

(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2004-185343
(P2004-185343A)

(43) 公開日 平成16年7月2日(2004.7.2)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
G06F 17/30	G06F 17/30 170B	5B029
G06K 9/20	G06F 17/30 380D	5B050
G06T 1/00	G06K 9/20 340L	5B075
	G06T 1/00 200D	

審査請求 未請求 請求項の数 4 O L (全 12 頁)

(21) 出願番号	特願2002-351921 (P2002-351921)	(71) 出願人	000001007 キヤノン株式会社 東京都大田区下丸子3丁目30番2号
(22) 出願日	平成14年12月4日 (2002.12.4)	(74) 代理人	100086818 弁理士 高梨 幸雄
		(72) 発明者	後藤 裕典 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
		Fターム(参考)	5B029 CC29 5B050 AA08 BA10 FA02 FA05 FA13 GA08 5B075 ND06 NK02 NK37 PP03 PP13 PQ02 PQ32

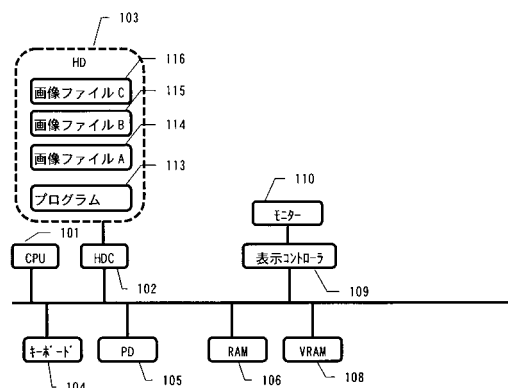
(54) 【発明の名称】 ファイル閲覧装置、ファイル閲覧方法、ファイル閲覧プログラムおよび記憶媒体

(57) 【要約】

【課題】 画像ファイルの内容から、その内容を表わす重要な単語を抽出し、表示することができるファイル閲覧装置、方法、プログラム、及び記憶媒体を実現することにより、ファイル閲覧プログラムから複数の画像ファイルの内容を確認することを容易にする。

【解決手段】 映像観察手段、入力手段、演算手段、該演算手段の使用状況を測定する測定手段、記憶手段、画像データから文字領域を分離する領域分離、文字を認識する文字認識手段、文字認識手段で得られた文字データからキーワードを抽出するキーワード抽出手段、抽出されたキーワードと画像ファイルの最終更新日時を保存する保存手段、抽出されたキーワードを表示するキーワード表示手段を持ったファイル閲覧装置、方法、プログラム及び記憶媒体で構成される。さらに、選択された複数の画像ファイルの内容から抽出した単語を表示する手段を設けた。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

映像観察手段、入力手段、演算手段、該演算手段の使用状況を測定する測定手段、記憶手段、画像ファイルの画像データから文字領域を分離する領域分離手段、前記領域分離手段で分離した文字領域から文字を切り出し、その文字を認識する文字認識手段、文字認識手段で得られた文字データからキーワードを抽出するキーワード抽出手段、前記キーワード抽出手段で抽出されたキーワードと画像ファイルの最終更新日時を保存する保存手段、前記キーワード抽出手段で抽出されたキーワードを表示するキーワード表示手段を有するファイル閲覧装置において、入力手段によって複数の画像ファイルが選択されたときに、表示手段にて、各々の画像ファイルから抽出されたキーワードを同時に表示することを特徴とするファイル閲覧装置 10

【請求項 2】

前記表示手段で表示された複数の画像ファイルの内容から抽出されたキーワードを、ファイル名順に並び替えるファイル名順並び替え手段を有することを特徴とする請求項 1 に記載のファイル閲覧装置

【請求項 3】

前記表示手段で表示された複数の画像ファイルの内容から抽出されたキーワードを、キーワード抽出日時順に並び替えるキーワード抽出日時順並び替え手段を有することを特徴とする請求項 1 または 2 にファイル閲覧装置

【請求項 4】

前記表示手段で複数の画像ファイルの内容から抽出されたキーワードを表示する際に、通常の並び順をファイル名順にするかキーワード抽出日時順にするか設定する設定手段を有することを特徴とする請求項 1 から 3 の何れかにファイル閲覧装置 20

【発明の詳細な説明】**【0001】****【発明の属する技術分野】**

本発明は、ファイル閲覧装置、ファイル閲覧方法、ファイル閲覧プログラムおよび記憶媒体に関するものである。

【0002】**【従来の技術】**

従来、ファイル閲覧アプリケーションは、Microsoft社製のエクスプローラなど存在している。それらが表示するファイルに関する情報は、ファイル名、ファイルのパス（ファイルシステム内の格納されている位置）、作成日、更新日、アクセス日等である。

