

(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2015-133085

(P2015-133085A)

(43) 公開日 平成27年7月23日(2015.7.23)

(51) Int.Cl.

G06T 7/00 (2006.01)
H04N 7/18 (2006.01)

F 1

G06T 7/00
G06T 7/00
H04N 7/18250
350B
D

テーマコード(参考)

5C054

5L096

審査請求 未請求 請求項の数 18 O L (全 25 頁)

(21) 出願番号
(22) 出願日特願2014-5444 (P2014-5444)
平成26年1月15日 (2014.1.15)(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
(74) 代理人 100076428
弁理士 大塚 康徳
(74) 代理人 100112508
弁理士 高柳 司郎
(74) 代理人 100115071
弁理士 大塚 康弘
(74) 代理人 100116894
弁理士 木村 秀二
(74) 代理人 100130409
弁理士 下山 治
(74) 代理人 100134175
弁理士 永川 行光

最終頁に続く

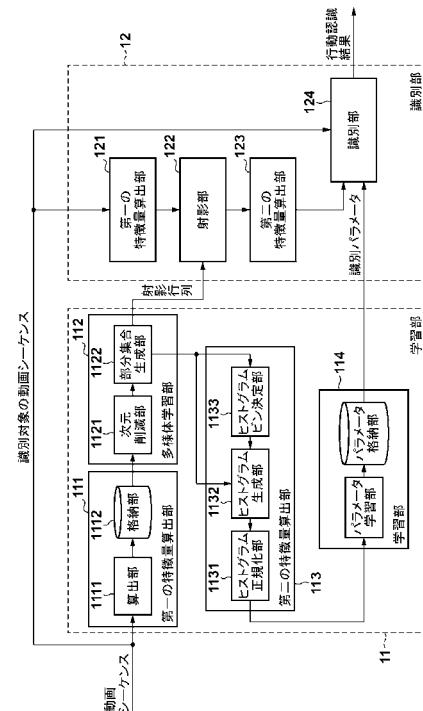
(54) 【発明の名称】情報処理装置およびその方法

(57) 【要約】

【課題】 映像中の物体のカテゴリを識別する際の、似て非なるカテゴリの識別率の向上を目的とする。

【解決手段】 第一の特徴量算出部111は、動画シーケンスから第一の特徴量を抽出する。多様体学習部112は、第一の特徴量を次元圧縮し、次元圧縮された第一の特徴量の部分集合と、当該次元圧縮に対応する射影行列を生成する。第二の特徴量算出部113は、部分集合の代表点に基づき、次元圧縮された第一の特徴量の分布を示す第二の特徴量を算出する。学習部114は、第二の特徴量を用いて識別パラメータを学習する。識別部12は、射影行列および識別パラメータを用いて、映像中の物体のカテゴリを識別する。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

動画シーケンスから第一の特徴量を抽出する第一の抽出手段と、
前記第一の特徴量を次元圧縮する圧縮手段と、
前記次元圧縮された第一の特徴量の部分集合と、当該次元圧縮に対応する射影行列とを生成する生成手段と、
前記部分集合の代表点に基づき、前記次元圧縮された第一の特徴量の分布を示す第二の特徴量を算出する第一の算出手段と、
前記第二の特徴量を用いて識別パラメータを学習する学習手段と、
前記射影行列および前記識別パラメータを用いて、映像中の物体のカテゴリを識別する識別手段とを有する情報処理装置。
10

【請求項 2】

前記第一の抽出手段は、前記第一の特徴量として複数種類の特徴量を抽出する請求項1に記載された情報処理装置。

【請求項 3】

前記生成手段は、前記次元圧縮された特徴量を真値として、前記特徴量を低次元空間に射影する射影行列を生成する請求項1または請求項2に記載された情報処理装置。

【請求項 4】

前記生成手段は、前記次元圧縮された特徴量のベクトル量子化データを前記部分集合とする請求項1から請求項3の何れか一項に記載された情報処理装置。
20

【請求項 5】

前記生成手段は、前記次元圧縮された特徴量のクラスタリングデータを前記部分集合とする請求項1から請求項3の何れか一項に記載された情報処理装置。

【請求項 6】

前記第一の算出手段は、各部分集合の代表点を算出する手段と、
BINとして前記代表点を原点からの距離の順に配列したヒストグラムを決定する手段と、
前記次元圧縮された第一の特徴量それぞれについて、当該特徴量との間の距離が最短のBINに投票したヒストグラムを前記第二の特徴量として生成する手段とを有する請求項1から請求項5の何れか一項に記載された情報処理装置。
30

【請求項 7】

前記第一の算出手段は、さらに、前記生成されたヒストグラムを正規化する手段を有する請求項6に記載された情報処理装置。

【請求項 8】

前記学習手段は、前記第二の特徴量を用いて、識別対象カテゴリごとの識別パラメータを学習する請求項1から請求項7の何れか一項に記載された情報処理装置。

【請求項 9】

前記識別手段は、識別対象の動画シーケンスから輝度の勾配に関係する第三の特徴量を抽出する第二の抽出手段と、

前記射影行列を用いて、前記第三の特徴量を低次元空間に射影する射影手段と、
前記低次元空間に射影された第三の特徴量の分布を示す第四の特徴量を算出する第二の算出手段と、
前記識別パラメータおよび前記第四の特徴量に基づき、前記識別対象の動画シーケンスから前記物体のカテゴリを識別する認識手段を有する請求項1から請求項8の何れか一項に記載された情報処理装置。
40

【請求項 10】

動画シーケンスから第一の特徴量を抽出する抽出手段と、
前記第一の特徴量を次元圧縮する圧縮手段と、
前記次元圧縮された第一の特徴量の部分集合を生成する生成手段と、
前記部分集合の境界を学習して、前記部分集合ごとの識別パラメータを生成する学習手
50

段と、

前記第一の特徴量に対する、前記部分集合に属す識別カテゴリの重みを行列要素とする重み行列を算出する算出手段と、

前記識別パラメータおよび前記重み行列を用いて、映像中の物体のカテゴリを識別する識別手段とを有する情報処理装置。

【請求項 1 1】

前記識別手段は、識別対象の動画シーケンスから第二の特徴量を抽出する第二の抽出手段と、

前記識別パラメータ、前記重み行列および前記第二の特徴量に基づき、前記識別対象の動画シーケンスから前記物体のカテゴリを識別する認識手段を有する請求項10に記載された情報処理装置。

【請求項 1 2】

前記識別手段は、前記物体のカテゴリとして前記物体の行動カテゴリを識別する請求項1から請求項11の何れか一項に記載された情報処理装置。

【請求項 1 3】

前記第二の特徴量は輝度の勾配に関係する特徴量である請求項11または請求項12に記載された情報処理装置。

【請求項 1 4】

前記第位一の特徴量は輝度の勾配に関係する特徴量である請求項1から請求項13の何れか一項に記載された情報処理装置。

【請求項 1 5】

抽出手段、圧縮手段、生成手段、算出手段、学習手段、識別手段を有する情報処理装置の情報処理方法であって、

前記抽出手段が、動画シーケンスから第一の特徴量を抽出し、

前記圧縮手段が、前記第一の特徴量を次元圧縮し、

前記生成手段が、前記次元圧縮された第一の特徴量の部分集合と、当該次元圧縮に対応する射影行列とを生成し、

前記算出手段が、前記部分集合の代表点に基づき、前記次元圧縮された第一の特徴量の分布を示す第二の特徴量を算し、

前記学習手段が、前記第二の特徴量を用いて識別パラメータを学習し、

前記識別手段が、前記射影行列および前記識別パラメータを用いて、映像中の物体のカテゴリを識別する情報処理方法。

【請求項 1 6】

抽出手段、圧縮手段、生成手段、学習手段、算出手段、識別手段を有する情報処理装置の情報処理方法であって、

前記抽出手段が、動画シーケンスから第一の特徴量を抽出し、

前記圧縮手段が、前記第一の特徴量を次元圧縮し、

前記生成手段が、前記次元圧縮された第一の特徴量の部分集合を生成し、

前記学習手段が、前記部分集合の境界を学習して、前記部分集合ごとの識別パラメータを生成し、

前記算出手段が、前記第一の特徴量に対する、前記部分集合に属す識別カテゴリの重みを行列要素とする重み行列を算出し、

前記識別手段が、前記識別パラメータおよび前記重み行列を用いて、映像中の物体のカテゴリを識別する情報処理方法。

【請求項 1 7】

コンピュータを請求項1から請求項14の何れか一項に記載された情報処理装置の各手段として機能させるためのプログラム。

【請求項 1 8】

請求項17に記載されたプログラムが記録されたコンピュータが読み取り可能な記録媒体。

10

20

30

40

50

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、映像中の物体のカテゴリを識別する情報処理に関する。

【背景技術】**【0002】**

映像中の人物の複数カテゴリの認識手法として、例えば特許文献1に示されるような手法が提案されている。特許文献1の手法は、画像データに対して次元圧縮を施し、低次元データを抽出し、抽出した低次元データをクラスタリングして、映像中の複数カテゴリ(人物の年齢)を認識する。

10

【0003】

しかし、特許文献1の手法は主成分分析、局所性保存射影、線形判別分析など、入力データの特徴空間における線形性を仮定した次元圧縮を用いるため、似て非なるカテゴリを識別することが困難である。

【先行技術文献】**【特許文献】****【0004】**

【特許文献1】特開2010-257409号公報

【非特許文献】**【0005】**

20

【非特許文献1】Heng Wang、Alexander Klaser、Cordelia Schmid、Cheng-Lin Liu「Action recognition by dense trajectories」CVPR 2011: 3169-3176頁

