

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5084798号  
(P5084798)

(45) 発行日 平成24年11月28日(2012.11.28)

(24) 登録日 平成24年9月14日(2012.9.14)

(51) Int.Cl.	F I		
<b>G06F 3/048 (2006.01)</b>	G06F	3/048	653A
<b>G06F 3/14 (2006.01)</b>	G06F	3/14	330A
<b>G06F 17/30 (2006.01)</b>	G06F	17/30	320Z
<b>G06F 11/30 (2006.01)</b>	G06F	11/30	320E
	G06F	17/30	170B

請求項の数 5 (全 9 頁)

(21) 出願番号	特願2009-193015 (P2009-193015)	(73) 特許権者	000004226
(22) 出願日	平成21年8月24日(2009.8.24)		日本電信電話株式会社
(65) 公開番号	特開2011-44074 (P2011-44074A)		東京都千代田区大手町二丁目3番1号
(43) 公開日	平成23年3月3日(2011.3.3)	(74) 代理人	100069981
審査請求日	平成23年9月5日(2011.9.5)		弁理士 吉田 精孝
		(74) 代理人	100087860
			弁理士 長内 行雄
		(74) 代理人	100166224
			弁理士 角田 成夫
		(72) 発明者	西川 健一
			東京都千代田区大手町二丁目3番1号 日
			本電信電話株式会社内
		(72) 発明者	川端 祐人
			東京都千代田区大手町二丁目3番1号 日
			本電信電話株式会社内

最終頁に続く

(54) 【発明の名称】 アプリケーション状態認識方法、装置及びプログラム

(57) 【特許請求の範囲】

【請求項1】

コンピュータ上で動作するアプリケーションの動作状態をアプリケーション状態認識装置が認識する方法において、

アプリケーション状態認識装置の画像データ取得手段が、前記アプリケーションが表示装置に出力している情報を画像データとして取得するステップと、

アプリケーション状態認識装置の前処理手段が、取得された画像データに対して該画像データの可変領域を所定の固定値又は固定パターンに置換する処理を含む前処理を行うステップと、

アプリケーション状態特徴量抽出手段が、前処理された画像データの特徴量を抽出するステップと、

アプリケーション状態の識別手段が、識別辞書として記憶手段に予め記憶されている、画像データの特徴量からアプリケーションの動作状態を識別するための学習データと、前記画像データの特徴量とに基づきアプリケーションの動作状態を識別するステップとを備えた

ことを特徴とするアプリケーション状態認識方法。

【請求項2】

前記識別辞書には、画像データに対して該画像データの可変領域を所定の固定値又は固定パターンに置換したものの特徴量と置換していないものの特徴量の双方について学習データを記憶しておく

ことを特徴とする請求項 1 記載のアプリケーション状態認識方法。

【請求項 3】

前記置換有りの学習データには可変領域の座標情報を含み、

前記前処理ステップでは、識別辞書の学習データを参照して可変領域の座標情報を取得して該座標情報で示された領域を所定の固定値又は固定パターンに置換する

ことを特徴とする請求項 2 記載のアプリケーション状態認識方法。

【請求項 4】

コンピュータ上で動作するアプリケーションの動作状態を認識するアプリケーション状態認識装置において、

前記アプリケーションが表示装置に出力している情報を画像データとして取得する画像データ取得手段と、

取得された画像データに対して該画像データの可変領域を所定の固定値又は固定パターンに置換する処理を含む前処理を行う前処理手段と、

前処理された画像データの特徴量を抽出する特徴量抽出手段と、

画像データの特徴量からアプリケーションの動作状態を識別するための学習データを識別辞書として予め記憶した記憶手段と、

抽出した特徴量に基づき前記識別辞書からアプリケーションの動作状態を識別する識別手段とを備えた

ことを特徴とするアプリケーション状態認識装置。

【請求項 5】

コンピュータを、請求項 4 記載のアプリケーション状態認識装置の各手段として機能させる

ことを特徴とするアプリケーション状態認識プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、コンピュータ上で動作するアプリケーション（AP）を通じて操作を行うユーザに対して、ユーザが操作を円滑に行うことを支援する技術に関するものである。

