

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5765452号
(P5765452)

(45) 発行日 平成27年8月19日(2015.8.19)

(24) 登録日 平成27年6月26日(2015.6.26)

(51) Int. Cl. F 1
G 0 6 F 17/22 (2006.01) G O 6 F 17/22 6 4 7
G 0 6 F 17/24 (2006.01) G O 6 F 17/24 6 1 0

請求項の数 4 (全 16 頁)

(21) 出願番号	特願2014-8130 (P2014-8130)	(73) 特許権者	000005223
(22) 出願日	平成26年1月20日 (2014.1.20)		富士通株式会社
(62) 分割の表示	特願2010-75512 (P2010-75512) の分割		神奈川県川崎市中原区上小田中4丁目1番 1号
原出願日	平成22年3月29日 (2010.3.29)	(74) 代理人	100107766
(65) 公開番号	特開2014-81958 (P2014-81958A)		弁理士 伊東 忠重
(43) 公開日	平成26年5月8日 (2014.5.8)	(74) 代理人	100070150
審査請求日	平成26年1月20日 (2014.1.20)		弁理士 伊東 忠彦
		(74) 代理人	100146776
			弁理士 山口 昭則
		(72) 発明者	糸 照宣
			神奈川県川崎市中原区上小田中4丁目1番 1号 富士通株式会社内

最終頁に続く

(54) 【発明の名称】 アノテーション付与復元方法及びアノテーション付与復元装置

(57) 【特許請求の範囲】

【請求項1】

アノテーション付与復元方法であって、
コンピュータに、

構造を有するコンテンツ上のいずれかの位置についてアノテーションを付与する際に、
前記コンテンツにおける前記位置の上位側の構造の記憶情報及び下位側の構造の記憶情報
を前記アノテーションと関連づけて記憶する処理と、

前記コンテンツについてアノテーションが付与された位置の上位側の構造の記憶情報に
基づいて、アノテーション復元対象のコンテンツ中に、記憶された前記上位側の構造に合
致する構造が含まれることを検出できなければ、アノテーションが付与された位置の下位
側の構造の記憶情報に基づいて、アノテーション復元対象のコンテンツ中に、記憶された
前記下位側の構造に合致する構造が含まれるかの検出を行い、記憶された前記下位側の構
造に合致する構造が含まれることを検出すると、前記位置に対応づけてアノテーションを
出力する処理とを実行させ、

記憶される前記位置の下位側の構造は、前記位置の下位側の複数のパスの中から出現頻
度に応じて選択された1つのパスに関する構造であること
を特徴とするアノテーション付与復元方法。

【請求項2】

記憶される前記位置の下位側の構造は、前記位置の下位側の複数のパスの中から出現頻
度が最も少ない1つのパスに関する構造である、こと

を特徴とする請求項 1 記載のアノテーション付与復元方法。

【請求項 3】

記憶される前記位置の下位側の構造は、前記位置を頂点とする木構造の一つであることを特徴とする請求項 1 記載のアノテーション付与復元方法。

【請求項 4】

構造を有するコンテンツ上のいずれかの位置についてアノテーションを付与する際に、前記コンテンツにおける前記位置の上位側の構造の記憶情報及び下位側の構造の記憶情報を前記アノテーションと関連づけて記憶部に記憶させる手段と、

前記コンテンツについてアノテーションが付与された位置の上位側の構造の記憶情報に基づいて、アノテーション復元対象のコンテンツ中に、記憶された前記上位側の構造に合致する構造が含まれることを検出できなければ、アノテーションが付与された位置の下位側の構造の記憶情報に基づいて、アノテーション復元対象のコンテンツ中に、記憶された前記下位側の構造に合致する構造が含まれるかの検出を行い、記憶された前記下位側の構造に合致する構造が含まれることを検出すると、前記位置に対応づけてアノテーションを表示部に出力する制御を行う手段と、を備え、

記憶される前記位置の下位側の構造は、前記位置の下位側の複数のパスの中から出現頻度に応じて選択された 1 つのパスに関する構造であること

を特徴とするアノテーション付与復元装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、利用者からの付与要求に基づき、コンテンツに基づく表示にアノテーションを付与し、利用者からの参照要求に基づき、コンテンツに基づく表示にアノテーションを復元させるアノテーション付与復元方法及びアノテーション付与復元装置に関する。

【背景技術】

【0002】

近年、インターネットやイントラネット等のネットワーク上のコンテンツ（例えば HTML データ）に基づく表示（例えば Web ページ）に二次情報としてアノテーション（注釈）を付与する Web アノテーション技術が、よく利用されるようになった。なお、アノテーションとは、コンテンツ（HTML データ）に基づく表示（Web ページ）に貼り付けられる内容をいう。

【0003】

Web アノテーション技術は、ネットワーク上のコンテンツに対してコンテンツの作者と異なる別の第三者が、コンテンツに変更を加えることなく、Web ページに対して付箋（コメントや意見など）、下線、マーカ等のアノテーションを付与し、Web ページの閲覧（参照）時に、Web ページに付与したアノテーションを復元させるものである。

【0004】

図 1 は付箋紙のような形式のアノテーションが表示された Web ページの一例の画面イメージ図である。図 1 中、アノテーション 1 は Web ページ上に付箋紙のような形式で表示されている。

【0005】

Web アノテーション技術はアノテーション 1 が付与されたときに、アノテーション 1 が付与された Web ページの所在情報（例えば URL）を保持しておき、その URL で指定される Web ページの閲覧時に、Web ページに付与したアノテーション 1 を復元することにより、時間をおいての Web ページの閲覧や、アノテーション 1 を付与したアノテータと異なる第三者による Web ページの閲覧であっても、アノテーション 1 を復元できる。このような Web アノテーション技術は、情報共有、校正支援、レビュー支援、調査支援、広告システムなど、広く利用されている。

