

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7212771号
(P7212771)

(45)発行日 令和5年1月25日(2023.1.25)

(24)登録日 令和5年1月17日(2023.1.17)

(51)国際特許分類 F I
G 0 6 F 16/957(2019.01) G 0 6 F 16/957

請求項の数 11 (全23頁)

(21)出願番号	特願2021-516984(P2021-516984)	(73)特許権者	512015127
(86)(22)出願日	令和2年7月14日(2020.7.14)		バイドゥ オンライン ネットワーク テクノロジー(ペキン)カンパニー リミテッド
(65)公表番号	特表2022-512056(P2022-512056 A)		中華人民共和国 ペキン ハイディエン ディストリクト シャンディー テンス ストリート ナンバー 10 バイドゥ キャンパス 3エフ
(43)公表日	令和4年2月2日(2022.2.2)	(74)代理人	100209048 弁理士 森川 元嗣
(86)国際出願番号	PCT/CN2020/101910	(72)発明者	レイ、シャオチエン
(87)国際公開番号	WO2021/098242		中華人民共和国、ベイジン、ハイディエン・ディストリクト、シャンディ・テンス・ストリート、ナンバー 10、バイドゥ・キャンパス 3 /フロア
(87)国際公開日	令和3年5月27日(2021.5.27)		最終頁に続く
審査請求日	令和3年3月25日(2021.3.25)		
(31)優先権主張番号	202010125624.1		
(32)優先日	令和2年2月27日(2020.2.27)		
(33)優先権主張国・地域又は機関	中国(CN)		

(54)【発明の名称】 ページ処理方法、デバイス、電子デバイス及びコンピュータ読み取り可能な記憶媒体

(57)【特許請求の範囲】

【請求項1】

ページ処理デバイスによって、取得したハイパーテキストマークアップ言語HTMLファイルに基づいて、ページの複数のレイアウトオブジェクトノードを確定するステップと、
前記ページ処理デバイスによって、前記ページの複数のレイアウトオブジェクトノードに対してレイアウト過程を実行した後、所定のリコール規則を用いて、前記複数のレイアウトオブジェクトノードをスクリーニングして、前記リコール規則に適合するレイアウトオブジェクトノードを取得するステップと、
前記ページ処理デバイスによって、前記リコール規則に適合するレイアウトオブジェクトノードが指定ターゲットノードであるか否かを予測するステップと、
前記ページ処理デバイスによって、前記指定ターゲットノードをマスク処理し、マスク処理後に残ったレイアウトオブジェクトノードを用いて、前記マスク処理されたページを生成するステップと、を含み、
前記ページ処理デバイスによって、前記リコール規則に適合する前記レイアウトオブジェクトノードが前記指定ターゲットノードであるか否かを予測するステップは、
前記ページ処理デバイスによって、前記リコール規則に適合する前記レイアウトオブジェクトノードの属性情報に基づいて、前記リコール規則に適合する前記レイアウトオブジェクトノードのノード特徴を計算するステップと、
前記ページ処理デバイスによって、所定のノード予測モデルを用いて前記ノード特徴を処理し、前記リコール規則に適合する前記レイアウトオブジェクトノードが前記指定ター

10

20

ゲットノードである確率値を取得するステップと、

前記ページ処理デバイスによって、前記確率値に基づいて前記リコール規則に適合する前記レイアウトオブジェクトノードが前記指定ターゲットノードであるか否かを確定するステップと、を含む、

ページ処理方法。

【請求項2】

前記ページ処理デバイスによって、ページの複数のレイアウトオブジェクトノードを確定する前記ステップの後に、

前記ページ処理デバイスによって、前記複数のレイアウトオブジェクトノードの中にスクリプトファイルでロードされたレイアウトオブジェクトノードが含まれると確定されたことに応答して、前記スクリプトファイルでロードされたレイアウトオブジェクトノードを前記リコール規則に適合するレイアウトオブジェクトノードとするステップをさらに含む、
請求項1に記載のページ処理方法。

10

【請求項3】

前記ページ処理デバイスによって、前記ページの複数のレイアウトオブジェクトノードに対してレイアウト過程を実行した後、所定のリコール規則を用いて前記レイアウトオブジェクトノードをスクリーニングして、前記複数のレイアウトオブジェクトノードのうちの前記リコール規則に適合するレイアウトオブジェクトノードを取得する前記ステップは、前記ページ処理デバイスによって、前記ページのレイアウトオブジェクトノードをレイアウトし、前記レイアウトされたレイアウトオブジェクトノードの属性情報を取得するステップと、

20

前記ページ処理デバイスによって、前記属性情報が前記リコール規則に限定されたノードリコール条件を満たすかどうかを判断するステップと、

前記ページ処理デバイスによって、前記ノードリコール条件を満たすレイアウトオブジェクトノードを、前記リコール規則に適合するレイアウトオブジェクトノードとするステップとを含む、

請求項1に記載のページ処理方法。

【請求項4】

前記リコール規則は、ノードのアスペクト比、ノードの埋め込み形式、ノードの位置的特徴、ノードの内容、ノードの生成メカニズム、ノードの構造の少なくとも1つに応じて設けられる規則を含む、

30

請求項3に記載のページ処理方法。

【請求項5】

前記ページ処理デバイスによって、前記リコール規則に適合する前記レイアウトオブジェクトノードが指定ターゲットノードであるかどうかを予測する前記ステップの前に、

前記ページ処理デバイスによって、前記リコール規則に適合するレイアウトオブジェクトノードを、初回スクリーニングで得られたレイアウトオブジェクトノードとし、前記初回スクリーニングで得られたレイアウトオブジェクトノードのノード状態を確定するステップと、

前記ページ処理デバイスによって、前記ページの全てのレイアウトオブジェクトノードでレイアウトが完了した後、ノード状態が変化するレイアウトオブジェクトノードを取得するステップと、

40

前記ページ処理デバイスによって、所定のリコール規則を利用して、ノード状態が変化するレイアウトオブジェクトノードを再度スクリーニングするステップと、

前記ページ処理デバイスによって、前記初回スクリーニングで得られたレイアウトオブジェクトノードと前記再度スクリーニングで得られたレイアウトオブジェクトノードとを、前記リコール規則に適合するレイアウトオブジェクトノードとするステップをさらに含む、
請求項1から4のいずれか1項に記載のページ処理方法。

【請求項6】

前記リコール規則に適合する前記レイアウトオブジェクトノードは、前記ページのレイ

50

アウトオブジェクトツリーにおけるノードであり、
前記ページ処理デバイスによって、前記リコール規則に適合する前記レイアウトオブジェクトノードの属性情報に基づき、前記リコール規則に適合する前記レイアウトオブジェクトノードのノード特徴を計算するステップには、
前記ページ処理デバイスによって、前記リコール規則に適合する前記レイアウトオブジェクトノードの属性情報を取得し、前記属性情報は前記レイアウト過程で取得された情報であり、
前記ページ処理デバイスによって、深さを優先的にトラバースする方式を用い、前記属性情報を利用して、前記レイアウトオブジェクトツリーにおける前記リコール規則に適合するレイアウトオブジェクトノードに対し、トップダウン特徴計算を行い、前記リコール規則に適合するレイアウトオブジェクトノードのノード特徴を得ることが含まれる、
請求項1に記載のページ処理方法。

【請求項7】

前記ノード予測モデルは、マーキングされたオフラインレンダリング済みの静的ページデータを用いて予めトレーニングされたモデルであり、指定された深さ及び指定された数の確定ツリーを有する勾配強調確定ツリーモデルである、
請求項1から6のいずれか1項に記載のページ処理方法。

【請求項8】

前記ページ処理デバイスによって、前記指定ターゲットノードをマスク処理する前記ステップには、
前記ページ処理デバイスによって、前記指定ターゲットノードの属性情報に基づいて、前記ページ内の位置、幅、高さ、主題のコンテンツ内であるか否か、及び前記ページ内の面積占有率のうち少なくとも1つを含む対応するノード特性情報を計算することと、
前記ページ処理デバイスによって、前記ノード特性情報が所定のマスク閾値に達したと確定されたことに応答して、前記指定ターゲットノードの状態を非表示に設定してマスク処理を行うことが含まれる、
請求項1から7のいずれか1項に記載のページ処理方法。

【請求項9】

取得されたハイパーテキストマークアップ言語HTMLファイルに基づき、ページの複数のレイアウトオブジェクトノードを確定するノード確定モジュールと、
前記ページの複数のレイアウトオブジェクトノードをレイアウトした後、所定のリコール規則を用いて、前記複数のレイアウトオブジェクトノードをスクリーニングして、前記リコール規則に適合するレイアウトオブジェクトノードを取得するノードスクリーニングモジュールと、
前記リコール規則に適合するレイアウトオブジェクトノードが指定ターゲットノードであるか否かを予測する予測モジュールと、
前記指定ターゲットノードをマスク処理し、マスク処理後に残ったレイアウトオブジェクトノードを用いて前記マスク処理後のページを生成するマスク処理モジュールと、を含み、