【非特許文献2】N. Dalal、B. Triggs「Histograms of oriented gradients for human detection」In CVPR、2005年

【非特許文献3】I. Laptev、M. Marsza.ek、C. Schmid、B. Rozenfeld「Learning realistic human actions from movies」In CVPR、2008年

【非特許文献4】N. Dalal、B. Triggs、C. Schmid「Human detection using oriented histograms of flow and appearance」In ECCV、2006年

【非特許文献5】S. Roweis、L. Saul「Non linear dimensionality reduction by locally linear embedding」Science、Vol. 290、No. 5500、2323-2326頁、2000年12月22日

30

【非特許文献6】J. B. Tenenbaum、V. deSilva、J. C. Langford「A global geometric framework for nonlinear dimensionality reduction」Science、290: 2319-2323頁、2000年

【非特許文献7】P. S. Bradley、U. M. Fayyad「Refining Initial Points for K-Means Clustering」in Proceedings of the 15th International Conference on Machine Learning、91-99頁、1998年

【非特許文献8】Y. Cheng「Mean shift mode seeking, and clustering」IEEE Trans. Pattern Anal. and Machine Intell.、17(8): 790-799頁、1995年

【非特許文献9】Kohonen T.「Self-organizing formation of topologically correct feature maps」Biol Cybern 43: 59-69頁、1982年

40

【非特許文献10】Y. Linde、A. Buzo、R. M. Gray「An Algorithm for Vector Quantizer Design」IEEE Trans. on Communications、Vol. 28、No. 1、84-95頁、1980年

【非特許文献11】V. Vapnik「The Nature of Statistical Learning Theory」Springer、1995年

【非特許文献12】Minsky M.、Papert S.「Perceptrons」MIT Press、1969年

【発明の概要】**【発明が解決しようとする課題】****【0006】**

本発明は、映像中の物体のカテゴリを識別する際の、似て非なるカテゴリの識別率の向上を目的とする。

50

【課題を解決するための手段】**【0007】**

本発明は、前記の目的を達成する一手段として、以下の構成を備える。

【0008】

本発明にかかる情報処理は、動画シーケンスから第一の特徴量を抽出し、前記第一の特徴量を次元圧縮し、前記次元圧縮された第一の特徴量の部分集合と、当該次元圧縮に対応する射影行列とを生成し、前記部分集合の代表点に基づき、前記次元圧縮された第一の特徴量の分布を示す第二の特徴量を算出し、前記第二の特徴量を用いて識別パラメータを学習し、前記射影行列および前記識別パラメータを用いて、映像中の物体のカテゴリを識別する。

10

【発明の効果】**【0009】**

本発明によれば、映像中の物体のカテゴリを識別する際に、似て非なるカテゴリの識別率を向上することができる。

【図面の簡単な説明】**【0010】**

【図1】実施例1の行動認識処理を行う情報処理装置（行動認識装置）の構成例を示すブロック図。

20

【図2】学習部11と識別部12の処理の概要を説明するフローチャート。

【図3】特徴量の次元削減（次元圧縮）を示す図。

【図4】部分集合の一例を示す図。

【図5】部分集合の代表点の一例を示す図。

【図6】ヒストグラムの一例を示す図。

【図7】射影部による射影を説明する図。

【図8】実施例2の行動認識装置の構成例を示すブロック図。

【図9】次元削減（次元圧縮）を示す図。

【図10】部分集合の一例を示す図。

【図11】部分集合の代表点の一例を示す図。

【図12】ヒストグラムの一例を示す図。

30

【図13】連結されたヒストグラムの一例を示す。

【図14】射影部による射影を説明する図。

【図15】実施例3の行動認識装置の構成例を示すブロック図。

【図16】部分集合の一例を示す図。

【発明を実施するための形態】**【0011】**

以下、本発明にかかる実施例の情報処理を図面を参照して詳細に説明する。

【実施例1】**【0012】****[装置の概要]**

図1のブロック図により実施例1の行動認識処理を行う情報処理装置（以下、行動認識装置）の構成例を示す。なお、行動認識装置は半導体集積回路（LSI）を用いて実現される。

40

【0013】

行動認識装置は、学習部11と識別部12を有し、これら構成は行動認識装置の機能にそれぞれ対応する。図2のフローチャートにより学習部11と識別部12の処理の概要を説明する。なお、図2に示す処理を実行するプログラムをコンピュータ機器に供給して行動認識装置を実現することも可能である。

【0014】

行動認識装置は、時間的に連続する複数の静止画像（以下、動画シーケンス）を入力して（S11）、動画シーケンスから例えば輝度の勾配に関係する第一の特徴量を算出する（S12）。そして、学習モードにおいて（S13）、第一の特徴量の次元を圧縮し（S14）、次元圧縮によ

50

つて得られる低次元空間の第一の特徴量の部分集合に基づき第二の特徴量を算出する(S15)。また、行動認識装置は、低次元空間の第一の特徴量を真値とする射影行列を算出し、所定の記憶領域に格納する(S16)。さらに、行動認識装置は、第二の特徴量に基づき識別対象カテゴリごとの識別パラメータを学習し(S18)、学習した識別パラメータを所定の記憶領域に格納する(S19)。以上が学習部11の処理の概要である。

【0015】

一方、識別モードにおいて(S13)、行動認識装置は、所定の記憶領域に格納された射影行列を用いて第一の特徴量の次元を圧縮し(S17)、次元圧縮によって得られる低次元空間の第一の特徴量の部分集合に基づき、第二の特徴量を算出する(S15)。また、行動認識装置は、所定の記憶領域に格納された識別対象カテゴリごとの識別パラメータを用いて識別を行う(S20)。以上が認識部12の処理の概要である。

10

【0016】

行動認識装置によって得られる映像中の物体(例えば人物)の行動認識結果は、行動認識装置の上位階層(アプリケーション層)に伝達される。そして、行動認識結果は、監視カメラに実装される異常行動検出機能や、DVDレコーダや家庭用ビデオカメラなどに実装されるビデオセグメンテーション機能など、様々な映像情報処理アプリケーションに利用される。

【0017】

識別対象カテゴリとしては、例えば、人間の「走る」「ジョギングする」「歩く」「這う」「立つ」「座る」「うずくまる」「倒れる」などの行動カテゴリが挙げられる。

20

【0018】

学習部11

図1に示すように、学習部11は、学習部11の各処理に対応する第一の特徴量算出部111、多様体学習部112、第二の特徴量算出部113、学習部114を有する。

【0019】

第一の特徴量算出部111において、算出部1111は、識別対象カテゴリごとに、時間的に連続した複数の静止画像(以下、動画シーケンス)を入力し、所定の特徴量を抽出する。特徴量としては、例えば、非特許文献1のHistograms of Oriented Gradients (HOG)、非特許文献2のHistograms of Flow (HOF)、非特許文献3のMotion Boundary Histogram (MBH)などがある。

30

【0020】

HOGは、物体を認識するために局所領域における輝度の勾配方向をヒストグラム化した勾配ベースの特徴量(輝度勾配ヒストグラム)である。HOGは、映像中から歩行者や人工物などの物体を検出するために使われ、輝度勾配ヒストグラム(物体のみえから得られる角度ごとのヒストグラム)のBINに輝度強度が累積される。つまり、画像の局所領域の輝度の勾配方向ごとの強度に着目する特徴量である。

【0021】

HOFは、カメラと物体の距離に応じて観測される画像上の動きベクトルを局所領域においてヒストグラム化し、近隣の領域で作成したヒストグラムとの差分を抽出した特徴量であり、輝度の時系列の勾配強度に着目する特徴量である。

40

【0022】

MBHは、オプティカルフローの水平、垂直各成分の局所勾配が表す動きの境界のヒストグラムを生成して、動きの軌道の時空間的コンテキストを表現する特徴量である。なお、第一の特徴量としては、輝度の勾配に限らず、輝度値そのものを用いてもよいし、色に基づく特徴量などを用いてもよい。

【0023】

算出部1111が抽出した特徴量は、第一の特徴量として、特徴量格納部1112に一時格納される。特徴量格納部1112は、多様体学習部112の要求に応じて、第一の特徴量を多様体学習部112に出力する。

【0024】

50

多様体学習部112において、次元削減部1121は、第一の特徴量算出部111が算出した第一の特徴量を入力し、第一の特徴量に多様体学習による次元削減を施し、次元削減した第一の特徴量を部分集合生成部1122に出力する。多様体学習による次元削減には、例えば、非特許文献5のLLEや非特許文献6のisomapを用いる。図3により特徴量の次元削減（次元圧縮）を示す。図3(a)は第一の特徴量算出部111が抽出した第一の特徴量を示し、図3(b)は次元が圧縮（削減）された第一の特徴量を示す。

【0025】

また、多様体学習部112は、第一の特徴量算出部111が抽出した第一の特徴量を真値として、別の新たな第一の特徴量が入力された場合、多様体学習を行わずに行列演算のみで次元削減結果が得られるように射影行列を算出する。つまり、図3(a)に対する図3(b)の射影行列が算出され、射影行列は識別部12に出力される。なお、射影行列は、例えば線形回帰分析を用いて算出する。

10

【0026】

部分集合生成部1122は、次元削減部1121から低次元空間の第一の特徴量を入力し、機械学習によるクラスタ解析やベクトル量子化、あるいは、自己組織化写像などを適用して、複数の部分集合を生成する。部分集合は、例えば、低次元空間の第一の特徴量のベクトル量子化データまたはクラスタリングデータである。機械学習の手法には、例えば、非特許文献7のK-means、非特許文献8のMean shift、非特許文献9のSOM、非特許文献10のLinde-Buzo-Grayアルゴリズムを用いる。図4に部分集合の一例を示す。図4(a)は次元圧縮後の第一の特徴量を示し、図4(b)はクラスタ解析により第一の特徴量が複数の部分集合に区分された状態を示す。