【0003】

また、例えば、特開平10-247156で提案されているように、イメージファイルのイメージを縮小した形で表示したり、または画像ファイルの文章全体をイメージとして捉えそれを縮小した形で表示したりする方法があった。

【0004】**【発明が解決しようとする課題】**

しかしながら、上記従来例では、画像ファイルのファイル名が的確に内容を表わす名前である場合はファイル名から内容を確認することが容易であるが、しかし、ファイル名が的確に内容を表わさない場合、例えば、image1.gif、photo1.jpgなど内容に関係なくファイル名が付けられていたり、あるいはS-P-1.gifのようなファイルを作成した人にしかわからない省略形でファイル名を付けられたりする場合などはファイル名から内容を確認することは難しかった。 40

【0005】

また、特開平10-247156で提案されているような手段では、画像ファイルを縮小した形で表示するため、その表示は小さく、わかり辛かった。

【0006】

そこで、本発明は、画像ファイルからキーワードを抽出することができるファイル閲覧装 50

置、ファイル閲覧方法、ファイル閲覧プログラムおよび記憶媒体を提供することを目的とする。

【0007】

【課題を解決するための手段】

上記目的を達成するために、本発明の請求項1に記載のファイル閲覧装置は、映像観察手段、入力手段、演算手段、該演算手段の使用状況を測定する測定手段、記憶手段、画像ファイルの画像データから文字領域を分離する領域分離手段、前記領域分離手段で分離した文字領域から文字を切り出し、その文字を認識する文字認識手段、文字認識手段で得られた文字データからキーワードを抽出するキーワード抽出手段、前記キーワード抽出手段で抽出されたキーワードと画像ファイルの最終更新日時を保存する保存手段、前記キーワード抽出手段で抽出されたキーワードを表示するキーワード表示手段を有するファイル閲覧装置において、入力手段によって複数の画像ファイルが選択されたときに、表示手段にて、各々の画像ファイルから抽出されたキーワードを同時に表示することを特徴とする。

10

【0008】

上記目的を達成するために、本発明の請求項2に記載のファイル閲覧装置は、前期表示手段で表示された複数の画像ファイルの内容から抽出されたキーワードを、ファイル名順に並び替えるファイル名順並び替え手段を有することを特徴とする。

【0009】

上記目的を達成するために、本発明の請求項3に記載のファイル閲覧装置は、前期表示手段で表示された複数の画像ファイルの内容から抽出されたキーワードを、キーワード抽出日時順に並び替えるキーワード抽出日時順並び替え手段を有することを特徴とする。

20

【0010】

上記目的を達成するために、本発明の請求項4に記載のファイル閲覧装置は、前期表示手段で複数の画像ファイルの内容から抽出されたキーワードを表示する際に、通常の並び順をファイル名順にするかキーワード抽出日時順にするか設定する設定手段を有することを特徴とする。

【0011】

【発明の実施の形態】

以下、本発明の実施例を図面に基づいて詳細に説明する。

【0012】

(実施例)

本発明は、ファイル閲覧アプリケーションの一つの機能であり、本説明はファイル閲覧アプリケーションの本発明に関わる画像ファイルから、その内容を複数のキーワードを抽出し表示する機能について説明する。

30

【0013】

図1に本実施例のブロック図を示す。101はCPUで本実施例全体の制御を行なう。102はハードディスクコントローラ(HDC)で103のハードディスク(HD)内のデータ・プログラムの制御を行なう。ハードディスク内には、本発明に関わる機能を持ったファイル閲覧プログラム113、画像ファイルA114、画像ファイルB115、画像ファイルC116が格納されている。104はキーボード、105は例えばマウスやディジタイザなどのポインティングデバイス(PD)でプログラム開始などの指示を出す。106はRAMでプログラムやデータを格納する。109は表示コントローラでVRAM108に格納された映像データを映像信号としてモニター110に出力する制御を行なう。

40

【0014】

図2は、本発明に関わる機能を持ったファイル閲覧アプリケーションのブロック図である。ただし、本発明に直接関わらない部分は簡略化している。

【0015】

図2において、メイン処理部201は、ファイル閲覧アプリケーションの全体の制御を行なう部分である。ファイルマネージャ処理部202は、HD(103)に階層構造に格納されたファイルのファイル名を表示するための処理や階層間の移動などの処理を行なっ

