

(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2017-182437

(P2017-182437A)

(43) 公開日 平成29年10月5日(2017.10.5)

(51) Int.Cl. F I テーマコード (参考)
G06T 7/20 (2017.01) G06T 7/20 200Z 5L096
 G06T 7/20 A

審査請求 未請求 請求項の数 12 O L (全 24 頁)

(21) 出願番号 特願2016-68436 (P2016-68436)
 (22) 出願日 平成28年3月30日 (2016. 3. 30)

(71) 出願人 591261509
 株式会社エクス・リサーチ
 東京都千代田区外神田1丁目18番13号
 (71) 出願人 504174135
 国立大学法人九州工業大学
 福岡県北九州市戸畑区仙水町1番1号
 (74) 代理人 100096655
 弁理士 川井 隆
 (74) 代理人 100091225
 弁理士 仲野 均
 (72) 発明者 山田 英夫
 東京都千代田区外神田1丁目18番13号
 秋葉原ダイビル7階 株式会社エクス
 ・リサーチ内

最終頁に続く

(54) 【発明の名称】 画像認識装置、移動体装置、及び画像認識プログラム

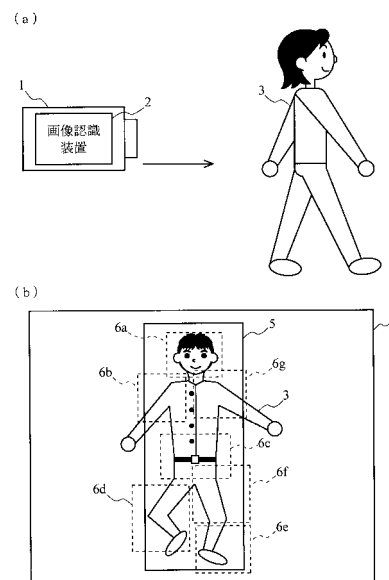
(57) 【要約】

【課題】 頑健性の高い画像認識を行う。

【解決手段】 画像認識装置 2 は、対象者 3 の全身を囲む全体観測領域 5 と、対象者 3 の特徴的な部分を囲む部分観測領域 6 a ~ 6 g を、画像 4 で対象者 3 が写っていると推測される箇所に設定する。画像認識装置 2 は、全体観測領域 5 と部分観測領域 6 a ~ 6 g の画像を切り取り、これらと予め学習した画像との類似度を HOG 特徴量と色分布特徴量の 2 つの画像特徴量を組合せて計算する。画像認識装置 2 は、HOG 特徴量と色分布特徴量を組合せる最適な割合 (比率) を全体観測領域 5 と部分観測領域 6 a ~ 6 g について個別に計算する。この割合の決定は、HOG 特徴量と色分布特徴量を組合せる重みを設定する重みパラメータ i を状態ベクトルに含めてパーティクルフィルタによる全数探索の対象とすることにより行う。

【選択図】 図 1

【図 1】



【特許請求の範囲】**【請求項 1】**

画像を取得する画像取得手段と、

前記取得した画像において追跡対象を観測する観測領域を、一時刻前の画像で前記追跡対象の検出に用いた観測領域に基づいて設定する観測領域設定手段と、

前記設定した観測領域から複数種類の画像特徴量を取得する画像特徴量取得手段と、

前記取得した画像特徴量ごとに基準特徴量との特徴相関係数を取得する特徴相関係数取得手段と、

前記取得した画像特徴量ごとの特徴相関係数を組合せる割合を、一時刻前の画像で前記追跡対象の検出に用いた割合に基づいて設定する割合設定手段と、

前記設定した割合で前記特徴相関係数を組合せて前記観測領域に対する領域相関係数を取得する領域相関係数取得手段と、

前記観測領域設定手段が設定する観測領域と、前記割合設定手段が設定する割合を変化させながら、前記領域相関係数取得手段が取得する領域相関係数を用いて前記観測領域で観測している対象の尤度を繰り返し取得する尤度取得手段と、

前記尤度取得手段で得られた複数の尤度のうち、所定の条件を満たす尤度の取得元となった観測領域で前記追跡対象を検出する検出手段と、

を具備したことを特徴とする画像認識装置。

【請求項 2】

前記観測領域設定手段は、前記観測領域として、前記追跡対象の全体を観測する全体観測領域と、前記追跡対象の部分を観測する複数の部分観測領域の組を設定し、

前記画像特徴量取得手段は、前記全体観測領域、及び前記部分観測領域ごとに前記複数種類の画像特徴量を取得し、

前記特徴相関係数取得手段は、前記全体観測領域、及び前記部分観測領域ごとに前記画像特徴量ごとの特徴相関係数を取得し、

前記割合設定手段は、前記全体観測領域、及び前記部分観測領域ごとに前記割合を設定し、

前記領域相関係数取得手段は、前記全体観測領域、及び前記部分観測領域ごとに前記領域相関係数を取得し、

前記尤度取得手段は、前記全体観測領域、及び前記部分観測領域ごとの領域相関係数を用いて前記尤度を取得することを特徴とする請求項 1 に記載の画像認識装置。

【請求項 3】

前記全体観測領域、及び前記部分観測領域ごとの領域相関係数に、一時刻前の画像で前記追跡対象の検出に用いた重みに基づいて重みを設定する重み設定手段を具備し、

前記尤度取得手段は、前記重みも変化させながら、前記尤度を繰り返し取得することを特徴とする請求項 2 に記載の画像認識装置。

【請求項 4】

前記尤度取得手段は、前記領域相関係数が所定の閾値を下回る場合は、当該領域相関係数を除外して前記尤度を取得することを特徴とする請求項 3 に記載の画像認識装置。

【請求項 5】

前記重みの総和が小さい前記全体観測領域と前記部分観測領域の組ほど前記尤度を小さく補正する尤度補正手段を具備したことを特徴とする請求項 3、又は請求項 4 に記載の画像認識装置。

【請求項 6】

前記割合設定手段が設定する割合が一方に偏るほど前記尤度が小さくなるように補正を行う偏り補正手段を具備したことを特徴とする請求項 3、請求項 4、又は請求項 5 に記載の画像認識装置。

【請求項 7】

前記割合設定手段と前記重み設定手段は、前記追跡対象の検出に失敗した場合に、それぞれ前記割合と前記重みを一時刻前の画像とは無関係な任意の値に設定することを特徴と

10

20

30

40

50

する請求項 3 から請求項 6 までのうちの何れか 1 の請求項に記載の画像認識装置。

【請求項 8】

前記複数種類の画像特徴量は、前記取得した画像の輝度勾配分布特徴量と、色分布特徴量であることを特徴とする請求項 1 から請求項 7 に記載の画像認識装置。

【請求項 9】

前記尤度が満たす所定の条件は、尤度の最大値であることを特徴とする請求項 1 から請求項 8 までのうちの何れか 1 の請求項に記載の画像認識装置。

【請求項 10】

前記領域相関係数取得手段は、前記設定した割合による線形和によって、前記取得した特徴相関係数を組合せることを特徴とする請求項 1 から請求項 9 までのうちの何れか 1 の請求項に記載の画像認識装置。

10

【請求項 11】

請求項 1 から請求項 10 までのうちの何れか 1 の請求項に記載の画像認識装置を用いて対象を検出する検出手段と、

前記検出した対象に対して所定の位置に移動する移動手段と、
を具備したことを特徴とする移動体装置。

【請求項 12】

画像を取得する画像取得機能と、

前記取得した画像において追跡対象を観測する観測領域を、一時刻前の画像で前記追跡対象の検出に用いた観測領域に基づいて設定する観測領域設定機能と、

20

前記設定した観測領域から複数種類の画像特徴量を取得する画像特徴量取得機能と、

前記取得した画像特徴量ごとに基準特徴量との特徴相関係数を取得する特徴相関係数取得機能と、

前記取得した画像特徴量ごとの特徴相関係数を組合せる割合を、一時刻前の画像で前記追跡対象の検出に用いた割合に基づいて設定する割合設定機能と、

前記設定した割合で前記特徴相関係数を組合せて前記観測領域に対する領域相関係数を取得する領域相関係数取得機能と、

前記観測領域設定機能が設定する観測領域と、前記割合設定機能が設定する割合を変化させながら、前記領域相関係数取得機能が取得する領域相関係数を用いて前記観測領域で観測している対象の尤度を繰り返し取得する尤度取得機能と、

30

前記尤度取得機能で得られた複数の尤度のうち、所定の条件を満たす尤度の取得元となった観測領域で前記追跡対象を検出する検出機能と、

をコンピュータで実現する画像認識プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、画像認識装置、移動体装置、及び画像認識プログラムに関し、例えば、画像に含まれる追跡対象を認識するものに関する。

【背景技術】

【0002】

近年、画像に写った対象を認識し、その移動を追跡する画像認識装置の需要が高まっている。

40

このような画像認識装置は、例えば、車両やロボットに搭載されて、歩行者や障害物を避けたり、あるいは、監視カメラに搭載されて、不審者を追跡したりするのに用いられる。

【0003】

画像中に対象を認識・追跡する技術は、各種提案されているが、特に優れたものに特許文献 1 の「物体識別装置、方法及びプログラム」で使用されている DPM (Deformable Part Model) がある。

この技術は、追跡対象人物の全身を認識するための全体的な領域と、追跡対象人物の特

50

微的な一部分（頭部、腰部、・・・など）を認識するための複数の部分的な領域を組み合わせることにより、例えば追跡対象人物が姿勢変化して全体的な領域での認識が困難になった場合でも、部分的な領域での認識でこれを補い、画像中の追跡対象人物を特定するものである。

