

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5861539号
(P5861539)

(45) 発行日 平成28年2月16日 (2016. 2. 16)

(24) 登録日 平成28年1月8日 (2016. 1. 8)

(51) Int. Cl.	F I
G06T 7/00 (2006.01)	G06T 7/00 300F
G06T 1/00 (2006.01)	G06T 1/00 200E
G06F 17/30 (2006.01)	G06F 17/30 170B
	G06F 17/30 220B

請求項の数 10 (全 21 頁)

(21) 出願番号	特願2012-77535 (P2012-77535)	(73) 特許権者	000005223
(22) 出願日	平成24年3月29日 (2012. 3. 29)		富士通株式会社
(65) 公開番号	特開2012-216207 (P2012-216207A)		神奈川県川崎市中原区上小田中4丁目1番1号
(43) 公開日	平成24年11月8日 (2012. 11. 8)	(74) 代理人	100070150
審査請求日	平成27年1月6日 (2015. 1. 6)		弁理士 伊東 忠彦
(31) 優先権主張番号	201110085705.4	(74) 代理人	100146776
(32) 優先日	平成23年3月31日 (2011. 3. 31)		弁理士 山口 昭則
(33) 優先権主張国	中国 (CN)	(72) 発明者	ツァオ・チョン
			中国, 100025, ペイジン, チャオヤン ディストリクト, ジョオン ロード, ドン ス ホアヌ ナンバー56, オーシャン インターナショナル センター, タワー エイ 15エフ 富士通研究開発中心有限公司内
			最終頁に続く

(54) 【発明の名称】 複数の画像の記述的情報を取得する方法及び装置、並びに画像マッチング方法

(57) 【特許請求の範囲】

【請求項 1】

複数の画像の記述的情報を取得する方法であって、

前記複数の画像の各画像に関して特徴点検出を実行し、各画像の複数の特徴点を取得する検出ステップと、

前記複数の画像の第0レベルの記述的情報を取得する第1取得ステップであり、前記複数の画像の各画像の第0レベルの記述的情報が、該画像の前記複数の特徴点に基づいて得られる該画像の複数の第1ローカル特徴記述子と、前記複数の画像の前記複数の第1ローカル特徴記述子に基づいて得られる複数の第1ビジュアルワードとに従って生成される、第1取得ステップと、

Kは正の整数であるとして、 $n = 2, 3, \dots, K + 1$ のそれぞれについて、前記複数の画像の第 $(n - 1)$ レベルの記述的情報を取得する第n取得ステップであり、各画像に関して、該画像の $(n - 1)$ 回目の分割を実行し、該画像の複数の第 $(n - 1)$ レベルサブ画像を取得する第 $(n - 1)$ 分割サブステップ、及び該画像の第 $(n - 1)$ レベルの記述的情報が、該画像の前記複数の特徴点に基づいて得られる該画像の複数の第nローカル特徴記述子と、前記複数の画像の前記複数の第nローカル特徴記述子に基づいて得られる複数の第nビジュアルワードとに従って生成され、該画像の前記第nローカル特徴記述子は、それぞれの前記特徴点が位置する対応する第 $(n - 1)$ レベルサブ画像に従って得られる、第n取得サブステップ、

10

20

が実行される、第 n 取得ステップと、
を有する方法。

【請求項 2】

各画像の前記第 0 レベルの記述的情報の生成は、

該画像の前記複数の特徴点の各々に関して 1 つの第 1 ローカル特徴記述子を計算し、該画像の前記複数の第 1 ローカル特徴記述子を取得する第 1 計算サブステップと、

該画像の前記複数の第 1 ローカル特徴記述子をそれぞれ、前記複数の第 1 ビジュアルワードのうちの対応する第 1 ビジュアルワードに対応付け、前記複数の第 1 ビジュアルワードの各々の該画像内での出現頻度の統計を介して該画像の前記第 0 レベルの記述的情報を生成する第 1 記述的情報生成サブステップと、

を有する、請求項 1 に記載の方法。

【請求項 3】

各画像に関する前記第 n 取得サブステップは、

該画像の前記複数の特徴点の各特徴点に関して 1 つの第 n ローカル特徴記述子を、該特徴点が位置する第 $(n - 1)$ レベルサブ画像に基づいて計算し、該画像の前記複数の第 n ローカル特徴記述子を取得する第 n 計算サブステップと、

該画像の前記複数の特徴点の各々の前記第 n ローカル特徴記述子を、前記複数の第 n ビジュアルワードのうちの対応する第 n ビジュアルワードに対応付け、該画像の各第 $(n - 1)$ レベルサブ画像に関して、前記複数の第 n ビジュアルワードの各々の該第 $(n - 1)$ レベルサブ画像内での出現頻度の統計を介して該第 $(n - 1)$ レベルサブ画像の記述的情報を生成し、且つ、前記複数の第 $(n - 1)$ レベルサブ画像の記述的情報を結合して、該画像の前記第 $(n - 1)$ レベルの記述的情報を生成する、第 n 記述的情報生成サブステップと、

を有する、請求項 1 又は 2 に記載の方法。

【請求項 4】

前記複数の第 1 ビジュアルワードを得ることは、前記複数の画像の前記複数の第 1 ローカル特徴記述子を集合化することを含み、前記複数の第 n ビジュアルワードを得ることは、前記複数の画像の前記複数の第 n ローカル特徴記述子を集合化することを含む、請求項 1 乃至 3 の何れか一項に記載の方法。

【請求項 5】

入力画像を複数の画像の各々とそれぞれマッチングすることによって複数の類似度を取得することを有する画像マッチング方法であって、

各画像に関する類似度の取得は、

K は正の整数であるとして、前記入力画像の第 0 レベルの記述的情報から第 K レベルの記述的情報を、それぞれ、該画像の第 0 レベルの記述的情報から第 K レベルの記述的情報と比較し、第 0 レベル合致度から第 K レベル合致度を含む複数の合致度を取得することと、

前記複数の合致度の加重和に従って該類似度を取得することと

を有し、

前記複数の画像の各画像の前記第 0 レベルの記述的情報から第 K レベルの記述的情報は、請求項 1 乃至 4 の何れか一項に記載の方法によって取得され、

前記入力画像の前記第 0 レベルの記述的情報は、前記入力画像に含まれる複数の特徴点に基づいて得られる前記入力画像の複数の第 0 レベルローカル特徴記述子と前記複数の第 1 ビジュアルワードとに従って取得され、

$n = 2, 3, \dots, K + 1$ のそれぞれについて、前記入力画像の第 $(n - 1)$ レベルの記述的情報は、前記入力画像に関して、

前記入力画像の $(n - 1)$ 回目の分割を実行し、複数の第 $(n - 1)$ レベルサブ画像を取得するステップと、

前記複数の第 n ビジュアルワードと、前記入力画像の前記複数の特徴点に基づいて得られる前記入力画像の複数の第 n レベルローカル特徴記述子と、に従って前記入力画像の

10

20

30

40

50

前記第 $(n - 1)$ レベルの記述的情報を生成するステップであり、前記複数の第 n レベルローカル特徴記述子の各々は、それぞれの特徴点が位置する対応する第 $(n - 1)$ レベルサブ画像に従って得られる、ステップと、

を実行することによって取得され、

第 1 レベルの記述的情報の比較は、前記第 0 レベル合致度が所定の閾値以上であるときにのみ実行され、第 n レベルの記述的情報の比較は、第 $(n - 1)$ レベル合致度が前記所定の閾値以上であるときにのみ実行される、

画像マッチング方法。

【請求項 6】

前記入力画像の前記第 0 レベルの記述的情報の取得は、

10

前記入力画像に含まれる前記複数の特徴点の各々に関して 1 つの第 0 レベルローカル特徴記述子を計算し、前記入力画像の前記複数の第 0 レベルローカル特徴記述子を取得することと、

前記入力画像の前記複数の第 0 レベルローカル特徴記述子を、前記複数の第 1 ビジュアルワードのうちの対応する第 1 ビジュアルワードに対応付け、前記複数の第 1 ビジュアルワードの各々の前記入力画像内での出現頻度の統計を介して前記入力画像の前記第 0 レベルの記述的情報を生成することと、

を有し、

前記入力画像の前記第 $(n - 1)$ レベルの記述的情報を生成するステップは、

前記入力画像の前記複数の特徴点の各特徴点に関して 1 つの第 n レベルローカル特徴記述子を、該特徴点が位置する前記第 $(n - 1)$ レベルサブ画像に基づいて計算し、前記入力画像の前記複数の第 n レベルローカル特徴記述子を取得することと、

20

前記入力画像の前記複数の特徴点の各々の前記第 n レベルローカル特徴記述子を、前記複数の第 n ビジュアルワードのうちの対応する第 n ビジュアルワードに対応付け、前記入力画像の各第 $(n - 1)$ レベルサブ画像に関して、前記複数の第 n ビジュアルワードの各々の該第 $(n - 1)$ レベルサブ画像内での出現頻度の統計を介して該第 $(n - 1)$ レベルサブ画像の記述的情報を生成することと、

