

(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2004-282535
(P2004-282535A)

(43) 公開日 平成16年10月7日(2004.10.7)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
HO4N 5/232	HO4N 5/232 C	5C022
GO6T 7/00	GO6T 7/00 100C	5L096

審査請求 未請求 請求項の数 9 O L (全 28 頁)

(21) 出願番号	特願2003-73037 (P2003-73037)	(71) 出願人	000006013 三菱電機株式会社 東京都千代田区丸の内二丁目2番3号
(22) 出願日	平成15年3月18日 (2003.3.18)	(74) 代理人	100099461 弁理士 溝井 章司
		(74) 代理人	100111800 弁理士 竹内 三明
		(74) 代理人	100114878 弁理士 山地 博人
		(72) 発明者	田口 進也 東京都千代田区丸の内二丁目2番3号 三 菱電機株式会社内
		(72) 発明者	田中 昭二 東京都千代田区丸の内二丁目2番3号 三 菱電機株式会社内

最終頁に続く

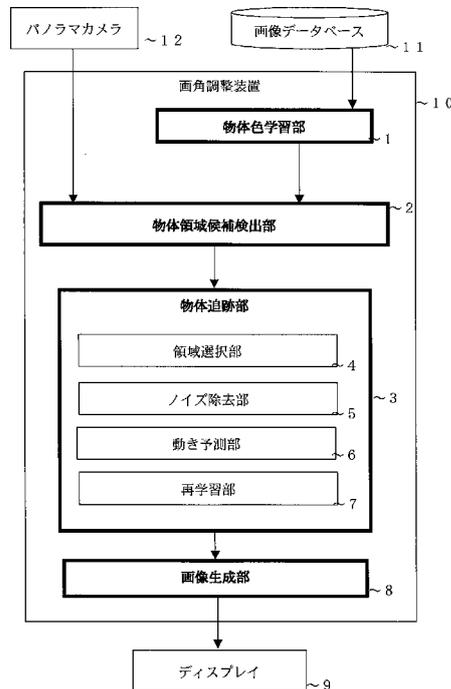
(54) 【発明の名称】 画角調整装置

(57) 【要約】

【課題】 手持ちカメラ等による瞬時変動が大きい画像入力に対しても対象物体を安定して抽出し、画像表示を行う。

【解決手段】 広角で撮影した入力画像から所定の対象部分を中心とする画像範囲を抽出表示する構成において、入力画像の特定部分を色相と彩度を基に色を特定する物体色学習部1と、この特定した色情報を参照して入力画像のそれぞれの物体領域範囲を推定する物体領域候補検出部2と、入力画像の時間変動を抑える変動除去部5と、時間変動を抑えた物体領域の動きを予測して特定物体領域を追跡する物体追跡部3とを備えて、この追跡する特定物体を中心として所定の画角範囲を画像表示するようにした。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

広角で撮影した入力画像から所定の対象部分を中心とする画像範囲にある部分を抽出表示する構成において、

上記入力画像の特定部分を色相と彩度を基に色を特定する物体色学習部と、

上記特定した色情報を参照して上記入力画像のそれぞれの物体領域範囲を推定する物体領域候補検出部と、

上記入力画像の時間変動を抑える変動除去部と、

上記時間変動を抑えた物体領域の動きを予測して特定物体領域を追跡する物体追跡部とを備えて、上記追跡する特定物体を中心として所定の画角範囲を画像表示するようにしたことを特徴とする画角調整装置。

10

【請求項 2】

複数の物体の色相と彩度をサンプルとして記憶する画像データベースを備え、物体色学習部は上記画像データベースを参照して特定部分の色を識別するようにしたことを特徴とする請求項 1 記載の画角調整装置。

【請求項 3】

物体色学習部は、画像データベースに記憶している識別基準の物体の画像サンプルを色相と彩度で表現し、該色相と彩度で入力画像を識別し、物体領域候補検出部は、上記識別結果に基づいて物体領域を定めて候補とするようにしたことを特徴とする請求項 2 記載の画角調整装置。

20

【請求項 4】

物体領域候補検出部は、入力画像から切り出して複数の物体領域を定める際に、所定色における確率密度が閾値以上であるものを物体領域として検出し、かつ所定の大きさの領域に分割 / 合併を行なって物体領域の数を定めるようにしたことを特徴とする請求項 1 記載の画角調整装置。

【請求項 5】

物体領域候補検出部は、求めた候補領域数と、確率を規範とする方法とで混合数を決定し、改めて候補領域を求めるようにしたことを特徴とする請求項 4 記載の画角調整装置。

【請求項 6】

動き予測は、前フレームの入力画像において物体追跡部が特定した特定物体領域の位置情報と、現フレームの入力画像における上記特定物体領域の位置情報とから速度情報とを求めて、該求めた位置と速度情報とに基づいて次フレームの画像範囲を定めるようにしたことを特徴とする請求項 1 記載の画角調整装置。

30

【請求項 7】

変動除去部は、特定物体領域の位置情報と、速度情報とを用いて次フレームの特定物体領域の位置をフィルタリングして推定し、物体追跡部は、該推定結果に基づいて動き予測して画像範囲を定めるようにしたことを特徴とする請求項 6 記載の画角調整装置。

【請求項 8】

特定物体領域における色の色相と彩度を再学習する再学習部を備えて、該再学習により得られた色を以後のフレームにおける特定物体領域の識別基準としたことを特徴とする請求項 1 記載の画角調整装置。

40

【請求項 9】

特定物体領域を指定する物体指定部を備えて、識別基準として上記特定物体領域を指定されると、物体追跡手段は、上記追跡する特定物体領域を中心として所定の画像範囲を画像表示するようにしたことを特徴とする請求項 1 記載の画角調整装置。

【発明の詳細な説明】**【0001】****【発明の属する技術分野】**

この発明は、カメラで撮像した画像から通話者あるいは特定の物体の位置を抽出して検出し、常にこの検出された部分が画面の中心となるように画角を調整する画角調整に関する

50

ものである。

【0002】

【従来の技術】

携帯電話やPDAを使用したテレビ電話では、ユーザは端末を手で持ちながら、さらには移動中に通話を行う場合がある。このような状況では、ともすれば捕らえられた通話者の顔が画面のあちこちに移動するので、この場合は顔を常に画面の中心とするようにカメラの画角を調整する必要がある。

この画角調整の技術として、従来から、移動物体の動きにカメラを追従させる方法が用いられてきた。しかしカメラを追従させる方法では、カメラの画角内に移動物体が存在しないと機能しないため、目的とする移動物体の移動速度が速く、カメラの画角から外れてしまった場合には、自動追跡ができなくなってしまう。

10

【0003】

これを解決する別の方法として、パノラマカメラにより周囲360°の全方位画像を取り込み、その撮像画像から物体を検出して追跡する方法が提案されている。

例えば、第1の従来例として特許文献1では、1台の静止した魚眼レンズカメラを用いて複数の人物像を同時に追跡し、多画面表示やシーケンス表示により追跡した人物をわかりやすく表示する。

図18に、この監視装置の構成を示す。この監視装置は、全方位画像の移動領域を検出する移動領域検知手段51と、検出された移動領域の画素を移動物体ごとにグループ化してラベリングするラベリング手段52と、ラベリングされた物体の移動を追跡する移動領域追跡手段53と、全方位画像からラベリングされた移動物体の画像領域を切り出し、画像領域の画像データを補正する正像変換手段54と、ラベリングされた移動物体の数に応じて表示画面を制御する表示制御手段55とを備える。

20

【0004】

この監視装置では、移動物体を検出するアルゴリズムとして、キャプチャされた画像データと背景画像との差分を計算する方法を採用している。図19は、この監視装置の移動領域検知手段が行う機能を説明する図である。この移動領域検知手段は、まず入力する画像に(1 -)を乗じ、また背景画像に を乗じてこれらを加算部45で加算する。このことで、入力画像の動きを示す部分が減衰された画像データを得ることができる。次に、背景映像出力部47は、加算部45の出力を取得すると、減算部48に出力し、減算部48

30

【0005】

また、例えば、第2の従来例として特許文献2では、1台の静止したパノラマカメラにより周囲360°の画像を撮像し、画像処理技術により周囲の移動物体を検知して追尾する。この移動物体追尾装置では、移動物体を検出するアルゴリズムとして、256階調の画像を背景と移動物体(追尾対象)との2階調に階調変換し、フレーム差分をとる方法を採用している。この処理により、移動量の無い背景は輝度差の値が0となり、移動物体は前フレームと現フレームとの間で輝度差ができるため、この輝度差の値がある値以上のものを

40

【0006】

【特許文献1】

特開2001-333422号公報

【特許文献2】

特開2002-64812号公報

【0007】

【発明が解決しようとする課題】

しかしながら、上記の従来例1又は従来例2に示される装置は、例えば、携帯電話やPDAにおけるテレビ電話など、カメラが固定されていない状況では使用不可能である。なぜ

50

なら、携帯電話やPDAを使用したテレビ電話では、通常ユーザは端末を手で持ちながら、さらには移動中に通話を行うため、このような状況ではカメラが細かく振動して、撮像した全体の画像にずれが生じる。このような状況では、上記のようなフレーム間の差分を取るアルゴリズムでは全てが移動物体となり、正しい移動物体部分の特定や検出ができないという課題がある。

【0008】

本発明は、上記の課題を解決するためになされたもので、カメラを手で持つ場合や、さらには移動中にカメラを使用する場合のように、カメラが固定されていない状況でも、パノラマカメラにより周囲360°の全方位画像を撮像し、その撮像画像情報から画像処理技術により画像のブレを補正し、適切に移動物体を検出して追跡し、その物体が常に画像の中心となるように画角調整を行うことを目的とする。

