

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3725996号
(P3725996)

(45) 発行日 平成17年12月14日(2005.12.14)

(24) 登録日 平成17年9月30日(2005.9.30)

(51) Int. Cl.⁷

F I

HO4N 7/32
 GO6T 7/20
 // HO4N 7/15
 HO4N 11/00
 HO4N 11/24

HO4N 7/137 Z
 GO6T 7/20 A
 HO4N 11/00
 HO4N 7/15 630Z

請求項の数 11 (全 23 頁)

(21) 出願番号 特願平11-163935
 (22) 出願日 平成11年6月10日(1999.6.10)
 (65) 公開番号 特開2000-354247(P2000-354247A)
 (43) 公開日 平成12年12月19日(2000.12.19)
 審査請求日 平成14年1月25日(2002.1.25)

(73) 特許権者 000005049
 シャープ株式会社
 大阪府大阪市阿倍野区長池町22番22号
 (74) 代理人 100064746
 弁理士 深見 久郎
 (72) 発明者 斎藤 修
 大阪府大阪市阿倍野区長池町22番22号
 シャープ株式会社内
 (72) 発明者 小田 守
 大阪府大阪市阿倍野区長池町22番22号
 シャープ株式会社内

審査官 松尾 俊介

最終頁に続く

(54) 【発明の名称】 画像処理装置

(57) 【特許請求の範囲】

【請求項1】

入力される動画像から人物の顔画像を抽出する画像処理装置であって、
 前記入力される動画像に基づいて動き物体領域からなる矩形領域を抽出する矩形領域抽出手段と、

前記矩形領域抽出手段によって抽出された矩形領域内の顔画像の特徴による精密座標を抽出する顔画像特徴抽出手段とを備え、

前記矩形領域抽出手段は、前フレームと現フレームの画像の差に基づいて矩形領域を抽出するとき、前フレーム矩形領域外の動き成分を判定するためのしきい値を矩形領域内の動き成分を判定するためのしきい値よりも感度を低く設定したことを特徴とする、画像処理装置。

【請求項2】

前記矩形領域抽出手段は、前記前フレームと現フレーム間の動き画素量が少ないときに、前フレームの画像により矩形領域を抽出して出力することを特徴とする、請求項1に記載の画像処理装置。

【請求項3】

前記顔画像特徴抽出手段は、前記矩形領域内の顔画像の特徴を抽出するために顔幅を判定するとき、顔幅検索範囲を頭頂座標から画面最下座標までの1/2の位置の矩形座標の左右幅に設定することを特徴とする、請求項1に記載の画像処理装置。

【請求項4】

10

20

前記顔画像特徴抽出手段は、顔幅を判定するとき、鼻成分を避けて実際の顔幅よりも狭い領域を精密座標として判定しないことを特徴とする、請求項 3 に記載の画像処理装置。

【請求項 5】

前記顔画像特徴抽出手段は、顔画像特徴抽出による精密座標抽出ができなかったとき、ある決められた大きさの領域を矩形領域とすることを特徴とする、請求項 1 に記載の画像処理装置。

【請求項 6】

前記顔画像特徴抽出手段は、前記顔画像特徴抽出による精密座標抽出ができなかったとき、画面全体を矩形領域とすることを特徴とする、請求項 1 に記載の画像処理装置。

【請求項 7】

さらに、前記矩形領域抽出手段によって現フレームで動き画素より算出された矩形領域座標と、前フレームの矩形領域座標の値にローパスフィルタをかけて動き物体が画面内で急激に移動した場合でも滑らかに矩形領域が動き物体に追従させるようにしたことを特徴とする、請求項 1 に記載の画像処理装置。

【請求項 8】

前記矩形領域抽出手段は、画面内の端の領域に有効な顔画像が存在しないとして求められた矩形領域座標値に対してクリッピング処理を行なうことを特徴とする、請求項 1 に記載の画像処理装置。

【請求項 9】

前記顔画像特徴抽出手段は、顔画像特徴抽出による精密座標判定のための顔幅を判定するとき、前フレームで求めた精密座標と現フレームで求めた精密座標の値の差があるしきい値内であれば前フレームで算出した精密座標を現フレームの精密座標とすることを特徴とする、請求項 1 に記載の画像処理装置。

【請求項 10】

前記顔画像特徴抽出手段は、顔画像特徴抽出による精密座標判定のための顔幅を判定するとき、現フレームで求めた精密座標の値と過去数フレームの精密座標の値の平均値の差があるしきい値内であれば過去数フレームの精密座標の値の平均値を現フレームの精密座標値とすることを特徴とする、請求項 1 に記載の画像処理装置。

【請求項 11】

前記顔画像特徴抽出手段によって抽出された精密座標内の色分布標準偏差に基づいて前記入力される動画像の肌色領域を前記人物の顔画像として抽出する肌色領域抽出手段をさらに備える、請求項 1 ~ 10 のいずれか 1 項に記載の画像処理装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

この発明は画像処理装置に関し、特に、動画像符号化技術を使用し、動画像符号化技術の中でも TV 電話などのような対象画像内に人物などの顔画像が含まれるような画像を処理する画像処理装置に関する。

【0002】

【従来の技術】

画像処理装置による顔画像特徴抽出技術としては、たとえば特開平 7 - 29014 号公報に記載されているように、動きオブジェクトを人物の領域として矩形領域を抽出し、抽出した矩形領域内から顔画像の特徴によりさらに顔領域の肌色領域を抽出するものがある。

【0003】

【発明が解決しようとする課題】

しかしながら、上述の特開平 7 - 29014 号公報に記載された発明では、動き要素判定による動き領域抽出というアルゴリズムから生じる、人物領域以外の動き成分を誤って抽出してしまうおそれがある。また、顔画像内の縦エッジ成分を顔の両幅とすることから、鼻領域の縦エッジを顔の右頬もしくは左頬として誤って認識してしまうこともある。さらに、ハードウェアで顔画像領域抽出技術を実現する場合でなく、プロセッサなどのソフト

10

20

30

40

50

ウェアで実現する場合の不要演算によりシステム全体のパフォーマンスが低下するなどの問題点があった。

【0004】

それゆえに、この発明の主たる目的は、動き画素による動き矩形領域抽出や矩形領域内の顔特徴抽出による顔の精密座標抽出技術において、抽出性能の向上と誤抽出の低減を図られるような画像処理装置を提供することである。

【0005】

【課題を解決するための手段】

請求項1に係る発明は、入力される動画像から人物の顔画像を抽出する画像処理装置であって、入力される動画像に基づいて動き物体領域からなる矩形領域を抽出する矩形領域抽出手段と、抽出された矩形領域内の顔画像の特徴による精密座標を抽出する顔画像特徴抽出手段とを備え、矩形領域抽出手段は前フレームと現フレームの画像の差に基づいて矩形領域を抽出するとき、前フレーム矩形領域外の動き成分を判定するためのしきい値を矩形領域内の動き成分を判定するためのしきい値よりも感度を低く設定することにより、矩形領域の判定精度を向上させる。

10

【0006】

請求項2に係る発明では、請求項1の矩形領域抽出手段は前フレームと現フレーム間の動き画素量が少ないときは前フレームの画像により矩形領域を抽出して出力することにより、対象人物領域外に矩形領域が移動しにくくすることで矩形領域の判定精度を向上させる。

20

【0007】

請求項3に係る発明では、請求項1の顔画像特徴抽出手段は、矩形領域内の顔画像の特徴を抽出するために顔幅を判定するとき、顔幅検索範囲を頭頂座標から画面最下座標までの1/2の位置の矩形座標の左右幅に設定することにより、精密座標検索の精度と演算の高効率化を図る。

