

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3680588号

(P3680588)

(45) 発行日 平成17年8月10日(2005.8.10)

(24) 登録日 平成17年5月27日(2005.5.27)

(51) Int. Cl.⁷

G06F 17/30

F I

G06F 17/30 419B

G06F 17/30 380D

請求項の数 31 外国語出願 (全 12 頁)

(21) 出願番号	特願平10-300275	(73) 特許権者	000005496
(22) 出願日	平成10年9月16日(1998.9.16)		富士ゼロックス株式会社
(65) 公開番号	特開平11-242549		東京都港区赤坂二丁目17番22号
(43) 公開日	平成11年9月7日(1999.9.7)	(74) 代理人	100079049
審査請求日	平成14年9月19日(2002.9.19)		弁理士 中島 淳
(31) 優先権主張番号	929426	(74) 代理人	100084995
(32) 優先日	平成9年9月15日(1997.9.15)		弁理士 加藤 和詳
(33) 優先権主張国	米国 (US)	(72) 発明者	ウィリアム ノア シリット
			アメリカ合衆国 94304 カリフォルニア州 パロ アルト ヒルビュー アベニュー 3400 ビルディング 4 ケアオブエフエックス パロ アルト ラボラトリー インコーポレイテッド

最終頁に続く

(54) 【発明の名称】 ドキュメント表示方法及び電子ドキュメントシステム

(57) 【特許請求の範囲】

【請求項1】

第1のドキュメントをディスプレイに表示する際に、前記第1のドキュメントに関連した少なくとも1つの別のドキュメントを示唆するための電子ドキュメントシステムにおいて、第1のドキュメントを表示する際に、別のドキュメントに対する少なくとも1つのリンクを表示し、別のドキュメントがそれぞれ前記第1のドキュメントに関連している第1のドキュメント表示方法であって、

前記電子ドキュメントシステムが

前記第1のドキュメントの少なくとも1つのユーザによる注記事項付きセグメントを識別するステップと、

前記第1のドキュメントの少なくとも1つの注記事項付きセグメントに関連した少なくとも1つの第2のドキュメントを識別するステップと、

前記第2のドキュメントのそれぞれに対し選択可能なリンクを前記第1のドキュメントの余白に表示するステップとの各処理を含む、

電子ドキュメントシステムのドキュメント表示方法。

【請求項2】

前記選択可能なリンクが、前記第1のドキュメントに対する後注として表示されることを特徴とする、請求項1に記載の方法。

【請求項3】

少なくとも1つの第2のドキュメントを識別するステップに、前記第1のドキュメント

10

20

に関連した少なくとも1つの第2のドキュメントの少なくとも1つの部分を識別するステップが含まれ、選択可能なリンクが、識別された少なくとも1つの部分を参照することを特徴とする、請求項1に記載の方法。

【請求項4】

前記識別された少なくとも1つの部分が、少なくとも1つの注記事項付きセグメントに関連するものとして識別され、前記選択可能なリンクが、前記識別された少なくとも1つの注記事項付きセグメントの近くに表示されることを特徴とする、請求項3に記載の方法。

【請求項5】

識別ステップが、前記第1のドキュメントの前記少なくとも1つのセグメントの注記事項10
項に回答することを特徴とする、請求項4に記載の方法。

【請求項6】

前記少なくとも1つの第2のドキュメントの識別ステップが、前記第1のドキュメントの前記少なくとも1つの注記事項付きセグメントに回答することを特徴とする、請求項5に記載の方法。

【請求項7】

前記選択可能なリンクが、前記少なくとも1つの注記事項付きセグメント付近の欄外に表示されることを特徴とする、請求項5に記載の方法。

【請求項8】

前記少なくとも1つの第2のドキュメントを識別するステップが、ユーザが識別する用語、及び、関連性フィードバック技法を用いて識別される用語に基づいて、関連性を判定するステップを含むことを特徴とする、請求項1に記載の方法。 20

【請求項9】

前記識別ステップが、重み付きの和による間合せを利用することを特徴とする、請求項8に記載の方法。

【請求項10】

さらに、
前記選択可能なリンクが選択されたか否かを判定するステップと、
前記選択可能なリンクの選択に回答して、前記識別された少なくとも1つの第2のドキュメントを表示するステップが含まれていることを特徴とする、 30
請求項1に記載の方法。