10

【0009】

【課題を解決するための手段】

この発明に係る画角調整装置は、広角で撮影した入力画像から所定の対象部分を中心とする画像範囲を抽出表示する構成において、入力画像の特定部分を色相と彩度を基に色を特定する物体色学習部と、この特定した色情報を参照して入力画像のそれぞれの物体領域範囲を推定する物体領域候補検出部と、入力画像の時間変動を抑える変動除去部と、時間変動を抑えた物体領域の動きを予測して特定物体領域を追跡する物体追跡部とを備えて、この追跡する特定物体を中心として所定の画角範囲を画像表示するようにした。

20

【0010】

【発明の実施の形態】

実施の形態1

以下に、図1から図14を用いて、本実施の形態における画角調整装置の構成と動作を説明する。

図1は、本実施の形態における画角調整装置の構成を示す図である。また、図2は、その動作を示したフローチャートである。更に、図3は、本実施の形態における動作を説明するための具体事例を示す説明図である。

本実施の形態1では、パノラマ画像からある一つの特定物体又は部分を検出し、常にその物体又は部分が画面の中心に表示されるように画角を調整する動作を説明する。

30

【0011】

図1示す画角調整装置10は、予め用意した画像データベース11からある特定物体の色を学習する物体色学習部1と、パノラマカメラ12によりキャプチャした全方位画像から、物体色学習部1により学習した特定物体が存在する候補領域を求める物体領域候補検出部2と、手ブレに伴う画像中の振動ノイズを除去して、特定物体を追跡する物体追跡部3と、特定物体を中心とする画像を切り出す画像生成部8とを備える。

【0012】

次に、図1に示す画角調整装置10の各構成要素の機能を説明する。

この装置は、物体領域候補検出部2で検出した複数の特定物体の候補領域から、前フレームのノイズ除去部5において推定した特定物体位置座標に最も近い領域を求め、求めた領域の中心位置座標と画素数を取得する領域選択部4と、物体領域候補検出部2により取得した特定物体中心位置座標と前フレームで予測した特定物体中心位置座標に基づき、現フレームの特定物体中心位置座標を推定することにより、手ブレに伴うカメラの振動ノイズを除去する変動除去部の一種であるノイズ除去部5と、前記ノイズ除去部において特定物体中心位置座標を推定するために必要な、次フレームの特定物体中心位置座標の予測をする、同じく組み合わせて変動除去を行う、動き予測部6と、現時点での照明条件下における物体の色を再学習する再学習部7とを備えている。

40

【0013】

次に、図2に示すフローチャートを用いてパノラマカメラによる画角調整装置10の概略動作を説明する。

50

本実施の形態における画角調整装置の動作は、物体の色を画像データベースに基づき各物体の色を学習するフェイズ1と、フェイズ1で学習した色を持つ特定物体が常に画角に入るよう調整するフェイズ2の2段階に分かれる。

【0014】

(フェイズ1)

まず、物体色学習部1は、予め用意した画像データベース11を用いて、追跡すべき特定物体の色を学習する(ステップS110)。学習した色は、キャプチャ画像から特定物体の候補領域を検出する際(ステップS130)の基準として利用する。なお、画像データベース11は、全色の色相と彩度を細かく基準として持つてもよいし、幾つかを間引き、省略した簡略色彩のデータベースであってもよい。

10

(フェイズ2)

まず、パノラマカメラ等の広角カメラを使用して全方位画像をキャプチャする(ステップS120)。

【0015】

続くステップS130では、物体領域候補検出部2において、キャプチャした全方位画像から、ステップS110で学習した特定物体が存在する候補領域を検出する。一般に複数の候補領域が検出される。例えば、テレビ電話において人物を検出する場合には、通話者の他に、通話者の周りに存在する人物や、壁にかけられた人物のポスター等を同時に検出される場合があるが、これらすべてを人物の候補領域とする。

20

【0016】

続くステップS140では、領域選択部4において、前のステップS130で検出した複数の特定物体の候補領域から、前フレームのノイズ除去部5において推定した特定物体位置座標に最も近い領域を求め、求めた領域の中心位置座標を取得する。

ただし、初期フレームにおいては、領域選択部4はステップS130で検出した複数の特定物体の候補領域の中から、予め決められた画像中のある位置(例えば、Xとする)に最も近い領域を求め、求めた領域の中心位置座標を取得する。例えば、テレビ電話において常に通話者が中心となる画像を切り出す場合には、始めに通話者が画像中の上記のXの位置に映るように設定して、以降、画角調整装置10の動作を開始する。

【0017】

続くステップS150では、ノイズ除去部5において、前のステップS140で取得した領域中心位置座標と前フレームで予測した特定物体中心位置座標に基づき、現フレームの特定物体中心位置座標を推定する。

30

続くステップS160では、動き予測部6は、前のステップS150で推定した特定物体中心位置を用いて、物体が等速直線運動するという仮定から次フレームにおける特定物体中心位置の予測を行う。この予測値は、次フレームでノイズ除去部5において物体位置を正確に求めるために使用する。

【0018】

続くステップS170では、照明条件等により毎フレームに変化する物体の色に適應するために、再学習部7において特定物体の色を再学習する。

続くステップS180では、画像生成部8は、ステップS160で推定した特定物体中心位置を中心とする画像を切り出す。

40

最後に、ステップS120に戻り、上記の処理を繰り返す。

上記の処理をフレーム毎に繰り返すことで、動画中の物体検出を行い、その物体が中心となる動画を生成する。

【0019】

図3は、実施の形態1における動作を説明するための説明図である。

始めに、図3(a)に示すように、パノラマカメラ13で周囲360°の全方位の画像を取得する。次に、撮像した(b)に示す全方位画像14から、画像処理によって通話者の位置を検出し、図5(c)のように、通話者の顔が常に中心になる画像15を切り出す。

切り出した中心になる画像15は相手端末16に送信される。切り出された物体画像をテ

50

レビ電話の送信画像として使用することで、通話者は画角を気にしなくとも常に自分を撮像でき、通話相手は、常に人物の顔が画面の中心となる安定した映像を受信することができる。

【0020】

次に、図1に記載の画角調整装置10の各構成要素の動作を詳細に説明する。物体色学習部1)

まず、図4から図6を用いて、物体色学習部1の動作について説明する。

図4は、物体色学習部1の動作を説明するためのフローチャートである。また、図5は、肌色領域からサンプリングを行い、サンプリングされた肌色画素の色相と彩度をプロットしたグラフである。更に、図6は、物体の色を認識する閾値を示す図である。

10

ここでは、特定物体を記号Aで表す。本実施の形態では、特定物体Aの色の特徴として、色相と彩度に着目する。これは、どの物体も色相と彩度の空間で固有の分布を持ち、物体認識の基準として使用できるためである。図5に、肌色領域からサンプリングを行い、サンプリングされた画素の色相と彩度をプロットしたグラフを示す。

【0021】

まず、ステップS200で特定の物体画像を抽出し、その画像から均一に十分な数だけ画素をサンプリングする。次いで、ステップS210で、このサンプリングされたデータから特定物体Aの色相と彩度に関する確率分布 $P_c(f|A)$ を求め、物体認識のための基準として利用する。ここで f は色相(H)と彩度(S)のベクトル $f=(H, S)$ である。色相と彩度は赤(R)と緑(G)と青(B)の情報から、例えば、次式に従って求める。

20

【0022】

【数1】

$$Y = 0.3R + 0.59G + 0.11B \quad (1)$$

$$C_1 = R - Y \quad (2)$$

$$C_2 = B - Y \quad (3)$$

30

$$H = \tan^{-1}(C_1/C_2) \quad (4)$$

$$S = \sqrt{C_1^2 + C_2^2} \quad (5)$$

【0023】

この確率分布 $P_c(f|A)$ は、物体領域候補検出部2における検出基準となり、例えば、ある適当な閾値 $P_{c.o.l.o.r}$ を設定し、ある画素の f が次の式(6)を満たすならば、その画素はその物体領域に属するとみなすことができる。

40

$$P_c(f|A) > P_{c.o.l.o.r} \quad (6)$$

反対に式(6)を満たさないならば、その画素はその物体領域に属しないと判断することができる。図6に、この上記式(6)による物体認識の概念を示す。図6では、簡略化のため色相Hの一次元確率分布を図示している。図6で、ある色相 h_1 における確率密度は閾値 $P_{c.o.l.o.r}$ よりも小さいため、この色相 h_1 を持つ画素は物体に属すると判断されない。一方、ある色相 h_2 における確率密度は閾値 $P_{c.o.l.o.r}$ よりも大きいため、この色相 h_2 を持つ画素は物体としてその領域に存在していると判断される。

【0024】

次に、確率分布 $P_c(f|A)$ を正規混合モデル(ガウス分布に重みを掛けて重ね合わせたもの)で近似的に学習する。正規混合モデルは次式(7)で与えられる。

【0025】

50

【数 2】

$$P_c(f|A) = \sum_{i=1}^K \omega_i^c N(f|\mu_i^c, \Sigma_i^c) \quad (7)$$

【0026】

ここで、色相と彩度の2次元空間をUとし、Nとして次式(8)を定義する。

【0027】

【数 3】

$$N(f|\mu_i^c, \Sigma_i^c) = \exp\left[-\frac{1}{2}(f - \mu_i^c)^T \Sigma_i^{c-1} (f - \mu_i^c) - \frac{1}{2} \log|\Sigma_i^c| - \log(2\pi)\right] \quad (8)$$

10

$$\mu_i^c \in U \quad (9)$$

$$\Sigma_i^c \in U \times U \quad (10)$$

$$\omega_i^c \quad (11)$$

20

【0028】

上記式(8)において、Nは平均が μ で、分散が Σ の正規分布である。なお、式(8)において、 T は行列の転置を表し、 $|\Sigma|$ は Σ の行列式を表すものとする。また、平均 μ と分散 Σ の次元は、それぞれ式(9)と式(10)で定義される。また、式(7)におけるKは重ね合わせる正規分布の数、つまり、混合数を表し、更に、式(11)の ω_i^c は混合比(i番目の正規分布から色相と彩度のデータが発生する比率)である。

