

(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2006-179003

(P2006-179003A)

(43) 公開日 平成18年7月6日(2006.7.6)

(51) Int. Cl.	F I	テーマコード (参考)
G06F 17/30 (2006.01)	G06F 17/30 220A	5B050
G06T 11/60 (2006.01)	G06T 11/60 100A	5B075
	G06F 17/30 170B	
	G06F 17/30 380F	

審査請求 未請求 請求項の数 30 O L (全 25 頁)

(21) 出願番号	特願2005-368491 (P2005-368491)	(71) 出願人	000006747
(22) 出願日	平成17年12月21日 (2005.12.21)		株式会社リコー
(31) 優先権主張番号	11/023142		東京都大田区中馬込1丁目3番6号
(32) 優先日	平成16年12月22日 (2004.12.22)	(74) 代理人	100070150
(33) 優先権主張国	米国 (US)		弁理士 伊東 忠彦
		(72) 発明者	キャスリン パークナー
			アメリカ合衆国, カリフォルニア 940
			25, メンロ・パーク, サンド・ヒル・ロ
			ード 2882番, スイート 115 リ
			コー イノベーション インク内
		(72) 発明者	ジンガオ ミアオ
			アメリカ合衆国, カリフォルニア 940
			25, メンロ・パーク, サンド・ヒル・ロ
			ード 2882番, スイート 115 リ
			コー イノベーション インク内

最終頁に続く

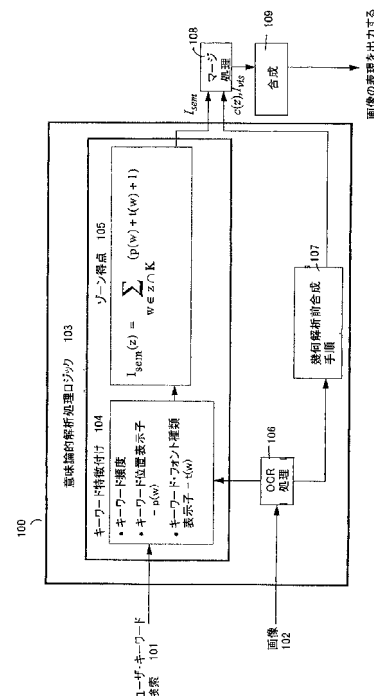
(54) 【発明の名称】 意味論的文書スマートネール

(57) 【要約】

【課題】 文書の視覚化を意味情報に基づいて生成する方法及び装置を開示する。

【解決手段】 一実施例では、方法は、意味論的重要度を複数の文書画像セグメント毎に含む複数の属性を有する複数の画像セグメントを選択する工程と、複数の文書画像セグメント上の複数の位置と、複数の文書画像セグメントに対して行う対象の複数の幾何演算とを複数の属性と、表示の制約とによって判定する工程とを備える。

【選択図】 図1



【特許請求の範囲】

【請求項1】

方法であって、

意味論的重要度情報を複数の文書画像セグメント毎に含む複数の属性を有する複数の画像セグメントを選択する工程と、

前記複数の文書画像セグメント上の複数の位置と、前記複数の文書画像セグメントに対して行う対象の複数の幾何演算とを前記複数の属性と、表示の制約とによって判定する工程とを備えることを特徴とする方法。

【請求項2】

請求項1記載の方法であって、

文書の文書画像セグメント群に対して意味論的文書解析を行う工程を更に備え、

前記複数の画像セグメントは、前記文書画像セグメント群から前記意味論的重要度情報に基づいて選択されることを特徴とする方法。

【請求項3】

請求項2記載の方法であって、

意味論的文書解析を行う工程は、前記文書画像セグメントの各々からキーワード属性を抽出する工程を備えることを特徴とする方法。

【請求項4】

請求項1記載の方法であって、

前記意味論的重要度情報を生成する工程と、

視覚重要度情報を文書画像セグメント毎に生成する工程と、

前記意味論的重要度情報及び前記視覚重要度情報を文書画像セグメント毎にマージして、マージされた重要度情報を文書画像セグメント毎に生成する工程と、

前記複数の文書画像セグメントを文書の文書画像セグメント群から前記マージされた重要度情報に基づいて選択する工程とを更に備えることを特徴とする方法。

【請求項5】

請求項4記載の方法であって、

前記意味論的重要度情報が意味論的重要度値を備え、

前記視覚重要度情報が視覚重要度値を備えることを特徴とする方法。

【請求項6】

請求項4記載の方法であって、

大文字表示子を画像セグメント毎に生成する工程を更に備え、

意味論的重要度情報及び視覚重要度情報を画像セグメント毎にマージする工程が、前記意味論的重要度情報と、視覚重要度情報と、前記大文字表示子とをマージして、マージされた重要度情報を文書画像セグメント毎に生成する工程を備えることを特徴とする方法。

【請求項7】

請求項1記載の方法であって、

前記複数の文書画像セグメントを再フォーマットして前記文書画像セグメントを目標キャンバス上に表示する工程を更に備えることを特徴とする方法。

【請求項8】

1つ又は複数の記録可能媒体を有する製造物品であって、該1つ又は複数の記録可能媒体は、システムによって実行されると該システムに方法を行わせる命令を前記1つ又は複数の記録可能媒体内に記憶させており、

前記方法は、

意味論的重要度情報を複数の文書画像セグメント毎に含む複数の属性を有する複数の画像セグメントを選択する工程と、

前記複数の文書画像セグメント上の複数の位置と、前記複数の文書画像セグメントに対して行う対象の複数の幾何演算とを前記複数の属性と、表示の制約とによって判定する工程とを備えることを特徴とする製造物品。

【請求項9】

10

20

30

40

50

請求項8記載の製造物品であって、

前記方法は、文書の文書画像セグメント群に対して意味論的文書解析を行う工程を更に備え、

前記複数の画像セグメントは、前記文書画像セグメント群から前記意味論的重要度情報に基づいて選択されることを特徴とする製造物品。

【請求項10】

請求項9記載の製造物品であって、

意味論的文書解析を行う工程は、前記文書画像セグメントの各々からキーワード属性を抽出する工程を備えることを特徴とする製造物品。

【請求項11】

10

請求項8記載の製造物品であって、前記方法は、

前記意味論的重要度情報を生成する工程と、

視覚重要度情報を文書画像セグメント毎に生成する工程と、

意味論的重要度情報及び視覚重要度情報を文書画像セグメント毎にマージして、マージされた重要度情報を文書画像セグメント毎に生成する工程と、

前記複数の画像セグメントを文書の文書画像セグメント群から前記マージされた重要度情報に基づいて選択する工程とを更に備えることを特徴とする製造物品。

【請求項12】

請求項11記載の製造物品であって、

前記意味論的重要度情報が意味論的重要度値を備え、

20

前記視覚重要度情報が視覚重要度値を備えることを特徴とする製造物品。

【請求項13】

請求項11記載の製造物品であって、

前記方法は、大文字表示子を文書画像セグメント毎に生成する工程を更に備え、

意味論的重要度情報及び視覚重要度情報を文書画像セグメント毎にマージする工程が、前記意味論的重要度情報と、視覚重要度情報と、前記大文字表示子とをマージして、マージされた重要度情報を文書画像セグメント毎に生成する工程を備えることを特徴とする製造物品。

【請求項14】

請求項8記載の製造物品であって、前記方法は、

30

前記複数の文書画像セグメントを再フォーマットして前記文書画像セグメントを目標キャンバス上に表示する工程を更に備えることを特徴とする製造物品。

【請求項15】

装置であって、

意味論的重要度情報を複数の文書画像セグメント毎に含む複数の属性を有する複数の画像セグメントを選択する解析器と、

前記複数の文書画像セグメント上の複数の位置と、前記複数の文書画像セグメントに対して行う対象の複数の幾何演算とを前記複数の属性と、表示の制約とによって判定するマージ装置とを備えることを特徴とする装置。

【請求項16】

40

方法であって、

意味論的文書解析を文書に対して行う工程と、

意味情報を前記文書から前記意味論的文書解析の結果に基づいて抽出する工程と、

抽出された意味情報を、抽出された視覚情報とマージする工程と、

前記文書の視覚表現のレイアウトをマージの結果に基づいて作成する工程とを備えることを特徴とする方法。

【請求項17】

1つ又は複数の記録可能媒体を有する製造物品であって、該1つ又は複数の記録可能媒体は、システムによって実行されると該システムに方法を行わせる命令を前記1つ又は複数の記録可能媒体内に記憶させており、

50

前記方法は、
意味論的文書解析を文書に対して行う工程と、
意味情報を前記文書から前記意味論的文書解析の結果に基づいて抽出する工程と、
抽出された意味情報を、抽出された視覚情報とマージする工程と、
前記文書の視覚表現のレイアウトをマージの結果に基づいて作成する工程とを備えることを特徴とする製造物品。

【請求項 18】

方法であって、
意味論的重要度表示情報を複数の画像セグメント毎に含む複数の属性を有する複数の画像セグメントを生成する工程と、
関連した意味論的重要度情報に基づいて選択される前記複数の画像セグメントの1つ又は複数の画像セグメントに対して複数の演算を行う工程と、
処理された画像セグメントを目標キャンバス上に表示する工程とを備えることを特徴とする方法。