【請求項11】

前記識別ステップが、
前記第1のドキュメントに関連した複数の第2のドキュメントを識別するステップと、
前記識別された複数の第2のドキュメントをクラスタ化するステップと、
個々のクラスタ内の前記識別された複数の第2のドキュメントの全てを代表する、前記識別された複数の第2のドキュメントの少なくとも1つを各クラスタ毎に選択するステップを含み、前記選択可能なリンクが、前記識別された複数の第2のドキュメントから選択された少なくとも1つを参照することを特徴とする、
請求項1に記載の方法。 40

【請求項12】

前記関連性が、前記第1のドキュメントに対する前記少なくとも1つの第2のドキュメントの内容の類似性に基づいて判定されることを特徴とする、請求項1に記載の方法。

【請求項13】

さらに、前記識別された少なくとも1つの第2のドキュメントにフィルタリングを施して、所定の類似性しきい値未満である前記第2のドキュメントを排除するステップを含むことを特徴とする、請求項12に記載の方法。

【請求項14】

前記識別された少なくとも1つの第2のドキュメントのタイプを表すアイコンが表示されることを特徴とする、請求項1に記載の方法。 50

【請求項 15】

前記識別された少なくとも1つの第2のドキュメントのタイトルが表示されることを特徴とする、請求項1に記載の方法。

【請求項 16】

前記識別された少なくとも1つの第2のドキュメントの要約が表示されることを特徴とする、請求項1に記載の方法。

【請求項 17】

第1のドキュメントを表示する際に、前記第1のドキュメントに関連した少なくとも1つの第2のドキュメントを示唆するための電子ドキュメントシステムであって、

前記第1のドキュメントの少なくとも1つのユーザによる注記事項付きセグメントを識別し、前記第1のドキュメントの注記事項付きセグメントに関連した少なくとも1つの第2のドキュメントを識別するプロセッサと、

前記第1のドキュメントを表示する際に、前記識別された少なくとも1つの第2のドキュメントを参照する選択可能なリンクを前記第1ドキュメントの余白に表示するディスプレイを含むことを特徴とする、

電子ドキュメントシステム。

【請求項 18】

前記プロセッサが、前記第1のドキュメントに対する前記少なくとも1つの第2のドキュメントの少なくとも1つの部分の関連性に基づいて、前記少なくとも1つの第2のドキュメントを識別し、前記選択可能なリンクが、前記識別された少なくとも1つの部分を参照することを特徴とする、請求項17に記載のシステム。

【請求項 19】

前記プロセッサが、前記第1のドキュメントの少なくとも1つの注記事項付きセグメントを識別し、前記識別された少なくとも1つの注記事項付きセグメントに関連した前記少なくとも1つの部分を識別して、前記選択可能なリンクが、前記識別された少なくとも1つの注記事項付きセグメントの近くに表示されることを特徴とする、請求項18に記載のシステム。

【請求項 20】

前記選択可能なリンクが、前記第1のドキュメントに対する後注として表示されることを特徴とする、請求項17に記載のシステム。

【請求項 21】

さらに、ユーザ入力インターフェイスを含み、前記プロセッサが、ユーザによる前記第1のドキュメントの1つのセグメントに関する注記事項に回答して、前記少なくとも1つの第2のドキュメントを識別することを特徴とする、請求項17に記載のシステム。

【請求項 22】

前記選択可能なリンクが、少なくとも1つの注記事項付きセグメントに隣接した欄外に表示されることを特徴とする、請求項20に記載のシステム。

【請求項 23】

前記プロセッサが、ユーザの識別した用語、及び、関連性フィードバック技法に基づいて識別された用語に基づいて、前記少なくとも1つの第2のドキュメントを識別することを特徴とする、請求項17に記載のシステム。

【請求項 24】

前記プロセッサが、また、重み付けされた間合わせに基づいて少なくとも1つの第2のドキュメントを識別することを特徴とする、請求項23に記載のシステム。

【請求項 25】

さらに、ユーザ・インターフェイスを含み、前記ディスプレイが、ユーザによる前記選択可能なリンクの選択に回答して、前記識別された少なくとも1つの第2のドキュメントを表示することを特徴とする、請求項17に記載のシステム。