前記複数の第 $(n - 1)$ レベルサブ画像の記述的情報を結合して、前記入力画像の前記第 $(n - 1)$ レベルの記述的情報を生成することと

を有する、

30

請求項 5 に記載の方法。

【請求項 7】

複数の画像の記述的情報を取得する装置であって、

前記複数の画像の各画像に関して特徴点検出を実行し、各画像の複数の特徴点を取得する検出ユニットと、

前記複数の画像の第 0 レベルの記述的情報を取得する第 1 取得ユニットであり、各画像に関して、該画像の前記複数の特徴点に基づいて得られる該画像の複数の第 1 ローカル特徴記述子と、前記複数の画像の前記複数の第 1 ローカル特徴記述子に基づいて得られる複数の第 1 ビジュアルワードとに従って、該画像の前記第 0 レベルの記述的情報を生成する第 1 取得ユニットと、

40

K は正の整数であるとして、 $n = 2, 3, \dots, K + 1$ のそれぞれについて、前記複数の画像の第 $(n - 1)$ レベルの記述的情報を取得する第 n 取得ユニットであり、

各画像に関して、該画像の 第 $(n - 1)$ 回目の分割を実行し、該画像の複数の第 $(n - 1)$ レベルサブ画像を取得する第 $(n - 1)$ 分割サブユニット、及び

各画像に関して、該画像の複数の第 n ローカル特徴記述子と複数の第 n ビジュアルワードとに従って該画像の第 $(n - 1)$ レベルの記述的情報を生成する第 n 取得サブユニットであり、前記複数の第 n ローカル特徴記述子の各々は、該画像のそれぞれの特徴点と該それぞれの特徴点が位置する対応する第 $(n - 1)$ レベルサブ画像とに基づいて得られ、前記複数の第 n ビジュアルワードは、前記複数の画像の前記複数の第 n ローカル特徴記述子に基づいて得られる、第 n 取得サブユニット、

50

を有する、第 n 取得ユニットと、
を有する装置。

【請求項 8】

入力画像の第 0 レベルの記述的情報を取得する第 1 情報取得ユニットであり、前記入力画像に含まれる複数の特徴点に基づいて得られる前記入力画像の複数の第 0 レベルローカル特徴記述子と前記複数の第 1 ビジュアルワードとに従って前記入力画像の前記第 0 レベルの記述的情報を取得するように構成された第 1 情報取得ユニットと、

$n = 2, 3, \dots, K + 1$ のそれぞれについて、前記入力画像の第 $(n - 1)$ レベルの記述的情報を取得する第 n 情報取得ユニットであり、

前記入力画像の $(n - 1)$ 回目の分割を実行し、複数の第 $(n - 1)$ レベルサブ画像を取得する第 $(n - 1)$ 分割サブユニット、及び

前記複数の第 n ビジュアルワードと、前記入力画像の前記複数の特徴点に従って得られる前記入力画像の複数の第 n レベルローカル特徴記述子と、に従って前記入力画像の前記第 $(n - 1)$ レベルの記述的情報を生成する第 n 情報生成サブユニットであり、前記複数の第 n レベルローカル特徴記述子の各々は、それぞれの特徴点が位置する対応する第 $(n - 1)$ レベルサブ画像に従って得られる、第 n 情報生成サブユニット、

を有する第 n 情報取得ユニットと

前記入力画像を前記複数の画像の各々とそれぞれマッチングすることによって複数の類似度を取得する画像マッチングユニットであり、

前記複数の画像の各画像に関して、前記入力画像の第 0 レベルの記述的情報から第 K レベルの記述的情報を、それぞれ、該画像の第 0 レベルの記述的情報から第 K レベルの記述的情報と比較し、第 0 レベル合致度から第 K レベル合致度を含む複数の合致度を取得する比較サブユニットと、

各画像の類似度を、該画像の前記複数の合致度の加重和に従って取得する類似度生成サブユニットと

を有する画像マッチングユニットと、

を更に有し、

第 1 レベルの記述的情報の比較は、前記第 0 レベル合致度が所定の閾値以上であるときにのみ実行され、第 n レベルの記述的情報の比較は、第 $(n - 1)$ レベル合致度が前記所定の閾値以上であるときにのみ実行される、

請求項 7 に記載の装置。

【請求項 9】

機械読み取り可能命令コードを有するプログラムであり、該命令コードは、機械によって読み取られて実行されるときに、請求項 1 乃至 6 の何れか一項に記載の方法を実行する、プログラム。

【請求項 10】

請求項 9 に記載のプログラムを格納した記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は概して画像処理の分野に関し、特に、画像の記述的情報を取得する方法及び装置に関する。

【背景技術】

【0002】

情報技術の発展に伴い、例えば画像ファイル管理、産業画像／ビデオ監視、医療画像診断などのマルチメディアアプリケーションの分野で、大規模データベースがますます見受けられるようになってきている。従って、大規模データベースに関する迅速な画像検索及び画像認識が今日の重要な課題となっている。

【0003】

“単語（ワード）バッグ（bag of words）”に基づく伝統的な画像記述方法は、この

10

20

30

40

50

課題を達成するのに良好に機能するものとして広く適用されてきた。例えば、そのような単語バッグに基づく画像記述方法は、非特許文献 1 にて紹介されており、そこでは、画像の多数の局所的特徴から代表的な“ビジュアルワード（視覚語）”が抽出され、それらのビジュアルワードがその画像内で現れる頻度が、その画像を記述するために使用される。また、空間的な情報に関する特徴の記述が限られていることを解決するために、マルチレイヤ（複数層）の画像空間構造に基づく記述方法も提案されている。さらに、非特許文献 2 には、空間的なピラミッドの形態で画像を記述する方法が開示されている。しかしながら、これらの方法においては、各特徴点に対して 1 つの局所的特徴記述子が計算されるのみであり、そのため、生成されるものは静的な固定のビジュアルワードと見なし得るものであり、そのような静的ビジュアルワードは、異なる空間レイヤにある画像特徴の記述には適用可能でないことがある。

10

【先行技術文献】

【非特許文献】

【0004】

【非特許文献 1】C. Wallraven 等、「Recognition with local features: the kernel recipe」、Proc. ICCV、2003 年、第 1 巻、pp.257-264

【非特許文献 2】K. Grauman 等、「Pyramid match kernel: Discriminative classification with sets of image features」、Proc. ICCV、2005 年

【発明の概要】

【発明が解決しようとする課題】

20

【0005】

従来技術における上述の問題に鑑み、開示の技術は、複数の異なる空間レイヤで画像を正確に記述することが可能な動的単語（ダイナミックワード）バッグに基づく新たなマルチレイヤ画像記述方法を提供する。

【0006】

開示の技術はまた、画像を柔軟にマッチングすることが可能な画像マッチング方法を提供する。

【課題を解決するための手段】

【0007】

本発明の一態様によれば、複数の画像の記述的情報を取得する方法が提供される。当該方法は、

30

前記複数の画像の各画像に関して特徴点検出を実行し、各画像の複数の特徴点を取得する検出ステップと、

前記複数の画像の第 0 レベルの記述的情報を取得する第 1 取得ステップであり、前記複数の画像の各画像の第 0 レベルの記述的情報が、該画像の前記複数の特徴点に基づいて得られる該画像の複数の第 1 ローカル特徴記述子と、前記複数の画像の前記複数の第 1 ローカル特徴記述子に基づいて得られる複数の第 1 ビジュアルワードとに従って生成される、第 1 取得ステップと、

K は正の整数であるとして、 $n = 2, 3, \dots, K + 1$ のそれぞれについて、前記複数の画像の第 $(n - 1)$ レベルの記述的情報を取得する第 n 取得ステップであり、各画像に関して、

40

該画像の $(n - 1)$ 回目の分割を実行し、該画像の複数の第 $(n - 1)$ レベルサブ画像を取得する第 $(n - 1)$ 分割サブステップ、及び

該画像の第 $(n - 1)$ レベルの記述的情報が、該画像の前記複数の特徴点に基づいて得られる該画像の複数の第 n ローカル特徴記述子と、前記複数の画像の前記複数の第 n ローカル特徴記述子に基づいて得られる複数の第 n ビジュアルワードとに従って生成され、該画像の前記第 n ローカル特徴記述子は、それぞれの前記特徴点が位置する対応する第 $(n - 1)$ レベルサブ画像に従って得られる、第 n 取得サブステップ、

が実行される、第 n 取得ステップと、を含む。

【0008】

50

本発明の他の一態様によれば、画像マッチング方法が提供される。当該方法は、入力画像を複数の画像の各々とそれぞれマッチングすることによって複数の類似度を取得することを含み、各画像に関する類似度の取得は、 K は正の整数であるとして、前記入力画像の第0レベルの記述的情報から第 K レベルの記述的情報を、それぞれ、該画像の第0レベルの記述的情報から第 K レベルの記述的情報と比較し、第0レベル合致度から第 K レベル合致度を含む複数の合致度を取得することと、前記複数の合致度の加重和に従って該類似度を取得することとを有し、

