

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4817994号
(P4817994)

(45) 発行日 平成23年11月16日(2011.11.16)

(24) 登録日 平成23年9月9日(2011.9.9)

(51) Int.Cl.

G06F 3/12 (2006.01)
H04N 1/00 (2006.01)

F 1

G06F 3/12
H04N 1/00 1 O 7 Z

請求項の数 10 (全 18 頁)

(21) 出願番号 特願2006-183015 (P2006-183015)
 (22) 出願日 平成18年7月3日 (2006.7.3)
 (65) 公開番号 特開2008-15579 (P2008-15579A)
 (43) 公開日 平成20年1月24日 (2008.1.24)
 審査請求日 平成21年6月25日 (2009.6.25)

(73) 特許権者 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
 (74) 代理人 100126240
 弁理士 阿部 琢磨
 (74) 代理人 100124442
 弁理士 黒岩 創吾
 (72) 発明者 岡部 康矢
 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
 審査官 内田 正和

最終頁に続く

(54) 【発明の名称】データ管理システム

(57) 【特許請求の範囲】

【請求項 1】

クライアント装置と画像処理サーバとデータベースとを含むデータ管理システムであつて、

前記クライアント装置は、

複数種類のアプリケーションを実行するCPUと、

管理設定画面を介して為されたユーザの操作に基づいて、各アプリケーションによる印刷処理時に出力される印刷描画命令からテキストデータを抽出するか否かを、当該クライアント装置のCPUで実行されるアプリケーションの種類ごとに設定する設定手段と、

前記設定手段でテキストデータを抽出する設定が為されている種類のアプリケーションによる印刷処理を実行する場合、当該アプリケーションによる印刷処理時に出力される印刷描画命令からテキストデータを抽出する抽出手段と、

前記設定手段でテキストデータを抽出する設定が為されている種類のアプリケーションによる印刷処理を実行する場合は、前記抽出手段で抽出したテキストデータと当該印刷処理時に出力される印刷描画命令から生成される画像データとを関連付けて前記画像処理サーバに送信し、一方、前記設定手段でテキストデータを抽出する設定が為されていない種類のアプリケーションによる印刷処理を実行する場合は、当該印刷処理時に出力される印刷描画命令から生成される画像データを前記画像処理サーバに送信するように制御する送信手段と、を有し、

前記画像処理サーバは、

10

20

前記クライアント装置の前記送信手段からテキストデータが送信されてこなかったと判断した場合に、前記クライアント装置の前記送信手段から送信された画像データに対してOCR処理を実行してテキストデータを抽出するOCR手段と、

前記クライアント装置の前記送信手段からテキストデータが送信されてこなかったと判断した場合は、前記OCR手段で抽出されたテキストデータと前記送信されてきた画像データとを関連付けて前記データベースへ登録し、一方、前記クライアント装置の前記送信手段から前記テキストデータと前記画像データとが送信されてきたと判断した場合は、当該送信されてきた前記テキストデータと前記画像データとを関連付けて前記データベースへ登録する登録手段と、を有する

ことを特徴とするデータ管理システム。

10

【請求項2】

前記クライアント装置の抽出手段は、前記CPUがプリントドライバを実行することによって実現されることを特徴とする請求項1に記載のデータ管理システム。

【請求項3】

前記設定手段は、前記アプリケーションによる印刷処理時に出力される印刷描画命令からテキストデータを抽出することをデフォルト設定として、前記アプリケーションによる印刷処理時に出力される印刷描画命令からテキストデータを抽出しないように設定するアプリケーションの種類を前記管理設定画面において前記ユーザに指定させることによって前記設定を行うことを特徴とする請求項1または2のいずれかに記載のデータ管理システム。

20

【請求項4】

前記データ管理システムは更に画像処理装置を含み、

前記画像処理装置は、当該画像処理装置で実行したジョブに関する画像データを前記画像処理サーバに送信する送信手段を有し、

前記画像処理サーバの前記OCR手段は、前記画像処理装置から送信されてきた画像データに対してOCR処理を実行してテキストデータを抽出し、更に、前記画像処理サーバの前記登録手段は、当該抽出されたテキストデータと前記画像データとを関連付けて前記データベースへ登録することを特徴とする請求項1に記載のデータ管理システム。

【請求項5】

前記クライアント装置は、前記アプリケーションによる印刷処理時のログ情報を抽出するログ情報抽出手段を更に備え、

30

前記画像処理サーバの前記登録手段は、前記クライアント装置の前記送信手段からテキストデータが送信されてこなかったと判断した場合、前記OCR手段で抽出されたテキストデータと前記送信されてきた画像データと前記ログ情報を関連付けて前記データベースへ登録する一方、前記クライアント装置の前記送信手段から前記テキストデータと前記画像データとが送信されてきたと判断した場合、当該送信されてきた前記テキストデータと前記画像データと前記ログ情報を関連付けて前記データベースへ登録することを特徴とする請求項1に記載のデータ管理システム。

【請求項6】

複数種類のアプリケーションを実行するCPUと、

40

管理設定画面を介して為されたユーザの操作に基づいて、各アプリケーションによる印刷処理時に出力される印刷描画命令からテキストデータを抽出するか否かを、当該クライアント装置のCPUで実行されるアプリケーションの種類ごとに設定する設定手段と、

前記設定手段でテキストデータを抽出する設定が為されている種類のアプリケーションによる印刷処理を実行する場合、当該アプリケーションによる印刷処理時に出力される印刷描画命令からテキストデータを抽出する抽出手段と、

前記設定手段でテキストデータを抽出する設定が為されている種類のアプリケーションによる印刷処理を実行する場合は、前記抽出手段で抽出したテキストデータと当該印刷処理時に出力される印刷描画命令から生成される画像データとを関連付けてサーバへ送信することにより、前記サーバが当該送信されたテキストデータと画像データとを関連付けて

50

データベースに格納するように制御する一方、前記設定手段でテキストデータを抽出する設定が為されていない種類のアプリケーションによる印刷処理を実行する場合は、当該印刷処理時に出力される印刷描画命令から生成される画像データを前記サーバへ送信することにより、前記サーバが当該送信された画像データに対してOCR処理を実行してテキストデータを抽出し更に当該抽出されたテキストデータと当該送信された画像データとを関連付けて前記データベースに格納するように制御する送信手段と、

を有することを特徴とするクライアント装置。

【請求項7】

クライアント装置と画像処理サーバとデータベースとを含むデータ管理システムにおける制御方法であって、

10

複数種類のアプリケーションを実行するCPUを備える前記クライアント装置において、

前記クライアント装置の設定手段が、管理設定画面を介して為されたユーザの操作に基づいて、各アプリケーションによる印刷処理時に出力される印刷描画命令からテキストデータを抽出するか否かを、当該クライアント装置のCPUで実行されるアプリケーションの種類ごとに設定する設定ステップと、