【請求項 26】

前記プロセッサが、前記第1のドキュメントに対する関連性に基づいて複数の第2のド

キュメントを識別し、前記識別された複数の第2のドキュメントをクラスタ化し、個々のクラスタ内における識別された複数の第2のドキュメントの全てを代表する、識別された複数の第2のドキュメントの少なくとも1つを各クラスタ毎に選択して、前記選択可能なリンクが、前記複数の第2のドキュメントから選択された少なくとも1つを参照することを特徴とする、請求項17に記載のシステム。

【請求項27】

前記プロセッサが、内容の類似性に基づいて、第1のドキュメントに対する前記少なくとも1つの第2のドキュメントの関連性を判定することを特徴とする、請求項17に記載のシステム。

【請求項28】

さらに、前記複数の第2のドキュメントのうち識別された少なくとも1つにフィルタリングを施して、所定の類似性しきい値未満である前記第2のドキュメントを排除するためのフィルタを含むことを特徴とする、請求項27に記載のシステム。

【請求項29】

前記ディスプレイが、前記識別された少なくとも1つの第2のドキュメントのタイプを表すアイコンをも表示することを特徴とする、請求項17に記載のシステム。

【請求項30】

前記ディスプレイが、前記識別された少なくとも1つの第2のドキュメントのタイトルをも表示することを特徴とする、請求項17に記載のシステム。

【請求項31】

前記ディスプレイが、前記識別された少なくとも1つの第2のドキュメントの要約をも表示することを特徴とする、請求項17に記載のシステム。

【発明の詳細な説明】

【0001】

〔発明の背景〕

【発明の属する技術分野】

本発明は、一般に、電子ドキュメント（文書）読み取りシステムに関するものである。特に、本発明は、第1のドキュメントを表示する際に、他の関連ドキュメントを示唆する電子ドキュメント読み取りシステムに関するものである。

【0002】

【従来の技術及び発明が解決しようとする課題】

ユーザが関連していると識別したドキュメントに類似したドキュメントの検索は、関連性フィードバック（relevance feedback）として知られている。関連性フィードバックについては、参考として本明細書においてその全容が援用されている、“Introduction to Modern Information Retrieval”, G. Salton et al., McGraw Hill, (1983)に記載がある。関連性フィードバックを支援するインターフェイスは、従来、読者側で行う明示的アクションを必要とし、関連ドキュメントを自発的に示唆するものではない。ウィンドウ・ベースのコンピューティング環境に合わせて設計された情報検索インターフェイスは、一般に、独立したウィンドウにおけるリストを介して、あるいは、可視ドキュメントを検索結果に置き換えることによって、他の関連ドキュメントに関する検索結果を提示する。これらのシステムは、極めて邪魔であり、読み取りプロセスを中断する。

【0003】

ハイパーテキスト・インターフェイスは、リンクを含んでいる欄外（余白）（margin）を設けるか、「Hyperties」が先駆けとなったやり方で、ソース・ドキュメントのテキストにリンクを埋め込むことによって、ソース・ドキュメントに関連したドキュメントに対するリンクを表示する。このシステムについては、参考として本明細書においてその全容が援用されている、“User Interface Design for the Hyperties Electronic Encyclopedia” by

10

20

30

40

50

Shneiderman, Proceedings of Hypertext'87, November 1987, Chapel Hill, NC に記載がある。しかし、これらのリンクは、静的であり、ハイパーテキストの作者によってソース・ドキュメントと共に生成される。Trellis のように、システムによっては、リンクを動的に表示するものもあるが、あらかじめ定義されたリンクの固定された集合からのものに限られる。Trellis については、参考として本明細書においてその全容が援用されている、“Programmable Browsing Semantics and Trellis” by R. Furuta et al. Proceedings of Hypertext'89, November 1989, Pittsburgh, PA, ACM Press に記載がある。