50

ている。

【0016】

GUI (Graphical User Interface) 処理部 203 は、ファイルマネージャー処理部 202 で得られたある階層内に存在するファイルの情報等を表示し、また、利用者からのキーボード 104、あるいはポインティングデバイス 105 などの操作により、ファイルやディレクトリが選択された時にその情報を、メイン処理部 201 に送る。

【0017】

キーワード抽出処理部 204 では、ファイルが選択された情報を GUI 処理部 203 からメイン処理部 201 を経由して渡されたときに、選択されたファイルが画像ファイルであるか判断し、画像ファイルである場合は、画像読込部 205 で画像データを読み込み、キーワード抽出部 206 で画像ファイルの内容を的確に表す複数のキーワードを抽出する。その後キーワード抽出処理部 204 は抽出されたキーワードをキーワードデータベース 208 に格納し、それを管理する。

【0018】

また、キーワード抽出処理部 204 は抽出されたキーワードの情報をメイン処理部 201 に返す。

【0019】

さらに、キーワード抽出処理部 204 は、CPU 監視部 207 に CPU の使用状況を監視させ、同一コンピュータ上で動作している別のアプリケーションに対する CPU の負荷が増大したときにはキーワード抽出部 206 にキーワード抽出処理を中止させ、中止したキーワード抽出処理の情報 (ファイルのパス等) をバックグラウンド処理登録データベース 209 に格納する。

【0020】

その後、同一コンピュータ上で動作している別のアプリケーションに対する CPU の負荷が軽減されたときは、CPU 監視部 207 はキーワード抽出処理部 204 に対しその事を通知する。

【0021】

その通知を受けたキーワード抽出処理部 204 は、バックグラウンド処理登録データベース 209 に格納してある中止していたキーワード抽出処理の情報を読み出し、その情報をキーワード抽出部 206 に送り、キーワード抽出処理を再開させる。

【0022】

図 3 は、本発明の機能を持ったファイル閲覧アプリケーションの GUI 画面である。

【0023】

301 には、現在このファイル閲覧アプリケーションがファイルシステム内のどのディレクトリを表示しているかを、path で示している。303 は、現在表示しているディレクトリに存在するファイルやディレクトリを示している。図中では、今 image1.gif と image2.gif が選択されており、その選択されていることを示すために、そのファイル名とアイコン (絵文字) を反転表示で示している。302 は本発明の機能を使用して、その選択されたファイル (image1.gif、image2.gif) の画像データから抽出した複数のキーワードを表示している。

【0024】

図 4 は、図 3 の 303 で表示しているファイルが、利用者からのキーボード 104、あるいはマウスなどのポインティングデバイス 105 などの操作によって複数選択されたときに、本発明の機能がキーワード抽出を行い、抽出されたキーワードを表示する流れを示した図であり、以下にその処理を詳細に説明する。

【0025】

S401 にて、選択された全てのファイル名と選択されたファイルの総数などの情報を取得する。また、ここで、RAM 106 にある処理回数カウンターを 0 にリセットする。

【0026】

10

20

30

40

50

S 4 0 2 では、キーワード情報表示部 3 0 2 に現在表示されている情報を消去する。

【 0 0 2 7 】

以降、S 4 0 3 から S 4 1 4 までの処理の流れは、各ファイルごとに実施される。S 4 0 3 では、選択されたファイルが、画像ファイルかどうかを判断する。判断方法は、ファイル名の拡張子で、画像ファイルの拡張子（例えば、j p g、g i f、b m p）かどうかで判断する。

【 0 0 2 8 】

S 4 0 4 では、キーワード抽出処理部 2 0 4 からキーワードデータベース 2 0 8 にアクセスし、対象となる画像ファイルのキーワードが登録されているかを確認する。確認した結果、キーワードが登録されていないければ、S 4 0 5 でキーワード抽出処理が未完であるメッセージを図 3 の 3 0 2 に表示する。また S 4 0 4 で確認した結果、すでにキーワードが登録されている場合、S 4 1 2 に行きそのキーワードを図 3 の 3 0 2 に表示する。さらに S 4 1 3 で画像ファイルの最終更新日時をキーワードデータベースに登録されている最終更新日時情報と比較することでキーワードが抽出・登録された後に画像ファイルが更新されているかを確認する。更新されていないければ S 4 1 4 に進む。S 4 0 4 で確認した結果キーワードが登録されていない、もしくは S 4 1 3 で画像ファイルが更新されていると判断された場合は、S 4 0 6 に進む。