【0029】

以上のモデル化により、次の式(12)で表されるパラメータ集合 θ を求めて、物体領域検出の基準となる確率分布 $P_c(f|A)$ を求めることができる。

【0030】

【数 4】

$$\theta = \{\omega_i^c, \mu_i^c, \Sigma_i^c\}_{i=1}^K \quad (12)$$

30

【0031】

次に、図4のフローチャートを用いて、予め用意した物体の画像に基づいてパラメータ集合を決定する方法を説明する。

まず、予め用意した物体の画像からランダムに画素をサンプリングし、画素のRGB値を、色相(H)と彩度(S)に変換する。サンプリングをJ回繰り返し、次式(13)で表される色相と彩度の2次元データ集合Dを得る(ステップS200)。

40

【0032】

【数 5】

$$D = \{f_n\}_{n=1}^J \quad (13)$$

【0033】

ここで、 f_n はn番目にサンプリングされた画素の色相と彩度のベクトルである。図5に、肌色領域のサンプリングの例を示す。

50

次に、パラメータ集合を、例えば、最尤推定によって決定する（ステップ S 2 1 0）。即ち、式（14）に記載のデータ集合 D に基づいて対数尤度を最大化するようなパラメータ集合を求める。

【0034】

【数6】

$$\log P_c(D|A) = \sum_{i=1}^J \log P_c(f_i|A) \quad (14)$$

10

【0035】

例えば、最尤推定によりパラメータを求める方法として、EMアルゴリズム（A. O. Dempster, N. M. Laird, and D. B. Rubin, Maximum likelihood from incomplete data via the em algorithm, Journal of the Royal Statistical Society, Series B (Methodological), vol. 39, no. 1, pp. 1-38, 1977.)が使用できる。EMアルゴリズムとは、次に示すE (Estimation、推定) - stepとM (Maximization、最大化) - stepを繰り返すことで、対数尤度を増加させ、漸近的に最尤推定値を求める方法である。パラメータ集合を上

20

【0036】

E - step

バーを現在の推定パラメータとする。このとき、観測値 f_n が i 番目の正規分布より発生する確率を次式（15）に従って求める。

【0037】

【数7】

$$P_c(i|f_n, \bar{\theta}) = \frac{P_c(f_n, i|\bar{\theta})}{\sum_i^K P_c(f_n, i|\bar{\theta})} \quad (15)$$

30

M - step

$$Q(\theta|\bar{\theta}) = \sum_{n=1}^J \sum_{i=1}^K P_c(i|f_n, \bar{\theta}) \log P_c(f_n, i|\theta) \quad (16)$$

【0038】

上記式（15）で求めた確率から、上式（16）に示す対数尤度の期待値 Q を計算し、これを最大化するようなパラメータを求める。即ち、これがM - stepである。

40

バー = とおいて、E - step（式（15）を求めるステップ）に戻る。

【0039】

本発明におけるE - step, M - stepアルゴリズムの動作を以下に示す。このとき、混合数を K とする。E - stepとM - stepを収束するまで繰り返すことにより、逐次的にパラメータ集合と、確率 i, n を次式（17）で表して求める。

【0040】

【数8】

$$\pi_{i,n} = \left\{ \pi_{i,n}^c \right\}_{i=1, n=1}^{K, J} \quad (17)$$

但し、

$$\pi_{i,n}^c = P_c(i | f_n, \theta) \quad (18)$$

【0041】

即ち、E - s t e pでは、現在推定しているパラメータ集合 から、確率 $\pi_{i,n}$ を式 (19) に基づいて計算する。

10

【0042】

【数9】

$$\pi_{i,n}^c = \frac{\omega_i^c N(f_n | \mu_i^c, \Sigma_i^c)}{\sum_{i=1}^K \omega_i^c N(f_n | \mu_i^c, \Sigma_i^c)} \quad (19)$$

【0043】

ここで、確率 $\pi_{i,n}$ はデータ f_n が与えられたときに、それが i 番目の正規分布に 20 属する確率である。

M - s t e pでは、E - s t e pで計算した確率を使用して期待値や分散を計算することで、パラメータ集合を次式により求める。ここで、 T は行列の転置を表す。

【0044】

【数10】

$$\omega_i^c = \frac{\langle 1 \rangle_i}{J} \quad (20)$$

$$\mu_i^c = \frac{\langle f \rangle_i}{\langle 1 \rangle_i} \quad (21)$$

$$\Sigma_i^c = \frac{\langle (f_n - \mu_i^c)(f_n - \mu_i^c)^T \rangle_i}{\langle 1 \rangle_i} \quad (22)$$

$$\langle 1 \rangle_i \equiv \sum_{n=1}^J \pi_{i,n}^c \quad (23)$$

$$\langle f \rangle_i \equiv \sum_{n=1}^J f_n \pi_{i,n}^c \quad (24)$$

である。

また、 μ_i^c は正規分布の中心、

Σ_i^c は正規分布の分散、

ω_i^c は混合比 (i番目の正規分布から色相と彩度のデータが発生する比率) である。

【0045】

パラメータを求める際には、混合数 K を予め与えておく。ここで、例えば、肌色を学習する場合は、その確率分布は 1 つの正規分布で十分近似できるため、 $K = 1$ と予め決めてからパラメータを決定することができる。

【0046】

物体領域候補検出部 2)

次に、図 7 から図 13 を用いて、物体領域候補検出部 2 の動作の詳細を説明する。

図 7 は、物体領域候補検出部の動作を示すフローチャートである。図 8 は、正規混合モデルによって 1 次元の確率分布を近似する例を示した説明図である。図 9 は、物体領域候補検出部で全方位画像から肌色を検出する例を示した説明である。図 10 は、混合数決定における正規分布生成の一例を示す図である。図 11 は、混合数決定における正規分布消滅の一例を示す図である。図 12 は、混合数決定における正規分布分割の一例を示す図である。図 13 は、混合数決定における正規分布合併の一例を示す図である。

【0047】

本手段では、画像中から特定物体 A が存在する候補領域を求める。そのために、まず、画像上のある位置 $u = (x, y)$ に物体 A が存在する条件付確率分布 $P(u | A)$ を正規混合モデルによって求める。ここで、 $u = (x, y)$ は画像の左上を原点とし、下方向に Y 軸を、右方向に X 軸を取る場合の 2 次元ベクトルである。

【0048】

正規混合モデルを使用することで、複数の物体領域候補を同時に検出することができる。例えば、テレビ電話において人物を検出する場合には、通話者の他に、通話者の周りに存在する人物や、壁にかけられた人物のポスター等が存在する。このような場合、1 つの顔に 1 つの正規分布を対応させて表現できれば、確率分布 $P(u | A)$ は各正規分布を統合

10

20

30

40

50

したものとして表現される。

図 8 に、正規混合モデルによる 1 次元の確率分布を近似する際の例を示す。肌色領域からサンプリングを行い 1 次元のデータ集合（図 8 (a) では、ヒストグラムとして表示している）を得る。このデータ集合を発生させる確率分布 $P(u | A)$ を 3 つの正規分布で近似する。このとき、正規分布の平均 μ_i を各物体の中心とみなすことができ、次式 (25) で表される楕円の面積（画素数）を物体の大きさとみなすことができる。

【 0 0 4 9 】

【 数 1 1 】

$$(u - \mu_i)^T \Sigma_i^{-1} (u - \mu_i) = 1 \quad (25)$$

10

【 0 0 5 0 】

また、この方法を用いれば、例えば、図 8 のように 2 つの顔が重なっていても、2 つの正規分布の重ね合わせとして推定できる。

正規混合モデルは次式で与えられる。

【 0 0 5 1 】

【 数 1 2 】

$$P(u | A) = \sum_{i=1}^m \omega_i N(u | \mu_i, \Sigma_i) \quad (26)$$

20

【 0 0 5 2 】

ここで $N(u | \mu_i, \Sigma_i)$ は平均が μ_i で分散が Σ_i の正規分布であり、 m は正規分布の混合数（正規分布の数）であり、 ω_i は混合比（ i 番目の正規分布からデータが発生する比率、あるいは、画像中にある m 個の物体のうち i 番目の物体の存在比率）である。

以上のモデル化によって、物体の中心と大きさを求めることは、次式 (27) で表されるパラメータ集合を求めることに帰着する。

【 0 0 5 3 】

【 数 1 3 】

$$\{\omega_i, \mu_i, \Sigma_i\}_{i=1}^m \quad (27)$$

30

【 0 0 5 4 】

以下では、図 7 に示すフローチャートとともに、キャプチャした画像からパラメータ集合を決定する方法を説明する。

始めに、パノラマカメラによりキャプチャされた全方位画像から均一に十分な数だけ画素をサンプリングする。サンプリングされた画素の RGB 値を色相と彩度 $f = (H, S)$ に変換する。物体色学習部 1 において求めた確率分布 $P_c = (f | A)$ を基準として物体領域の検出を行う。ここで、閾値 $P_{c.o.l.o.r}$ を設定し、もし、次式 (28) を満足するなら、データとして採用し、それ以外ならば棄却する。このようにして得られたデータ集合を式 (29) とする（ステップ S 3 1 0）。ここで、 B はデータ数、 u_n はデータとして採用された各画素の位置である。

40

【 0 0 5 5 】

【 数 1 4 】

$$P_c(f | A) > P_{\text{color}} \quad (28)$$

$$U = \{u_n\}_{n=1}^B \quad (29)$$

【 0 0 5 6 】

この操作により、サンプリングした画素のうち、学習した物体の色に近いものだけをデータとすることができる。

次に、E - step , M - step アルゴリズムにより、次式 (3 0) で表されるパラメータ集合を求める (ステップ S 3 2 0 , ステップ S 3 3 0) 。