【0006】

従来、Web ページに対してアノテーション 1 を付与し、アノテーション 1 を複数の利

10

20

30

40

50

ユーザーで共有する場合は、仲介サーバにWebページの所在情報と、Webページに付与されたアノテーション1の構造位置情報及び内容を登録しておく。構造位置情報はアノテーション1が付与されたWebページのコンテンツ上の位置を表している。なお、構造位置情報は木構造のパス表現(XPath)で表現される。

【0007】

利用者が所在情報を指定してWebページを要求すると、仲介サーバはWebサーバからコンテンツを取得する。仲介サーバは利用者により指定された所在情報に対応するアノテーション1が登録されていれば、構造位置情報で表現されるコンテンツ上の位置にアノテーション1を復元して利用者のクライアントに送信する。このように、従来のWebアノテーション技術では、Webページに付与されたアノテーション1を複数のユーザーで共有していた(例えば引用文献1参照)。

10

【先行技術文献】

【特許文献】

【0008】

【特許文献1】特開2003-150542号公報

【発明の概要】

【発明が解決しようとする課題】

【0009】

従来のWebアノテーション技術は、アノテーションを付与したWebページのコンテンツ上の位置をXPathで表現している。従来のWebアノテーション技術では、アノテーションを復元する場合、登録しておいたXPathを利用して、アノテーションを復元するコンテンツ上の位置を特定する。

20

【0010】

ところで、Webアノテーション技術はコンテンツの作者と異なる別の第三者がWebページにアノテーションを付与するものである。したがって、作者は付与されたアノテーションに関係なく、コンテンツの内容を変更することもある。コンテンツの内容が変更されると、コンテンツの木構造が変化してしまい、登録しておいたXPathの意味が無くなってしまふ。XPathの意味が無くなってしまふと、従来のWebアノテーション技術ではアノテーションを復元するコンテンツ上の位置を特定できなくなるという問題があった。

30

【0011】

図2はXPath取得時のコンテンツ及び木構造を表した説明図である。図2(A)はXPath取得時のコンテンツを表している。図2(B)は図2(A)に示したコンテンツの木構造を表している。図2のコンテンツにおいて、タグ<TABLE>を表すXPathは「/HTML/BODY/TABLE」となる。

【0012】

図3は内容が変更されたコンテンツ及び木構造を表した説明図である。図3(A)は内容が変更されたコンテンツを表している。図3(B)は図3(A)に示したコンテンツの木構造を表している。図3のコンテンツにおいて、タグ<TABLE>を表すXPathは「/HTML/BODY/DIV/TABLE」となる。

40

【0013】

このように、図2に示すコンテンツが図3に示す内容に変更された場合、XPath取得時に登録しておいたXPath「/HTML/BODY/TABLE」は図3のコンテンツに存在しなくなるため、アノテーションを復元するコンテンツ上の位置を特定できなくなる。

【0014】

従来のWebアノテーション技術では、タグ<TABLE>が変更により削除されてしまった場合を除き、コンテンツ内に存在しているにもかかわらず、アノテーションを復元するコンテンツ上の位置を特定できないという問題があった。

【0015】

50

なお、Webアノテーション技術において、アノテーション付与とアノテーション復元との間には時間的なずれがある。このため、アノテーションを付与してから復元するまでの間にコンテンツの構造が変化することは、Webアノテーション技術において避けることができない。

【0016】

本発明は、コンテンツが変更されたとしても、アノテーションを復元するコンテンツ上の位置を特定できるアノテーション付与復元方法及びアノテーション付与復元装置を提供することを目的とする。

【課題を解決するための手段】

【0017】

上記課題を解決するため、本発明の一実施形態は、アノテーション付与復元方法であって、コンピュータに、構造を有するコンテンツ上のいずれかの位置についてアノテーションを付与する際に、前記コンテンツにおける前記位置の上位側の構造の記憶情報及び下位側の構造の記憶情報を前記アノテーションと関連づけて記憶する処理と、前記コンテンツについてアノテーションが付与された位置の上位側の構造の記憶情報に基づいて、アノテーション復元対象のコンテンツ中に、記憶された前記上位側の構造に合致する構造が含まれることを検出できなければ、アノテーションが付与された位置の下位側の構造の記憶情報に基づいて、アノテーション復元対象のコンテンツ中に、記憶された前記下位側の構造に合致する構造が含まれるかの検出を行い、記憶された前記下位側の構造に合致する構造が含まれることを検出すると、前記位置に対応づけてアノテーションを出力する処理とを実行させ、記憶される前記位置の下位側の構造は、前記位置の下位側の複数のパスの中から出現頻度に応じて選択された1つのパスに関する構造であることを特徴とする。

【0018】

なお、本発明の一実施形態の構成要素、表現または構成要素の任意の組合せを、方法、装置、システム、コンピュータプログラム、記録媒体、データ構造などに適用したのもも本発明の態様として有効である。

【発明の効果】

【0019】

上述の如く、本発明の一実施形態によれば、コンテンツが変更されたとしても、アノテーションを復元するコンテンツ上の位置を特定できる。

【図面の簡単な説明】

【0020】

【図1】付箋紙のような形式のアノテーションが表示されたWebページの一例の画面イメージ図である。

【図2】X Path取得時のコンテンツ及び木構造を表した説明図である。

【図3】内容が変更されたコンテンツ及び木構造を表した説明図である。

【図4】本実施例のシステムの一例の構成図である。

【図5】本実施例のシステムの他の例の構成図である。

【図6】利用者端末の一例の構成図である。

【図7】利用者端末の一例の処理ブロック図である。

【図8】アノテーションDBに格納されるアノテーション情報の一例の構成図である。

【図9】特徴X Pathについて説明する為の説明図である。

【図10】特徴情報について説明する為の説明図である。

【図11】X Pathにより対象ノードが特定できない例を表した説明図である。

【図12】特徴X Pathを用いたワイルドカード検索により複数の対象ノードが特定される例を表した説明図である。

【図13】アノテーション付与処理の一例のフローチャートである。

【図14】アノテーション復元処理の一例のフローチャートである。

【図15】本実施例のアノテーションプログラムを適用したブラウザの一例の構成図である。

10

20

30

40

50

【発明を実施するための形態】**【0021】**

次に、本発明を実施するための形態を、以下の実施例に基づき図面を参照しつつ説明していく。

【0022】

(システム構成)