10

【 0 0 2 9 】

S 4 0 6 では、RAM 1 0 6 にあるバックグラウンド処理終了フラグが H I G H になっているかチェックする。もし、H I G H になっていれば、後述のバックグラウンド処理の中止命令がすでになされていることなので、S 4 0 8 に進む。L O W のままの場合、S 4 0 7 にて、バックグラウンド終了フラグを H I G H にすることで、バックグラウンド処理の中止命令を出す。

20

【 0 0 3 0 】

バックグラウンド処理を中止させるのは、利用者からのキーボード 1 0 4、あるいはマウス 1 0 5 などの操作により、ファイルが選択された時の本説明の処理が優先的に行われなければならないためである。

【 0 0 3 1 】

S 4 0 8 にてキーワードを抽出する。キーワードの抽出方法について図 5 および図 9 を基に以下に説明する。

30

【 0 0 3 2 】

S 5 0 1 で CPU 監視部 2 0 7 で取得した CPU 使用率データを確認し、CPU 使用率が図 6 の CPU 使用率上限値設定画面であらかじめ設定された値を超えていなければ、S 5 0 2 に進む。もし、超えていれば、画像ファイルのパスをバックグラウンド処理登録データベースに登録し、キーワード抽出処理を終了する。

【 0 0 3 3 】

またこのキーワード抽出処理がバックグラウンド処理の S 7 0 3 であった場合、さらにここで、RAM 1 0 6 にあるバックグラウンド処理終了フラグを確認し、バックグラウンド処理終了フラグが H I G H になっている場合、画像ファイルのパスをバックグラウンド処理登録データベースに登録し、キーワード抽出処理を終了する。

40

【 0 0 3 4 】

S 5 0 2 では画像読込部 2 0 5 にて画像ファイルを読み込み、RAM 1 0 6 に格納する。S 5 0 3 では S 5 0 1 と同様に CPU 使用率とバックグラウンド処理終了フラグを確認する。

【 0 0 3 5 】

S 5 0 4 では、読み込んだ画像データの文字領域を抽出する。その手法は、画像データを適当なしきい値で 2 値化した後に、図 1 0 に示すような拡大処理を数回行った後ラベリング処理をする。次に、図 1 2、1 3、1 4 に示すように各領域に対しその最小外接矩形をもとめ、その幅 W、長さ L を求める。この W、L の関係性から、文字領域かそうでないかを判断する。この手法は特開平 6 - 1 5 0 0 5 7 で詳細に述べられている。

50

【0036】

S505では、S504にて一つでも文字領域と判断された領域があれば、S506に進み、そうでなければ、S516に進む。S506ではS501と同様にCPU使用率とバックグラウンド処理終了フラグを確認する。

【0037】

S507では、S504で抽出された文字領域に対して文字認識を行う。S508ではS501と同様にCPU使用率とバックグラウンド処理終了フラグを確認する。

【0038】

S509では、S507で文字認識された全文字領域の文字を1つの文書と捕らえ、その文書を形態素解析を用いて最小単位の語に分割し、各語の文書内での出現頻度、出現位置、主格や目的格などの格情報等によりキーワードを抽出する従来のキーワード抽出法（たとえば、特開平6-282572など）を用いてキーワード抽出を行う。S510ではS501と同様にCPU使用率とバックグラウンド処理終了フラグを確認する。

【0039】

S511はS509でのキーワード抽出結果および文字の大きさ、色、文字領域において文章内の語であったかそれとも単独の単語であったか等によって評価点を与える。まず、S509のキーワード抽出法で抽出された文書内の語に評価点を与える。次にS502で求めた各文字領域の最小外接矩形の面積を計算し、それをその文字領域に属する文字の総数で割ることで、各文字領域の一文字当たりの面積を求める。その面積の大きさは文字の大きさを意味しており、一文字当たりの面積が大きいほど大きな文字で書かれていることになる。文字領域の文章が大きな文字で書かれているということは、その画像のなかで、重要な文章であるので、その文章に含まれる語に評価点を与える。

【0040】

さらに各文字の色を測定し一番使用されている文字数が少ない色を使用している語に評価点を与える。最後に、各文字領域に書かれている文が、文章ではなく単語である場合（つまり、文字領域の文の中に助詞・助動詞が含まれていない）

【0041】

助詞・助動詞が含まれているか・いないかはS509で形態素解析した結果で判別可能である。)にその単語に含まれる語に評価点を与える。S512はS501と同様にCPU使用率とバックグラウンド処理終了フラグを確認する。

【0042】

S513では、S511で与えられた評価点の合計が大きい語から図8のキーワード数設定画面であらかじめ設定された数をキーワードとし、決定する。S514ではS501と同様にCPU使用率とバックグラウンド処理終了フラグを確認する。