【0008】

請求項4に係る発明では、請求項1の顔画像特徴抽出手段は顔幅を判定するとき、鼻成分を避けて実際の顔幅よりも狭い領域を精密座標として判定しないようにする。

【0009】

請求項5に係る発明では、請求項1の顔画像特徴抽出手段は顔画像特徴抽出による精密座標抽出が失敗した場合、ある決められた大きさの領域を矩形領域とすることにより、矩形領域外に顔画像が存在することを低減する。

30

【0010】

請求項6に係る発明では、顔画像特徴抽出による精密座標抽出が失敗した場合、画面全体を矩形領域とすることで矩形領域外に顔画像が存在することをなくす。

【0011】

請求項7に係る発明では、矩形領域抽出手段による現フレームで動き画素より算出された矩形領域座標と、前フレームの矩形領域座標の値にローパスフィルタをかけ、動き物体が画面内で急激に移動した場合でも滑らかに矩形領域が動き物体に追従するようにする。

【0012】

請求項8に係る発明では、画面内の端の領域には有効な顔画像が存在しないとして求められた矩形領域座標値に対してクリッピング処理を行なうことにより、後の精密座標抽出時の精度向上と演算量削減を行なう。

40

【0013】

請求項9に係る発明では、顔画像特徴抽出による精密座標判定のための顔幅判定時に、前フレームで求めた精密座標と現フレームで求めた精密座標の値の差があるしきい値内であれば前フレームで算出した精密座標を現フレームの精密座標として採用することにより、精密座標が見失われることを避ける。

【0014】

請求項10に係る発明では、顔画像特徴抽出による精密座標判定のための顔幅を判定す

50

るとき、現フレームで求めた精密座標の値と過去数フレームの精密座標の値の平均値の差があるしきい値内であれば、過去数フレームの精密座標の値の平均値を現フレームの精密座標値として採用することにより、精密座標が見失われるのを避ける。

請求項 1 1 に係る発明では、顔画像特徴抽出手段によって抽出された精密座標内の色分布標準偏差に基づいて入力される動画像の肌色領域を人物の顔画像として抽出する肌色領域抽出手段をさらに備え、顔画像特徴抽出手段によって抽出できなかった肌色領域をも含む顔画像を抽出できる。

【 0 0 1 5 】

【 発明の実施の形態 】

図 1 は H . 2 6 1 / H . 2 6 3 画像エンコーダと顔領域抽出部の関係を示すブロック図である。図 1 において、符号化器に入力された C I F 画像データは、顔領域抽出部 1 0 1 と減算器 1 0 3 とスイッチ 1 0 4 の一方の入力端 a と予測メモリ 1 1 2 とにそれぞれ与えられる。減算器 1 0 3 は入力された画像データから、ループ内フィルタ 1 1 1 の出力するビデオ信号を減算し、その差分データはスイッチ 1 0 4 の他方の入力端 b に与えられる。スイッチ 1 0 5 は 2 つの入力端 a と b とを有し、入力端 b にもループ内フィルタ 1 1 1 の出力が与えられる。スイッチ 1 0 5 はスイッチ 1 0 4 とともに符号化制御部 1 0 2 の切換制御信号によって同期して切換えられる。

10

【 0 0 1 6 】

変換器 1 0 7 はスイッチ 1 0 4 で切換えられたフレーム内のビデオ信号とフレーム間のビデオ信号のいずれかを D C T (Discrete Cosine Transform : 離散コサイン変換) し、その出力は量子化器 1 0 8 に与えられる。量子化器 1 0 8 は変換器 1 0 7 のデータを量子化し、量子化インデックス q を出力するとともに、この量子化インデックス q を逆量子化器 1 0 9 に出力する。

20

【 0 0 1 7 】

逆量子化器 1 0 9 は生成された量子化インデックス q を逆量子化し、逆変換器 1 1 0 に与える。逆変換器 1 1 0 は逆量子化器 1 0 9 で逆量子化されたデータを逆変換し、その出力を加算器 1 0 6 に与える。

【 0 0 1 8 】

加算器 1 0 6 はスイッチ 1 0 5 を介して与えられるループ内フィルタ 1 1 1 の出力である前フレームのビデオ信号に差分データを加算し、その出力を予測メモリ 1 1 2 に与える。予測メモリ 1 1 2 は数フレーム分の画像データを保持するとともに、前フレームの画像に対する各ブロックの画像の動きを動きベクトル v として出力する。ループ内フィルタ 1 1 1 は予測メモリ 1 1 2 に保持された画像における歪みをスムージングにより除去するフィルタであり、その動作の有無を示すオン / オフ信号 f を出力する。

30

【 0 0 1 9 】

符号化制御部 1 0 2 はスイッチ 1 0 4 , 1 0 5 に切換制御信号を出力するとともに、量子化器 1 0 8 に対して量子化特性指定情報 q z を指示し、フレーム間 / フレーム内符号化識別フラグ p と伝送 / 非伝送識別フラグ t をそれぞれ発生する。顔領域抽出部 1 0 1 では、フレーム内の画像データから顔領域の特徴となる領域を抽出し、座標を符号化制御部 1 0 2 に出力する。

40

【 0 0 2 0 】

以下に、顔領域の判定方法について説明する。顔領域抽出は大きく分けて物体の動きに着目した動きベースによる顔抽出と、顔の色からの色ベースによる肌色抽出の 2 つの処理で構成される。

【 0 0 2 1 】

図 2 は顔領域抽出の大まかな流れを示すフローチャートである。図 2 において、まず始めにステップ (図示では S P と略称する) S P 1 において、動きベースによる顔領域で抽出する Still Count を 0 に初期化する。また、PreNoMv を 1 0 0 0 に初期化する。ステップ S P 2 において画像フレームが取込まれ、ステップ S P 3 で顔・肌色抽出処理の第 1 次段階として、動きを利用して顔領域が抽出される。ここで得られた顔領域の情報が次の処理

50

の色ベースによる肌色抽出で使用される。そして、ステップSP4で肌色の抽出の実施であるか否かが判別され、そうであればステップSP5で色ベースによる肌色が抽出され、ステップSP6でプリフィルタ処理がされた後、ステップSP7で1フレームの画像符号化が行なわれる。

【0022】

次に、動きベースによる顔領域の位置をもとに、動きベースによる顔抽出において抽出できなかった顔以外の肌色領域、たとえば、手、腕、首などの領域を動き物体の中から抽出する動作について説明する。

【0023】

ここで、動きベースによる顔領域抽出の具体的な動作について説明すると、抽出対象となる顔領域は動き物体であり、動き領域の頂上が頭頂であるという特徴を用いて顔領域の抽出が行なわれる。

【0024】

図3および図4は動きベースによる顔抽出の動作を示すフローチャートであり、図5は縮小画像の作成を示す図である。

【0025】

はじめに、処理の高速化および細かな動きを除去して有効な動きを得るために、図3のステップSP11において現フレームおよび前フレームの入力画像の輝度成分からなるY画像を縮小する。縮小画像の大きさは入力画像の大きにかかわらず、表1に示すように44×36画素である。

【0026】

【表1】

入力画像の大きさ	縮小後の大きさ	縮小率
CIF (352×288)	44×36	縦、横 1/8
QCIF (176×144)	44×36	縦、横 1/4
SQCIF (128×96)	32×24	縦、横 1/4

【0027】

図5に示すように、入力画像がCIFのときは8×8画素の平均、QCIFおよびSQCIFのときは4×4画素の平均を求め、縮小画像の1画素とする。ただし、SQCIFの場合は縮小後の大きさが32×24画素であるので、その外側に相当する部分に0を入れて縮小画像の大きさを44×36画素にする。現在のフレームの1つ前のフレームを前フレームとすることで、前フレームの縮小画像とする。すなわち、現フレームの縮小画像と前フレームの縮小画像はダブルバッファ構成でピンポン動作する。