【 0 0 0 4 】

ところで、移動体（例えば人物）を追跡する場合、追跡対象が人混みに紛れたり、照明環境が変化したりなど、追跡対象の撮影状態は時間と共に様々に変化する。

一方、追跡対象を画像で認識する方法には、HOG特徴量と呼ばれる対象のエッジ検出を用いた方法や対象の色分布特徴量を用いた方法など各種のものがあるが、これらは認識が得意なシーン（場面）と不得意なシーンがあり、これらの方法を単独で用いると、追跡対象の撮影状態の変化に対応できず、追跡が破綻するという問題があった。

【 先行技術文献 】

【 特許文献 】

【 0 0 0 5 】

【 特許文献 1 】 特開 2 0 1 5 - 2 3 0 5 5 9 号 公 報

【 発明の概要 】

【 発明が解決しようとする課題 】

【 0 0 0 6 】

本発明は、頑健性の高い画像認識を行うことを目的とする。

【 課題を解決するための手段 】

【 0 0 0 7 】

(1) 本発明は、前記目的を達成するために、請求項 1 に記載の発明では、画像を取得する画像取得手段と、前記取得した画像において追跡対象を観測する観測領域を、一時刻前の画像で前記追跡対象の検出に用いた観測領域に基づいて設定する観測領域設定手段と、前記設定した観測領域から複数種類の画像特徴量を取得する画像特徴量取得手段と、前記取得した画像特徴量ごとに基準特徴量との特徴相関係数を取得する特徴相関係数取得手段と、前記取得した画像特徴量ごとの特徴相関係数を組合せる割合を、一時刻前の画像で前記追跡対象の検出に用いた割合に基づいて設定する割合設定手段と、前記設定した割合で前記特徴相関係数を組合せて前記観測領域に対する領域相関係数を取得する領域相関係数取得手段と、前記観測領域設定手段が設定する観測領域と、前記割合設定手段が設定する割合を変化させながら、前記領域相関係数取得手段が取得する領域相関係数を用いて前記観測領域で観測している対象の尤度を繰り返し取得する尤度取得手段と、前記尤度取得手段で得られた複数の尤度のうち、所定の条件を満たす尤度の取得元となった観測領域で前記追跡対象を検出する検出手段と、を具備したことを特徴とする画像認識装置を提供する。

(2) 請求項 2 に記載の発明では、前記観測領域設定手段が、前記観測領域として、前記追跡対象の全体を観測する全体観測領域と、前記追跡対象の部分を観測する複数の部分観測領域の組を設定し、前記画像特徴量取得手段は、前記全体観測領域、及び前記部分観測領域ごとに前記複数種類の画像特徴量を取得し、前記特徴相関係数取得手段は、前記全体観測領域、及び前記部分観測領域ごとに前記画像特徴量ごとの特徴相関係数を取得し、前記割合設定手段は、前記全体観測領域、及び前記部分観測領域ごとに前記割合を設定し、前記領域相関係数取得手段は、前記全体観測領域、及び前記部分観測領域ごとに前記領域相関係数を取得し、前記尤度取得手段は、前記全体観測領域、及び前記部分観測領域ごとの領域相関係数を用いて前記尤度を取得することを特徴とする請求項 1 に記載の画像認識装置を提供する。

(3) 請求項 3 に記載の発明では、前記全体観測領域、及び前記部分観測領域ごとの領域相関係数に、一時刻前の画像で前記追跡対象の検出に用いた重みに基づいて重みを設定する重み設定手段を具備し、前記尤度取得手段は、前記重みも変化させながら、前記尤度を繰り返し取得することを特徴とする請求項 2 に記載の画像認識装置を提供する。

(4) 請求項 4 に記載の発明では、前記尤度取得手段が、前記領域相関係数が所定の閾値

10

20

30

40

50

を下回る場合は、当該領域相関係数を除外して前記尤度を取得することを特徴とする請求項 3 に記載の画像認識装置を提供する。

(5) 請求項 5 に記載の発明では、前記重みの総和が小さい前記全体観測領域と前記部分観測領域の組ほど前記尤度を小さく補正する尤度補正手段を具備したことを特徴とする請求項 3、又は請求項 4 に記載の画像認識装置を提供する。

(6) 請求項 6 に記載の発明では、前記割合設定手段が設定する割合が一方に偏るほど前記尤度が小さくなるように補正を行う偏り補正手段を具備したことを特徴とする請求項 3、請求項 4、又は請求項 5 に記載の画像認識装置を提供する。

(7) 請求項 7 に記載の発明では、前記割合設定手段と前記重み設定手段は、前記追跡対象の検出に失敗した場合に、それぞれ前記割合と前記重みを一時刻前の画像とは無関係な任意の値に設定することを特徴とする請求項 3 から請求項 6 までのうちの何れか 1 の請求項に記載の画像認識装置を提供する。

(8) 請求項 8 に記載の発明では、前記複数種類の画像特徴量が、前記取得した画像の輝度勾配分布特徴量と、色分布特徴量であることを特徴とする請求項 1 から請求項 7 に記載の画像認識装置を提供する。

(9) 請求項 9 に記載の発明では、前記尤度が満たす所定の条件は、尤度の最大値であることを特徴とする請求項 1 から請求項 8 までのうちの何れか 1 の請求項に記載の画像認識装置を提供する。

(10) 請求項 10 に記載の発明では、前記領域相関係数取得手段が、前記設定した割合による線形和によって、前記取得した特徴相関係数を組合せることを特徴とする請求項 1 から請求項 9 までのうちの何れか 1 の請求項に記載の画像認識装置を提供する。

(11) 請求項 11 に記載の発明では、請求項 1 から請求項 10 までのうちの何れか 1 の請求項に記載の画像認識装置を用いて対象を検出する検出手段と、前記検出した対象に対して所定の位置に移動する移動手段と、を具備したことを特徴とする移動体装置を提供する。

(12) 請求項 12 に記載の発明では、画像を取得する画像取得機能と、前記取得した画像において追跡対象を観測する観測領域を、一時刻前の画像で前記追跡対象の検出に用いた観測領域に基づいて設定する観測領域設定機能と、前記設定した観測領域から複数種類の画像特徴量を取得する画像特徴量取得機能と、前記取得した画像特徴量ごとに基準特徴量との特徴相関係数を取得する特徴相関係数取得機能と、前記取得した画像特徴量ごとの特徴相関係数を組合せる割合を、一時刻前の画像で前記追跡対象の検出に用いた割合に基づいて設定する割合設定機能と、前記設定した割合で前記特徴相関係数を組合せて前記観測領域に対する領域相関係数を取得する領域相関係数取得機能と、前記観測領域設定機能が設定する観測領域と、前記割合設定機能が設定する割合を変化させながら、前記領域相関係数取得機能が取得する領域相関係数を用いて前記観測領域で観測している対象の尤度を繰り返し取得する尤度取得機能と、前記尤度取得機能で得られた複数の尤度のうち、所定の条件を満たす尤度の取得元となった観測領域で前記追跡対象を検出する検出機能と、をコンピュータで実現する画像認識プログラムを提供する。

【発明の効果】

【0008】

本発明によれば、複数種類の画像特徴量を組合せることにより、頑健性の高い画像認識を行うことができる。

【図面の簡単な説明】

【0009】

【図 1】画像認識装置の使用形態や観測領域について説明するための図である。

【図 2】HOG 特徴量の概念を説明するための図である。

【図 3】色分布特徴量の概念を説明するための図である。

【図 4】実験結果を説明するための図である。

【図 5】重みパラメータ W_i について説明するための図である。

【図 6】追跡の破綻とその復帰を説明するための図である。

【図 7】画像認識装置のハードウェア構成を説明するための図である。

【図 8】画像認識処理の手順を説明するためのフローチャートである。

【図 9】尤度計算処理を説明するための図である。

【図 10】各実施形態で使用する数式を表した説明図である。

【発明を実施するための形態】

【0010】

(1) 実施形態の概要

図 1 (b) に示したように、画像認識装置 2 は、対象者 3 の全身を囲む全体観測領域 5 と、対象者 3 の特徴的な部分を囲む部分観測領域 6 a ~ 6 g を、画像 4 で対象者 3 が写っていると推測される箇所に設定する。

10

画像認識装置 2 は、全体観測領域 5 と部分観測領域 6 a ~ 6 g の画像を切り取り、これらと予め学習した画像との類似度を HOG 特徴量と色分布特徴量の 2 つの画像特徴量を組合せて計算する。これら 2 つの画像特徴量は、得意・不得意とするシーンが異なるが、これを組合せることにより両者を相補的に作用させることができる。

【0011】

また、画像認識装置 2 は、HOG 特徴量と色分布特徴量を組合せる最適な割合 (比率) を全体観測領域 5 と部分観測領域 6 a ~ 6 g について個別に計算する。

この割合の決定は、HOG 特徴量と色分布特徴量を組合せる重みを設定する重みパラメータ i を状態ベクトルに含めてパーティクルフィルタ (粒子フィルタ) による全数探索の対象とすることにより行う。

20

このように、画像認識装置 2 は、2 つの画像特徴量をシーンに適した重み付けで組合せることにより、例えば、明暗の変化や対象者 3 の周囲の歩行者密度の変化などの撮影環境の変化に対して頑健性を有する画像認識、及び追跡を行うことができる。

【0012】

(2) 実施形態の詳細

(第 1 実施形態)

図 1 (a) は、本実施の形態に係る画像認識装置 2 の使用形態を説明するための図である。

カメラ 1 は、例えば、ハンディカムなどで構成された携帯可能な動画撮影カメラである。ユーザは、カメラ 1 を把持し、追跡対象である対象者 3 の後について移動しながら、カメラ 1 で対象者 3 を動画撮影する。