【背景技術】

【0002】

ユーザによる操作を支援するには、現在のユーザやアプリケーションの状態を知る必要がある。現状態を知らないと、状態に依らずに適用可能な汎用的な支援に留まり、個別の状況に適合した具体的で個別的な支援を行うことができないためである。

【0003】

ユーザやアプリケーションの状態を知る従来手法として、アプリケーションプログラムやOSが提供する情報を利用する方法がある。例えば、Microsoft（登録商標）社製OSであるWindows（登録商標）では、「AutoIt」や「Spy++」というツールが知られている。これらのツールでは、OS上でのウィンドウの基本的な情報（例えばタイトル、クラス、サイズなど）の取得、ウィンドウやプロセス等とOSとの間でやりとりされるメッセージの取得などを行うことができる。そこで、これらの情報を複

【先行技術文献】

【非特許文献】

【0004】

【非特許文献 1】"AutoIt Script Home Page"、[online]、[平成 21 年 8 月 13 日検索]、URL<<http://www.autoitscript.com/>>

【非特許文献 2】"Home Page: Spy++"、[online]、[平成 21 年 8 月 13 日検索]、URL<[http://msdn.microsoft.com/en-us/library/aa264396\(VS.60\).aspx](http://msdn.microsoft.com/en-us/library/aa264396(VS.60).aspx)>

【発明の概要】

10

20

30

40

50

**【発明が解決しようとする課題】****【0005】**

しかし上記の方法では、ユーザの操作する環境によっては、ユーザ支援に必要な情報を充分に取得することができないという問題がある。例えば、セキュリティ強化の目的で導入されることが多くなっているシンクライアント環境では、多くの場合非常に限られた情報しか取得することができない。シンクライアントの実装方式として比較的ポピュラーな画面情報を転送する方式では、従来手法を使うとアプリケーションの充分な情報を取得することができない。

**【0006】**

本発明は、上記事情に鑑みてなされたものであり、その目的とするところは、従来手法ではアプリケーションやOSから充分な情報を取得することができないような環境下であっても、アプリケーションの動作状態を認識することが可能なシステムを提供することにある。

10

**【課題を解決するための手段】****【0007】**

上記目的を達成するために、本発明は、コンピュータ上で動作するアプリケーションの動作状態を認識するアプリケーション状態認識装置において、前記アプリケーションが表示装置に出力している情報を画像データとして取得する画像データ取得手段と、取得された画像データに対してノイズ除去や正規化等の前処理を行う前処理手段と、前処理された画像データの特徴量を抽出する特徴量抽出手段と、画像データの特徴量からアプリケーションの動作状態を識別するための学習データを識別辞書として予め記憶した記憶手段と、抽出した特徴量に基づき前記識別辞書からアプリケーションの動作状態を識別する識別手段とを備えたことを特徴とする。

20

**【0008】**

本発明によれば、アプリケーションの動作により表示装置に画面表示される画像データに対して画像認識を行うことにより該アプリケーションの動作状態を識別するので、従来手法のようにカメラ等の特別なデバイスを別途用意することや複雑な初期設定を行うことなどは必要ない。したがって汎用的な技術として種々の環境下でアプリケーションの動作状態を取得することができる。

**【0009】**

また、本発明の好適な態様としては、前記前処理部において、取得された画像データに対して該画像データの可変領域を所定の固定値又は固定パターンに置換することを特徴するものが挙げられる。すなわち、単純に画像データに対する画像認識を行っただけでは細かな動作状態による違いを検出しすぎて誤認識しがちなところ、具体的には入力画像と識別辞書中の画像において動作状態によって可変な画像部分（可変領域）を、一律的なパターンで塗り潰してから認識させる。これにより、データの入力状態などにより生じる画像上の違いを無視したアプリケーションの動作状態を、より確実に認識することを可能にする。

30

**【発明の効果】****【0010】**

以上説明したように本発明によれば、特別な機器や設定を必要とせずに、画面上の画像データに対する画像認識により、アプリケーションの現在の動作状態を把握することができる。また、従来技術ではアプリケーションの動作状態の違いが画像上の違いに表れることで誤認識しやすいところを、前処理段階において無視してよい状態による画像状態の違いを除去することで、より確実な認識を可能にした。これにより、ユーザによる次操作の提案やエラー状態からの回復、自律動作など様々な応用が可能になり、ユーザによるアプリケーション操作をより強力に支援できるようになる。

