

(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2017-182436

(P2017-182436A)

(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 7/62	(2017.01)	G06T 7/60	150J			5B057
G06T 3/00	(2006.01)	G06T 3/00	775			5L096

審査請求 未請求 請求項の数 9 O L (全 16 頁)

(21) 出願番号 特願2016-68435 (P2016-68435)
 (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
 (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
 (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

最終頁に続く

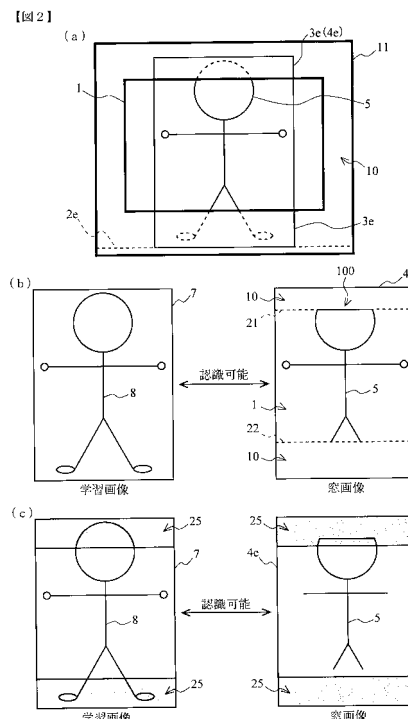
(54) 【発明の名称】 画像認識装置、及び画像認識プログラム

(57) 【要約】

【課題】 認識対象の一部が画像の外側に出ている場合でも認識対象を検出する。

【解決手段】 画像認識装置は、撮影画像1の周囲に所定の画像データで埋めた余白領域10を追加して撮影画像1よりも大きい拡張画像11を作成する。人物がカメラに接近しすぎると、人物像5の一部が撮影画像1からはみ出すが、画像認識装置は、余白領域10も含む大きな検出窓3eを設定することにより、はみ出た領域をも含む窓画像4eを拡張画像11から取り出す。窓画像4eは、人物像5のはみ出た部分を欠いてはいるが、人物像5の全身を写した画像となる。画像認識装置は、人物8が種々の状態(歩行、走行、ジャンプ、座り等の各姿勢)に対して、手や足の位置や向き等が異なる状態)を想定した多数の基準画像7を記憶している。画像認識装置は、窓画像4eから特徴を抽出し、これを基準画像7の特徴と比較することにより人物像5を認識する。

【選択図】 図2



【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、

前記画像取得手段から被写体までの遠近に対応し、少なくとも一部が前記取得した画像の外側の領域に位置する基準線を、前記取得した画像に対して水平方向に設定する基準線設定手段と、

前記基準線の位置に応じた大きさの検出窓を前記基準線に沿って設定する検出窓設定手段と、

前記設定した検出窓に含まれる領域の窓画像を取得する窓画像取得手段と、

前記画像内の認識対象の特徴を表した基準特徴データを取得する基準特徴データ取得手段と、

前記取得した基準特徴データを用いて前記取得した窓画像に前記認識対象が含まれるかを判定する判定手段と、

前記判定による判定結果を出力する出力手段と、

を具備したことを特徴とする画像認識装置。

10

【請求項 2】

前記取得した窓画像から当該窓画像の特徴を表す窓画像特徴データを抽出する窓画像特徴データ抽出手段を具備し、

前記判定手段は、前記取得した基準特徴データと前記抽出した窓画像特徴データを比較することにより前記判定を行うことを特徴とする請求項 1 に記載の画像認識装置。

20

【請求項 3】

前記基準線設定手段は、前記取得した画像の下側の当該画像の外側の領域に前記基準線を設定することを特徴とする請求項 1、又は請求項 2 に記載の画像認識装置。

【請求項 4】

前記基準線設定手段は、前記取得した画像の水平方向の外側に前記基準線を設定することを特徴とする請求項 1、請求項 2、又は請求項 3 に記載の画像認識装置。

【請求項 5】

前記取得した窓画像が前記取得した画像の外側の領域を含む場合に、当該外側の領域を所定の画像データで補完する補完手段を具備したことを特徴とする請求項 1 から請求項 4 までのうちの何れか 1 の請求項に記載の画像認識装置。

30

【請求項 6】

前記補完手段は、前記取得した画像の周囲に前記所定の画像データからなる余白領域を設定することにより、前記外側の領域を前記所定の画像データで補完することを特徴とする請求項 5 に記載の画像認識装置。

【請求項 7】

前記補完手段は、前記取得した窓画像が前記外側の領域を含む場合に、当該外側の領域に前記所定の画像データを追加することにより補完することを特徴とする請求項 5 に記載の画像認識装置。

【請求項 8】

前記判定手段は、比較に際して、前記窓画像に含まれる前記画像の外側の領域の部分の重み付けを他の部分の重み付けよりも小さく設定することを特徴とする請求項 2 に記載の画像認識装置。

40

【請求項 9】

画像を取得する画像取得機能と、

被写体までの遠近に対応し、少なくとも一部が前記取得した画像の外側の領域に位置する基準線を、前記取得した画像に対して水平方向に設定する基準線設定機能と、

前記基準線の位置に応じた大きさの検出窓を前記基準線に沿って設定する検出窓設定機能と、

前記設定した検出窓に含まれる領域の窓画像を取得する窓画像取得機能と、

前記画像内の認識対象の特徴を表した基準特徴データを取得する基準特徴データ取得機

50

能と、

前記取得した基準特徴データを用いて前記取得した窓画像に前記認識対象が含まれるか否かを判定する判定機能と、

前記判定による判定結果を出力する出力機能と、
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、及び画像認識プログラムに関し、例えば、画像に写っている対象物を認識するものに関する。