【0043】

S515ではバックグラウンド処理登録データベース209に本画像ファイルが登録されていないか確認し、すでに登録されていればそれを抹消する。

S516で抽出されたキーワードを画像ファイルのパスと最終更新日時データと共にキーワードデータベース208に登録する。

【0044】

図9を用いてS511で説明した評価点の与え方の具体的な例を示す。ここでは評価点は従来のキーワード抽出法で抽出された語でキーワードとして重要度の高いほうから3点、2点、1点を与え、また、文字の大きさで大きいほうから2点と1点を与え、文字の色で一番使われていない色で書かれた語に2点を与え、最後に文字領域に書かれている語が単語である語に2点を与える。

【0045】

その結果を図15に示す。ただし、評価点が見つかなかった語の掲載は省略した。

【0046】

まず、「日曜日」「運動会」「応援」の順で従来のキーワード抽出法で重要であるとされたので、それぞれ3, 2, 1点を与えた。次に文字の大きさで大きい順に「小学校運動会

10

20

30

40

50

」「応援をよろしく」に含まれる語「小学校」「運動会」「応援」「よろしく」にそれぞれ2, 2, 1, 1点を与えた。さらに「応援をよろしく」の文章が赤い文字色でかかれており、その他は、黒で書かれているので、「応援」、「よろしく」に2点を与えた。

【0047】

最後に、「第10回」、「小学校運動会」がそれぞれの文字領域において文章ではなく単語であるので、「運動会」「小学校」「第10回」に2点を与えた。

【0048】

以上の結果、最終的に評価点が高かった順に「運動会」「応援」「小学校」「日曜日」「よろしく」「第10回」という結果が得られ、S513で上位から図8のキーワード数設定画面であらかじめ設定された数をキーワードと決定する。

10

【0049】

図4のS409では、キーワードデータベースにアクセスし、S405にてキーワード抽出が完了したかを登録されている更新日時データと文書ファイルの更新日時情報を比較し一致しているかで確認する。ただし、S501、S503、S505、S508、S510、S512、S514でCPU利用率が設定値を超えた場合にS517にてキーワード抽出処理がバックグラウンド処理に移され、抽出が完了してない場合がある。よって完了している、つまり一致している場合はS410に進み、完了していない、つまり一致していない場合S414に進む。

【0050】

S410では、S408で抽出されたキーワードを図3の302に表示する。

20

【0051】

S414ではRAM106にある処理回数カウンターを1カウントアップし、その結果S401で取得した選択ファイルの総数に一致するか確認する。一致していれば、すべてのファイルへの処理が終了したことを示すので、S415に進む。まだすべてのファイルへの処理が終了してない場合、S403に戻り、次のファイルへの処理を行う。

【0052】

最後にS415でバックグラウンド処理終了フラグをLOWにしたのち、バックグラウンド処理を再開させる。

【0053】

図7は、バックグラウンド処理の流れ図である。

30

【0054】

まず、S701ではバックグラウンド処理登録データベース209にアクセスし、バックグラウンド処理待ち状態のデータ(画像ファイルへのパスなど)があるか確認する。もしなければ、S704に進む。

【0055】

S701でバックグラウンド処理待ちデータがバックグラウンド処理登録データベース209に存在するならば、S702に進み、バックグラウンド処理待ちデータをバックグラウンド処理登録データベース209から読み込む。その時、読み込んだデータの登録はバックグラウンド処理登録データベース209から抹消しておく。

【0056】

次にS703でキーワードを抽出する。キーワード抽出処理の説明は、上記において図5を用いた説明に準じるので、ここでは省略する。

40

【0057】

最後に、S704にてRAM106にあるバックグラウンド処理終了フラグを確認し、バックグラウンド処理終了フラグがHIGHになっている場合、処理を終了する。また、LOWである場合、S701に戻りバックグラウンド処理を繰り返す。

【0058】

バックグラウンド処理の開始手段は、本発明の係わるファイル閲覧プログラムの起動時であってもかまわない。また、バックグラウンド処理の開始手段は、図3に示す本発明の係わるファイル閲覧プログラムのGUI画面からキーボード104もしくはポインティング

50

デバイス105からの入力、メニュー選択、ボタン押下、等であってもかまわない。

【0059】

上記説明ではS501、S503、S505、S508、S510、S512、S514でキーワード抽出処理部204がCPU監視部207が取得したCPU使用率データを確認して処理の終了を判断したが、その判断をCPU監視部207に判断させ、CPU使用率が設定値を超えたときにCPU監視部207からキーワード抽出処理部204に割り込みをかける形で処理し、割り込みがかかった時点でS517を実施し、キーワード抽出処理を終了してもかまわない。