30

【0013】

後述するように、画像認識装置 2 は、例えば、対象者 3 の人混みとの錯綜や明暗の変化などの状態変化に対して頑健であるため、自らも移動しながらこれらの変化を受けつつ、対象者 3 を追跡することができる。

そのため、画像認識装置 2 は、車両やロボットなどの移動体に搭載するのに適している。

【0014】

画像認識装置 2 は、カメラ 1 の内部に内蔵されており、撮影された映像の各フレームの画像において対象者 3 を連続的に検出することにより対象者 3 を画像上で追跡する。

40

なお、画像認識装置 2 をカメラ 1 の外部に設置し、カメラ 1 から画像認識装置 2 に動画を転送してもよい。

以下では、一例として、追跡対象を人物とするが、これは追跡対象を限定するものではなく、任意の移動体を追跡対象とすることができる。

【0015】

図 1 (b) は、画像 4 で対象者 3 を認識する方法を説明するための図である。

画像 4 は、カメラ 1 が撮影して画像認識装置 2 に出力した動画のあるフレームを構成する画像を示している。

このように、画像認識装置 2 は、画像 4 を取得する画像取得手段を備えている。

画像認識装置 2 は、前時刻の画像 4 (一時刻前の画像、ここでは、1 つ前のフレームの

50

画像)で検出された対象者3に基づいて、現時刻の画像4で対象者3が写っていると推測される領域に対象者3の全身を囲む矩形の全体観測領域5と、対象者3の特徴的な部分を囲む矩形の部分観測領域6a~6gを設定する。

【0016】

図の例では、対象者3の頭部、左側上肢部、腰部、左側下肢部、下肢部の下側、右側下肢部、右側上肢部に部分観測領域6a~6gが設定されている。

以下では、部分観測領域6a~6gを特に区別しない場合は、単に部分観測領域6と記し、全体観測領域5と部分観測領域6を特に区別しない場合は、単に観測領域と呼ぶことにする。

【0017】

画像認識装置2は、前時刻の画像4で対象者3を検出した全体観測領域5と部分観測領域6に基づいて、その付近に位置や大きさの異なる全体観測領域5と部分観測領域6をパーティクルフィルタ手法を用いて多数発生させるのであるが、図に示したのは、そのうちの1つの組合せである。

全体観測領域5はルートフィルタ、部分観測領域6はパートフィルタと呼ばれている。

【0018】

このように、画像認識装置2は、取得した画像4において追跡対象を観測する観測領域を、一時刻前の画像で追跡対象の検出に用いた観測領域に基づいて設定する観測領域設定手段を備えている。

そして、当該観測領域設定手段は、観測領域として、追跡対象の全体を観測する全体観測領域5と、追跡対象の部分を観測する複数の部分観測領域6の組を設定している。

【0019】

画像認識装置2は、画像4から、これらの領域で囲まれた部分を切り取り(抽出し)、切り取られた個々の画像が、予め学習した画像とどの程度類似しているかの類似度を表す相関係数(特徴相関係数)を後述の画像特徴量空間における内積によって計算する。

そして、画像認識装置2は、全体観測領域5と部分観測領域6について計算された相関係数を所定の演算により合成(統合)して全体観測領域5、部分観測領域6の組による相関係数(領域相関係数)を計算し、全体観測領域5、部分観測領域6の組による相関係数を用いて尤度を計算する。

尤度は、全体観測領域5と部分観測領域6の組によって観測している対象が、どの程度対象者3であるか、という程度を示す量である。

【0020】

画像認識装置2は、パーティクルフィルタ手法を用いて、大きさや位置などのパラメータが少しずつ異なる全体観測領域5と部分観測領域6a~6gを多数発生させ、色々な組合せに対して尤度を計算する。これにより尤度分布が得られる。

画像認識装置2は、最大尤度となる全体観測領域5と部分観測領域6a~6gの組によって観測している対象を対象者3とすることにより、画像4から対象者3を検出する。

なお、尤度に閾値を設け、尤度が当該閾値を超えた場合に、そのときの全体観測領域5と部分観測領域6a~6gの組によって観測している対象を対象者3とするなど、他の基準を用いてもよい。

【0021】

次に、パーティクルフィルタ手法の概要について説明する。

前時刻の画像4で、全体観測領域5と部分観測領域6a~6gが定まっており、これを初期状態とする。

この状態を、一例として、状態ベクトル $X = (全体観測領域5の位置、全体観測領域5の大きさ、各部分観測領域6の位置、\dots)$ で表す。

【0022】

現時刻の画像4での全体観測領域5、部分観測領域6は、前時刻のものとはさほど異なっていないと考えられるため、状態ベクトル X の各状態変数(状態ベクトルの成分)について、前時刻での値の付近にホワイトノイズを発生させ(前時刻での値を平均値とする正規

10

20

30

40

50

性白色雑音を発生させる)、これを前時刻での値に重畳することにより現時刻の値であると推測する。ホワイトノイズを重畳した各値がパーティクルに相当する。

【0023】

このように画像認識装置2は、対象者3の状態の遷移を、前時刻の値から決まり、それ以前の履歴を引きずらないマルコフ過程と仮定し、上記の状態遷移モデルにより、前時刻の状態ベクトルを遷移させて現時刻の状態ベクトルを推定する。

【0024】

画像認識装置2は、ホワイトノイズによってパーティクルを無数に発生させることにより、前時刻の観測領域とは少しずつ異なる全体観測領域5と部分観測領域6を無数に生成する。

10

これによって、画像認識装置2は、この中で尤度が最大となる全体観測領域5と部分観測領域6の組合せを手当たり次第に探索(探査)し、探索された全体観測領域5と部分観測領域6の組によって対象者3を検出する。これは、マルコフ性を用いた逐次的なベイズ推定問題となっている。

【0025】

次に、画像認識装置2は、探索した全体観測領域5と部分観測領域6を初期状態としてパーティクルを発生させて、次の画像4で同様の処理を行う。これを逐次的に繰り返すことにより、画像認識装置2は、動画の連続する各画像4において対象者3を認識・追跡する。

このように、画像認識装置2の用いるパーティクルフィルタ手法では、状態遷移(観測領域を多数発生させる) 観測(最大尤度の組合せを探索) 複製(探索された組合せを初期値にする)のサイクルを繰り返しながら対象を追跡する。

20

【0026】

基本的には、学習画像と類似する画像を画像4の全体から総当たりで探索するのであるが、対象者3が写っている場所の当たりをつけて、その近辺で探索することにより、画像認識処理を効率化することができる。

【0027】

次に、画像認識装置2が画像認識に用いる画像特徴量について説明する。

画像認識装置2は、HOG(Histogram of Oriented Gradients)特徴量と色分布特徴量の組合せを用いる。

30

これは、HOG特徴量は、明暗の変化に対して頑健であるが、周囲の人物との区別が困難であるのに対して、色分布特徴量は、明暗の変化に対して弱い、周囲の人物との区別が容易であるなど、得意とするシーンが異なるため、これらを組合せることにより相補的な効果が期待できるからである。

このように、画像認識装置2は、画像の輝度勾配分布特徴量と、色分布特徴量から成る複数種類の画像特徴量を用いる。

【0028】

図2は、HOG特徴量の概念を説明するための図である。

HOG特徴量は、輝度勾配分布を用いた画像特徴量であって、対象のエッジを検出する技術である。例えて言えば、対象をエッジによるシルエットで認識するものである。

40

【0029】

HOG特徴量は、次の手順により画像から抽出される。

図2(a)左図に示した画像101を全体観測領域5や部分観測領域6などの注目画像領域とする。

まず、画像101を矩形のセル102a、102b、・・・に分割する。

次に、図2(a)右図に示したように、セル102ごとに各画素(ピクセル)の輝度勾配方向(低輝度から高輝度に向かう方向)を例えば8方向に量子化する。

【0030】

次に、図2(b)に示したように、量子化した輝度勾配の方向を階級とし、出現回数を度数とするヒストグラムを生成することにより、セル102に含まれる輝度勾配のヒスト

50

グラム 106 をセル 102 ごとに作成する。

そして、セル 102 をいくつか集めたブロック単位でヒストグラム 106 の合計度数が 1 となるように正規化する。

【0031】

図 2 (a) 左図の例では、セル 102 a、102 b、102 c、102 d から 1 ブロックが形成されている。

このようにして正規化したヒストグラム 106 a、106 b、・・・を図 2 (c) のように一列に並べたヒストグラム 107 が画像 101 の HOG 特徴量である。

【0032】

HOG 特徴量を用いた画像の類否程度の判断は、次のようにして行う。

10

まず、HOG 特徴量の度数 (M 個あるとする) を成分とするベクトル (x) を考える。ここで、x は、画像 101 を表すベクトルであり、 $x = (\text{第 1 番目の画素の輝度、第 2 番目の画素の輝度、} \dots)$ である。

なお、ベクトルは太字などで表すが、文字化け防止のため、以下では、通常の文字で表す。

【0033】

図 2 (d) は、HOG 特徴量空間を表しており、画像 101 の HOG 特徴量は、M 次元空間のベクトル (x) に写像される。

なお、図では簡単化のため HOG 特徴量空間を 2 次元空間で表してある。

一方、F は、人物画像の学習によって得た重みベクトルであり、多数の人物画像の HOG 特徴量を平均化したベクトルである。

20

【0034】

画像 101 が学習した画像に類似する場合、(x) は、ベクトル 109 のように F の周辺に分布し、類似しない場合は、ベクトル 110、111 のように F とは異なる方向に分布する。