10

【背景技術】

【0002】

近年、ロボットや車両にカメラを搭載し、撮影された画像から対象物を認識して制御に用いたり、あるいは、カメラを街路に設置して人を画像認識し、これによって人の往来を観察したりなど、画像認識技術が産業や社会の各方面で利用されるようになってきた。

これら画像認識技術には、対象物や非対象物（対象物ではないと認識するのに必要）の特徴を表す基準画像を画像認識装置に記憶させておき、基準画像とカメラ画像の特徴を比較することにより行うものがある。

【0003】

このような画像認識技術として、特許文献1の「物体検出装置」がある。

20

この技術は、HOG (Histogram of Oriented Gradients) 特徴量を用いて画像に写った物体を検出するものである。

【0004】

図7は、従来例の画像認識方法を説明するための図である。

図7(a)に示したように、カメラで撮影した画像1に認識対象である人物像5が写っているとする。

画像認識装置は、この人物像5を次のようにして検出する。

まず、画像認識装置は、人物の画像1での立ち位置に応じた複数の基準線（ベースライン）を設定する。

【0005】

30

図の例では、カメラから1m、5m、10m、15mの位置に基準線2a、2b、2c、2dが設定されている。

人物は、カメラから遠いほど画像1の上側に写るため、基準線2（基準線2a～2dを特に区別しない場合は、単に基準線2と記し、他の要素についても同様とする）は、画像1の下からこの順に設定される。

【0006】

次に、画像認識装置は、基準線2を下端とする検出窓3a～3dを基準線2ごとに設定する。

検出窓3の大きさは、基準線2で規定される距離に人物が立っていた場合に、全身が収まる程度の大きさに設定されている。

40

これは、人物の全身が写った基準画像を使用しているため、基準画像との整合性を図るためである。

【0007】

次に、画像認識装置は、検出窓3を基準線2に沿って移動（シフト）させながら検出窓3で囲まれた窓画像4の特徴を抽出し、これを予め記憶している基準画像の特徴と比較する。そして両者の類似性（相関）により人物を認識する。

この例では、図7(b)に示したように人物像8が写った基準画像7の特徴と、検出窓3cで囲まれた領域の窓画像4cの特徴を比較して人物像5が認識される。

【0008】

ところで、人物がカメラに接近しすぎると、図7(c)に示したように、人物像5が画

50

像 1 からはみ出てしまい、最も大きい検出窓 3 a を用いたとしても、人物像 5 の頭部や下肢などを欠いた窓画像 4 a が得られる。

【 0 0 0 9 】

これにより、画像認識装置は、図 7 (d) に示したように、人物像 8 の全体が写った基準画像 7 の特徴と、人物像 5 の一部が画面全体に写った窓画像 4 a の特徴を比較することになる。

この場合、画像における人物の形、大きさ、構図などが全く異なるため、基準画像 7 の特徴から窓画像 4 a で人物像 5 を認識することは困難である。

【 0 0 1 0 】

更に、画像の特徴として H O G などの輝度勾配を用いる手法を用いる場合、頭部が寸断されたエッジ部分 1 0 0 では、画像の端部であるため、輝度勾配の計算、即ち特徴の抽出が困難であり、計算のための何らかの工夫を施す必要がある。

【 0 0 1 1 】

このような問題に対して、人物の一部が画像の外側に出ている状態の基準画像を作り画像認識装置に記憶しておく方法も考えられるが、これを実行するには様々な態様で人物の一部が画像の外側に出ている基準画像を記憶する必要があり、記憶すべきデータ量が膨大となって実現は困難である。

具体的には、1 の対象物に対する基準画像（全体が写った画像）が複数（数千枚程度）存在し、その各々に対して上下が欠けた各種状態の画像、右側が欠けた各種状態の画像、左側が欠けた各種状態の画像、上下と左右の一方が欠けた各種状態の画像を準備、記憶する必要があり膨大な量になってしまう。

【 0 0 1 2 】

このように、従来技術は、認識対象の一部が画像の外に出ている場合にこれを認識できないという問題があった。

特に、画像認識装置をロボットや車両などの移動体に搭載する場合、遠くの対象物を認識する場合よりも、例えば、対象物の回避動作を素早く行うためにも、接近した対象物（人や物）を適確に認識することが重要になる場面が多い。

このように、近距離に対する認識要請が高い場合にも係わらず、認識対象の全体がカメラの画角に入りきらないことが原因で認識できなかった。

【 先行技術文献 】

【 特許文献 】

【 0 0 1 3 】

【 特許文献 1 】 特開 2 0 1 6 - 1 2 3 1 1 号公報

【 発明の概要 】

【 発明が解決しようとする課題 】

【 0 0 1 4 】

本発明は、認識対象の一部が画像の外側に出ている場合でも認識対象を検出することを目的とする。

【 課題を解決するための手段 】

【 0 0 1 5 】

(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、前記画像取得手段から被写体までの遠近に対応し、少なくとも一部が前記取得した画像の外側の領域に位置する基準線を、前記取得した画像に対して水平方向に設定する基準線設定手段と、前記基準線の位置に応じた大きさの検出窓を前記基準線に沿って設定する検出窓設定手段と、前記設定した検出窓に含まれる領域の窓画像を取得する窓画像取得手段と、前記画像内の認識対象の特徴を表した基準特徴データを取得する基準特徴データ取得手段と、前記取得した基準特徴データを用いて前記取得した窓画像に前記認識対象が含まれるか否かを判定する判定手段と、前記判定による判定結果を出力する出力手段と、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、前記取得した窓画像から当該窓画像の特徴を表す窓画

10

20

30

40

50

像特徴データを抽出する窓画像特徴データ抽出手段を具備し、前記判定手段は、前記取得した基準特徴データと前記抽出した窓画像特徴データを比較することにより前記判定を行う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、前記基準線設定手段が、前記取得した画像の下側の当該画像の外側の領域に前記基準線を設定することを特徴とする請求項 1、又は請求項 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、前記基準線設定手段が、前記取得した画像の水平方向の外側に前記基準線を設定することを特徴とする請求項 1、請求項 2、又は請求項 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、前記取得した窓画像が前記取得した画像の外側の領域を含む場合に、当該外側の領域を所定の画像データで補完する補完手段を具備したことを特徴とする請求項 1 から請求項 4 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。