10

【0004】

HieNet システムでは、ノード間類似性測度 (inter-node similarity measure) を利用し、ハイパーテキスト作者によってあらかじめ生成されたリンクに基づいて、ハイパーテキスト・リンクが生成される。このシステムについては、参考として本明細書においてその全容が援用されている、“Hienet: A User-Centered Approach for Automatic Link Generation”, D. T. Chang, Proceedings of Hypertext'93, November 1993, Seattle, WA, ACM Press に記載がある。作者がドキュメント A からドキュメント B へのリンクを生成する場合、システムは、ドキュメント A に類似した全てのドキュメントからドキュメント B に類似した全てのドキュメントへのリンクを自動的に追加する。これらの自動生成されたリンクのアンカ (anchor) は、さまざまなドキュメントの欄外のアイコンによって体现される。アイコンをクリックすると、問合せ (query) に対する関連性によってランク付けされた、可能性のある目的ドキュメントのリストを含んだポップ・メニューが表示される。やはり、このシステムも、作者によってあらかじめ生成されたリンクに依存する。

20

【0005】

他の従来システムは、検索結果を表示するハイパーテキストに似た方法に関連する。HieNet は、欄外に自動リンクを表示するが、欄外におけるアンカは、アンカ付近の一節 (passage) の内容とは無関係である。HieNet は、ドキュメント間リンクと一節・ドキュメント間リンクとを区別しない。さらに、HieNet は、欄外のリンクを介して到達可能なドキュメントの数及び性質を表示しない。

30

【0006】

視覚化情報検索システム (今後は VOIR と呼ぶ) については、それぞれ、参考として本明細書においてその全容が援用されている、“Queries? Links? Is There a Difference?”, Proceedings of CHI'97, G. Golovinsky, March 1997, Atlanta, GA, ACM Press、及び、“What the Query Told the Link: The Integration of Hypertext and Information Retrieval”, Proceedings of Hypertext'97, G. Golovinsky, April 1997, Southampton, UK, ACM Press に記載がある。VOIR は、選択されたアンカの周りのテキストから算定される問合せとのハイパーテキスト・リンクを動的に生成し、解決するメカニズムである。VOIR は、問合せを利用して、選択されたアンカを含む一節に関連したドキュメントの集合を検索する。VOIR は、あらかじめ設定された関係を有するユーザ・リンクを示さない。それどころか、問合せを提示して、関係を確立するため、ユーザは、休止して、アンカを選択しなければならない。VOIR は、読み取りプロセスを容易化するよりも、特に、対話式情報検索を支援するように設計された。従って、VOIR の焦点は、ドキュメント間のナビゲーションを支援することである。従って、ユーザが多量の認知努力をブラウジング (閲覧) に傾注することが期待される。さらに、VOIR の場合、ユーザはドキュメントに注記 (annotate) またはタグを付けることができない。VO

40

50

IRは、特定の表示を生成するために、どのリンクが選択されたかについても指示しない。

【0007】

リメンブランス・エージェント（今後はRAと呼ぶ）と呼ばれるバックグラウンド情報検索プロセスについては、参考として本明細書においてその全容が援用されている、“A Continuously Running Automated Information Retrieval System”, B. J. Rhodes et al. Proceedings of The First International Conference on the Practical Application of Intelligent Agents in Multi-Agent Technology, PAAM'96, April, 1997, London, UKに記載がある。RAは、EMACSテキスト・ウィンドウにおいて動作し、ユーザがタイプしたテキストの最後の数行に関連したドキュメントを示唆する。RAは、ユーザの私有データを検索し、タイプされているテキストに関連したドキュメントを示唆するように設計されている。しかし、これらの示唆は、束の間のものであり、現在書かれているテキストだけに関するものである。RAは、ユーザがドキュメントを編集する際、示唆を絶えず置換するので、読み取りタスクの助けにはならない。

10

【0008】

QRLは、テキスト上へのインク状のマークを利用して、ブール代数による問合せを行う、問合せベースの情報検索インターフェイスである。このシステムについては、参考として本明細書においてその全容が援用されている、“Queries - R - Links: Graphical Markup for Text Navigation”, by G. Golovchinsky et al., Proceedings of INTERCHI'93, April 1993, Amsterdam, The Netherlands, ACM Pressに記載がある。問合せ用語は、矩形を用いて選択される。ラインが矩形を接続して、ブール代数でのAND演算子を表す。

20

【0009】

これらのシステムは、全て、関連ドキュメントに対するリンクの生成、または、書き込みの支援だけに大量のユーザ対話を必要とする。受動的に、邪魔にならないように関連ドキュメントに対するリンクを生成して、読み取りを支援する、電子ドキュメント読み取りシステムが必要である。