【0028】

次に、前フレームと現フレームの縮小画像の差を検出することで、動き画像を抽出する。雑音による動き画素の抽出を防ぐために、過去の数フレームの動き画素の履歴を参照し、あるしきい値上であったとき有意と判断し、その画素を動き画素として抽出する。

【0029】

図6は図3のステップSP12に示すオブジェクトマスク作成の動作をより具体的に示すフローチャートであり、図7は動き履歴のアップデートを示す図であり、図8は動き画素判定しきい値適応領域を示す図であり、図9は3×3画素のウィンドウによる拡大処理を示す図であり、図10は3×3画素のウィンドウによる縮小処理を示す図である。

【0030】

次に、図7～図10を参照して、オブジェクトマスクの作成について説明する。図6のステップSP41において、前フレームの縮小画素の輝度値を $PrevY_{i,j}$ とし、現フレームの縮小画像の輝度値を $CurrY_{i,j}$ としたとき、次の第(1)式のようにそれぞれの画素の差 $ABS_{i,j}$ を求める。

【0031】

$$ABS_{i,j} = | PrevY_{i,j} - CurrY_{i,j} | \dots (1)$$

次に、ステップSP42において、縮小画像における画素の位置 i, j に対する各Nフレームの動きの履歴 $HIS_{i,j}$ をアップデートする。 $HIS_{i,j}$ が過去の1フレーム単位で1ビットごとに記憶されているとしてアップデート前の $HIS_{i,j}$ を図7(a)とする。このとき、現フレームTに最も時間的に近い過去のフレームは $T-1$ であり、Nフレームの過去のフレームは $T-N$ となる。ここで、動き画素を1とし、静止画素を0とすると、 $T-N$ フレームから $T-1$ フレームまでの履歴は順に、“1, 0, 1, ..., 0, 1, 1”となる。

10

【0032】

これに対して、アップデートを行なうことは左へ1ビットシフトすることであり、その結果図7(b)に示すようにこれから処理を行なう現フレームのビット位置Tに新しい動き情報が入力できるように空きができ、過去の履歴は $T-N$ フレーム目の履歴画素であり、Tフレームから $T-N+1$ フレームまでの“0, 1, 0, ..., 1, 1, X”のビット列となる。なお、この動きの履歴は各画素を8ビットとする。

【0033】

このようにして得られた $ABS_{i,j}$ をステップSP43でそのしきい値 TH_y によりその画素が動き画素か、動き画素でない(静止画素)かを判定する。

20

【0034】

この発明の実施形態では、画面内の動き画素に、符号化すべき最も重要な要素(すなわち話者)があるとしている。よって、話者の背後で、話者とは別な動き対象があった場合は、話者とは別な対象にトラッキングしてしまう。

【0035】

これを避けるために、以下の2つの操作が行なわれる。

すなわち、図8に示すように、画面の周囲には通話者がいることはないとし、画面の周囲のオブジェクトマスクを作成しない(上下左右とも、約1/10の領域)。

30

【0036】

動き判定しきい値を2種設ける。そして、話者の含まれる領域には感度の良いしきい値を適用し、それ以外の領域は感度を落としたしきい値を適用する。また、話者の含まれる領域とは、前フレームにて抽出された矩形領域をX方向に $\exp \text{const } X$, Y方向に $\exp \text{const } Y$ だけ拡げた領域とする。ここで、 $\exp \text{const } X$, $\exp \text{const } Y$ は例として、それぞれ $\exp \text{const } X=4$, $\exp \text{const } Y=8$ とする。すなわち、次の第(2)式を満足するときは、ステップSP44で $CurrY_{i,j}$ を動き画素とし、満足しない場合はステップSP45で静止画素とする。

【0037】

$$\begin{aligned} ABS_{i,j} > TH_y \text{ sense} & \text{ (前フレーム矩形領域内画素)} \\ ABS_{i,j} > TH_y \text{ insense} & \text{ (前フレーム矩形領域外画素)} \\ \dots & (2) \end{aligned}$$

40

なお、 $TH_y \text{ sense}$, $TH_y \text{ insense}$ は例として、実際のフレームレートが7より小さいときは $TH_y \text{ sense}$ を3とし、 $TH_y \text{ insense}$ を12とし、7以上のときは $TH_y \text{ sense}$ を5とし、 $TH_y \text{ insense}$ を20とする。

【0038】

判定された結果は、動きの履歴の現フレームの位置(LSBの位置、図7(b)のXで示す部分)に1あるいは0を書込む。

【0039】

次に、 $CurrY_{i,j}$ の画素が過去Nフレームにおいて動き画素と判定された数 $MvFrame_{i,j}$

50

を求めるために $HIS_{i,j}$ の N フレーム分の和を算出する。これは図 7 (b) において T から $T - N + 1$ までの $HIS_{i,j}$ の位置の合計を算出して $MvFrame_{i,j}$ とする。ここで、 N は 8 である。

【 0 0 4 0 】

次に、過去 N フレームにおける $MvFrame_{i,j}$ の値からオブジェクト画素抽出のしきい値 TH_{obj} により、

$$MvFrame_{i,j} > TH_{obj} \quad \dots (3)$$

第 (3) 式を満足する場合は、有意な動きがあると判定してオブジェクト画素として抽出する。なお、オブジェクト画素には、ステップ $SP48$ で 1 を与え、それ以外にはステップ $SP49$ で 0 を与える。また、 TH_{obj} は例として 3 とする。ステップ $SP50$ で 1 フレーム分の処理が終了したことを判別すると、図 3 のステップ $SP13$ に戻り、オブジェクトマスク抽出において抽出されたオブジェクト画素の数をカウントして $NoMv$ とする。

10

【 0 0 4 1 】

たとえば、 TV 電話などの利用形態において、顔領域は通常は画面のほぼ中央に位置し、 TV 電話の利用中には大きく位置が変化することはない。よって、画面全体の中で画素の動き量 $NoMv$ が小さくなった場合には、顔領域抽出処理のオブジェクトマスクの拡大処理以降のラベル付けやエッジ抽出処理などといった演算をスキップし、矩形領域座標、オブジェクトマスク、エッジ画像を前回のフレームで求めた値を使うことで演算量の大幅な削減ができ、フレームレートの向上を図ることができる。ただし、最初の 10 フレーム目までは顔位置が完全に定まっていなるとし、スキップを行なわない。すなわち、ステップ $SP14$ で次の第 (4) 式を満足することを判別した場合は、オブジェクトマスク拡大処理以降をスキップする。

20

【 0 0 4 2 】

$$NoMv < TH_{judge} \text{ and } FrameCount > 10 \quad \dots (4)$$

ここで、 TH_{judge} は例として 100 とする。

【 0 0 4 3 】

また、矩形座標抽出処理以降の精密座標抽出、FaceMap 作成、HueMap 作成処理のため、スキップ演算判定フラグ “ Bypass flag ” を設ける。すなわち、

演算スキップ判別されたとき Bypass flag = 1

演算スキップ判定されなかったとき Bypass flag = 0

30

... (5)

ステップ $SP15$ で Bypass flag が 1 の場合、矩形座標抽出処理、精密座標抽出、FaceMap 作成、HueMap 作成処理では、前フレームで求めた値を返り値として使用する。

【 0 0 4 4 】

ステップ $SP16$ で Bypass flag = 0 を 0 に設定し、オブジェクトマスクには、ホールや欠けが存在するため、それらを埋める処理として、ステップ $SP17$ において 3×3 画素のウィンドウを用いた拡大を実施する。これは、図 9 に示すように、オブジェクトマスクの各画素に 3×3 画素のウィンドウを設定して中心の画素がオブジェクト画素であるときにその 8 近傍をすべてオブジェクト画素とすることにより拡大が行なわれる。