20

【0027】

第二の特徴量算出部113において、ヒストグラム bin 決定部1131は、部分集合生成部112が生成した部分集合を入力する。そして、各部分集合の代表点（重心、部分集合内の単位面積当たりの点群の密度が最大点、または、部分集合内から無作為に選択した点など）の座標を算出して、ヒストグラム生成部1132に出力する。図5に部分集合の代表点の一例を示す。図5(a)は複数の部分集合を示し、図5(b)は算出された部分集合の代表点（×記号）を示す。

20

【0028】

ヒストグラム生成部1132は、ヒストグラム bin 決定部1131から入力される部分集合の代表点の座標を例えば原点からの距離の昇順（または降順）に配列してヒストグラム bin とするヒストグラムを生成する。図6にヒストグラムの一例を示す。図6(a)は部分集合の代表点 $G(C_1)$ 、 $G(C_2)$ 、…、 $G(C_7)$ を示し、図6(b)は生成されるヒストグラムを示す。詳細な手順は次のとおりである。

30

手順1：学習用として与えられた動画シーケンスに第一の特徴量算出部111および多様体学習部112を適用して、低次元空間の第一の特徴量を算出する。

手順2：算出した低次元空間の第一の特徴量と全部分集合の代表点の座標の間のユークリッド距離を算出する。

手順3：算出したユークリッド距離が最小の代表点（最短の代表点）の座標を特定し、この代表点の座標に対応するヒストグラム bin に投票する。例えば、第一の特徴量が図6(a)に示す点 P で与えられた場合、代表点 $G(C_5)$ との間のユークリッド距離が最小であり、図6(b)に示すヒストグラム bin $G(C_5)$ に投票する。

40

手順4：手順1～3を、学習用に与えられたすべての動画シーケンスに適用して、複数のヒストグラムを生成する。

【0029】

ヒストグラム正規化部1133は、各学習用動画シーケンスに対応するヒストグラムについて、その頻度値の合計が1になるようにヒストグラムを正規化したヒストグラムを第二の特徴量として学習部114に出力する。

【0030】

学習部114において、パラメータ学習部1141は、第二の特徴量算出部113から第二の特徴

50

量を入力し、識別対象カテゴリごとの識別パラメータを学習する。ある識別対象カテゴリの識別パラメータの学習は、当該カテゴリの第二の特徴量をポジティブデータ、その他の第二の特徴量をネガティブデータとし、例えば非特許文献11のサポートベクトルマシン(SVM)や非特許文献12のパーセプトロンを用いて行われる。パラメータ格納部1142は、識別対象カテゴリごとの識別パラメータを一時格納し、識別部12からの要求に応じて、識別対象カテゴリごとの識別パラメータを出力する。

【0031】

認識部12

図1に示すように、識別部12は、識別部12の各処理に対応する第一の特徴量算出部121、射影部122、第二の特徴量算出部123、識別部124を有す。

10

【0032】

第一の特徴量算出部121は、学習部11の第一の特徴量算出部111と同様の構成を有し、識別対象の動画シーケンスから第一の特徴量を抽出する。抽出された特徴量は、学習部11と同様に、第一の特徴量として、特徴量格納部1112と同様のメモリ領域に一時格納され、射影部122の要求に応じて、第一の特徴量が射影部122に出力される。

【0033】

射影部122は、学習部11の多様体学習部112において予め算出された射影行列を用いて、第一の特徴量算出部121が抽出した第一の特徴量を低次元空間に射影する。図7により射影部122による射影を説明する。図7(a)は第一の特徴量算出部121が抽出した第一の特徴量を示し、図7(b)は射影により次元が圧縮(削減)された第一の特徴量を示す。

20

【0034】

第二の特徴量算出部123は、学習部11の第二の特徴量算出部113と同様の構成を有し、射影部122から入力した低次元空間の第一の特徴量から、例えば、第一の特徴量の分布を示すヒストグラムを算出する。そして、当該ヒストグラムを第二の特徴量として識別部124に出力する。

【0035】

識別部124は、第二の特徴量算出部123から第二の特徴量を入力し、学習部11の学習部114において予め学習された識別対象カテゴリごとの識別パラメータを入力する。そして、第二の特徴量と識別パラメータを用いて、識別部12に入力される、カテゴリが未知の識別対象の動画シーケンスのカテゴリを識別する。なお、識別パラメータは、例えば、サポートベクトルマシンやパーセプトロンを用いる場合は識別関数の重みベクトルである。

30

【0036】

識別部124の行動認識結果は、映像中の物体(例えば人物)の動作や行動を示し、行動認識装置の上位階層(アプリケーション層)に伝達される。例えば、監視カメラ、DVDレコーダ、家庭用ビデオカメラにおいて、映像中の物体の動作や行動を認識し、予め用意されたデータベースと比較して、特定シーンを検出する機能を想定する。この場合、上位階層は、行動認識装置を制御するCPUや、上記機能のプログラムを実行するCPUなどである。また、上記機能としては、監視カメラに実装される異常行動検出機能、DVDレコーダや家庭用ビデオカメラなどに実装されるビデオセグメンテーション機能などが挙げられる。

40

【0037】

このように、学習部11により、学習用として与えられた動画シーケンスの第二の特徴量を生成し、識別対象カテゴリごとの識別パラメータを学習する。その後、識別部12により、学習部11において学習された識別対象カテゴリごとの識別パラメータを用いて、カテゴリが未知の動画シーケンスのカテゴリを識別する。従って、例えば、歩く、ジョギング、走るなどの、似て非なる識別カテゴリの識別率を向上させることができる。

【実施例2】

【0038】

以下、本発明にかかる実施例2の情報処理を説明する。なお、実施例2において、実施例1と略同様の構成については、同一符号を付して、その詳細説明を省略する。

【0039】

50

図8のブロック図により実施例2の行動認識装置の構成例を示す。なお、行動認識装置は半導体集積回路(LSI)を用いて実現される。

【0040】

実施例2の行動認識装置が実施例1の行動認識装置と異なるのは、第一の特徴量算出部111の算出部1111が動画シーケンスから特徴量を抽出する際に、複数種類の特徴量を抽出する点である。以下では、特徴量としてHOG、HOF、MBHの三種類の特徴量が抽出されることを想定する。抽出された特徴量はそれぞれ、第一の特徴量として、特徴量格納部1112a、1112b、1112cに一時格納される。特徴量格納部1112a、1112b、1112cは、多様体学習部112の要求に応じて、第一の特徴量を多様体学習部112に出力する。

【0041】

多様体学習部112において、次元削減部1121は、第一の特徴量算出部111が算出した第一の特徴量を入力し、第一の特徴量に多様体学習による次元削減を施し、次元削減した第一の特徴量を部分集合生成部1122に出力する。図9により次元削減(次元圧縮)を示す。図9(a)はHOGを示し、図9(b)は次元圧縮後のHOGを示す。図9(c)はHOFを示し、図9(d)は次元圧縮後のHOFを示す。図9(e)はMBHを示し、図9(f)は次元圧縮後のMBHを示す。

【0042】

多様体学習部112は、第一の特徴量算出部111が抽出した第一の特徴量を真値として、別の新たな第一の特徴量が入力された場合、多様体学習を行わずに行列演算のみで次元削減結果が得られるように射影行列を算出する。つまり、図9(a)に対する図9(b)の射影行列、図9(c)に対する図9(d)の射影行列、図9(e)に対する図9(f)の射影行列が算出され、射影行列は識別部12に出力される。

【0043】

部分集合生成部1122は、次元削減部1121から低次元空間のHOG、HOF、MBHを入力し、機械学習によるクラスタ解析やベクトル量子化、あるいは、自己組織化写像などを適用して、複数の部分集合を生成する。図10により部分集合の一例を示す。図10(a)(c)(e)は次元圧縮後のHOG、HOF、MBHを示し、図10(b)(d)(f)はクラスタ解析によりHOG、HOF、MBHが複数の部分集合に区分された状態を示す。

【0044】

第二の特徴量算出部113において、ヒストグラム bin 決定部1131は、部分集合生成部112が生成した部分集合を入力する。そして、HOG、HOF、MBHごとに、各部分集合の代表点(重心、部分集合内の単位面積当たりの点群の密度が最大点、または、部分集合内から無作為に選択した点)の座標を算出して、ヒストグラム生成部1132に出力する。図11により部分集合の代表点の一例を示す。図11(a)(c)(e)はHOG、HOF、MBHの複数の部分集合を示し、図11(b)(d)(f)は算出された部分集合の代表点を示す。

【0045】

ヒストグラム生成部1132は、HOG、HOF、MBHごとに、ヒストグラム bin 決定部1131から入力される部分集合の代表点の座標を例えば原点からの距離の昇順(または降順)に配列してヒストグラム bin とするヒストグラムを生成する。図12によりヒストグラムの一例を示す。図12(a)(c)(e)はHOG、HOF、MBHの部分集合の代表点 $G(C_1), G(C_2), \dots, G(C_{20})$ を示し、図12(b)(d)(f)は生成されるヒストグラムを示す。

【0046】

ヒストグラムの生成手順は、実施例1で説明した手順3の後、HOG、HOF、MBHのヒストグラムを連結して一つのヒストグラムを生成し、実施例1で説明した手順4を実行する。図13により連結されたヒストグラムの一例を示す。