前記クライアント装置の抽出手段が、前記設定ステップでテキストデータを抽出する設定が為されている種類のアプリケーションによる印刷処理を実行する場合、当該アプリケーションによる印刷処理時に出力される印刷描画命令からテキストデータを抽出する抽出ステップと、

20

前記クライアント装置の送信手段が、前記設定ステップでテキストデータを抽出する設定が為されている種類のアプリケーションによる印刷処理を実行する場合は、前記抽出ステップで抽出したテキストデータと当該印刷処理時に出力される印刷描画命令から生成される画像データとを関連付けて前記画像処理サーバに送信し、一方、前記設定ステップでテキストデータを抽出する設定が為されていない種類のアプリケーションによる印刷処理を実行する場合は、当該印刷処理時に出力される印刷描画命令から生成される画像データを前記画像処理サーバに送信する送信ステップと、を有し、

前記画像処理サーバにおいて、

前記画像処理サーバのOCR手段が、前記クライアント装置の前記送信ステップでテキストデータが送信されてこなかったと判断した場合に、前記クライアント装置の前記送信ステップで送信された画像データに対してOCR処理を実行してテキストデータを抽出するOCRステップと、

30

前記画像処理サーバの登録手段が、前記クライアント装置の前記送信ステップでテキストデータが送信されてこなかったと判断した場合は、前記OCRステップで抽出されたテキストデータと前記送信されてきた画像データとを関連付けて前記データベースへ登録し、一方、前記クライアント装置の前記送信ステップで前記テキストデータと前記画像データとが送信されてきたと判断した場合は、当該送信されてきた前記テキストデータと前記画像データとを関連付けて前記データベースへ登録する登録ステップと、を有することを特徴とするデータ管理システムにおける制御方法。

【請求項8】

40

複数のアプリケーションを実行するコンピュータを備えるクライアント装置と、画像処理サーバと、データベースとを含むデータ管理システムを制御するためのコンピュータプログラムであって、

前記クライアント装置のコンピュータを、

管理設定画面を介して為されたユーザの操作に基づいて、各アプリケーションによる印刷処理時に出力される印刷描画命令からテキストデータを抽出するか否かを、当該クライアント装置のコンピュータで実行されるアプリケーションの種類ごとに設定する設定手段、

前記設定手段でテキストデータを抽出する設定が為されている種類のアプリケーションによる印刷処理を実行する場合、当該アプリケーションによる印刷処理時に出力される印

50

刷描画命令からテキストデータを抽出する抽出手段、

前記設定手段でテキストデータを抽出する設定が為されている種類のアプリケーションによる印刷処理を実行する場合は、前記抽出手段で抽出したテキストデータと当該印刷処理時に出力される印刷描画命令から生成される画像データとを関連付けて前記画像処理サーバに送信し、一方、前記設定手段でテキストデータを抽出する設定が為されていない種類のアプリケーションによる印刷処理を実行する場合は、当該印刷処理時に出力される印刷描画命令から生成される画像データを前記画像処理サーバに送信するように制御する送信手段、

として機能させるためのプログラムと、

前記画像処理サーバのコンピュータを、

10

前記クライアント装置の前記送信手段からテキストデータが送信されてこなかったと判断した場合に、前記クライアント装置の前記送信手段から送信された画像データに対してOCR処理を実行してテキストデータを抽出するOCR手段、

前記クライアント装置の前記送信手段からテキストデータが送信されてこなかったと判断した場合は、前記OCR手段で抽出されたテキストデータと前記送信されてきた画像データとを関連付けて前記データベースへ登録し、一方、前記クライアント装置の前記送信手段から前記テキストデータと前記画像データとが送信されてきたと判断した場合は、当該送信されてきた前記テキストデータと前記画像データとを関連付けて前記データベースへ登録する登録手段、

として機能させるためのプログラムと、

20

によって構成されるコンピュータプログラム。

【請求項 9】

複数種類のアプリケーションを実行するクライアント装置のコンピュータを、

管理設定画面を介して為されたユーザの操作に基づいて、各アプリケーションによる印刷処理時に出力される印刷描画命令からテキストデータを抽出するか否かを、当該クライアント装置のコンピュータで実行されるアプリケーションの種類ごとに設定する設定手段、

前記設定手段でテキストデータを抽出する設定が為されている種類のアプリケーションによる印刷処理を実行する場合、当該アプリケーションによる印刷処理時に出力される印刷描画命令からテキストデータを抽出する抽出手段、

30

前記設定手段でテキストデータを抽出する設定が為されている種類のアプリケーションによる印刷処理を実行する場合は、前記抽出手段で抽出したテキストデータと当該印刷処理時に出力される印刷描画命令から生成される画像データとを関連付けてサーバへ送信することにより、前記サーバが当該送信されたテキストデータと画像データとを関連付けてデータベースに格納するように制御する一方、前記設定手段でテキストデータを抽出する設定が為されていない種類のアプリケーションによる印刷処理を実行する場合は、当該印刷処理時に出力される印刷描画命令から生成される画像データを前記サーバへ送信することにより、前記サーバが当該送信された画像データに対してOCR処理を実行してテキストデータを抽出し更に当該抽出されたテキストデータと当該送信された画像データとを関連付けて前記データベースに格納するように制御する送信手段、

40

として機能させるためのコンピュータプログラム。

【請求項 10】

請求項8または9のいずれかに記載のコンピュータプログラムを格納した、コンピュータ読み出し可能な記録媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、画像処理装置において行った処理（ジョブ）の履歴を管理するシステムに関する。特に、履歴情報、画像データ、テキストデータを管理サーバに効率よく格納、管理する技術に関する。

50

【背景技術】**【0002】**

デジタル複合機やプリンタなどの普及に従って、誰でも原稿の印刷、複写や送信が容易に行なうことが可能になっている。しかし、これらのデバイスの普及は、ユーザの利便性が向上する反面、機密原稿の印刷や複写、送信などが容易にできてしまい、情報漏洩しやすいという問題が発生している。

【0003】

このような問題への対策として、印刷や複写、FAX、E-mail送信等の処理を行う際に、処理を行ったジョブデータ（画像データやテキストデータ）を全てサーバ（記憶装置）に蓄積しておく文書管理システムが存在する。このシステムでは、管理者が前記蓄積したデータから所望の文書を検索することによって、いつ、どのデバイスで、どのような処理がなされたかを追跡可能になる。したがって、情報漏洩した場合などは、当該情報漏洩対象文書を検索することで、どこから情報漏洩したか特定（追跡）することができるようになるので、情報漏洩の抑止効果が見込める。10