30

【0010】

【課題を解決するための手段】

〔発明の概要〕

本発明によれば、読み取りプロセスを妨害することなく、読者に関連したドキュメントを受動的に示す方法及びシステムが得られる。

【0011】

本発明によれば、さらに、読者が読み取っているソース・ドキュメントに与える作用に基づいて、読者にとって興味があるかもしれないドキュメントを自動的に検出することによって、直観的な読み取り支援が可能になる。人がテキストを読む場合には、興味のある、あるいは、議論の余地のある一節及び用語を強調するために、注記事項を付けることが多い。こうしたマーク及び走り書きの存在あるいは密度は、読者が特定の一節に抱く相対的な関心の度合を示すインジケータとして用いることが可能である。読み取られているドキュメントに関連した多量のドキュメントが利用可能ならば、読者は、読み取りプロセスの一部として関連ドキュメントを見出すことに興味を持つ可能性がある。

40

【0012】

ユーザにとって重要な特定の一節に関連したドキュメントに対する参照が、ソース・ドキュメントの欄外に配置され、全体的にソース・ドキュメントに類似したドキュメントに対する参照が、後注として挿入される。本発明のシステム及び方法は、リンクが識別される

50

と、それを維持して、非線形読み取り及び走り読みを容易にする。

【0013】

ユーザの関心は、ソース・ドキュメントを読んでいる間になされる注記事項から推断される。従って、本発明のシステム及び方法によれば、次の2つのやり方で認知オーバーヘッドが最小限に抑えられる：1) ソース・ドキュメントに関連したドキュメントを識別するのに表現力に富んだ問合せを必要としない、2) 関連ドキュメントに対する選択可能なリンクが、邪魔にならないようにドキュメントの欄外及び終わりに設けられる。これが、それぞれ、図2及び3に示されている。

【0014】

システムには、さらに、ユーザに一方的な対話をする重荷を負わせるのではなく、他の対話と互換性のあるやり方で、読者に対する示唆を導入する。示唆されたドキュメントは、選択可能なリンクに従うことによってアクセス可能である。しかし、ユーザは、示唆がなされても、それに対してアクションを起こす必要はない。そうではなく、ユーザは、そうするのが道理にかなう場合に、その示唆に対するアクションをとることができる。本発明のシステム及び方法は、アイコンで参照ドキュメントのタイプを表し、アイコンにテキスト・ラベルを付与して、ユーザがリンクのターゲットをよりよく理解できるようにする。

【0015】

本発明の以上の及びその他の特徴及び利点については、望ましい実施形態に関する下記の詳細な説明において示されるか、あるいは、前記説明から明らかになるであろう。

【0016】

【発明の実施の形態】

〔好ましい実施の形態の詳細な説明〕

図1には、本発明によるドキュメント読み取りシステム10のブロック図が示されている。ドキュメント読み取りシステム10には、ディスプレイ18においてユーザが現在読み取り中のソース・ドキュメント16を記憶して、第1のメモリ14との通信を行うプロセッサ12が含まれている。プロセッサ12は、関連している可能性のあるターゲット・ドキュメント22を記憶している第2のメモリ20との通信も行う。ユーザは、マウス26、キーボード28、または、ペン・ベースのインターフェイス30といった、任意の数の従来の入力/出力装置24を介して、ドキュメント読み取りシステム10と対話し、制御を行う。入力/出力装置24は、入力/出力インターフェイス31と通信し、前記インターフェイスは、さらに、プロセッサ12と通信する。

【0017】

図1に示すように、システム10は、プログラムされた汎用コンピュータ上で実施するのが望ましい。しかし、システム10は、専用コンピュータ、プログラムされたマイクロプロセッサ、または、マイクロコントローラと必要な周辺集積回路素子、ASICまたは他の集積回路、個別素子回路のような配線による(hard-wired)電子または論理回路、PLD、PLA、FPGA、または、PAL等のようなプログラム可能な論理素子を用いて実施可能である。一般に、図4に示すフローチャートを実施することが可能な有限な状態を持つマシン(finite state machine)を有する装置を、システム10を実施するのに利用することもできる。