【0047】

このようにして、学習用に与えられた動画シーケンスそれぞれに対応するヒストグラムが生成され、各ヒストグラムの頻度値の合計が1になるようにヒストグラムが正規化され、正規化されたヒストグラムが第二の特徴量として学習部114に出力される。

【0048】

以降の学習部11の処理は、実施例1と同様であり、詳細説明を省略する。

10

20

30

40

50

【0049】

識別部12の射影部122は、学習部11の多様体学習部112においてHOG、HOF、MBHごとに予め算出された射影行列を用いて、第一の特徴量算出部121が抽出した第一の特徴量を低次元空間に射影する。図14により射影部122による射影を説明する。図14(a)(c)(d)はHOG、HOF、MBHを示し、図14(b)(d)(f)は射影により次元が圧縮（削減）されたHOG、HOF、MBHを示す。

【0050】

識別部12の他の処理は、実施例1と同様であり、詳細説明を省略する。

【実施例3】

【0051】

10

以下、本発明にかかる実施例3の情報処理を説明する。なお、実施例3において、実施例1、2と略同様の構成については、同一符号を付して、その詳細説明を省略する。

【0052】

図15のブロック図により実施例3の行動認識装置の構成例を示す。なお、行動認識装置は半導体集積回路(LSI)を用いて実現される。

【0053】

学習部11

図15に示すように、実施例3の学習部11の多様体学習部112の構成は、図1に示す実施例1の構成と異なる。

【0054】

20

多様体学習部112は、次元削減部1121、部分集合生成部1122、境界学習部1123、重み行列生成部1124、パラメータ格納部1125を有する。次元削減部1121と部分集合生成部1122の動作、機能は実施例1と同様であり、低次元空間の第一の特徴量の部分集合を示す情報が境界学習部1123と重み行列生成部1124に出力される。

【0055】

境界学習部1123は、低次元空間の第一の特徴量の部分集合を入力して、それら部分集合の境界を学習し、部分集合ごとの識別パラメータを生成する。図16により部分集合の一例を示す。例えば図16(a)に示す部分集合 c_1 の学習は、部分集合 c_1 に属す第一の特徴量をポジティブデータ、他の集合に属す第一の特徴量をネガティブデータとし、例えばSVMやパーセプトロンを用いて行われる。部分集合 c_1 以外の部分集合 c_2 から c_7 についても同様に学習することができる。

30

【0056】

重み行列生成部1124は、低次元空間の第一の特徴量の部分集合を入力して、第一の特徴量に対する、部分集合 n ($n=0, 1, \dots, N$)に属すカテゴリ m ($m=0, 1, \dots, M$)の重み w_{mn} を行列要素とする重み行列 W を算出する。行列要素 w_{mn} は次式によって算出される。

$$w_{mn} = \text{Num}_{nm} / \text{Num}_n \quad \dots (1)$$

ここで、 Num_{nm} は部分集合 n に属す特徴量のうちカテゴリ m に属す特徴量の数、

Num_n は部分集合 n に属する特徴量の総数。

【0057】

40

パラメータ格納部1125には、境界学習部1123から部分集合ごとの識別パラメータが格納され、重み行列生成部1124から重み行列 W が格納される。パラメータ格納部1125は、これらパラメータを一時格納し、識別部32からの要求に応じて、これらパラメータを識別部32に出力する。

【0058】

識別部12

図15に示すように、実施例3の識別部12の構成は、図1に示す実施例1の構成と異なり、識別部12は、第一の特徴量算出部121と識別部124を有する。第一の特徴量算出部121の動作、機能は実施例1と同様である。

【0059】

識別部324は、第一の特徴量段出部121が算出した第一の特徴量、並びに、学習部11の多

50

様体学習部112が算出した部分集合ごとの識別パラメータおよび重み行列Wを入力する。そして、識別部12に入力されるカテゴリが未知の識別対象の動画シーケンスのカテゴリを識別する。

【 0 0 6 0 】

ある動画シーケンスから抽出されたL個の第一の特徴量を X ($=x_1, x_2, \dots, x_n, \dots, x_L$) とする。この場合、 n 番目の第一の特徴量 x_n が属す部分集合 n は、部分集合ごとの識別パラメータを用いて、それぞれ($n_1, n_2, \dots, n_n, \dots, n_L$)のように算出される。第一の特徴量 X とカテゴリmの類似度 S_m は、L個の部分集合($n_1, n_2, \dots, n_n, \dots, n_L$)のカテゴリmに対する重みの合計値 S_m として次式により算出される。

$$S_m = \mathcal{L}_{W_{m,n}} \dots (2)$$

【 0 0 6 1 】

さらに、識別結果 m' は次式により算出される。

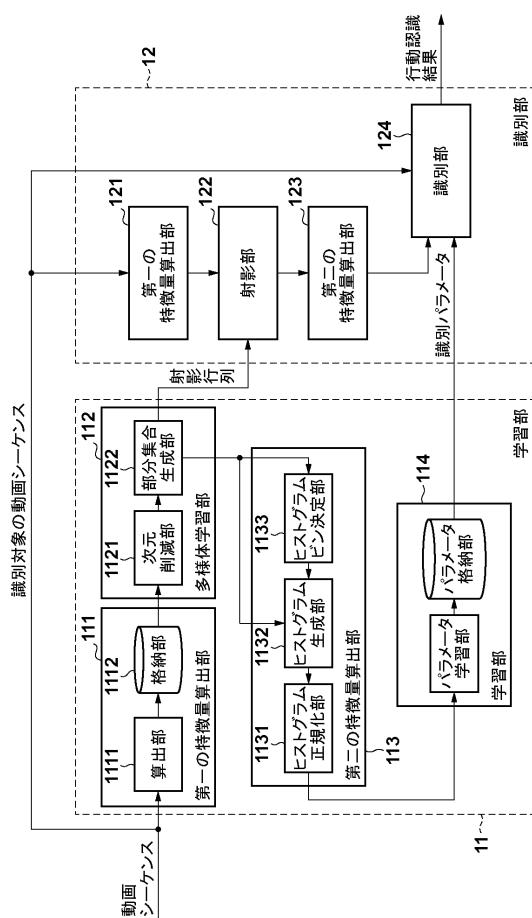
$$m' = \arg \max_m S_m \quad \dots (3)$$

【 0 0 6 2 】

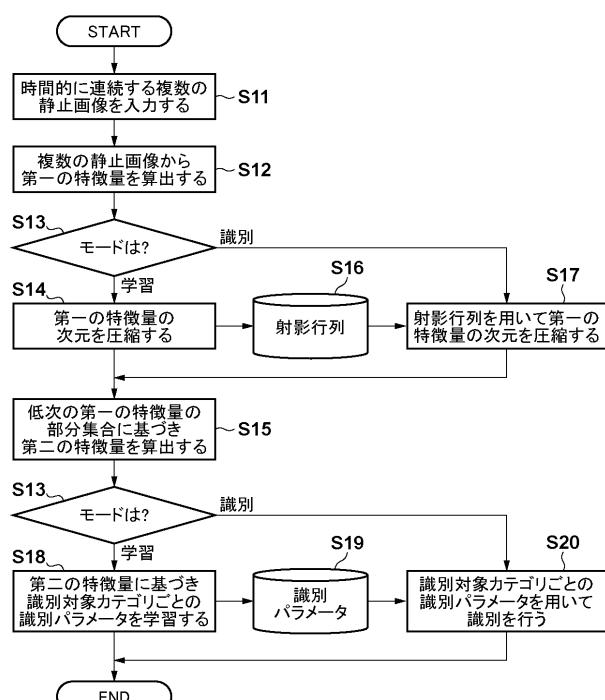
[その他の実施例]

また、本発明は、以下の処理を実行することによっても実現される。即ち、上述した実施形態の機能を実現するソフトウェア（プログラム）を、ネットワーク又は各種記録媒体を介してシステム或いは装置に供給し、そのシステムあるいは装置のコンピュータ（又はCPUやMPU等）がプログラムを読み出して実行する処理である。