【 0 0 4 5 】

40

また、SubQCIF の縮小画像は 32×24 であり、CIF / QCIF の縮小画像 (44×36) よりも小さい。このため、SubQCIF のオブジェクト画素 1 画素が実画像で対応する領域は、CIF / QCIF のオブジェクト画素の 1 画素が実画像で対応する領域よりも大きくなる。よって、SubQCIF の場合だけ、拡大処理後に 3×3 画素ウィンドウを用いた縮小を実施する。

【 0 0 4 6 】

これは、図 10 に示すように、オブジェクトマスクの各画素に 3×3 画素のウィンドウを設定して、中心の画素がオブジェクト画素でなかったときに、その 8 近傍すべてをオブジェクト画素としないことから縮小を行なう。

【 0 0 4 7 】

50

この発明の一実施形態のアルゴリズムでは、動き物体を顔領域と定義しているため、顔が静止したときにはこれらの領域の抽出ができなくなる。したがって、その場合には前フレームの領域を顔領域として使用することで動きがなくなった場合でも顔領域の消失を防ぐようにする必要がある。そのために、前フレームの動き量と現フレームの動き量との関係から動きしきい値 TH_{mv} を算出する。

【0048】

図11は図3に示すステップSP18の TH_{mv} 算出の動作をより具体的に示すフローチャートであり、図12は最大領域の抽出を示す図であり、図13は頭頂の検出を示す図である。

【0049】

TH_{mv} からオブジェクトマスクの論理和処理判定として動きが少なくなったときのオブジェクトマスクの消失を防ぐために、 $NoMv$ の値から次の第(6)式によってオブジェクトマスクの論理和を判定する。

【0050】

$$NoMv < TH_{mv} \dots (6)$$

ステップSP19において第(6)式を満足するとき、ステップSP20で前フレームのオブジェクトマスクと現フレームのオブジェクトマスクの論理和をとり、これを新たな現フレームのオブジェクトマスクとする。第(6)式を満足しない場合は、論理和処理を行わず、次に示すオブジェクトマスクの抽出処理に進む。

【0051】

画面には多くの動物が存在する場合があります、その中から抽出の対象だけを選択する必要があります。たとえば、テレビ電話などにおいては、対象者がカメラに最も近い位置に立つ確率が高いので、ここでは最も大きい連結領域をオブジェクトマスクから切り離す。

【0052】

この発明の一実施形態では、図12に示すように、縦方向に4画素ごとに区切ったスリット内における画素の合計のx軸への投影を行なってラベル付けを実施することにより、ステップSP21で最大領域を抽出する。さらに、最大領域の左端、右端のスリットにおいて左端のスリットのときはそのスリットの左端の座標を求め、右端のスリットのときはそのスリットの右端の座標をそれぞれ Xa 、 Xb として求めておく。このとき、最大領域の画素数が同じものが複数あった場合は、一番左端のものを最大領域とする。

【0053】

現フレームのオブジェクトマスクに存在するホールや欠けを埋めるために、ステップSP22で 3×3 画素のウィンドウによる拡大処理を実施する。拡大の方法は前述の図9と同様である。なお、拡大前のオブジェクトマスクは次のフレームでの処理に利用するため、拡大されたオブジェクトマスクを別のエリアに書込む。したがって、拡大前のオブジェクトマスクが次フレーム処理で前フレームのオブジェクトマスクとなる。そのために、現フレームのオブジェクトマスクと前フレームのオブジェクトマスクは、縮小画像と同様にダブルバッファ構成でピンポン動作となる。

【0054】

前述のごとく求めた最大領域の左端、右端の座標 Xa 、 Xb の間においてステップSP23で頭頂を検出する。図13において頭頂がオブジェクトマスクの最上部なので、まず Xa と Xb に挟まれるオブジェクトマスクの横方向の4画素ごとの画素の合計をY軸方向に投影する。次いで、投影した画像を上から調べ、最初にしきい値 TH_{top} 以上となるスリットにおいて、そのスリットの左上から右下へ走査し、最初に検出されるY座標を頭頂のY座標 $HeadTopY$ とする。さらに、 $HeadTopY$ において、左側から走査して最初に検出されるオブジェクトのX座標を $X1$ とする。同様に、 $HeadTopY$ において、右側から走査して最初に検出されるオブジェクトの座標を $X2$ とする。そして、 $X1$ と $X2$ の中心を頭頂のX座標 $HeadTopX$ とする。なお、 TH_{top} は、ここでは4とする。

【0055】

図14は図4のステップSP24における顔幅検出の具体的な動作を示すフローチャート

10

20

30

40

50

であり、図 1 5 は顔幅検出範囲の決定を示す図であり、図 1 6 は顔幅の検出を示す図である。

【 0 0 5 6 】

次に、図 1 4 ~ 図 1 6 を参照して、顔幅の検出処理について説明する。顔領域は通常、画面のほぼ中央に位置し、画面の下方には肩の領域がある。これを利用し、ステップ S P 7 1 で顔領域の検出範囲を定める。

【 0 0 5 7 】

頭頂位置HeadTopYから画面の下までの領域の 1 / 2 のラインに注目し、このラインを検出ラインとする。すなわち、検出ラインの Y 座標sLineYは以下ようになる。

【 0 0 5 8 】

$$sLineY = HeadTopY + (Ymax - HeadTopY) / 2$$

Ymax = 35 : CIF, QCIF

Ymax = 23 : SubQCIF

... (7)

検出ラインの左から右へオブジェクト画素を検索する。最初にオブジェクト画素が見つかった X 座標を左側の検出範囲の最大maxXa とする。同様に、検出ラインの右から左へオブジェクト画素を検索し、最初にオブジェクト画素が見つかった X 座標を右側の検出範囲の最大maxXb とする。

【 0 0 5 9 】

ここで、求めたmaxXa が第 (8) 式を満たすときに、maxXa を左側の最大検出範囲とする。これを満たさないときは、Xaを左側の最大検出範囲とする。

【 0 0 6 0 】

$$Xa < maxXa < HeadTopX \quad \dots (8)$$

また、同様にmaxXb が第 (9) 式を満たすときに、図 1 5 に示すように、maxXb を右側の最大検出範囲とする。これを満たさないときは、Xbを右側の最大検出範囲とする。

【 0 0 6 1 】

$$HeadTopX < maxXb < Xb \quad \dots (9)$$

次に、ステップ S P 7 2 において画面を 4 画素ごとに縦方向のスリットに分割し、図 1 6 に示すようにHeadTopXを含むエリアを求める。図 1 6 ではエリア b のHeadTopXが含まれている。次に、ステップ S P 7 3 および S P 7 4 においてエリア b から左方向に隣接するエリア a との画素数の合計の差を求める。この図 1 6 では、エリア b の画素数が 1 2 0 であり、エリア a の画素数が 8 0 である。したがって、その差PelSubは、第 (1 0) 式のようになる。

【 0 0 6 2 】

$$PelSub = 120 - 80 = 40 \quad \dots (1 0)$$

ここで、ステップ S P 7 5 で画素の差のしきい値THsub と比較して第 (1 1) 式を満足したとき、ステップ S P 7 6 および S P 7 7 の処理により左の顔幅の座標FaceWideX1とする。