前記予測モジュールは、

前記リコール規則に適合する前記レイアウトオブジェクトノードの属性情報に基づいて、前記リコール規則に適合する前記レイアウトオブジェクトノードのノード特徴を計算する特徴計算ユニットと、
所定のノード予測モデルを用いて前記ノード特徴を処理し、前記リコール規則に適合する前記レイアウトオブジェクトノードが前記指定ターゲットノードである確率値を取得する確率計算ユニットと、
前記確率値に基づいて前記リコール規則に適合する前記レイアウトオブジェクトノードが前記指定ターゲットノードであるか否かを確定するターゲットノード確定ユニットと、を含む、

ページ処理デバイス。

10

20

30

40

50

【請求項 10】

1つ又は複数のプロセッサと、

1つ又は複数のプログラムが記憶され、前記1つ又は複数のプログラムが前記1つ又は複数のプロセッサで実行されると、前記1つ又は複数のプロセッサが請求項1～8のいずれか1項に記載のページ処理方法を実現するメモリと、

前記プロセッサとメモリとの間に接続され、前記プロセッサとメモリとの情報のやりとりを実現するように構成された1つ又は複数のI/Oインターフェースとを備える、電子デバイス。

【請求項 11】

プロセッサによって実行されるときに、前記プロセッサが請求項1～8のいずれか1項に記載のページ処理方法を実行するコンピュータプログラムが記憶された

10

コンピュータ読み取り可能な記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本願実施例は、深層学習、スマート検索技術分野に関し、特にページ処理方法、デバイス、電子デバイス及びコンピュータ読み取り可能な記憶媒体に関するものである。

【背景技術】

【0002】

モバイルインターネットの全体的な普及に伴い、益々多くのサイトがモバイルシーンで広告マーケティング及びアプリケーションや普及を進めている。移動デバイスの画面によって制限されることから、広告等の要素がユーザのブラウジング体験に与える影響は益々顕著になっている一方で、短期的な利益を最大化するために、一部のサイトは、サイトに多数の虚偽、ポルノ及びユーザを欺く広告要素を掲載しており、このことは、ユーザのブラウジング体験に深刻な影響を与え、モバイルエコセキュリティを損なっている。

20

【0003】

よって、モバイル検索のエコセキュリティを確保するためにウェブサイトに表示されるページコンテンツをフィルタリングして、ユーザのブラウジング体験を向上させるべきである。

【発明の概要】

30

【0004】

本願実施例は、ページ処理方法、デバイス、電子デバイスとコンピュータ読み取り可能な記憶媒体を提供する。

【0005】

第1の態様において、本願実施例は、取得したハイパーテキストマークアップ言語HTMLファイルに基づいて、ページの複数のレイアウトオブジェクトノードを確定するステップと、ページの複数のレイアウトオブジェクトノードをレイアウトした後、所定のリコール規則を用いて、前記複数のレイアウトオブジェクトノードをスクリーニングして、リコール規則に適合するレイアウトオブジェクトノードを取得するステップと、リコール規則に適合するレイアウトオブジェクトノードが指定ターゲットノードであるか否かを予測するステップと、指定ターゲットノードをマスク処理し、マスク処理後に残ったレイアウトオブジェクトノードを用いて、マスク処理されたページを生成するステップと、を含むページ処理方法を提供する。

40

【0006】

第2の態様において、本願実施例は、取得されたHTMLファイルに基づき、ページの複数のレイアウトオブジェクトノードを確定するノード確定モジュールと、ページの複数のレイアウトオブジェクトノードをレイアウトした後、所定のリコール規則を用いて、前記複数のレイアウトオブジェクトノードをスクリーニングして、リコール規則に適合するレイアウトオブジェクトノードを取得するノードスクリーニングモジュールと、リコール規則に適合するレイアウトオブジェクトノードが指定ターゲットノードであるか否かを予測

50

する予測モジュールと、指定ターゲットノードをマスク処理し、マスク処理後に残ったレイアウトオブジェクトノードを用いてマスク処理後のページを生成するマスク処理モジュールと、を含むページ処理デバイスを提供する。

【0007】

第3の態様において、本願実施例は、1つ又は複数のプロセッサと、1つ又は複数のプログラムが記憶され、1つ又は複数のプログラムが前記1つ又は複数のプロセッサで実行されると、1つ又は複数のプロセッサが上述したページ処理方法のいずれか1つを実現するメモリと、プロセッサとメモリとの間に接続され、プロセッサとメモリとの情報のやりとりを実現するように構成された1つ又は複数のI/Oインターフェースとを備える、電子デバイスを提供する。

10

【0008】

第4の態様において、本願実施例は、プロセッサによって実行されるときに、上述したページ処理方法のいずれか1つを実行するコンピュータプログラムが記憶されたコンピュータ読み取り可能な記憶媒体を提供する。

【0009】

本願実施例が提供するページ処理方法、デバイス、電子デバイス及びコンピュータ読み取り可能な記憶媒体は、リコール規則とノード予測モデルとの結合方式を採用してページを処理し、リコール規則によってスクリーニングされたレイアウトオブジェクトノードに対して、ノード予測モデルを再利用してブラウジング体験に影響するか否かを確定することにより、ブラウジング体験に影響すると予測されたレイアウトオブジェクトノードに対してマスク処理を行い、全体としてページのブラウジング体験を最適化し、モバイル検索のエコセキュリティに対する保障を提供する。

20

【0010】

図面は、本願実施例に対するさらなる理解を提供するために用いられ、明細書の一部を構成し、本願実施例とともに本願を説明するために用いられるものであり、本願を限定するものではない。上記特徴及び利点並びに他の特徴及び利点は、図面を参照して詳細な例示的な実施例を説明することにより、当業者にとってさらに明らかとなるであろう。図面において、

【図面の簡単な説明】

【0011】

- 【図1】図1は、本願実施例によるページ処理のアーキテクチャの概略図であり、
【図2】図2は、本願の一実施例におけるページ処理方法のフローチャートであり、
【図3】図3は、本願の例示的な実施例におけるリコール規則の概略図であり、
【図4】図4は、本願の別の実施例によるページ処理方法のフローチャートであり、
【図5(a)】図5(a)は、本願によるページ処理方法の効果を示す図であり、
【図5(b)】図5(b)は、本願によるページ処理方法の効果を示す図であり、
【図5(c)】図5(c)は、本願によるページ処理方法の効果を示す図であり、
【図6】図6は、本願実施例によるページ処理デバイスの構成ブロック図であり、
【図7】図7は、本願実施例による電子デバイスの構成ブロック図であり、

30

【図8】図8は、本願実施例によるコンピュータ読み取り可能な記憶媒体の構成ブロック図である。

40

【具体的な実施形態】

【0012】

当業者が本願の技術案をよりよく理解できるように、以下では図面を組み合わせる本願によるページ処理方法、デバイス、電子デバイス及びコンピュータ読み取り可能な記憶媒体について詳細に説明する。

【0013】

以下、図面を参照して例示的な実施例についてより詳細に説明するが、例示的な実施例は異なる形式で示されてもよく、本文に記載の実施例に限定されるものではない。むしろ、これらの実施例を提供する目的は、本願を徹底的かつ完全にし、当業者に本願の範囲を

50

十分に理解させることにある。本願の各実施例及び実施例における各特徴は、矛盾しない限り、互いに組み合わせることができる。

【0014】

図1は、本願の一実施例によるページ処理のアーキテクチャの概略図であり、図1に示すように、このアーキテクチャは、移動デバイス20とウェブサイト30を含んでもよく、移動デバイス20は、ブラウザカーネル21、メモリ22、及び表示画面23を含んでもよく、ウェブサイト30は、複数のページ31を含んでもよい。

【0015】

ここで、移動デバイス20は、パーソナルコンピュータ、スマートフォン、タブレット、携帯端末、サーバー等を含んでもよいが、これらに限定されない。それらには、例えば、メールボックスApp等の様々なアプリケーション（App）をインストールすることができる。

10

【0016】

本願実施例におけるページ31は、ランディングページを含んでもよいが、これに限定されない。ランディングページは、独立したウェブページを表示するために使用されてもよく、例えば、ユーザ又は訪問者がクリックして検索された広告、有料検索チャンネルをクリックして入ったページといったマーケティング又は広告活動のために使用されてもよい。

