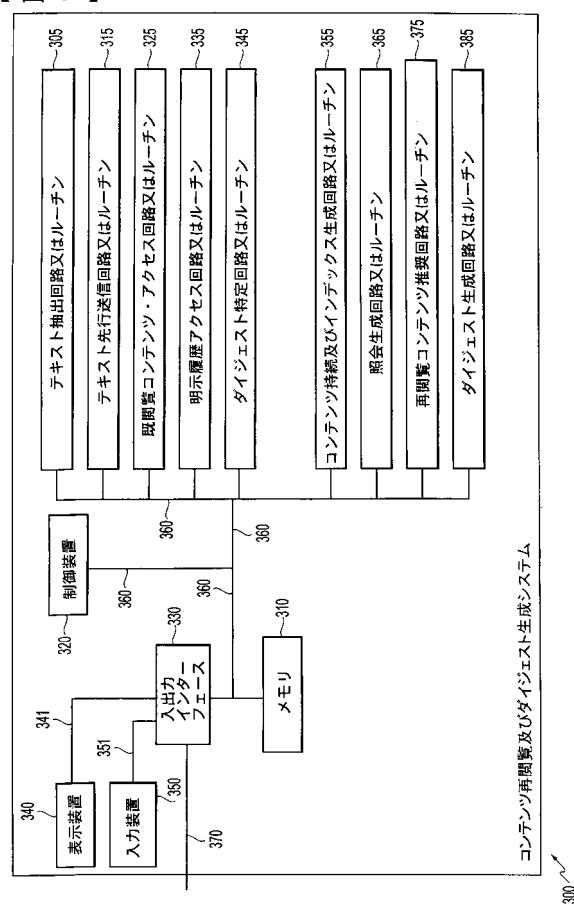
310 メモリ

320 制御装置

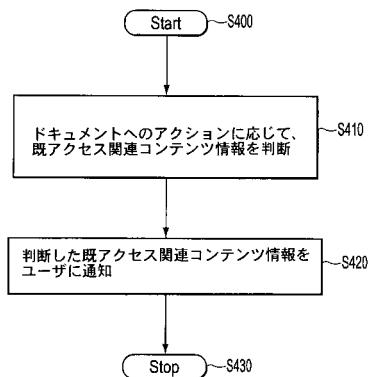
【図2】



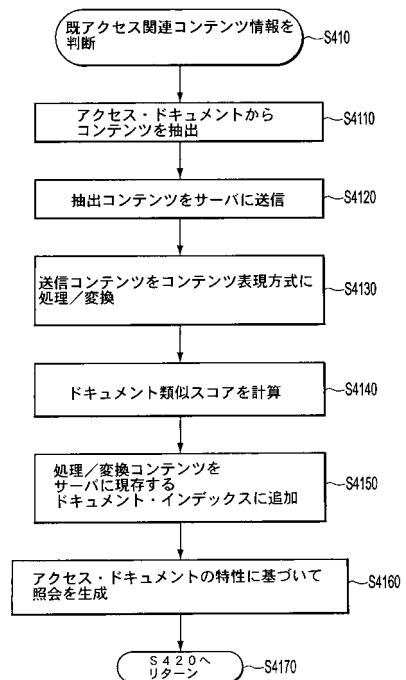
【図3】



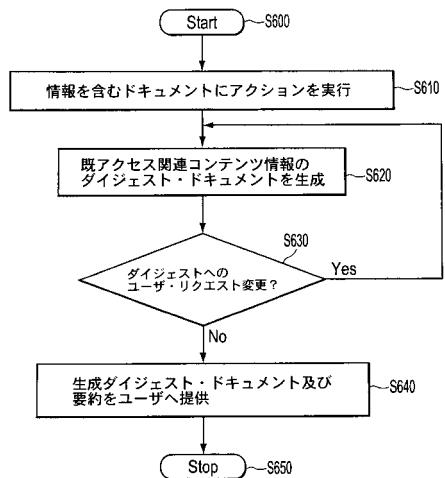
【図4】



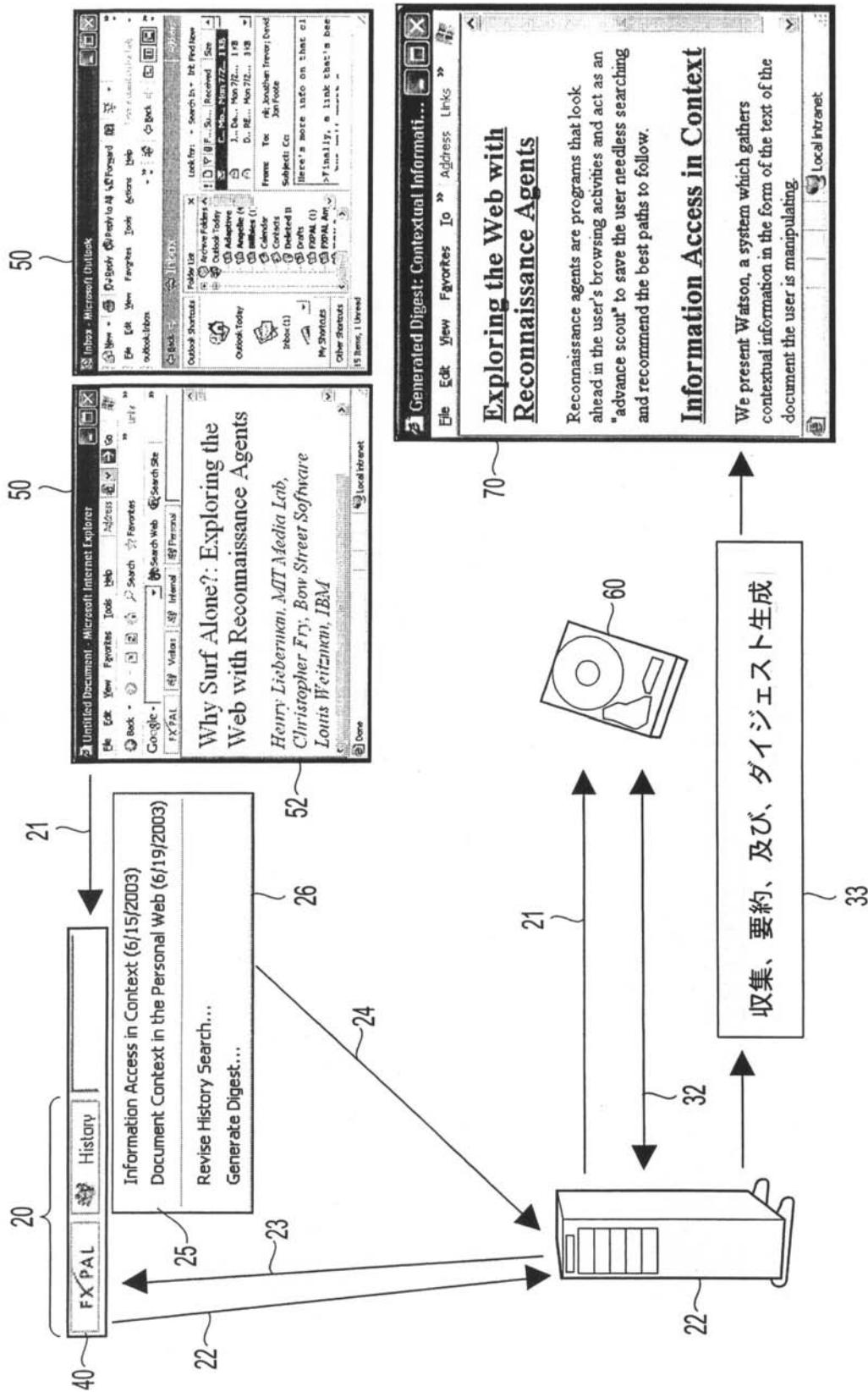
【図5】



【図6】



【図1】



フロントページの続き

(72)発明者 ジョナサン ジェイ . トレヴァー

アメリカ合衆国 95051 カリフォルニア州 サンタ クララ コズメル サークル 281
4

(72)発明者 クリストファー ディー . カリー

アメリカ合衆国 94943 カリフォルニア州 マウンテン ビュー ネドソン コート 25
13

(72)発明者 ローラン ドゥヌ

アメリカ合衆国 94313 カリフォルニア州 パロ アルト フィールディング 885

(72)発明者 ジーン ゴロブチンスキー

アメリカ合衆国 94306 カリフォルニア州 パロ アルト エル カミノ レアル 425
0 ナンバー シー327

F ターム(参考) 5B075 KK33 KK37 PQ02 PQ32 PR03

5B085 BA06 BG07