10

【請求項 19】

1つ又は複数の記録可能媒体を有する製造物品であって、該1つ又は複数の記録可能媒体は、システムによって実行されると該システムに方法を行わせる命令を前記1つ又は複数の記録可能媒体内に記憶させており、

前記方法は、
意味論的重要度表示情報を複数の画像セグメント毎に含む複数の属性を有する複数の画像セグメントを生成する工程と、

20

関連した意味論的重要度情報に基づいて選択される前記複数の画像セグメントの1つ又は複数の画像セグメントに対して複数の演算を行う工程と、

処理された画像セグメントを目標キャンバス上に表示する工程とを備えることを特徴とする製造物品。

【請求項 20】

方法であって、
目標表示に含めるよう、文書における領域群を、該領域群における各領域と関連した意味論的重要度情報に基づいて選ぶ工程と、

目標表示に含めるよう選ばれる前記領域群を更新することとするか否かを判定する工程と、

30

前記領域群を更新することとしない場合に、絶対位置領域又は相対位置領域を前記目標表示上の絶対位置に変換する工程と、

前記文書に対して処理する対象の命令の出力リストを生成する工程とを備えることを特徴とする方法。

【請求項 21】

1つ又は複数の記録可能媒体を有する製造物品であって、該1つ又は複数の記録可能媒体は、システムによって実行されると該システムに方法を行わせる命令を前記1つ又は複数の記録可能媒体内に記憶させており、

前記方法は、
目標表示に含めるよう、文書における領域群を、該領域群における各領域と関連した意味論的重要度情報に基づいて選ぶ工程と、

40

目標表示に含めるよう選ばれる前記領域群を更新することとするか否かを判定する工程と、

前記領域群を更新することとしない場合に、絶対位置領域又は相対位置領域を前記目標表示上の絶対位置に変換する工程と、

前記文書に対して処理する対象の命令の出力リストを生成する工程とを備えることを特徴とする製造物品。

【請求項 22】

請求項15記載の装置であって、

50

文書の文書画像セグメント群に対して意味論的文書解析を行う手段を更に備え、
前記複数の画像セグメントは、前記文書画像セグメント群から前記意味論的重要度情報に基づいて選択されることを特徴とする装置。

【請求項 23】

請求項22記載の装置であって、
意味論的文書解析を行う手段は、前記文書画像セグメントの各々からキーワード属性を抽出する手段を備えることを特徴とする装置。

【請求項 24】

請求項15記載の装置であって、
前記意味論的重要度情報を生成する手段と、
視覚重要度情報を文書画像セグメント毎に生成する手段と、
前記意味論的重要度情報及び前記視覚重要度情報を文書画像セグメント毎にマージして、マージされた重要度情報を文書画像セグメント毎に生成する手段と、
前記複数の文書画像セグメントを文書の文書画像セグメント群から前記マージされた重要度情報に基づいて選択する手段とを更に備えることを特徴とする装置。

【請求項 25】

請求項24記載の装置であって、
前記意味論的重要度情報が意味論的重要度値を備え、
前記視覚重要度情報が視覚重要度値を備えることを特徴とする装置。

【請求項 26】

請求項24記載の装置であって、
大文字表示子を画像セグメント毎に生成する手段を更に備え、
意味論的重要度情報及び視覚重要度情報を画像セグメント毎にマージする手段が、前記意味論的重要度情報と、視覚重要度情報と、前記大文字表示子とをマージして、マージされた重要度情報を文書画像セグメント毎に生成する手段を備えることを特徴とする装置。

【請求項 27】

請求項15記載の装置であって、
前記複数の文書画像セグメントを再フォーマットして前記文書画像セグメントを目標キャンバス上に表示する手段を更に備えることを特徴とする装置。

【請求項 28】

装置であって、
意味論的文書解析を文書に対して行う手段と、
意味情報を前記文書から前記意味論的文書解析の結果に基づいて抽出する手段と、
抽出された意味情報を、抽出された視覚情報とマージする手段と、
前記文書の視覚表現のレイアウトをマージの結果に基づいて作成する手段とを備えることを特徴とする装置。

【請求項 29】

装置であって、
意味論的重要度表示情報を複数の画像セグメント毎に含む複数の属性を有する複数の画像セグメントを生成する手段と、
関連した意味論的重要度情報に基づいて選択される前記複数の画像セグメントの1つ又は複数の画像セグメントに対して複数の演算を行う手段と、
処理された画像セグメントを目標キャンバス上に表示する手段とを備えることを特徴とする装置。

【請求項 30】

装置であって、
目標表示に含めるよう、文書における領域群を、該領域群における各領域と関連した意味論的重要度情報に基づいて選ぶ手段と、
目標表示に含めるよう選ばれる前記領域群を更新することとするか否かを判定する手段と、

10

20

30

40

50

前記領域群を更新することとしない場合に、絶対位置領域又は相対位置領域を前記目標表示上の絶対位置に変換する手段と、

前記文書に対して処理する対象の命令の出力リストを生成する手段とを備えることを特徴とする装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は一般的に画像処理の分野に関し、特にコンテンツが意味情報に基づいている文書に対するサムネールの生成に関する。

【背景技術】

【0002】

表示技術の向上、モバイル・コンピューティング機能及び記憶機能の拡充に伴って、ハンドヘルド型機器、携帯電話機、ブラウザ・ウィンドウやサムネールなどの一層小型化された物理表示上及び仮想表示上に示す画像情報の量が増加する傾向にある。そのような画像情報表示には、再目的化と呼ばれることの多い、元々記憶された画像の再フォーマッティングが必要である。

【0003】

表示寸法が制約されたPDAや携帯電話機などのハンドヘルド型機器の普及の進展によって、ユーザが効果的かつ効率的に取り出すことができるような適切な方法で文書を表現する技術が要求されている。伝統的なサムネールは元の文書を小型にスケーリングするだけで作成され、この文書に含まれている情報の読みやすさを考慮に入れていないという大きな課題がある。よって、文書が走査されることを前提とすれば、文書全体を（以降キャンバスと呼ぶ）目標表示にダウンサンプリングして寸法変更することによって、テキストの読みやすさと画像特徴の認識可能性が失われてしまうことが多い。

【0004】

小型表示機器に収めるよう文書及び画像の再フォーマッティングを行うことに関する従来技術として多くのものがある。従来技術には、PDA表示の幅にコンテンツ情報を収め、コンテンツを垂直方向にスクロールすることを可能にすることによって文書画像の単一ページのPDAへのリフローを行うブレイエル(Breuel)の手法が含まれる。更なる情報については、Breuelによる「Paper to PDA」 Proceedings of IAPR 16th Int'l Conf. on Pattern Recognition, pp. 476-479, 11-15 August, Quebec City, Canada参照。しかし、この手法は、サムネールの作成にはいくつかの理由で問題がある。まず、この手法は、文書をスクロールする必要があるため固定サイズのサムネールを生成するのに用いることができない。スクロールを可能にすることは、垂直方向には制約が何ら課されないということを意味する。よって、全てのコンテンツを示すことができ、構成部分を選択する必要はない。第2に、この手法は、意味論的文書解析の結果を含まない。

【0005】

一方、一部の研究者は、意味論的文書情報を表現することに焦点を当てた。ウッドラフ(Woodruff)は、キーワード情報を位置特定し、大きなポイントサイズでレンダリングされた色コード・テキストをサムネールにおけるキーワード位置にペーストし、キーワードが存在した元の画像セグメントの外観を変える拡張サムネール手法を開発した。しかし、ウッドラフは、キーワード又は、キーワードの周辺の文を含む元の文書画像からカットした画像セグメントの読みやすさの問題を解決していない。更なる情報は、Woodruffによる「Using Thumbnails to Search the Web」 Proceedings of SIGCHI '2001, March 31-April 4, 2001, Seattle, Washingtonを参照。

【0006】

ウッドラフ手法に基づいて、ソー(Suh)は、文書を提示する、概要プラス詳細のインタフェースを紹介した。このインタフェースによれば、ユーザは、概要としてのダウンサンプル文書画像も、キーワードを含む特定部分へのズームインを行う詳細表示としてのポップアップ・ウィンドウも得る。キーワードは、文書概要画像上にもペーストされる。よっ

10

20

30

40

50

て、概要表現から詳細表現への遷移が達成される。文書画像の概要では、ユーザは、詳細ポップアップ表示の助けを借りて、キーワードを含む文書の一部を読み取ることができるので、読みやすさの問題には対処していない。文書の再フォーマットは何ら行われない。スーの手法は、ポップアップが多すぎることによって表示が雑然としてしまい、又、固定サイズのサムネールの作成に適していないという点で問題がある。更なる情報については、Suhによる「Popout Prism: Adding Perceptual Principles to Overview + Detail Document Interfaces」Proceedings of CHI '2002, Minneapolis, MN, April 20-25, 2002を参照。