F と (x) は、規格化されており、F と (x) の内積で定義される相関係数は、画像 101 が学習画像に類似するほど 1 に近づき、類似程度が低いほど -1 に近づく。

このように、類似判断の対象となる画像を HOG 特徴量空間に写像することにより、学習画像に類似している画像と類似していない画像を輝度勾配分布により分離することができる。

30

【0035】

輝度勾配分布を用いた特徴量抽出技術には、HOG のほかに、HOG を改良した CoHOG (Co-occurrence HOG) や、これを更に強化した MRCoHOG (Multi Resolution CoHOG) などがあり、何れを用いてもよい。

【0036】

図 3 は、色分布特徴量の概念を説明するための図である。

図 3 (a) に示した注目領域の画像 101 は、色々な色成分 (色 1、色 2、・・・) を有する画素から構成されている。

これら色成分の出現頻度からヒストグラム 121 を作成すると、その度数を成分とするベクトル q が得られる。

40

一方、図 3 (b) に示したように、追跡対象モデルについても同様のヒストグラム 122 を作成し、その度数を成分とするベクトル p を作成する。

【0037】

これら色成分が N 個あるとすると、画像 101 と追跡対象モデルの色分布は、図 3 (c) に示したように、N 次元の色分布特徴量空間ベクトル p、q に写像される。

画像 101 の画像が追跡対象モデルに類似する場合、q は、ベクトル 123 のように p の周辺に分布し、類似しない場合は、ベクトル 124、125 のように p とは異なる方向に分布する。

【0038】

q と p は、規格化されており、q と p の内積で定義される相関係数は、画像 101 が追

50

跡対象モデルに類似するほど 1 に近づき、類似程度が低いほど - 1 に近づく。

このように、類似判断の対象となる画像を色分布特徴量空間に写像することにより、追跡対象モデルに類似している画像と類似していない画像を色分布特徴量により分離することができる。

【0039】

次に、これらの要素技術を用いた尤度計算方法について説明する。

図10は、実施形態で使用する各計算式を表したものである。

画像認識装置2は、HOG特徴量を用いた観測と色分布特徴量を用いた観測を観測領域ごとに行い、HOG特徴量の相関係数と色分布特徴量の相関係数を動的に組合せて尤度を計算する。具体的には、尤度 $L(x)$ の計算式を図10の式(1)で定義する。

10

【0040】

式(1)の右辺第1項目は、 $i = 0$ から n まで $R(x_i)$ の総和を取る。

なお、 n は、部分観測領域6の個数であり、 $i = 0, 1, 2, \dots$ は、全体観測領域5、部分観測領域6a、6b、 \dots に対応する。

また、文字化け防止のため、下付小文字を全角文字で記すことにする。

【0041】

$R(x_i)$ は、図10の式(2)で示したように、 i 番目の観測領域におけるHOG特徴量の相関係数(項2a)と色分布特徴量の相関係数(項2b)の線形和となっている。

i は、当該線形和における重みパラメータであり、0から1までの値を取ることができるが、画像認識装置2では、0.25から0.75までの値とし、必ず両者が $R(x_i)$ に寄与するように設定している。

20

【0042】

ここで、項(2a)、(2b)の (x_i) と、 q_i は、 i 番目の観測領域のHOG特徴量と色分布特徴量であるため、画像認識装置2は、設定した観測領域から複数種類の画像特徴量を取得する画像特徴量取得手段を備えており、当該画像特徴量取得手段は、全体観測領域5、及び部分観測領域6ごとに複数種類の画像特徴量を取得している。

また、項(2a)、(2b)の F_i 、 p_i は、基準特徴量として機能している。そして、項(2a)、(2b)は、それぞれ、 i 番目の観測領域におけるHOG特徴量と色分布特徴量の相関係数(類似度)を表しており、特徴相関係数として機能している。

そのため、画像認識装置2は、取得した画像特徴量ごとに基準特徴量との特徴相関係数を取得する特徴相関係数取得手段を備え、当該特徴相関係数取得手段は、全体観測領域5、及び部分観測領域6ごとに画像特徴量ごとの特徴相関係数を取得している。

30

【0043】

図10の式(2)によると、 i が大きいほど $R(x_i)$ に対するHOG特徴量の寄与が大きくなり、 i が小さくなるほど色分布特徴量の寄与が大きくなる。

そのため、 i を適当に選択することにより、シーンに適した値を得ることができ、頑健性が向上する。

i の値は、基本的には0.25から0.75まで全数探索するが、隣接するフレームでHOG特徴量と色分布特徴量が極端に急変することは考えにくいいため、前時刻での値付近を重点的に探索することにより効率化を図る。

40

より具体的には、後述するように、画像認識装置2は、 i も状態ベクトルに含め、パーティクルフィルタ手法により動的に推定する。

【0044】

ここで、 i は、取得した画像特徴量ごとの特徴相関係数を組合せる割合として機能しており、画像認識装置2は、これを前時刻で値に基づいて設定するため、これを一時刻前の画像4で追跡対象の検出に用いた割合に基づいて設定する割合設定手段を備えている。そして、当該割合設定手段は、全体観測領域5、及び部分観測領域6ごとに当該割合を設定している。

【0045】

また、 $R(x_i)$ は、割合設定手段で設定した割合で特徴相関係数を組合せることによ

50

り i 番目の観測領域に対して設定された領域相関係数として機能しており、画像認識装置 2 は、当該領域相関係数を取得する領域相関係数取得手段を備えている。

そして、当該領域相関係数取得手段は、一例として、設定した割合による線形和によって、特徴相関係数を組合せている。

更に、当該領域相関係数取得手段は、 i 番目の $R(x_i)$ を計算することにより、全体観測領域 5、及び部分観測領域 6 ごとに領域相関係数を取得している。

【0046】

図 10 の式 (1) の第 2 項目は、ペナルティと呼ばれ、項 (1b) を $j = 1$ から n まで、即ち、全ての部分観測領域 6 に対して加算する。このように、ペナルティには、全体観測領域 5 は、含まれていない。

10

ペナルティは、例えば、歩行している対象者 3 の頭部が突然腹部の辺りに現れるなど、全体観測領域 5 に対して部分観測領域 6 が不自然な態様を取った場合に、そのような全体観測領域 5 と部分観測領域 6 の組合せを除去するための量である。ペナルティが大きいほど不自然で排除すべき対象となる。

本実施の形態では、従来技術で利用されているペナルティを用いた。

【0047】

図 10 の式 (1) に示した項 (1b) におけるベクトル d_j は、変形コストと呼ばれ、式 (3) で表される。この量は、部分観測領域 6 の不自然な態様を学習させたものである。

項 (1b) のベクトル d は、部分観測領域 6 の遷移状態を表しており、式 (4) で表される。このベクトル d の成分 (d_x 、 d_y) は、部分観測領域 6 の位置に対応し、(d_x の自乗、 d_y の自乗) は、部分観測領域 6 の移動距離に対応している。即ち、 d は、部分観測領域 6 の位置と移動距離を合わせた量である。

20

この d と d_j との内積を取ることにより不自然度が計算され、部分観測領域 6 の位置や移動距離が不自然なほど、尤度 $L(x)$ から大きなペナルティが減算されることになる。

【0048】

次に、本実施の形態で用いた状態ベクトルについて説明する。

画像認識装置 2 は、全体観測領域 5 と部分観測領域 6 の組を多数発生させて、尤度を最大化するものを探索するが、 s 番目の組の状態ベクトル X_s は、例えば、 $X_s = ((x$ 、 $y)$ 、大きさ、(d_{xj} 、 d_{yj})、 i) によって表される。

30

【0049】

$j = 1, 2, \dots, n$ 、 $i = 0, 1, \dots, n$ とし、 n は部分観測領域 6 の個数である。

(x 、 y) は、全体観測領域 5 の中心座標、大きさは全体観測領域 5 の大きさ、(d_{xj} 、 d_{yj}) は、部分観測領域 6 の配置、 i は、式 (2) で用いられる重みパラメータである。

【0050】

画像認識装置 2 では、前時刻のこれらの値に対して、例えば、正規性白色雑音 (ガウスノイズ; 正規分布乱数) を重畳したパーティクルを次々に発生させることにより前時刻での値近辺で観測領域の複数の組を次々に作成して状態を遷移させる。

40

なお、正規性白色雑音は、正規分布 (ガウス分布) に従って発生する白色雑音 (ホワイトノイズ) である。

【0051】

例えば、 i 番目の領域の時刻 t における X を (i, t) と表すと、 $(i, t) = N((i, (t-1)), \Sigma)$ の自乗) と表される。即ち、遷移後の X の分布は、前時刻の X を平均とし、分散が Σ の自乗である正規分布となる。

このように、画像認識装置 2 は、各状態変数を、前時刻の値を平均とし、状態変数ごとの分散で遷移させる状態遷移モデルを用いている。

【0052】

50

そして、画像認識装置 2 は、重みパラメータ w_i を状態ベクトルに含め、パーティクルフィルタ手法によりこれも動的に推定する。

即ち、パーティクルフィルタ手法の逐次推定のアルゴリズムの中に w_i を含めて w_i も逐次ベイズ推定で有効な値を推定させるのである。

そして、画像認識装置 2 は、尤度 $L(x)$ を最大とする全体観測領域 5 と部分観測領域 6 の組を特定することにより、即ち、尤度 $L(x)$ を最大にする状態ベクトルを特定することにより対象者 3 を検出する。