【 0 0 5 7 】

【 数 1 5 】

$$\{\omega_i, \mu_i, \Sigma_i\}_{i=1}^m \quad (30)$$

$$U = \{u_n\}_{n=1}^B \quad (31)$$

$$\{\pi_{i,n}\}_{i=1, n=1}^{m, B} \quad (32)$$

【 0 0 5 8 】

本手段における E - step , M - step アルゴリズムの動作を以下に示す。

上記ステップ S 3 1 0 で求めたデータ集合である式 (3 1) を基に、次の E - step と M - step を 3 ~ 5 回交互に繰り返し、式 (3 0) で表されるパラメータ集合と、式 (3 2) で表される確率を求める。パラメータの初期値としては、前フレームで計算された値を使用する。

【 0 0 5 9 】

・ E - step (ステップ S 3 2 0)

現在推定しているパラメータ集合から、次式 (3 3) により、確率を求める。

【 0 0 6 0 】

【 数 1 6 】

$$\pi_{i,n} = \frac{\omega_i N(u_n | \mu_i, \Sigma_i)}{\sum_{i=1}^m \omega_i N(u_n | \mu_i, \Sigma_i)} \quad (33)$$

【 0 0 6 1 】

・ M - step (ステップ S 3 3 0)

式 (3 3) の確率を使用して期待値や分散を計算することで、式 (3 0) のパラメータ集合を次式 (3 4) ないし式 (3 8) により求める。ここで、T は行列の転置を表す。

【 0 0 6 2 】

【 数 1 7 】

10

20

30

40

$$\omega_i = \frac{\langle 1 \rangle_i}{B} \quad (34)$$

$$\mu_i = \frac{\langle u \rangle}{\langle 1 \rangle_i} \quad (35)$$

$$\Sigma_i = \frac{\langle (u_n - \mu_i)(u_n - \mu_i)^T \rangle_i}{\langle 1 \rangle_i} \quad (36)$$

$$\langle 1 \rangle_i \equiv \sum_{n=1}^B \pi_{i,n} \quad (37)$$

$$\langle u \rangle_i \equiv \sum_{n=1}^B u_n \pi_{i,n} \quad (38)$$

10

【0063】

また、 $\pi_{i,n}$ はデータ u_n が i 番目の正規分布に属する確率である。
 上記アルゴリズムにより求めた正規分布の平均 μ_i を物体の中心とし、式(25)の楕円の面積(画素数)を物体の大きさとする。式(30)のパラメータ集合をフレーム毎に求めることで、動画中の物体検出ができる。

20

図9に、全方位画像から肌色領域をサンプリングし、そのデータが発生する確率分布を正規混合モデルで求めた例を示す。図9(c)には、各正規分布の輪郭と平均を図示している。

【0064】

混合数の決定法を説明する。

上記の計算では、混合数(物体の数)は固定であったが、動画において物体の数は変化する。そこで、E-step, M-stepの次に混合数 m を決定する。一般にE-step, M-stepアルゴリズムで正規混合分布のパラメータ推定を行う場合、混合数(重ね合わせる正規分布の数)を予め決めておく必要がある。しかし、例えば、テレビ電話のような状況では入退場などにより人物の数が時間的に変化するため、予め混合数を決めておくことはできない。そこで、プロップ(連結した物体の色の領域)数の検出と、例えば、確率を規範とする方法で混合数を決定する。

30

【0065】

まず、プロップ数を求める手順を示す。始めに、各画素の色相と彩度 $f = (H, S)$ を求める。物体色学習部もしくは再学習部で求めた $P_c(f|A)$ を使用し、もし、次式(39)が成立するなら1、それ以外ならば0とすることで、キャプチャ画像と同じサイズの2値のマスクを作る。

40

$$P_c(f|A) > P_{c.o.l.o.r} \quad (39)$$

ここで、 $P_{c.o.l.o.r}$ は適切な閾値である。マスクの中で1が連結した領域数を求め、プロップ数とする。

【0066】

プロップ数と、ある確率を規範とし混合数を決定する動作手順を説明する。

(初期フレーム)プロップ数を検出し、混合数とする(ステップS350)。

(2フレーム目以降)以下の手順に従い各フレームにおいて正規分布の生成(ステップS360)、消滅(ステップS370)、分割(ステップS380)、合併(ステップS390)を、この順番で実行する。

生成(ステップS360): 図10に正規分布が生成される例を示す。例えば、物体が急

50

に画面に現れた場合には、この物体を検出するために混合数を増やす必要がある。そこで、あるサンプリングデータ u_n がどの正規分布の中心からも離れた位置座標にあるならば、新たに物体が現れたと判断し、再びプロップ数の検出を行い新しい混合数とする。サンプリングデータ u_n と正規分布の中心の距離を判断する基準として、例えば、 i 番目の正規分布からあるサンプリングデータ u_n が発生する次式 (40) で表される確率を使用する。具体的な手順として、先ず十分に小さい閾値 P_{gene} を設定し、次にあるサンプリングデータ u_n に対して次式 (41) を評価する。次式 (41) が成り立つ場合には、このサンプリングデータ u_n はどの正規分布からも離れた位置座標に存在することを意味する。そこで、再びプロップ数を検出し、新しい混合数として設定する。

【0067】

【数18】

$$P(u_n, i | A) = \omega_i N(u_n | \mu_i, \Sigma_i) \quad (40)$$

$$\max_{i=1}^m P(u_n, i | A) < P_{gene} \quad (41)$$

【0068】

消滅 (ステップ S370) : 図11に正規分布が消滅する例を示す。例えば、物体が急に画面から消えた場合には、即ち、対応部分の入力画像がなくなるので、この物体に対応する正規分布は余計となり、消滅する必要がある。そこで物体が画面から消えたことを判断するために、正規分布の中心から一定の距離に存在するデータ量を計算し、そのデータ量が少なければ余計な正規分布とみなして消滅させる。

ある i 番目の正規分布の中心から一定の距離に存在するデータ量の大小を測る基準として、例えば、 i 番目の正規分布からデータが発生する割合 ω_i を使う。具体的な手順として、先ず、式 (42) を満たす閾値 P_{del} を設定する。ここで、 m は現フレームにおける合数であり、 $1/m$ は混合比の平均値である。

$$0 < P_{del} \leq 1/m \quad (42)$$

【0069】

次に、次式 (43) を評価する。もし、次式 (43) が成立するならば、例えば、物体が急に画面から消えたために、これに対応していた i 番目の正規分布は必要でないことを意味するので、これを消滅させることで混合数を1つ減らす。

$$\omega_i < P_{del} \quad (43)$$

分割 (ステップ S380) : 図12に正規分布が分割される例を示す。即ち、大きい面積を持つ領域は、1つの物体領域として認識されない。そこで、次式 (44) で表される楕円の面積 (画素数) がある程度大きければ、2つの正規分布に分割することで混合数を1つ増やす。

【0070】

【数19】

$$x^T \Sigma_i^{-1} x = 1 \quad (44)$$

【0071】

合併 (ステップ S390) : 図13に正規分布が合併される例を示す。小さい面積を持つ領域は、独立の物体領域として認識されない。そこで、楕円の面積 (画素数) がある程度小さければ、近くの正規分布と統合することで混合数を1つ減らす。

以上のように、物体のプロップ検出と確率による判断を合わせた混合数の計算により、物体の数が突然変化する場合でも、オンラインで物体の数を推定することが可能である。

【0072】

物体追跡部3)

次に、その動作を説明するためのフローチャートである図14を用いて、物体追跡部3が

10

20

30

40

50

行う動作について説明する。これは、更に以下の要素で構成される。

領域選択部 4)

領域選択部では、物体領域候補検出部 2 で検出した複数の候補領域から、次に示す方法で、特定領域（特定物体を表現している正規分布）を選択し、選択した領域の中心位置座標 $u_t \text{ バー} = (x_t \text{ バー}, y_t \text{ バー})$ と、選択した領域の中心速度ベクトル $v_t \text{ バー} = (x_t \text{ バー} - x_{t-1} \text{ バー}, y_t \text{ バー} - y_{t-1} \text{ バー})$ を取得する。ここで、 t はフレーム番号である。

【0073】

領域選択部では、始めに、式(25)で示される楕円の1.5倍から2倍の範囲を探索領域に設定する(ステップS400)。ここで i は、前フレームの領域候補検出部において、特定物体を表現している正規分布の番号であり、 μ_i はその正規分布の平均であり、 σ_i はその正規分布の分散である。即ち、これは現フレームの物体領域候補検出部で求められた式(45)で示される複数の物体領域候補の中心において、この探索領域の中に入るものを調べる。

【0074】

【数20】

$$\{u_i\}_{i=1}^m \quad (45)$$

【0075】

このうち、探索領域の中心と最も距離が小さい物体領域中心（すなわち正規分布の中心）を、観測データ $u_t \text{ バー}$ とする(ステップS410)。

しかし、初期フレームにおいては、領域選択部 4 は、物体領域候補検出部 2 で検出した複数の候補領域の中から、予め決められた画像中のある位置（例えば、 X とする）に最も近い領域を求め、求めた領域の中心位置座標を取得する。

なお、例えば、物体が一時的に遮蔽されている場合など、探索領域の中に物体候補領域の中心が入らない場合は、観測データは無しとみなし、後述する予測ステップを実行する(ステップS490)。

さらに、もし、一定の時間、観測データ無しの場合は、現フレームの探索領域の面積を大きくし、次フレームの探索領域として設定する(ステップS510)。

【0076】

観測データ $u_t \text{ バー}$ を取得後、 t 番目のフレームにおける式(46)で与えられる速度ベクトルを計算する。ここで、2つのベクトル $u_t \text{ バー}$ と $v_t \text{ バー}$ を合わせて式(47)の d_t と表現する。なお、 T は行列の転置を表す。

【0077】

【数21】

$$\bar{v}_t = (\bar{x}_t - \bar{x}_{t-1}, \bar{y}_t - \bar{y}_{t-1}) \quad (46)$$

$$d_t = (\bar{u}_t, \bar{v}_t)^T \quad (47)$$

【0078】

ノイズ除去部 5 と動き予測部 6)