(6) 請求項 6 に記載の発明では、前記補完手段が、前記取得した画像の周囲に前記所定の画像データからなる余白領域を設定することにより、前記外側の領域を前記所定の画像データで補完することを特徴とする請求項 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、前記補完手段が、前記取得した窓画像が前記外側の領域を含む場合に、当該外側の領域に前記所定の画像データを追加することにより補完することを特徴とする請求項 5 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、前記判定手段が、比較に際して、前記窓画像に含まれる前記画像の外側の領域の部分の重み付けを他の部分の重み付けよりも小さく設定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。

(9) 請求項 9 に記載の発明では、画像を取得する画像取得機能と、被写体までの遠近に対応し、少なくとも一部が前記取得した画像の外側の領域に位置する基準線を、前記取得した画像に対して水平方向に設定する基準線設定機能と、前記基準線の位置に応じた大きさの検出窓を前記基準線に沿って設定する検出窓設定機能と、前記設定した検出窓に含まれる領域の窓画像を取得する窓画像取得機能と、前記画像内の認識対象の特徴を表した基準特徴データを取得する基準特徴データ取得機能と、前記取得した基準特徴データを用いて前記取得した窓画像に前記認識対象が含まれるか否かを判定する判定機能と、前記判定による判定結果を出力する出力機能と、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】

【0016】

本発明によれば、検出窓領域を画像の外側に拡張することにより、認識対象の一部が画像の外側に出ている場合でも認識対象を検出することができる。

【図面の簡単な説明】

【0017】

【図 1】画像認識装置が行う検出窓の設定方法を説明するための図である。

【図 2】人物が画像からはみ出た場合の処理を説明するための図である。

【図 3】人物が画像からはみ出た場合の処理を説明するための図である。

【図 4】人物が画像からはみ出た場合の処理を説明するための図である。

【図 5】画像認識装置のハードウェア構成の一例を示した図である。

【図 6】画像認識装置の動作を説明するためのフローチャートである。

【図 7】従来例の画像認識方法を説明するための図である。

【発明を実施するための形態】

【0018】

(1) 実施形態の概要

図 2(a) に示したように、画像認識装置は、カメラで撮影した撮影画像 1 の周囲に所定の画像データで埋めた余白領域 10 を追加して撮影画像 1 よりも大きい拡張画像 11 を作成する。

10

20

30

40

50

人物がカメラに接近しすぎると、人物像 5 の一部が撮影画像 1 からはみ出すが、画像認識装置は、余白領域 10 も含む大きな検出窓 3 e を設定することにより、はみ出た領域をも含む窓画像 4 e を拡張画像 11 から取り出す。

図 2 (b) に示したように、窓画像 4 e は、人物像 5 のはみ出た部分 (頭部と下肢の一部) を欠いてはいるが、人物像 5 の全身を写した画像となる。

【 0 0 1 9 】

画像認識装置は、人物 8 が種々の状態 (歩行、走行、ジャンプ、座り等の各姿勢に対して、手や足の位置や向き等が異なる状態等) を想定した多数の基準画像 7 を記憶している。

なお、本実施形態では、基準画像 7 そのものの画像データではなく、各基準画像 7 から予め抽出した特徴量を記憶することで記憶すべきデータ量を減らすと共に、基準画像 7 の特徴量を抽出する処理負荷を減らして認識速度を向上させているが、基準画像 7 を記憶するようにしてもよい。

【 0 0 2 0 】

そして、画像認識装置は、窓画像 4 e から特徴を抽出し、これを基準画像 7 の特徴量 (以下たんに基準画像 7 という。) と比較することにより人物像 5 を認識する。

基準画像 7 と窓画像 4 e は、余白領域 10 に相当する部分は似ていないものの、撮影画像 1 に相当する部分は類似するため、基準画像 7 の特徴によって人物像 5 を認識することが可能となる。

【 0 0 2 1 】

(2) 実施形態の詳細

図 1 は、本実施の形態に係る画像認識装置が行う検出窓の設定方法を説明するための図である。

画像認識装置は、カメラから撮影画像 1 を得ると、撮影画像 1 の外周部の全周に渡って白色の画像データで埋められた余白領域 10 を追加する。

これにより、画像認識装置は、撮影画像 1 の上下端部を y 方向 (縦方向) に拡張し、左右端部を x 方向 (横方向) に拡張した拡張画像 11 を作成する。

【 0 0 2 2 】

ここで拡張画像 11 による上下方向拡張量 y と横方向拡張量 x について説明する。

拡張量については、カメラの撮影状態や撮像画像 1 の画角やサイズ等により任意に決定される。

例えば、撮像画像 1 のサイズが横 1280 ピクセル、縦 720 ピクセルである場合、上下左右の各々に対して 200 ピクセル分の余白領域 10 を追加する。

また、一例として、撮影画像 1 の最下部に基準線 2 を設定した場合の検出窓 3 のサイズを基準サイズとし、上下方向の各々に基準サイズの n % を拡張し、左右方向の各々に基準サイズの m % とする。

n と m の値は、認識対象に対応した基準画像 7 に応じて決定され、例えば、25 %、30 %、50 % 等が使用され、n と m は同じ値であっても異なる値であってもよい。また、基準画像の縦横比に基づいて n と m の比率を決定するようにしてもよい。

【 0 0 2 3 】

このように、画像認識装置は、画像を取得する画像取得手段と、当該取得した画像の周囲に所定の画像データからなる余白領域 10 を設定することにより、当該外側の領域を当該所定の画像データで補完する補完手段を備えている。