【0053】

このように、画像認識装置 2 は、観測領域設定手段が設定する観測領域と、割合設定手段が設定する割合を変化させながら、領域相関係数取得手段が取得する領域相関係数を用いて観測領域で観測している対象の尤度を繰り返し取得する尤度取得手段を備えている。

そして、当該尤度取得手段は、全体観測領域 5、及び部分観測領域 6 ごとの領域相関係数を用いて尤度 $L(x)$ を取得している。

更に、画像認識装置 2 は、最大尤度を与える全体観測領域 5 と部分観測領域 6 の組によって対象者 3 を検出するため、当該尤度取得手段で得られた複数の尤度のうち、所定の条件を満たす尤度の取得元となった観測領域で追跡対象を検出する検出手段を備えており、当該所定の条件は、尤度の最大値となっている。

【0054】

画像認識装置 2 は、一例として、パーティクルの数に上限を設け、全てのパーティクルについて尤度 $L(x)$ を求めて最大値を探索することにより対象者 3 を探索するが、他の条件を用いることも可能である。

例えば、尤度 $L(x)$ に所定の閾値を設定し、これを満たした時点で探索を終了したり、あるいは、尤度 $L(x)$ の分布を取得する過程で、極大点が現れてきたら、極大を与える尤度 $L(x)$ を採用して探索を終了する、といったことも可能である。

【0055】

図 4 (a) は、対象者 3 を追跡する実験での w_0 の併用状況を記録したグラフを表している。

縦軸は、全体観測領域 5 の重みパラメータ w_0 を表している。この重みパラメータ w_0 は、色分布特徴量と HOG 特徴量による認識の割合が規定され、両者の認識割合の合計は 1 である。例えば、重みパラメータ w_0 が 0.4 である場合、色分布特徴量による認識の割合が 0.4、HOG 特徴量による認識の割合が 0.6 となる。

このため、重みパラメータ w_0 の値は 0 から 1 の範囲とすることが可能であるが、本実施形態では、0.2 から 0.75 の範囲となっている。

なお、重みパラメータ w_i の取り得る範囲は、追跡対象の種類や大きさ特徴的な形状などによって変更するようにしてもよい。

【0056】

横軸は、追跡対象を撮影した動画のフレーム番号を表しており、時間の推移を示している。

実験では、大型建築物の構内で対象者 3 を歩行させ、その後を追いながら対象者 3 をカメラ 1 で撮影した。

対象者 3 は、周囲から目立つ赤い服を着ており、図 4 (b) に示した照明の明るい地点 31 から、図 4 (c) に示した影の多い薄暗い地点 32 に向けて進んだ。

【0057】

地点 31 では、環境が明るく色がよく観察できるため、色分布特徴量による認識が優位と考えられる。

一方、地点 32 では、環境が暗く、色による識別が困難なため、明暗の変化に対して強い HOG 特徴量が優位になると考えられる。

【0058】

実験結果は、グラフに示したように、地点 31 では、 w_0 が 0.4 程度で色分布特徴量による認識が優位となっているが、対象者 3 が暗がりに向かって進むにつれて w_0 の値が

10

20

30

40

50

徐々に上昇してHOG特徴量による認識の割合が増えていき、地点32では、0が0.7程度になった。

この実験により、環境に応じたHOG特徴量と色分布特徴量のシーンに応じた役割分担がねらい通りに機能していることが確認できた。

部分観測領域6a、6b、・・・におけるパラメータ1、2、・・・も同様である。

【0059】

以上に説明した第1実施の形態により、次のような効果を得ることができる。

(1) HOG特徴量と色分布特徴量を組合せ、状況に応じてこれらの重み付けを変化させることにより、状況変化に対して頑健(ロバスト)な画像認識を行うことができる。

(2) 重みパラメータ w_i を状態ベクトルに含めることにより、パーティクルフィルタ手法を用いて最適な値を動的に推定することができる。

(3) 状況の変化に対して頑健性を有するため、画像認識装置2をロボットや車両などの移動体に搭載し、移動体周辺の追跡対象を移動しながら追跡したり、あるいは、移動体に接近する対象を避けたりすることが可能となる。

【0060】

(第2実施形態)

本実施の形態では、対象者3の部分隠れに対する頑健性を向上させる。

対象者3の追跡中に、他の歩行者や障害物などの遮蔽物が観測領域に侵入し、これによって尤度が低下して追跡が破綻する場合がある。

そこで、第2実施形態では、図10の式(5)に示したように、各観測領域の相関係数 $R(x_i)$ を重みパラメータ w_i で重み付けし、隠れが生じている観測領域の寄与度を低く調節する。

【0061】

図10の式(5)の w_i は、 i 番目の観測領域に対する重みパラメータであり、0から1までの値をとる。

具体的には、 w_0 は、全体観測領域5の重みパラメータである。また、 w_1 、 w_2 、・・・は、部分観測領域6a、6b、・・・の重みパラメータである。

重みパラメータ w_i の値は、基本的に全数探索で決定するが、対象者3の一部が隠れる場合、時系列的に徐々に隠れていく。

これにより、重みパラメータ w_i の値は、前時刻の値とそれほど差がないと考えられるので、その付近を中心に探索する。

そこで、重みパラメータ w_i を状態ベクトルに X_s に含めて、 $X_s = ((x, y)$ 、大きさ、 (dx_j, dy_j) 、 i 、 w_i)とし、パーティクルフィルタ手法により探索する。

【0062】

このように、画像認識装置2は、全体観測領域5、及び部分観測領域6ごとの領域相関係数 $R(x_i)$ に、一時刻前の画像で追跡対象の検出に用いた重みに基づいて重みを設定する重み設定手段を備えており、尤度取得手段は、当該重みも変化させながら、尤度を繰り返し取得する。

【0063】

重みパラメータ w_i の推定方法の基本は、上記の通りであるが、更なる性能の向上を図るため、画像認識装置2は、 $R(x_i)$ の値が所定の閾値より小さい場合は、その観測領域で隠れが生じていると判断して重みパラメータ $w_i = 0$ とする。

これは、 $R(x_i)$ は、-1から1までの値を取るが、隠れが生じた場合、 $R(x_i)$ は、-1に近い値となるため、尤度 $L(x)$ から大きく減算されて検出が不安定になるのを抑制するためである。

【0064】

そこで、画像認識装置2は、 $R(x_i)$ が所定の閾値を下回った場合は重みパラメータ w_i を0とする補正を行うことにした。閾値は、実験によって求めればよいが、ここでは

、一例として、 $R(x_i)$ が負の場合に重みパラメータ $W_i = 0$ とする。

これにより、隠れが生じている観測領域を尤度計算から切り離し、対象者 3 を認識している観測領域によって対象者 3 を認識することができる。

このように、尤度取得手段は、領域相関係数 $R(x_i)$ が所定の閾値を下回る場合は、当該領域相関係数を除外して尤度を取得する。

【0065】

図 5 (a) は、隠れが生じていない場合の重みパラメータ W_i を示した例である。

対象者 3 は、歩行者 200 と共に歩行しているが、両者は、カメラ 1 の視界で重なっていない。

この場合、全体観測領域 5 と部分観測領域 6 には隠れが生じていないため、 $W_0 \sim W_6$ は、何れも 1.0 となる。

【0066】

図 5 (b) は、隠れが生じている場合の重みパラメータ W_i を示した例である。

対象者 3 の一部が歩行者 200 と重なり、隠れが生じている。

部分観測領域 6 b、6 c は、隠れが殆どないため、 W_2 、 $W_3 = 1.0$ となっている。

部分観測領域 6 a、6 d は、隠れが軽微なため、 $W_4 = 0.8$ 、 $W_5 = 0.9$ となっている。

【0067】

全体観測領域 5 と、部分観測領域 6 e は、半分ほど隠れているため、 $W_0 = 0.7$ 、 $W_5 = 0.6$ となっている。

部分観測領域 6 f は、殆ど隠れてしまって $R(x_6)$ が負となるため、 $W_6 = 0.0$ とし、尤度 $L(x)$ の計算から部分観測領域 6 f を分離する。

【0068】

以上に説明した第 2 実施形態により、次のような効果を得ることができる。

(1) 観測領域の組合せにおいて、尤度計算に寄与する最適な値に重み付けを各観測領域に対して設定することができる。

(2) 重みパラメータ W_i を状態ベクトルに含めることにより、パーティクルフィルタ手法を用いて最適な値を動的に推定することができる。

(3) 隠れを含むような尤度の低下に繋がる観測領域の重みパラメータ W_i が小さくなり、隠れによるノイズを低減した尤度計算が可能になる。

(4) 隠れが大きい観測領域に関しては、これを尤度の計算から切り離すことができ、対象者 3 を観測している観測領域によって適切な尤度を得ることができる。

(5) 対象者 3 に部分的な隠れが発生しても追跡を継続することができるため、人の往来や交通の激しい場所でも対象者 3 を追跡することができる。

【0069】

(第 3 実施形態)

本実施の形態では、重みパラメータ W_i に基づいて尤度の信頼性を評価し、これを尤度に反映させる。

例えば、重みパラメータ (W_0 、 W_1 、 W_2 、 W_3 、 W_4 、 W_5 、 W_6) が (1.0、1.0、1.0、1.0、1.0、1.0、1.0) の場合の尤度と、(0.4、1.0、0.3、0.2、0.1、0.0、0.0) の場合の尤度が同じだったとする。

【0070】

この場合、重みパラメータ W_i の値が全体的に大きい前者の方が対象者 3 を全体的に観察しているため、信頼性が高いと考えられる。

即ち、一部の局所領域に偏った観測では、部分的に類似した特徴を持つ物体を誤って追跡する可能性が高くなるため、より多くの観測結果を反映した尤度の方が信頼できると考えられる。