【0017】

一実施例において、ユーザ10が移動デバイス20を介してウェブサイト30にアクセスし、ウェブサイト30内のページ31のURLをクリックすると、ブラウジングカーネル21はURL（Uniform Resource Locator）に応じてHTMLファイルダウンロードを開始し、ダウンロードされたHTMLファイルを解析してDOM（Document Object Model）ツリーを取得すると同時に、HTMLファイル上のCSS（CaScading Style SheetS）及びJS（Java Script（登録商標））ファイル等のリソースリンクを解析した場合、CSSファイルダウンロード及びJSファイルダウンロードを開始し、ダウンロードされたCSSファイル及びJSファイルはメモリ22に保存される。

20

【0018】

ウェブページのサイトの挙動は非常に速く変化するため、ルール集合を配置することですべてのタイプ及びページを網羅することはできない。また、すべての広告がユーザのブラウジング体験に影響を与えるのではなく、広告要素がページの主要コンテンツをブラウジングするのに影響しない場所にあり、かつ誘導等の行為がない場合は、通常の商業的行為であり、大面積にわたって誤って損なわれれば、通常のインターネットエコロジーを損なうこともある。しかし、現在の多くの技術案は、上記のような通常の商業的行動の広告とユーザのブラウジング体験に影響を与える広告とを区別できない。ルール集合に基づいてウェブサイト内のページ要素をフィルタリングした場合、ルール集合が大き過ぎれば、ウェブページのロードの速度は著しく影響を受ける。

30

【0019】

本願実施例は、ページ処理方法を提供することができ、この方法は、移動デバイス20の表示画面23にページ31を表示する前に、ブラウザカーネル21のレンダリング段階でページ31内のページ要素の種類をインテリジェントに識別し、且つ自動的にユーザのブラウジング体験に影響するページ要素をマスクし、ページ31のレンダリングが完了した後に、ユーザ10が目にするのは最適化されたページであり、ユーザのブラウジング体験を大幅に改善し、且つモバイル検索のエコセキュリティに対する保護を提供する。

40

【0020】

以下の各実施例はいずれも本実施例のシステムアーキテクチャに応用できる。説明を簡潔にするため、以下の各実施例を相互に参照し引用することができる。

【0021】

図2は本願の一実施例におけるページ処理方法のフローチャートである。図2に示すように、当該ページ処理方法は、以下のステップを含んでもよい。

【0022】

50

S110は、取得されたHTMLファイルに基づいて、ページの複数のレイアウトオブジェクトノードを確定する。

【0023】

S120は、ページの複数のレイアウトオブジェクトノードをレイアウトした後、所定のリコール規則を用いて、複数のレイアウトオブジェクトノードをスクリーニングし、リコール規則に適合するレイアウトオブジェクトノードを取得する。

【0024】

S130は、リコール規則に適合するレイアウトオブジェクトノードが指定ターゲットノードであるか否かを予測する。

【0025】

S140では、指定ターゲットノードをマスク処理し、マスク処理後に残ったレイアウトオブジェクトノードを用いてマスク処理後のページを生成する。

【0026】

本願実施例のページ処理方法に基づけば、リコール規則とノード予測モデルを組み合わせる方式を用いてページを処理する。リコール規則によってスクリーニングされたレイアウトオブジェクトノードに対し、ノード予測モデルを用いてブラウジング体験に影響するか否かを確定し、ブラウジング体験に影響すると予測されたレイアウトオブジェクトノードをマスク処理し、マスク処理されたページを生成することにより、全体的にページのブラウジング体験を最適化し、モバイル検索のエコセキュリティに対する保障を提供する。

【0027】

本願実施例において、カーネルをレンダリングしてウェブページを処理するプロセスは非常に複雑なもので、処理性能及びユーザ体験の観点から、適切なタイミングを選択してターゲットノードを非表示にすることは非常に重要である。レイアウトオブジェクトノードのレイアウトとは、レイアウトオブジェクトノードの幅、高さ、位置等の幾何情報を割り当て算出するプロセスを示す。単にページ全体のレイアウトが完了する度に広告を識別し、ページ全体を再レイアウトすることで識別を完了することはできるが、ページを表示する際に数十乃至は百回を超えてレイアウトを行う必要があるとともにページ全体をトラバースしてターゲットノードを識別する必要があり、トラバースして再レイアウトするのに時間がかかり、ページ全体のロード時間に非常に大きな影響を与え、ページ全体のロードに対する感知が遅くなってしまう。

【0028】

そこで、本願実施例によるページ処理方法は、最も高い性能とユーザ体験を達成するために、ページ全体をトラバースすることなく、かつページ全体を再レイアウトすることなく、能動的に一部のレイアウトをトリガーできるようにしている。具体的に、上記ステップS120でページの複数のレイアウトオブジェクトノードをレイアウトした後、所定のリコール規則を用いてレイアウトオブジェクトノードをスクリーニングすることができる。

【0029】

すなわち、本願実施例において、ページ上の各ノードは、レイアウトする度に自身のレイアウト方法呼び出し、DOMツリーをトラバースするのが回避され、当該ノードのレイアウトが完成した後、当該ノードがブラウジング体験に影響を与えるターゲットノードとして識別されると、当該ターゲットノードにマスク処理を行い、例えば、ターゲットノードの状態を非表示に設定し、カーネルのレイアウト状態を再設定し、カーネルの再レイアウトを能動的に開始し、これにより、当該ノードを局所的に直接レイアウトし、ページレベル全体での再レイアウトを回避することができる。

【0030】

一実施例において、ステップS110は、具体的には、HTMLファイルを解析して、ドキュメントオブジェクトモデルDOM及びカスケードリング・スタイル・シートCSSを取得するS21と、CSSを解析して、DOMにおけるHTML要素ノードのスタイルデータを取得するS22と、DOMにおけるレンダリングを必要とするHTML要素ノード及びスタイルデータに従って、ページの複数のレイアウトオブジェクトノードを確定するS23と、を含んでよい。

10

20

30

40

50

【 0 0 3 1 】

ここで、各レイアウトオブジェクトノードは、レンダリングを必要とするHTML要素ノードに対応し、各レイアウトオブジェクトノードのスタイルデータは、対応するHTML要素ノードのスタイルデータである。

【 0 0 3 2 】

当該実施例において、ドキュメントオブジェクトモデルDOMは、ツリー構造のDOM、つまり、DOMツリーであってもよく、複数のレイアウトオブジェクトノードは、レイアウトオブジェクトLayout Objectツリー内のノードであってもよく、Layout Objectツリーを構築してレイアウトした後、Layout Objectツリーのノードは、座標、幅、高さの一連の属性情報を有することができる。

10

【 0 0 3 3 】

即ち、当該実施例では、Layout Objectツリー内の各ノードは、DOM内のレンダリングを必要とするHTML要素ノードに対応しており、新たに作成されたLayout Objectツリー内のレイアウトオブジェクトノードに対して、DOMツリー内のHTML要素ノードを記述するためのCSS属性オブジェクトが設定され、CSS内のスタイルデータに従ってLayout Objectツリー内のレイアウトオブジェクトノードが描画されるようになっている。

【 0 0 3 4 】

一実施例において、HTMLファイルを解析してスクリプトファイルリンクを取得する場合、ステップS23の前に、スクリプトファイルリンクに対応するスクリプトファイルをダウンロードして実行し、スクリプトファイルに対応するHTML要素ノードを取得するステップS31と、当該スクリプトファイルに対応するHTML要素ノードを、当該リコール規則に適合するレイアウトオブジェクトノードとするステップS32と、をさらに含んでもよい。

20

【 0 0 3 5 】

すなわち、幾つかの実施例において、ページの複数のレイアウトオブジェクトノードを確定した後、当該ページ処理方法は、前記複数のレイアウトオブジェクトノードの中にスクリプトファイルによりロードされたレイアウトオブジェクトノードが含まれている場合、スクリプトファイルによりロードされた前記レイアウトオブジェクトノードを、前記リコール規則に適合するレイアウトオブジェクトノードとして使用するステップをさらに含んでもよい。

【 0 0 3 6 】

当該実施例において、ブラウジング体験に影響を与える多くのターゲットノードはJSで動的にロードされるので、ノードがJSによりロードされるか否かの特徴に応じて、スクリプトファイルに対応するHTML要素ノードをリコール規則に適合するレイアウトオブジェクトノードとして確定して、識別を必要とするノードの一次フィルタリングを行い、これにより非同期的にロードされるJSリソースによってノードの再レイアウトをトリガーし、後でノード予測モデルを用いてブラウジング体験に影響を与えるノードの識別の予測に要する時間を効果的に減少させる。

30

【 0 0 3 7 】

図3は、本願の例示的な実施例におけるリコール規則の概略図である。本願実施例において、レイアウトオブジェクトノードを所定のリコール規則でスクリーニングすることを、規則に基づく粗リコールと称することもできる。