【 0 0 2 4 】

ここで、一例として余白領域 10 を白色の画像データで埋めたのは、後ほど画像の輝度勾配の分布による、基準画像の特徴データとの比較を行うため、余白領域 10 においてノイズとなる輝度勾配が発生するのを抑制するためである。

そのため、認識精度に影響しないなら他の色や模様の画像データによって余白領域 10 を補完してもよい。

【 0 0 2 5 】

10

20

30

40

50

画像認識装置は、このようにして拡張画像 1 1 を形成した後、1 m、5 m、1 0 m、1 5 m に対応する基準線 2 a ~ 2 d を x 方向に延長して、撮影画像 1 の外側の余白領域 1 0 まで設定する。

更に、画像認識装置は、カメラから 5 0 c m の距離に対応する基準線 2 e を撮影画像 1 の下方外側の余白領域 1 0 に追加する。

なお、これら基準線 2 を設定する距離や本数については、一例であって、画像認識装置の用途や認識対象によって各種の設定が可能である。

例えば、本実施形態の場合、撮影画像 1 内に 8 本を設定し、更に拡張した余白領域 1 0 に 2 本を設定するようにしている。但し、図面では複雑化を避けるために減らして表示している。

10

【0026】

このように、画像認識装置は、被写体の遠近に対応し、少なくとも一部が撮影画像 1 の外側の領域に位置する基準線 2 a ~ 2 e を、撮影画像 1 に対して水平方向に設定する基準線設定手段を備えている。

そして、当該基準線設定手段は、撮影画像 1 の下側の当該撮影画像 1 の外側の領域に基準線 2 e を設定し、撮影画像 1 の水平方向の外側に基準線 2 a ~ 2 d を設定している。

【0027】

画像認識装置は、基準線 2 a ~ 2 e を設定すると、これらに対応する矩形形状の検出窓 3 a ~ 3 e を設定する。

検出窓 3 a ~ 3 d は、従来例のものと同様であり、人物の全身が検出窓 3 に程よく収まるように基準線 2 の位置に応じて（カメラからの距離が遠くなるにつれて）窓枠が小さくなるように形成されている。

20

そして、検出窓 3 a ~ 3 d は、延長された基準線 2 に沿って、撮影画像 1 の水平方向外側に位置する余白領域 1 0 まで移動できるようになっている。

【0028】

検出窓 3 e は、基準線 2 e の新設に伴って新たに追加されたものである。

検出窓 3 e の窓枠は、人物がカメラから 5 0 c m 程度の距離に立っている場合に、撮影画像 1 からはみ出る部分も含めて当該人物の全身が程よく収まる程度の大きさに形成されている。

【0029】

30

この例では、検出窓 3 e は、撮影画像 1 の上下端の外側の余白領域 1 0 に渡って設定されている。

検出窓 3 e は、基準線 2 e に沿って水平方向に移動し、余白領域 1 0 を含めた拡張画像 1 1 を走査することができる。

【0030】

このように、拡張画像 1 1 の余白領域 1 0 もカバーするように検出窓 3 を設定したため、人物が左右に移動して撮影画像 1 の左右端からはみ出した場合でも検出窓 3 a ~ 3 d で捕捉することができ、また、人物がカメラに近づきすぎて撮影画像 1 の上下端からはみ出した場合でも検出窓 3 e で捕捉することができる。

以上のように、画像認識装置は、基準線 2 a ~ 2 e に応じた大きさの検出窓 3 a ~ 3 e を基準線 2 に沿って設定する検出窓設定手段を備えている。

40

【0031】

図 2 の各図は、人物がカメラに接近して撮影画像 1 からはみ出した場合の処理を説明するための図である。

図 2 (a) の波線で示したように、人物がカメラから 5 0 c m ~ 1 m 程度に近づくと、当該人物の人物像 5 は、頭部が撮影画像 1 の上端からはみ出し、下肢が撮影画像 1 の下端からはみ出して撮影されない。

【0032】

しかし、検出窓 3 e は、撮影画像 1 の上端側と下端側で余白領域 1 0 まで拡張されているため、画像認識装置は、図 2 (b) に示したように、頭部と下肢は欠けているものの人

50

物像 5 の全身を写した窓画像 4 e を得ることができる。

ここで、窓画像 4 e は、検出窓 3 e によって拡張画像 1 1 から取得された画像であって、境界線 2 1、2 2 は、撮影画像 1 と余白領域 1 0 の境界を表している。

【 0 0 3 3 】

このように、画像認識装置は、検出窓 3 e に含まれる領域の窓画像 4 e を取得する窓画像取得手段を備えている。

そして、補完手段によって、余白領域 1 0 が所定の画像データで補完されているため、窓画像 4 e が撮影画像 1 の外側の領域を含む場合に、当該外側の領域が補完手段によって所定の画像データで補完されている。

【 0 0 3 4 】

また、この例では、予め所定の画像データで余白領域 1 0 を補完しておいて検出窓 3 e で拡張画像 1 1 を取り出したが、検出窓 3 e で余白領域 1 0 の部分の欠損した撮影画像 1 のサイズの窓画像 4 e を取り出して、その後、余白領域 1 0 を所定の画像データで補完することも可能である。

この場合、補完手段は、窓画像 4 e が撮影画像 1 の外側の領域を含む場合に、当該外側の領域に所定の画像データを追加することにより補完している。

あるいは、予め余白の設定値に基づく（例えば、白色の）拡張画像 1 1 を用意しておき、その拡張画像 1 1 の中央部に画像 1 を書き込むことで画像 1 の周囲に余白が設定された拡張画像 1 1 を作成するようにしてもよい。

【 0 0 3 5 】

窓画像 4 e を基準画像 7 と比べてみると、人物像 5 の頭部と下肢が欠損しているものの、両者の構図は近く、頭部と下肢以外の部分は、特徴がよく似ている。

そのため、基準画像 7 から抽出した特徴（特徴量）と窓画像 4 e から抽出した特徴（特徴量）の類似度（相関）が大きくなり、窓画像 4 e の人物像 5 を認識することが可能となる。実験によってもそのような結果が得られている。