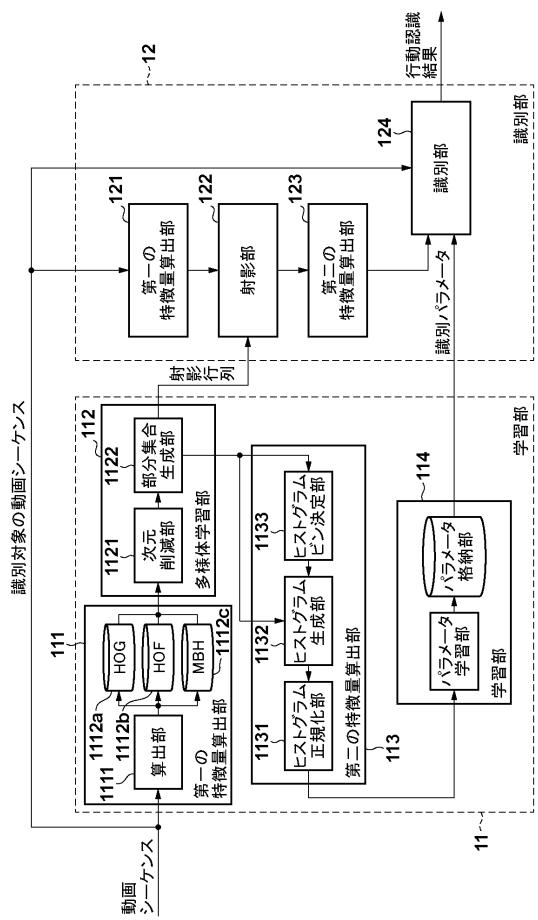
〔 义 1 〕



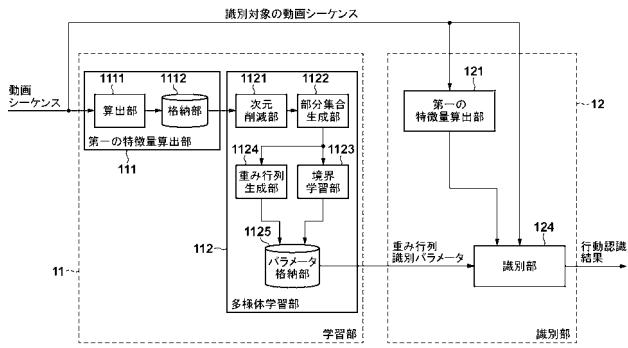
【図2】



【図 8】



【図 15】



【図 3】

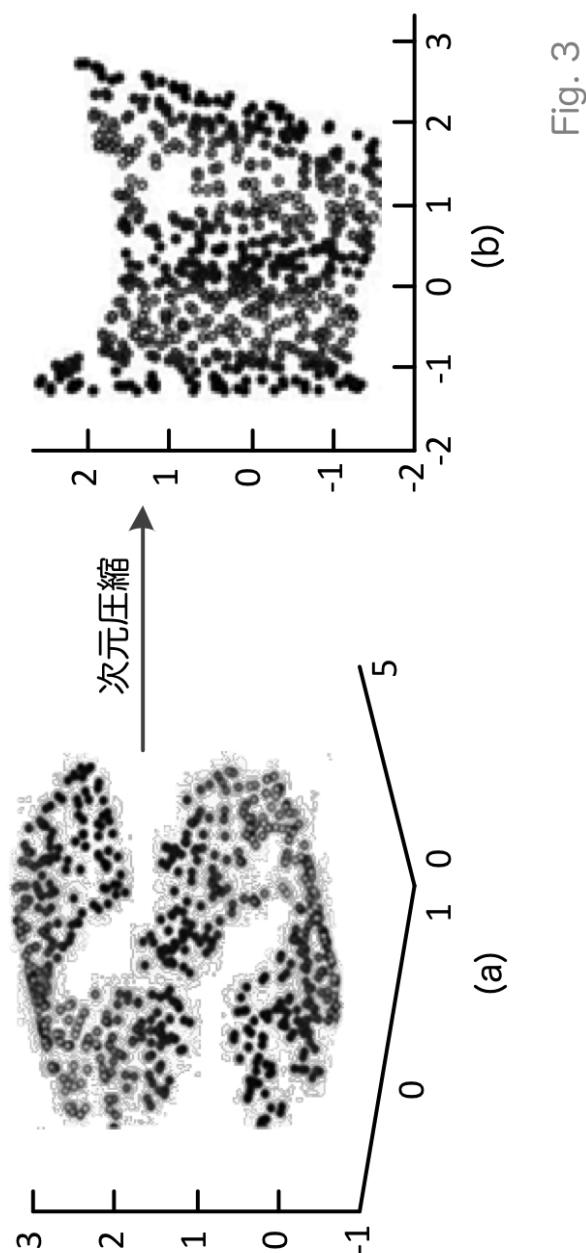


Fig. 3

【図4】

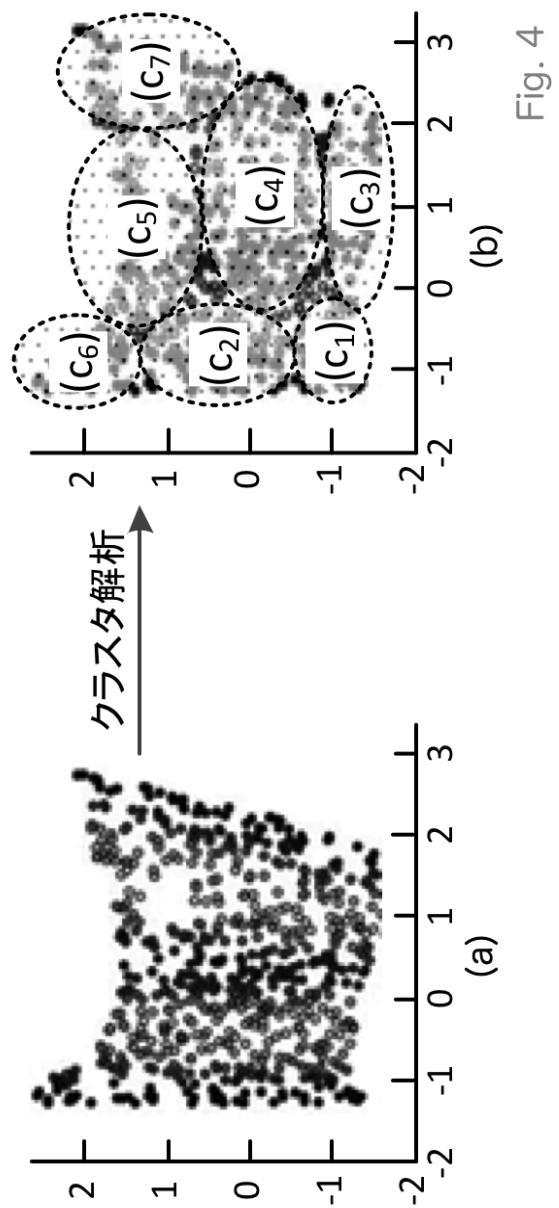


Fig. 4

【図 5】