40

【 0 0 3 8 】

図3に示すように、規則に基づく粗リコールでは、ノードのアスペクト比、ノードの埋め込み形態、ノードの位置的特徴、ノードの内容、ノードの生成メカニズムとノードの構造等の面からノードのリコール条件を設定することができる。

【 0 0 3 9 】

すなわち、リコール規則は、ノードのアスペクト比、ノードの埋め込み形態、ノードの位置的特徴、ノードの内容、ノードの生成メカニズムとノードの構造のうち少なくとも1つに応じて事前に設定する規則を含んでもよい。

【 0 0 4 0 】

50

一実施例において、ステップS120は、具体的には、ページの複数のレイアウトオブジェクトノードをレイアウトし、レイアウトされたレイアウトオブジェクトノードの属性情報を取得するステップS41と、属性情報がリコール規則に定義されたノードリコール条件を満たすか否かを判断するステップS42と、ノードリコール条件を満たすレイアウトオブジェクトノードをリコール規則を満たすレイアウトオブジェクトノードとするステップS43と、を含んでもよい。

【0041】

一例として、ノードアスペクト比に従って設定された規則は、ノード高さ占有率が高さ占有率閾値より小さく、及び/又は、幅占有率が幅占有率閾値より小さいノードを、リコール規則に適合するノードとして設定することを含む。この例では、ブラウジング体験に影響を与えるノードは画面全体を占めることは少なく、ページ内に散在して又は浮遊して存在することがほとんどであり、画面の高さ占有率75%のノードがターゲットノードではない確率は非常に高く、幅占有率の閾値より小さい幅占有率を持つ他のノードのサブノードはフィルタリングすることができる。

10

【0042】

一例として、ノードの埋め込みに応じて設定される規則には、リコール規則に適合するノードとして、埋め込みが指定されたノードが含まれる。例えば、データ解析によれば、埋め込みフレームiframeノードは、ターゲットノードによく見られる寄生サイトであり、ノードには、多くの広告会社のインテグラルデータが含まれているため、iframeのあるノードも、擬似ターゲットノードの集合に含まれることが分かっている。

20

【0043】

一例として、ノードの位置的特徴に応じて設定される規則は、例えば、ノード位置的特徴に浮遊型のノードを含むノードを、リコール規則に適合するノードとする。この例では、ターゲットノードは、ページに対して固定、埋め込み、浮遊等の形式を有し、浮遊型のターゲットノードがブラウジング体験に最も悪影響を及ぼし、有効な情報を遮り、ユーザを強制的にオフにさせ、したがって、浮揚型のノードも、すべて、擬似ターゲットノードの集合に列挙される。

【0044】

一例として、ノードの内容の特徴に応じて設定される規則は、所定の種類の内容を有するノードをリコール規則に適合するノードとすることを含む。この例では、ノード内のテキスト、写真、対話型等のコンテンツが豊富である場合、非ターゲットノードである可能性が高い。

30

【0045】

一例として、ノード生成メカニズムに応じて設定される規則は、指定された生成メカニズムを有するノードを、リコールの規則に適合するノードとすることを含む。この例では、ページ内のノードがHTMLソースコードとJSで動的に生成されるノードとを含む場合、JSで生成されるノードはフレキシブルで変化しやすく、大部分のページの主なコンテンツがHTML内にあり、動的に変化する広告、関連推奨等を必要とする他のコンテンツがJSを用いて生成される。このため、JSで生成されたノードがターゲットノードである可能性が高い。

40

【0046】

一例として、ノード構造特徴に基づいて設定される規則は、指定された構造を有するノードをリコール規則に適合するノードとすることを含む。この例では、DOMツリー上にノードがある構造特徴をフィルタリングの基礎として使用することもできる。例えば、DOMツリー構造では、純粋なテキストのノードのみが大部分は非ターゲットノード（リコール規則に適合しないノード）、及びdiv/a/img形式を有するブロックレベルノードは、写真を通じて広く普及しているノードである可能性が高い。

【0047】

本願実施例のページ処理方法に基づけば、規則に基づく粗呼びリコールにおいて、リコール規則のうちの任意の1つに限定されるノードリコール条件にヒットする限り、当該ノ

50

ードが疑似ターゲットノードの特性を有していることを示すことができ、後続のターゲットノード判断ロジックが行われてもよく、全ての規則がヒットしない場合は、非ターゲットノードであると見なされ、こうして規則をリコールするフィルタリングポリシーによってブラウジング体験に影響を与えない多数の正常なノードがフィルタリングされる。

【0048】

一実施例において、前記ステップS130の前、さらに、

リコール規則に適合するレイアウトオブジェクトノードを初回スクリーニングしたレイアウトオブジェクトノードとし、初回スクリーニングしたレイアウトオブジェクトノードのノード状態を確定するS51と、

ページの全てのレイアウトオブジェクトノードのレイアウトが完了した後、ノード状態が変化したレイアウトオブジェクトノードを取得するS52と、

所定のリコール規則を再度用いて、ノード状態が変化したレイアウトオブジェクトノードをスクリーニングするS53と、

初回スクリーニングしたオブジェクトノードと再度スクリーニングしたオブジェクトノードを、リコール規則に適合するレイアウトオブジェクトノードとするS54とを含んでよい。

【0049】

当該実施例では、ノードのレイアウト過程において、ノードのいくつかは互いに依存関係を有するため、初回レイアウト時に正確なノード視覚情報が算出されておらず、これらのノードは粗リコールのポリシーを満たすことが困難である。したがって、全体のレイアウトが完了した後に、ノード状態（例えば、ノード視覚情報）が変化したノードをチェックし、当該ノードに対して規則に基づき再度粗リコールを行い、これにより、Recheckメカニズムによって、レイアウト過程で状態変化してリコール規則に適合するようになったノードをひとまとめにすることで、リコール規則に適合するノードをより多くリコールして、ターゲットノードの見落としを防止する。

【0050】

一実施例において、ステップS130は、具体的に、

リコール規則に適合するレイアウトオブジェクトノードの属性情報に基づいて、リコール規則に適合するレイアウトオブジェクトノードのノード特徴を算出するS61と、

所定のノード予測モデルを用いてノード特徴を処理し、リコール規則に適合するレイアウトオブジェクトノードが指定ターゲットノードである確率値を求めるS62と、

確率値に基づいて、リコール規則に適合するレイアウトオブジェクトノードが指定ターゲットノードであるか否かを確定するS63とを含んでよい。

【0051】

当該本実施例において、機械学習モデルを用いて、リコール規則に適合するノードがブラウジング体験に影響を与える指定ターゲットノードであるか否かを確定することができる。

【0052】

一実施例において、リコール規則に適合するレイアウトオブジェクトノードは、ページのレイアウトオブジェクトツリー内のノードである。S61は、具体的に、

リコール規則に適合するレイアウトオブジェクトノードの属性情報を取得するステップであって、属性情報はレイアウト過程で取得された情報であるS71と、

深さを優先的にトラバースする方式を採用し、属性情報を利用して、レイアウトオブジェクトツリーのリコール規則に適合するレイアウトオブジェクトノードについて、トップダウン特徴計算を行い、リコール規則に適合するレイアウトオブジェクトノードのノード特徴を取得するS72とを含んでよい。

【0053】

一実施例において、当該ノード特徴は、ノード視覚情報、ノード内容、及びノード構造等から抽出して計算される特定次元特徴である。この特定次元特徴は、例えば10以上の指定次元等の、実際の計算上の必要性に応じて設定されてもよく、本願実施例は、指定次元

10

20

30

40

50

を具体的に限定しない。

【0054】

この実施例では、レイアウトオブジェクトツリーを構築する際にトップダウンの特徴計算を用いることで各ノードのノード特徴を計算して親ノードに伝達することができ、このモードでは、ほとんどのページノードが特徴計算に参加する必要がある。ノードのレイアウト時に、リコール規則によって正常ノードのフィルタリングを先に行っているため、深さを優先的にトラバースする方式を採用し、トップダウンの特徴計算を用いることで、リコール規則に適合するレイアウトオブジェクトノード（すなわち、疑似ターゲットノード）のノード特徴を選択的に計算して、特徴を計算するノードの数を減少させ、ノード特徴の計算速度を向上させることができる。

10

【0055】

一実施例において、ノード予測モデルは、タグ付けされたオフラインレンダリング済みの静的ページデータを用いて事前にトレーニングされたモデルであり、指定された深さ及び指定された数の確定ツリーを有する勾配強調確定ツリーモデルである。