【0071】

そこで、本実施の形態では、重みパラメータ W_i の総和によって信頼性 (信頼度) を評価する信頼性因子を図 10 の式 (6) の項 (6 a) で定義し、これを尤度 $L(x)$ に乗じ

10

20

30

40

50

ることにより、信頼性が低い場合は、尤度が小さく補正されるようにした。

【0072】

項(6a)は、重みパラメータ W_i の和を $n+1$ で除して正規化されており、信頼性が低いほど0に近い値をとり、信頼性が高いほど1に近い値をとる。

乗するのは、所謂ガンマ補正であり、総和が尤度計算に与える影響力の特性を調節するパラメータである。

【0073】

例えば、 α が2の平方根(ルート2)の場合は、0付近での立ち上がりが大きく、1に近づくにつれて緩やかに変化し、 α が2の場合は、0付近での立ち上がり小さく1に近づくにつれて急激に変化する。最適な α は、実験により決定される。

【0074】

図10の式(6)により、信頼性が低いほど(重みパラメータ W_i の総和が小さいほど)尤度 $L(x)$ は、小さく補正される。

これにより、極端に信頼性の低い観測領域の組を探索対象から除外することができる。

このように、画像認識装置2は、重みパラメータ W_i の総和が小さい全体観測領域5と部分観測領域6の組ほど尤度を小さく補正する尤度補正手段を備えている。

【0075】

第3実施の形態によって、次のような効果を得ることができる。

(1) 重みパラメータ W_i の総和に基づいた重みによる信頼性因子によって、尤度の信頼性を尤度計算に反映させることができる。

(2) 尤度分布を得る際に、信頼性の低い観測領域の組を淘汰し、信頼性の高い観測領域の組を生き残らせることができる。

(3) 観測する領域が一部分に極端に偏った観測領域の組を選択的に無視し、全体的にバランスよく観測している組を優遇することにより、できるだけ多くの全体観測領域5と部分観測領域6による観測結果を尤度計算に反映させることができる。

(4) 補正によって感度を調節することができる。

【0076】

(第4実施形態)

本実施の形態では、HOG特徴量と色分布特徴量の何れかに偏った尤度計算に対して負荷を与え、両者をバランスよく用いたものの尤度を大きくする。

HOG特徴量、あるいは色分布特徴量の一方に極端に偏った尤度も信頼性が低いと言えるが、これを是正するために、本実施の形態では、図10の式(6)に式(7)を乗じる。

先に説明した項(6a)は、重みパラメータ W_i の極端な偏りに起因する信頼性の低下を評価する信頼性因子であったが、式(7)は、 i の極端な偏りに起因する信頼性の低下を評価する信頼性因子である。

【0077】

\max は、 i の取りうる最大値であって、ここでは、0.75に設定されている。 u 乗、 v 乗するのは、ガンマ補正であり、 i の偏りが尤度計算に与える影響力の特性を調節するパラメータである。

図10の式(7)は、本願発明者によって経験的に得られたものであり、各観測領域の i が0.5に近いほど1に近くなり、0.5から離れるほど小さくなる。

【0078】

即ち、HOG特徴量と色分布特徴量のバランスが等しいときに1となり、バランスが崩れるほど0に近づく。

そのため、HOG特徴量と色分布特徴量のバランスが悪いものほど尤度が低く補正される。

図10の式(7)は、式(1)や式(5)に乗じることもできる。

このように、画像認識装置2は、割合設定手段が設定する割合 i が一方に偏るほど尤度が小さくなるように補正を行う偏り補正手段を備えている。

10

20

30

40

50

【 0 0 7 9 】

第 4 実施形態により、次のような効果を得ることができる。

(1) H O G 特徴量と色分布特徴量を同じ程度バランスよく用いている傾向のある観測領域の組ほど組合せの試行で生き残り易い。

(2) 何れかの画像特徴量に偏りすぎて誤認識する可能性を低減することができる。

(3) i は、前時刻での i の付近で探索するため、 i の極端な偏りが一度発生すると、それが後続の i に影響するが、後続の i が不適切な i に引きずられるのを抑制することができる。

【 0 0 8 0 】

(第 5 実施形態)

本実施の形態では、追跡が破綻した場合に、状態遷移モデルを切り替えることにより追跡対象を発見して追跡を再開する。

図 6 (a) は、画像認識装置 2 が対象者 3 を認識し、全体観測領域 5 と図示しない部分観測領域 6 によって対象者 3 を追跡しているところを示している。

図 6 (b) は、追跡中に対象者 3 が歩行者 2 0 0 の前に回り込んでカメラ 1 の視界から消えたところを示している。

対象者 3 がカメラ 1 の画像から消えると、各観測領域の相関係数が次々に低下し、追跡が破綻する。

【 0 0 8 1 】

画像認識装置 2 は、各観測領域の相関係数 $R (x_i)$ や重みパラメータ W_i の変化から追跡破綻を検出する。

例えば、画像認識装置 2 は、尤度 $L (x)$ が所定の閾値 T_{h1} を下回った場合、あるいは、重みパラメータ W_i の合計が所定の閾値 T_{h2} を下回った場合に、追跡が破綻したと判断する。

前者の条件では、相関係数 $R (x_i)$ が著しく高い観測領域がある場合には、これによって追跡が可能となる。

後者の条件では、大半の観測領域が隠れた場合はノイズが多いため追跡を中断する。

このように追跡が破綻した判断すると、画像認識装置 2 は、状態遷移モデルを次のように切り替えて探索範囲を広げ、対象者 3 の発見に努める。

まず、観測領域の座標と観測領域の大きさなどについては、正規性白色雑音の分散値を追跡時よりも大きく設定する。これにより、追跡が破綻している間に状態が大きく変化する対象者 3 を観測しやすくなる。

なお、分散値を時間の関数とし、対象者 3 を見失っている時間が長いほど探索範囲を広げるように時間と共に広げてよい。

【 0 0 8 2 】

重みパラメータ i と重みパラメータ W_i については、追跡時は、それぞれ、前時刻の値を中心とするガウス分布にて遷移後の値を推測していたが、破綻後は、これらの値がどのような値になるか不明であるため、これらを乱数 (ランダム) として前時刻の状態を考慮しない遷移を行う。

対象者 3 の全体が隠れた状態が解消するときには、どの観測領域から観測可能になるかわからないため、乱数による予測であらゆる状態を考慮するものである。

【 0 0 8 3 】

また、追跡時は、状態遷移 観測 複製、のサイクルを繰り返していたが、破綻時は、状態遷移 観測、のサイクルを繰り返す。

画像認識装置 2 は、図 6 (c) に示したように、対象者 3 を発見すると、再び全体観測領域 5 と図示しない部分観測領域 6 を設定して追跡を再開する。

このように、割合設定手段と重み設定手段は、追跡対象の検出に失敗した場合に、それぞれ割合と重みを一時刻前の画像とは無関係な任意の値に設定している。

【 0 0 8 4 】

第 5 実施形態により、次のような効果を得ることができる。

10

20

30

40

50

(1) 対象者3の全身、あるいは大部分が隠れることによって追跡が破綻した場合でも、対象者3が現れ始めた際に、速やかに復帰することができる。

(2) 追跡時の状態遷移モデルと、追跡破綻時の状態遷移モデルを用意し、追跡状況に応じて、使用する状態遷移モデルを切り替えることができる。

(3) 追跡破綻時の状態遷移モデルでは、位置などに関しては破綻時の位置を基準に探索範囲を広げ、重みパラメータ w_i 、重みパラメータ W_i に関してはランダムに変化させるなど、対象者3の発見に適した値を設定することができる。

(4) 追跡が破綻した場合でも復帰するためのモードが用意されているため、人や車両の往来が激しい場所でも使用に耐えることができる。

【0085】

図7は、画像認識装置2のハードウェア構成を説明するための図である。この構成は、第1実施形態～第5実施形態に対して共通である。

画像認識装置2は、CPU (Central Processing Unit) 21、ROM (Read Only Memory) 22、RAM (Random Access Memory) 23、撮影装置24、記憶装置25、入力装置26、出力装置27、インターフェース28などバスラインで接続されて構成されている。

【0086】

CPU 21は、記憶装置25に記憶された画像認識プログラムに従って画像認識処理を行うほか、画像認識装置2の各部の制御などを行う。

より詳細には、CPU 21は、撮影装置24から入力される画像4に対して観測領域を設定して対象者3を認識し、追跡する。

【0087】

ROM 22は、CPU 21が画像認識装置2を動作させるための基本的なプログラムやパラメータなどを記憶した読み取り専用のメモリである。

RAM 23は、CPU 21が上記処理を行うためのワーキングメモリを提供する読み書きが可能なメモリである。

撮影装置24から入力された画像4は、RAM 23に展開されてCPU 21により利用される。

【0088】

撮影装置24は、被写体を投影する光学系と、投影された投影像を電気信号に変換する撮像面を用いて構成されており、画像4をフレームとする動画データを生成して出力する。

また、撮影装置24は、マイクロフォンも備えており、これで音声を検出して、音声の含まれた動画データを生成する。

【0089】

記憶装置25は、例えば、ハードディスクや半導体メモリなどの記憶媒体を用いた記憶装置であり、CPU 21に上記の画像認識処理を行わせるための画像認識プログラムなどを記憶している。