次に、ノイズ除去部 5 と動き予測部 6 の動作について詳しい説明をする。

ノイズ除去部 5 は、領域選択部 4 で取得したデータ d_t と、次の式(48)で示す前フレームで予測した特定物体中心位置座標 $u_t \text{ チルダ}$ と、特定物体中心速度ベクトル $v_t \text{ チルダ}$ に基づき、現フレームの特定物体中心位置座標 $u_t \text{ ハット}$ と特定物体中心速度ベクトル $v_t \text{ ハット}$ を推定する(ステップS440)。ここで、予測した値には“ \sim ”(チルダ)を、推定した値には“ $\hat{}$ ”(ハット)をつけて区別している。

10

20

30

40

50

【 0 0 7 9 】

【 数 2 2 】

$$\tilde{u}_t = (\tilde{x}_t, \tilde{y}_t) \quad (48)$$

$$\tilde{v}_t = (\tilde{x}_{t1} - \tilde{x}_{t-1}, \tilde{y}_t - \tilde{y}_{t-1}) \quad (49)$$

【 0 0 8 0 】

以下では、位置と速度の推定ベクトルをまとめて式 (50) と表記し、同様に、位置と速度の予測ベクトルを式 (51) と表記する。 10

【 0 0 8 1 】

【 数 2 3 】

$$\hat{z}_t = (\hat{u}_t, \hat{v}_t)^T \quad (50)$$

$$\tilde{z}_t = (\tilde{u}_t, \tilde{v}_t)^T \quad (51)$$

【 0 0 8 2 】

この表記を用いると、ノイズ除去部 5 の動作は、 z_t チルダと d_t から z_t ハットを推定することである、と言い換えることができる。 20

このように、ノイズ除去部 5 は、観測したデータ d_t だけでなく、前フレームで予測したベクトル z_t チルダを用いることで正確に物体位置の推定が行える。

【 0 0 8 3 】

続いて、動き予測部 6 は、ノイズ除去部 5 において推定した z_t ハットの情報を用いて、次フレームにおける特定物体中心位置と特定物体中心速度の予測ベクトル z_{t+1} チルダを計算する (ステップ S 4 5 0)。計算の際に、物体のダイナミクスを記述した状態遷移行列 A を使用する。

例えば、物体が等速直線運動する場合、 A は次式で与えられる。 30

【 0 0 8 4 】

【 数 2 4 】

$$A = \begin{pmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (52)$$

【 0 0 8 5 】

ここで、 t は各フレームの時間間隔である。 40

ノイズ除去部 5 と、動き予測部 6 の計算は、例えば、パーティクルフィルタ (Doucet, A., Freitas, N., and Gordon, N., editors. Sequential Monte Carlo Methods in Practice. Springer-Verlag, 2001.) によって実行することができる。パーティクルフィルタとは、後述する推定ステップと予測ステップを実行し、時間的に伝播する確率分布を求めることで、フィルタリングや物体運動の追跡を行うアルゴリズムである。

ノイズ除去部 5 の計算は、後述する推定ステップに相当する。また、動き予測部 6 の計算は後述する予測ステップに相当する。

【 0 0 8 6 】

以下で使用する記号の説明をする。

特定物体中心位置座標と特定物体中心速度ベクトルを確率変数とみなし、それぞれの値を u_t と v_t で表記する。

また、これらを合わせた式 (53) で示されるベクトル z_t を定義する。更に、式 (54) で示される P は領域選択部 4 で求めたデータ d_t に基づいた次フレームの特定物体中心位置座標と特定物体中心速度ベクトルの確率分布であり、式 (55) の P は領域選択部 4 で求めたデータ d_t に基づいた現フレームの特定物体中心位置座標と特定物体中心速度ベクトルの確率分布である。

【0087】

【数25】

$$z_t = (u_t, v_t)^T \quad (53)$$

$$P(z_{t+1} | d_t) \quad (54)$$

$$P(z_t | d_t) \quad (55)$$

10

【0088】

・推定ステップ

次式 (56) を用いて、領域選択部で求めたデータ d_t と、前フレームで求めた確率分布から、現フレームの確率分布を計算する。

【0089】

【数26】

$$P(z_t | d_t) = \frac{P(d_t | z_t)P(z_t | d_{t-1})}{P(d_t)} \quad (56)$$

20

【0090】

ここで、 $P(d_t | z_t)$ は尤度と呼ばれ、システムの観測モデルを確率の形で表現したものである。例えば、観測モデルを次式 (57) で与える。

$$d_t = z_t + n_1 \quad (57)$$

【0091】

ここで、 n_1 は平均が 0 で、分散が σ_e の、物体領域の観測に伴う正規ノイズである。これは、例えば、手ブレに伴うカメラの振動ノイズ等に相当する。このとき、尤度は式 (58) で表現される。

$$P(d_t | z_t) = N(d_t | z_t, \sigma_e) \quad (58)$$

式 (56) の計算で求めた確率分布 $P(z_t | d_t)$ から、次式 (59) に示すように、特定物体領域位置座標と特定物体領域速度ベクトルの期待値を計算し、特定物体中心位置座標の推定値 \hat{u}_t ハットと、特定物体中心速度ベクトル \hat{v}_t ハットを求める (ステップ S440)。

30

40

【0092】

【数27】

$$\hat{z}_t = (\hat{u}_t, \hat{v}_t)^T = \int z_t P(z_t | d_t) dz_t \quad (59)$$

【0093】

画像生成部 8 は、この \hat{u}_t ハットが中心となるような画像を全方位画像から切り出すこ

50

とで、自動画角調整を実現する。

【0094】

・予測ステップ

次式(60)を用いて、推定ステップで求めた現フレームの確率分布 $P(z_t | d_t)$ から次フレーム用の確率分布 $P(z_{t+1} | d_t)$ を計算する。

【0095】

【数28】

$$P(z_{t+1} | d_t) = \int P(z_{t+1} | z_t) P(z_t | d_t) dz_t \quad (60)$$

10

【0096】

ここで、 $P(z_{t+1} | z_t)$ は状態遷移確率であり、物体の運動を確率分布で表現したものである。例えば、物体の運動を次式(61)でモデル化する。

$$z_{t+1} = A z_t + n_2 \quad (61)$$

ここで、 A は状態遷移行列であり、前出の式(52)と同一である。式(61)に示す n_2 は平均が0で、分散が σ_p の、物体の移動に伴う正規ノイズとする。このとき、状態遷移確率は次式(62)となる。

$$P(z_{t+1} | z_t) = N(z_{t+1} | A z_t, \sigma_p) \quad (62)$$

式(60)により、前フレームで求めた確率分布 $P(z_t | d_{t-1})$ から、次式(63)で表される期待値 z_t チルダを計算する。

【0097】

【数29】

$$\tilde{z}_t = (\tilde{u}_t, \tilde{v}_t)^T = \int z_t P(z_t | d_{t-1}) dz_t \quad (63)$$

【0098】

u_t チルダを特定物体中心位置座標の予測値とし、 v_t チルダを特定物体中心速度ベクトルの予測値とする。物体が遮蔽されると、画像生成部は、この予測値 u_t チルダが中心となるような画像を全方位画像から切り出すことで、自動画角調整を実現する(ステップS490)。

【0099】

再学習部7)

次に、現フレームの照明条件下における物体の色を再学習する再学習部7の動作を説明する。

物体色学習部1において学習した確率分布 $P_c(f | A)$ の再学習を行う手順を以下に示す。

【0100】

【数30】

$$(x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i) = 1 \quad (64)$$

40

【0101】

式(64)の楕円の1.5~2倍の領域からサンプリングを行い、画素のRGB値を、色相(H)と彩度(S)に変換する。

ここで i は、物体領域候補検出部2において、特定物体を表現している正規分布の番号であり、 μ_i はその正規分布の平均であり、 Σ_i はその正規分布の分散である。サンプリングを J 回繰り返し、式(65)の色相と彩度の2次元データ集合 D を得る。

【0102】

50

【数 3 1】

$$D = \{f_n\}_{n=1}^J \quad (65)$$

【0 1 0 3】

ここで、 f_n は n 番目にサンプリングされた画素の色相と彩度のベクトルである。

【0 1 0 4】

次に、確率分布 $P_c(f|A)$ を正規混合モデルによってモデル化し、上記式 (65) のデータ D を基に、物体色学習部 1 と同様の方法で式 (66) のパラメータ集合を求める。

【0 1 0 5】

【数 3 2】

$$\theta = \{\omega_i^c, \mu_i^c, \Sigma_i^c\}_{i=1}^K \quad (66)$$

【0 1 0 6】

このとき、前フレームで学習したパラメータ集合を初期値として使用する (初期フレームのみ、前記物体色学習部 1 により予め学習しておいたパラメータ集合を使用する)。以上のように、物体色の学習をフレーム毎に行うため、照明条件の変化等による物体色の变化に適応することができる。

【0 1 0 7】

実施の形態 2 .