【 0 0 3 6 】

なお、画像認識装置は、予め各基準画像 7 から抽出した特徴を記憶しており、これを窓画像 4 e から抽出した特徴と比較して比較結果を出力するようになっている。

このように、画像認識装置は、認識対象の特徴を表した基準特徴データ（基準画像 7 から抽出した特徴データ）を取得する基準特徴データ取得手段と、当該取得した基準特徴データを用いて窓画像 4 e に認識対象（人物像 5）が含まれるか否かを判定する判定手段と、当該判定による判定結果を出力する出力手段を備えている。

【 0 0 3 7 】

更に、画像認識装置は、窓画像 4 e から当該窓画像 4 e の特徴を表す窓画像特徴データ（窓画像 4 e の特徴）を抽出する窓画像特徴データ抽出手段を備えており、判定手段は、当該取得した基準特徴データと当該抽出した窓画像特徴データを比較することにより判定を行っている。

【 0 0 3 8 】

窓画像 4 e から人物像 5 を認識する技術は、既存の各種のものが使用可能であるが、本実施の形態では、一例として、M R C o H O G (M u l t i R e s o l u t i o n C o H O G) 技術を用いた。

M R C o H O G 技術は、輝度勾配を用いて画像認識する系統の技術であり、H O G 技術を改良した C o H O G (C o - o c c u r r e n c e H O G) 技術を更に改良して、処理演算量を低減させつつ、頑健性（ロバスト性）を維持したものである。

【 0 0 3 9 】

H O G 技術では、画像をセルと呼ばれる小領域に区分する。そして、セル内の各画素の輝度勾配を、例えば、8 方向に量子化し、どの方向の輝度勾配がどの頻度で出現しているか、のヒストグラムを作成する。

そして、複数のセルを含むブロック単位でヒストグラムを正規化した後、全セルのヒストグラムを一行に並べて出力したものが当該画像の特徴となる。

10

20

30

40

50

これを基準画像 7 の特徴と比較することにより、画像中の認識対象を認識することができる。認識対象の認識については、例えば、所定閾値以上の尤度であるか否かによる。

【 0 0 4 0 】

C o H O G 技術では、H O G 技術と同様に画面をセル（ブロックとも呼ばれる）に分割し、セル内の各画素の輝度勾配を量子化する。

そして、セル内のある画素に注目し、この画素の輝度勾配方向と予め指定された周辺の画素の輝度勾配方向との組合せ（共起）により、ヒストグラムへの投票を行う。これをセル内の全画素について行う。

【 0 0 4 1 】

例えば、注目画素の輝度勾配方向が右方向で、隣の画素の輝度勾配が上方向であったら、右行上列の行列要素に一票を投じる。このようにしてセルごとに共起行列が作成される。

10

そして、全セルの共起行列要素を一列に並べて出力したものが当該画像の特徴となる。基準画像 7 の特徴量と比較することにより、画像中の対象を認識することができる。

【 0 0 4 2 】

M R C o H O G 技術では、同じ対象を写した異なる解像度の画像間で共起をとる。

例えば、高解像度画像から中解像度画像と低解像度画像を作成することにより、高中低の 3 種類の解像度の画像を用意する。

そして、高解像度画像のある画素に注目し、当該画素の輝度勾配の方向と、これに隣接する中解像度、及び低解像度の画素の輝度勾配の方向の組合せにより共起行列に投票する。

20

これが高解像度画像の全ての画素について行う。場合によっては、中解像度画像や低解像度画像の画素を注目画素として同様の投票を行う。

そして、全共起行列要素を一列に並べて出力したものが当該画像の特徴となる。

これを基準画像 7 の特徴と比較することにより、画像中の対象を認識することができる。

【 0 0 4 3 】

なお、C o H O G 技術や H O G 技術を使用する場合、複数の基準画像 7 から抽出した各特徴データから 1 の統一基準特徴データを作成（学習）し、この統一基準特徴データを認識対象毎に記憶するようにしてもよい。これにより、窓画像 4 との比較量を大幅に削減することができる。

30

【 0 0 4 4 】

本実施の形態では、このように画像認識に輝度勾配を用いるため、頭部のエッジ部分 1 0 0 に着目すると、余白領域 1 0 がなかった場合、隣接する画素がないため輝度勾配が計算できない。

しかし、窓画像 4 e では、余白領域 1 0 の追加によりエッジ部分 1 0 0 の隣接画素が形成されているため、この部分での輝度勾配を計算することが可能であり、これら輝度勾配を用いた特徴抽出技術を利用することができる。

【 0 0 4 5 】

図 2 (c) は、基準画像 7 と窓画像 4 e の相関を更に高めるための試みを説明するための図である。

40

基準画像 7 と窓画像 4 e を比較すると、余白領域 1 0 の部分が明らかに似ていない。そのため、この例では、基準画像 7 と窓画像 4 e の類似を判定する際に、余白領域 1 0 に相当する領域 2 5 の重み付けを小さくする。

【 0 0 4 6 】

このように基準画像 7 と窓画像 4 e の撮影画像 1 に相当する部分に比べて余白領域 1 0 に相当する部分の重み付けを小さくして両者の特徴の類似を判断することにより、より認識精度を高めることができる。

この例では、判定手段は、比較に際して、窓画像 4 e に含まれる撮影画像 1 の外側の領域の部分の重み付けを他の部分の重み付けよりも小さく設定している。

50

【 0 0 4 7 】

なお、領域 2 5 は、余白領域 1 0 と厳密に一致している必要はなく、実験や状況に応じてどの程度の範囲に領域 2 5 を設定したらよいか決めればよい。

例えば、輝度勾配による特徴量を抽出する場合、撮影画像 1 と余白領域 1 0 の境界（頭部や下肢の切断箇所）で、本来の人物像 5 とは異なった輝度勾配が生じるため、これがノイズになる可能性がある。