【0018】

さらに、図1に示すように、記憶装置すなわちメモリ14及び20は、スタティックまたはダイナミックRAMを利用して実施されるのが望ましい。しかし、装置14及び20は、フロッピ・ディスク及びディスク・ドライブ、書き込み可能光ディスク及びディスク・ドライブ、ハード・ドライブ、フラッシュ・メモリ等を用いて実施することも可能である。また、認識しておくべき点は、装置14及び20は、単一メモリの別個の部分であっても良いし、あるいは、物理的に別個のメモリであっても良いことである。

【0019】

さらに、認識しておくべき点は、装置14及び20とプロセッサ12を接続するリンク15及び17は、ネットワーク(図示せず)に対する配線式リンクとすることもできるし、

10

20

30

40

50

あるいは、無線式リンクとすることも可能であるという点である。ネットワークは、ローカル・エリア・ネットワーク、広域ネットワーク、イントラネット、インターネット、または、他の任意の分散処理及び記憶ネットワークとすることが可能である。この場合、電子ドキュメント16は、リンク15を介して物理的に遠隔のメモリ装置14から引き出され、プロセッサ12において、概要を後述する方法に従って処理される。この場合、電子ドキュメント16は、システム10の他のメモリ装置の一部(図示せず)に局部的に記憶することが可能である。

【0020】

本発明の方法では、各ソース・ドキュメント16毎に、2種類のターゲット・ドキュメント22を識別する。2つのタイプのターゲット・ドキュメントとは、1)特に注記事項を付けられた一節に特に関連したターゲット・ドキュメントと、2)ソース・ドキュメント全体に一般に関連したターゲット・ドキュメントである。ソース・ドキュメントとターゲット・ドキュメント22の間に、ある関係が確立されると、表示されたドキュメント16内の選択可能なリンクをクリックすることによって、ターゲット・ドキュメントを表示することが可能になる。

【0021】

図2には、2つのタイプのターゲット・ドキュメント22への参照が示されている。ソース・ドキュメント16における特定の一節32に関連したターゲット・ドキュメント22が、ソース・ドキュメント16の関連する一節32付近の欄外に配置された欄外表示34によって識別される。図3に示すように、全体としてソース・ドキュメント16に関連したターゲット・ドキュメント22は、注記事項が付けられ、ソース・ドキュメントに対する後注36として示される。後注には、タイプ、タイトル及び要約の情報が含まれる。

【0022】

図4は、本発明の方法の1つの実施形態に関する制御ルーチンの概略を示すフロー・チャートである。制御ルーチンは、ステップS100から開始され、引き続きステップS105に移行する。ステップS105において、制御ルーチンは、ユーザが注記事項を施したか否かを確認する。注記事項が施されていないければ、制御ルーチンはループを辿ってステップS105に戻る。注記事項が施されていれば、制御ルーチンは引き続きステップS110に移行する。ステップS110において、制御ルーチンは、ユーザによるソース・ドキュメントの注記事項を判定する。次に、ステップS120において、制御ルーチンは、ソース・ドキュメントのテキスト及び注記事項を分析して、注記事項が施されている一節を確認する。一節には、欄外・バーでマーキングされたパラグラフ、下線を引いた文または句、または、1つ以上の円で囲んだ用語の内容が含まれる可能性がある。次に、ステップS130において、制御ルーチンは、その一節から問合せを生成する。問合せには、円で囲んだワード(word)に重要性を付与するために重み付けがなされ、識別された一節からの内容を備える用語が含まれている。次に、ステップS140において、制御ルーチンは、問合せを利用して、ターゲット・ドキュメントを検索し、その一節に関連したドキュメントを識別する。次に、ステップS150において、検索結果がクラスタ化される。クラスタ化は、参考として本明細書においてその全容が援用されている、“Reexamining the Cluster Hypothesis: Scatter/Gather on Retrieval Results”, M. A. Hearst et al. Proceedings of ACM SIGIR'96, August 1996, Zurich, Switzerlandに記載のものと同様の方法で実施するのが望ましい。

【0023】

次に、ステップS160において、制御ルーチンは、各クラスタから典型的なドキュメントを選択する。これらのドキュメントは、ステップS170において、ユーザ指定の類似性しきい値(similarity threshold)によってフィルタリングされる。次に、ステップS180において、選別されて残ったドキュメントが、ソース・ドキュメントの問合せが生