【0060】

上記説明では、画像から文字を切り出す手段は特開平6-150057を用いて説明したが、それ以外の手段を用いても構わない。 10

【0061】

上記説明では、キーワード抽出手段は特開平6-82572を用いて説明したが、それ以外の手段を用いても構わない。

【0062】

図3の304、305は、それぞれ「名前」ボタン、「キーワード抽出日時」ボタンである。「名前」ボタン304を一度押下することで、キーワード情報表示部302の表示がファイル名の昇順で表示される。さらに「名前」ボタン304を一度押下すると、キーワード情報表示部302の表示はファイル名の降順で表示される。「名前」ボタン304を押下することで、キーワード情報表示部302の表示はファイル名の昇順/降順のトグル式で切り替わる。 20

【0063】

これにより、例えば作成した画像ファイルのバージョンをファイル名で管理する場合(例えば、「お知らせ1.gif」「お知らせ2.gif」「お知らせ3.gif」...)にバージョン順の内容の移り変わりを簡単に確認することができる。

【0064】

「キーワード抽出日時」ボタン305を一度押下することで、キーワード情報表示部302の表示がキーワード抽出日時の昇順で表示される。さらに「キーワード抽出日時」ボタンを一度押下すると、キーワード情報表示部302の表示はキーワード抽出日時の降順で表示される。「キーワード抽出日時」ボタンを押下することで、キーワード情報表示部302の表示はキーワード抽出日時の昇順/降順のトグル式で切り替わる。 30

【0065】

これにより、例えば選択した文書ファイルのファイル名はそれぞれまったく違っていても書かれている内容は近いと判断されるときに、キーワード抽出日時順の内容の移り変わりを簡単に確認することができる。

【0066】

図17は、複数ファイル選択時にキーワード情報表示部302において、デフォルトでどの順序で表示するかを設定する画面である。図17の1703~1706のチェックボタンのどれかを選択することで名前の昇順/降順・キーワード抽出日時順の昇順/降順のうちのどれかを選択し、「OK」ボタン1701を押下することで、デフォルトでどの順序で表示するかを設定することができる。 40

【0067】

以上により、本実施例では、ファイル閲覧アプリケーションにて、複数の選択された画像ファイルにおいて、各々の画像ファイルから複数のキーワードを抽出し表示することで、容易に複数の画像ファイルの内容を確認することが可能である。

【0068】

【発明の効果】

以上説明したように、請求項1のファイル閲覧装置によれば、ファイル閲覧アプリケーションにおいて、複数の画像ファイルが選択されたときに、選択されたすべての画像ファイルに対して、その内容から内容を表すのにふさわしい複数のキーワードを抽出し表示する 50

ことで、容易に複数の画像ファイルの内容を確認することができる。

【0069】

また、請求項2のファイル閲覧装置によれば、ファイル閲覧アプリケーションにおいて、複数の画像ファイルの内容から抽出した複数のキーワードの表示をファイル名の昇順および降順で並び替えることができる。

【0070】

また、請求項3のファイル閲覧装置によれば、ファイル閲覧アプリケーションにおいて、複数の画像ファイルの内容から抽出した複数のキーワードの表示をキーワード抽出日時の昇順および降順で並び替えることができる。

【0071】

また、請求項4のファイル閲覧装置によれば、ファイル閲覧アプリケーションにおいて、複数の画像ファイルの内容から抽出した複数のキーワードを表示する際に、ファイル名の昇順・降順およびキーワード抽出日時の昇順・降順の中からどの順で表示させるか選択することができる。