40

**【図面の簡単な説明】****【0011】**

**【図1】** アプリケーション状態認識システムの構成図

50

【図2】アプリケーション状態認識システムの動作を説明するフローチャート

【図3】ユーザ支援アプリケーションの動作を説明するフローチャート

【図4】可変領域の置換処理を説明する図

【図5】可変領域の置換処理を説明する図

【発明を実施するための形態】

【0012】

本発明の一実施の形態について図面を参照して説明する。図1は、本発明のアプリケーション状態認識システムの構成図である。本実施の形態に係るアプリケーション状態認識システムでは、ユーザ支援アプリケーションに対して他のアプリケーションの動作状態を認識・提供するサービスを行う。すなわち、図1に示すように、コンピュータ10にはOS11上で複数の既存アプリケーションAP1, AP2, ...が動作している。各既存アプリケーションAP1, AP2は、ユーザ1に対するユーザインタフェースとして、キーボードやマウス等の入力デバイス15及び表示装置であるディスプレイ16を利用する。

10

【0013】

ユーザ支援アプリケーション20は、既存アプリケーションAP1, AP2, ...を操作するユーザ1に対して、その操作を支援するアプリケーションである。該ユーザ支援アプリケーション20には様々な形態が考えられ、例えば、現在の画面上の既存アプリケーションAP1, AP2, ...の状態をもとに次操作を提案するものや、ユーザ1がエラー状態と判断される状態に陥った時に回復手順を示し回復支援をするもの、自動操作システムに状況を判断しつつ自律的な動作を行う機能を追加するものなど挙げられる。本発明に係るアプリケーション状態認識システム100は、該ユーザ支援アプリケーション20に対して既存アプリケーションAP1, AP2, ...の動作状態を認識し、該認識結果を提供するものである。

20

【0014】

以下にアプリケーション状態認識システム100について詳述する。本発明に係るアプリケーション状態認識システム100は、構成自体は一般的な画像認識システムと同等のものであり、図1に示すように、画像取得部110, 前処理部120, 特徴抽出部130, 識別部140, 識別辞書150とを備えている。また、本発明に係るアプリケーション状態認識システム100は、コンピュータとプログラムによっても実現でき、プログラムを記録媒体に記録することも、ネットワークを通して提供することも可能である。さらに、本発明に係るアプリケーション状態認識システム100やユーザ支援アプリケーション20の動作環境や実装形態は不問である。すなわち、既存アプリケーションAP1, AP2, ...が動作するコンピュータと同一コンピュータ上で動作してもよいし、例えばシンクライアント環境のように他のコンピュータ上で動作してもよい。さらに、アプリケーション状態認識システム100の各部も単一のコンピュータではなく複数のコンピュータに分散させて実装してもよい。

30

【0015】

画像取得部110は、OS11やそれに付随するシステムライブラリ等の提供する機能を用いてディスプレイ16上に表示されている、ウィンドウやスクリーン全体などを画像データとして取得する。

40

【0016】

前処理部120では、画像データに対して認識処理をしやすくするための処理を行う。具体的には、認識の際に余計なノイズなどをできるだけ除去するために行う二値化や画像サイズなどを一定の値に揃える正規化などの処理を行う。さらに、前処理部120では、必要に応じて画像データから特定される可変領域の置換処理を行う。可変領域及び置換処理については後述する。

【0017】

特徴抽出部130では、画像データを認識するために、該画像データをより少数の情報で表す特徴量を特徴ベクトルとして抽出する。特徴量を算出する方法については、様々な形で検討されている既存手法が豊富に存在するため、それらを利用する。具体的には、S

50

I F T 特徴量や H o G 特徴量、H a a r - l i k e 特徴量などの様々な特徴量が利用可能である。また、後述する識別部 1 4 0 でテンプレートマッチ法を利用する場合は、前処理部 1 2 0 で画像サイズの正規化など適切な前処理が行ってあれば、画像のデータをそのまま特徴量とする。