【0007】

チェン(Chen)他は、要約文を統計モデルに基づいて形成することによって意味情報を文書画像から抽出した。更なる情報については、Chen他による「Extraction of Indicative Summary Sentences from Imaged Documents」Proceedings of ICDAR '97, August 18-20, 1997, Ulm, Germanyを参照。同様に、メーダーレクナー(Maderlechner)他は、情報を文書画像からレイアウト情報に基づいて抽出した。更なる情報については、Maderlechner他による「Information Extraction from Document Images using Attention Based Layout Segmentation」Proceedings of DLIA '99, September 18, 1999, Bangalore, India参照。しかし、これらの研究内容は何れも、サムネールの生成に用いることはできないが、それは、画像の構成部分を制約された表示に再フォーマットする機能を欠いているからである。

【0008】

スマートネール手法と呼ぶ手法がこの問題に対応するものとして開発された。本発明の企業譲受人に譲渡された、「Resolution Sensitive Layout of Document Regions」と題する西暦2003年5月9日付願の、西暦2004年7月29日付公開（公開番号第20040145593号）の米国特許出願第10/435,300号参照。スマートネールは、文書トピックスなどの意味論を無視する、画像ベースの解析を用いる。文書画像における特定の構成部分は、視覚重要度の尺度によって、文書から抽出される何れかの重要な意味情報を組み入れることなく判読可能なサイズに縮小される。

【発明の開示】

【発明が解決しようとする課題】

【0009】

文書検索アプリケーションでは、ユーザは、通常はキーワードの組み合わせである検索を通常、実行依頼し、文書が表示されるのを待つ。情報検索の観点から、ユーザは表示文書について、各々のユーザ検索への適切性の程度に関するすばやい判断を下さなければならない。よって、文書に対する検索情報を組み入れることなく、純粋に画像をベースとしたサムネール生成を行うことは十分なものでない。

【課題を解決するための手段】

【0010】

本明細書及び特許請求の範囲では、文書の視覚化を意味情報に基づいて生成する方法及び装置を開示する。一実施例では、方法は、意味論的重要度情報を複数の文書画像セグメント毎に含む複数の属性を有する複数の画像セグメントを選択する工程と、複数の文書画像セグメント上の複数の位置と、複数の文書画像セグメントに対して行う対象の複数の幾何演算とを複数の属性と、表示の制約とによって判定する工程とを備える。

【発明を実施するための最良の形態】

【0011】

本発明は、以下に表す詳細説明から、かつ本発明の種々の実施例の添付図面からなお十分に分かるものであるが、それらは本発明を特定の実施例に限定するものとして解釈されるべきでなく、説明を行い、理解をするためのものに過ぎない。

【0012】

図式的文書表現（例えば、サムネール）を作成するよう、レイアウトを意味情報に基づいて生成する方法及び装置を説明する。一実施例では、本明細書及び特許請求の範囲記載

10

20

30

40

50

の手法は、文書画像におけるテキストの内容を、文書画像の図式表現を生成するうえで考慮に入れる。よって、本明細書及び特許請求の範囲記載の手法によって、文書の内容の意味論的解析に基づいて文書画像の図式表現を生成することが可能になる。

【0013】

意味情報は意味論的解析から生じるものであり得る。意味論的解析はテキスト・ゾーンに対して行ってもよく、キーワード・ベースの属性抽出に基づくものであり得る。一実施例では、図式表現の合成を促進するよう、テキスト・ゾーンの視覚重要度属性及び意味論的重要度属性を合成するのにメトリックを用いる。

【0014】

以下の開示は、解像度感度の高いセグメント化結果を、入力としての意味論的文書情報と合成することを前提とした2次元表示のレイアウト問題の枠組みを備える。本明細書及び特許請求の範囲記載の手法は文書レイアウトを行って文書セグメントの適切な解像度及び配置を判定する。セグメントは、スケーリング、クロッピング及び、セグメント間の空間関係を変える処理を可能にする制約された表示に配置される。一実施例では、レイアウト・アルゴリズムを調節してテキスト・ゾーンの意味論的に適切な部分を視覚化させることが行われる。一実施例では、レイアウト処理はキーワード依存クロッピングを含めるよう修正される。

一実施例では、図式表現は、特定のユーザ検索に応じて生成される。図式表現のレンダリングは、検索に関係した文書コンテンツを示す。よって、同じ文書に対する別々のユーザ検索によって通常、異なる図式表現が生成されることになる。

【0015】

下記明細書では、多くの詳細を表して本発明を更に徹底的に説明している。しかし、当業者には、本発明をこれらの特定の詳細なしで実施し得ることが明らかとなるものである。別の場合には、周知の構造及び装置を、詳細形式ではなく構成図形式で示し、本発明を不明瞭にすることがないようにしている。

【0016】

以下の詳細説明の一部は、コンピュータ・メモリ内部のデータ・ビットに対する演算のアルゴリズム及び記号表現によって示す。これらのアルゴリズムの記述と表現は、当該データ処理技術分野における当業者が別の当業者に自分達の作業内容の本質を最も効果的に伝えるのに用いる手段である。アルゴリズムは本明細書及び特許請求の範囲においても一般的にも、所望の結果につながる首尾一貫した一連の工程として考案されている。当該工程は、物理的数量を物理的に操作することを必要とするものである。この数量は、記憶、転送、合成、比較やその他の方法による操作ができる電気信号又は磁気信号の形式を通常は呈するものであるが、必ずしもそうでなくてもよい。主に慣用的な理由で、これらの信号をビット、値、要素、記号、文字、項、数字などとして表すことが場合によっては好都合であることが証明されている。

【0017】

これらの語や同様な語は、適切な物理的数量と関連付けられるべきであり、単に、この数量に付される好都合なラベルであることを、しかし、念頭におくべきである。以下の記載から明らかであるように別途明記しない限り、本明細書及び特許請求の範囲を通して、「処理」若しくは「計算」若しくは「算出」若しくは「判定」若しくは「表示」などの語を用いる記載は、コンピュータ・システムのレジスタ内部及びメモリ内部で物理的（電子的）数量として表すデータを操作し、該データを、コンピュータ・システムのメモリ若しくはレジスタや他のそのような情報記憶デバイス、情報伝送デバイス又は情報表示デバイスの内部で物理的数量として同様に表す別のデータに変換するコンピュータ・システムや同様な電子計算機装置の動作及び処理を表す。

【0018】

本発明は、本明細書及び特許請求の範囲記載の動作を行う装置にも関する。この装置は、要求される目的用に特に構築されてよく、コンピュータに記憶されるコンピュータ・プ

プログラムによって選択的に起動又は再構成される汎用コンピュータを備えてもよい。そのようなコンピュータ・プログラムは、フロッピー（登録商標）・ディスク、光ディスク、CD-ROM、及び光磁気ディスクを含む何れかの種類のディスク、読み取り専用メモリ (ROM)、ランダム・アクセス・メモリ (RAM)、EPROM、EEPROM、磁気カード若しくは光カードなどであるがそれらには限定されないコンピュータ判読可能記憶媒体、又は電子命令を記憶させるのに適した何れかの種類の媒体で、各々がコンピュータ・システム・バスに結合されているものに記憶させ得る。

【0019】

本明細書及び特許請求の範囲に示すアルゴリズム及び表示は、何れかの特定のコンピュータや他の装置に固有に関連する訳でない。種々の汎用システムを本明細書及び特許請求の範囲記載の教示によるプログラムとともに用いる場合があり、更に特化した装置を構築して要求方法工程を行うことが好都合であることが証明される場合もある。これらの各種システムに要求される構造は下記明細書から明らかとなるものである。更に本発明は、何れかの特定のプログラミング言語を参照しながら記載されている訳でない。各種プログラミング言語を用いて本明細書及び特許請求の範囲記載の本発明の教示を実施し得ることが分かるものである。

10

【0020】

マシン判読可能媒体は、マシン（例えば、コンピュータ）によって判読可能な形式での情報を記憶又は伝送する何れかの機構を含む。例えば、マシン判読可能媒体は読み取り専用メモリ（「ROM」）；ランダム・アクセス・メモリ（「RAM」）；磁気ディスク記憶媒体；光記憶媒体；フラッシュ・メモリ・デバイス；及び、電気的形式、光学的形式、音響的形式や他の形式の伝搬信号（例えば、搬送波、赤外線信号、デジタル信号等）；などを含む。

20

【実施例】

【0021】

概要

本発明の実施例は意味情報の抽出と画像の再フォーマットとを組み合わせ、文書を表すうえで縮小されたサイズの画像を備える。本明細書及び特許請求の範囲記載の手法は、縮小サイズの画像を生成するうえで意味論的文書解析を用いる。

【0022】

一実施例では、このことは、小型表示装置における文書のブラウジングを拡充させるのに有用である。一実施例では、本明細書及び特許請求の範囲記載の手法の実用的な用法としては、ユーザが検索をタイプ入力すると、意味論的に適切な文書情報（例えば、キーワードが生起するコンテンツ）を含む縮小サイズの画像をシステムが戻すというものがある。一実施例では、本発明は、ユーザの検索タームを取り込み、文書の適切な部分を識別し、検索に対する適切性によって導き出される重要な属性を割り当てる。重要度属性の割り当ては、検索に対する適切性の順序で文書をランク付けすることを可能にするのみならず、文書部分の、スケーリングと、クロッピングと、位置の再設定との処理にペナルティを科すコスト関数を最適化させるレイアウト処理も可能にする。一実施例では、コスト関数の最適化は、本発明の企業譲受人に譲渡された、「Resolution Sensitive Layout of Document Regions」と題する西暦2003年5月9日出願の、西暦2004年7月29日付公開（公開番号第20040145593号）の米国特許出願第10/435,300号参照。