前記複数の画像の各画像の前記第0レベルの記述的情報から第 K レベルの記述的情報は、本発明に係る方法によって取得され、

前記入力画像の前記第0レベルの記述的情報は、前記入力画像に含まれる複数の特徴点に基づいて得られる前記入力画像の複数の第0レベルローカル特徴記述子と前記複数の第1ビジュアルワードとに従って取得され、

$n = 2, 3, \dots, K + 1$ のそれぞれについて、前記入力画像の第 $(n - 1)$ レベルの記述的情報は、前記入力画像に関して、

前記入力画像の $(n - 1)$ 回目の分割を実行し、複数の第 $(n - 1)$ レベルサブ画像を取得するステップと、

前記複数の第 n ビジュアルワードと、前記入力画像の前記複数の特徴点に基づいて得られる前記入力画像の複数の第 n レベルローカル特徴記述子と、に従って前記入力画像の前記第 $(n - 1)$ レベルの記述的情報を生成するステップであり、前記複数の第 n レベルローカル特徴記述子の各々は、それぞれの特徴点が位置する対応する第 $(n - 1)$ レベルサブ画像に従って得られる、ステップと、

を実行することによって取得され、

第1レベルの記述的情報の比較は、前記第0レベル合致度が所定の閾値以上であるときにのみ実行され、第 n レベルの記述的情報の比較は、第 $(n - 1)$ レベル合致度が前記所定の閾値以上であるときにのみ実行される。

【0009】

本発明の他の一態様によれば、複数の画像の記述的情報を取得する装置が提供される。当該装置は、

前記複数の画像の各画像に関して特徴点検出を実行し、各画像の複数の特徴点を取得する検出ユニットと、

前記複数の画像の第0レベルの記述的情報を取得する第1取得ユニットであり、各画像に関して、該画像の前記複数の特徴点に基づいて得られる該画像の複数の第1ローカル特徴記述子と、前記複数の画像の前記複数の第1ローカル特徴記述子に基づいて得られる複数の第1ビジュアルワードとに従って、該画像の前記第0レベルの記述的情報を生成する第1取得ユニットと、

K は正の整数であるとして、 $n = 2, 3, \dots, K + 1$ のそれぞれについて、前記複数の画像の第 $(n - 1)$ レベルの記述的情報を取得する第 n 取得ユニットであり、

各画像に関して、該画像の $(n - 1)$ 回目の分割を実行し、該画像の複数の第 $(n - 1)$ レベルサブ画像を取得する第 $(n - 1)$ 分割サブユニット、及び

各画像に関して、該画像の複数の第 n ローカル特徴記述子と複数の第 n ビジュアルワードとに従って該画像の第 $(n - 1)$ レベルの記述的情報を生成する第 n 取得サブユニットであり、前記複数の第 n ローカル特徴記述子の各々は、該画像のそれぞれの特徴点と該それぞれの特徴点が位置する対応する第 $(n - 1)$ レベルサブ画像とに基づいて得られ、前記複数の第 n ビジュアルワードは、前記複数の画像の前記複数の第 n ローカル特徴記述子に基づいて得られる、第 n 取得サブユニット、

を有する、第 n 取得ユニットと、を含む。

【0010】

本発明の更なる他の一態様によれば、機械読み取り可能命令コードを有するプログラムが提供される。該命令コードは、機械によって読み取られて実行されるときに、本発明の実施形態に係る複数の画像の記述的情報を取得する方法を実行することができる。

【 0 0 1 1 】

本発明の更なる一態様によれば、上述のプログラムを担持した記憶媒体が提供される。

【発明の効果】

【 0 0 1 2 】

本発明の実施形態は、各特徴点に対して1つのローカル特徴記述子が得られるのみの伝統的な方法と異なり、各特徴点に対して一連の記述子を生成する。故に、画像を異なる複数の空間レイヤで正確に記述することができる。

【 0 0 1 3 】

また、粗いレイヤから精細なレイヤまでのシーケンスで画像をマッチングすることができる。そして、画像は、第 $m - 1$ レイヤで所定の閾値を上回る程度に別の画像と合致する場合にのみ、第 m レイヤでマッチングされる。故に、画像のマッチングを柔軟性高く行うことができる。

【図面の簡単な説明】

【 0 0 1 4 】

以下の図面を参照しての本発明の実施形態の詳細な説明により、本発明の上述及びその他の目的、態様及び利点が更に明らかになるであろう。図面において、同一あるいは同様の技術的特徴若しくは構成要素は、同一あるいは類似の参照符号を用いて表記することとする。

【図1】本発明の一実施形態に係る複数の画像の記述的情報を取得する方法を示すフローチャートである。

【図2】図1に示した第 n 取得ステップを実行する一具体例を示すフローチャートである。

【図3】画像の2レベル記述的情報を例示する模式図である。

【図4】本発明の一実施形態に係る画像マッチング方法を示すフローチャートである。

【図5】本発明の一実施形態に係る複数の画像の記述的情報を取得する装置を示すブロック図である。

【図6】本発明の実施形態に係る方法及び装置が具現化され得るコンピュータを示すブロック図である。

【発明を実施するための形態】

【 0 0 1 5 】

以下、図面を参照して、本発明の実施形態を説明する。図面中の1つの図又は本発明の1つの実施形態にて説明される要素及び特徴が、1つ以上のその他の図又は実施形態にて説明される要素及び特徴と組み合わせられてもよい。図面及び以下の説明において、本発明に特に関連せず且つ当業者に周知の要素及び処理については、明瞭さのため、その表示及び説明を省略することとする。

【 0 0 1 6 】

本発明に一実施形態において、十分な画像特徴情報を収集するため、画像は複数のレイヤに分割され、それぞれの空間レイヤでローカル画像記述子が計算され、粗いレイヤから精細なレイヤまでのそれぞれの空間レイヤに関して、それぞれのビジュアルワードが、対応するローカル特徴記述子から生成され、そして、それぞれの空間レイヤにおいて、対応するビジュアルワードから、該画像の妥当なマルチレイヤ記述子が取得される。

【 0 0 1 7 】

図1は、本発明の一実施形態に係る複数の画像の記述的情報を取得する方法100を示す簡略化したフローチャートである。図1に示すように、先ず、検出ステップS110にて、複数の画像の各画像に関して特徴点検出が実行され、各画像の複数の特徴点を取得される。特徴点の数は、例えば、画像の特性や、記述的情報を用いて画像の検索及び認識を行うプロセッサの性能などに依存する。特徴点の数は、ユーザによって、メニューから、あるいは用途の要求に応じて選択されることができる。他の例では、記述的情報を取得する時間の長さがユーザの許容範囲に入るように、記述的情報を用いるプロセッサの性能に応じて、また、サンプル抽出及びテストを介して、特徴点の妥当な数が決定され得る。特

10

20

30

40

50

徴点は、例えばハリス (Harris) のコーナー検出器やエッジ画素のサンプリングなどの周知の方法で検出することができる。

【 0 0 1 8 】

次に、第 1 取得ステップ S 1 3 0 にて、複数の画像の第 0 レベルの記述的情報が取得される。具体的には、複数の画像の各画像の第 0 レベル記述的情報が、該画像の複数の特徴点に基づいて得られる該画像の複数のローカル特徴記述子 (第 1 ローカル特徴記述子) と、複数の画像それぞれのローカル特徴記述子に基づいて得られるビジュアルワードのライブラリ (複数の第 1 ビジュアルワード) とに従って生成される。

【 0 0 1 9 】

一実施形態において、画像の複数の特徴点の各々に対して 1 つのローカル特徴記述子が計算されて、該画像の複数のローカル特徴記述子が得られる。例えば、スケール不変特徴変換 (Scale Invariant Feature Transform; S I F T) 抽出法で S I F T 特徴点を抽出し、該特徴点の形状コンテキスト記述子を計算することができる。あるいは、画像内の特徴点とその他の画素点との間の相対的な位置・角度関係の行列の統計量を生成することができる。そして、それぞれの画像のローカル特徴記述子がまとめられて、ビジュアルワードのライブラリが生成される。ビジュアルワードのライブラリは、例えば、k 平均クラスタリング法で取得することができる。例えば、画像のそれぞれの特徴点のローカル特徴記述子が k 平均クラスタリングされ、各クラスタ中心の値が 1 つのビジュアルワードに対応し、k 個のクラスタ中心の値及び対応するビジュアルワードのシリアル番号がビジュアルワードのライブラリとして格納される。ただし、k は、典型的に 5 0 0 から 1 0 0 0 の範囲の自然数である。次に、画像の複数のローカル特徴記述子が、ビジュアルワードのライブラリ内の対応するビジュアルワードにそれぞれ対応付けられ、ビジュアルワードのライブラリ内のそれぞれのビジュアルワードが該画像内に現れる頻度の統計を生成することによって該画像の第 0 レベルの記述的情報が生成される。各画像にビジュアルワードマッピング法を適用して、該画像のそれぞれの特徴点のローカル特徴記述子に対応するビジュアルワードにマッピングすることができる。例えば、特徴点ごとに、該特徴点のローカル特徴記述子とビジュアルワードライブラリ内のそれぞれのビジュアルワードに対応するそれぞれのクラスタ中心の値との間のユークリッド距離を計算して、該特徴点を表す最小のユークリッド距離を有するビジュアルワードのシリアル番号を見出し、且つ、それぞれのビジュアルワードのシリアル番号がその画像内に現れる回数の統計を生成することで、該