【0056】

一例としては、ブラウザカーネルによって処理されるノードの特徴は動的に変化するもので、トレーニングデータの選択時に、オフラインレンダリングが完成した静的データをマーキングし、高精度の自動マーキングツールが手動マーキングを支援するように設定し、最終的にトレーニングデータを構成することができる。

20

【0057】

一例としては、機械学習により得られたノード予測モデルは、勾配強調確定ツリーモデル（Gradient Boosted Decision Tree、GBDT）を含み、マーキングされたトレーニングデータを利用して当該GBDTモデルを予めトレーニングして得られ、指定された深さと指定された数の確定ツリー、例えば深さが4の100本のツリーのモデルファイルを得て、その後、直接当該モデルファイルを利用してリコール規則に適合するレイアウトオブジェクトノードが指定ターゲットノードであるか否かを予測する。

【0058】

上記トレーニングで得られたノード予測モデルの深さと確定ツリーの個数は例示的な数値であり、実際の応用シナリオにおいて、ユーザの実際の需要に基づいてモデルトレーニングを完成することができ、本願実施例は具体的に限定されないと理解されるべきである。

30

【0059】

一実施例において、ステップS140において、指定ターゲットノードに対してマスク処理を行うステップは、具体的に、

指定ターゲットノードの属性情報に基づいて対応するノード特徴情報を算出するステップであって、ノード特徴情報は、ページ内における位置、幅、高さ、主題コンテンツであるか否か、ページにおける面積占有率のうち少なくとも1つを含むS81と、

ノード特徴情報が対応する所定のマスク閾値に達する場合、指定ターゲットノードの状態を非表示に設定することにより、ブラウジング体験に影響を与えるレイアウトオブジェクトノードをマスク処理するS82とを含んでよい。

【0060】

40

本願実施例によるページ処理方法は、指定ターゲットノードの特徴に対し特別なオブジェクト処理メカニズムを適用することができるターゲットノードのマスクポリシーを提供する。ターゲットノードを識別した後、ページ全体のターゲットノードの特徴及び面積占有率を計算し、その後、設定可能なマスク閾値に従って、例えばページ内のノードの位置、アスペクト、主題のコンテンツ内にあるか否か等をマスクし、これにより、指定ターゲットノードを柔軟にマスクし、モバイル検索のエコセキュリティを維持及び保証し、全体的にページのブラウジング体験を最適化することを達成することができる。

【0061】

当該本実施例において、マスクはユーザのブラウジング体験に影響を与え、ページがレンダリング及び描画を完了した後、ユーザが見るのは最適化されたページであり、ユーザ

50

のブラウジング体験を大幅に改善し、モバイル検索のエコセキュリティに対する保護を提供する。

【0062】

本願実施例におけるページ処理方法は、指定ターゲットノードに対してマスク処理を行い、例えば、ノード状態を非表示に設定し、カーネルレイアウト状態を再設定し、カーネルの再レイアウトを能動的に開始し、ページ処理プロセス全体がノード描画の前に発生し、それによってユーザがページをブラウジングする際にページノード非表示のいかなるディザリング知覚もないことを保証し、それによって全体としてページブラウジング体験を最適化する。

【0063】

本願のページ処理方法をよりよく理解するために、以下、本願の他の実施例に係るページ処理フローを図4を参照して説明する。図4は本願の別の実施例によるページ処理方法のフローチャートである。図4に示すように、ページ処理方法は、以下のステップを含んでよい。

【0064】

S201において、ページURLに基づいてHTMLファイルをダウンロードする。

【0065】

S202において、パーサによりHTMLファイルを解析してDOMツリーを取得し、HTMLファイル上のCSS及びJSファイルリソースのリンクを解析するとき、CSSをダウンロード及び解析し、JSファイルをダウンロードして実行する。

【0066】

当該ステップにおいて、CSSをダウンロードして解析し、DOMツリー中のノードのスタイルデータを取得し、JSファイルをダウンロードして実行すると、JSにより動的にロードされたノードが得られ、当該動的にロードされたノードがDOMツリーに挿入/追加される。

【0067】

S203において、DOMツリーのレンダリングを必要とするHTML要素ノードと、DOMツリーのノードのスタイルデータとから、レイアウトオブジェクトLayout Objectツリーが構築される。

【0068】

S204において、Layout Objectツリーの構築が完了した後、レイアウト層Layout Layerツリーが作成される。

【0069】

当該ステップにおいて、層の位置決め及びレイアウトがLayout Layerツリーに基づいて実現され得る。

【0070】

S205において、JSの動的ロードによって生成された、レイアウトオブジェクトツリー内のノードをフィルタリングし、当該動的ロードによって生成されたノードの再レイアウトをトリガーするためにS209を実行する。

【0071】

図4において、JSの動的ローディングは非同期資源ローディングであるため、動的にロードされて生成されたノードの再レイアウトを行う過程は、非同期資源ローディングをトリガーとするノードの再レイアウトともいうことができる。

【0072】

S206において、Layout Objectツリー中のノードをレイアウトする過程で、レイアウトオブジェクトノードの属性情報を収集する。

【0073】

S207において、所定のノード予測モデルに基づいて、レイアウトオブジェクトノードがブラウジング体験に影響を与える指定ターゲットノードであるか否かを採点し、採点した結果から当該レイアウトオブジェクトノードがブラウジング体験に影響を与えるか否かを予測する。

10

20

30

40

50

【0074】

当該ステップにおいて、レイアウトオブジェクトノードのスコアは、そのレイアウトオブジェクトノードがブラウジング体験に影響を与える指定ターゲットノードであるか否かの確率値である。

【0075】

いくつかの実施例において、Layout Objectツリーにおける任意のノードをレイアウトした後、所定のリコール規則を用いて、任意のレイアウトオブジェクトノードをスクリーニングして、Layout Objectツリーにおける前記リコール規則に適合するレイアウトオブジェクトノードを得て、これにより、上記ステップS207において、所定のノード予測モデルに基づいて、前記リコール規則に適合するレイアウトオブジェクトノードがブラウジ
10

【0076】

S208において、ブラウジング体験に影響を与える指定ターゲットノードと予測される場合、ブラウザカーネルはレイアウト状態設定を行い、S209を実行して、当該レイアウトオブジェクトノードの再レイアウトを能動的にトリガーする。

【0077】

ステップS208では、指定ターゲットノードを再レイアウトすることにより、指定ターゲットノードをマスク処理（例えば、ノードステータスを非表示に設定する）することができる。

【0078】

S209において、レイアウトオブジェクトノードの再レイアウトを実行し、再レイアウトのマスク処理後のレイアウトオブジェクトノードを得る。
20

【0079】

S210において、マスク処理後のレイアウトオブジェクトノードに基づいてページを描画し、指定された表示画面に表示する。

【0080】

本願実施例のページのページ処理方法に基づけば、レンダリングされるノードのフィルタリングは、ページ内のブラウジング体験に影響を及ぼす要素をマスクするために、リコール規則及びポリシーに基づく前処理と機械学習モデルとの組合せの手法を用いて達成される。
30

【0081】

図5(a)～図5(c)は、本願実施例によるページ処理の効果を示す図である。図5(a)に示すように、ページ1には、複数のHTMLオブジェクト要素に対応する複数のレイアウトオブジェクトノード、例えば、ノード1、ノード2、ノード3、ノード4が含まれている。

【0082】

図5(a)におけるページ1の各レイアウトオブジェクトノードは、DOMツリーのトラバースを回避するために、レイアウト時に自身のレイアウト方法呼び出す。各レイアウトオブジェクトノードは、以下のステップを行うことができる。

【0083】

図5(b)のS301「規則による粗リコール」に示すように、ページの複数のレイアウトオブジェクトノードをレイアウトした後、所定のリコール規則を用いて、レイアウトオブジェクトノードをスクリーニングし、ページ中のリコール規則に適合するレイアウトオブジェクトノードを得る。
40

【0084】

ステップS301において、上述の実施例のステップS120と同一の処理プロセスを有し、本願実施例では改めて説明しない。

【0085】

図5(b)のS302「Recheckメカニズム」に示すように、ページの全てのレイアウトオブジェクトノードがレイアウトされた後、所定のリコール規則を再度用いて、ノード状態が変化したレイアウトオブジェクトノードをスクリーニングする。
50

【 0 0 8 6 】

ステップS302は、上述の実施例におけるステップS53と同じ処理プロセスを有し、本願実施例では説明を省略する。

【 0 0 8 7 】

図5 (b) のS303「モデルに基づくリコール」に示すように、所定のノード予測モデルに基づき、リコール規則に適合するレイアウトオブジェクトノードが指定ターゲットノードであるか否かを予測する。