【0004】

また、サーバに蓄積されたジョブデータを容易に検索するために、処理を行った画像データとテキストデータとを対応付けて蓄積しておき、当該テキストデータを全文検索することで検査することができる。システム監査者は例えば「機密情報」という単語を検索語として、蓄積されたデータに対して全文検索を実施し、検索結果としてその単語を含むジョブデータを得ることができる。20

【0005】

一方、サーバに蓄積されるジョブデータ（文書データ）は、プリントジョブ、コピージョブ、ファックスジョブなど様々な形式のデータである。本発明では、このような形式の異なるジョブ（文書）の入力／管理方法について注目する。

【0006】

また、特許文献1では、データ形式の異なる複数の文書を入力して一元管理し、シームレスに検索できるようにする技術が開示されている。特許文献1では、データ形式の異なる文書（例えばアプリ文書、WWW文書、FAX文書など）を処理対象として、それぞれの文書から予め定めた文書構造ファイルを生成して保管している。この文書構造ファイルは、オリジナル文書ファイルと、オリジナル文書ファイルから生成したテキストファイルと、サムネールファイルと、個々の文書構造ファイルを管理するための文書管理ファイルとから構成されている。そして、この文書構造ファイルを用いることにより、データ形式の異なる複数の文書を統合的に管理することを特徴としている。30

【特許文献1】特開平11-120202**【発明の開示】****【発明が解決しようとする課題】****【0007】**

特許文献1では、オリジナル文書ファイルから生成したテキストファイルを含む文書構造ファイルを、文書構造ファイル保管手段で保管している。更に、文書構造ファイル保管手段に保管されている文書構造ファイルからテキストファイルのみを抽出して作成した1つの検索用テキストファイルを、文書管理手段で格納している。40

【0008】

オリジナル文書としてテキストファイルが入力された場合、オリジナル文書ファイルとオリジナル文書から生成されたテキストファイルとを含む文書構造ファイルを保管する。更に、検索用テキストファイルを生成して保管している。つまり、テキストファイルの生成が冗長となってしまい、また、テキストファイルが二重に保存されることになってしまったという問題点があった。

【0009】

また本発明は、デジタル複合機やプリンタに対しユーザが実行させたジョブを蓄積するようなシステムに関するものであるので、処理（ジョブ）を実行すると共に画像データや50

テキストデータを生成して蓄積する必要がある。

【0010】

このようなシステムにおいて、アプリケーションデータをプリンタで印刷実行する場合、プリンタドライバにおいて文字描画命令からテキストデータを抽出し蓄積する方法が考えられる。しかしながら、テキストを文字描画命令として印刷命令せずに、テキストを画像として印刷命令してくるようなアプリケーションの場合、この手法ではテキストを抽出することができない。そのため、生成した画像データにOCRをかけてテキストを抽出したほうがより良いテキストデータを得ることができる場合がある。すなわちアプリケーションに応じて適切なテキスト抽出方法を設定できることが望ましい。

【課題を解決するための手段】

10

【0011】

上記課題を解決するために、本発明のデータ管理システムは、クライアント装置と画像処理サーバとデータベースとを含むデータ管理システムであって、

前記クライアント装置は、

複数種類のアプリケーションを実行するCPUと、

管理設定画面を介して為されたユーザの操作に基づいて、各アプリケーションによる印刷処理時に出力される印刷描画命令からテキストデータを抽出するか否かを、当該クライアント装置のCPUで実行されるアプリケーションの種類ごとに設定する設定手段と、

前記設定手段でテキストデータを抽出する設定が為されている種類のアプリケーションによる印刷処理を実行する場合、当該アプリケーションによる印刷処理時に出力される印刷描画命令からテキストデータを抽出する抽出手段と、

前記設定手段でテキストデータを抽出する設定が為されている種類のアプリケーションによる印刷処理を実行する場合は、前記抽出手段で抽出したテキストデータと当該印刷処理時に出力される印刷描画命令から生成される画像データとを関連付けて前記画像処理サーバに送信し、一方、前記設定手段でテキストデータを抽出する設定が為されていない種類のアプリケーションによる印刷処理を実行する場合は、当該印刷処理時に出力される印刷描画命令から生成される画像データを前記画像処理サーバに送信するように制御する送信手段と、を有し、

前記画像処理サーバは、

前記クライアント装置の前記送信手段からテキストデータが送信されてこなかったと判断した場合に、前記クライアント装置の前記送信手段から送信された画像データに対してOCR処理を実行してテキストデータを抽出するOCR手段と、

前記クライアント装置の前記送信手段からテキストデータが送信されてこなかったと判断した場合は、前記OCR手段で抽出されたテキストデータと前記送信してきた画像データとを関連付けて前記データベースへ登録し、一方、前記クライアント装置の前記送信手段から前記テキストデータと前記画像データとが送信されてきたと判断した場合は、当該送信されてきた前記テキストデータと前記画像データとを関連付けて前記データベースへ登録する登録手段と、を有する

ことを特徴とする。

【発明の効果】

30

【0012】

ジョブに応じてテキストの抽出方法を変えることができるので、検索に適したテキストデータを抽出することができる。また、データベースサーバ上に、1つのジョブに対して冗長にテキストデータを登録せずに済む。

【発明を実施するための最良の形態】

【0013】

(実施例1)

図1は本発明の実施の一形態に係るシステム構成図を表す図である。ここではネットワーク上にデジタル複合機101、画像処理サーバ102、データベースサーバ103、ユーザPC(クライアントPC)104、プリントサーバ106、プリンタ107、管理者

40

50

P C 1 0 8 が接続されている。

【 0 0 1 4 】

デジタル複合機 1 0 1 はスキャナ・プリンタ・コピー・メール・ファックスなどの機能を有している。またデジタル複合機 1 0 1 は実行したジョブに関して、そのジョブの電子的な画像データを生成し、画像データに解像度変換処理を行い、当該デジタル複合機の記憶装置に画像データを一時保存する。デジタル複合機 1 0 1 はその後、定期的な間隔で画像データを画像処理サーバ 1 0 2 へ転送する。

【 0 0 1 5 】

ユーザ P C 1 0 4 にはプリンタドライバがインストールされている。印刷処理が行われる（印刷ジョブの実行）と、プリンタドライバはそのジョブの電子的な画像データを生成したり、文字描画命令から文字情報を抽出してテキストデータを生成したりする。ユーザ P C 1 0 4 は、これらの画像データ、テキストデータを画像処理サーバ 1 0 2 へすぐに転送するか、当該ユーザ P C の記憶装置に一時保存して定期的な間隔で画像処理サーバ 1 0 2 へ転送する。