また、記憶装置25は、撮影した動画データを記憶するデータ記憶部も備えている。

【0090】

入力装置26は、画像認識装置2に各種の情報を入力する装置であり、ユーザが画像認識装置2を操作するための操作ボタンなどの入力デバイスで構成されている。

出力装置27は、画像認識装置2が各種の情報を出力する装置であり、例えば、操作画面や、撮影中・撮影済みの動画を表示する液晶ディスプレイなどの出力デバイスで構成されている。

インターフェース28は、画像認識装置2を外部機器に接続するインターフェースであって、例えば、インターフェース28を介して画像認識装置2をロボットや車両、あるいは、パーソナルコンピュータに接続し、動画データや追跡データを転送するのに用いられる。

【0091】

10

20

30

40

50

図 8 は、第 5 実施形態の画像認識装置 2 が画像認識処理を行う手順を説明するためのフローチャートである。

以下の処理は、記憶装置 2 5 に記憶した画像認識プログラムに従って CPU 2 1 が行うものである。

まず、画像認識装置 2 は、撮影装置 2 4 から画像 4 を取得して RAM 2 3 に記憶する（ステップ 5）。

【0092】

次に、画像認識装置 2 は、状態ベクトルを構成する各状態変数（全体観測領域 5 の位置、全体観測領域 5 の大きさ、各部分観測領域 6 の位置、各 i 、各重みパラメータ W_i など）の前時刻の値に正規性白色雑音を重畳するなどして状態を遷移させる（ステップ 10）。

10

次に、画像認識装置 2 は、遷移した状態で定まる観測領域の組に対して尤度計算処理を行う（ステップ 15）。

【0093】

次に、画像認識装置 2 は、対象者 3 の探索が終了したか否かを判断する（ステップ 20）。

この判断は、例えば、パーティクルの上限である N 個の場合について全て探索したか否かにより判断する。

【0094】

探索が終了していない場合（ステップ 20； N ）、画像認識装置 2 は、ステップ 10 に戻って、更に状態ベクトルを遷移させて尤度計算を続行する。

20

一方、探索が終了した場合（ステップ 20； Y ）、画像認識装置 2 は、探索した尤度のうち、最大尤度を与える観測領域の組から対象者 3 を検出する（ステップ 25）。

【0095】

次に、画像認識装置 2 は、追跡を続行するか否かを判断する（ステップ 30）。

この判断は、例えば、ユーザが追跡終了ボタンを押すなど、所定の追跡終了操作がなされたか否かにより判断する。

【0096】

追跡を続行する場合（ステップ 30； Y ）、画像認識装置 2 は、ステップ 5 に戻り、動画フレームを構成する次の画像に対して同様の処理を行って対象者 3 を検出する。

30

一方、追跡を終了する場合（ステップ 30； N ）、画像認識装置 2 は、処理を終了する。

以上のようにして、画像認識装置 2 は、連続する動画のフレーム画像で対象者 3 を連続的に検出することにより、対象者 3 を追跡する。

【0097】

図 9 は、ステップ 15 の尤度計算処理を説明するための図である。

画像認識装置 2 は、観測領域を識別するパラメータ i を 0 に初期化する（ステップ 105）。

ここでは、部分観測領域 6 が 6 個あるものとする。そのため、 $i = 0$ は、全体観測領域 5 に対応し、 $i = 1 \sim 6$ は、部分観測領域 6 a ~ 6 f に対応する。

40

【0098】

次に、画像認識装置 2 は、 i 番目の観測領域の画像を取得する（ステップ 110）。例えば、 $i = 0$ の場合は、全体観測領域 5 の画像を取得する。

そして、画像認識装置 2 は、 i 番目の観測領域の画像から HOG 特徴量（ x_i ）を取得し（ステップ 115）、図 10 の式（2）の項（2 a）により HOG 特徴量の相関係数を計算する（ステップ 120）。

【0099】

次に、画像認識装置 2 は、 i 番目の観測領域の画像から色分布特徴量 q_i を取得し（ステップ 125）、式（2）の項（2 b）により色分布特徴量の相関係数を計算する（ステップ 130）。

50

次に、画像認識装置 2 は、両相関係数を重みパラメータ w_i に基づく割合で加算し、式 (2) の $R(x_i)$ を計算する (ステップ 135)。

【0100】

次に、画像認識装置 2 は、 $R(x_i)$ に重みパラメータ w_i を乗じて $R(x_i)$ の重みを補正する (ステップ 140)。

次に、画像認識装置 2 は、 i を 1 だけインクリメントし (ステップ 145)、 i が 6 より大きいか否かを判断する (ステップ 150)。

i が 6 以下の場合は (ステップ 150; N)、まだ未処理の部分観測領域 6 があるため、ステップ 110 に戻って、次の部分観測領域 6 について処理する。

【0101】

一方、 i が 6 より大きい場合は (ステップ 150; Y)、全体観測領域 5 と全ての部分観測領域 6 について相関係数を計算したので、図 10 の式 (1) の項 (1b) によって各部分観測領域 6 についてのペナルティを計算し、その合計を $R(x)$ から減算する (ステップ 155)。

次に、画像認識装置 2 は、図 10 の式 (6) の項 (6a) によって、重みパラメータ w_i の信頼性に基づいて尤度 $L(x)$ を補正する (ステップ 160)。

【0102】

更に、画像認識装置 2 は、図 10 の式 (7) によって i の偏りによる補正項を計算し、これをステップ 160 で計算した式 (6) の値に乗じて尤度 $L(x)$ を補正する (ステップ 165)。

次に、画像認識装置 2 は、各種の補正を施した尤度 $L(x)$ を出力する (ステップ 170)。

【0103】

(第 6 実施形態)

本実施の形態では、画像認識装置 2 を備えたカメラ 1 を移動体に搭載する。

カメラ 1 は、環境変化に対して頑健に対象を追跡できるため、移動体に搭載して様々な環境の中を移動しながら対象を追跡することができる。

例えば、移動体が車両の場合は、カメラ 1 を車両前方に設置し、車両の前景を撮影する。

これにより、車両は、カメラ 1 の映像から前方を走行する他車両を追跡することにより、当該他車両に追従して所謂コンボイ走行を行ったり、前方を移動する車両や歩行者の追跡軌跡から衝突の危険を判断した場合には、ブレーキを作動させたり、走行方向を変えて回避動作を行ったりなどする。このように、画像認識装置 2 は、自動運転技術やその他の技術に応用することができる。

【0104】

また、移動体がロボットの場合は、ロボットの目として利用し、例えば、家庭でサービスを提供したり (例えば、荷物を持って主人の後を追跡する)、動いている対象を追跡しながらロボットアームでこれを把持したり、家畜を追跡しながら誘導したりなど、民生分野、工業分野、農業分野など、産業の各分野で用いることができる。

このように、本実施の形態では、画像認識装置 2 を用いて対象を検出する検出手段と、検出した対象に対して所定の位置に移動する移動手段と、を備えた移動体装置を提供することができる。

【0105】

以上に説明した実施の形態は、一例であって、各種の変形が可能である。

(1) 画像認識装置 2 は、一例として 2 つの画像特徴要素 (HOG 特徴量と色分布特徴量) の組合せを用いて画像認識するが、これに限定するものではなく、他の画像特徴要素を用いてもよく、また、3 個以上の画像特徴要素を組合せてもよい。

例えば、3 種類組合せる場合は、第 1 の特徴量、第 2 の特徴量、第 3 の特徴量を w_1 、 w_2 、 w_3 の割合で加算する。ただし、 $0 < w_1$ 、 $w_2 < 1$ 、及び $w_1 + w_2 + w_3 < 1$ である。全特徴要素を最低でもある割合で必ず含めるために、更に w_1 、 w_2 の値を

10

20

30

40

50

制限してもよい。

HOG特徴要素、色分布特徴要素の他の画像特徴要素を使用できる。

【0106】

(2) 画像認識装置2は、一例として、HOG特徴量の相関係数と色分布特徴量の割合 i による線形和で相関係数 $R(x_i)$ を計算したが、これは、 $R(x_i)$ の計算方法を限定するものではなく、HOG特徴量、色分布特徴量、 i の組合せから任意の計算方法によって算出するものとしてよい。

【0107】

(3) 画像認識装置2は、全体観測領域5と複数の部分観測領域6の組合せによって対象者3を検出するが、本実施の形態の特徴である複数種類の画像特徴データを組合せて対象を認識する技術は、単一の検出窓にて対象を認識するシステムに適用することもできる。

10

【符号の説明】

【0108】

1 カメラ

2 画像認識装置

3 対象者

4 画像

5 全体観測領域

6 部分観測領域

21 CPU

22 ROM

23 RAM

24 撮影装置

25 記憶装置

26 入力装置

27 出力装置

28 インターフェース

31、32 地点

101 画像

102 セル

106、107 ヒストグラム

109、110、111、123、124、125 ベクトル

121、122 ヒストグラム

200 歩行者

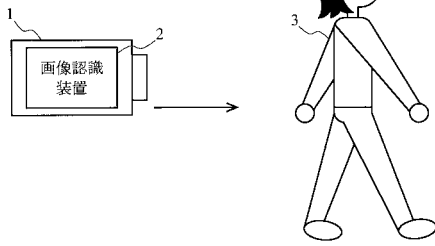
20

30

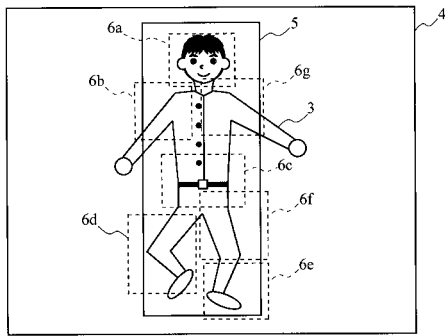
【 図 1 】

【 図 1 】

(a)