【 0 0 2 0 】

処理は第 n 取得ステップ S 1 5 0 へと進み、上記複数の画像の第 (n - 1) レベルの記述的情報が取得される。ただし、 $n = 2, 3, \dots, K + 1$ であり、K は正の整数である。具体的には、各画像に対して以下のステップが実行される。すなわち、該画像の (n - 1) 回目の分割を実行し、該画像の複数の第 (n - 1) レベルサブ画像を取得する (n - 1) 番目の分割サブステップと、該画像の第 (n - 1) レベルの記述的情報が、該画像の複数の特徴点に基づいて得られる該画像の複数の第 n ローカル特徴記述子と複数の画像の第 n ローカル特徴記述子に基づいて得られる複数の第 n ビジュアルワードとに従って生成される n 番目の取得サブステップと、が実行される。ただし、該画像の第 n ローカル特徴記述子の各々は、それぞれの特徴点の位置する対応する第 (n - 1) レベルサブ画像に従って得られる。

【 0 0 2 1 】

続いて、図 2 を参照して、図 1 の第 n 取得ステップ S 1 5 0 ($n = 2, 3, \dots, K + 1$) を達成する一具体例 S 2 5 0 の簡略化したフローチャートを説明する。すなわち、具体例 S 2 5 0 は、第 2 取得ステップから第 (K + 1) 取得ステップまでの K ステップを含む。図 2 に示すように、ステップ S 2 5 2 にて、各画像は、第 1 グリッドレベルから第 K グリッドレベルの K 組のグリッド要素 (サブ画像) を取得するよう、K 回にわたって、一連の徐々に細くなる空間グリッドへと繰り返し分割される。分割する回数は、ユーザによって決定されてもよいし、実際の用途に従って決定されてもよい。例えば、 4×4 の長

方形グリッド又は円形グリッドが1つの空間グリッドとして使用され得る。次に、ステップS 2 5 4にて、各グリッドレベルにて、該グリッドレベルのそれぞれの画像のローカル特徴記述子を取得するため、各特徴点に対応するローカル特徴記述子が、該特徴点が位置するグリッド要素から計算される。例えば、特徴点の形状コンテキスト記述子をグリッド要素内で計算することができ、あるいは、グリッド要素内の特徴点とその他の画素点との間の相対的な位置・角度関係の行列の統計量を生成することができる。

【0022】

好適な一実施形態において、第2取得ステップから第(K+1)取得ステップまでの一部の取得ステップで、より正確に画像を記述するための、更なる特徴点を検出する追加検出ステップがステップS 2 5 2の前に追加され得る。

【0023】

処理はステップS 2 5 6へと進み、同じグリッドレベルの全ての画像のローカル特徴記述子をクラスタ化(集合化;例えば、k平均クラスタリングなど)し、該グリッドレベルのビジュアルワードのライブラリを生成する。故に、異なるグリッドレベルでは異なるビジュアルワードのライブラリが存在する。

【0024】

そして、ステップS 2 5 8にて、各画像に関して、第1から第Kレベルの画像記述子が、それぞれのグリッドレベルでの該画像のローカル特徴記述子とビジュアルワードのライブラリとから生成される。具体的には、各グリッドレベルにて、該グリッドレベルの画像のローカル特徴記述子がそれぞれ、該グリッドレベルのビジュアルワードのライブラリ内の対応するビジュアルワードに対応付けられる。例えば、そのグリッドレベルの画像のローカル特徴記述子が、ビジュアルワードマッピング法にて、該グリッドレベルのビジュアルワードのライブラリ内の対応するビジュアルワードにマッピングされる。次に、そのグリッドレベルの画像の各グリッド要素に関して、ビジュアルワードのライブラリ内のそれぞれのビジュアルワードが該グリッド要素内で現れる頻度の統計を生成することによって、該グリッド要素の記述的情報(例えば、ビジュアルワードのヒストグラムなど)が生成される。そして、そのグリッドレベルの全てのグリッド要素の記述的情報が結合され(例えば、該全てのグリッド要素それぞれのビジュアルワードのヒストグラムが、1つの長いヒストグラムへと順次接続され)、該グリッドレベルの画像の記述的情報が取得される。他の例では、そのグリッドレベルの画像の記述的情報は、該レベルのビジュアルワードのライブラリ内のそれぞれのビジュアルワードが該画像内に現れる頻度の統計を直接的に生成することによって生成され得る。

【0025】

図3は、画像の2レベル記述的情報を例示する模式図である。図3において、円内の記述的情報は、1つのグリッド要素(サブ画像)のビジュアルワードのヒストグラムである。

【0026】

一方、本発明の他の一実施形態は更に、入力画像を複数の画像内のそれぞれの画像とマッチングする方法を提供する。図4は、方法400の簡略化したフローチャートを示している。図4に示すように、ステップS 4 1 0にて、複数の画像の各々の第0レベルから第Kレベルの記述的情報と入力画像とが取得される。複数の画像の各々の記述的情報は、図1に示した方法で得ることができる。

【0027】

その一方で、入力画像に含まれる複数の特徴点に基づいて得られる入力画像の複数の第0レベルのローカル特徴記述子と図1の第1取得ステップS 1 3 0にて得られたビジュアルワードのライブラリとに従って、入力画像の第0レベル記述的情報を取得することができる。一実施形態において、入力画像の複数の特徴点の各々に対して1つのローカル特徴記述子が計算されて、入力画像の複数の第0レベルローカル特徴記述子が取得される。例えば、スケール不変特徴変換(SIFT)抽出法でSIFT特徴点を抽出し、該特徴点の形状コンテキスト記述子を計算することができる。あるいは、入力画像内の特徴点とその

10

20

30

40

50

他の画素点との間の相対的な位置・角度関係の行列の統計量を生成することができる。そして、入力画像の複数の第0レベルローカル特徴記述子が、ビジュアルワードのライブラリ内の対応するビジュアルワードにそれぞれ対応付けられ、ビジュアルワードのライブラリ内のそれぞれのビジュアルワードが入力画像内に現れる頻度の統計を生成することによって入力画像の第0レベルの記述的情報が生成される。

【0028】

さらに、 $n = 2, 3, \dots, K + 1$ として、入力画像に以下のステップを実行することによって、入力画像の第 $(n - 1)$ レベルの記述的情報が取得される。すなわち、入力画像が $(n - 1)$ 回にわたって分割されて、複数の第 $(n - 1)$ レベルサブ画像が取得され、且つ、入力画像の第 $(n - 1)$ レベルの記述的情報が、入力画像の複数の特徴点に基づいて得られる入力画像の複数の第 n ローカル特徴記述子と図1の第 n 取得ステップS150にて得られる複数の第 n ビジュアルワードとに従って生成される。ただし、複数の第 n ローカル特徴記述子の各々は、それぞれの特徴点が位置する入力画像の対応する第 $(n - 1)$ レベルサブ画像に従って得られる。一実施形態において、入力画像は、第1グリッドレベルから第 K グリッドレベルの K 組のグリッド要素（サブ画像）を取得するよう、 K 回にわたって、一連の徐々に細くなる空間グリッドへと繰り返し分割される。さらに、各グリッドレベルにて、該グリッドレベルの入力画像のローカル特徴記述子を取得するため、各グリッド要素に含まれる特徴点のローカル特徴記述子が該グリッド要素から計算される。次に、各グリッドレベルにて、該グリッドレベルにおける入力画像の記述的情報が、該グリッドレベルの入力画像のローカル特徴記述子と図1の第 n 取得ステップS150で得られた同じグリッドレベルのビジュアルワードのライブラリとから生成される。

【0029】

処理はステップS430へと進み、複数の画像の各々の第0レベルから第 K レベルの記述的情報が、それぞれ、入力画像の第0レベルから第 K レベルの記述的情報と比較され、第0レベルから第 K レベルの合致（マッチング）度が取得される。

【0030】

そして、ステップS450にて、複数の画像の各々に関して、第0レベルから第 K レベルの合致度の加重和が計算され、該画像の入力画像に対する類似度が取得される。それぞれのレベルの重み係数は、それぞれのレベルの重要度に従って割り当てられ得る。

【0031】

方法400において、第1レベルの記述的情報の比較は、第0レベルの合致度が所定の閾値以上のときにのみ実行され、第 n レベルの記述的情報の比較は、第 $(n - 1)$ レベルの合致度が所定の閾値以上のときにのみ実行される。