【 0 0 1 6 】

なおプリンタドライバはプリントサーバ 1 0 6 にインストールされているプリンタドライバを共有化してインストールしたものであっても良い。この場合、ユーザ P C 1 0 4 で印刷指示された電子データの追跡用データ（画像データ、テキストデータ）は、プリントサーバ 1 0 6 に保存される。プリントサーバ 1 0 6 はその後、定期的な間隔で追跡用データ（画像データとテキストデータ）を画像処理サーバ 1 0 2 に転送する。なお、以下の本実施例においては、プリントサーバ 1 0 6 を介さない構成について説明していくが、これに限るものではなく、プリントサーバを介しても構わない。

【 0 0 1 7 】

画像処理サーバ 1 0 2 は、デジタル複合機 1 0 1 やプリントサーバ 1 0 6 などから転送されてきた画像データに対し、データベースサーバ 1 0 3 に格納できる形式に画像変換処理や解像度変換処理を行う。また画像処理サーバ 1 0 2 では、管理者により O C R を行う設定がされていると、O C R 処理を行い、画像データからテキストデータを抽出する。画像処理サーバ 1 0 2 は処理し終わった画像データから順にデータベースサーバ 1 0 3 へ登録するとともに、当該画像データに対応するテキストデータを該画像の全文検索用テキストとしてデータベースサーバ 1 0 3 に設定する。

【 0 0 1 8 】

管理者 P C 1 0 8 は、システム管理者（監査者）がデータベースサーバ 1 0 3 へアクセスし、情報漏洩が発生していないか監査を行うためのクライアント P C である。1 0 9 はネットワーク（L A N ）である。

【 0 0 1 9 】

図 2 は、図 1 のユーザ P C 1 0 4 、画像処理サーバ 1 0 2 、データベースサーバ 1 0 3 、プリントサーバ 1 0 6 、および管理者 P C 1 0 8 のハードウェア構成を示すブロック図である。本実施例では、いずれも I B M - P C / A T 互換機などの汎用 P C を用いることが可能があるので、同じブロック図としている。

【 0 0 2 0 】

C P U 2 0 1 は、内部バスで接続される各デバイス（後述の R O M 、 R A M など）を直接或いは間接的に制御し、本発明を実現するためのプログラムを実行する C P U である。

【 0 0 2 1 】

R O M 2 0 2 は、B I O S などの基本ソフトウェアを格納している。R A M 2 0 3 は、C P U 2 0 1 のワーク領域として利用される。また、R A M 2 0 3 は本発明を実現するためのプログラムをロードするための一時記憶として利用される。

【 0 0 2 2 】

H D D 2 0 4 （ハードディスクドライブ）は前記プログラムを格納している。入力装置 2 0 5 は、前記プログラムで、ユーザからの操作入力を受け付けるための装置である。モニタ 2 0 6 は、操作画面を表示する。L A N I / F 2 0 7 はネットワークに接続するた

10

20

30

40

50

めのネットワークインターフェースである。

【0023】

図3は、図1に示したデジタル複合機101の構成を示すブロック図の例である。

【0024】

コントローラユニット300は、画像入力デバイスであるスキャナ部302や画像出力デバイスであるプリンタ部303と接続されている。また、ネットワーク(LAN)304や公衆回線(WAN)305と接続されている。これらのインターフェースを用いて、画像情報やデバイス情報の入出力を行う。

【0025】

コントローラユニット300において、CPU310はシステム全体を制御する。RAM311は、CPU310が動作するためのシステムワークメモリであり、入力された画像データを一時記憶するための画像メモリ(バッファメモリ)でもある。ROM312は、ブートROMであり、システムのブートプログラムが格納されている。ハードディスクドライブ(HDD)313は、システムソフトウェア、ジョブのログ情報、画像データなどを格納する。

【0026】

操作部I/F314は、操作部301とのインターフェース部であり、操作部301に表示する画面データを出力する。また、操作部I/F314は、操作部301からオペレータが入力した情報を、CPU310に伝える役割をする。

【0027】

ネットワーク部(Network)315は、ネットワーク(LAN)304に接続し、ネットワークに接続された外部装置との間で情報の入出力を行う。316はモデム(MODEM)で、公衆回線(WAN)305に接続し、外部装置との間で情報の入出力を行う。以上のデバイスがシステムバス327上に配置される。

【0028】

イメージバスインターフェース(Image Bus I/F)317は、システムバス327と画像データを高速で転送する画像バス328とを接続し、データ構造を変換するバスプリッジである。画像バス328は、PCIバスまたはIEEE1394で構成される。画像バス328上には以下のデバイスが配置される。

【0029】

ラスタイメージプロセッサ(RIP)318は、PDLコードをピットマップイメージに展開する。デバイスI/F319は、画像入力デバイスであるスキャナ部302・プリンタ部303とコントローラ300とを、画像入力部インターフェース329・印刷部インターフェース330を介して接続し、画像データの同期系/非同期系の変換を行う。

【0030】

スキャナ画像処理部320は、入力画像データに対し補正、加工、編集を行う。また、スキャナ画像処理部320は、入力された画像データがカラー原稿か白黒原稿かを画像データの彩度信号から判断しその結果を保持する機能を有する。プリンタ画像処理部321は、出力画像データに対し補正、加工、編集を行う。

【0031】

画像回転部322は、スキャナ画像処理部320と連携して、スキャナ部302からの画像読み込みと同時に画像データを回転しメモリ上に格納することができる。また画像回転部322は、メモリ上にある画像データを回転してメモリ上に格納し直したり、メモリ上にある画像データをプリンタ画像処理部321と連携して回転しながら印字出力したりすることができる。

【0032】

画像圧縮部323は、多値画像データに対してはJPEG、2値画像データに対してはJBIG、MMR、MR、MHなどの圧縮伸長処理を行う。解像度変換部324は、メモリ上にある画像データを解像度変換処理し、メモリ上に格納する。

【0033】

10

20

30

40

50

色空間変換部 325 は、マトリクス演算により、たとえばメモリ上にある YUV 画像データを Lab 画像データに変換し、メモリ上に格納する。階調変換部 326 は、たとえばメモリ上にある 8bit, 256 階調の画像データを誤差拡散処理などの手法により 1bit, 2 階調に変換し、メモリ上に格納する。