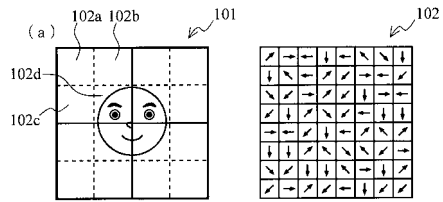
(b)



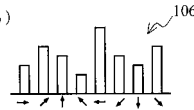
【 図 2 】

【 図 2 】

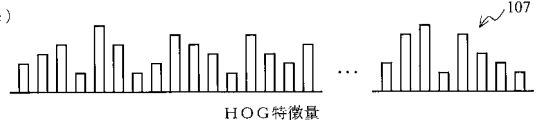
(a)



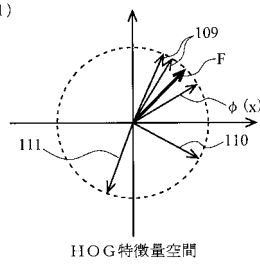
(b)



(c)



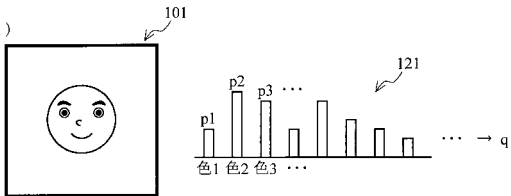
(d)



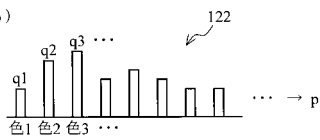
【 図 3 】

【 図 3 】

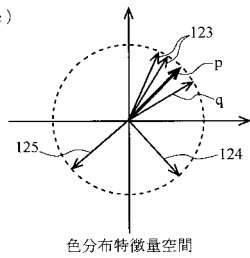
(a)



(b)



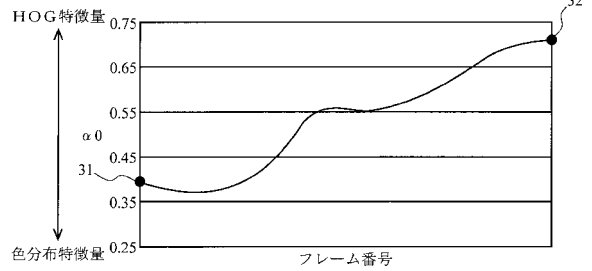
(c)



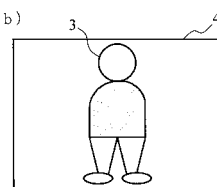
【 図 4 】

【 図 4 】

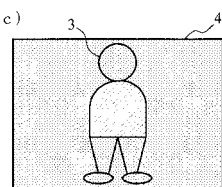
(a)



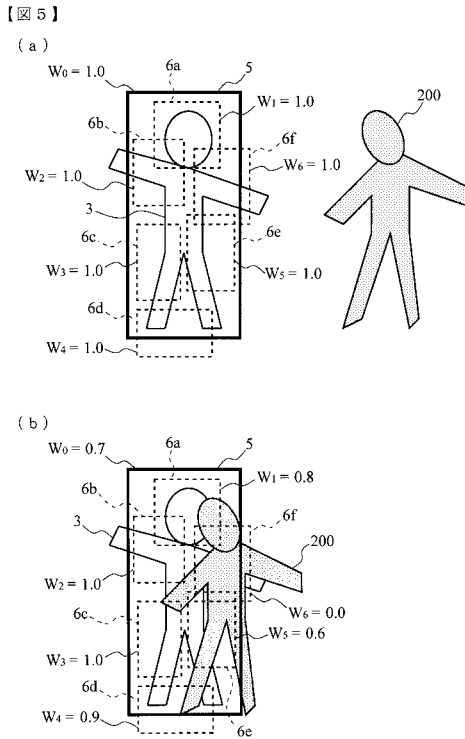
(b)



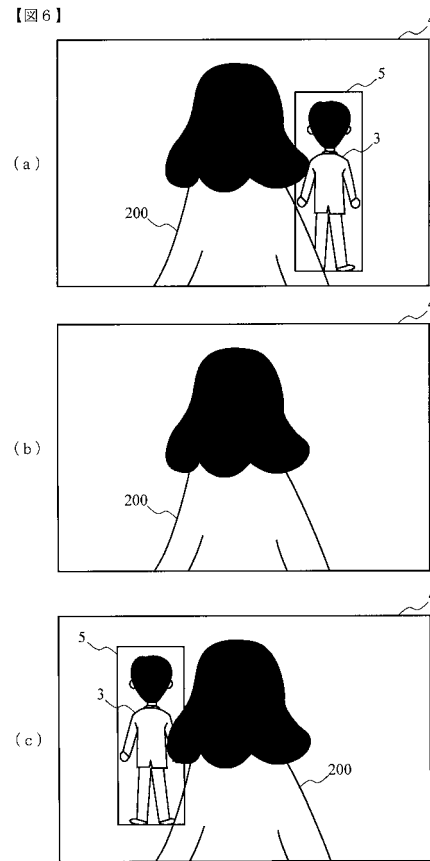
(c)



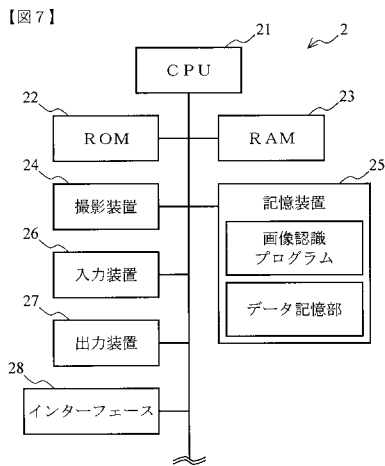
【図5】



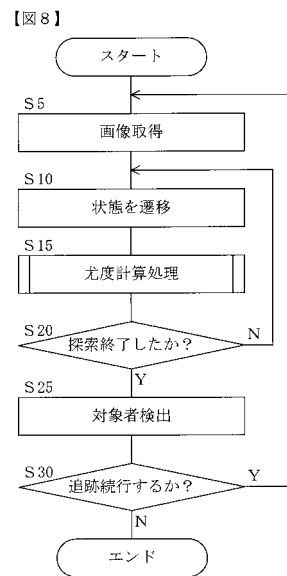
【図6】



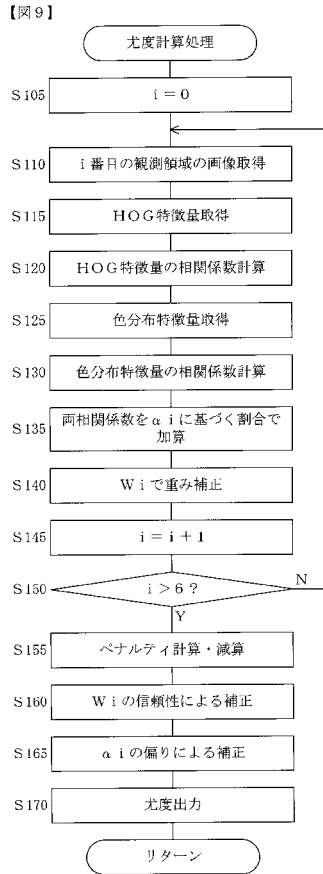
【図7】



【図8】



【 図 9 】



【 図 1 0 】

【 図 1 0 】

$$L(x) = \sum_{i=0}^n R(x_i) - \sum_{j=1}^n \underbrace{d_j \cdot \phi_d(dx_j, dy_j)}_{(1b)} \dots (1)$$

$$R(x_i) = \alpha_i \times \underbrace{F_i \cdot \phi(x_i)}_{(2a)} + (1 - \alpha_i) \times \underbrace{p_i \cdot q_i}_{(2b)} \dots (2)$$

$$d_j = (a_{x_i}, a_{y_i}, b_{x_i}, b_{y_i}) \dots (3)$$

$$\phi_d(dx_i, dy_i) = (dx_i, dy_i, dx_i^2, dy_i^2) \dots (4)$$

$$L(x) = \sum_{i=0}^n R(x_i) \times W_i - \sum_{j=1}^n d_j \cdot \phi_d(dx_j, dy_j) \dots (5)$$

$$L(x) = \left\{ \sum_{i=0}^n R(x_i) \times W_i - \sum_{j=1}^n d_j \cdot \phi_d(dx_j, dy_j) \right\} \times \underbrace{\left\{ \left(\sum_{i=0}^n W_i \right) / (n + 1) \right\}^y}_{(6a)} \dots (6)$$

$$\left\{ \sum_{i=0}^n \frac{\max -(0.5 - \alpha_i)}{\max} \right\}^u \times \left\{ \sum_{i=0}^n \frac{\max -(0.5 - \alpha_i)}{\max} \right\}^v \dots (7)$$

フロントページの続き

(72)発明者 久野 和宏

東京都千代田区外神田1丁目18番13号 秋葉原ダイビル7階 株式会社エクス・リサーチ内

(72)発明者 榎田 修一

福岡県飯塚市大字川津680-4 国立大学法人九州工業大学 情報工学研究院 知能情報工学科
内

(72)発明者 橋本 達哉

福岡県飯塚市大字川津680-4 国立大学法人九州工業大学 情報工学研究院 知能情報工学科
内

Fターム(参考) 5L096 AA06 BA04 BA05 CA04 DA02 EA39 FA06 FA34 FA35 FA77
GA07 GA51 HA05 JA05 JA11 JA16 JA18