10

20

30

40

50

成された一節付近の欄外に、該ドキュメントに対するリンクを表示することによって識別される。各選択可能なリンクは、選択され、フィルタリングを施されたターゲット・ドキュメントのタイプ及び短いタイトルを表すアイコンとすることが可能である。

【0024】

次に、ステップS190において、制御ルーチンは、ユーザが、現在のソース・ドキュメント内における選択可能なリンクを選択したか否かを確認する。ステップS190において、ユーザが選択可能なリンクを選択した場合、制御ルーチンはステップS200に進む。ステップS200において、ターゲット・ドキュメントが、新たな現在のソース・ドキュメントとして表示されると、制御ルーチンは、引き続きステップS105に戻り、第2の注記事項が施されるのを待つ。あるいはまた、ステップS190において、選択可能なリンクが選択されていない場合、制御ルーチンは、ジャンプして、直接ステップS105に戻る。制御は、ユーザが、ディスプレイ18に表示されている全ての開いたソース・ドキュメントを閉じるまで、続行される。

10

【0025】

後注を算定するためには、わずかに修正を加えた、図4のフロー・チャートを用いることが可能である。制御ルーチンは、ステップS100からステップS120までの、欄外の注の生成に関する説明と同じように進行する。しかし、ステップS130では、重み付けされた和による問合せが生成される。ステップS130では、読者によって明示的に識別される用語、及び、標準的な関連性フィードバック技法によって識別される用語を用いて、ステップS130の重み付けされた和による問合せが構成される。識別された用語は、ドキュメントに対して施された注記事項に基づいて重み付けされる。例えば、ユーザによって特にはっきりと選択されたワードは、最高の重み付けがなされ、選択されたパラグラフに生じるワードには、ソース・ドキュメントの残りの用語よりも高い重み付けがなされる。

20

【0026】

ステップS130において生成された重み付けされた和による問合せを利用して、そのドキュメントに関連するものとして識別されたドキュメントは、リンクが、ステップS180において、欄外の注としてではなく、後注として表示される点を除けば、残りのステップS140～S200と同様のやり方で処理が施される。

【0027】

これらの制御ルーチンの一方または両方は、本発明のドキュメント読み取りシステムのバックグラウンドで実行されている可能性があることは理解されるべきである。

30

【0028】

オプションにより、本発明のシステム及び方法は、参考として本明細書においてその全容が援用されている、“A Trainable Document Summarizer”, J. Kupiec et al., Proceedings of SIGIR'95, Pittsburgh, PA, ACM Pressに記載のものと同様のやり方で、自動テキスト要約プロセスを介して、ドキュメントから要約を導き出すことが可能である。そこでは、要約は後注として表示される。

【0029】

本明細書において用いられる注記事項という用語は、テキスト、デジタル・インク、オーディオ、ビデオ、または、ドキュメントに関連した他の任意の入力を含めることを意図したものであることは理解されるべきである。また、ドキュメントという用語は、テキスト、ビデオ、オーディオ、及び、他の任意の媒体及び任意の媒体の組み合わせを含めることを意図したものであることも理解されるべきである。さらに、テキストという用語は、テキスト、デジタル・インク、オーディオ、ビデオ、または、ドキュメント構造を含むドキュメントの他の任意の内容を含めることを意図したものであることも理解されるべきである。

40

【0030】

本発明の説明は、上記概要を示した特定の実施形態に関連して行ってきたが、当業者には

50

、多くの代替案、修正、及び、変更が明らかであり、また、明らかになるであろう。従って、上述の望ましい実施形態は、例証となるものであって、制限するものではない。前記請求項に記載の本発明の精神及び範囲を逸脱することなく、さまざまな変更を加えることが可能である。

【図面の簡単な説明】

【図1】本発明の電子ドキュメント読み取りシステムの実施形態の1つに関するブロック図である。

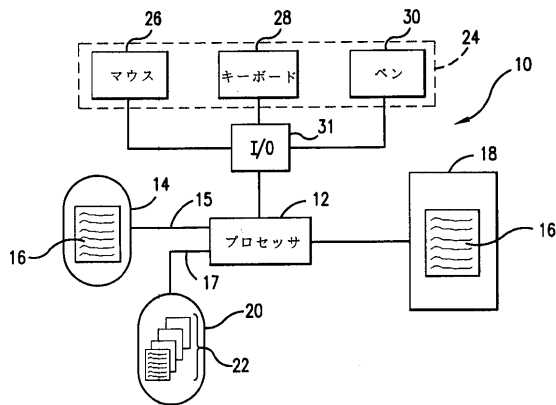
【図2】注記事項付きの一節に隣接して欄外内にアイコンを備えるソース・ドキュメントを示す図である。