図4は本実施例のシステムの一例の構成図である。図4に示すシステムは、利用者端末10、コンテンツサーバ20、アノテーション情報管理サーバ30、ネットワーク50を有する。利用者端末10、コンテンツサーバ20及びアノテーション情報管理サーバ30はネットワーク50を介してデータ通信可能に接続されている。ネットワーク50は、例えばインターネットやLANなど、データ通信可能な如何なる通信網であってもよい。

10

【0023】

利用者端末10はブラウザ11がインストールされている。ブラウザ11はプラグイン12により拡張機能が追加されている。コンテンツサーバ20はWebページ等のコンテンツ21を格納している。アノテーション情報管理サーバ30はアノテーションDB31を有し、アノテーション情報を管理している。

【0024】

図5は本実施例のシステムの他の例の構成図である。図5に示すシステムは、図4のシステムと同様、利用者端末10、コンテンツサーバ20、アノテーション情報管理サーバ30及びネットワーク50を有する。なお、図5のブラウザ11はプラグイン12により拡張機能が追加されていない。また、図5のシステムは、コンテンツサーバ20に、スクリプト22が一部を構成するコンテンツ21を格納している点で、図4のシステムと異なっている。

20

【0025】

図4のシステムはプラグイン12により本実施例に掛かるアノテーション機能を実現している。図5のシステムはスクリプト22により本実施例に掛かるアノテーション機能を実現している。

【0026】

(利用者端末10の構成)

図6は利用者端末の一例の構成図である。利用者端末10は、それぞれバスBで相互に接続された入力装置61、出力装置62、ドライブ装置63、補助記憶装置64、主記憶装置65、演算処理装置66及びインターフェース装置67を有する。

30

【0027】

入力装置61はキーボードやマウスなどである。入力装置61は各種信号を入力するために用いられる。出力装置62はディスプレイ装置などである。出力装置62は各種ウィンドウやデータ等を表示するために用いられる。インターフェース装置67はモデム、LANカードなどである。インターフェース装置67はネットワーク50に接続する為に用いられる。

【0028】

本実施例に掛かるアノテーションプログラムは利用者端末10を制御する各種プログラムの少なくとも一部である。アノテーションプログラムは、例えばアノテーションプログラムを記録した記録媒体68の配布やネットワーク50からのプログラムのダウンロードなどによって提供される。アノテーションプログラムを記録した記録媒体68は、CD-ROM、フレキシブルディスク、光磁気ディスク等の様に情報を光学的、電氣的或いは磁氣的に記録する記録媒体、ROM、フラッシュメモリ等の様に情報を電氣的に記録する半導体メモリ等、様々なタイプの記録媒体を用いることができる。

40

【0029】

アノテーションプログラムを記録した記録媒体68がドライブ装置63にセットされるとアノテーションプログラムは、記録媒体68からドライブ装置63を介して補助記憶装置64にインストールされる。ネットワーク50からダウンロードされたアノテーション

50

プログラムは、インターフェース装置 67 を介して補助記憶装置 64 にインストールされる。補助記憶装置 64 はインストールされたアノテーションプログラムを格納すると共に必要なファイル、データ等を格納する。

【0030】

主記憶装置 65 には、補助記憶装置 64 から読み出されたアノテーションプログラムが格納される。そして、演算処理装置 66 は主記憶装置 65 に格納されたアノテーションプログラムに従って、後述するような各種処理を実現している。

【0031】

(利用者端末 10 の処理ブロック)

図 7 は利用者端末の一例の処理ブロック図である。利用者端末 10 は、Web 資源取得部 71、Web 資源解析結果準備部 72、アノテーション情報格納部 73、アノテーション付与部 74、アノテーション構造位置取得部 75、アノテーション埋め込み部 76、アノテーション情報取得部 77、アノテーション復元部 78、アノテーション構造位置取得部 79、アノテーション埋め込み部 80 を有している。

10

【0032】

Web 資源取得部 71 は、Web 資源であるコンテンツの所在情報 (URL) を指定されたときに、指定された URL のコンテンツを、コンテンツサーバ 20 から取得する。Web 資源解析結果準備部 72 は、取り出されたコンテンツをそのまま Web 資源解析結果とすることも可能であるが、通常、コンテンツをパージングした結果である木構造を Web 資源解析結果として保持する。

20

【0033】

アノテーション付与部 74 は、利用者からコンテンツに基づく表示 (Web ページ) に対するアノテーション付与の操作を受け付ける。アノテーション構造位置取得部 75 はアノテーション付与の操作に基づき、利用者がアノテーションを付与した Web ページのコンテンツ上の位置を表す構造位置情報 (XPath) を取得する。アノテーション埋め込み部 76 はアノテーション構造位置取得部 75 が取得した構造位置情報に基づき、アノテーション内容を Web 資源解析結果の指定された位置に埋め込む。アノテーション付与部 74 は、Web 資源解析結果をレンダリングエンジンに渡し、描画させることで、アノテーションが付与された Web ページの表示を行うことができる。

【0034】

また、アノテーション情報格納部 73 は利用者に付与されたアノテーションを復元する為のアノテーション情報をアノテーション DB 31 に格納させる。アノテーション情報には、アノテーションを付与した Web ページのコンテンツの URL、アノテーションの構造位置情報である XPath、後述の特徴 XPath、後述の特徴情報、アノテーション内容が含まれる。

30

【0035】

アノテーション情報取得部 77 は取得したコンテンツに対応するアノテーション情報がアノテーション DB 31 に格納されていれば、アノテーション DB 31 からコンテンツに対応するアノテーション情報を取得する。