【 0 0 6 3 】

$$PelSub > THsub \quad \dots (1 1)$$

この例では、エリア b とエリア a との画素数の差が第 (1 1) 式を満足するので、比較エリアであるエリア a の左端の座標がFaceWideX1となる。

【 0 0 6 4 】

左の顔幅検出終了後、右方向についても同様にステップ S P 7 9 ~ S P 8 3 の処理を行ない、右の顔幅の座標FaceWideXrを得る。

【 0 0 6 5 】

図 1 7 は顔幅補正の流れを示すフローチャートである。次に、図 1 7 を参照して、顔幅の補正処理について説明する。上述のごとく検出された顔幅は、雑音や動き不十分により検出を失敗している可能性がある。そのため、顔幅が小さ過ぎたり、顔幅がHeadTopXを基準として偏り過ぎたりしている場合がある。それを補正するために、FaceWideX1, FaceWide

10

20

30

40

50

Xrを図17のフローチャートのステップSP91～SP101に従って処理を行なう。

【0066】

次に、顔の下座標を頭頂座標HeadTopY, 顔幅の座標FaceWideX1, FaceWideXrから顔の下の座標FaceBottomYを次の第(12)式から求める。

【0067】

$FaceBottomY = HeadTopY + (FaceWideXr - FaceWideX1) * 1.5 \dots (12)$

さらに、これらの値から顔領域を表わす矩形の座標を顔矩形領域として、左上および右下の座標を

$(FaceWideX1, HeadTopY), (FaceWideXr, FaceBottomY)$

と定義する。

10

【0068】

以上のようにして、動きベースによる矩形領域が求められる。

次に、ステップSP25およびSP26で矩形領域内の顔領域抽出による精密座標の検出処理を行なう。抽出対象となる顔領域は、頬部分の縦線が強い、目のまわりの横線に強いという特徴を使い、顔領域の抽出を行なう。ただし、処理のスキップ判定がされていた場合(Bypass flag = 1)は、この処理を行わずに前フレームで求めた精密座標を現フレームの精密座標とする。

【0069】

図18は精密座標の検出処理の動作を示すフローチャートである。

人間の顔において頬は縦線が強く、目, 鼻, 口は横線が強いという特徴がある。この性質を利用して顔の判定が可能となる。したがって、図18のステップSP111～SP120の処理により縦および横エッジを顔矩形領域内のオブジェクトマスク内に存在する現画像のYの画素から抽出する。その際、顔矩形領域は縮小画像の大きさであるため、それぞれの画像フォーマットの大きさに対応させて使用する。ただし、処理量を軽減するという観点からCIFの場合は縦横1画素おきの処理とする。

20

【0070】

図19はエッジの検出動作を示す図であり、図20は顔幅精密座標検出動作を示す図であり、図21は横エッジの特徴量を示す図である。

【0071】

まず、顔矩形領域内においてオブジェクトマスクに存在する画素を注目画素として、図20に示すように3×3画素のウィンドウを設定する。このとき、縦エッジUX, 横エッジUYは第(13)式で表わされる。

30

【0072】

$DX = |C+2F+I-A-2D-G|$

$DY = |G+2H+I-A-2B-C|$

$UX = \text{fix}(DX-DY)$

$UY = \text{fix}(DY-DX)$

ただし、 $\text{fix}(a) = a (a \geq 0), 0 (a < 0)$

... (13)

さらに、UX, UYをしきい値THedgeで2値化し、縦エッジ画素VEdge, 横エッジ画素HEdgeを第(14)式によって得る。なお、THedgeは例として60とする。

40

【0073】

$UX > THedge \quad VEdge = 1$

$UX \leq THedge \quad VEdge = 0$

$UY > THedge \quad HEdge = 1$

$UY \leq THedge \quad HEdge = 0$

... (14)

まず、縦・横エッジ画像における顔幅の長さXlenを顔幅の座標FaceWideX1, FaceWideXrから第(15)式で求める。

【0074】

50

$Xlen = (FaceWideXr - FaceWideXl + 1) * 4 \quad \dots (15)$

次に、Xlenの正方形を縦エッジ画像の上部に設定して、図19に示すようにこの内部を縦エッジの探索領域とする。このとき、中心付近には鼻が存在するとし、鼻の幅を以下の式で求める。

【0075】

$NoseWidth = Xlen/4 \quad \dots (16)$

また、エッジ画像の座標系である頭頂HeadCenterを、矩形座標の頭頂HeadTopYから第(17)式により求める。

【0076】

$HeadCenter = HeadTopX \times 4 \quad \dots (17)$

そして、頭頂位置HeadCenterからNoseWidthの分だけ離れた左の位置を探索開始位置として、ここから左の領域を順次探索し、縦方向の画素の累積値が最初にしきい値THpeakより大きくなる位置をPeakLとする。

【0077】

また、同様に、頭頂位置HeadCenterからNoseWidthの分だけ離れた右の位置を探索開始位置として、ここから右の領域を順次探索し、縦方向の画素の累積値が最初にしきい値THpeakより大きくなる位置をPeakRとする。

【0078】

ここで、THpeakは次の第(18)式より求めることができる。

$THpeak = Xlen \cdot PeakRathio/10$

ただし、

$PeakRathio = 2 \quad \dots (18)$

PeakL, PeakRがともにTHpeak以上であり、かつPeakLとPeakRのX座標の値がemnTHsub以上であったとき、顔幅が存在したと判断して精密座標抽出成功とし、後述する横エッジからの特徴抽出を行なう。また、PeakL, PeakRのX座標をそれぞれEMNI, EMNrとする。

【0079】

もし、上記の条件を満足しない場合は、顔幅精密座標抽出失敗と判断して、前フレームで求めたEMNI, EMNrを現フレームのEMNI, EMNrとする。前フレームで求めたEMNI, EMNrは現フレームの矩形座標に対して大きく外れている場合がある。この場合は、前フレームのEMNI, EMNrは不適切と判断する。現フレームの矩形座標から大きく外れているかどうかの判定には、現フレームの矩形領域の幅の左右parm x%の領域に前フレームのEMNI, EMNrが入っているかを判定する。すなわち、前フレームのEMNI, EMNrをそれぞれpreEMNI, preEMNrとしたとき、第(19)式を判定する。

【0080】

$preEMNI > FaceWideXl \times 4 - (FaceWideXr - FaceWideXl) \times 4 \times parm_x \div 100$

かつ

$preEMNr > FaceWideXr \times 4 + (FaceWideXr - FaceWideXl) \times 4 \times parm_x \div 100$

$\dots (19)$

第(19)式が満たされるときは、preEMNI, preEMNrを現フレームのEMNI, EMNrとし、満たされない場合は、精密座標抽出失敗とし、後述する顔検出の失敗処理を行なう。ここで、parm xは例として12とする。

【0081】

人間の顔は顔幅の間に横線特徴をもつ目・鼻・口が存在することを利用し、顔幅間で横エッジの分布を調べることで顔らしさの判断を行なう。図21のように太線で示す横エッジ探索領域を横エッジ画像の大きさに対応させた顔矩形領域の下部に設定し、この探索領域内においてEMNIとEMNrを3等分した領域を決定する。各領域ごとの横エッジの画素数の累積を求め、それぞれSY0, SY1, SY2とする。

【0082】

SY0, SY1, SY2の関係が第(20)式を満足していたとき顔領域と判定し、縦方向の精密座標検出処理を行なう。満足しない場合は後述する顔検出の失敗処理を行なう。

10

20

30

40

50

【 0 0 8 3 】

SY0 < SY1 かつ SY2 < SY1 ... (2 0)

顔領域の判定に成功した場合は、縦方向の精密座標の検出処理を行なう。図 2 2 に示すように横エッジ画素を Y 軸方向に累積したときの分布を求め、上方から走査したとき初めてしきい値 THemn 以上となる Y 座標を顔面の上部の座標 EMNtop とする。また、下方から走査して初めて THemn 以上となる Y 座標を顔面の下部の座標 EMNbottom とする。