本発明の実施の形態 2 では、全方位画像から複数の物体を検出し、その中からユーザによって選択された物体が中心となる画像を切り出す装置を説明する。

本発明の実施の形態 2 における画角調整装置を図 15 から図 17 を用いて説明する。

図 15 は、実施の形態 2 における画角調整装置の構成図であり、図 16 は、実施の形態 2 における画角調整装置の動作手順を示すフローチャートである。図 17 は、実施の形態 2 における動作を説明するための具体例を用いた説明図である。

【0 1 0 8】

実施の形態 2 における画角調整装置 70 は、予め用意した画像データベースから物体の色を学習する物体色学習部 61 と、キャプチャした全方位画像から物体色学習部により学習した物体が存在する領域の候補を求める物体領域候補検出部 62 と、物体領域候補検出部 62 により検出された複数の候補領域の中から、ユーザが指定した物体を選択する物体指定部 60 と、カメラの手ブレに伴う画像中の振動ノイズを除去し物体を追跡する物体追跡部 63 と、特定物体を中心とする画像を切り出す画像生成部 68 とを備える。

実施の形態 2 における画角調整装置 70 は、実施の形態 1 における画角調整装置 10 に、物体指定部 60 を加えた装置である。即ち、実施の形態 2 における物体色学習部 61 と、物体領域候補検出部 62 と、物体追跡部 63 と、画像生成部 68 は、それぞれ実施の形態 1 に記載の物体色学習部 1、物体領域候補検出部 2、物体追跡部 3、画像生成部 8 と同一機能の要素である。

【0 1 0 9】

次に、図 16 を用いてパノラマカメラによる画角調整装置 70 の動作を説明する。

本実施の形態における画角調整装置 70 の動作は、物体の色を画像データベースに基づき学習するフェイズ 1 と、フェイズ 1 で学習した色を持つ特定物体が常に画角に入るよう自動調整するフェイズ 2 の 2 段階に分かれる。

【0 1 1 0】

(フェイズ 1)

まず、物体色学習部 61 は、予め用意した画像データベース 75 を用いて、追跡すべき物体の色を学習する (ステップ S590)。学習した色は、キャプチャ画像から特定物体の候補領域を検出する際 (ステップ S610) の基準として利用する。

10

20

30

40

50

【0111】

(フェイズ2)

まず、パノラマカメラ74を使用して全方位画像をキャプチャする(ステップS600)

。続くステップS610では、物体領域候補検出部62において、キャプチャした全方位画像から、ステップS590で学習した物体が存在する候補領域を抽出する。一般に複数の候補領域が検出され抽出される。例えば、テレビ電話において人物を検出する場合には、通話者の他に、通話者の周りに存在する人物や、壁にかけられた人物のポスター等を同時に検出される場合があるが、ここではすべて人物の候補領域となる。

【0112】

続くステップS615では、物体指定部60において、まず、ユーザによる物体指定要求があればそれを受け取り、ステップS610で検出した複数の領域候補から、ユーザにより指定された領域を求め、その領域中心位置座標を取得する。

次に、物体指定部60は、前フレームのノイズ除去部65において推定した特定物体位置座標の値を、この取得した領域中心位置座標の値に置き換える。この処理により、追跡対象となる特定物体をユーザにより指定された物体に変更する。

もし、ユーザによる物体指定要求がない場合、物体指定部60は処理を行わずにステップS620に進む。

【0113】

続くステップS620では、領域選択部64において、ステップS610で検出した複数の特定物体の候補領域から、前フレームのノイズ除去部65において推定した特定物体位置座標に最も近い領域を求め、求めた領域の中心位置座標を取得する。

しかし、初期フレームにおいては、領域選択部64は、ステップS610で検出した複数の特定物体の候補領域の中から、予め決められた画像中のある位置(例えば、Xとする)に最も近い領域を求め、求めた領域の中心位置座標を取得する。例えば、テレビ電話において常に通話者が中心となる画像を切り出す場合には、始めに通話者が画像中のXの位置に映るように設定して、次に画角調整装置70の動作を開始する。

続くステップS630では、ノイズ除去部65において、ステップS620で取得した特定物体中心位置座標と前フレームで予測した特定物体中心位置座標に基づき、現フレームの特定物体中心位置座標を推定する。

【0114】

続くステップS640では、動き予測部66は、ステップS150で推定した特定物体中心位置を用いて、物体が等速直線運動するという仮定から次フレームにおける特定物体中心位置の予測を行う。この予測値は、次フレームでノイズ除去部65において物体位置を正確に求めるために使用する。

続くステップS650では、照明条件等により毎フレームに変化する物体の色に適應するために、再学習部67において特定物体の色を再学習する。

続くステップS660では、画像生成部68は、ステップS640で推定した特定物体中心位置を中心とする画像を切り出す。

最後に、ステップS600に戻り、上記の処理を繰り返す。

上記の処理をフレーム毎に繰り返すことで、映像中からユーザが指定した物体検出を行い、その物体が中心となる動画の生成を実現する。

【0115】

図17に、上記で説明した実施の形態2における具体例を示す。例えば、物体領域候補検出部62により求められた複数の顔領域候補のうち、現在画角の中心にいる人物を“C”、その左右に存在する人物を“L”、“R”とする(図17(a))。

例えば、図17に示すように、携帯電話(端末)の十字キー43を右側に押す操作で、ユーザが“R”の人物を指定すれば、物体指定部60は“R”の人物領域を指定し、物体追跡部63に信号を送る。続く物体追跡部63と画像生成部68の処理により、物体指定部60で指定した人物が中心となる画像が生成され、携帯電話の端末41に映し出される。

10

20

30

40

50

上記の例のように、本発明を携帯電話やPDAのテレビ電話において人物の検出に使用すれば、画像を受信しているユーザは、通話先にいる複数の人物の中から選択した人物が中心となる映像を自分の端末に取得することが可能となる。

【0116】

【発明の効果】

以上のようにこの発明によれば、物体色学習部と物体領域候補検出部と変動除去部と物体追跡部とを備えたので、手持ちカメラ等による瞬時変動が大きい画像入力に対しても、対象物体を安定して抽出し、画像表示できる効果がある。

【図面の簡単な説明】

【図1】この発明の実施の形態における画角調整装置の構成を示す図である。

10

【図2】実施の形態1における画角調整装置の動作を示すフロー図である。

【図3】実施の形態1における画角調整装置の具体的な動作を説明するための説明図である。

【図4】実施の形態1における物体色学習部の動作を示すフロー図である。

【図5】実施の形態1において肌色のサンプル部分における画素の色相と彩度を2次元化表現した図である。

【図6】実施の形態1における物体認識の概念を示す図である。

【図7】実施の形態1における物体領域候補検出部の動作を示すフロー図である。

【図8】実施の形態1において、各画素を正規混合モデルによって1次元の確率分布として近似する例を示す図である。

20

【図9】実施の形態1における物体領域候補検出部が全方位画像から肌色領域を検出する例を示す図である。

【図10】実施の形態1における物体領域候補検出部が行う正規分布生成の例を示す図である。

【図11】実施の形態1における物体領域候補検出部が行う正規分布消滅の例を示す図である。

【図12】実施の形態1における物体領域候補検出部が行う正規分布分割の例を示す図である。

【図13】実施の形態1における物体領域候補検出部が行う正規分布合併の例を示す図である。

30

【図14】実施の形態1における物体追跡部の動作を示すフロー図である。

【図15】この発明の実施の形態2における画角調整装置の構成を示す図である。

【図16】実施の形態2における画角調整装置の動作を示すフロー図である。

【図17】実施の形態2における画角調整装置の具体的な動作を説明するための説明図である。

【図18】第1の従来例としての画像監視装置の構成図である。

【図19】第1の従来例としての画像監視装置における物体移動検知部の動作フロー図である。

【符号の説明】

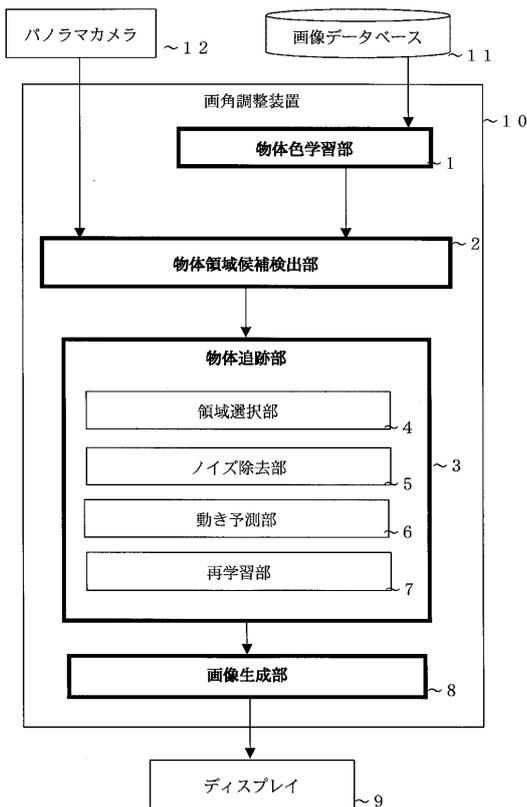
1 物体色学習部、2 物体領域候補検出部、3 物体追跡部、4 領域選択部、5 ノイズ除去部、6 動き予測部、7 再学習部、8 画像生成部、9 ディスプレイ、10 画角調整装置、11 画像データベース、12 パノラマカメラ、13 パノラマカメラ、14 全方位画像、15 中心になる画像、16 相手端末、17 パノラマカメラ、18 ディスプレイ、19 画像データベース、20 予め用意した画像、21 色相と彩度の空間、22 ガウス分布の輪郭、23 色相Hの一次元確率分布、24 顔、25 正規分布、26 パノラマ画像、27 肌色のサンプリング結果、28 サンプリングデータを正規混合分布で近似、29 各ガウス分布の輪郭と中心、30 新規データの発生、31 正規分布の生成、32 データ発生源の消滅、33 正規分布の消滅、34 2つのデータ発生源を1つの正規分布で近似している例、35 正規分布の分割、36 1つのデータ発生源を2つの正規分布で近似している例、37 正規分布の合併、38