【 0 0 1 8 】

識別部 1 4 0 では、事前に学習した情報を保管するなどして作成した識別辞書 1 5 0 の情報（学習データ）を参照し、入力された画像の特徴ベクトルを元に入力画像の識別・分類を行う。識別・分類は、入力画像が、認識したい状態を示すクラスのいずれに属するかを決定する処理である。クラスは、区別して認識したい状態の数だけ予め用意しておく。通常、識別辞書 1 5 0 中の全てのデータをその特徴ベクトルを元に、幾つかの有限のクラスに分割する。その上で、クラスを代表する特徴ベクトル（別途算出する）や辞書中の全データの特徴ベクトルと、入力画像の特徴ベクトルを比較することで、入力画像の属するクラスを識別する。比較する際はユークリッド距離が最小となるベクトルを選ぶやり方やニューラルネットワークを利用するやり方、サポートベクタマシンを利用するやり方、テンプレートマッチ法を利用する方法など、様々な既存の手法があり、それらを利用する。

【 0 0 1 9 】

次に、図 2 を参照して本発明に係るアプリケーション状態認識システムの動作について説明する。

【 0 0 2 0 】

図 2 に示すように、まず画像取得部 1 1 0 が、サービス提供先のアプリケーション（本実施の形態ではユーザ支援アプリケーション 2 0 ）等からの要求に応じて、ディスプレイ 1 6 上に表示されている画像データを取得する（ステップ S 1 ）。なお、前記要求には必要に応じて、認識対象となる既存アプリケーション A P 1 , A P 2 , ... や更にそのウィンドウ等を指定するようにしてもよい。次に、前処理部 1 2 0 が前処理として、画像データに対して後段の識別処理における画像認識が行いやすくするよう画像処理を行う（ステップ S 2 ）。次に、特徴抽出部 1 3 0 が、前処理された画像データから特徴量として特徴ベクトルを抽出する（ステップ S 3 ）。そして、識別部 1 4 0 が、抽出された特徴量と識別辞書 1 5 0 中のデータを比較して、画像データの属するクラスを判別して、該クラスをアプリケーションの動作状態としてサービス提供先のアプリケーションに返す（ステップ S 4 ）。

【 0 0 2 1 】

次に、図 3 を参照して本発明に係るアプリケーション状態認識システムを用いたユーザ支援アプリケーションの動作について説明する。ここでは、ユーザ支援アプリケーション 2 0 として、既存アプリケーション A P 1 , A P 2 , ... を操作するユーザ 1 に対してその操作を支援するものを例にとって説明する。

【 0 0 2 2 】

図 3 に示すように、ユーザ支援アプリケーション 2 0 は、動作状態の認識対象である既存アプリケーション A P 1 , A P 2 , ... の画像領域（ウィンドウなど）を指定して、アプリケーション状態認識システム 1 0 0 に対して動作状態の認識を依頼する（ステップ S 1 1 ）。アプリケーション状態認識システム 1 0 0 は、図 2 を参照して前述した動作にしたがって既存アプリケーション A P 1 , A P 2 , ... の動作状態を識別・応答する（ステップ S 1 2 ）。ユーザ支援アプリケーション 2 0 は、取得した動作状態に基づき次操作の支援などのユーザ支援処理を行う（ステップ S 1 3 ）。なお、該ユーザ支援処理については従来周知のものと同様なので詳述は省略する。

【 0 0 2 3 】

ところで本願発明では既存アプリケーション A P 1 , A P 2 , ... の動作状態を把握する手段として、表示画面の画像データを画像認識技術により解析するという手法をとっている。しかし、単純に画面画像に対する画像認識を行っただけでは、テキストエリアへの入力状況などによる違いを過度に検出してしまい、誤認識の可能性が高くなるという課題がある。

10

20

30

40

50

## 【 0 0 2 4 】

そこで本願発明では、前処理部 1 2 0 において、必要に応じて可変領域を所定の固定値又は固定パターンに置換する処理を行う。ここで、「可変領域」とは、アプリケーションの動作状態が属するクラスは同じであるが、その動作状態によって内容が変わる可能性のある領域を意味する。例えば、住所入力画面における住所入力用のテキスト領域などが該当する。

## 【 0 0 2 5 】

具体的には前処理部 1 2 0 は、入力文字列の違いやラジオボタンなどの選択状況の違いなどを区別して認識したくない場合は、細かな状態の違いを区別したくない可変領域を一定の色やパターンで塗り潰す。その後、特徴量の計算を行うことで、文字列の違いなどの状態の違いを無視した確度の高い認識が可能になる。