【0032】

本発明の実施形態に係る画像マッチング方法において、画像は、粗いレイヤから精細なレイヤまでのシーケンスにて柔軟にマッチングされることができ、 $(m - 1)$ 番目のレイヤにおける他の画像との合致度が所定の閾値より大きい場合にのみ m 番目のレイヤでマッチングされることになる。

【0033】

図5は、本発明の一実施形態に係る複数の画像の記述的情報を取得する装置500を示す簡略化したブロック図である。図5に示すように、装置500は、検出ユニット510と、第1取得ユニット530と、第2取得ユニット550₁から第 $(K + 1)$ 取得ユニット550_Kまでとを含んでいる。具体的には、第2取得ユニット550₁から第 $(K + 1)$ 取得ユニット550_Kまでの各々は、対応する分割サブユニットと取得サブユニットとを含んでいる。図5に示すように、第2取得ユニット550₁は、第1分割サブユニット550_{1A}と第2取得サブユニット550_{1B}とを含み、第 $(K + 1)$ 取得ユニット550_Kは、第 K 分割サブユニット550_{KA}と第 $(K + 1)$ 取得サブユニット550_{KB}とを含んでいる。動作時、検出ユニット510は、入力された複数の画像に関して特徴点検出を実行し、これらの画像の各々に関して複数の特徴点を抽出する。第1取得ユニット530は、入力された複数の画像の各々に関して該画像の抽出された特徴点に従ってのロー

カル特徴記述子を取得し、それぞれの画像のローカル特徴記述子に従って複数の第1ビジュアルワードを取得し、且つ、各画像に関して、該画像のローカル特徴記述子と複数の第1ビジュアルワードとに従って該画像の第0レベルの記述的情報を生成する。第1分割サブユニット550_{1A}から第K分割サブユニット550_{KA}の各々は入力された複数の画像を分割し、それにより、各画像が複数の第1レベルから第Kレベルのサブ画像に分割される。第2取得サブユニット550_{1B}から第(K+1)取得サブユニット550_{KB}の各々は、検出ユニット510によって抽出された特徴点を受信する。第2取得サブユニット550_{1B}は、第1分割サブユニット550_{1A}から第1レベルサブ画像を受信する。同様に、第(K+1)取得サブユニット550_{KB}は、第K分割サブユニット550_{KA}から第Kレベルサブ画像を受信する。各画像に関して、第2取得サブユニット550_{1B}は、受信した該画像の特徴点と該画像の第1レベルサブ画像とに従って、該画像の第1レベルのローカル特徴記述子を取得する。そして、第2取得サブユニット550_{1B}は、それぞれの画像の第1レベルローカル特徴記述子に従って複数の第2ビジュアルワードを取得し、各画像に関して、該画像の第1レベルローカル特徴記述子と複数の第2ビジュアルワードとに従って該画像の第1レベルの記述的情報を生成する。同様に、第(K+1)取得サブユニット550_{KB}は、複数の画像の第Kレベルの記述的情報を取得する。

【0034】

図5に示した装置500、及びそれに含まれるそれぞれのユニット510乃至550は、図1及び図2を参照して説明した様々な処理を実行するように構成され得る。これらの処理の更なる詳細については、上述のそれぞれの実施形態、実装例及び例を参照することができ、ここでは繰り返しての説明は省略する。

【0035】

以上のブロック図、フローチャート及び/又は実施形態の詳細な説明にて、本発明の実施形態に係る装置及び/又は方法の様々な実装例を説明した。1つ以上の機能及び/又は処理がこれらのブロック図、フローチャート及び/又は実施形態に含まれるとき、当業者は、これらのブロック図、フローチャート及び/又は実施形態におけるそれぞれの機能及び/又は処理が、様々なハードウェア、ソフトウェア、ファームウェア、又は実質的にそれらの組み合わせにて、別々に、且つ/或いは組み合わせて具現化され得ることを認識するであろう。一実装例において、本明細書に記載された事項の一部は、特定用途向け集積回路(ASIC)、フィールド・プログラマブル・ゲート・アレイ(FPGA)、デジタル信号プロセッサ(DSP)又はその他の集積形態にて具現化され得る。しかしながら、当業者に認識されるように、本明細書に記載された一部の態様の実装例は全体的あるいは部分的に事実上、1つ以上のコンピュータ上で実行される1つ以上のコンピュータプログラムの形態(例えば、1つ以上のコンピュータシステム上で実行される1つ以上のコンピュータプログラムの形態)、1つ以上のプロセッサ上で実行される1つ以上のプログラムの形態(例えば、1つ以上のマイクロプロセッサ上で実行される1つ以上のプログラムの形態)、ファームウェアの形態、又は実質的にそれらの組み合わせの形態、をした集積回路にて具現化され得る。また、当業者は、本明細書中の開示を受けて、本開示に係る回路を設計し、且つ/或いは本開示に係るソフトウェア及び/又はファームウェアのコードを記述することができるであろう。

【0036】

例えば、上述の装置500、及びそれぞれの構成モジュール、ユニット及びサブユニットは、ソフトウェア、ファームウェア、ハードウェア、又はこれらの組み合わせにて構成されることができる。ソフトウェア又はファームウェアにて具現化される場合、そのソフトウェアを構成するプログラムが、記憶媒体又はネットワークから、専用のハードウェア構成を有するコンピュータ(例えば、図6に例示する汎用コンピュータ600など)にインストールされることができ、該コンピュータは、様々なプログラムをインストールされるとき、様々な機能を実行することができる。

【0037】

図6は、本発明の実施形態に係る方法及び装置が具現化されることが可能なコンピュー

10

20

30

40

50

タの模式ブロック図を示している。

【 0 0 3 8 】

図 6 において、中央処理装置（CPU）601 は、読み出し専用メモリ（ROM）602 に格納されたプログラム、又は記憶部 608 からランダムアクセスメモリ（RAM）603 にロードされたプログラムに従って様々な処理を実行する。RAM 603 にはまた、必要に応じて、CPU 601 が様々な処理を実行するのに必要なデータが格納され得る。CPU 601、ROM 602 及び RAM 603 は、バス 604 によって互いに接続されている。バス 604 にはまた、入力／出力インタフェース 605 が接続されている。

【 0 0 3 9 】

入力／出力（I/O）インタフェース 605 には、入力部 606（キーボードやマウスなどを含む）、出力部 607（例えば陰極線管（CRT）若しくは液晶ディスプレイ（LCD）などの表示装置やスピーカなどを含む）、記憶部 608（ハードディスクなどを含む）、及び通信部 609（例えば LAN カードやモデムなどのネットワークインタフェースカードを含む）が接続されている。通信部 609 は、例えばインターネットなどのネットワーク上での通信処理を実行する。ドライバ 610 も入力／出力インタフェース 605 に接続され得る。ドライバ 610 には、必要に応じて、例えば磁気ディスク、光ディスク、光磁気ディスク又は半導体メモリなどの、リムーバブルメディア（取り外し可能媒体）611 が取り付けられることができ、それにより、必要に応じて、それから読み出されたコンピュータプログラムが記憶部 608 にインストールされる。

【 0 0 4 0 】

上述の一連の処理がソフトウェアにて実行される場合、該ソフトウェアを構成するプログラムが、例えばインターネットなどのネットワーク又は例えばリムーバブルメディア 611 などの記憶媒体からインストールされる。

【 0 0 4 1 】

当業者に認識されるように、記憶媒体は、プログラムが格納され且つ該プログラムをユーザに提供するために装置とは別に配布される図 6 に示したリムーバブルメディア 611 に限定されない。リムーバブルメディア 611 の例には、磁気ディスク（フロッピー（登録商標）ディスクを含む）、光ディスク（コンパクトディスク ROM（CD-ROM）及びデジタル多用途ディスク（DVD）を含む）、光磁気ディスク（ミニディスク（MD）（登録商標）を含む）、及び半導体メモリが含まれる。代替的に、記憶媒体は、プログラムが格納され且つ当該記憶媒体を含んだ装置とともにユーザに配布される、ROM 602、又は記憶部 608 に含まれるハードディスクであってもよい。

【 0 0 4 2 】

故に、本発明によって更に、機械読み取り可能な命令コードが格納されたプログラムが提案される。該命令コードは、機械によって読み出されて実行されるときに、上述の本発明の実施形態に係る様々な方法を実行することが可能である。従って、そのようなプログラムを担持する様々な記憶媒体も本発明の開示に包含される。

【 0 0 4 3 】

以上の本発明の実施形態の説明において、1つの実施形態に関して説明／図示された特徴は、同一あるいは同様のようにして1つ以上のその他の実施形態で使用されてもよいし、その他の実施形態の特徴と組み合わせられてもよいし、あるいはその他の実施形態の特徴を置き換えるように使用されてもよい。

【 0 0 4 4 】

なお、文脈中の用語“含む／有する”及びそれらの如何なる変形も、特徴、要素、ステップ又は構成要素の存在を意味するものであり、1つ以上のその他の特徴、要素、ステップ又は構成要素の存在若しくは付加を排除するものではない。