【図面の簡単な説明】

【図1】本発明の第1の実施の形態に係るファイル閲覧装置の概略構成を示すブロック図である。

【図2】本発明の第1の実施の形態に係るファイル閲覧装置のプログラムのモジュール関係を示す説明図である。

【図3】本発明の第1の実施の形態に係るファイル閲覧装置のGUI画面である。

【図4】本発明の第1の実施の形態に係るファイル閲覧装置の処理の流れ図である。

【図5】本発明の第1の実施の形態に係るファイル閲覧装置のキーワード抽出処理の流れ図である。

【図6】本発明の第1の実施の形態に係るファイル閲覧装置のCPU使用率上限値設定画面である。

【図7】本発明の第1の実施の形態に係るファイル閲覧装置のバックグラウンド処理の流れ図である。

【図8】本発明の第1の実施の形態に係るファイル閲覧装置のキーワード数設定画面である。

【図9】画像ファイルの画像データの一例を示す模式図である。

【図10】拡大処理を説明した説明図である。

【図11】文字領域を分離するために画像データに画像処理を施した結果を示す図である。

【図12】領域分離をした結果を示す図である。

【図13】最小外接矩形を説明する図である。

【図14】文字領域判定処理に使われる幅W，長さLを説明するための図である。

【図15】評価点を与えた結果を示す図である。

【図16】画像ファイルの画像データの一例を示す模式図である。

【図17】本発明の第1の実施の形態に係るファイル閲覧装置のキーワード表示設定画面である。

【符号の説明】

101 CPU

102 ハードディスクコントローラ

103 ハードディスク

104 キーボード

105 ポインティングデバイス

106 RAM

108 ビデオRAM

109 表示コントローラ

110 モニター

10

20

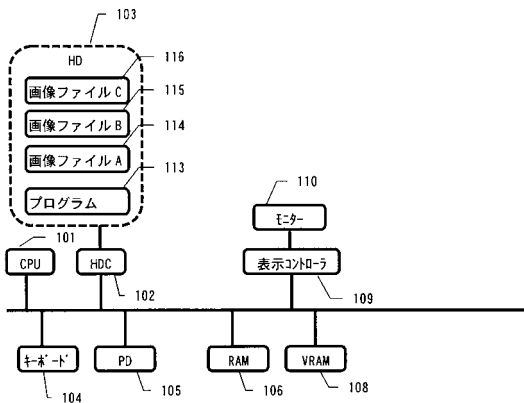
30

40

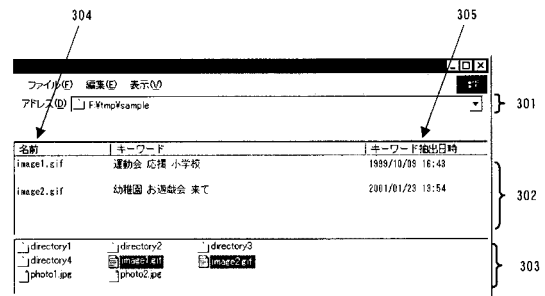
50

- 1 1 3 プログラム
- 1 1 4 画像ファイル A
- 1 1 5 画像ファイル B
- 1 1 6 画像ファイル C
- 2 0 1 メイン処理部
- 2 0 2 ファイルマネージャー処理部
- 2 0 3 GUI処理部
- 2 0 4 キーワード抽出処理部
- 2 0 5 画像読込部
- 2 0 6 キーワード抽出部
- 2 0 7 CPU監視部
- 2 0 8 キーワードデータベース
- 2 0 9 バックグラウンド処理登録データベース