【0036】

そして、アノテーション構造位置取得部 79 はアノテーション情報に指示されているアノテーションの構造位置を探し出す。アノテーション埋め込み部 80 はアノテーション構造位置取得部 79 が探し出した構造位置に基づき、アノテーション内容を Web 資源解析結果の指定された位置に埋め込む。アノテーション復元部 78 は、Web 資源解析結果をレンダリングエンジンに渡すことで、アノテーションが復元された Web ページを表示させることができる。

40

【0037】

(アノテーション DB の構成)

図 8 はアノテーション DB に格納されるアノテーション情報の一例の構成図である。図 8 のアノテーション DB 31 は、URL、XPath、特徴 XPath、特徴情報、アノ

50

テーション内容が対応付けられたレコードを有している。アノテーションDB31に含まれるURLはアノテーションを付与したWebページのコンテンツのURLである。XPathはアノテーションが付与されたWebページのコンテンツ上の構造位置である。

【0038】

特徴XPath及び特徴情報は、XPathによりアノテーションが付与されたWebページのコンテンツ上の構造位置が特定できなかったときに、アノテーションが付与されたWebページのコンテンツ上の構造位置を検索する為に利用するものである。特徴XPath及び特徴情報の詳細は後述する。アノテーション内容はアノテーションに記述された情報である。

【0039】

10

(特徴XPath及び特徴情報)

図9は特徴XPathについて説明する為の説明図である。図9はアノテーションが付与されたWebページのコンテンツ上の構造位置がXPath「/HTML/BODY/TABLE」である例を表している。アノテーション構造位置取得部75はXPath「/HTML/BODY/TABLE」により表される対象ノードを頂点とし、末端までの全てのXPathを抽出する。

【0040】

図9の例では、XPath「/TABLE/TR/TD」が2回、XPath「/TABLE/TR/TD/P」が1回、抽出される。アノテーション構造位置取得部75は抽出されたXPathの内、最も出現回数の少ないXPathを特徴XPathとする。図9の例では、XPath「/TABLE/TR/TD/P」が特徴XPathとなる。

20

【0041】

図9に示すように、アノテーション構造位置取得部75は対象ノードに含まれるXPathの内、最も出現回数の少ないXPathを特徴XPathとして抽出する。最も出現回数の少ないXPathを特徴XPathとして抽出する理由は、対象ノードの特徴を最も表していると考えられるからである。

【0042】

なお、図9では対象ノードに含まれるXPathの内、最も出現回数の少ないXPathを特徴XPathとして抽出したが、最も出現回数の多いXPathや平均的な出現回数のXPathなど、様々なバリエーションも考えられる。

30

【0043】

XPathとして、出現回数が少ないものを用いる場合は、再現位置を特定しやすくなるが、コンテンツの内容が変更された時に、特徴パスに関連する部分に変更される可能性が高くなり、特徴パスを発見できない可能性が上がる。XPathとして、出現回数が多いものを用いる場合は、逆にコンテンツの内容が変更された場合でも特徴パスを発見することが可能であるが、他にも頻繁に使われるパスである可能性が高く、正しい再現位置を特定しにくくなる。平均値を使う場合は、上記特徴の中間となる。

【0044】

例えば、アノテーションシステムを使い画面構成を複数人で検討する場合など、正しい再現位置に確実に戻すことが必要な場面では、少ないものを特徴パスとすることで、精度の高い情報提供ができる。また、足跡システムのように、再現位置のズレよりも情報があることを通知することが重要なシステムでは、多いものを選択すると良い。

40

【0045】

図10は特徴情報について説明する為の説明図である。図10はアノテーションが付与されたWebページのコンテンツ上の構造位置がXPath「/HTML/BODY/TABLE」である例を表している。アノテーション構造位置取得部75は対象ノードに含まれるタグの種類と数とを特徴情報として抽出する。

【0046】

図10の例では、タグ<TABLE>が1回、タグ<TR>が1回、タグ<TD>が3回、タグ<P>が1回、抽出される。アノテーション構造位置取得部75は抽出されたタ

50

グの種類と数とを特徴情報とする。なお、アノテーション構造位置取得部75と同様な手順にてアノテーション構造位置取得部79も特徴情報を抽出する。

【0047】

図11はX Pathにより対象ノードが特定できない例を表した説明図である。ここで言う対象ノードとは、アノテーション情報に含まれているX Pathで表されるノードである。

【0048】

図11(A)は、アノテーション内容を付与したときのコンテンツを表している。図11(A)は図9に示すコンテンツと同様である。したがって、図11(A)に示すコンテンツはアノテーション情報に含まれるX Pathが「/HTML/BODY/TABLE」となる。図11(A)に示すコンテンツはアノテーション情報に含まれる特徴X Pathが「/TABLE/TR/TD/P」となる。また、図11(A)に示すコンテンツはアノテーション情報に含まれる特徴情報が「タグ<TABLE>1回、タグ<TR>1回、タグ<TD>3回、タグ<P>1回」となる。

10

【0049】

図11(B)は内容が変更されたコンテンツを表している。図11(B)のコンテンツは図11(A)に示す変更前のコンテンツから取得したX Path「/HTML/BODY/TABLE」が存在していない。したがって、アノテーション構造位置取得部79は図11(A)に示す変更前のコンテンツから取得したX Pathで検索したとしても、図11(B)のコンテンツに含まれる対象ノードを特定できない。

20

【0050】

本実施例のシステムでは、アノテーション情報に含まれる特徴X Path「/TABLE/TR/TD/P」でワイルドカード検索「//特徴X Path」を行うことにより対象ノードを特定できる。

【0051】

なお、本実施例のシステムでは特徴X Path「/TABLE/TR/TD/P」でワイルドカード検索「//特徴X Path」を行うことにより対象ノードを特定しようとした場合、複数の対象ノードが特定されることがある。特徴X Pathを用いたワイルドカード検索「//特徴X Path」により複数の対象ノードが特定された場合は、図12に示すように特徴情報を用いて1つの対象ノードを特定する。