【 0 0 8 4 】

ここで、もしも EMNtop と EMNbottom の差が emnTHsub より小さかった場合は、精密座標抽出失敗とし、後述する顔検出の失敗処理を行なう。

【 0 0 8 5 】

顔領域の検出に成功した場合は、顔領域検出成功フラグ FindFace を第 (2 1) 式として精密座標検出処理を終了する。

【 0 0 8 6 】

FindFace = 1 ... (2 1)

また、顔領域の検出に成功した場合は、顔領域検出成功フラグ FindFace を第 (2 2) 式とし、精密座標検出処理を終了する。

【 0 0 8 7 】

FindFace = 0 ... (2 2)

顔検出失敗時は、抽出された領域以外に顔の存在する可能性がある。そのときは、顔領域抽出処理の後段で行なわれる符号化制御処理やフィルタ処理で、実際の顔の部分が顔領域以外の処理が行なわれ、顔の画質が低下するおそれがある。これを避けるために、顔検出失敗時には、矩形領域を実際より大きな領域とするかまたは画面全体を領域とすることで顔が顔領域処理の対象領域から外れないようにする。

【 0 0 8 8 】

まず、現フレームで求められた矩形領域 X 座標の中心を第 (2 3) 式のように FaceCenter とする。

【 0 0 8 9 】

FaceCenter = (FaceWideXr - FaceWideXl) / 2 + FaceWideXl ... (2 3)

画面の中に 3 つの大きな領域 (area0, area1, area2) を設定し、FaceCenter の位置より、area0, area1, area2 のうちのどれかを現フレームの矩形領域とする。すなわち、第 (2 4) 式により現フレームの置換えを行なう。

【 0 0 9 0 】

0 FaceCenter < Border01 矩形領域を area0 に置換え
Border01 < FaceCenter Border12 矩形領域を area1 に置換え
Border12 < FaceCenter X 座標の最大 矩形領域を area2 に置換え

... (2 3)

表 2 に置換え領域 (area0, area1, area2) 選択の判定値 Border01, Border12 の値と、(area0, area1, area2) の置換え値を示す。

【 0 0 9 1 】

【表 2】

	Border01	Border12		FaceWideXl	FaceWideXr	HeadTopY	FaceBottomY
CIF QCIF	14	29	area0	2	28	3	28
			area1	9	34	3	28
			area2	18	41	3	28
SubQCIF	10	21	area0	1	21	2	18
			area1	10	21	2	18
			area2	11	30	2	18

【 0 0 9 2 】

顔検出失敗時の処理は、顔領域抽出関数をコールする符号化制御部により、FindFace フラ

10

20

30

40

50

グにより判定されて行なわれる。これは、顔領域抽出部でフレーム間処理の連続性を断ち切らないためである。これにより、次フレームで顔検出に成功した場合は、直ちに正しい矩形領域に復帰することができる。

【 0 0 9 3 】

次に、前記矩形領域の抽出処理で求めた矩形領域の拡大処理を行なう。動きが少なくなったときには全体的な符号量が減少するので、その分、顔領域を拡大することができる。その判定には第(25)式からどれだけ動きが少ないフレームが連続したかをカウントする Still Count を求め、ステップ S P 2 7、S P 2 8 および S P 2 9 の処理を行なう。

【 0 0 9 4 】

```
NoMv < THstill
then : StillCount ++
NoMv > THstill
then : StillCount = 0
StillCount > THstill max
then : StillCount = 0
... ( 2 5 )
```

10

そして、ステップ S P 3 0 で次のようにして領域の拡大領域 RgnExp を求め、FaceWideXl, FaceWideXr, HeadTopY を拡大し、それぞれ FaceWideXle, FaceWideXre, HeadTopYe とする。なお、THstill はここでは 20, THexp は C I F のとき 20 とし、Q C I F のとき 50 とし、S Q C I F のとき 70 とした。

20

【 0 0 9 5 】

```
RgnExp = StillCount / THexp
FaceWideXle = FaceWideXl - RgnExp
FaceWideXre = FaceWideXr + RgnExp
HeadTopYe = HeadTopY - RgnExp
ただし、
FaceWideXle < 0 のとき、FaceWideXle = 0
FaceWideXre > 43 のとき、FaceWideXre = 43
HeadTopYe < 0 のとき、HeadTopYe = 0
... ( 2 6 )
```

30

また、前フレームの顔幅 PreFaceWideXle, PreFaceWideXre, 頭頂 PreHeadTopYe と比較し、差が THdif より小さいときは頻繁な変動を防ぐために前フレームの値を使用する。なお、THdif はここでは 3 とした。

【 0 0 9 6 】

```
| Pre FaceWideXle - FaceWideXle | < THdif
then: FaceWideXle = PreFaceWideXle
| PreFaceWideXre - FaceWideXre | < THdif
then: FaceWideXre = PreFaceWideXre
| PreHeadTopYe - HeadTopYe | < THdif
then: HeadTopYe = PreHeadTopYe
... ( 2 7 )
```

40

さらに、移動量に第(28)式に示すローパスフィルタをかけ、前フレーム位置からの急峻な変動を抑制する。

【 0 0 9 7 】

```
FaceWideXle = PreFaceWideXle × k + FaceWideXle × ( 1-k )
FaceWideXre = PreFaceWideXre × k + FaceWideXre × ( 1-k )
HeadTopYe = PreHeadTopYe × k + HeadTopYe × ( 1-k )
k = 1/2
... ( 2 8 )
```

顔領域は、一般に画面に対して中心付近に位置する。そこで、ステップ S P 3 1 で顔の長

50

さを算出する。すなわち、画面の上下左右の位置に対して第(29)式に示すクリッピングを施し、画面の端に矩形領域が片寄らないようにする。

【0098】

```
FaceWideXle < minX
then: FaceWideXle = minX
FaceWideXre > maxX
then: FaceWideXle = maxX
HeadTopY < minY
then: HeadTopY = minY
```

ただし、

```
minX = 縮小画像の横幅 × 1/10
maxX = 縮小画像の横幅 - (縮小画像の横幅 × 1/10)
minY = 縮小画像の縦幅 × 1/12
FaceWideXre < FaceWideXle
then: FaceWideXre = FaceWideXle
... (29)
```

最後に、次のフレームのために顔幅と頭頂を記憶しておく。

【0099】

```
PreFaceWideXle = FaceWideXle
PreFaceWideXre = FaceWideXre
PreHeadTopYe = HeadTopYe
```

... (30)

顔幅と頭頂位置から顔の下の座標FaceBottomYeを推定する。

【0100】

```
FaceBottomYe = HeadTopYe + (FaceWideXre - FaceWideXle + 1) * 15
... (31)
```

FaceWideXle, FaceWideXre, HeadTopYe と同様に第(32)式によりクリッピングを施し、画面の端に矩形領域が偏らないようにする。

【0101】

```
FaceBottomYe > maxY
then: FaceBottomYe = maxY
```

ただし、

```
maxY = 縮小画像の縦幅 - (縮小画像の縦幅 × 1/12)
FaceBottomYe < HeadTopYe
then: FaceBottomYe = HeadTopYe
```

... (32)

次に、ステップSP32で抽出された顔幅、頭頂、顔の長さにより得られる顔領域の矩形領域から符号化制御の際に使用するFaceMapを作成する。これは、顔幅、頭頂、顔の長さの各座標は44×36画素の縮小画像に対応しているが、この領域内に50%以上含まれるマクロブロック1とし、それ以外のマクロブロックには0を表わすラベルを付ける。FaceMapは画像フォーマットごとに大きさが異なり、CIFのとき22×18となり、QCIFのとき11×9となり、SQCIFのとき8×6の大きさとなる。