その場合は、これら切断箇所を覆う位置まで領域 2 5 を拡張し、当該箇所の重み付けを小さくすればよい。

【 0 0 4 8 】

図 3 の各図は、人物がカメラに接近して撮影画像 1 の上下端からはみ出ると共に左端からもはみ出した場合の処理を説明するための図である。

図 3 (a) の波線で示したように、人物がカメラから 5 0 c m ~ 1 m 程度に近づき、画角の左端に寄ると、当該人物の人物像 5 は、頭部が撮影画像 1 の上端からはみ出し、下肢が撮影画像 1 の下端からはみ出し、左側の部分が撮影されない。

【 0 0 4 9 】

しかし、検出窓 3 e は、撮影画像 1 の上下端側、及び左端側で余白領域 1 0 まで拡張されているため、画像認識装置は、図 3 (b) に示したように、頭部、下肢、及び左部分は欠けているものの人物像 5 の全身を写した窓画像 4 e を得ることができる。

【 0 0 5 0 】

窓画像 4 e を基準画像 7 と比べてみると、人物像 5 の頭部、下肢、及び左部分が境界線 2 1 ~ 2 3 の外側で欠損しているものの、両者の構図は近く、欠損部分以外は、特徴がよく似ている。

そのため、基準画像 7 から抽出した特徴量と窓画像 4 e から抽出した特徴量が類似し、窓画像 4 e の人物像 5 を認識することが可能となる。

【 0 0 5 1 】

図 3 (c) は、基準画像 7 と窓画像 4 e の相関を更に高めるための試みを説明するための図である。

上の例と同様に、余白領域 1 0 に対応する部分（上下端部と左端部）に領域 2 5 を設定し、この部分における特徴量比較の重み付けを他の部分よりも小さく設定する。

これにより、認識精度の向上が期待できる。

【 0 0 5 2 】

図 4 の各図は、カメラから人物までの距離は、人物の全身が撮影画像 1 に写る距離であるものの、人物が撮影画像 1 の左端に寄ってしまったため、人物像 5 の左部分が撮影画像 1 の外側にはみ出した場合の処理を説明するための図である。

【 0 0 5 3 】

図 4 (a) で示したように、人物の左部分は、撮影画像 1 の外側に位置するためこの部分の画像は得られないが、検出窓 3 b は、延長した基準線 2 b 上を左側に移動可能なため、図 4 (b) に示したように、左部分が欠損するものの人物の全身が写った窓画像 4 b が得られる。

このように画像認識装置は、接近により撮影画像 1 からはみ出る人物のほか、カメラからの距離は適切であるが、カメラの画角からはみ出る人物に対しても画像認識を行うことができる。

【 0 0 5 4 】

窓画像 4 b を基準画像 7 と比べてみると、人物像 5 の左部分が欠損しており余白領域 1 0 で置き換えられているものの、両者の構図は近く、欠損部分以外は、特徴がよく似ている。

そのため、基準画像 7 から抽出した特徴量と窓画像 4 b から抽出した特徴量を比較することにより窓画像 4 の人物像 5 を認識することができる。

また、余白領域 1 0 に領域 2 5 を設定してこの部分の重み付けを小さくすることにより認識精度を高めることもできる。

10

20

30

40

50

【0055】

図5は、本実施の形態に係る画像認識装置30のハードウェア構成の一例を示した図である。

画像認識装置30は、CPU(Central Processing Unit)31、ROM(Read Only Memory)32、RAM(Random Access Memory)33、カメラ34、記憶装置35、入力装置36、出力装置37などがバスラインで接続されたパーソナルコンピュータを用いて構成されている。

【0056】

CPU31は、記憶装置35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30の各部の制御などを行う。

10

より詳細には、CPU31は、カメラ34から入力される画像フレームに対して余白領域10の設定による拡張画像11の形成、拡張画像11に対する基準線2a~2e、及び検出窓3a~3eの設定、窓画像4からの特徴量の抽出と、当該抽出した特徴量と学習データ(基準画像から抽出した特徴量)の相関による人物像5の認識などを行う。

【0057】

ROM32は、CPU31が画像認識装置30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。

RAM33は、CPU31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。

カメラ34による画像フレームや基準画像7の特徴量はRAM33に展開されてCPU31により利用される。

20

【0058】

カメラ34は、所定のインターフェースを介して画像認識装置30に接続されており、投影像を電気信号に変換する撮像面を有する撮像素子と、当該撮像面上に被写体を投影する光学系を用いて構成されている。

カメラ34は、所定のフレームレートで被写体(レンズの前方の風景であり、人物と背景から成る)の動画を出力し、各フレームを構成する画像が画像認識処理の対象となる。

【0059】

記憶装置35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、CPU31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を記憶している。

30

また、記憶装置35は、人物像5を認識するために大量の基準画像7に対する特徴量も記憶しており、これらを画像認識処理のためにCPU31に提供する。

【0060】

入力装置36は、画像認識装置30に各種の情報を入力する装置であり、例えば、キーボード、マウスなどの入力デバイスで構成されている。

ユーザは、キーボード操作やマウス操作によってコマンドを入力するなどして画像認識装置30を操作することができる。

【0061】

出力装置37は、画像認識装置30が各種の情報を出力する装置であり、例えば、ディスプレイ、プリンタなどの出力デバイスで構成されている。

40

ユーザは、ディスプレイに表示された操作画面に従いながら画像認識装置30を操作することができる。

以上の構成は、一例であって、画像認識装置30を画像キャプチャボードに組み込んでロボットに搭載するなど、実装形態に応じた構成を採用することができる。

【0062】

図6は、画像認識装置30の動作を説明するためのフローチャートである。

以下の動作は、画像認識装置30のCPU31が画像認識プログラムに従って行うものである。

まず、CPU31は、カメラ34から撮影画像1の画像データを取得してRAM33に

50

記憶する（ステップ5）。

【0063】

次に、CPU31は、撮影画像1に余白領域10を追加し、又は予め用意した拡張画像11の中央部に撮影画像1を書き込むことで、拡張画像11の画像データを作成し、RAM33に記憶する（ステップ10）。