【0034】

画像回転部 322、画像圧縮部 323、解像度変換部 324、色空間変換部 325、階調変換部 326 は、それぞれ連動して動作することが可能である。たとえば、メモリ上の画像データに対して画像回転と解像度変換処理を行う場合は、画像回転部 322 と解像度変換部 324 との両処理を、メモリを介すことなしに行うことができる。

【0035】

図4は、デジタル複合機101のコントローラユニットにおける、システムソフトウェアの機能構成を示すブロック図である。ユーザインタフェース(UI)部401は、オペレータからの指示をアプリケーション部402に供給するとともに、その処理結果をアプリケーション部402から受け取り、操作部301に表示する画面を生成する。

【0036】

アプリケーション部402は、ユーザインタフェース(UI)部401からの要求に従って処理を行う。アプリケーション部402は、たとえば、コピーが要求された場合、ジョブコントロールメイン部403に対して、指定されたコピー設定が為されたコピージョブを投入する。そして、機器の状態やジョブの状態などの情報をジョブコントロールメイン部403から受け取る。また、アプリケーション部402は、画像処理サーバ102に対して、ジョブのログ情報とログ画像データを送信する場合、ジョブコントロールメイン部403からジョブのログ情報とログ画像データを受け取る。その後、アプリケーション部402は、ジョブのログ情報と追跡用画像データ(ログ画像データ)とを送信する要求をサービス通信部404に供給し、その処理結果を受け取る。

【0037】

ジョブコントロールメイン部403は、スキャンジョブ、コピージョブ、プリントジョブ、FAXジョブなどの、複数のジョブの処理を司るものである。たとえば、コピージョブが投入された場合、ジョブコントロールメイン部403は、指定されたコピー設定に基づきスキャナ部302とプリンタ部303とを動作させ、原稿をスキャンして、当該スキャンで得た画像データを印刷する。

【0038】

また、こうしたジョブの処理において、複合機に画像を入力するジョブ(入力ジョブ)が実行された場合、該ジョブのログ情報とともに追跡用の画像データ(ログ保存用の画像データ)を生成し、ハードディスクドライブに保存する。なお、データベースサーバに格納される追跡用の画像データは、ジョブの画像データをそのまま複製するようにしてもよいし、所定の解像度・色空間・圧縮方式に変換するようにしてもよい。なお、本実施例では、スキャンジョブ、コピージョブ、プリントジョブ、FAXジョブなど、画像データが複合機に入力されるときのジョブを入力ジョブと総称する。また、本実施例の複合機は、HDD313にスキャンやFAX受信などで得た画像データを保持しておき、後で、ユーザーの出力指示により、当該保持しておいた画像を印刷したり、外部装置に送信したりすることができる。このようにデジタル複合機に格納されている画像を印刷や送信などの出力するように指示されたときに生じるジョブを出力ジョブと称する。また、デジタル複合機に格納されている画像を出力する出力ジョブが実行された場合、画像を複合機に格納する際(入力ジョブの実行の際)に追跡用画像データが生成されてデータベースサーバに格納済みである。したがって、出力ジョブの実行時には、追跡用画像データは生成せずに、ログ情報のみを生成し、データベースサーバに格納されている入力ジョブの追跡用画像データ及びログ情報と関連付けて保存するようにするようにしてもよい。

【0039】

サービス通信部404は、HTTPプロトコル上のSOAPプロトコルにより、ネットワークI/F405を介して、画像処理サーバ102内のデータ受信部601に対して要

求コマンドを送信し、その応答を受信する。ネットワーク I / F 4 0 5 は、ネットワーク 1 0 9 に接続し、TCP / IP などのネットワークプロトコルの処理を行う。

【 0 0 4 0 】

図 5 は、ユーザ PC 1 0 4 における印刷処理の一構成を示す図である。アプリケーション 5 0 1 、グラフィックエンジン 5 0 2 、プリンタドライバ 5 0 3 、およびシステムスプーラ 5 0 4 は、ユーザ PC の HDD に保存されたプログラムとして存在する。これらは OS などのモジュールによって RAM 2 0 3 にロードされて実行されるプログラムモジュールである。

【 0 0 4 1 】

また、アプリケーション 5 0 1 およびプリンタドライバ 5 0 3 は、不図示の FD や CD-ROM 、或いはネットワークを経由して HDD 2 0 4 に追加することが可能となっている。HDD 2 0 4 に保存されているアプリケーション 5 0 1 は RAM 2 0 3 にロードされて実行される。このアプリケーション 5 0 1 からプリンタ 1 0 7 に対して印刷を行う際には、同様に RAM 2 0 3 にロードされ実行可能となっているグラフィックエンジン 5 0 2 を利用して出力（描画）を行う。

【 0 0 4 2 】

グラフィックエンジン 5 0 2 は、印刷装置（プリンタ等）ごとに用意されたプリンタドライバ 5 0 3 を HDD 2 0 4 から RAM 2 0 3 にロードし、アプリケーション 5 0 1 から出力されたデータをプリンタドライバ 5 0 3 に渡す。なお、グラフィックエンジン 5 0 2 はアプリケーション 5 0 1 から受け取る GDI (Graphic Device Interface) 関数を DDI (Device Driver Interface) 関数に変換して、プリンタドライバ 5 0 3 へ出力する。

【 0 0 4 3 】

プリンタドライバ 5 0 3 は、グラフィックエンジン 2 0 2 から受け取った DDI 関数に基づいて、プリンタが認識可能な制御コマンド、例えば PDL (Page Description Language) に変換する。変換されたプリンタ制御コマンドは、システムスプーラ 5 0 4 を経てプリンタ 1 0 7 へ印刷データとして出力される仕組みとなっている。

【 0 0 4 4 】

本実施形態の印刷システムでは、プリンタドライバ 5 0 3 内に電子データ生成部 5 0 5 を有する。なお、電子データ生成部 5 0 5 はプリンタドライバ 5 0 3 のビルドインモジュールであってもよいし、個別のインストーレーションによって追加されるライブラリモジュールの形式であっても構わない。プリンタドライバ 5 0 3 では、電子データ生成部 5 0 5 により追跡用データの抽出・生成を行い、電子データ転送部 5 0 6 へ送る。ここで、追跡用データとして格納されるデータは、画像データ、テキストデータ、ログ情報である。また、電子データ生成部 5 0 5 はアプリケーション 5 0 1 に応じて、該 DDI 関数からテキストデータを抽出するかどうか判断することができる。

【 0 0 4 5 】

電子データ転送部 5 0 6 は追跡用データを受け取り、これを画像処理サーバ 1 0 2 へ転送する。なお電子データ転送部 5 0 6 は受け取った追跡用データを必要に応じて加工したり、選別したりしてもよい。また電子データ転送部は追跡用データを受け取ると同時に画像処理サーバ 1 0 2 へ送信しても良いし、ハードディスク等に一時保存し、別途指定される送信スケジュールに従い画像処理サーバ 1 0 2 へ送信しても良い。