30

40

【0023】

図1は、意味論的文書情報特徴付け（例えば、キーワード特徴付け）及びゾーン得点を用いて画像の表現を生成するシステムの一実施例を示す。各ブロックは、ハードウェア（例えば、回路、専用ロジック等）、（汎用コンピュータ・システム上又は専用マシン上で実行されるものなどの）ソフトウェア、又は両方の組み合わせを備え得る処理ロジックを備える。

【0024】

図1を参照すれば、意味論的解析処理ロジック103は、ユーザ・キーワード検索101及び

50

画像102を受信し、合成ブロック109を制御して画像の出力表現を作成する情報を生成するマージ処理ブロック108が用いる出力を生成する。一実施例では、意味論的解析処理ロジック103は、光学式文字認識を画像102に対して行って光学式文字認識結果を生成するOCR処理ブロック106を含み、光学式文字認識結果は、キーワード特徴付けブロック104及び前合成手順ブロック107に送信される。

【0025】

OCR処理ブロック106の出力に関して、一実施例では、図式表現を意味論的文書情報とともに生成することは、セグメント文書のゾーンへの、行間スペースに基づいた前処理から始まる。一実施例では、このことは、本発明の企業譲受人に譲渡された、「Reformatting Documents Using Document Analysis Information」と題する西暦2003年1月29日付
10
出願の、西暦2004年7月29日付公開（公開番号第20040146199号）の米国特許出願第10/354,811号記載のものと同様に行う。その情報を得るよう、文書は光学式文字認識（OCR）によって処理して、文書構造などの幾何学的文書レイアウト情報や、テキスト認識結果を抽出する。一実施例では、各文字の情報の供給と出力ファイルの作成を、エクスペリビジョン（www.expervision.com、カリフォルニア州フリーモント、クリッパー・コート4109番地）を用いて、次のフォーマットで行う。

【0026】

1)行表示子及びデータが、元の走査文書画像における行の座標を表す。

【0027】

2)ワード表示子と、データが、元の走査文書画像における行中のワードの座標を、ワード
20
がOCR辞書にあるか否かの表示とともに表す。

【0028】

3)文字表示子及びデータが、元の走査画像における行中のワードにおける各文字の座標を、文字が太字であるかイタリック文字であるかの表示、文字のフォント及びサイズの表示、文字のASCII、バウンディング・ボックスにおける文字の数、及び文字認識の信頼度の表示とともに表す。

【0029】

OCR処理を行った後、上記出力ビットマップに記述されているゾーンのテキストに関する解析が行われる。一実施例では、意味論的解析処理ロジック103は、キーワード情報を用いて意味論的解析を行う。キーワードはユーザ検索から得てもよい。キーワードは自動抽出してもよい。キーワードの自動抽出は、Salton, G., Buckley, C.による「Text-Weighting Approaches in Automatic Text Retrieval」Information Processing
30
and Management, vol. 24, pp. 513-523, 1988記載のように行ってよく、Matsuo, Y., Ishizuka, M.による「Keyword Extraction from a Single Document using Word Co-Occurrence Statistical Information」International Journal on Artificial Intelligence Tools, vol. 13, no. 1, pp. 157-169, 2004記載のように行ってもよい。

【0030】

段落情報及び/又はヘッダ・ゾーンは、意味論的解析処理ロジック103によっても用い得る。一実施例では、キー段落が自動抽出される。このことは、Fukumoto, F., Suzuki, Y., Fukumoto, J.による「An Automatic Extraction of Key Paragraphs Based on Context Dependency」Proceedings of Fifth Conference on Applied Natural Language Processing, pp. 91-298, 1997における記載によって行い得る。
40

【0031】

一実施例では、以下の2つの項目をキーワード画像特徴付けに用いる。

【0032】

キーワード・フォント情報（例えば、サイズ、種類、大文字など）

キーワード位置（例えば、ゾーンの始め、ゾーンの終わりなど）

意味論的前マージ処理ロジック103は、上記キーワード属性を、ゾーン毎に得点を計算するのに用いるメトリックに合成する。一実施例では、図式表現を生成するうえで、得点
50

が最高のゾーンが、図式表現への配置が検討される最初のゾーンである。

【0033】

一実施例では、意味論的前マージ処理ロジック103は、ユーザ・キーワード検索101と、OCR処理ブロック106からの出力とを受信する。これらの入力に応じて、キーワード特徴付けブロック104は、文書画像セグメントに存在するキーワードをユーザ・キーワード検索101に基づいて特徴付ける。一実施例では、キーワード特徴付けブロック104は、キーワード特徴付けを、キーワード頻度と、キーワード位置の表示（例えば、キーワード位置表示子 $p(w)$ ）と、キーワード・フォント種類（例えば、キーワード・フォント種類表示子 $t(w)$ ）とのうちの1つ又は複数のものに基づいて行う。

【0034】

一実施例では、キーワード特徴付けブロック104は、OCR処理ブロック106からの出力ファイルに関する解析からのいくつかの属性を抽出する。一実施例では、これらの属性は、キーワード位置表示子 $p(w)$ と、キーワード・フォント種類表示子 $t(w)$ と、ゾーン大文字表示子 $c(w)$ である。キーワード群における各キーワード w は、それに関連付けられる、属性、 $p(w)$ 、 $t(w)$ 及び $c(w)$ を有する。一実施例では、 $p(w)$ の値は、入っているゾーンに関する座標位置に基づいて判定される。

【0035】

キーワードがゾーンの上半分に存在する場合、 $p(w)=1$ であり、さもなければ $p(w)=0$ である。(1)

一実施例では、 $t(w)$ の値は、フォント種類に基づいて判定される。

【0036】

キーワードが太字又はイタリック文字である場合、 $t(w)=1$ であり、さもなければ $t(w)=0$ である。(2)

大文字表示属性 $c(w)$ は、非大文字ゾーンにおける大文字キーワードを区別するよう計算される。

【0037】

ワードが非大文字ゾーンで大文字になっている場合、 $c(w)=1$ であり、さもなければ $c(w)=0$ である。(3)

アラインメント情報を用い得る。「センタリング」などのアラインメント情報は、リフロー手順を行う際に用いるのに記憶される。一実施例では、アラインメント情報は、このゾーンのバウンディング・ボックスの中心位置と比較した、第1テキスト行のバウンディング・ボックスの中心位置をみることによって計算される。

【0038】

取得可能な他の属性で考えられるものとして、ワードのバウンディング・ボックスのサイズ、ワードのポイントサイズなどがある。これらを用いて視覚重要度測定（例えば、視覚重要度値）を導き出すことは周知技術である。例えば、本発明の企業譲受人に譲渡されており、内容を本明細書及び特許請求の範囲に援用する、「Reformatting of Scanned Documents Using Layout Analysis Information」と題する西暦2003年1月29日付出願の、西暦2004年7月29日付公開（公開番号第20040146199号）の米国特許出願第10/354,811号参照。

【0039】

一実施例では、式(1)-(3)に表す3つの属性のみを用いて意味論的文書情報を特徴付けている。

【0040】

キーワード特徴付けブロック104の出力は、ゾーン得点ブロック105に入力される。ゾーン得点ブロック105は、各ゾーンの意味論的重要度を判定する。一実施例では、意味論的重要度値 $I_{sem}(z)$ として本明細書及び特許請求の範囲において表すメトリックが導き出される。一実施例では、得点ブロックは、意味論的重要度値をゾーン毎に以下の式に基づいて生成する。

【0041】

10

20

30

40

50

【数 1】

$$I_{sem}(z) = \sum_{w \in z \cap K} (p(w) + t(w) + 1 + c(w)) \quad (4)$$

ゾーン得点ブロック101は、重要度値 I_{sem} を出力し、そのとき、 z はテキスト・ゾーンであり、 K はキーワード群である。

10

【0042】

一実施例では、キーワードのみならずキー段落を含むゾーンの意味論的重要度値を算出するよう、以下の式を用い得るものであり、

【0043】

【数 2】

$$I_{sem}(z) = \sum_{p \in z \cap P} 1$$

20

そのとき、 p は段落であり、 P は全キー段落群である。

【0044】

一実施例では、各テキスト・ゾーンは、関連した視覚重要度値 $I_{vis}(z)$ を有し、この値は、例えば、ページにおけるゾーンの位置や、ゾーン内のテキストの文字サイズなどの視覚特徴付けにのみ基づいている前合成数107によって計算される。

【0045】

前合成手順ブロック107は、前合成手順をOCR処理106の結果に基づいて行う。一実施例では、前合成手順は、視覚重要度情報及び大文字表示を生成する。一実施例では、大文字表示属性 $C(z)$ を計算して、大文字を含むゾーンを他のゾーンから区別する。このゾーンは、文書の、題名、キャプションや特定の他の有益な記述を通常、有するので、通常は重要である。一実施例では、 $C(z)$ の値は、以下の測定によって判定される。

30

【0046】

ゾーンが大文字の場合 $C(z) = 1$ であり、さもなければ、 $C(z) = 0$ である。 (5)