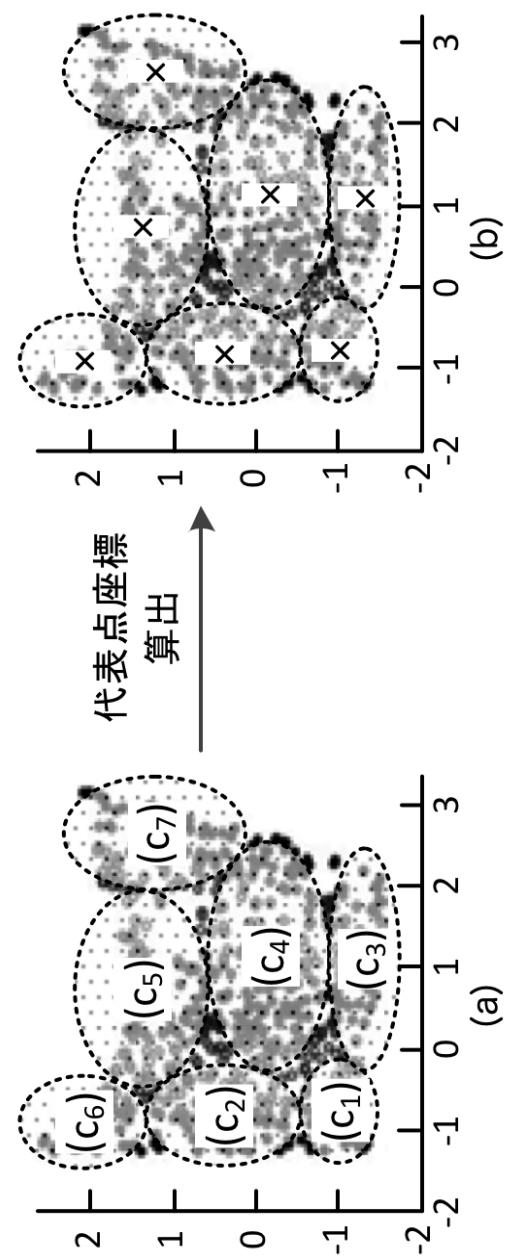


Fig. 5

【図 6】

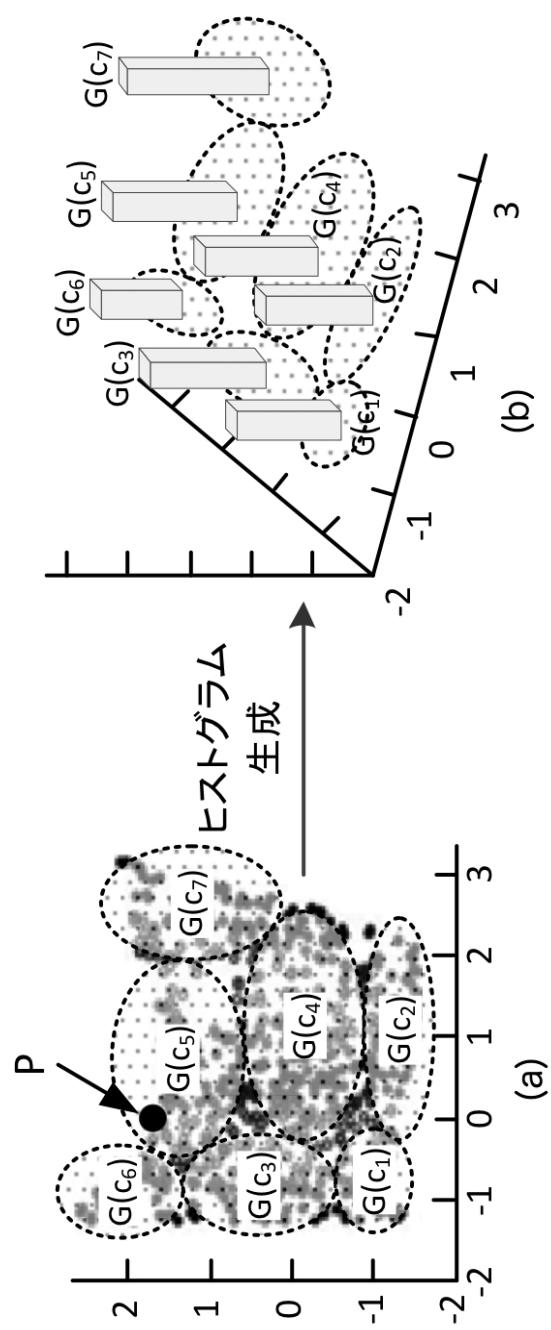


Fig. 6

【図 7】

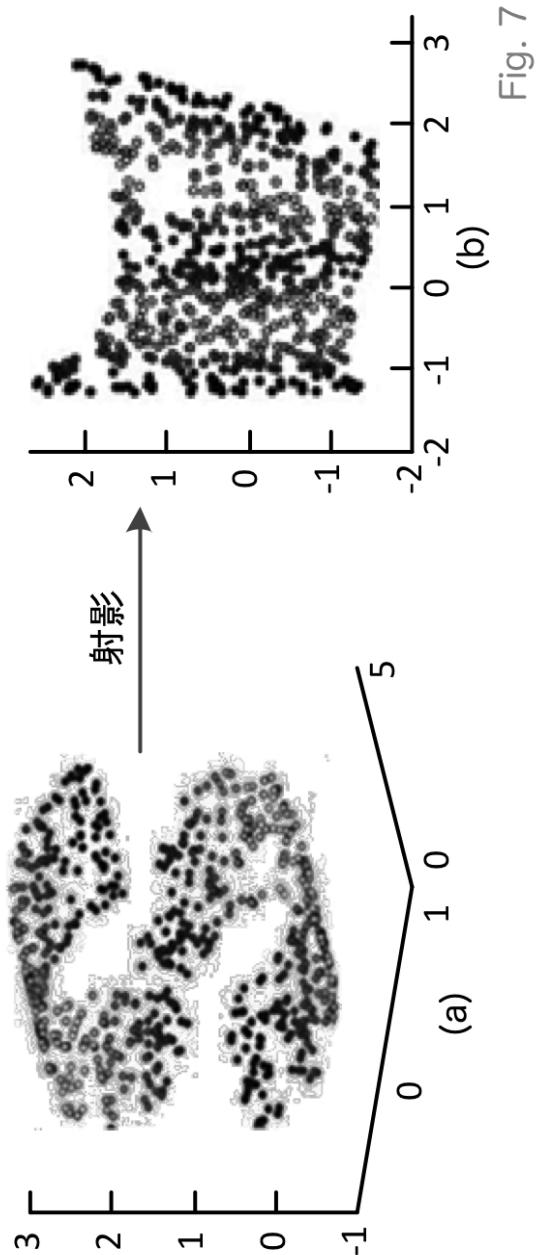


Fig. 7

【図 9】

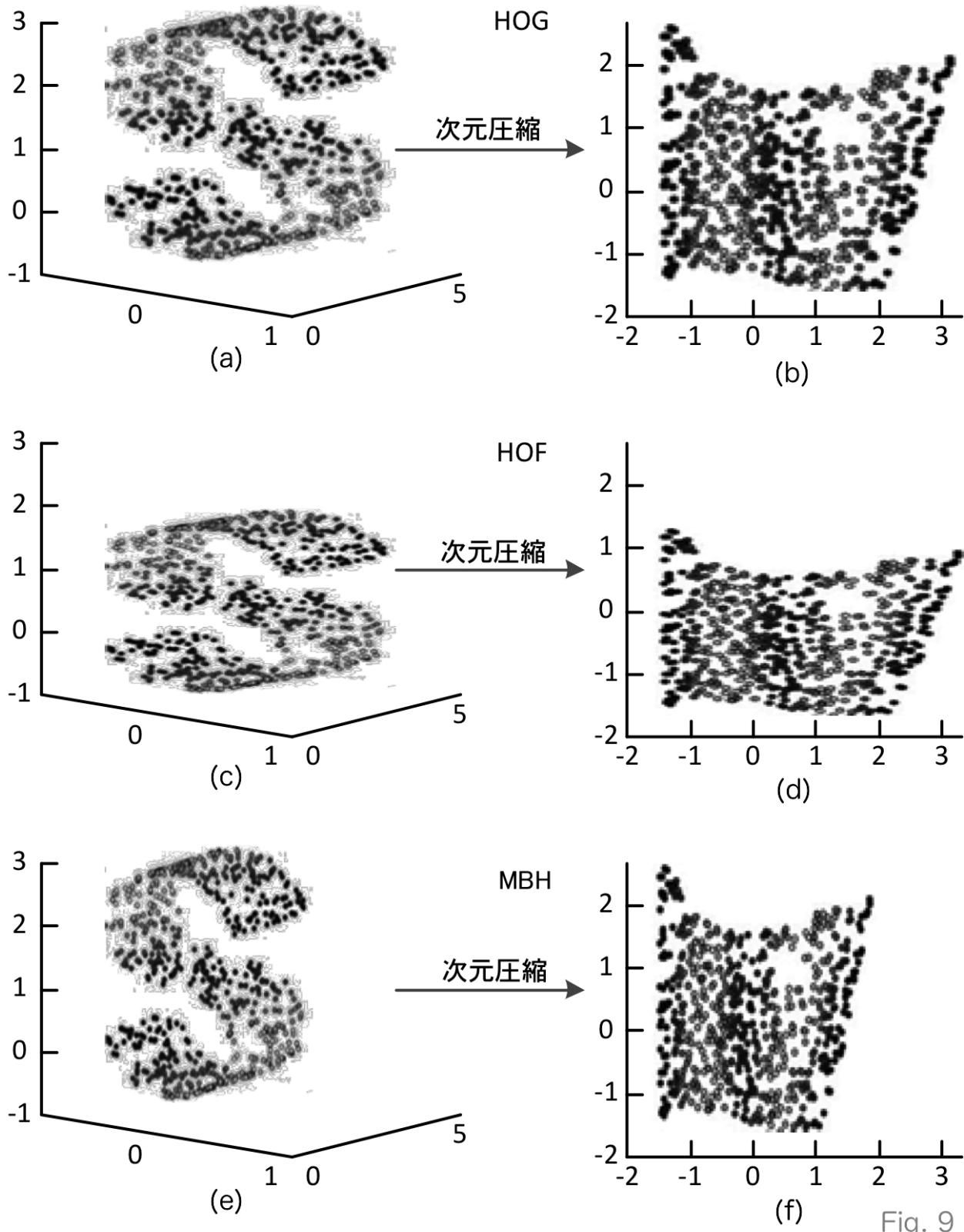


Fig. 9

【図 10】

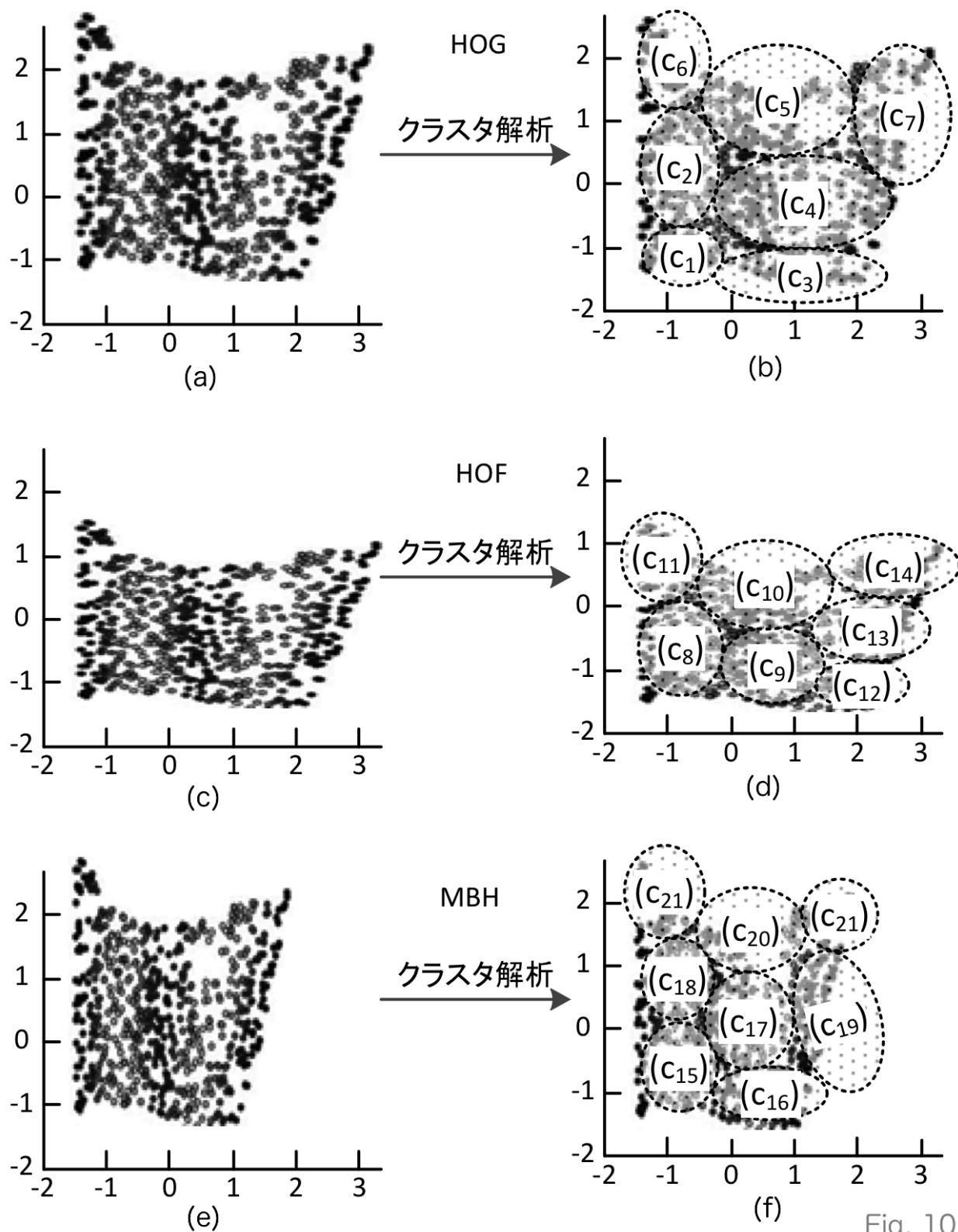


Fig. 10