30

【0052】

図12は特徴X Pathを用いたワイルドカード検索により複数の対象ノードが特定される例を表した説明図である。図12(A)は図11(A)と同様であるため、説明を省略する。図12(B)は内容が変更されたコンテンツを表している。図12(B)のコンテンツは、変更前のコンテンツから取得したX Path「/HTML/BODY/TABLE」が存在していない。したがって、アノテーション構造位置取得部79は変更前のコンテンツから取得したX Pathで検索したとしても、図12(B)のコンテンツに含まれる対象ノードを特定できない。

【0053】

そこで、アノテーション構造位置取得部79は変更前のコンテンツから抽出した特徴X Path「/TABLE/TR/TD/P」でワイルドカード検索「//特徴X Path」を行うが、2つの対象ノード100、101を特定してしまう。

40

【0054】

そこで、アノテーション構造位置取得部79は特定した対象ノード100、101に含まれるタグの種類と数とを特徴情報として抽出する。図12(B)の場合、対象ノード100は「タグ<TABLE>1回、タグ<TR>1回、タグ<TD>3回、タグ<P>1回」となる。また、対象ノード101は「タグ<TABLE>1回、タグ<TR>1回、タグ<TD>3回、タグ<P>3回」となる。

【0055】

アノテーション構造位置取得部79は対象ノード100、101から抽出した特徴情報

50

とアノテーション情報に含まれる特徴情報とを比較し、アノテーション情報に含まれる特徴情報との一致度が高い、言い換えれば一致する特徴情報の含有率が最も高い対象ノードを対象ノードとして特定する。

【 0 0 5 6 】

図 1 2 (B) の場合、アノテーション構造位置取得部 7 9 は対象ノード 1 0 0、1 0 1 から抽出した特徴情報とアノテーション情報に含まれる特徴情報とを比較し、対象ノード 1 0 0 を特定できる。

【 0 0 5 7 】

(利用者端末 1 0 の処理手順)

図 1 3 はアノテーション付与処理の一例のフローチャートである。アノテーション付与部 7 4 は利用者から Web ページに対するアノテーション付与要求があると、ステップ S 1 に進み、ブラウザ 1 1 に対するマウス操作で、利用者にアノテーションの付与位置を選択させる。ステップ S 2 に進み、アノテーション付与部 7 4 は例えばブラウザ 1 1 に対するキーボード操作で、利用者にアノテーション内容であるコメントを入力させる。

【 0 0 5 8 】

ステップ S 3 に進み、アノテーション構造位置取得部 7 5 はアノテーションの付与位置 (対象ノード) の X P a t h を取得する。ステップ S 4 に進み、アノテーション構造位置取得部 7 5 は対象ノードを頂点とし、末端までの全ての X P a t h を抽出する。

【 0 0 5 9 】

ステップ S 5 に進み、アノテーション構造位置取得部 7 5 はステップ S 4 で抽出した X P a t h の内、最も出現回数の少ない X P a t h を特徴 X P a t h として抽出する。アノテーション構造位置取得部 7 5 はステップ S 6 に進み、アノテーションの付与位置に含まれるタグの種類と数 (出現回数) とを特徴情報として抽出する。

【 0 0 6 0 】

そして、ステップ S 7 に進み、アノテーション情報格納部 7 3 は X P a t h、特徴 X P a t h、特徴情報及びアノテーション内容をアノテーション情報としてアノテーション DB 3 1 に格納する。ステップ S 8 に進み、アノテーション埋め込み部 7 6 はステップ S 3 でアノテーション構造位置取得部 7 5 が取得した X P a t h に基づき、アノテーション内容をコンテンツの Web 資源解析結果の指定された位置に埋め込む。

【 0 0 6 1 】

その後、アノテーション付与部 7 4 は、アノテーション内容が埋め込まれた Web 資源解析結果をレンダリングエンジンに渡すことで、アノテーションが付与されたコンテンツの Web ページを表示させることができる。

【 0 0 6 2 】

図 1 4 はアノテーション復元処理の一例のフローチャートである。Web 資源取得部 7 1 は例えばブラウザ 1 1 に対するキーボード操作で、使用者から Web ページの URL の指定があると、ステップ S 1 1 に進み、Web ページの URL をブラウザ 1 1 から取得する。

【 0 0 6 3 】

Web 資源取得部 7 1 は取得した URL の Web ページに対応したコンテンツをコンテンツサーバ 2 0 から取得する。Web 資源解析結果準備部 7 2 は、取得したコンテンツをパージングした結果である木構造を Web 資源解析結果として保持する。

【 0 0 6 4 】

ステップ S 1 2 に進み、アノテーション情報取得部 7 7 はステップ S 1 1 で取得した URL に対応するアノテーション情報をアノテーション DB 3 1 から取得する。ステップ S 1 3 に進み、アノテーション構造位置取得部 7 9 はアノテーション情報に含まれる X P a t h による復元位置 (対象ノード) の検索を行う。

【 0 0 6 5 】

ステップ S 1 4 に進み、アノテーション構造位置取得部 7 9 はステップ S 1 3 の検索により復元位置を特定できればステップ S 2 0 に進み、特定した復元位置に基づき、Web

10

20

30

40

50

資源解析結果の指定された位置にアノテーション内容を埋め込む。

【0066】

ステップS14において、アノテーション構造位置取得部79はステップS13の検索により復元位置を特定できなければステップS15に進み、アノテーション情報に含まれる特徴X P a t hによる復元位置のワイルドカード検索を行う。

【0067】

ステップS16に進み、アノテーション構造位置取得部79はステップS15のワイルドカード検索により復元位置を特定できればステップS20に進み、特定した復元位置に基づき、W e b資源解析結果の指定された位置にアノテーション内容を埋め込む。

【0068】

ステップS16において、アノテーション構造位置取得部79はステップS15のワイルドカード検索により複数の復元位置を特定してしまい、1つの復元位置を特定できなければステップS17に進み、特定した複数の復元位置のそれぞれについて、復元位置に含まれるタグ毎の出現回数を特徴情報として抽出する。アノテーション構造位置取得部79はステップS17で抽出した特徴情報とアノテーション情報に含まれる特定情報とを比較して、ステップS17で抽出した特徴情報の内、アノテーション情報に含まれる特定情報と一致する特定情報を特定できれば、ステップS20に進む。