一実施例では、前合成手順ブロック107の出力は、重要度値 I_{vis} と、 $C(z)$ である。

【0047】

意味論的前マージ処理ロジック103及び前合成手順ブロック107の出力は、マージ処理ブロック108に入力される。これらの入力に応じて、マージ処理ブロック108は重要度値をマージして、合成重要度値を作成する。マージ処理ブロック108の出力は、合成ブロック109に入力される。合成ブロック109は、画像の出力表現を、その入力に基づいて生成する。任意的には、合成ブロック109は、本発明の企業譲受人に譲渡された、「Reformatting Document Using Document Analysis Information」と題する西暦2003年1月29日付願の、西暦2004年7月29日付公開（公開番号第20040146199号）の米国特許出願第10/354,811号によって、画像ベース解析及びレイアウトの情報とのマージを行い得るものであり、かつ、画像作成を行う。

40

【0048】

図2は、前合成手順ブロック107の一実施例を示す。図2におけるブロックの各々は、ハードウェア（例えば、回路、専用ロジック等）、（汎用コンピュータ・システム上又は専用マシン上で実行されるものなどの）ソフトウェア、又は両方の組み合わせを備え得る処理ロジックによって行う。

50

【0049】

図2を参照すれば、前合成手順ブロック107は、画像102を受信し、画像の一部分（例えば、テキスト）に関する属性を抽出する、データの抽出及び作成のブロック201を含む。

【0050】

データの抽出及び作成のブロック201の出力に応じて、セグメント化装置202は、画像データをゾーンにセグメント化し、それらはゾーン評価装置202に出力される。ゾーン評価装置203は、セグメント化装置202によって生成されるゾーン各々を評価する。一実施例では、ゾーン評価装置203は、視覚重要度情報値（例えば、視覚重要度値 I_{vis} ）及び大文字表示値 $c(z)$ を出力し、それらは、前合成手順107の出力を表す。一実施例では、図2に関して説明した処理を、「Reformatting Documents Using Document Analysis Information」と題する西暦2003年1月29日付特許願の、西暦2004年7月29日付公開（公開番号第20040146199号）の米国特許出願第10/354,811号記載の手法によって行う。

【0051】

テキスト・ゾーン毎の重要度値と、意味論的重要度値とを生成した後、ゾーンの重要度値 $I(z)$ が、 $I_{vis}(z)$ と $I_{sem}(z)$ との結合を用いて生成される。一実施例では、単純な、 I_{vis} と I_{sem} との線形結合を

【0052】

【数3】

$$I(z) = \left[\alpha x \frac{I_{vis}(z)}{\max_{Z \in D} \{I_{vis}(A)\}} + (1 - \alpha) x \frac{I_{sem}(z)}{\max_{Z \in D} \{I_{sem}(Z)\}} \right] \cdot \max_{Z \in D} \{I_{vis}(z)\} \quad (5) \quad 20$$

として計算し、そのとき、 α は0と1との間のスカラーである。

【0053】

一実施例では、 α は、アプリケーションによって設定されるか、ユーザによって設定される。別の実施例では、 α はデータから導き出される。例えば、

【0054】

【数4】

$$\alpha = \frac{\sigma(I_{vis})}{\max_{z \in D} \{I_{vis}(z)\}}$$

であり、

【0055】

【数5】

$$\alpha(I_{vis}) = \sqrt{1/(|D|-1) \cdot \sum (I_{vis}(z) - \bar{I}_{vis})^2}, \quad 40$$

であり、そのとき、 $|D|$ は D におけるゾーン数であり、

【0056】

【数 6】

$$\overline{I_{vis}} = \frac{1}{|D|} \cdot \sum_{z \in D} I_{vis}(z)$$

である。

この場合、全てのゾーンが同様な視覚重要度値 I_{vis} を有する場合、標準偏差 $\sigma(I_{vis})$ はゾーンに近く、すなわち、 $\alpha = 0$ であり、意味論的重要度値 I_{sem} によって捕捉されるキーワードの存在は、合成重要度尺度において支配的な部分である。対照的に、ゾーンの存在に多くの変化がある場合、すなわち、 I_{vis} 値群の標準偏差が高い場合、 $\alpha = 1$ であり、意味論的重要度値 I_{sem} の重要度は低い。

【0057】

別の実施例では、 α は、論理的レイアウト解析情報などの他の意味論的文書情報を組み入れるように選ばれる。 α は、ゾーンが文書における題名か、普通の段落かによって変わってくる場合がある。題名ゾーンの場合、 α は1に設定してよい、すなわち、題名ゾーンの重要度はキーワード生起とは無関係である。普通の段落の場合、 α は、意味論的重要度が視覚重要度に対して支配的であるように0に設定し得る。

【0058】

図3は、文書解析情報を用いて文書を再フォーマットする処理ロジックの一実施例を示す。図3を参照すれば、意味論的前マージ処理ロジック103から出力される視覚重要度値及び意味論的重要度値をマージするマージ処理ブロック108を示す。一実施例では、マージ処理ブロック108は、視覚重要度値 $I_{vis}(z)$ と意味論的重要度値 $I_{sem}(z)$ と大文字情報 $C(z)$ とを、

【0059】

【数 7】

$$I(z) = \left[\alpha \times \frac{I_{vis}(z)}{\max_{z \in D} \{I_{vis}(z)\}} + (1-\alpha) \times \frac{I_{sem}(z) + C(z) \times \frac{I_{vis}(z)}{\max_{z \in D} \{I_{vis}(z)\}}}{\max_{z \in D} \left\{ I_{sem}(z) + C(z) \times \frac{I_{vis}(z)}{\max_{z' \in D} \{I_{vis}(z')\}} \right\}} \right] \cdot \max_{z \in D} \{I_{vis}(z)\}$$

によってマージする。

【0060】

マージ処理ブロック108の出力は、合成ブロック109に入力され、このブロックは画像の出力表現を生成する。

【0061】

あるいは、 $I_{vis}(z)$ と $I_{sem}(z)$ とを合成する他の方法を用い得る。一実施例では、 $I_{vis}(z)$ と $I_{sem}(z)$ は乗算される。この手法の欠点の1つとしては、値の一方が0の場合に他方の情報が失われるというものがある。

【0062】

前述のように、一実施例では、重要度値 $\{I(z) \mid z \in D\}$ 全てが算出され、重要度値が最高のゾーンを、レイアウト手順によって図式表現に入れられるものとみなされる。図式表現

はこの場合、（ $\alpha < 1$ の場合、）キーワード検索に基づいた特定の意味論的文書情報を含み得る。

【0063】

ゾーン・レイアウト処理

一実施例では、マージ処理ブロック108の出力は、キーワード情報を含む出力ファイルである。マージ処理ブロック108の出力は、ストリーミング出力であり得る。一実施例では、この出力ファイルは、本発明の企業譲受人に譲渡された、「Reformatting Documents Using Document Analysis Information」と題する西暦2004年1月29日付願の、西暦2004年7月29日付公開（公開番号第20040146199号）の米国特許出願公開第10/354,811号記載の画像解析からのセグメント記述とマージする対象の、マージ処理ロジック（例えば、マージ・プログラム）への入力である。出力ファイルは、（例えば、意味論的図式表現を生成する）レイアウト・プログラム合成ブロック109への入力である。一実施例では、出力ファイルの第1行は、

幅、高さ、エントリ数、ゾーン数、最大フラグメント数、
として定義され、そのとき、「幅」及び「高さ」は、元の走査文書画像の寸法である。ゾーン数は、テキスト・ゾーン数も画像ゾーン数も含む。フラグメントは通常、テキスト・ゾーンにおけるワードである。一実施例では、各エントリは、以下の規定に従う。

【0064】

列1：ゾーンの索引

列2：ゾーンにおける個々のワードの索引

列3：ダウンサンプルされたスカラー

列4：leftX

列5：upperY

列6：xの寸法

列7：yの寸法

列8：重要度値Iを100で乗算したもの

列9：マーク（1-テキスト、0-画像）

列10：アラインメント・フラッグ

列11：ゾーンにおけるフラグメントの数

列12：キーワード・ポインタ

上記出力ファイルでは、「キーワード・ポインタ」は行における第1ワードの索引によって特徴付けられ、キーワードはまず、元の走査文書画像に存在する。キーワードでない場合、この属性は0に設定される。

【0065】

リフロー処理中には、重要度値Iの順に図式表現を収めるよう候補ゾーンを選ぶ。利用可能なものとして存在する空間に応じて、ワードのクロッピングが必要になり得る。しかし、キーワードがゾーンの始めに生起しない場合、利用可能なものとして存在する空白空間に対する限度のために、キーワードがカットされ得る。このことは意味情報の喪失につながり得る。一実施例では、この欠点を解決するよう、ゾーン内でキーワードが最初に生起する行からゾーンが満たされる。別の実施例では、キーワードが最初に生起する文の始めから収めることによってクロッピングを行う。

【0066】

グレイアウトされたテキストを用いて、不適切なテキストを記号で表すか、キーワードを強調するなど、レイアウトを美しくするレイアウト手法として考えられる特定の他のものを採用し得る。これによって、適切なテキストに対して、より多くの空間を確保し得る。図7は、キーワード依存レイアウトである。図7を参照すれば、元のゾーンにおける行の始めは、文の始めと、キーワードとともに示し、それらは異なる色である。