40

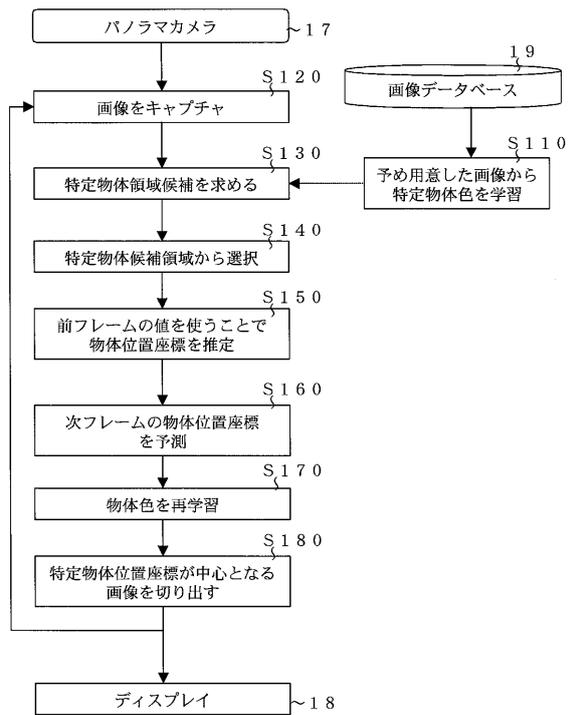
50

ラベル付けされたパノラマ画像、39 切り出された画像、40 ラベル付けされたパノラマ画像、41 携帯電話の端末、42 通話相手、43 携帯端末の十字キー、44 (1-)乗算部、45 加算部、46 乗算部、47 背景映像出力部、48 減算部、49 パノラマカメラ、50 A/D変換部、51 移動領域検知部、52 ラベリング部、53 移動領域追跡部、54 正像変換部、55 表示制御部、60 物体指定部、61 物体色学習部、62 物体領域候補検出部、63 物体追跡部、64 領域選択部、65 ノイズ除去部、66 動き予測部、67 再学習部、68 画像生成部、69 ディスプレイ、70 画角調整装置、71 画像データベース、72 パノラマカメラ。

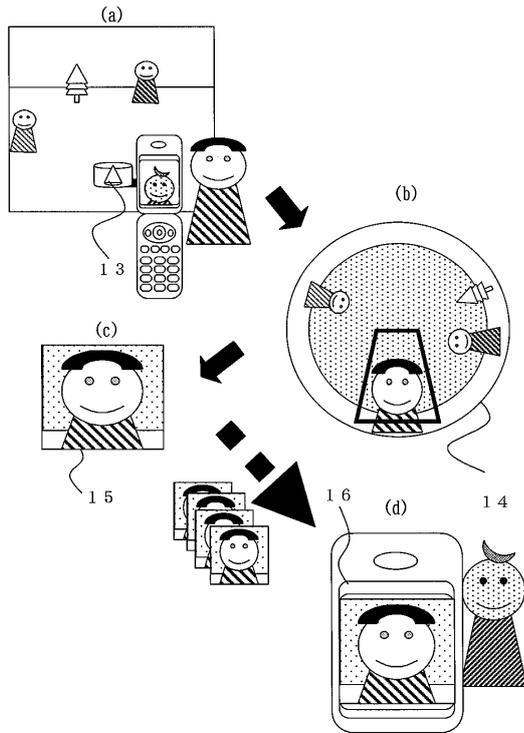
【図1】



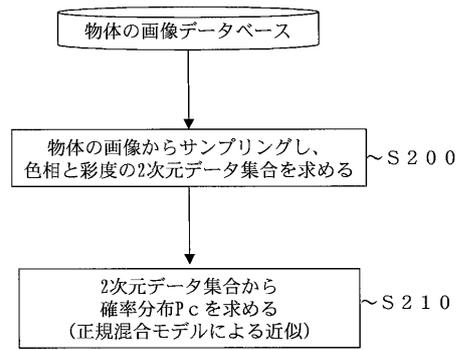
【図2】



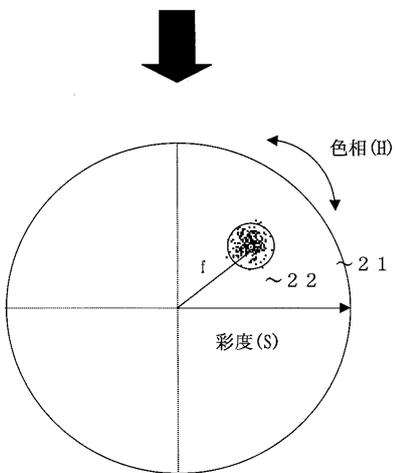
【 図 3 】



【 図 4 】

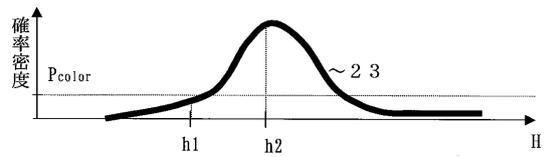


【 図 5 】

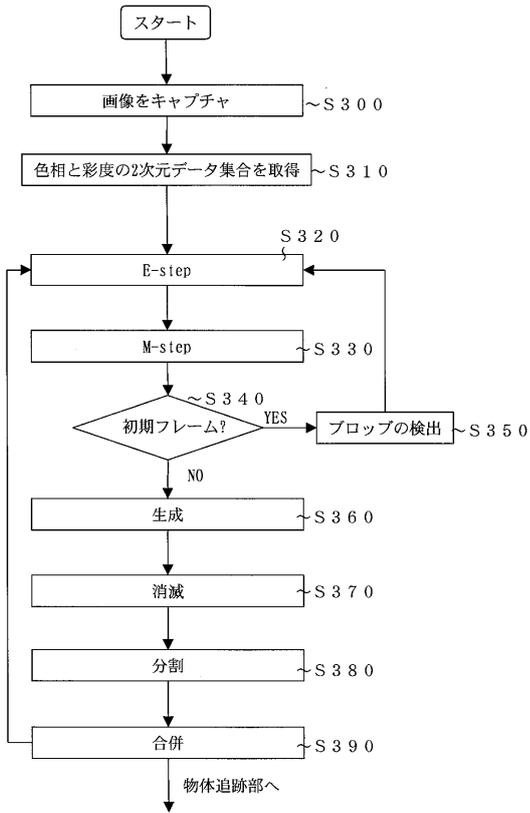


色相と彩度の空間

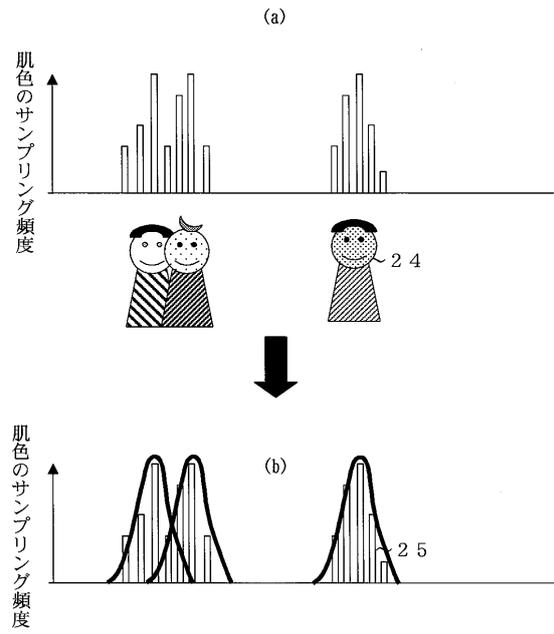
【 図 6 】



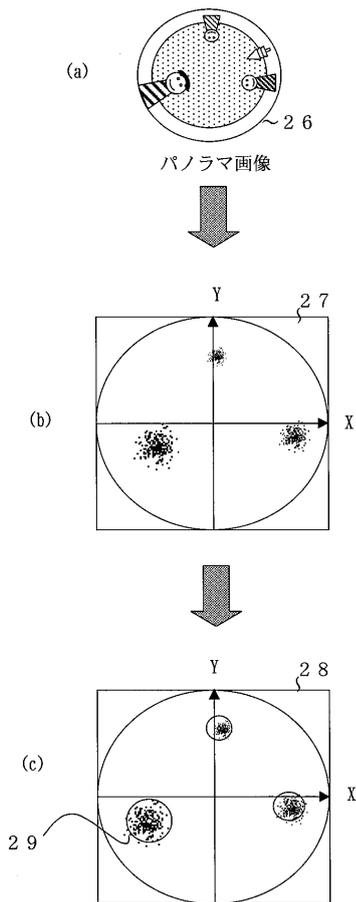
【 図 7 】



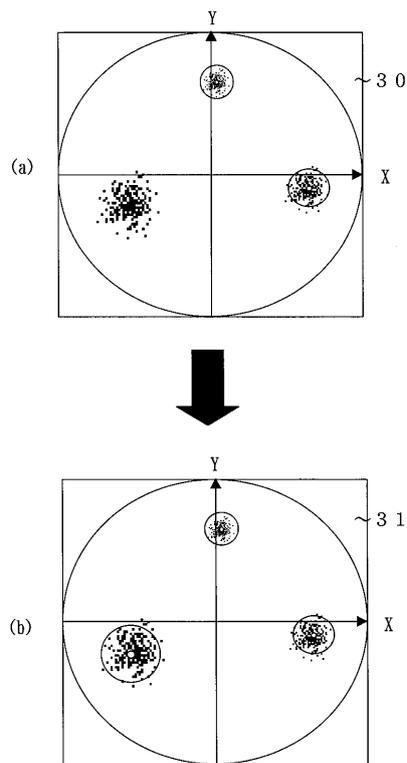
【 図 8 】



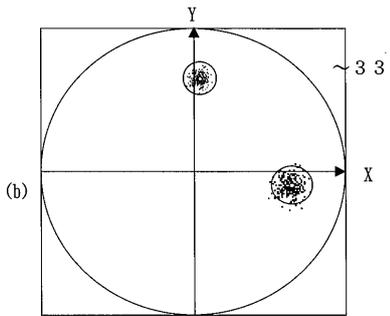
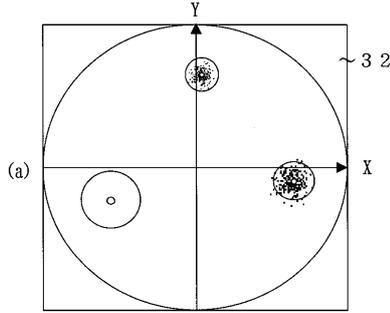
【 図 9 】