【図 11】

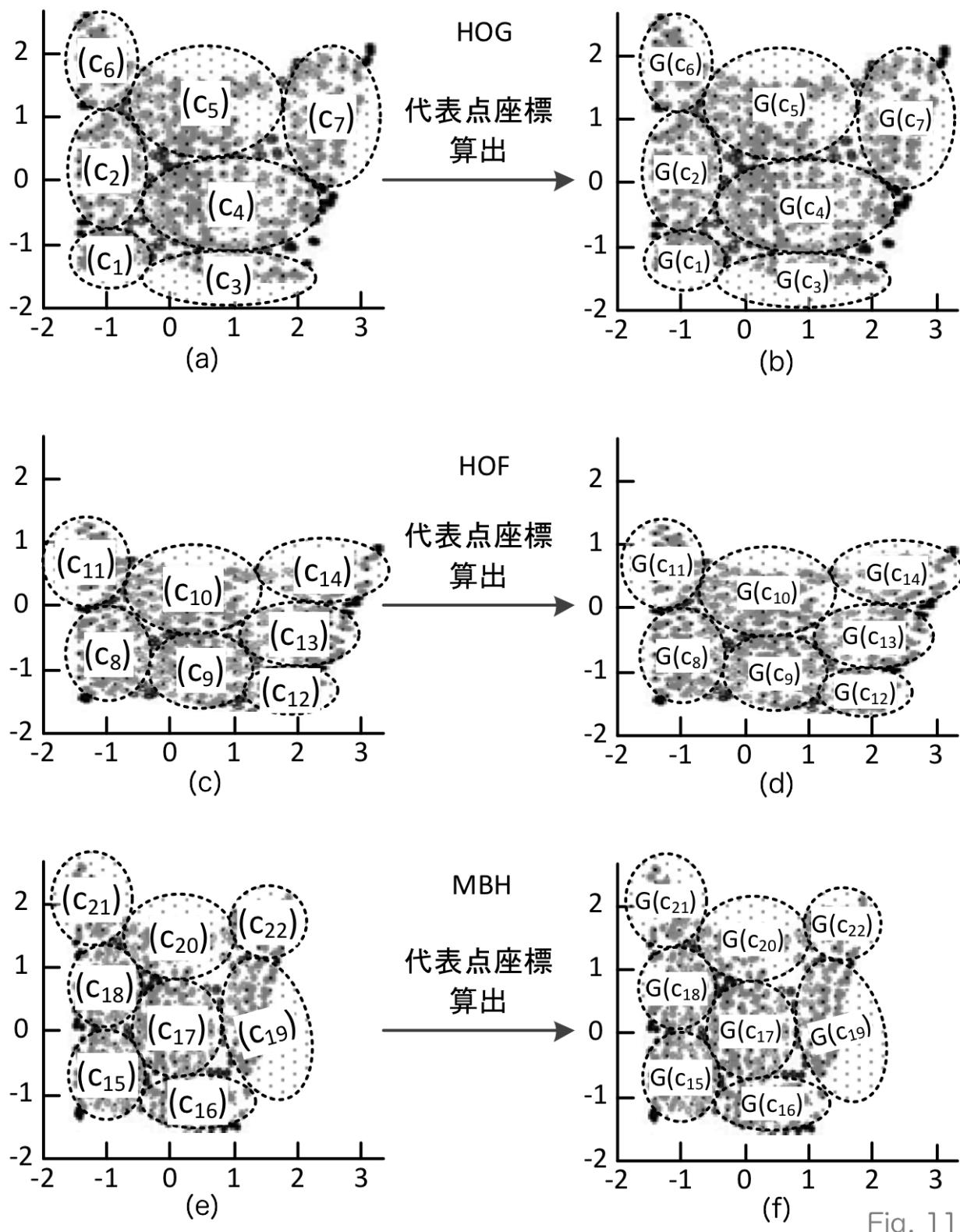


Fig. 11

【図 12】

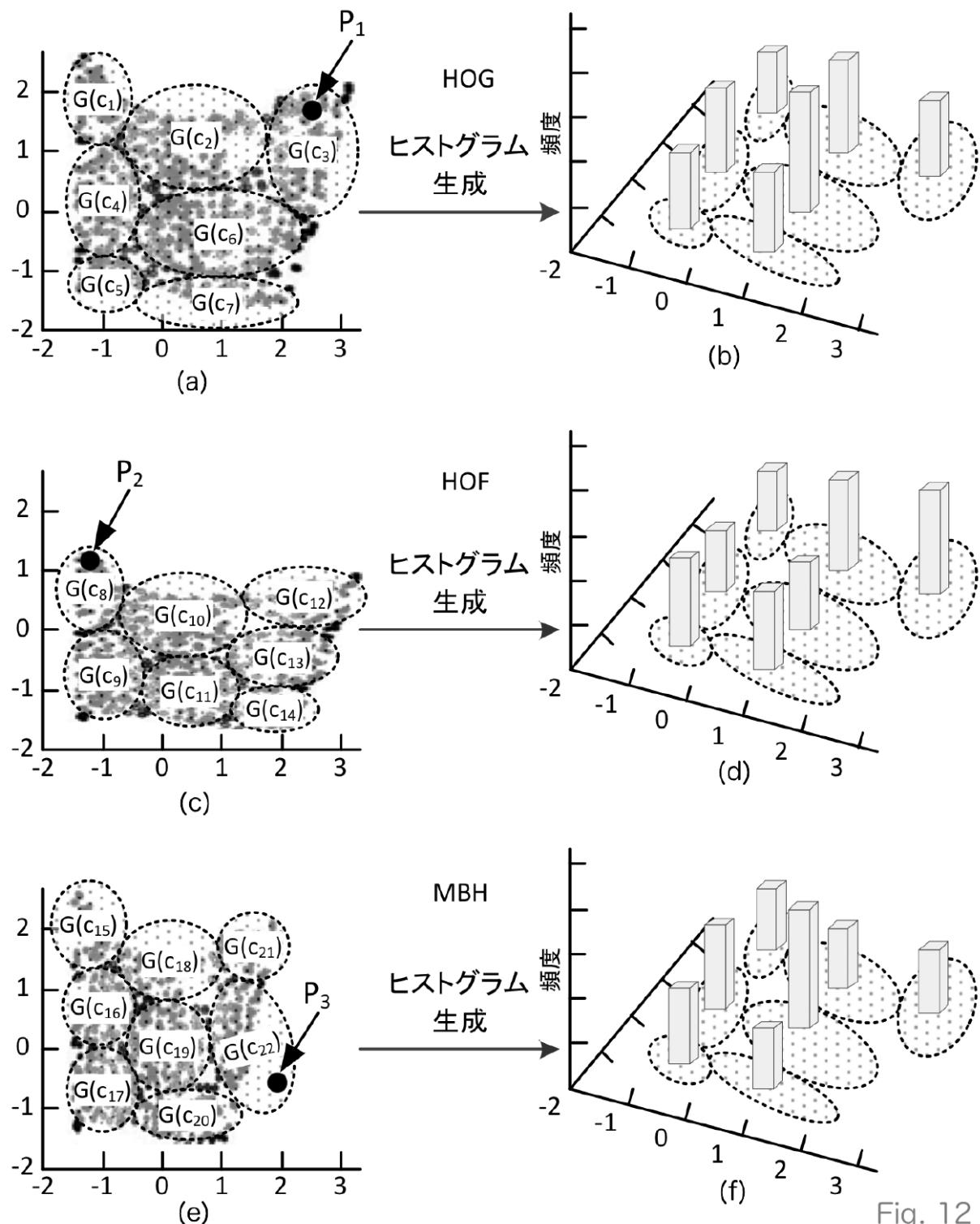


Fig. 12

【図 13】

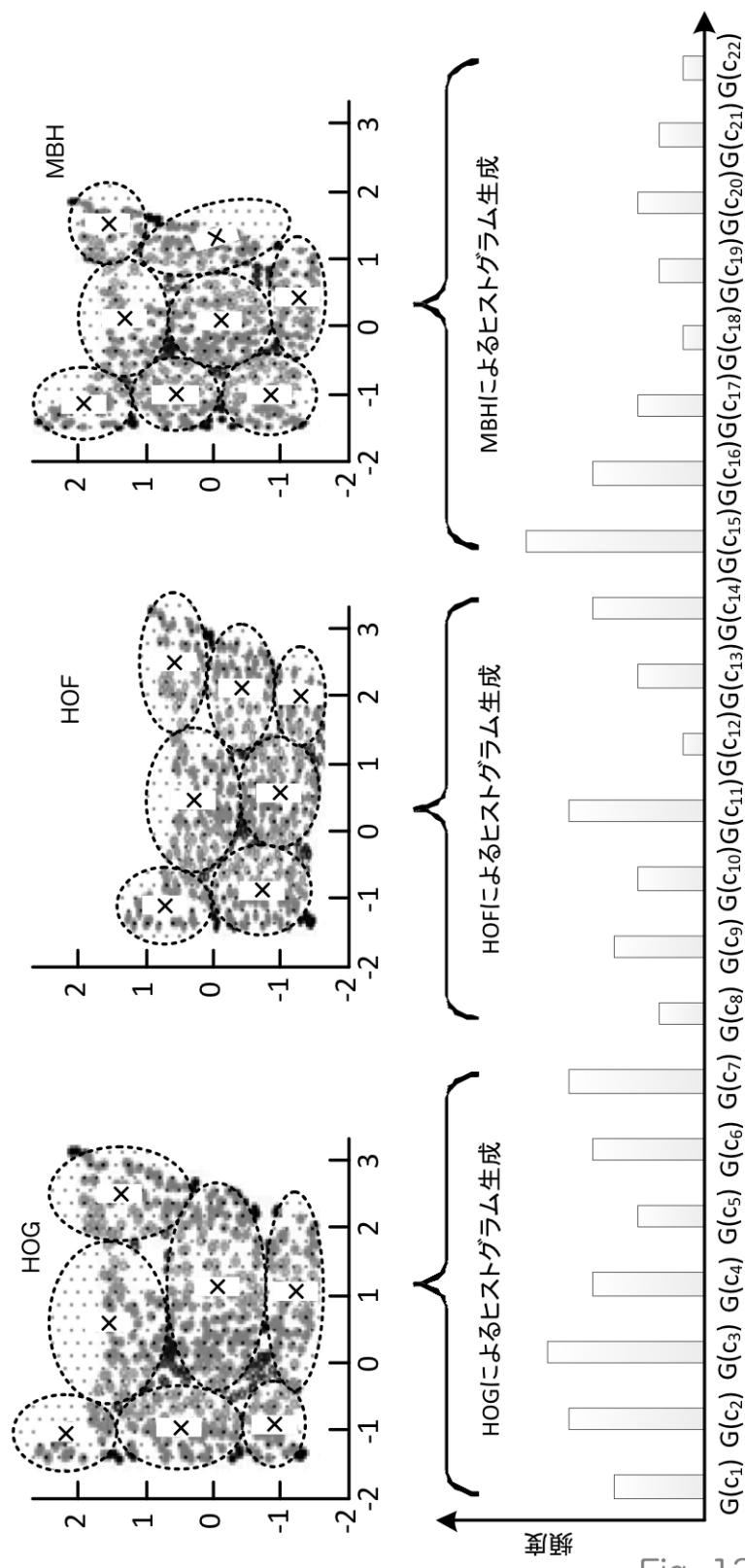


Fig. 13

【図 14】

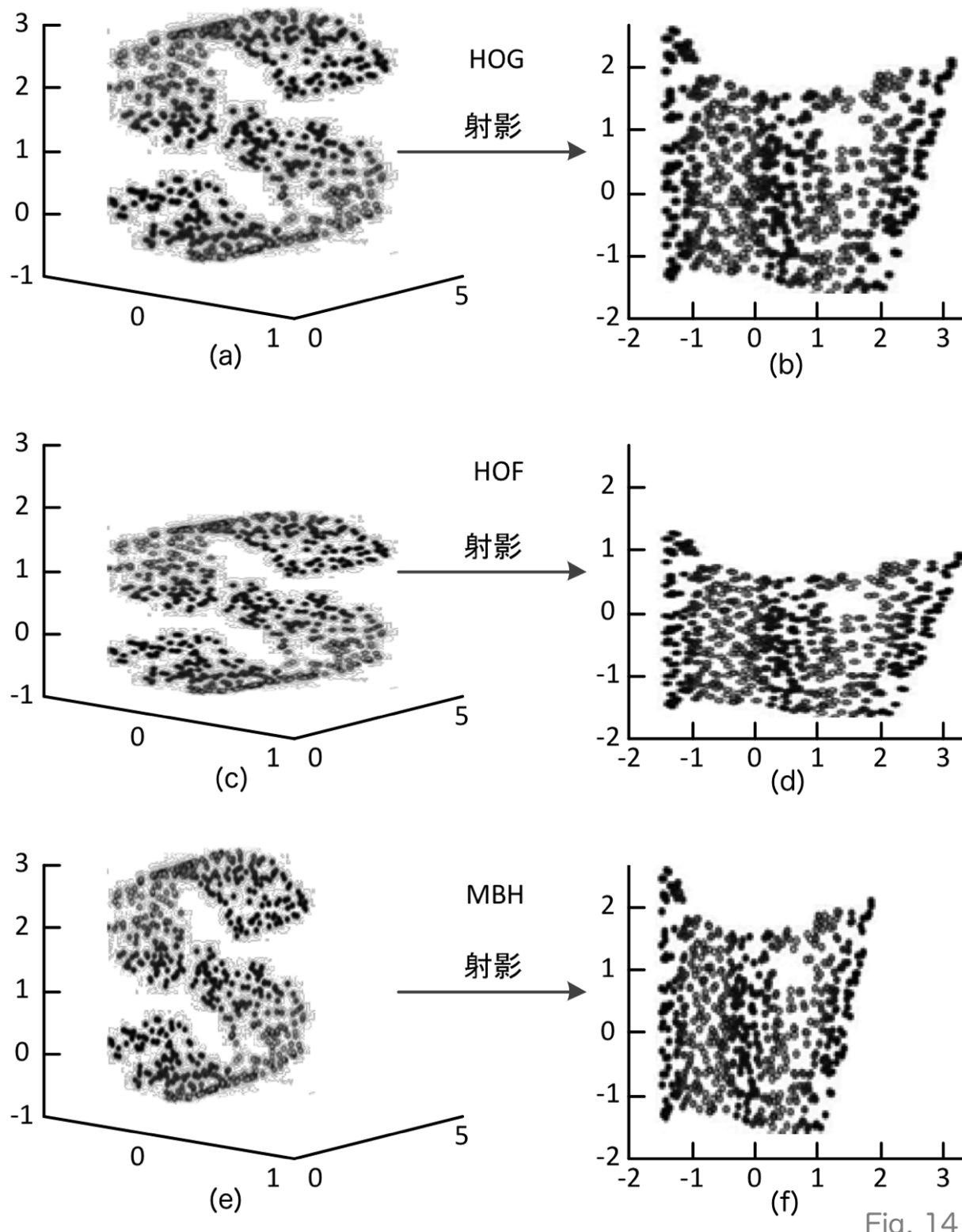
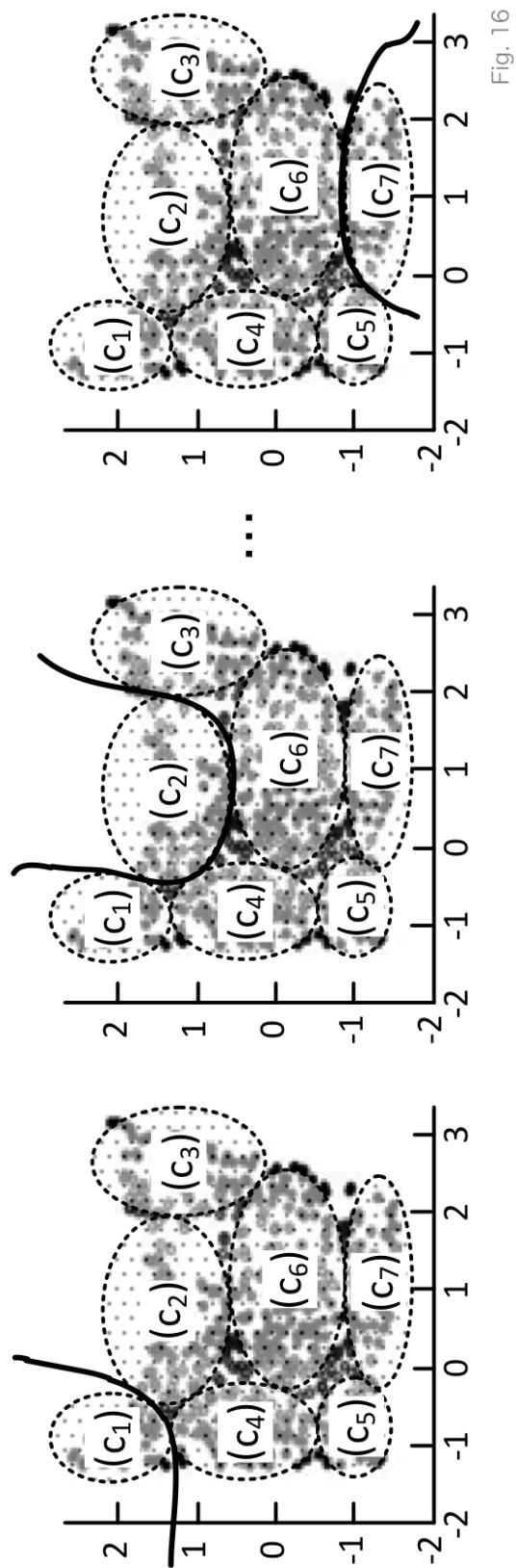


Fig. 14

【図 16】



フロントページの続き

(72)発明者 山中 正雄
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

(72)発明者 真継 優和
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) 5C054 CC02 EA01 FC13 FC16 HA19
5L096 EA11 EA39 FA35 FA38 FA66 FA67 JA11 JA22 KA04 MA07