【0067】

例示的な処理

図4Aは、画像セグメントの再フォーマットングを行う処理の一実施例の流れ図である

。当該処理は、ハードウェア（例えば、回路、専用ロジック等）、（汎用コンピュータ・システム上又は専用マシン上で実行されるものなどの）ソフトウェア、又は両方の組み合わせを備え得る処理ロジックによって行う。

【0068】

図4Aを参照すれば、当該処理は、処理ロジックが文書を文書画像セグメントに分割することによって始まる（処理ブロック401）。一実施例では、セグメント化は、本発明の企業譲受人に譲渡された、「Reformatting Documents Using Document Analysis Information」と題する西暦2003年1月29日付願の、西暦2004年7月29日付公開（公開番号第20040146199号、特に段落番号54、55及び75-78）の米国特許出願第10/354,811号記載のものなどのOCRレイアウト解析結果（例えば、行間スペース）に基づいている。

10

【0069】

文書画像セグメントを用いて、処理ロジックは、意味論的文書解析を文書の文書画像セグメント群に対して行って意味情報を文書から抽出する（処理ブロック402）。一実施例では、キーワード・リストがあれば、処理ロジックは、キーワード属性を文書画像セグメントの各々から抽出することによって意味論的文書解析を行う。一実施例では、処理ロジックは、画像セグメント各々に関するキーワード特徴付けを行うことによってキーワード属性を各文書画像セグメントから抽出する。

【0070】

図4Bは、意味論的文書解析を文書の文書画像セグメント群に対して行って意味情報を文書から抽出する処理の一実施例の流れ図である。当該処理は、ハードウェア（例えば、回路、専用ロジック等）、（汎用コンピュータ・システム上又は専用マシン上で実行されるものなどの）ソフトウェア、又は両方の組み合わせを備え得る処理ロジックによって行う。

20

【0071】

図4Bを参照すれば、当該処理は、処理ロジックが、キーワード・フォント情報（例えば、キーワードのフォントサイズ、キーワードのフォント種類、キーワードが大文字か否かなど）と、キーワード位置情報（例えば、キーワードは画像セグメントの始めに位置しているか画像セグメントの終わりに位置しているかなど）とのうちの1つ又は複数のものに基づいてキーワード特徴付けを行うことによって始まる（処理ブロック411）。一実施例では、キーワード属性は、キーワード位置表示子、キーワード・フォント種類表示子及び画像セグメント大文字表示子を備え、ワードのバウンディング・ボックスのサイズと、ワードのポイントサイズとを含み得る。一実施例では、処理ロジックは、画像セグメント毎に大文字表示子を生成する。

30

【0072】

キーワード属性抽出の後、処理ロジックは、意味論的重要度情報を生成する（処理ブロック412）。一実施例では、処理ロジックは、文書に対する情報を生成し、各画像セグメントは文書における意味論的重要度を示す。一実施例では、意味論的重要度は、キーワード検索に基づいている。一実施例では、意味論的重要度を示す情報は、意味論的重要度値である。

【0073】

処理ロジックは、視覚重要度情報の文書画像セグメント毎の生成も行う（処理ブロック413）。一実施例では、意味論的重要度情報は意味論的重要度値を備え、視覚重要度情報は視覚重要度値を備える。

40

【0074】

意味論的重要度情報及び視覚重要度情報を生成した後、処理ロジックは、意味論的重要度情報及び視覚重要度情報を、大文字表示情報とともに文書画像セグメント毎にマージして、マージ重要度情報を画像セグメント毎に生成する（処理ブロック414）。

【0075】

図4Aをもう一度参照すれば、抽出された意味情報に基づいて、処理ロジックは、目標キャンバスの、表示の制約を満たすよう、抽出された意味情報に関する画像部分の再フォ

50

ーマッティングを行う（処理ブロック403）。

【0076】

図4Cは、抽出された意味論的情報に基づいて画像の一部分の再フォーマッティングを行う処理の一実施例の流れ図である。当該処理は、ハードウェア（例えば、回路、専用ロジック等）、（汎用コンピュータ・システム上又は専用マシン上で実行されるものなどの）ソフトウェア、又は両方の組み合わせを備え得る処理ロジックによって行う。

【0077】

図4Cを参照すれば、処理は、処理ロジックが、文書の文書画像セグメント群から文書画像セグメントをマージ重要度情報に基づいて選択することによって始まる（処理ブロック421）。一実施例では、マージ重要度情報が意味論的重要度値に基づいているので、選択は意味論的重要度値に基づいている。別の実施例では、マージ重要度情報は、意味論的テキスト情報と、視覚テキスト情報と、視覚画像情報とに基づいている。

10

【0078】

次に、処理ロジックは、文書画像セグメント上の位置と、文書画像セグメントに対して行う対象の幾何演算とを、属性と、表示の制約とによって判定する（処理ブロック422）。

【0079】

位置情報及び幾何演算情報を用いて、処理ロジックは、絶対位置領域又は相対位置領域を目標表示上の絶対位置に変換し（処理ブロック423）、文書を処理する命令の出力リストを生成する（処理ブロック424）。

20

【0080】

最後に、処理ロジックは、画像セグメントに対する、クロッピング、スケーリングやペーストなどの処理を行い（処理ブロック425）、処理された画像セグメントを目標キャンバス上に表示する（処理ブロック426）。

【0081】

例示的なシステム

図5は、画像解析及びレイアウトのシステムの例示的な実施例を示す。システムにおける構成部分は、ハードウェア（例えば、回路、専用ロジック等）、（汎用コンピュータ・システム上又は専用マシン上で実行されるものなどの）ソフトウェア、又は両方の組み合わせを備え得る処理ロジックによって実施される。

30

【0082】

図5を参照すれば、システムは、レイアウト解析及び光学式文字認識511を画像501に対して行う。画像501は、JPEG2000符号化512にも備える。一実施例では、JPEG2000符号化512はJPEG2000符号化を2つの階層で行う。

【0083】

処理ロジックは更に、レイアウト解析及び光学式文字認識511からの出力と、JPEG2000復号化512からの結果とに対するデータ作成を行う（処理ブロック520）。データ作成は、キーワード特徴付け（104）、光学式文字認識及びレイアウト解析の情報の抽出（処理ブロック521）及びJPEG2000パケットヘッダ・データの抽出（処理ブロック522）を含む。これを行う方法の例は、各々、公に入手可能であり、内容を本明細書及び特許請求の範囲に援用し、本発明の譲受人に譲渡された、「Reformatting Documents Using Document Analysis Information」、「Content and Display Device Dependent creation of Smaller Representations of Images」及び「Header-Based Processing of Images Compressed Using Multi-Scale Transforms」と題する米国特許出願（米国特許出願番号第10/354,811号、第10/044,603号、第10/044,420号）で見つけ得る。処理ロジックは、光学式文字認識及びレイアウト解析の情報を光学式文字認識及びレイアウト解析の結果から抽出し、JPEG2000パケット・ヘッダ・データをJPEG2000符号化の結果から抽出する。

40

【0084】

処理ロジックは更に、データ作成からの結果を、連結した構成部分にセグメント化する

50

(処理ブロック530)。テキスト・ゾーンのセグメント化は、光学式文字認識及びレイアウト解析の情報の抽出からのデータに対して行い(処理ブロック531)、セグメント化を、抽出からのJPEG2000ヘッダ・データに対して行う(処理ブロック532)。更に、処理ロジックは、セグメント化処理からの各構成部分に属性を割り当てる(処理ブロック540)。最小限のスケーリング、テキスト・ゾーン及びフラグメントの重要度が、テキスト・ゾーンのセグメント化からの構成部分に割り当てられる(処理ブロック541)。最適な解像度と、最適な解像度でのビット数が、JPEG2000ヘッダ・データに対するセグメント化からの構成部分に割り当てられる(処理ブロック542)。

【0085】

処理ロジックは更に、視覚情報及び意味情報をマージする(処理ブロック543)。その後、属性を備えている構成部分をマージして入力リストを生成する(処理ブロック550)。

【0086】

検索依存レイアウト処理を入力リストに対して行い、その結果は解像度及び領域復号化に供給される(処理ブロック560)。2階層におけるJPEG2000符号化は更に、画像データ502を処理ブロック512で生成する。処理ロジックは、画像データ502と、解像度感度が高いレイアウト560からの出力とに対して解像度及び領域の復号化を処理ブロック570で行う。解像度及び領域の復号化570からの結果と、キャンバス503のサイズに関する情報とを用いて、処理ロジックは、最終のクロッピング及びスケーリングの処理を画像セグメントに対して行う(処理ブロック580)。

【0087】

処理ロジックは更に、クロッピングされ、スケーリングされたセグメントをキャンバス上にペーストして、画像コンテンツ及び表示装置によって変わってくる画像表現を生成する(処理ブロック590)。