10

## 【 0 0 2 6 】

さらに、テキスト入力中とテキスト未入力というような複数の状態を区別したい場合、なんらかのテキストが入力されている場合は、テキスト入力中を示す固定の文字列に置換するなどといった方法により、区別が可能である。可変領域に画像が描画されている場合でも同様に特定の画像に置換するなどの手法が適用できる。また、例えば未入力状態の画像における可変領域において、未入力を示す一定の固定パターンで塗り潰す必要があるケースもある。この手法を応用することで、任意の複数の状態を区別して認識することが可能である。

## 【 0 0 2 7 】

図 4 に前処理部における置換処理の一例を示す。図 4 は、あるアプリケーションの特定ウィンドウとして、テキスト領域、ラジオボタン、可変な図形領域を持つウィンドウに対して、入力状況に依らずにウィンドウの属するクラス（名前、住所など入力ウィンドウ）を識別するために、画像に施す前処理例の概要を示す図であり、処理を行う前と後とを示している。処理内容は、可変な領域を背景色などの固定色や固定パターンで塗り潰す処理である。

20

## 【 0 0 2 8 】

このような処理を施すためには、認識の精度を制御したい可変領域を指定する必要がある。一つの方法は、図 2 などで示した状態認識のフローと同様の画像の認識を行い、必要となる可変領域の認識を行う方法である。文字列の矩形領域の検出については、フーリエ記述子を用いた文字認識など、既存の手法を適用することが可能である。その他、一般の画像であっても、既存の S I F T 特徴量などを利用し、指定の可変領域を示す画像と同様の特徴量を持つ画像を検出することで、可変領域の認識が可能である。

30

## 【 0 0 2 9 】

ただし、上記の方法はどのような画像が現れるか予測ができないため、汎用的な適用には好ましいとは言えない。そこで、汎用的な方法の一つを図 5 を参照して説明する。図 5 は可変領域の検出例を図示したものである。

## 【 0 0 3 0 】

図 5 に示す検出方法では、文字領域や画像領域などの一般の可変領域の検出をするために、識別辞書 1 5 0 の学習データを作成する際に、学習対象の画像における可変領域を手動で指定し、可変領域を塗り潰して固定領域に変換し、特徴量を計算し、変換したものと変換していないもの双方を保存しておく。その際、変換した各固定領域（矩形であることを前提とする）のウィンドウや画面等の左上隅を原点として、左上隅と右下隅の座標を画像の特徴量と共に可変領域を示す座標として記録しておく。前処理部 1 2 0 では、入力画像に対して、識別辞書 1 5 0 に記録しておいた各データの可変領域の座標で表される領域を固定領域にそれぞれ置換する。特徴抽出部 1 3 0 は、置換後の各画像データについてそれぞれ特徴量を算出する。そして識別部 1 4 0 は、各特徴量と識別辞書 1 5 0 中の各データとを比較し、最も類似した特徴量が示す動作状態を識別結果とする。但し、この方法を適用する条件としては、画像の大きさが学習時と認識時で同じである、という条件がある。

40

50

【0031】

他に、より汎用的に適用できる方法としては、学習用の画像だけでなく、入力画像に対しても、可変領域を手動で指定し可変領域を塗り潰して固定領域に変換してから特徴量を計算し、認識させるという方法がある。

【0032】

このように本発明に係るアプリケーション状態認識システムによれば、既存アプリケーションAP1, AP2, ...が表示する表示画面の画像データに基づき当該アプリケーションの動作状態を認識するので、特別な機器や設定を必要とすることなく種々の動作環境下での認識処理が可能となる。また、従来技術ではアプリケーションの動作状態の違いが画像上の違いに表れることで誤認識しやすいところを、前処理段階において無視してよい状態による画像状態の違いを除去することで、より確実な認識を可能にした。これにより、ユーザによる次操作の提案やエラー状態からの回復、自律動作など様々な応用が可能になり、ユーザによるアプリケーション操作をより強力に支援できるようになる。