【 0 0 4 6 】

図 6 は画像処理サーバ 1 0 2 の構成と、データベースサーバ 1 0 3 内に保存されるデータ構成とを示す模式図である。

【 0 0 4 7 】

データ受信部 6 0 1 は、デジタル複合機 1 0 1 やユーザ PC 1 0 4 から転送されてきた画像データ、テキストデータ、及びログ情報を受信し、画像処理部 6 0 2 へ渡す。

【 0 0 4 8 】

10

20

30

40

50

画像処理部 602 は、受け取った画像データをデータベースサーバ 103 で格納できる形式に画像変換を行う。画像変換が必要ない場合には何も行わない。なお、画像データは 1 つのジョブに対して複数存在してもよい。それらは形式として PDF 形式のように 1 つにまとめてしまっても良いし、JPEG 形式などで複数に分かれて持っていても良く、最終的にデータベースサーバ 103 においてジョブ単位で認識できれば良い。本実施例ではわかりやすさのため 1 つのジョブに対し画像データ 1 つの場合について説明する。

【0049】

また、画像処理部 602 では、画像データに対応するテキストデータが存在するかどうか調べ、テキストデータのない画像データについては OCR 处理部 603 へ渡す。一方、テキストデータがある画像データについては、データ登録部 604 へ画像データとテキストデータを渡す。10

【0050】

OCR 处理部 603 では、画像データに OCR をかけてテキストデータを抽出し、画像データとテキストデータとをデータ登録部 604 に渡す。

【0051】

データ登録部 604 では、画像登録部 605 とテキスト登録部 606 とログ情報登録部 607 とから構成される。画像登録部 605 では、画像データをファイルとしてデータベースサーバ 103 へ登録する。テキスト登録部 606 では、テキストデータを全文検索用のテキストとしてデータベースサーバ 103 へ登録する。ログ情報登録部 607 では、ジョブの内容を示すログ情報をデータベースサーバ 103 へ登録する。20

【0052】

611 は、データ登録部 604 によって登録された画像データである。この画像データの持ち方は使用するデータベースシステムに応じて適宜定められるものとするが、例えば、各画像データに対して画像 ID が割り振られ、テーブル等で管理される。

【0053】

612 は、ジョブ管理テーブルであり、ジョブを識別する ID、ジョブに対応する画像データ ID、ジョブに対応するテキストデータからなる。このテキストデータは検索用インデックスをつけることによって全文検索に用いられる。前述したテキスト登録部 606 は、ジョブ管理テーブル 612 のテキスト部分に該テキストデータを登録する。なお、ジョブ管理テーブル 612 の属性はこれだけに限定されるものではなく、ジョブに関するその他の属性を持っていても問題はない。30

【0054】

またデータベースサーバ 103 は、ジョブに対するログ情報テーブル 613 を用いて、どのような処理が為されたかなどのログを管理しているが、本実施例では詳細な説明は省略する。

【0055】

図 7 は、本発明の実施の一形態に係るシステム管理設定を行う設定画面 UI の一例である。この UI は、例えば画像処理サーバ 102 やデータベースサーバ 103 上で動作しても良いし、或いは別の PC 上で動作しても良い。設定ウィンドウ 701 は、スタンドアロンアプリケーションや Web アプリケーションなどにより表示されるものであり、その実施形態は特に限定されない。このシステム設定では、例えば、デジタル複合機 101 やユーザ PC 104 が画像処理サーバ 102 へデータを送信するスケジュールや、画像の解像度、圧縮率、ビット数などを設定することができる。更に、システム設定 701 では、印刷を実行するアプリケーションに応じてプリンタドライバ 503 がテキストを抽出するか否かを設定することができる。本実施例では、プリンタドライバ 503 が文字印刷命令からテキストデータを抽出することをデフォルト設定としているが、アプリケーションの種類によってはテキスト抽出を行わないように設定できるようにしている。図 7 は、プリンタドライバ 503 でテキストを抽出せずに、画像処理サーバ 102 で画像データに対し OCR をかけてテキストを抽出するアプリケーションを例外アプリケーション（以後、例外アプリ）として、設定する UI である。システム設定ウィンドウ内の例外アプリケーショ4050

ン設定 702において、例外アプリを設定する。テキストボックス 703に表示されているアプリケーションが例外アプリとして設定されている。ここで設定した例外アプリは、全てのユーザ PC にインストールされているプリントドライバ 503の電子データ生成部へ設定が通知され、次の印刷処理時から設定が有効になる。通知された例外アプリのリストは、ユーザ PC 内でファイルとして持っても良いしレジストリに保存しても良い。なお、該設定はユーザ PC ごとやプリントドライバごとに設定を行えるようにしてよい。

【0056】

図 8 は、本実施例のユーザ PC 104における印刷ジョブ実行時に、追跡用データを保存する動作のフロー図である。

【0057】

ユーザにより印刷が指示されると、プリントドライバ 503では印刷処理を行うとともに、電子データ生成部 505で、アプリケーション 501のアプリケーション名（印刷実行モジュール名）を取得する（S801）。

【0058】

次に、電子データ生成部 505は、前述したように設定されている例外アプリ名を取得する（S802）。複数登録されている場合はすべて取得する。

【0059】

次に、電子データ生成部 505は、S801で取得した印刷実行アプリ名と S802で取得した例外アプリ名が等しいか判断する（S804）。印刷実行アプリ名と等しい例外アプリが見つかった場合はループを抜け、画像の抽出・生成を行う（S807）。S805では全ての例外アプリ名との比較が終わったか判断し、終わっていない場合は S803に戻る。一方、すべての例外アプリ名について比較したが、等しい例外アプリがなかった場合は S806に進んで、電子データ生成部 505は文字描画命令（印刷命令）からテキストデータを抽出する。その後、ログ画像（追跡用画像データ）の抽出・生成を行う（S807）。

【0060】

ステップ S808では、印刷時のログ情報（いつ、誰が、どこで印刷処理を実行させたかなどの情報）を、電子データ生成部 505で抽出する。

【0061】

ステップ S809において、電子データ生成部は、抽出した各データを内部記憶装置に一時保存して、予め設定されたスケジュールに従って該データを画像処理サーバ 102へ転送する。（即時送する設定が為されている場合は、抽出したデータを電子データ転送部 506に即渡して、電子データ転送部 506は画像処理サーバへ転送する。）