【 0 0 8 8 】

ステップS303は、上述の実施例におけるステップS130と同一の処理プロセスを有し、本願実施例では改めて説明しない。

【 0 0 8 9 】

図5 (b) のS304「マスク処理」に示すように、指定ターゲットノードに対してマスク処理を行い、マスク処理後のレイアウトオブジェクトノードを用いて、マスク処理後のページを生成する。

【 0 0 9 0 】

ステップS304は、上述の実施例におけるステップS140と同じ処理手順を有し、本願実施例ではその説明を省略する。

【 0 0 9 1 】

図 (c) 5に示すように、ページ1のレンダリングが完了した後、ユーザが目にするのは最適化されたページ2であり、ユーザのブラウジング体験を大幅に向上させ、モバイル検索のエコセキュリティを確保する。

【 0 0 9 2 】

図6は、本願実施例によるページ処理デバイスの構成ブロック図である。図6に示すように、当該ページ処理デバイスは、以下のモジュールを含む。

【 0 0 9 3 】

ノード確定モジュール610は、取得されたHTMLファイルに基づき、ページの複数のレイアウトオブジェクトノードを確定する。

【 0 0 9 4 】

ノードスクリーニングモジュール620は、ページの複数のレイアウトオブジェクトノードをレイアウトした後、所定のリコール規則を用いて、レイアウトオブジェクトノードをスクリーニングし、リコール規則に適合するレイアウトオブジェクトノードを取得する。

【 0 0 9 5 】

リコール規則に適合するレイアウトオブジェクトノードが指定ターゲットノードであるか否かを予測する予測モジュール630とを備えることができる。

【 0 0 9 6 】

いくつかの実施例において、予測モジュール630は、所定のノード予測モデルに基づき、リコール規則に適合するレイアウトオブジェクトノードが指定ターゲットノードであるか否かを予測する。

【 0 0 9 7 】

マスク処理モジュール640は、指定ターゲットノードをマスク処理し、マスク処理後に残ったレイアウトオブジェクトノードを用いてマスク処理後のページを生成する。

【 0 0 9 8 】

本願実施例のページ処理デバイスに基づけば、ウェブサイトに表示されたページコンテンツをフィルタリングし、モバイル検索のエコセキュリティを保障し、ユーザのブラウジング体験を向上させることができる。

【 0 0 9 9 】

一実施例において、ノード確定モジュール610は、HTMLファイルを解析して、ドキュメントオブジェクトモデルDOM及びカスケーディング・スタイル・シートCSSを得る第1の解析ユニットと、CSSを解析してDOM内のHTML要素ノードのスタイルデータを取得する第2解析ユニットと、具体的に、DOMにおけるレ

10

20

30

40

50

ンダリングが必要なHTML要素ノード及びスタイルデータから、ページの複数のレイアウトオブジェクトノードを確定するノード確定モジュール610とを含んでよい。

【0100】

ここで、各レイアウトオブジェクトノードはレンダリングを必要とするHTML要素ノードに対応し、各レイアウトオブジェクトノードのスタイルデータは、対応するHTML要素ノードのスタイルデータである。

【0101】

一実施例において、HTMLファイルを解析してスクリプトファイルリンクを取得する場合、ノード確定モジュール610は、スクリプトファイルリンクに対応するスクリプトファイルをダウンロード及び実行して、スクリプトファイルに対応するHTML要素ノードを取得するダウンロード実行ユニットと、具体的に、スクリプトファイルに対応するHTML要素ノードを、リコール規則に適合するレイアウトオブジェクトノードとするノード確定モジュール610とをさらに含んでもよい。

10

【0102】

一実施例において、当該ノードスクリーニングモジュール620は、前記ページの複数のレイアウトオブジェクトノードを確定した後、前記複数のレイアウトオブジェクトノードの中にスクリプトファイルによりロードされたレイアウトオブジェクトノードが含まれている場合、スクリプトファイルによりロードされた前記レイアウトオブジェクトノードを前記リコール規則に適合するレイアウトオブジェクトノードとするようにさらに構成されてもよい。

20

【0103】

一実施例において、ノードスクリーニングモジュール620は、具体的に、前記ページの任意の一つのレイアウトオブジェクトノードをレイアウトし、前記レイアウトされたレイアウトオブジェクトノードの属性情報を取得する属性情報取得ユニットと、属性情報がリコール規則に限定されるノードリコール条件に適合するか否かを確定する適合条件判定ユニットと、ノードリコール条件を満たすレイアウトオブジェクトノードをリコール規則に適合するレイアウトオブジェクトノードとするリコールノード確定ユニットとを含んでもよい。

【0104】

一実施例において、リコール規則は、ノードのアスペクト比、ノードの埋め込み形式、ノードの位置的特徴、ノードの内容、ノードの生成メカニズム、及びノードの構造のうちの少なくとも1つに従って事前に設定された規則を含んでもよい。

30

【0105】

一実施例において、ページ処理デバイスは、リコール規則に適合するレイアウトオブジェクトノードを初回スクリーニングしたレイアウトオブジェクトノードとして、初回スクリーニングしたレイアウトオブジェクトノードのノード状態を確定するノード状態確定モジュールと、ページの全てのレイアウトオブジェクトノードがレイアウトを完了した後、ノード状態が変化したレイアウトオブジェクトノードを取得する状態変化ノード取得モジュールと、ノード状態が変化したレイアウトオブジェクトノードを、所定のリコール規則を再度用いてスクリーニングするノード再スクリーニングモジュールと、初回スクリーニングしたレイアウトオブジェクトノードと再度スクリーニングしたレイアウトオブジェクトノードとを、リコール規則に適合するレイアウトオブジェクトノードとするスクリーニングノード確定モジュールとをさらに含んでもよい。

40

【0106】

一実施例において、予測モジュール630は、リコール規則に適合するレイアウトオブジェクトノードの属性情報に基づいて、リコール規則に適合するレイアウトオブジェクトノードのノード特徴を計算する特徴計算ユニットと、所定のノード予測モデルを用いてノード特徴を処理し、リコール規則に適合するレイアウトオブジェクトノードが指定ターゲットノードである確率値を取得する確率計算ユニットと、確率値に基づいてリコール規則に適合するレイアウトオブジェクトノードが指定ターゲットノードであるか否かを確定する

50

ターゲットノード確定ユニットとを含んでもよい。

【0107】

一実施例において、リコール規則に適合するレイアウトオブジェクトノードは、ページのレイアウトオブジェクトツリー内のノードである。

【0108】

当該実施例において、特徴計算ユニットは、リコール規則に適合するレイアウトオブジェクトノードの、レイアウト過程において取得された情報である属性情報を取得する属性情報収集サブユニットと、具体的に、深さを優先的にトラバースする方式を採用して、属性情報を利用して、レイアウトオブジェクトツリーにおけるリコール規則に適合するレイアウトオブジェクトノードに対してトップダウン特徴計算を行い、リコール規則に適合するレイアウトオブジェクトノードのノード特徴を取得する特徴計算ユニットと、を含んでもよい。

10

【0109】

一実施例において、ノード予測モデルは、タグ付けされたオフラインレンダリング済みの静的ページデータを用いて事前にトレーニングされたモデルであり、指定された深さ及び指定された数の確定ツリーを有する勾配強調確定ツリーモデルである。

【0110】

一実施例において、マスク処理モジュール640は、具体的に、指定ターゲットノードの属性情報に基づいて、対応するノード特性情報を計算し、ノード特性情報は、ページ内の位置、幅、高さ、主題コンテンツ内であるか否か、及びページ内の面積占有率のうちの少なくとも1つを含む特性計算ユニットと、ノード特性情報が対応する所定のマスク閾値に達すると、前記指定ターゲットノードの状態を非表示に設定することにより、指定ターゲットノードをマスク処理するノードマスクユニットと、を含んでもよい。

20

【0111】

一実施例において、マスク処理モジュール640は、具体的に、マスク処理後の残りのレイアウトオブジェクトノードで再レイアウトし、再レイアウト後のレイアウトオブジェクトノードで描画して、描画されたマスク処理後のページを得る描画ユニットをさらに含んでもよい。

【0112】

本願実施例のページ処理デバイスに基づき、規則によるリコールとモデルによる予測とを組み合わせたスキームを用いて、指定ターゲットノードに対してマスク処理を行うことにより、ページ処理全体のプロセスがノードの描画前に発生し、ページをブラウジングする際にユーザがページノードを非表示にする揺れが感知されないことを保証し、全体としてページのブラウジング体験を最適化する。

30

【0113】

図7は、本願実施例による電子デバイスの構成ブロック図である。図7に示すように、本願実施例は、

1つ又は複数のプロセッサ701と、

1つ又は複数のプログラムが記憶され、1つ又は複数のプログラムが1つ又は複数のプロセッサで実行されると、1つ又は複数のプロセッサが上記任意のページ処理方法を実現するメモリ702と、