【0088】

一実施例では、キーワード依存レイアウトは、検索に依存することなく生成される。キーワードは、自動キーワード抽出を用いて取得し得る。レイアウト例を図7に示す。

【0089】

例示的なコンピュータ・システム

図6は、本明細書及び特許請求の範囲記載の1つ又は複数の手順を行い得る例示的なコンピュータ・システムの構成図である。図6を参照すれば、コンピュータ・システム600は、例示的なクライアント・サーバ・コンピュータ・システムを備え得る。コンピュータ・システム600は、情報を伝達する通信機構すなわちバス611と、バス611に結合されて情報を処理するプロセッサ612とを備える。プロセッサ612は、例えば、ペンティウム(商標)、パワーPC(商標)、アルファ(商標)などのマイクロプロセッサを含むが、それらなどのマイクロプロセッサに限定されるものでない。

【0090】

システム600は、バス611に結合されて、情報及び、プロセッサ612が実行する対象の命令を記憶するランダム・アクセス・メモリ(RAM)や他の(主メモリとして呼ばれる)動的な記憶装置604を更に備える。主メモリ604は、プロセッサ612が命令を実行する間に一時変数や他の中間的情報を記憶するのにも用い得る。

【0091】

コンピュータ・システム600は更に、バス611に結合されて、静的情報とプロセッサ612用命令とを記憶する読み取り専用メモリ(ROM)及び/又は他の静的な記憶装置606と、磁気ディスク又は光ディスク、及び対応ディスク・ドライブなどのデータ記憶装置607とを備える。データ記憶装置607は、バス611に結合されて、情報及び命令を記憶する。

【0092】

コンピュータ・システム600は、バス611に結合されて、情報をコンピュータ・ユーザに向けて表示する、陰極線管(CRT)表示や液晶表示(LCD)などの表示装置621に更に結合されてもよい。英数字キーや他のキーを含む英数字入力装置622も、バス611に結合

されて、情報及びコマンドの選択をプロセッサ612に伝達し得る。別のユーザ入力装置としては、バス611に結合されて、方向情報及びコマンド選択をプロセッサ612に伝達し、表示621上のカーソルの動きを制御する、マウス、トラックボール、トラックパッド、スタイラスや、カーソル方向キーなどのカーソル制御部623がある。

【0093】

バス611に結合し得る別の装置としては、紙上、フィルム上や他の同様な種類の媒体上などの媒体上に命令、データや他の情報を印刷するのに用い得るハードコピー装置624がある。更に、スピーカ及び/又はマイクロフォンなどの音響記録及び再生装置を任意的に、バス611に結合して、コンピュータ・システム600とのオーディオ・インタフェースを行い得る。バス611に結合し得る別の装置としては、電話機やハンドヘルド型パーム機器に通信する無線/有線通信機構625がある。

10

【0094】

なお、システム600と、関連したハードウェアとの何れか又は全部の構成部分を本発明において用い得る。しかし、他のコンピュータ・システム構成が当該装置の一部又は全部を含み得るということは分かり得るものである。

【0095】

アプリケーション

意味論的文書図式表現を多くの文書検索システムにおいて用い得る。そのような場合には、ユーザは、サーチ検索に直接関係したサムネール（や他の図式表現を）検索の結果として得る。別の用法として考えられるものには、それらを文書管理システム（例えば、リ

20

【0096】

代替的な実施例

複数ページの文書を取り扱う場合、本明細書及び特許請求の範囲記載の意味論的手法を用いてサムネールを単一ページ毎に生成することは適切でない。この文書内のページ間の関係を示すことが重要である。この場合、本明細書及び特許請求の範囲記載の手法を拡張してこの目標を達成し得る。

【0097】

一実施例では、文書Dがnページ $P_1, P_2, \dots, P_n$ を備えていることとする。

$Z = \{z_1^*, z_2^*, \dots, z_n^*\}$ は、

30

【0098】

【数8】

$$(z_1^*, z_2^*, \dots, z_n^*) = \arg \max_{i, j \in \{1, \dots, n\}, z_j \in P_j, z_i \in P_i} \left(\prod_{i=1}^n I(z_i) \times \prod_{i, j=1, i \neq j}^n \text{sim}(z_i, z_j) \right) \quad (6)$$

の条件を満たすゾーン群とし、そのとき、 $I(z)$ は上記式で定義される情報尺度であり、 $\text{sim}(z_i, z_j)$ はゾーン z_i と z_j との間の類似度測定である。2つのゾーン間の類似度を計算するアルゴリズムには多くの種々のものがある（例えば、Nグラム手法。「Gauging Similarity with N-Grams: Language-Independent Categorization of Text」 Science, Vol. 267, pp. 843-848, February 10, 1995.）。

40

【0099】

一実施例では、一文書内のページ群を前提とすれば、当該処理は次の通り、進む。

【0100】

- 1) ループ：上記式(6)の条件を満たすゾーン群Zを見つける
- 2) Zにおける当該ゾーンをPにおけるページから取り除く
- 3) 何れのページにもゾーンがもうない状態になるまでループを続ける

50

よって、特定ページにおける各ゾーンは、式(6)における条件によって選ばれる順序で得点が割り当てられる。最初に選ばれるゾーンの得点は後で選ばれるものの得点よりも高い。

図式表現が個々のページ毎に生成される際に、現行のI(z)値を置き換えるよう得点が考慮される。

【0101】

図8は、ページ毎に無関係に作成される個々のページから構成される表現の2つの例を示す。図9は、ページ間のゾーン類似度によって変わってくる、個々のページの表現の2つの例を示す。図10は、キーワード検索の入力によって、従来技術のスマートネール処理又は、本発明による意味論的文書画像処理を介して、画像出力を行う4つの例を示す。

10

【0102】

本発明の多くの改変及び修正が確かに、当業者に本明細書を検討した後に明らかになる一方、例証の目的で表し、説明した特定の実施例は何れも、限定的なものと解釈されることを何ら意図していないものとする。したがって、種々の実施例の詳細への参照は、それ自体において本発明に本質的と思われる特徴のみを列挙する、その本特許請求の範囲記載の範囲を限定することを意図するものでないものとする。