【0102】

なお、動きベースによる顔抽出は動きの検出に数フレームの履歴を使用しているため、顔領域抽出処理の開始フレームからTHtrackまではFaceMapがオールゼロとする。

【0103】

また、処理のスキップ判定がなされていた場合(Bypass flag = 1)は、この処理を行わずに前フレームで求めたFaceMapを現フレームのFaceMapとする。

【0104】

次に、色ベースによる肌色抽出処理を行なう。色ベースによる肌色抽出は、前記の動きベ

10

20

30

40

50

ースによる顔抽出において抽出できなかった他の肌の領域、たとえば、手、腕、首などの領域を抽出するために実施され、HueMapを作成する。

【0105】

ただし、処理のスキップ判定がされていた場合 (Bypass flag =1) は、この処理を行わずに前フレームで求めたHueMapを現フレームのHueMapとする。

【0106】

図23は色ベースによる肌色抽出の処理の流れを示すフローチャートである。図23において、ステップSP121でBypass flag が1でないことを判別し、ステップSP122において顔面の精密座標から肌色サンプル領域を設定する。顔面の精密座標EMNtop, EMNbottom, EMNI, EMNr から得られる矩形領域をCb, Crに設定して肌色サンプル領域とする。ただし、QCIF, SQCIFの場合は精密座標を半分にした値を使用する。

10

【0107】

次に、ステップSP123において、肌色サンプル領域内の平坦領域におけるCb, Crの平均および標準偏差の算出を行なう。前述した動きベースによる顔抽出で作成した縦および横エッジ画像のエッジ画素を含まない肌色サンプル領域外のCb, Crの画素の値からCb, Crのそれぞれの平均 μ_u, μ_v , 標準偏差 σ_u, σ_v を算出する。

【0108】

色ベースによる肌色抽出は動きベースによる顔抽出結果に基づいて実施されるが、動きベースによる抽出が100%正確でないため、顔領域を誤って抽出していた場合には肌色抽出に悪影響を与える。したがって、サンプルしたCb, Crの分布により肌色抽出を実施するかどうかを判断する必要がある。これは、標準偏差 σ_u, σ_v から第(33)式を満足したときに単峰性のピークがあると考え、ステップSP124で肌色抽出処理を実施し、肌色画素の抽出処理を行なう。

20

【0109】

$$\sigma_u < TH \quad \text{かつ} \quad \sigma_v < TH \quad \dots (33)$$

第(33)式を満足したときに単峰性のピークがあると考え、ステップSP126~ステップSP128による肌色抽出処理を実施し、肌色画素の抽出処理を行なう。第(33)式を満足しないときには色にばらつきがあり安定した肌色抽出ができないので、ステップSP125で肌色抽出処理を中止し、FaceMapと同じ構成でHueMapを作成してオールゼロとし、処理を終了する。なお、TH はここでは20とする。

30

【0110】

ステップSP126~SP128の肌色画素の抽出では、始めにステップSP126でCb, Crの抽出範囲の値 [Cbl, Cbh], [Crl, Crh] を次のように決定する。

【0111】

$$[Cbl, Cbh] = [\mu_u - \sigma_u, \mu_u + \sigma_u]$$

$$[Crl, Crh] = [\mu_v - \sigma_v, \mu_v + \sigma_v]$$

...(34)

次に、ステップSP126で得られた抽出範囲に従い、Cb, Crがともにその範囲に属している画素をステップSP127で抽出する。さらに、現フレームのオブジェクトマスク内の画素のみを抽出して肌色領域の数を作成する。

40

【0112】

このようにして作成された肌色領域をそれぞれ画像フォーマット (CIF, QCIF, SQCIF) に対応させ、マクロブロック内に50%以上肌色画素を含むとき、そのマクロブロックを1とし、そうでないときは0のラベルを持つHueMapをステップSP128で作成する。なお、顔領域抽出処理の開始フレームからTHtrack まではHueMapをオールゼロとする。

【0113】

今回開示された実施の形態はすべての点で例示であって制限的なものではないと考えられるべきである。本発明の範囲は上記した説明ではなくて特許請求の範囲によって示され、特許請求の範囲と均等の意味および範囲内でのすべての変更が含まれることが意図される

50

。

【0114】

【発明の効果】

以上のように、この発明では、矩形領域抽出時に、前フレーム矩形領域外の動き成分判定しきい値を矩形領域内動き成分判定しきい値よりも感度を落とすことにより、対象人物周囲の動き成分雑音を矩形領域と判定しにくくし、対象人物顔領域外に矩形領域が移動しにくくすることで矩形領域の判定精度を向上できる。

【0115】

また、前フレームと現フレーム間の動き画素量により動き画素量が少ないときには、動き画素量算出以降の演算を行わないことで、矩形領域抽出処理の全体の演算量を軽減でき

10

【0116】

さらに、顔画像特徴抽出による精密座標判定のための顔幅判定時に、顔幅検索範囲を頭頂座標から画面最下座標までの1/2の位置の矩形座標の左右幅とすることで、精密座標検索の精度向上と演算の高効率化を行なうことができる。

【0117】

顔画像特徴抽出による精密座標判定のための顔幅判定時に、鼻成分を避けることにより、実際の顔幅よりも狭い領域を精密座標と判定しないようにする。

【0118】

さらに、顔画像特徴抽出による精密座標抽出が失敗した場合、ある決められた大きさの領域を矩形領域とすることで矩形領域外に顔画像が存在することを低減できる。

20

【0119】

さらに、ある決められた大きさの領域を矩形領域とすることに代えて、画面全体を矩形領域とすることで矩形領域外に顔画像が存在することを低減できる。

【0120】

さらに、現フレームで動き画素より算出された矩形領域座標と、前フレームの矩形領域座標の値にローパスフィルタをかけ、動き物体が画面内で急激に移動した場合でも滑らかに矩形領域が動き物体に追従することができる。

【0121】

さらに、画面内の端の領域には有効な顔画像が存在しないようにし、求められた矩形領域座標値に対してクリッピング処理を行なうことで、後の精密座標抽出時の精度向上と演算量削減を行なうことができる。

30

【0122】

さらに、顔画像特徴抽出による精密座標判定のための顔幅判定時に、前フレームで求めた精密座標と現フレームで求めた精密座標の値の差があるしきい値内であれば前フレームで算出した精密座標値を現フレームの精密座標として採用することにより、精密座標が見失われることを避けることができる。このとき、前フレームで求めた精密座標を利用することに代えて、過去数フレームの精密座標の値の平均値を用いることで同様の作用を得ることができる。

【図面の簡単な説明】

40

【図1】H.261/H.263画像エンコーダと顔領域抽出部の関係を示すブロック図である。

【図2】顔領域抽出の大まかな流れを示すフローチャートである。

【図3】動きベースによる顔抽出を示す前半のフローチャートである。

【図4】動きベースによる顔抽出を示す後半のフローチャートである。

【図5】縮小画像の作成例を示す図である。

【図6】オブジェクトマスク作成の流れを示すフローチャートである。

【図7】動き履歴のアップデートを示す図である。

【図8】動き画素判定しきい値適応領域を示す図である。

【図9】3×3画素のウィンドウによる拡大処理を示す図である。

50

- 【図10】 3×3 画素のウィンドウによる縮小処理を示す図である。
- 【図11】 $TH_{m,v}$ 算出の流れを示すフローチャートである。
- 【図12】 最大領域の抽出を示す図である。
- 【図13】 頭頂の検出例を示す図である。
- 【図14】 顔幅検出の流れを示すフローチャートである。
- 【図15】 顔幅検出範囲の決定を説明するための図である。
- 【図16】 顔幅の検出例を示す図である。
- 【図17】 顔幅補正の流れを示すフローチャートである。
- 【図18】 精密座標の検出処理の流れを示すフローチャートである。
- 【図19】 エッジ検出オペレータを示す図である。
- 【図20】 顔幅精密座標検出例を示す図である。
- 【図21】 横エッジの特徴量を示す図である。
- 【図22】 顔面の縦方向の精密座標算出例を示す図である。
- 【図23】 色ベースによる肌色抽出の処理の流れを示すフローチャートである。