次に、CPU31は、検出窓3a～3eを生成し、これらを拡張画像11の初期位置（左端）に設定する（ステップ15）。

【0064】

検出窓3を設定した後、CPU31は、各検出窓3に囲まれている画像（窓画像）の画像データを取得してRAM33に記憶する（ステップ20）。

そして、CPU31は、これらの画像データから、例えば、HOGなどの所定のアルゴリズムに従って特徴量を抽出してRAM33に記憶する（ステップ25）。

【0065】

次に、CPU31は、多数の基準画像7の特徴をRAM33から読み出し、窓画像の画像データから抽出した特徴と比較して、その類似性から窓画像に人物が写っているか否かを判定する（ステップ30）。

そして、CPU31は、判定結果を所定の出力先に出力した後（ステップ35）、検出窓3の右端が拡張画像11の右端に達したか否かを判断する（ステップ40）。

右端に達した検出窓3があった場合（ステップ40；Y）、その検出窓3に関しては、拡張画像11を全て走査したので画像認識処理を終了する。

【0066】

一方、まだ右端に達していない検出窓3があった場合（ステップ40；N）、その検出窓3に関しては、まだ画像認識処理が終了していないため、当該検出窓3を右水平方向に移動して（ステップ45）、ステップ20に戻る。

【0067】

以上の処理では、検出窓3a～3eによる人物像5の検索を平行して同時に行ったが、例えば、まず、検出窓3aによる検索を行った後、検出窓3bによる検索を行うなど、順に行ってもよい。

【0068】

以上に説明した実施の形態により、次のような効果を得ることができる。

（1）元画像の縦横に余白を追加することにより元画像を大きくすることができ、追加余白部分に対応した検出窓を追加した検出器に、その画像をかけることで、一部がカメラ外に出てしまった対象が認識可能となる。

（2）画像認識できる特徴量を有するにもかかわらず、撮影画像1の周辺部から一部がはみ出るため従来は捨てられていた画像に対しても画像認識を行うことができる。

（3）余白領域10を追加することにより、画像認識できる範囲を撮影画像1の外側に拡張することができる。

（4）撮影画像1からはみ出た人物像5を認識するために撮影画像1からはみ出た人物像5を大量に記憶する必要がなく、全身を撮影した標準の基準画像7を用いることができる。

【符号の説明】

【0069】

- 1 画像
- 2 基準線
- 3 検出窓
- 4 窓画像
- 5 人物像
- 7 基準画像
- 8 人物像
- 10 余白領域

10

20

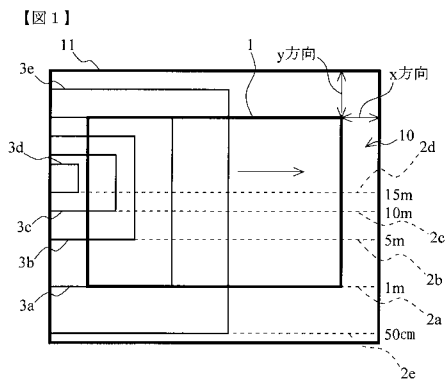
30

40

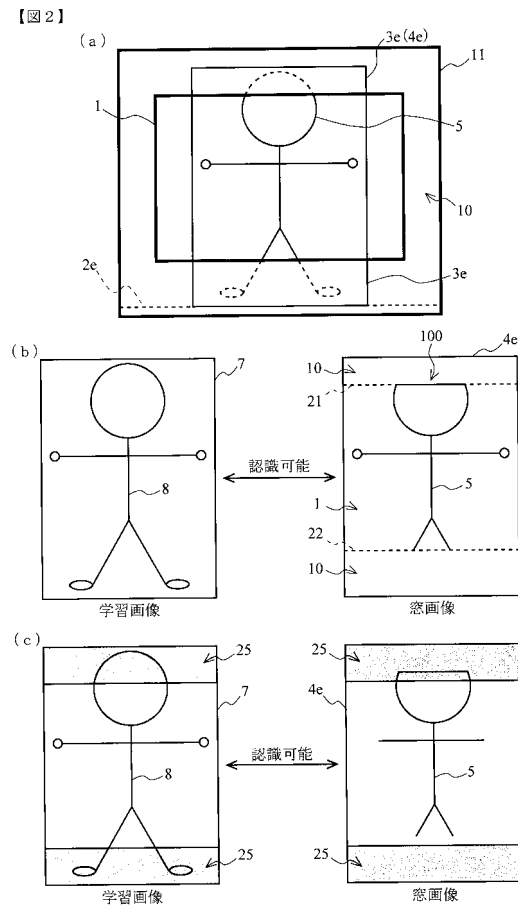
50

- 1 1 拡張画像
- 2 1、2 2、2 3 境界線
- 2 5 領域
- 3 0 画像認識装置
- 3 1 CPU
- 3 2 ROM
- 3 3 RAM
- 3 4 カメラ
- 3 5 記憶装置
- 3 6 入力装置
- 3 7 出力装置
- 1 0 0 エッジ部分