【0069】

アノテーション構造位置取得部79は特定した復元位置に基づき、W e b資源解析結果の指定された位置にアノテーション内容を埋め込む。また、アノテーション構造位置取得部79はステップS17で抽出した特徴情報の内、アノテーション情報に含まれる特定情報と一致する特定情報を特定できなければ、ステップS19に進み、アノテーション情報に含まれる特徴情報との一致度が高い、言い換えれば一致する特徴情報の含有率が最も高い特定情報を特定する。

【0070】

ステップS19に続いてステップS20に進み、アノテーション構造位置取得部79は特定した復元位置に基づき、W e b資源解析結果の指定された位置にアノテーション内容を埋め込む。

【0071】

その後、アノテーション復元部78は、アノテーション内容が埋め込まれたW e b資源解析結果をレンダリングエンジンに渡すことで、アノテーションが付与されたW e bページを表示させることができる。

【0072】

(アノテーションプログラムがブラウザ11に適用された例)

図15は、本実施例のアノテーションプログラムを適用したブラウザの一例の構成図である。図15のブラウザ200は、典型的な構成を示したものであり、説明に不要なC S Sの処理やP D Fプラグイン等の細かい部分を省略している。

【0073】

ネットワークを経由して受信されたH T M L文書150をH T M Lパーサ201は解析して、木構造(D O Mツリー)202を生成する。これと同時に、J a v a s c r i p t (登録商標)エンジン206はH T M L文書150中にあったJ a v a s c r i p tプログラム205を実行し、D O Mツリー202の中身を書き換える。書き換えられたD O Mツリー202はレンダリングエンジン203に送られる。

【0074】

レンダリングエンジン203は送られてきたD O Mツリー202を画面で見える形に変換する。そして、ユーザインタフェース204を介したユーザの操作等によってJ a v a s c r i p tエンジン206はD O Mツリー202を書き換え、表示させる内容を変化させる。

【0075】

アノテーションプログラム210は、典型的にブラウザ200の中に保持される(すな

10

20

30

40

50

わち、ブラウザ200に組込まれる、あるいは機能拡張として付け加えられる形で実現される)。

【0076】

そして、読み込まれたHTML文書150のJavaScriptプログラム205が実行されるのに追加して、アノテーションプログラム210は実行され、DOMツリー202を書き換える。なお、アノテーションプログラム210が参照するアノテーションDB等のデータはブラウザ200の外部に保持される。

【0077】

アノテーションプログラム210を保持する別の実施形態としては、読み込まれるHTML文書に、アノテーションプログラム210を保持しておくことが考えられる。この実施形態では、適用できるWebページが限定されるが、ブラウザ200にアノテーションプログラム210を保持するのと同じ効果を持たせることができる。

【0078】

本発明は、具体的に開示された実施例に限定されるものではなく、特許請求の範囲から逸脱することなく、種々の変形や変更が可能である。

【符号の説明】

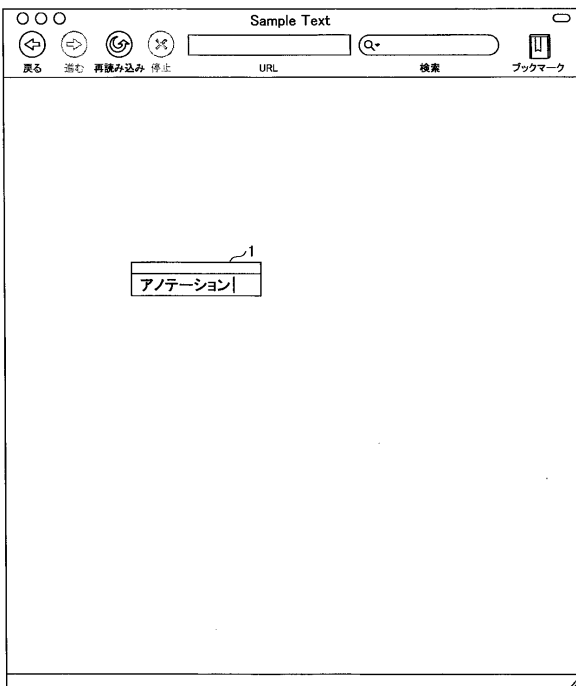
【0079】

1	アノテーション	
10	利用者端末	
11	ブラウザ	20
12	プラグイン	
20	コンテンツサーバ	
21	コンテンツ	
22	スクリプト	
30	アノテーション情報管理サーバ	
31	アノテーションDB	
50	ネットワーク	
61	入力装置	
62	出力装置	
63	ドライブ装置	30
64	補助記憶装置	
65	主記憶装置	
66	演算処理装置	
67	インターフェース装置	
68	記録媒体	
71	Web資源取得部	
72	Web資源解析結果準備部	
73	アノテーション情報格納部	
74	アノテーション付与部	
75	アノテーション構造位置取得部	40
76	アノテーション埋め込み部	
77	アノテーション情報取得部	
78	アノテーション復元部	
79	アノテーション構造位置取得部	
80	アノテーション埋め込み部	
100、101	対象ノード	
150	HTML文書	
200	ブラウザ	
201	HTMLパーサ	
202	DOMツリー	50

- 203 レンダリングエンジン
- 204 ユーザインタフェース
- 205 Javascriptプログラム
- 206 Javascriptエンジン
- 210 アノテーションプログラム