【 0 0 4 5 】

また、本発明に係る方法は、本明細書に記載された時系列で実行されるのみでなく、その他の時系列に従って、並行して、あるいは別々に実行されてもよい。故に、本発明の範囲は、本明細書に記載された方法の実行順序に限定されるものではない。

【 0 0 4 6 】

以上の本発明の実施形態の説明にて本発明を開示したが、認識されるように、当業者は、本発明の精神及び範囲を逸脱することなく、本発明に関する様々な変更、適応又は均等物を考え出し得る。それらの変更、適応又は均等物も、本発明の範囲に入るものとして見なされるべきである。

【 0 0 4 7 】

以上の説明に関し、更に以下の付記を開示する。

【 0 0 4 8 】

(付記 1)

複数の画像の記述的情報を取得する方法であって、

前記複数の画像の各画像に関して特徴点検出を実行し、各画像の複数の特徴点を取得する検出ステップと、

前記複数の画像の第 0 レベルの記述的情報を取得する第 1 取得ステップであり、前記複数の画像の各画像の第 0 レベルの記述的情報が、該画像の前記複数の特徴点に基づいて得られる該画像の複数の第 1 ローカル特徴記述子と、前記複数の画像の前記複数の第 1 ローカル特徴記述子に基づいて得られる複数の第 1 ビジュアルワードとに従って生成される、第 1 取得ステップと、

K は正の整数であるとして、 $n = 2, 3, \dots, K + 1$ のそれぞれについて、前記複数の画像の第 $(n - 1)$ レベルの記述的情報を取得する第 n 取得ステップであり、各画像に関して、

該画像の $(n - 1)$ 回目の分割を実行し、該画像の複数の第 $(n - 1)$ レベルサブ画像を取得する第 $(n - 1)$ 分割サブステップ、及び

該画像の第 $(n - 1)$ レベルの記述的情報が、該画像の前記複数の特徴点に基づいて得られる該画像の複数の第 n ローカル特徴記述子と、前記複数の画像の前記複数の第 n ローカル特徴記述子に基づいて得られる複数の第 n ビジュアルワードとに従って生成され、該画像の前記第 n ローカル特徴記述子は、それぞれの前記特徴点が位置する対応する第 $(n - 1)$ レベルサブ画像に従って得られる、第 n 取得サブステップ、

が実行される、第 n 取得ステップと、

を有する方法。

(付記 2)

各画像の前記第 0 レベルの記述的情報の生成は、

該画像の前記複数の特徴点の各々に関して 1 つの第 1 ローカル特徴記述子を計算し、該画像の前記複数の第 1 ローカル特徴記述子を取得する第 1 計算サブステップと、

該画像の前記複数の第 1 ローカル特徴記述子をそれぞれ、前記複数の第 1 ビジュアルワードのうちの対応する第 1 ビジュアルワードに対応付け、前記複数の第 1 ビジュアルワードの各々の該画像内での出現頻度の統計を介して該画像の前記第 0 レベルの記述的情報を生成する第 1 記述的情報生成サブステップと、

を有する、付記 1 に記載の方法。

(付記 3)

各画像に関する前記第 n 取得サブステップは、

該画像の前記複数の特徴点の各特徴点に関して 1 つの第 n ローカル特徴記述子を、該特徴点が位置する第 $(n - 1)$ レベルサブ画像に基づいて計算し、該画像の前記複数の第 n ローカル特徴記述子を取得する第 n 計算サブステップと、

該画像の前記複数の特徴点の各々の前記第 n ローカル特徴記述子を、前記複数の第 n ビジュアルワードのうちの対応する第 n ビジュアルワードに対応付け、該画像の各第 $(n - 1)$ レベルサブ画像に関して、前記複数の第 n ビジュアルワードの各々の該第 $(n - 1)$ レベルサブ画像内での出現頻度の統計を介して該第 $(n - 1)$ レベルサブ画像の記述的情報を生成し、且つ、前記複数の第 $(n - 1)$ レベルサブ画像の記述的情報を結合して、該画像の前記第 $(n - 1)$ レベルの記述的情報を生成する、第 n 記述的情報生成サブステップと、

を有する、付記 1 又は 2 に記載の方法。

(付記 4)

前記複数の第 1 ビジュアルワードを得ることは、前記複数の画像の前記複数の第 1 ローカル特徴記述子を集合化することを含み、前記複数の第 n ビジュアルワードを得ることは、前記複数の画像の前記複数の第 n ローカル特徴記述子を集合化することを含む、付記 1 乃至 3 の何れか一項に記載の方法。

(付記 5)

前記複数の第 1 ローカル特徴記述子及び前記複数の第 n ローカル特徴記述子は、スケール不変特徴変換 (SIFT) 記述子、形状コンテキスト記述子、及び相対位置・角度関係行列記述子のうちの 1 つを含む、付記 1 乃至 4 の何れか一項に記載の方法。

10

(付記 6)

入力画像を複数の画像の各々とそれぞれマッチングすることによって複数の類似度を取得することを有する画像マッチング方法であって、

各画像に関する類似度の取得は、

K は正の整数であるとして、前記入力画像の第 0 レベルの記述的情報から第 K レベルの記述的情報を、それぞれ、該画像の第 0 レベルの記述的情報から第 K レベルの記述的情報と比較し、第 0 レベル合致度から第 K レベル合致度を含む複数の合致度を取得することと、

前記複数の合致度の加重和に従って該類似度を取得することと

を有し、

20

前記複数の画像の各画像の前記第 0 レベルの記述的情報から第 K レベルの記述的情報は、付記 1 乃至 5 の何れか一項に記載の方法によって取得され、

前記入力画像の前記第 0 レベルの記述的情報は、前記入力画像に含まれる複数の特徴点に基づいて得られる前記入力画像の複数の第 0 レベルローカル特徴記述子と前記複数の第 1 ビジュアルワードとに従って取得され、

$n = 2, 3, \dots, K + 1$ のそれぞれについて、前記入力画像の第 $(n - 1)$ レベルの記述的情報は、前記入力画像に関して、

前記入力画像の $(n - 1)$ 回目の分割を実行し、複数の第 $(n - 1)$ レベルサブ画像を取得するステップと、

前記複数の第 n ビジュアルワードと、前記入力画像の前記複数の特徴点に基づいて得られる前記入力画像の複数の第 n レベルローカル特徴記述子と、に従って前記入力画像の前記第 $(n - 1)$ レベルの記述的情報を生成するステップであり、前記複数の第 n レベルローカル特徴記述子の各々は、それぞれの特徴点が位置する対応する第 $(n - 1)$ レベルサブ画像に従って得られる、ステップと、

30

を実行することによって取得され、

第 1 レベルの記述的情報の比較は、前記第 0 レベル合致度が所定の閾値以上であるときにのみ実行され、第 n レベルの記述的情報の比較は、第 $(n - 1)$ レベル合致度が前記所定の閾値以上であるときにのみ実行される、

画像マッチング方法。

(付記 7)

40

前記入力画像の前記第 0 レベルの記述的情報の取得は、

前記入力画像に含まれる前記複数の特徴点の各々に関して 1 つの第 0 レベルローカル特徴記述子を計算し、前記入力画像の前記複数の第 0 レベルローカル特徴記述子を取得することと、

前記入力画像の前記複数の第 0 レベルローカル特徴記述子を、前記複数の第 1 ビジュアルワードのうちの対応する第 1 ビジュアルワードに対応付け、前記複数の第 1 ビジュアルワードの各々の前記入力画像内での出現頻度の統計を介して前記入力画像の前記第 0 レベルの記述的情報を生成することと、

を有し、

前記入力画像の前記第 $(n - 1)$ レベルの記述的情報を生成するステップは、

50

前記入力画像の前記複数の特徴点の各特徴点に関して1つの第 n レベルローカル特徴記述子を、該特徴点が位置する前記第 $(n-1)$ レベルサブ画像に基づいて計算し、前記入力画像の前記複数の第 n レベルローカル特徴記述子を取得することと、

前記入力画像の前記複数の特徴点の各々の前記第 n レベルローカル特徴記述子を、前記複数の第 n ビジュアルワードのうちの対応する第 n ビジュアルワードに対応付け、前記入力画像の各第 $(n-1)$ レベルサブ画像に関して、前記複数の第 n ビジュアルワードの各々の該第 $(n-1)$ レベルサブ画像内での出現頻度の統計を介して該第 $(n-1)$ レベルサブ画像の記述的情報を生成することと、

前記複数の第 $(n-1)$ レベルサブ画像の記述的情報を結合して、前記入力画像の前記第 $(n-1)$ レベルの記述的情報を生成することと

10

を有する、

付記6に記載の方法。

(付記8)