このように、印刷ジョブが実行されると、当該ジョブに関する情報（印刷された画像を確認するための追跡用画像データ、ログ情報、テキストデータ）が画像処理サーバを介してデータベースサーバに格納されることになる。したがって、後にどのようなジョブが実行されたのか検証することができ、情報が漏洩した場合でもどのジョブが原因であるか追跡しやすくなる。また、アプリケーションの種類に応じて、テキストデータの抽出方法を変えることができるので、より正確なテキストデータを抽出できるようになる。

【0062】

図 9 は、本実施例のデジタル複合機 101におけるジョブ実行時の動作フロー図である。なお、このフロー図ではジョブの実行に伴って、追跡用データが抽出されてデータベースに保存される処理に関する説明を行う。

【0063】

デジタル複合機に画像が入力されるジョブ（スキャンなどの入力ジョブ）が実行された場合、ジョブコントロールメイン部 403は該ジョブの追跡用の画像データを抽出・生成する（S901）。

【0064】

ステップ S902では、ジョブコントロールメイン部 403において、ジョブ実行時のログ情報（いつ、誰が、どこでそのジョブを実行させたかを示す情報）を抽出する。こ

10

20

30

40

50

こでは、複合機に画像が入力されるジョブが実行された場合も、複合機に格納されている画像を出力するジョブが実行された場合も、ログ情報が抽出される。

【0065】

ステップS903では、ジョブコントロールメイン部403で抽出したデータ（追跡用画像データ、ログ情報）を内部記憶装置に一時保存しておき、予め設定されたスケジュールに従って画像処理サーバ102へ転送する。なお、即時転送する設定がなされている場合は、即時転送するようにしてもよい。また、出力ジョブである場合は、ログ情報を入力ジョブのIDとともに、画像処理サーバに転送する。

【0066】

このようにデジタル複合機で処理したジョブの情報も画像処理サーバを介してデータベースサーバに格納するようにしたので、後日、検証することが可能になる。なお、図9の例では、デジタル複合機で処理したジョブに対する処理について説明したが、デジタル複合機での処理に限るものではない。例えば、プリンタやFAXなどのその他の画像処理装置で実行するジョブに対しても同様の処理を行えばよい。

10

【0067】

図10は本実施例の画像処理サーバ102の処理フロー図である。

【0068】

ステップS1001では、データ受信部601において、デジタル複合機101やユーザPC104から転送されてきたデータ（追跡用画像データ、テキストデータ、ログ情報）を受信する。

20

【0069】

次に、S1002では、画像処理部602において、データ受信部601で受け取った追跡用画像データを、データベースサーバ102に登録できる画像形式に変換する。

【0070】

ステップS1003では、画像処理部602において、画像データに対応するテキストデータを受信したか判定する。受信していないければ、ステップS1004において、画像データに対してOCR（光学的文書認識）処理をかけてテキストを抽出し、データ登録部604へ渡す。一方、S1003においてテキストデータがあると判断された場合には、そのまま画像データとテキストデータをデータ登録部604へ渡す。なお、S1002の画像処理とS1004のOCR処理の順序は、逆であっても良い。

30

【0071】

ステップS1005では、画像登録部605が、追跡用の画像データをデータベースサーバ103へ登録する。

【0072】

ステップS1006では、テキスト登録部606が、受信したテキストデータ或いはOCR抽出したテキストデータを、S1005で登録した画像データに対応する全文検索用のテキストとして、データベースサーバ103のジョブ管理テーブル612へ登録する。

【0073】

ステップS1007では、ログ情報登録部607が、受信したログ情報をデータベースサーバ103のジョブ管理テーブル612へ登録する。

40

【0074】

このように、テキストデータが送信されてこなかった場合、画像データに対してOCR処理を行ってデータベースサーバに登録するようにしたので、検索を行いやすくなる。また、テキストデータが送信されてきた場合はOCR処理を行わないので、冗長にテキストデータを生成することがなくなる。

【0075】

なお、追跡用画像データに対して解像度変換等の画像変換処理を行って保存する場合、本発明の要旨を逸脱しない範囲内において、デジタル複合機101、ユーザPC104、画像処理サーバ102で適宜実行するようにしてもよい。

【0076】

50

以上、本発明の実施の一形態について説明したが、上記構成図や動作フローは一例であり、上記に限定されるものではない。

【0077】

尚、本発明は、前述した実施形態の機能を実現するソフトウェアのプログラム（実施形態では図に示すフローチャートに対応したプログラム）を、システムあるいは装置に直接あるいは遠隔から供給する。そして、そのシステムあるいは装置のコンピュータが該供給されたプログラムコードを読み出して実行することによっても達成される場合を含む。従って、本発明の機能処理をコンピュータで実現するために、該コンピュータにインストールされるプログラムコード自体も本発明を実現するものである。つまり、本発明は、本発明の機能処理を実現するためのコンピュータプログラム自体も含まれる。

10

【0078】

その場合、プログラムの機能を有していれば、オブジェクトコード、インタプリタにより実行されるプログラム、OSに供給するスクリプトデータ等の形態であっても良い。

【0079】

プログラムを供給するための記録媒体としては、例えば、磁気ディスク、光ディスク、光磁気ディスク、磁気テープ、不揮発性のメモリカードなどがある。

【0080】

その他、プログラムの供給方法としては、クライアントコンピュータのブラウザを用いてインターネットからハードディスク等の記録媒体にダウンロードすることによっても供給できる。すなわち、インターネットから本発明のコンピュータプログラムそのもの、もしくは圧縮され自動インストール機能を含むファイルをダウンロードする。また、本発明のプログラムを構成するプログラムコードを複数のファイルに分割し、それぞれのファイルを異なるサーバからダウンロードすることによっても実現可能である。つまり、本発明の機能処理をコンピュータで実現するためのプログラムファイルを複数のユーザに対してダウンロードさせるサーバも、本発明に含まれるものである。

20

【0081】

また、コンピュータが、読み出したプログラムを実行することによって、前述した実施形態の機能が実現される。その他にも、そのプログラムの指示に基づき、コンピュータ上で稼動しているOSなどが、実際の処理の一部または全部を行い、その処理によっても前述した実施形態の機能が実現され得る。

30

【0082】

さらに、記録媒体から読み出されたプログラムが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書き込まれた後にも前述した実施形態の機能が実現される。すなわち、そのプログラムの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行うことによっても前述した実施形態の機能が実現される。