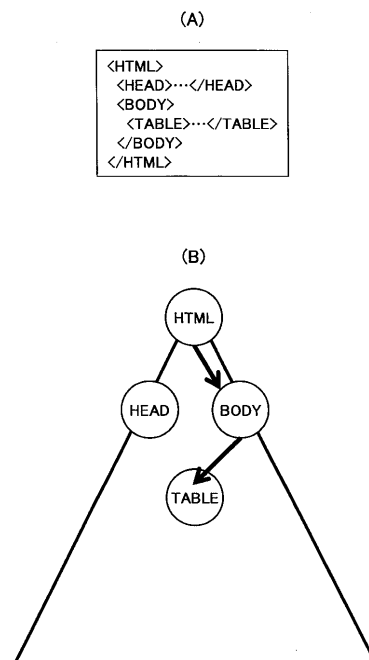
【図1】

付箋紙のような形式のアノテーションが
表示されたWebページの一例の画面イメージ図



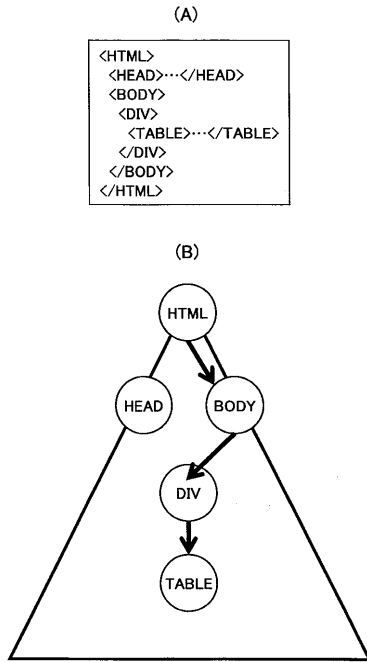
【図2】

XPath取得時のコンテンツ及び木構造を表した説明図



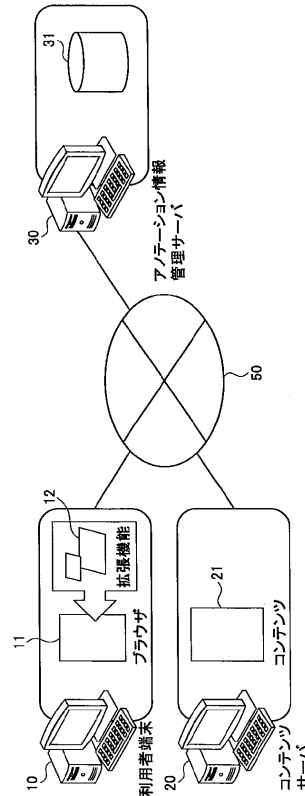
【図3】

内容が変更されたコンテンツ及び木構造を表した説明図



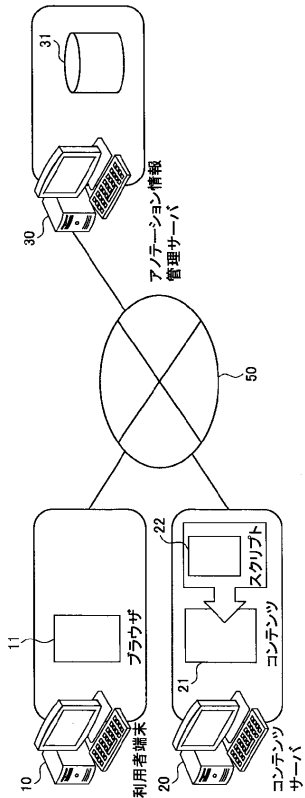
【図4】

本実施例のシステムの一例の構成図



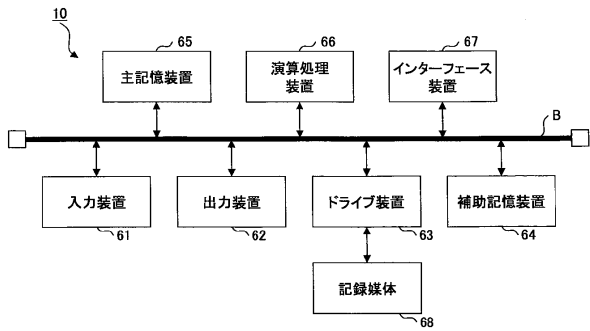
【図5】

本実施例のシステムの他の例の構成図

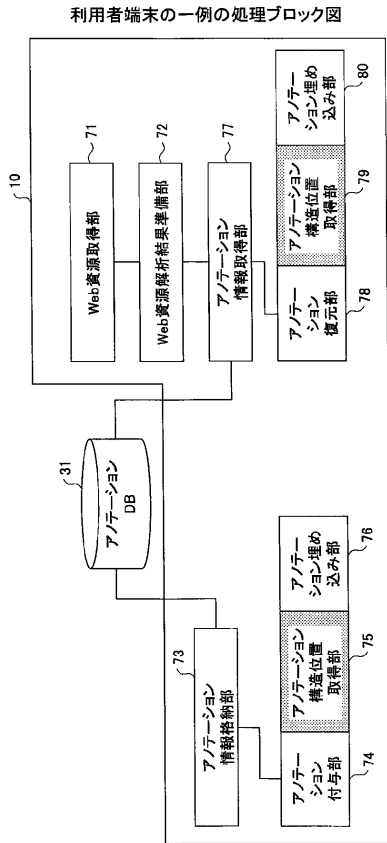


【図6】

利用者端末の一例の構成図



【 図 7 】



【 図 8 】

AnnotationDBに格納されるAnnotation情報の一例の構成図

URL	XPath	特徴XPath	特徴情報	Annotation内容
http://xxx/x.html	/HTML/BODY/DIV[2]	/DIV/FONT	DIV:1, FONT:1	X
http://yyy/y.html	/HTML/BODY/TABLE	/TABLE/TR/TD/P	TABLE:1, TR:1, TD:3, P:1	Y

【 図 9 】

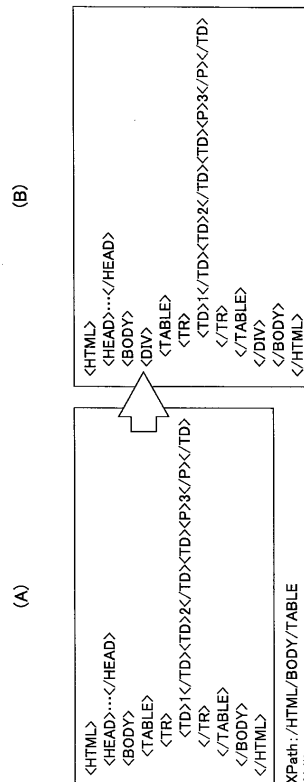
特徴XPathについて説明する為の説明図

```

<HTML>
<HEAD>...</HEAD>
<BODY>
<TABLE>
<TR>
<TD>1</TD><TD>2</TD><TD><P>3</P></TD>
</TR>
</TABLE>
</BODY>
</HTML>
  