複数の画像の記述的情報を取得する装置であって、

前記複数の画像の各画像に関して特徴点検出を実行し、各画像の複数の特徴点を取得する検出ユニットと、

前記複数の画像の第0レベルの記述的情報を取得する第1取得ユニットであり、各画像に関して、該画像の前記複数の特徴点に基づいて得られる該画像の複数の第1ローカル特徴記述子と、前記複数の画像の前記複数の第1ローカル特徴記述子に基づいて得られる複数の第1ビジュアルワードとに従って、該画像の前記第0レベルの記述的情報を生成する第1取得ユニットと、

20

K は正の整数であるとして、 $n = 2, 3, \dots, K+1$ のそれぞれについて、前記複数の画像の第 $(n-1)$ レベルの記述的情報を取得する第 n 取得ユニットであり、

各画像に関して、該画像の第 $(n-1)$ 回目の分割を実行し、該画像の複数の第 $(n-1)$ レベルサブ画像を取得する第 $(n-1)$ 分割サブユニット、及び

各画像に関して、該画像の複数の第 n ローカル特徴記述子と複数の第 n ビジュアルワードとに従って該画像の第 $(n-1)$ レベルの記述的情報を生成する第 n 取得サブユニットであり、前記複数の第 n ローカル特徴記述子の各々は、該画像のそれぞれの特徴点と該それぞれの特徴点が位置する対応する第 $(n-1)$ レベルサブ画像とに基づいて得られ、前記複数の第 n ビジュアルワードは、前記複数の画像の前記複数の第 n ローカル特徴記述子に基づいて得られる、第 n 取得サブユニット、

30

を有する、第 n 取得ユニットと、

を有する装置。

(付記9)

前記第1取得ユニットは、

各画像に関して、該画像の前記複数の特徴点の各々に関して1つの第1ローカル特徴記述子を計算し、該画像の前記複数の第1ローカル特徴記述子を取得する第1計算サブユニットと、

各画像に関して、該画像の前記複数の第1ローカル特徴記述子をそれぞれ、前記複数の第1ビジュアルワードのうちの対応する第1ビジュアルワードに対応付け、前記複数の第1ビジュアルワードの各々の該画像内での出現頻度の統計を介して該画像の前記第0レベルの記述的情報を生成する第1記述的情報生成サブユニットと、

40

を有する、付記8に記載の装置。

(付記10)

前記第 n 取得サブユニットは、

各画像に関して、該画像の前記複数の特徴点の各特徴点に関して1つの第 n ローカル特徴記述子を、該特徴点が位置する第 $(n-1)$ レベルサブ画像に従って計算し、該画像の前記複数の第 n ローカル特徴記述子を取得する第 n 計算サブユニットと、

各画像に関して、該画像の前記複数の特徴点の各々の前記第 n ローカル特徴記述子を、前記複数の第 n ビジュアルワードのうちの対応する第 n ビジュアルワードに対応付け、該

50

画像の各第 $(n - 1)$ レベルサブ画像の記述的情報を、前記複数の第 n ビジュアルワードの各々の該第 $(n - 1)$ レベルサブ画像内での出現頻度の統計を介して生成し、且つ、前記複数の第 $(n - 1)$ レベルサブ画像の記述的情報を結合して、該画像の前記第 $(n - 1)$ レベルの記述的情報を生成する、第 n 記述的情報生成サブユニットと、
を有する、付記 8 又は 9 に記載の装置。

(付記 1 1)

前記第 1 取得ユニットは、前記複数の画像の前記複数の第 1 ローカル特徴記述子を集合化することによって前記複数の第 1 ビジュアルワードを取得するように構成され、前記第 n 取得ユニットは、前記複数の画像の前記複数の第 n ローカル特徴記述子を集合化することによって前記複数の第 n ビジュアルワードを取得するように構成される、付記 8 乃至 10 の何れか一項に記載の装置。

(付記 1 2)

前記複数の第 1 ローカル特徴記述子及び前記複数の第 n ローカル特徴記述子は、スケール不変特徴変換 (SIFT) 記述子、形状コンテキスト記述子、及び相対位置・角度関係行列記述子のうちの 1 つを含む、付記 8 乃至 11 の何れか一項に記載の装置。

(付記 1 3)

入力画像の第 0 レベルの記述的情報を取得する第 1 情報取得ユニットであり、前記入力画像に含まれる複数の特徴点に基づいて得られる前記入力画像の複数の第 0 レベルローカル特徴記述子と前記複数の第 1 ビジュアルワードとに従って前記入力画像の前記第 0 レベルの記述的情報を取得するように構成された第 1 情報取得ユニットと、

$n = 2, 3, \dots, K + 1$ のそれぞれについて、前記入力画像の第 $(n - 1)$ レベルの記述的情報を取得する第 n 情報取得ユニットであり、

前記入力画像の $(n - 1)$ 回目の分割を実行し、複数の第 $(n - 1)$ レベルサブ画像を取得する第 $(n - 1)$ 分割サブユニット、及び

前記複数の第 n ビジュアルワードと、前記入力画像の前記複数の特徴点に従って得られる前記入力画像の複数の第 n レベルローカル特徴記述子と、に従って前記入力画像の前記第 $(n - 1)$ レベルの記述的情報を生成する第 n 情報生成サブユニットであり、前記複数の第 n レベルローカル特徴記述子の各々は、それぞれの特徴点が位置する対応する第 $(n - 1)$ レベルサブ画像に従って得られる、第 n 情報生成サブユニット、

を有する第 n 情報取得ユニットと

前記入力画像を前記複数の画像の各々とそれぞれマッチングすることによって複数の類似度を取得する画像マッチングユニットであり、

前記複数の画像の各画像に関して、前記入力画像の第 0 レベルの記述的情報から第 K レベルの記述的情報を、それぞれ、該画像の第 0 レベルの記述的情報から第 K レベルの記述的情報と比較し、第 0 レベル合致度から第 K レベル合致度を含む複数の合致度を取得する比較サブユニットと、

各画像の類似度を、該画像の前記複数の合致度の加重和に従って取得する類似度生成サブユニットと

を有する画像マッチングユニットと、

を更に有し、

第 1 レベルの記述的情報の比較は、前記第 0 レベル合致度が所定の閾値以上であるときにのみ実行され、第 n レベルの記述的情報の比較は、第 $(n - 1)$ レベル合致度が前記所定の閾値以上であるときにのみ実行される、

付記 8 乃至 12 の何れか一項に記載の装置。

(付記 1 4)

前記第 1 情報取得ユニットは、

前記入力画像に含まれる前記複数の特徴点の各々に関して 1 つの第 0 レベルローカル特徴記述子を計算し、前記入力画像の前記複数の第 0 レベルローカル特徴記述子を取得し、且つ

前記入力画像の前記複数の第 0 レベルローカル特徴記述子を、前記複数の第 1 ビジュ

アルワードのうちの対応する第 1 ビジュアルワードに対応付け、前記複数の第 1 ビジュアルワードの各々の前記入力画像内での出現頻度の統計を介して前記入力画像の前記第 0 レベルの記述的情報を生成する、

ように構成され、

前記第 n 情報生成サブユニットは、

前記入力画像の前記複数の特徴点の各特徴点に関して 1 つの第 n レベルローカル特徴記述子を、該特徴点が位置する前記第 $(n - 1)$ レベルサブ画像に基づいて計算して、前記入力画像の前記複数の第 n レベルローカル特徴記述子を取得し、

前記入力画像の前記複数の特徴点の各々の前記第 n レベルローカル特徴記述子を、前記複数の第 n ビジュアルワードのうちの対応する第 n ビジュアルワードに対応付け、前記入力画像の各第 $(n - 1)$ レベルサブ画像の記述的情報を、前記複数の第 n ビジュアルワードの各々の該第 $(n - 1)$ レベルサブ画像内での出現頻度の統計を介して生成し、且つ

前記複数の第 $(n - 1)$ レベルサブ画像の記述的情報を結合して、前記入力画像の前記第 $(n - 1)$ レベルの記述的情報を生成する、

ように構成される、

付記 13 に記載の装置。

(付記 15)

機械読み取り可能命令コードを有するプログラムであり、該命令コードは、機械によって読み取られて実行されるときに、付記 1 - 7 の何れか一項に記載の方法を実行する、プログラム。

(付記 16)