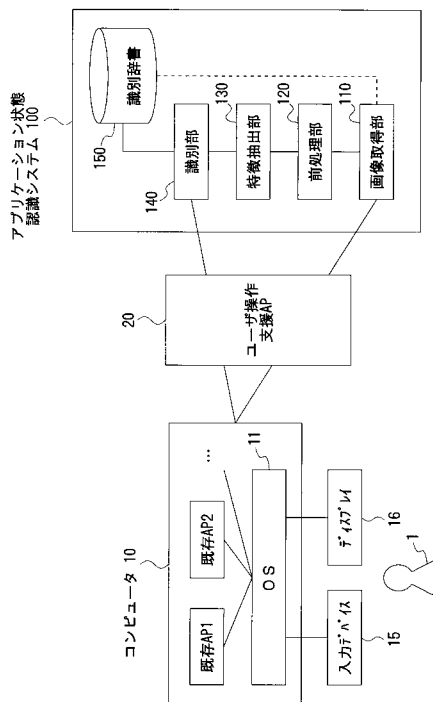
10

【符号の説明】

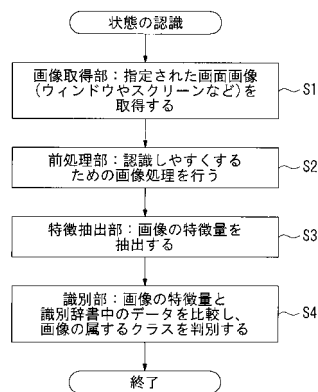
【0033】

1...ユーザ、10...コンピュータ、11...OS、15...入力デバイス、16...ディスプレイ、20...ユーザ支援アプリケーション、100...アプリケーション状態認識システム、110...画像取得部、120...前処理部、130...特徴抽出部、140...識別部、150...識別辞書。

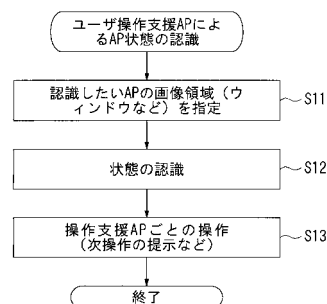
【図1】



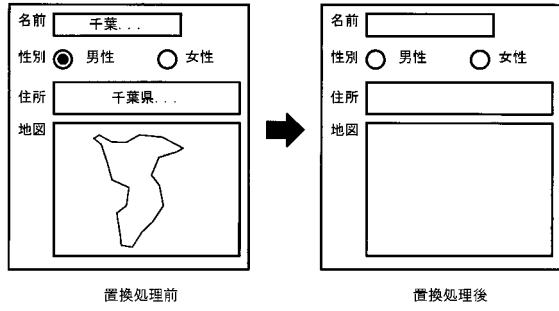
【図2】



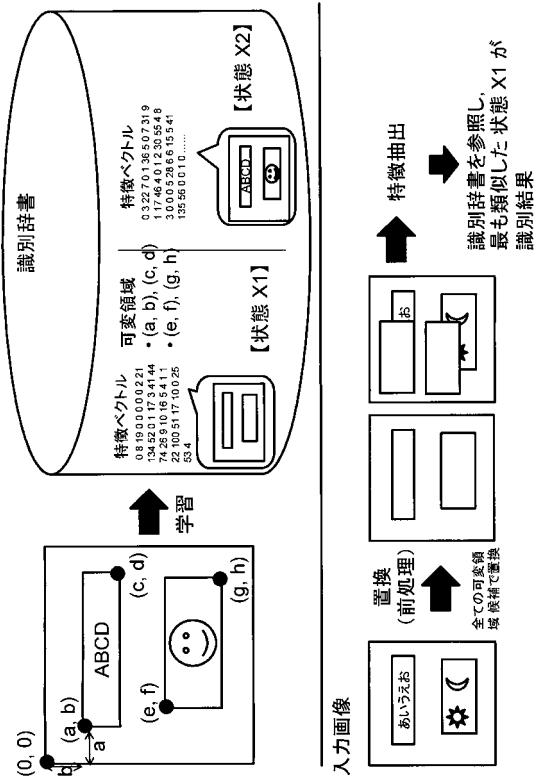
【図3】



【図4】



【図5】





---

フロントページの続き

(72)発明者 増田 健

東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内

審査官 森田 充功

(56)参考文献 特開2002-171506(JP,A)

特開2008-146387(JP,A)

特開2009-140268(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 3/14

G06F 11/30

G06F 17/30

G06F 3/048