40

プロセッサとメモリとの間に接続され、プロセッサとメモリとの情報のやりとりを実現する1つ又は複数のI/Oインターフェース703とを備える電子デバイス700を提供する。

【0114】

ここで、プロセッサ701は、データ処理能力を有する素子であり、これには、中央処理装置（CPU）等が含まれるがこれに限定されない。メモリ702は、データ記憶能力を有する素子であり、これには、ランダムアクセスメモリ（ROM、より具体的には、SDRAM、DDR等）、リードオンリーメモリ（ROM）、電氣的消去可能プログラマブルリードオンリーメモリ（EEPROM（登録商標））、フラッシュメモリ（FLASH（登録商標））が含まれるがこれらに限定されない。I/Oインターフェース（読み取り/書き込み用インターフェ

50

ース) 703は、プロセッサ701とメモリ702との間に接続され、プロセッサ701とメモリ702との情報のやりとりを可能にし、これには、データバス(BuS)等が含まれるがこれに限定されない。

【0115】

いくつかの実施例において、プロセッサ701、メモリ702及びI/Oインターフェース703は、バスを介して互いに接続され、電子デバイス700の他の構成要素と接続される。

【0116】

図8は、本願実施例によるコンピュータ読み取り可能な記憶媒体の構成ブロック図である。図8に示すように、本願実施例は、プロセッサによって実行されるときに、上記任意の一つのページ処理方法を実行するコンピュータプログラムが記憶されたコンピュータ読み取り可能な記憶媒体を提供する。

10

【0117】

当業者は、以上で公開されている方法におけるステップ、システム、デバイスにおける機能モジュール/ユニットの全部又は一部は、ソフトウェア、ファームウェア、ハードウェア、及びそれらの適切な組み合わせとして実施され得ることを理解するであろう。ハードウェアによる実施形態において、上記の説明で言及された機能モジュール/ユニット間の区分は、必ずしも物理的コンポーネントの区分に対応しない。例えば、1つの物理的コンポーネントは複数の機能を有してもよく、あるいは1つの機能又はステップは幾つかの物理的コンポーネントが連携して実行されてもよい。一部又は全ての物理的コンポーネントは、中央処理装置、デジタル信号プロセッサ、若しくはマイクロプロセッサ等のプロセッサによって実行されるソフトウェアとして実施されてもよく、ハードウェアとして、又は専用集積回路等の集積回路として実施されてもよい。このようなソフトウェアはコンピュータ読み取り可能な媒体上に配置されてもよく、コンピュータ読み取り可能な媒体は、コンピュータ記憶媒体(又は非一時的媒体)及び通信媒体(又は一時的媒体)を含んでもよい。当業者によく知られているように、コンピュータ記憶媒体という用語は、コンピュータ読み取り可能なコマンド、データ構造、プログラムモジュール又は他のデータ等の情報を記憶するための任意の方法又は技術で実施される揮発性及び不揮発性、取り外し可能及び取り外し不可能な媒体を含む。コンピュータ記憶媒体は、RAM、ROM、EEPROM、フラッシュメモリ若しくは他のメモリ技術、CD-ROM、デジタル多用途ディスク(DVD)若しくは他の光ディスクストレージ、磁気カセット、磁気テープ、磁気ディスクストレージ若しくは他の磁気ストレージ装置、又は所望の情報を記憶するために使用され得、コンピュータによってアクセスされ得る任意の他の媒体を含むが、これらに限定されない。また、通信媒体は通常コンピュータ読み取り可能なコマンド、データ構造、プログラムモジュール、又は搬送波若しくは他の搬送メカニズム等の変調データ信号内の他のデータを含み、任意の情報配信媒体を含み得ることが当業者に知られている。

20

30

【0118】

本文では例示的な実施例が公開されており、また具体的な用語が用いられているが、それらは、一般的な例示的な意味でのみ使用され、解釈されるべきであり、限定を目的としたものではない。一部の例では、特に明記しない限り、特定の実施例と組み合わせて記載された特徴、特性及び/又は要素は、単独で使用されてもよいし、他の実施例と組み合わせて記載された特徴、特性及び/又は要素と組み合わせて使用されてもよいというのは当業者にとって自明であろう。よって、添付の特許請求の範囲に記載される本願の範囲から逸脱しない場合、様々な形態及び詳細における変更が可能であることを当業者は理解するであろう。

40

以下に、出願当初の特許請求の範囲に記載の事項を、そのまま、付記しておく。

〔C1〕

取得したハイパーテキストマークアップ言語HTMLファイルに基づいて、ページの複数のレイアウトオブジェクトノードを確定するステップと、

前記ページの複数のレイアウトオブジェクトノードをレイアウトした後、所定のリコール規則を用いて、前記複数のレイアウトオブジェクトノードをスクリーニングして、前記

50

リコール規則に適合するレイアウトオブジェクトノードを取得するステップと、
 前記リコール規則に適合するレイアウトオブジェクトノードが指定ターゲットノードであるか否かを予測するステップと、
 前記指定ターゲットノードをマスク処理し、マスク処理後に残ったレイアウトオブジェクトノードを用いて、前記マスク処理されたページを生成するステップと、を含む
 ページ処理方法。

〔 C 2 〕

ページの複数のレイアウトオブジェクトノードを確定する前記ステップの後に、
 前記複数のレイアウトオブジェクトノードの中にスクリプトファイルでロードされたレイアウトオブジェクトノードが含まれると確定されたことに応答して、前記スクリプトファイルでロードされたレイアウトオブジェクトノードを前記リコール規則に適合するレイアウトオブジェクトノードとするステップをさらに含む、
 C 1に記載の方法。

〔 C 3 〕

前記ページの複数のレイアウトオブジェクトノードをレイアウトした後、所定のリコール規則を用いて前記レイアウトオブジェクトノードをスクリーニングして、前記複数のレイアウトオブジェクトノードのうちの前記リコール規則に適合するレイアウトオブジェクトノードを取得する前記ステップは、
 前記ページのレイアウトオブジェクトノードをレイアウトし、前記レイアウトされたレイアウトオブジェクトノードの属性情報を取得するステップと、
 前記属性情報が前記リコール規則に限定されたノードリコール条件を満たすかどうかを判断するステップと、
 前記ノードリコール条件を満たすレイアウトオブジェクトノードを、前記リコール規則に適合するレイアウトオブジェクトノードとするステップとを含む、
 C 1に記載の方法。

〔 C 4 〕

前記リコール規則は、ノードのアスペクト比、ノードの埋め込み形式、ノードの位置的特徴、ノードの内容、ノードの生成メカニズム、ノードの構造の少なくとも1つに応じて設けられる規則を含む、
 C 3に記載の方法。

〔 C 5 〕

前記リコール規則に適合する前記レイアウトオブジェクトノードが指定ターゲットノードであるかどうかを予測する前記ステップの前に、
 前記リコール規則に適合するレイアウトオブジェクトノードを、初回スクリーニングで得られたレイアウトオブジェクトノードとし、前記初回スクリーニングで得られたレイアウトオブジェクトノードのノード状態を確定するステップと、
 前記ページの全てのレイアウトオブジェクトノードでレイアウトが完了した後、ノード状態が変化するレイアウトオブジェクトノードを取得するステップと、
 所定のリコール規則を再度利用して、ノード状態が変化するレイアウトオブジェクトノードをスクリーニングするステップと、
 前記初回スクリーニングで得られたレイアウトオブジェクトノードと前記再度スクリーニングで得られたレイアウトオブジェクトノードとを、前記リコール規則に適合するレイアウトオブジェクトノードとするステップをさらに含む、
 C 1から4のいずれか1項に記載の方法。

〔 C 6 〕

前記リコール規則に適合するレイアウトオブジェクトノードが指定ターゲットノードであるか否かを予測する前記ステップは、
 前記リコール規則に適合する前記レイアウトオブジェクトノードの属性情報に基づいて、前記リコール規則に適合する前記レイアウトオブジェクトノードのノード特徴を計算するステップと、

10

20

30

40

50

所定のノード予測モデルを用いて前記ノード特徴を処理し、前記リコール規則に適合する前記レイアウトオブジェクトノードが前記指定ターゲットノードである確率値を取得するステップと、

前記確率値に基づいて、前記リコール規則に適合する前記レイアウトオブジェクトノードが前記指定ターゲットノードであるか否かを確定するステップと、を含む、

C 1から5のいずれか1項に記載の方法。

[C 7]

前記リコール規則に適合する前記レイアウトオブジェクトノードは、前記ページのレイアウトオブジェクトツリーにおけるノードであり、

前記リコール規則に適合する前記レイアウトオブジェクトノードの属性情報に基づき、前記リコール規則に適合する前記レイアウトオブジェクトノードのノード特徴を計算するステップには、