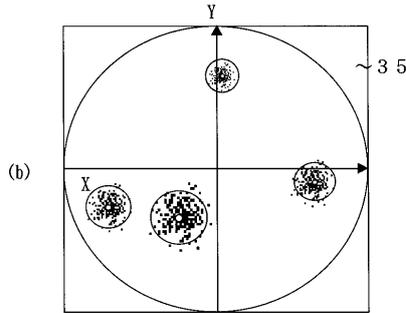
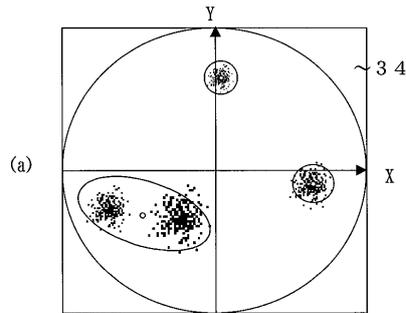
【 図 10 】



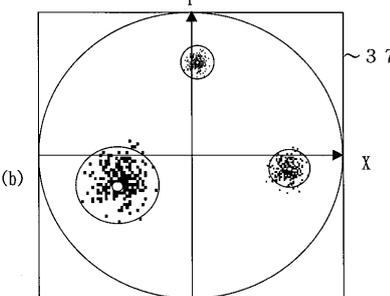
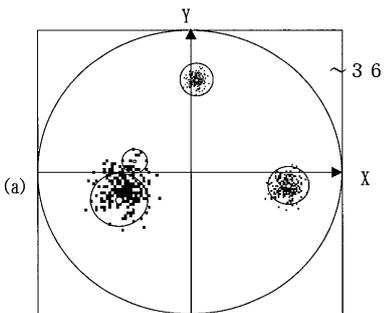
【図 1 1】



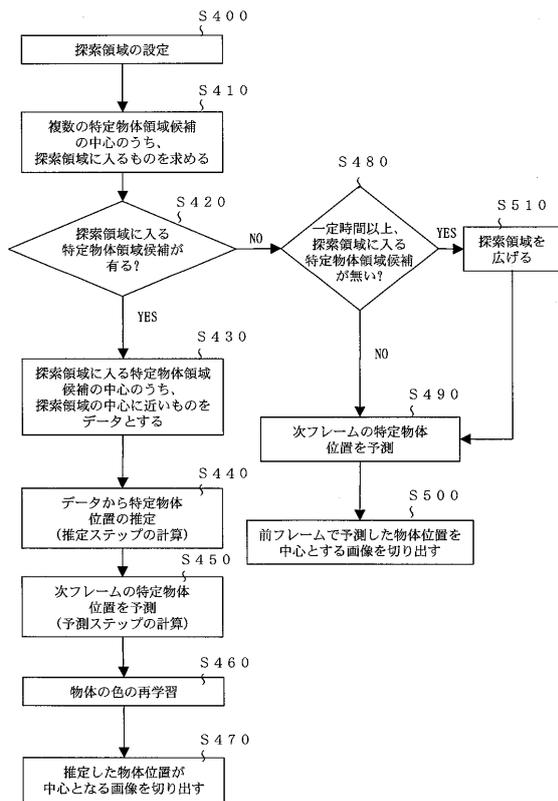
【図 1 2】



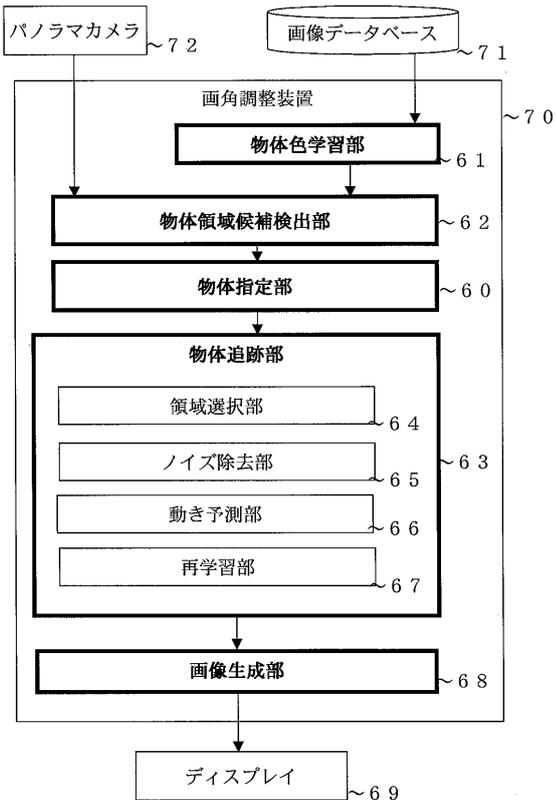
【図 1 3】



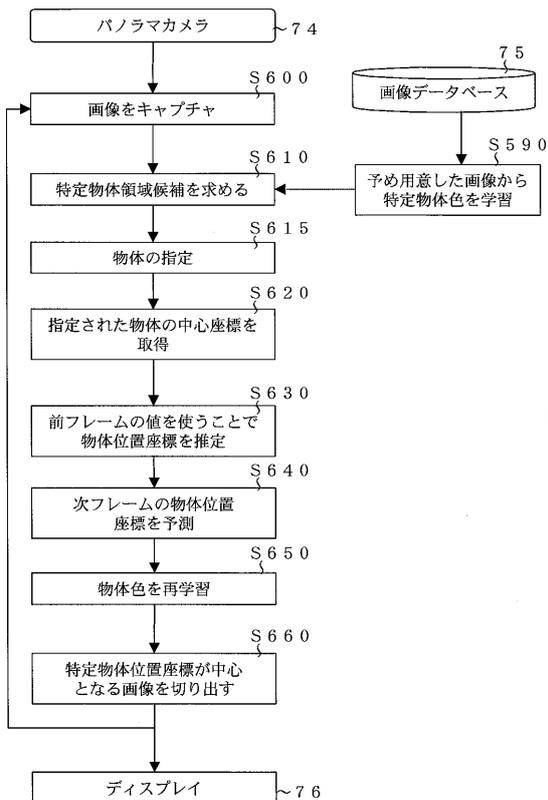
【図 1 4】



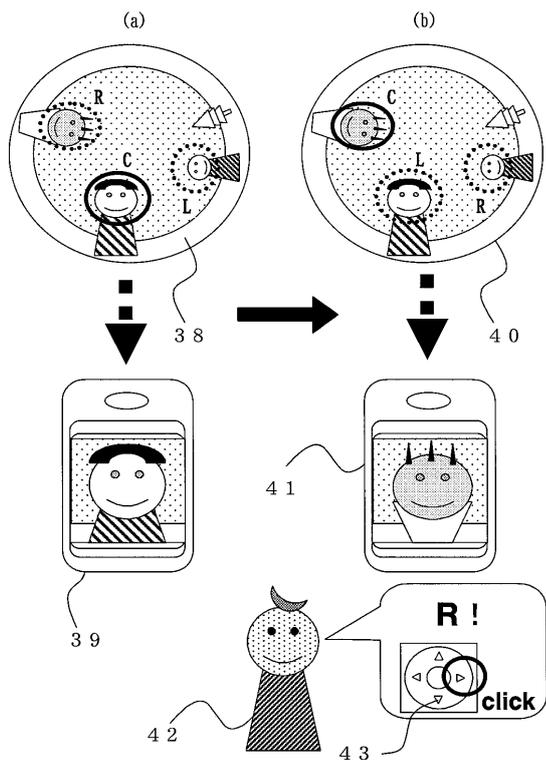
【図15】



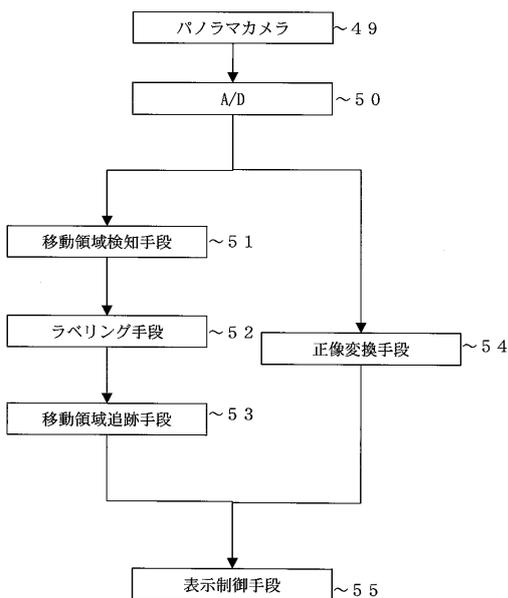
【図16】



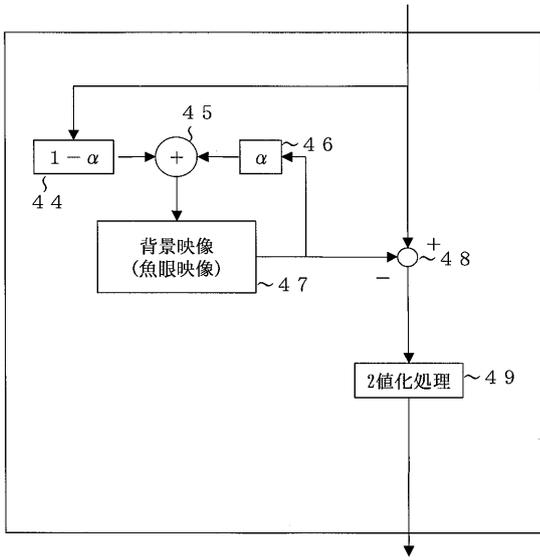
【図17】



【図18】



【 図 1 9 】



フロントページの続き

F ターム(参考) 5C022 AA01 AA12 AA13 AB36 AB62 AB63 AB66 AC27 AC69
5L096 AA02 CA04 DA05 EA05 FA04 FA15 FA32 FA33 FA59 GA34
GA41 HA04 HA05 KA04