【図3】後注を備えたもう1つのソース・ドキュメントの図である。

【図4】本発明の実施形態の1つに関する制御ルーチンの概要を示すフロー・チャートである。

10

【図1】



【図2】

has been demonstrated [28], it still relies on hierarchical structure of each document to support local navigation. SuperBook has been shown to be an effective interface for IR tasks when browsing structured collections [12], but alternatives to the book metaphor must be found to support browsing through loosely-structured hypertext collections.

One such alternative — the newspaper metaphor — is discussed in the following section, and VOIR, a prototype that implements it, is described. Some experimental results from an evaluation of VOIR are presented, and the paper concludes with a discussion of possible extensions and applications of this query-mediated hypertext interfaces.

2 VOIR

This section describes VOIR (Visualization of Information Retrieval), a prototype newspaper-based dynamic hypertext interface. The section first introduces the newspaper metaphor and discusses its implementation in VOIR. A description of VOIR's linking interface follows, and the discussion concludes with an overview of VOIR's visualization features.



32


【 図 3 】

5 CONCLUSIONS

This paper describes VOIR, a query-mediated hypertext interface designed to support information exploration tasks in very large text databases. It describes a technique for mediating links with passage-based relevance feedback queries. Experimental results indicate that link-based queries produced better performance than passage selections, and that subjects found dynamic hypertext interfaces intuitive to use. This research suggests that integrating hypertext interfaces with full-text search engines can produce effective solutions for a class of information exploration tasks.

This work also has implications for models of information exploration (e.g., [37]) that posit a distinction between selecting anchors and forming queries. Interfaces such as the one described here suggest that the distinction between hypertext and information retrieval can become progressively blurred. As more sophisticated (e.g., agent-based) techniques are integrated into information exploration interfaces (e.g., [26]), some distinctions between semantic and statistical links should also disappear.

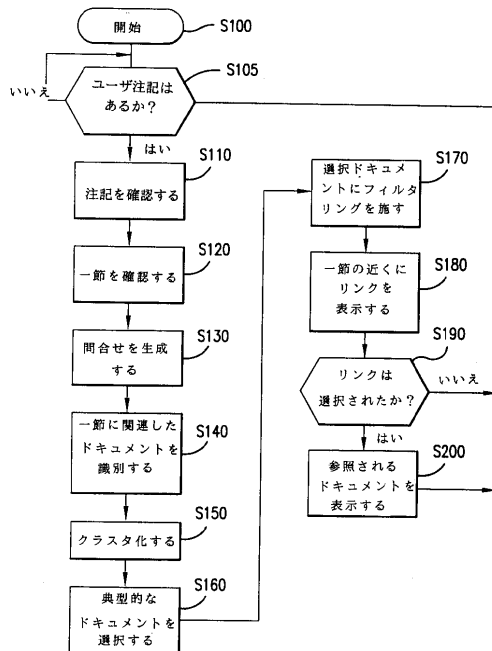
Related Documents

 G. Golovchinsky, *Queries? Links? Is There a Difference?*, In *Proceedings of CHI '97*, (March 1997, Atlanta, GA), ACM Press.

16

36

【 図 4 】



フロントページの続き

- (72)発明者 モーガン エヌ . プライス
アメリカ合衆国 94306 カリフォルニア州 パロ アルト パーク ブルバード 3875
- (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 リアル 425
0 ナンバーシー327
- (72)発明者 マーク ディー . ワイザー
アメリカ合衆国 94301 カリフォルニア州 パロ アルト グリーンウッド アベニュー
1144

審査官 野崎 大進

- (56)参考文献 特開平11-045269(JP,A)
特開平02-161543(JP,A)
特開平07-325827(JP,A)
特開平08-314790(JP,A)
特開平09-073463(JP,A)
特開平09-128403(JP,A)

- (58)調査した分野(Int.Cl.⁷, DB名)
G06F 17/30 419
G06F 17/30 380
JSTPlusファイル(JOIS)