【図面の簡単な説明】

【0103】

【図1】キーワード特徴付け及びゾーン得点を用いて画像の表現を生成するシステムの一実施例を示す図である。

20

【図2】幾何解析前合成手順ブロック107の一実施例を示す図である。

【図3】意味論的文書解析情報及び幾何学的文書解析情報を用いて文書の再フォーマットを行う処理ロジックの一実施例を示す図である。

【図4A】文書の特定画像セグメントの図式表現を生成する処理の一実施例を示す流れ図である。

【図4B】文書の特定画像セグメントの図式表現を生成する処理の一実施例を示す別の流れ図である。

【図4C】文書の特定画像セグメントの図式表現を生成する処理の一実施例を示す更に別の流れ図である。

【図5】画像解析及びレイアウトのシステムの例示的な実施例を示す図である。

30

【図6】本明細書及び特許請求の範囲記載の1つ又は複数の処理を行い得る例示的なコンピュータ・システムを示す構成図である。

【図7】キーワード依存レイアウトの例を示す図である。

【図8】ページ毎に無関係に作成される個々のページから構成される表現の2つの例を示す図である。

【図9】ページ間のゾーン類似度に応じた、個々のページの表現の2つの例を示す図である。

【図10】キーワード検索の入力によって、従来技術のスマートネール処理又は、本発明による意味論的文書画像処理を介して画像出力を行う4つの例を示す図である。

【符号の説明】

40

【0104】

100 システム

101 ユーザ・キーワード検索

102 画像

103 意味論的解析処理ロジック

104 キーワード特徴付け

105 ゾーン得点

106 OCR処理

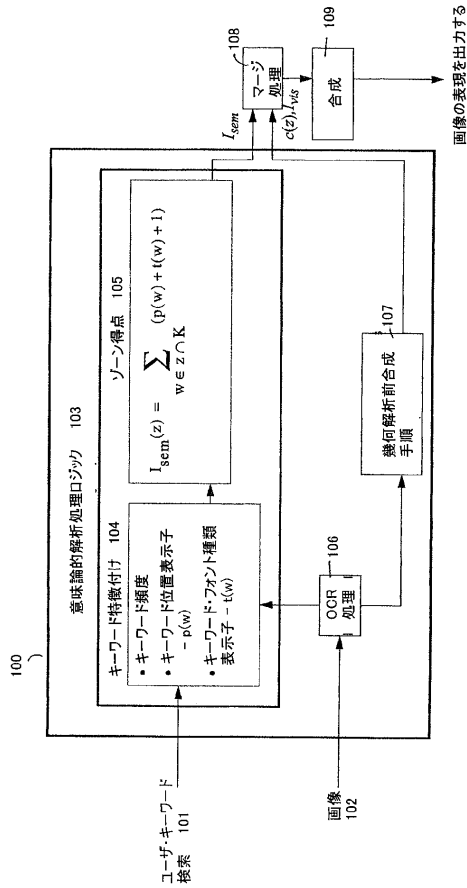
107 幾何解析前合成手順

108 マージ処理

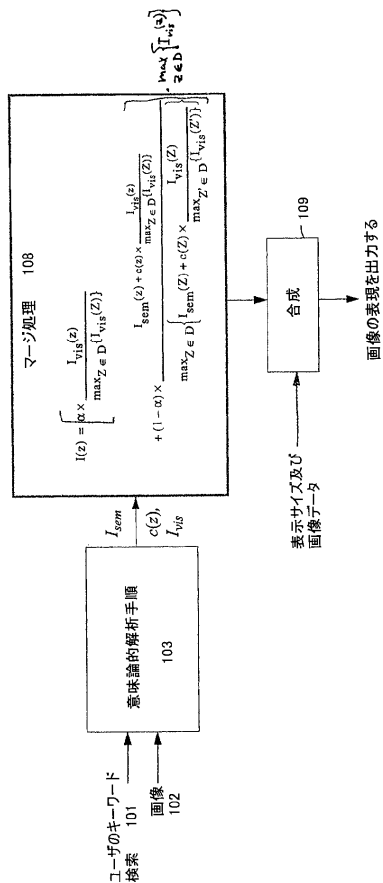
50

109	合成	
201	データの抽出及び作成	
202	連結されたゾーンへのセグメント化	
203	ゾーン評価	
501	画像	
502	画像データ	
503	キャンバス	
510	文書管理システム	
511	OCR解析	
512	JPEG2000	10
520	データ作成	
521	レイアウト抽出	
522	JPEG2000ヘッダ・データ	
530	セグメント	
531	テキスト・セグメント	
532	画像セグメント	
540	セグメントの属性	
541	テキスト属性	
542	画像属性	
550	属性を備えている構成部分のマージ	20
560	検索依存レイアウト	
570	実画像の抽出	
580	クロッピング及びスケーリング	
590	キャンバスへのペースト	
600	コンピュータ・システム	
604	主メモリ	
606	静的なメモリ	
607	大容量記憶メモリ	
611	バス	
612	プロセッサ	30
620	外部ネットワーク・インタフェース	
621	表示	
622	キーボード	
623	カーソル制御装置	
624	ハードコピー装置	

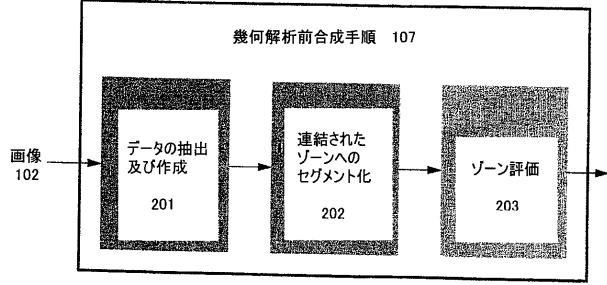
【図 1】



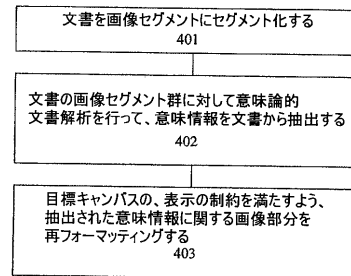
【図 3】



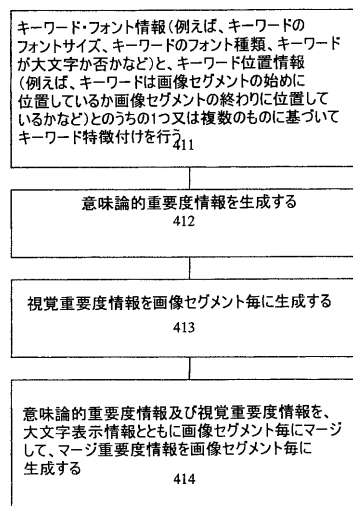
【図 2】



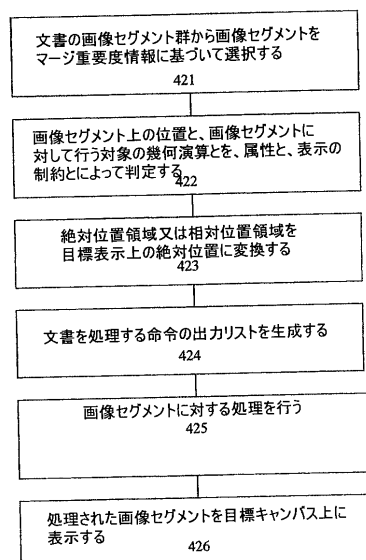
【図 4 A】



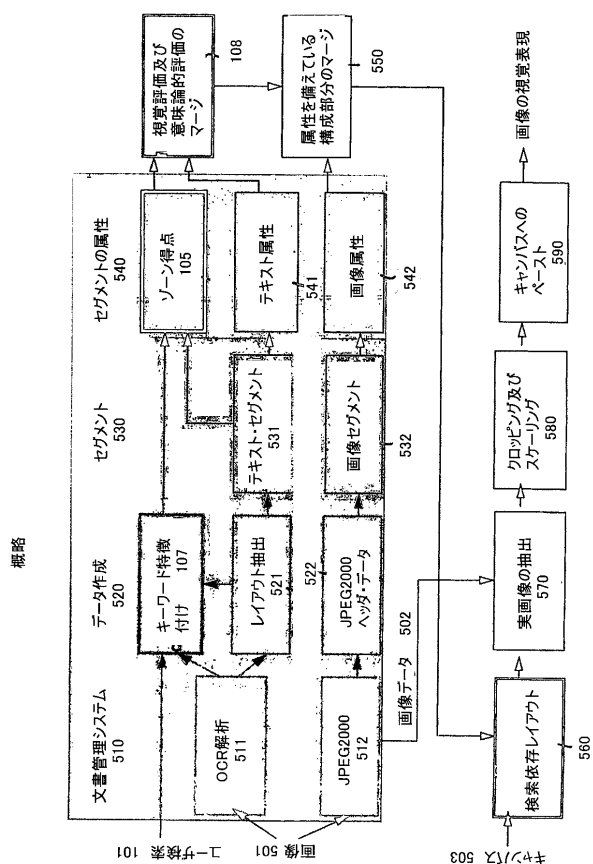
【图 4 B】



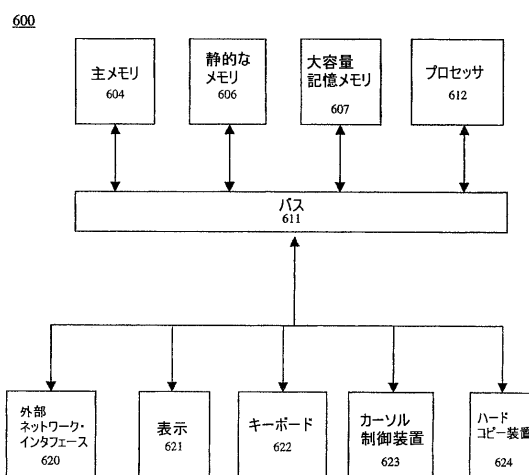
【図 4 C】



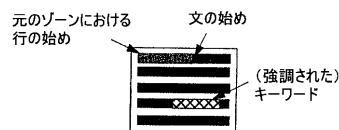
【图 5】



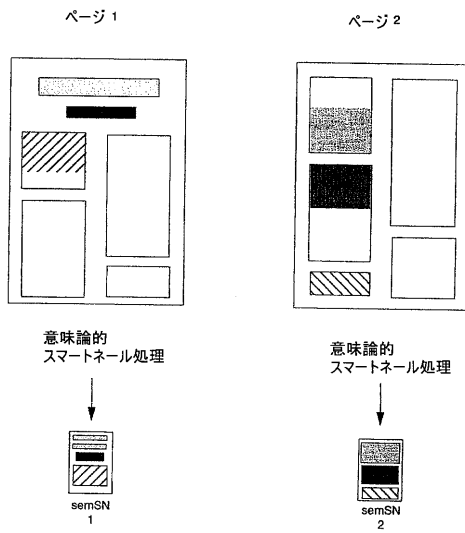
【图 6】



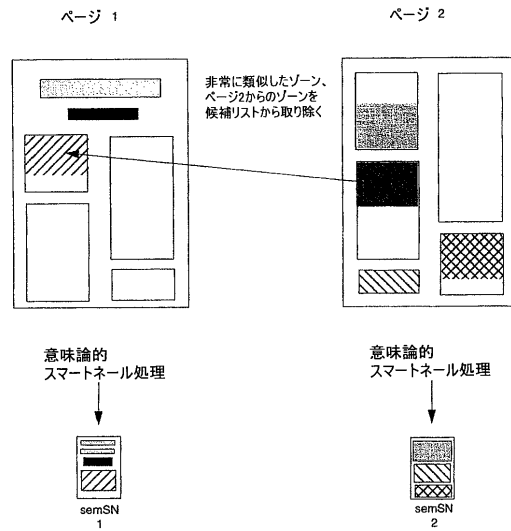
【图 7】



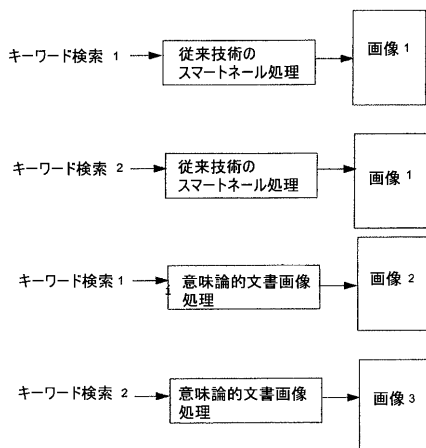
【図 8】



【図 9】



【図 10】



フロントページの続き

Fターム(参考) 5B050 BA06 BA16 EA12 EA19 FA02
5B075 ND06 ND07 PQ48 UU40