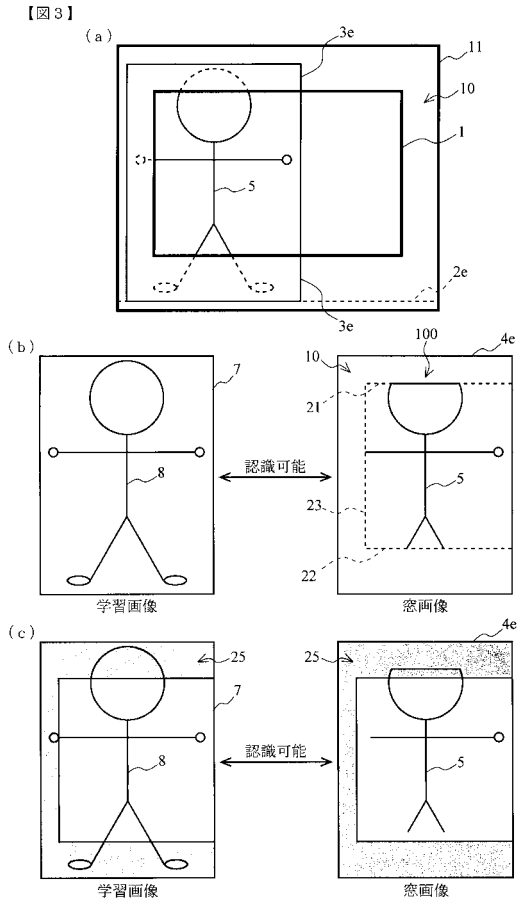
【図1】



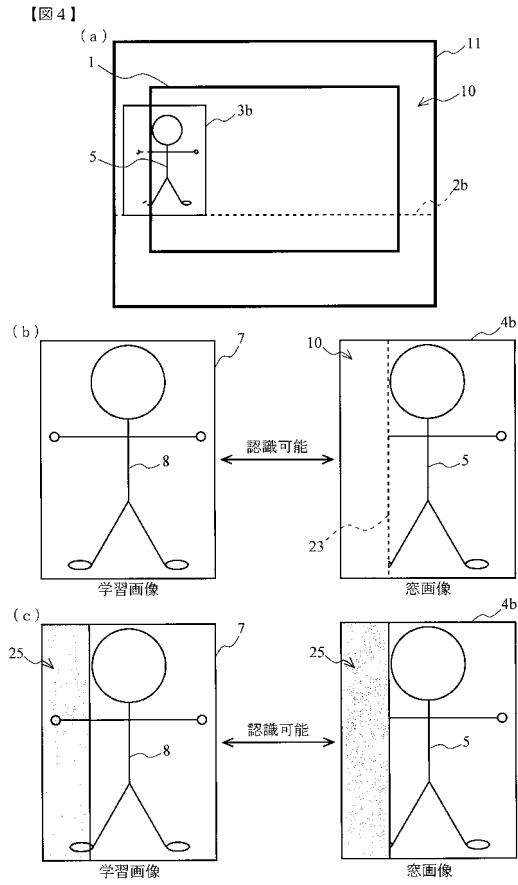
【図2】



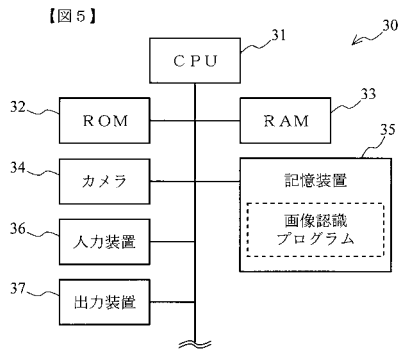
【図3】



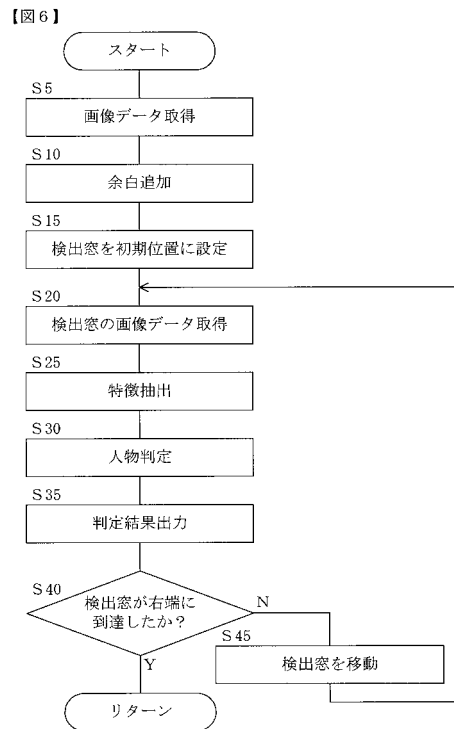
【図4】



【図5】

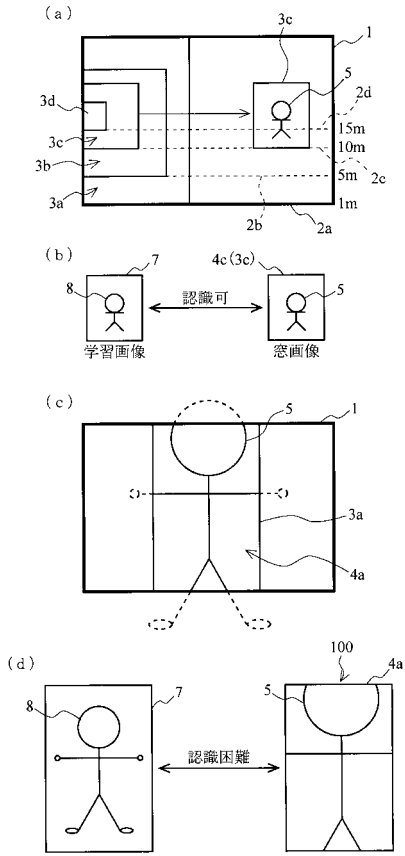


【図6】



【 図 7 】

【 図 7 】



フロントページの続き

Fターム(参考) 5B057 AA05 BA02 CA12 CA16 CB12 CB16 CC03 DA08 DC33
5L096 CA02 DA01 EA03 EA14 EA35 FA35 GA04 HA08 JA11 JA18