【図面の簡単な説明】

【0083】

【図1】システム構成図

40

【図2】各PCの構成図

【図3】デジタル複合機の構成図

【図4】デジタル複合機のソフトウェア機能構成図

【図5】ユーザPCのソフトウェア機能構成図

【図6】画像処理サーバの構成図とデータベースサーバのデータの保持

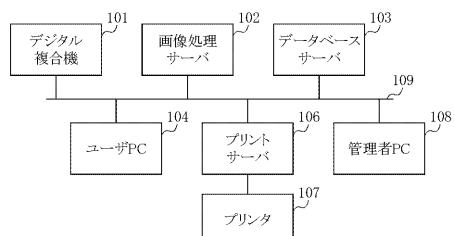
【図7】システム設定UI

【図8】ユーザPC104の印刷実行時におけるフロー

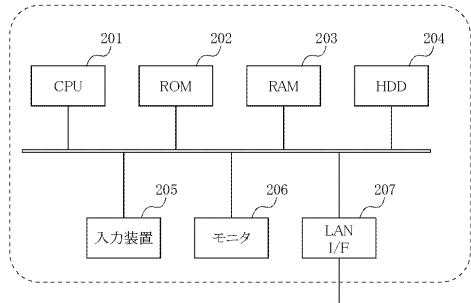
【図9】デジタル複合機のジョブ実行時におけるフロー

【図10】画像処理サーバの画像処理からデータ登録までのフロー

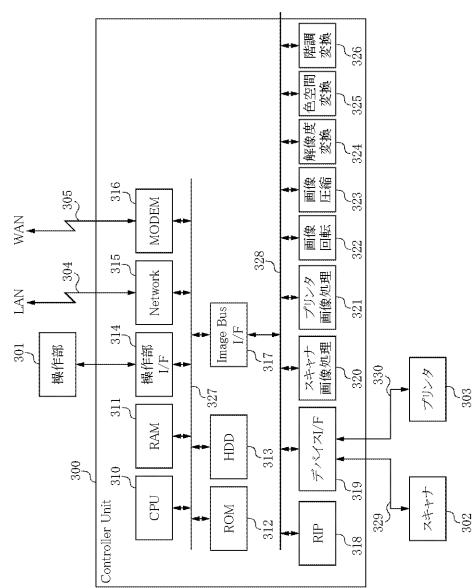
【図1】



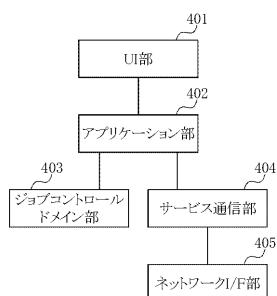
【図2】



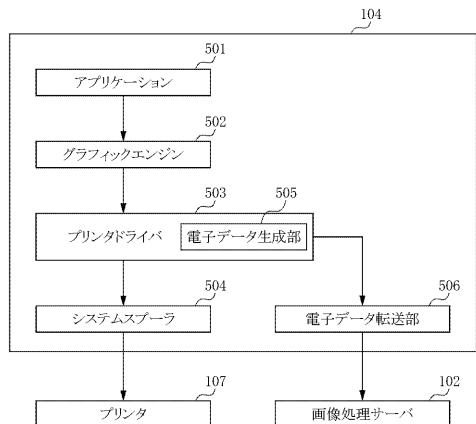
【図3】



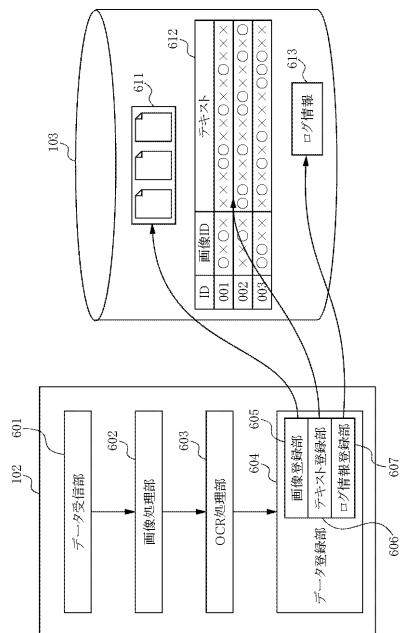
【図4】



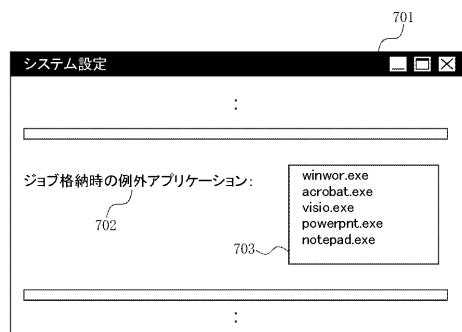
【図5】



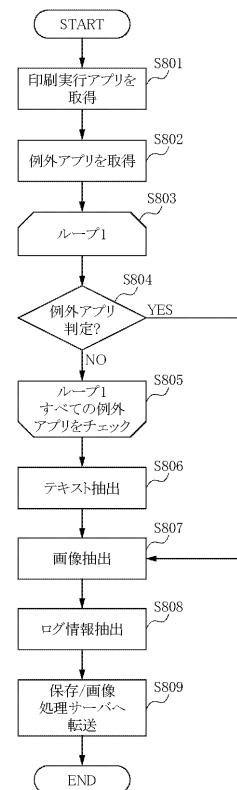
【図6】



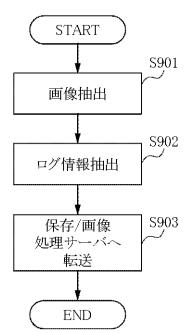
【図7】



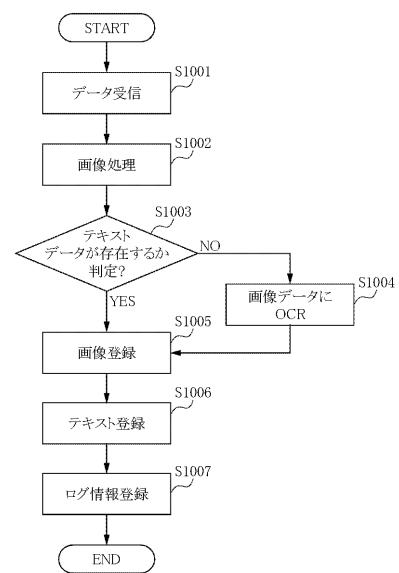
【図8】



【図9】



【図10】



フロントページの続き

(56)参考文献 特開平11-120202(JP,A)
特開2002-149371(JP,A)
特開2004-334341(JP,A)
特開2000-335059(JP,A)
特開2003-016070(JP,A)
特開2002-230016(JP,A)
特開2005-267057(JP,A)
特開2005-157568(JP,A)

(58)調査した分野(Int.Cl., DB名)

G 06 F	3 / 12
H 04 N	1 / 00