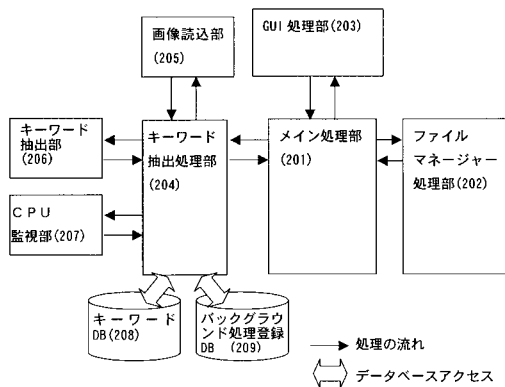
【図1】



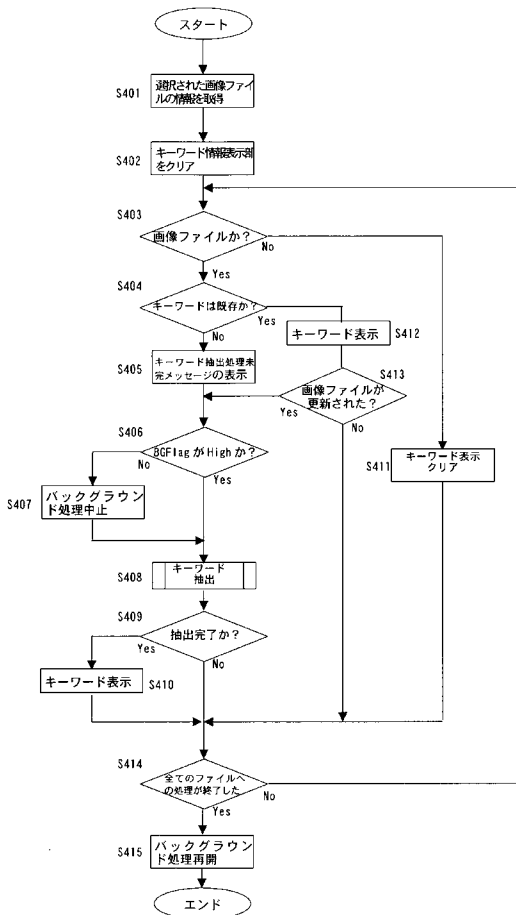
【図3】



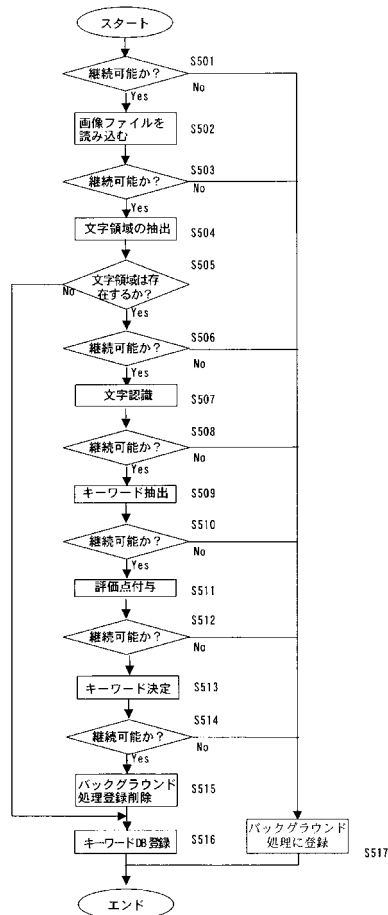
【図2】



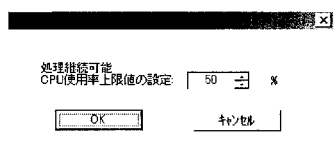
【 図 4 】



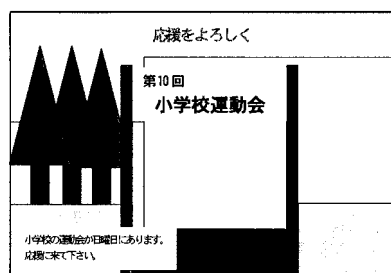
【 図 5 】



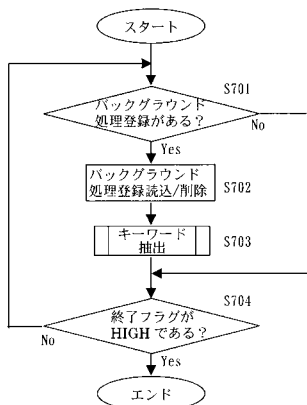
【 図 6 】



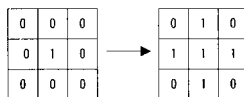
【 図 9 】



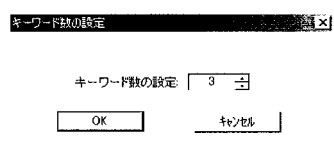
【 図 7 】



【 図 10 】



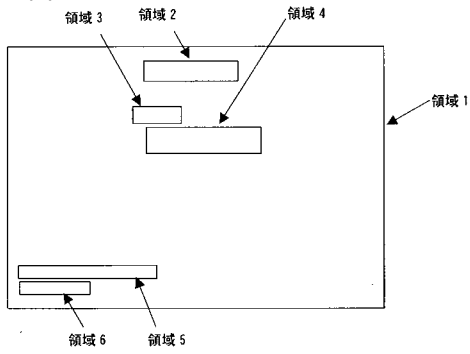
【 図 8 】



【 図 11 】



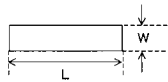
【図 1 2】



【図 1 3】



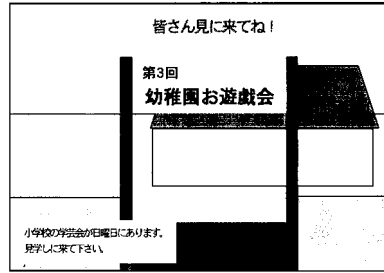
【図 1 4】



【図 1 5】

キーワード候補	従来のキーワードの結果	文字の大きさ	文字色	単語	合計
日曜日	3	0	0	0	3
運動会	2	2	0	2	6
応援	1	1	2	0	4
よろしく	0	1	2	0	3
小学校	0	2	0	2	4
第10回	0	0	0	2	2

【図 1 6】



【図 1 7】