前記リコール規則に適合する前記レイアウトオブジェクトノードの属性情報を取得し、前記属性情報は前記レイアウト過程で取得された情報であり、

深さを優先的にトラバースする方式を用い、前記属性情報を利用して、前記レイアウトオブジェクトツリーにおける前記リコール規則に適合するレイアウトオブジェクトノードに対し、トップダウン特徴計算を行い、前記リコール規則に適合するレイアウトオブジェクトノードのノード特徴を得ることが含まれる、

C 6に記載の方法。

[C 8]

前記ノード予測モデルは、マーキングされたオフラインレンダリング済みの静的ページデータを用いて予めトレーニングされたモデルであり、指定された深さ及び指定された数の確定ツリーを有する勾配強調確定ツリーモデルである、

C 6から7のいずれか1項に記載の方法。

[C 9]

前記指定ターゲットノードをマスク処理する前記ステップには、

前記指定ターゲットノードの属性情報に基づいて、前記ページ内の位置、幅、高さ、主題のコンテンツ内であるか否か、及び前記ページ内の面積占有率のうち少なくとも1つを含む対応するノード特性情報を計算することと、

前記ノード特性情報が所定のマスク閾値に達したと確定されたことに応答して、前記指定ターゲットノードの状態を非表示に設定してマスク処理を行うことが含まれる、

C 1から8のいずれか1項に記載の方法。

[C 1 0]

取得されたハイパーテキストマークアップ言語HTMLファイルに基づき、ページの複数のレイアウトオブジェクトノードを確定するノード確定モジュールと、

前記ページの複数のレイアウトオブジェクトノードをレイアウトした後、所定のリコール規則を用いて、前記複数のレイアウトオブジェクトノードをスクリーニングして、前記リコール規則に適合するレイアウトオブジェクトノードを取得するノードスクリーニングモジュールと、

前記リコール規則に適合するレイアウトオブジェクトノードが指定ターゲットノードであるか否かを予測する予測モジュールと、

前記指定ターゲットノードをマスク処理し、マスク処理後に残ったレイアウトオブジェクトノードを用いて前記マスク処理後のページを生成するマスク処理モジュールと、を含むページ処理デバイス。

[C 1 1]

1つ又は複数のプロセッサと、

1つ又は複数のプログラムが記憶され、前記1つ又は複数のプログラムが前記1つ又は複数のプロセッサで実行されると、前記1つ又は複数のプロセッサがC 1～9のいずれか1項に記載のデータ検索方法を実現するメモリと、

前記プロセッサとメモリとの間に接続され、前記プロセッサとメモリとの情報のやりと

10

20

30

40

50

りを実現するように構成された1つ又は複数のI/Oインターフェースとを備える、電子デバイス。

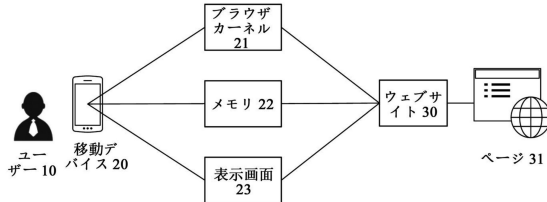
[C 1 2]

プロセッサによって実行されるときに、C 1 ~ 9のいずれか1項に記載のページ処理方法を実行するコンピュータプログラムが記憶された

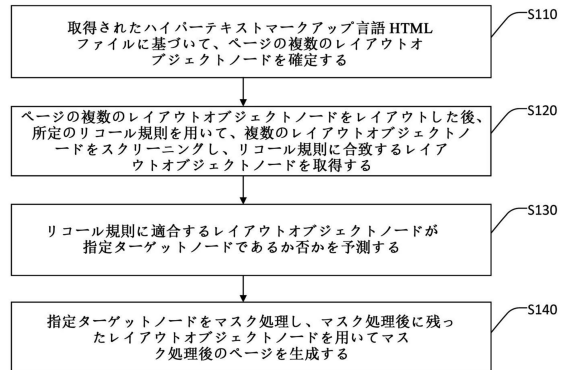
コンピュータ読み取り可能な記憶媒体。

【 図 面 】

【 図 1 】



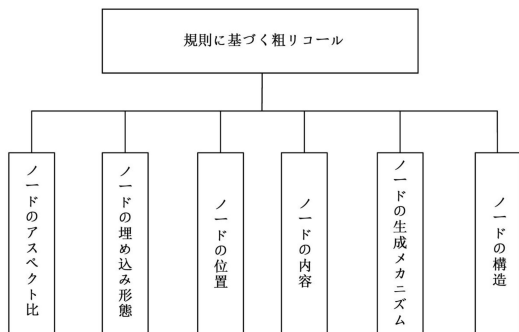
【 図 2 】



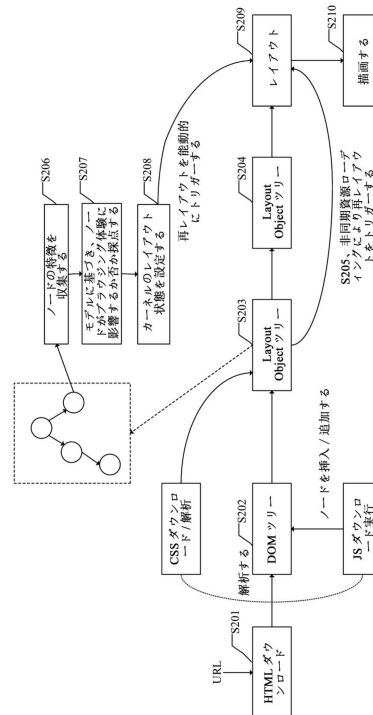
10

20

【 図 3 】



【 図 4 】

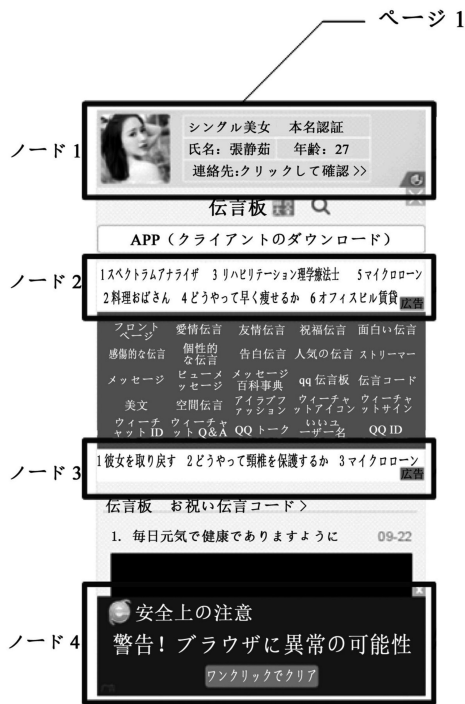


30

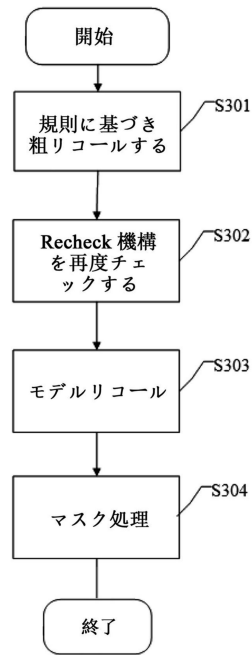
40

50

【図5(a)】



【図5(b)】



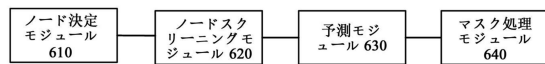
10

20

【図5(c)】



【図6】

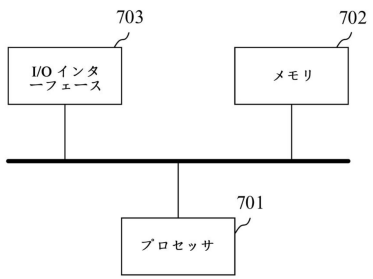


30

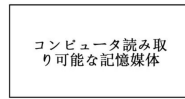
40

50

【図7】



【図8】



10

20

30

40

50

フロントページの続き

(72)発明者 ワン、インフェイ

中華人民共和国、ベイジン、ハイディエン・ディストリクト、シャンディ・テンス・ストリート、
ナンバー 10、パイドウ・キャンパス 3 / フロア

審査官 酒井 恭信

(56)参考文献 中国特許出願公開第110489636 (CN, A)

米国特許出願公開第2013/0145255 (US, A1)

米国特許出願公開第2018/0341986 (US, A1)

特開2009-199302 (JP, A)

特開2010-257412 (JP, A)

特開2011-044116 (JP, A)

(58)調査した分野 (Int.Cl., DB名)

G06F 16/00 - 16/958