付記 15 に記載のプログラムを格納した記憶媒体。

【符号の説明】

【0049】

500 記述的情報取得装置

510 検出ユニット

530 第 1 取得ユニット

550₁ 第 2 取得ユニット

550_{1A} 第 1 分割サブユニット

550_{1B} 第 2 取得サブユニット

550_K 第 $(K + 1)$ 取得ユニット

550_{KA} 第 K 分割サブユニット

550_{KB} 第 $(K + 1)$ 取得サブユニット

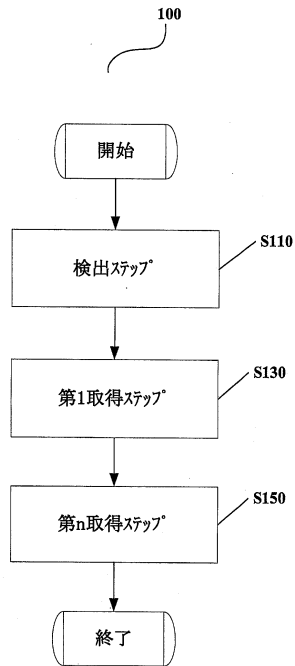
10

20

30

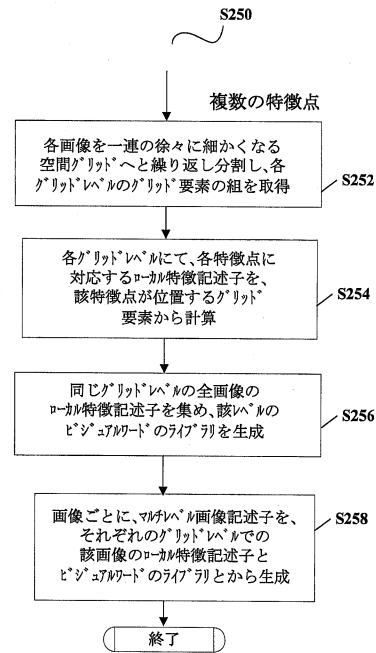
【図 1】

本発明の一実施形態に係る複数の画像の記述的情報を取得する方法を示すフローチャート



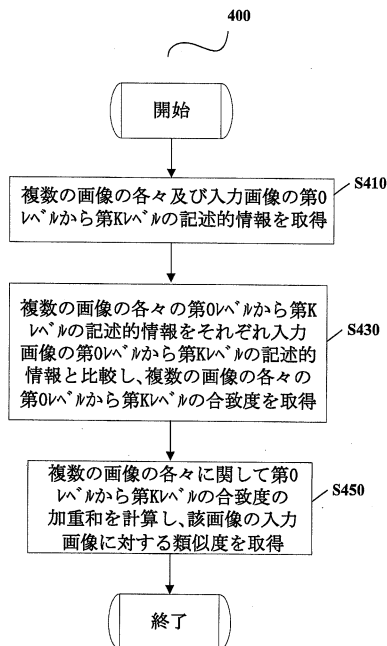
【図 2】

図 1 に示した第 n 取得ステップを実行する一具体例を示すフローチャート



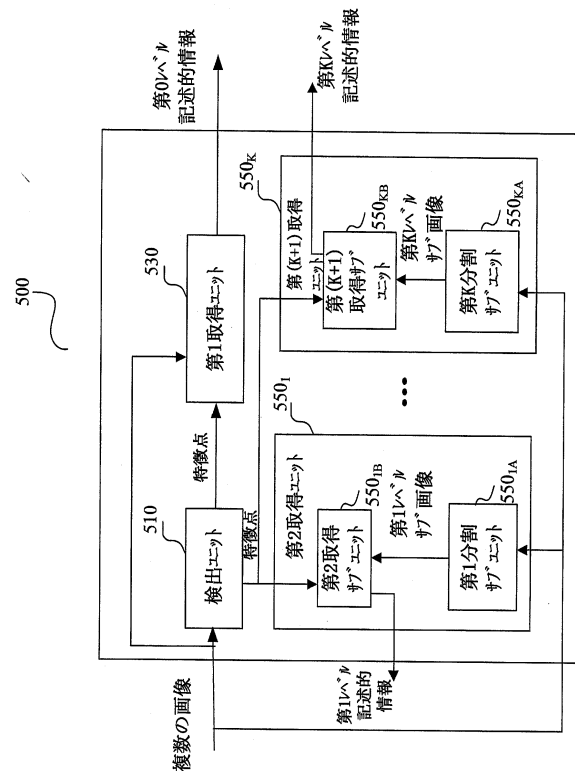
【図 4】

本発明の一実施形態に係る画像マッチング方法を示すフローチャート



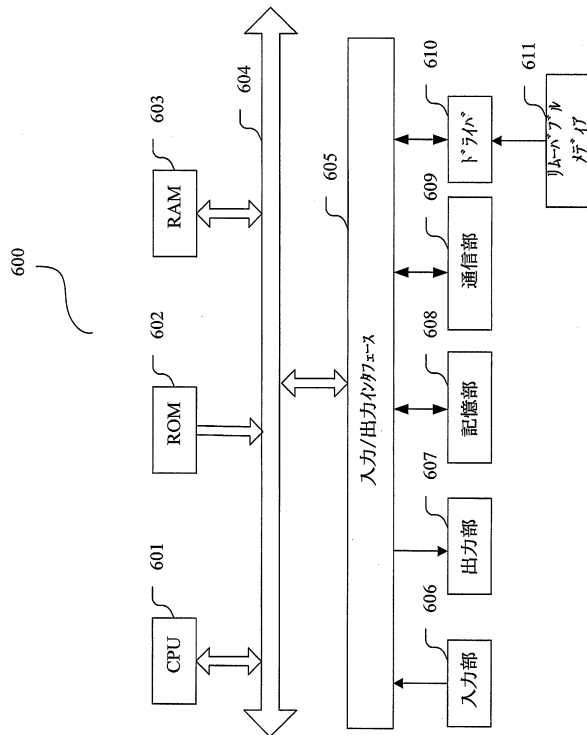
【図 5】

本発明の一実施形態に係る複数の画像の記述的情報を取得する装置を示す模式ブロック図



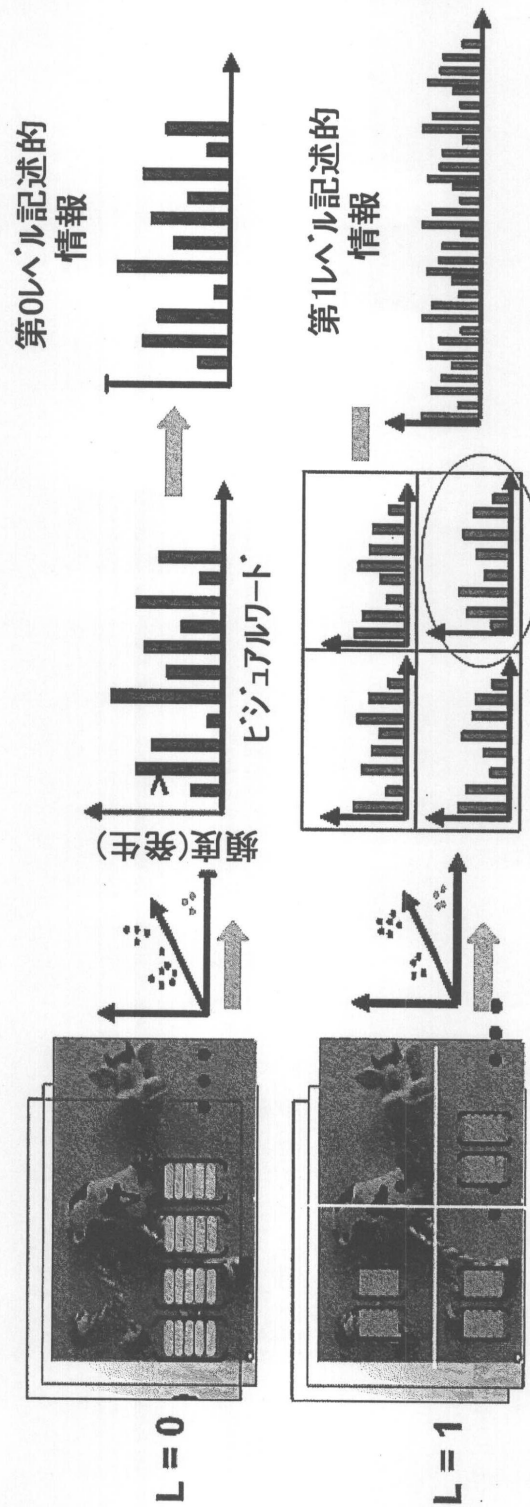
【図 6】

本発明の実施形態に係る方法及び装置が具現化され得るコンピュータを示す
模式ブロック図



【図3】

画像の2レベル記述的信息を例示する模式図



フロントページの続き

(72)発明者 リウ・ルウジエ

中国, 100025, ベイジン, チャオヤン ディストリクト, ジョオン ロード, ドン ス
ホ アヌ ナンバー 56, オーシャン インターナショナル センター, タワー エイ 15エフ 富
士通研究開発中心有限公司内

(72)発明者 于 浩

中国, 100025, ベイジン, チャオヤン ディストリクト, ジョオン ロード, ドン ス
ホ アヌ ナンバー 56, オーシャン インターナショナル センター, タワー エイ 15エフ 富
士通研究開発中心有限公司内

(72)発明者 杉村 昌彦

神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

(72)発明者 上原 祐介

神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

審査官 早川 学

(56)参考文献 Svetlana Lazebnik, Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Na
tural Scene Categories, 2006 IEEE Computer Society Conference on Computer Vision and P
attern Recognition, IEEE, 2006年, pp.2169-2178

(58)調査した分野(Int.Cl., DB名)

G06T 7/00

G06F 17/30

G06T 1/00