```

【 図 1 1 】

XPathにより対象ノードが特定できない例を表した説明図



【 図 1 0 】

特徴情報について説明する為の説明図

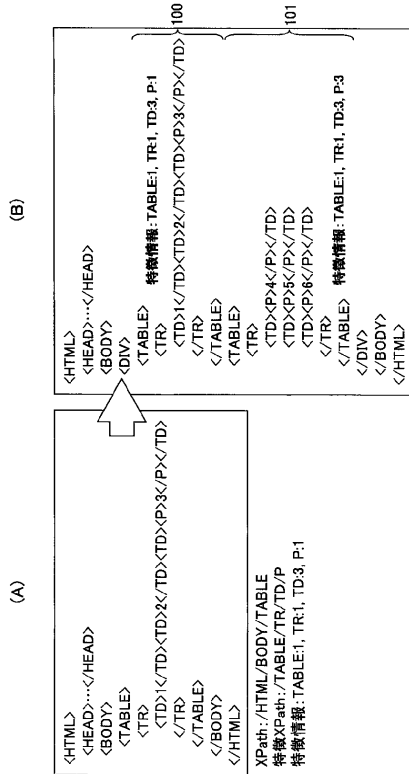
```

<HTML>
<HEAD>...</HEAD>
<BODY>
<TABLE>
<TR>
<TD>1</TD><TD>2</TD><TD><P>3</P></TD>
</TR>
</TABLE>
</BODY>
</HTML>
  
```

XPath: /HTML/BODY/TABLE
 特徴XPath: /TABLE/TR/TD/P
 特徴情報: TABLE:1, TR:1, TD:3, P:1

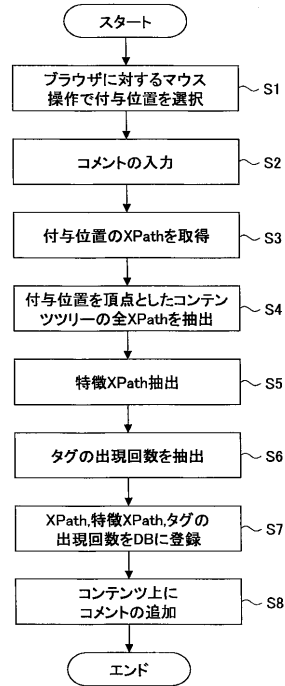
【図12】

特徴XPathを用いたワイルドカード検索により複数の対象ノードが特定される例を表した説明図



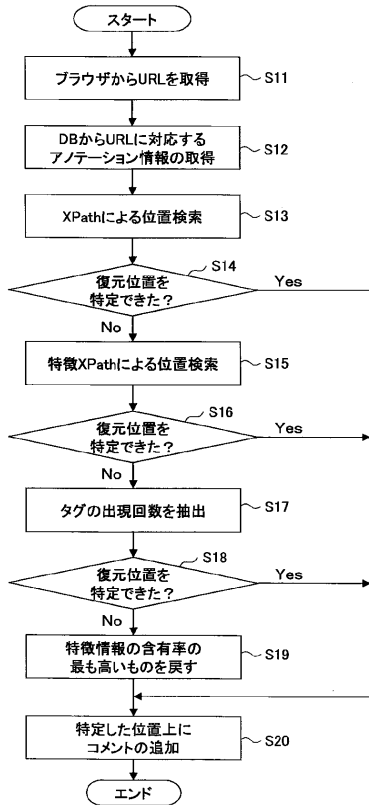
【図13】

アノテーション付与処理の一例のフローチャート



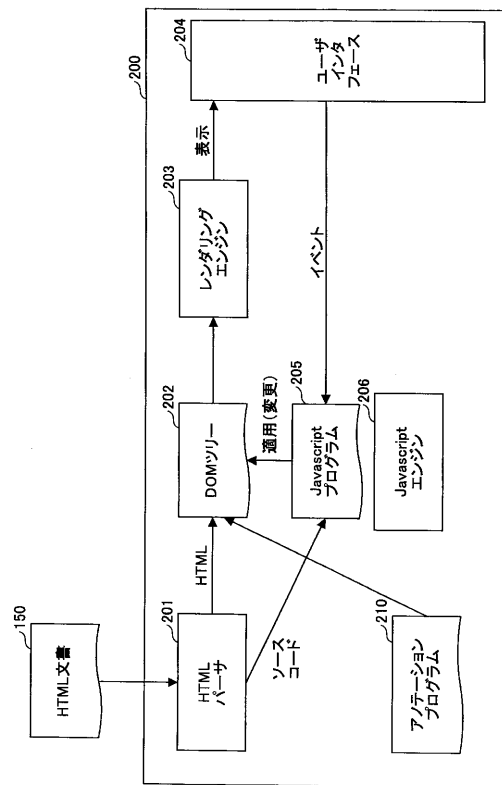
【図14】

アノテーション復元処理の一例のフローチャート



【図15】

本実施例のアノテーションプログラムを適用したブラウザの一例の構成図



フロントページの続き

(72)発明者 西野 文人

神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

審査官 成瀬 博之

(56)参考文献 特開2011-209886(JP,A)
特開2009-140020(JP,A)
特開2007-140836(JP,A)
特開2004-303097(JP,A)
特開2004-046745(JP,A)
特開2010-044708(JP,A)
特開平11-110384(JP,A)
国際公開第2007/119567(WO,A1)

(58)調査した分野(Int.Cl., DB名)

G06F 17/20 - 17/26