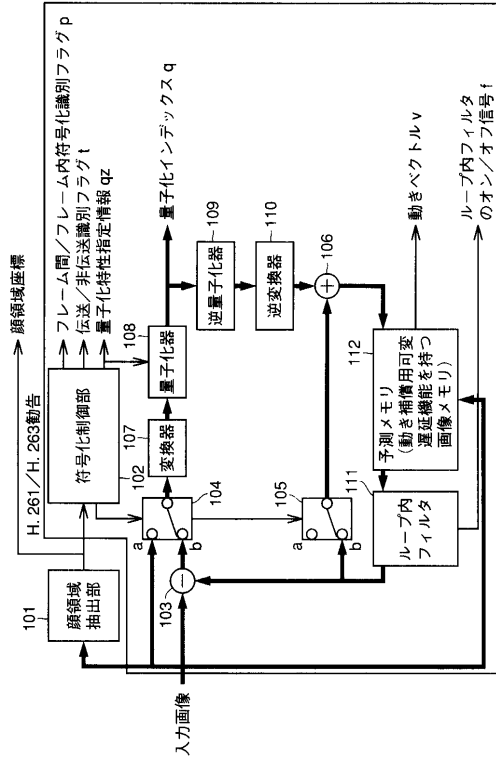
10

【符号の説明】

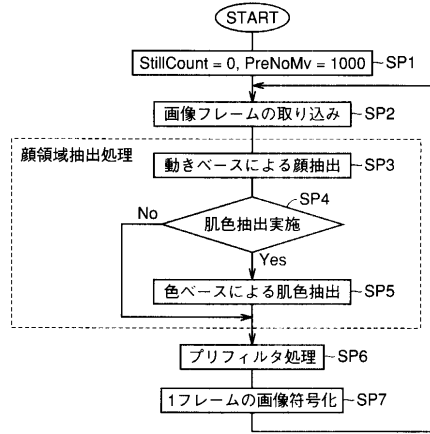
- 101 顔領域抽出部
- 102 符号化制御部
- 103 減算器
- 104, 105 スイッチ
- 106 加算器
- 107 変換器
- 108 量子化器
- 109 逆量子化器
- 110 逆変換器
- 111 ループ内フィルタ
- 112 予測メモリ

20

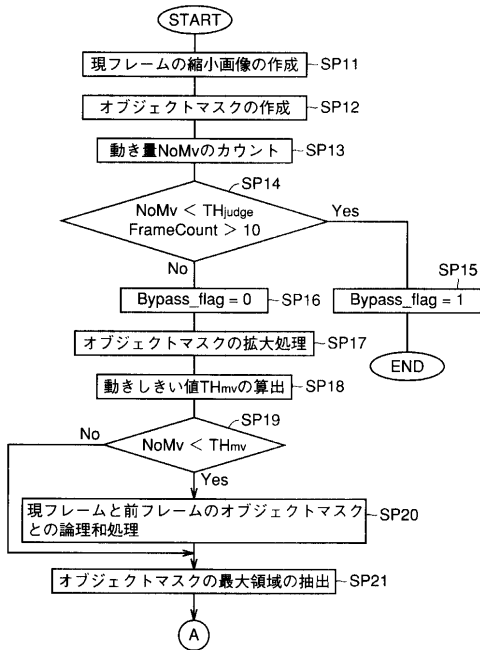
【 図 1 】



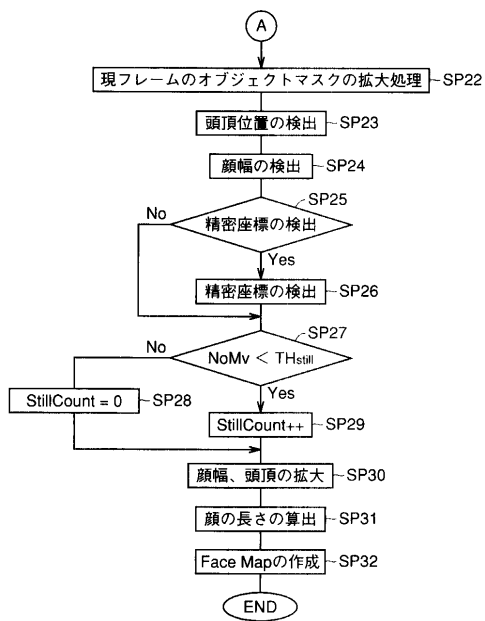
【 図 2 】



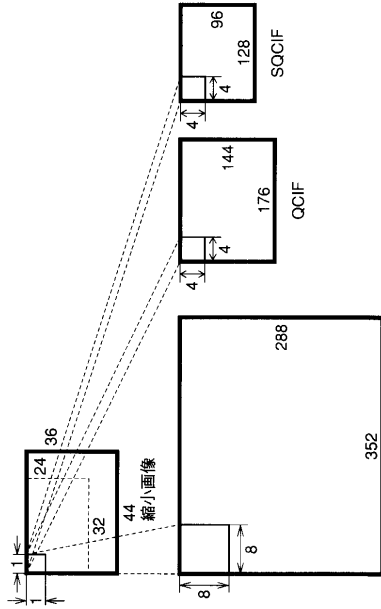
【 図 3 】



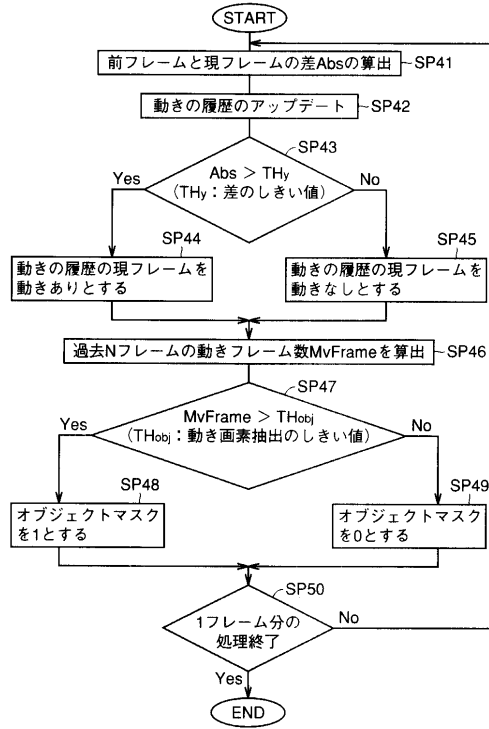
【 図 4 】



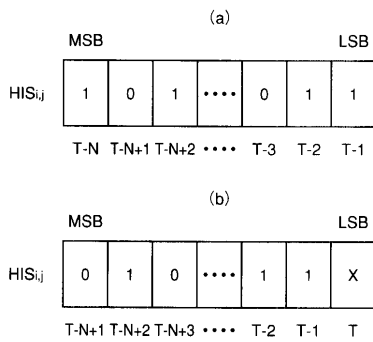
【 図 5 】



【 図 6 】

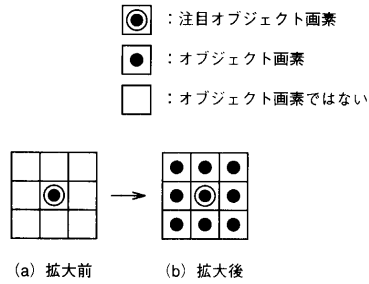


【 図 7 】

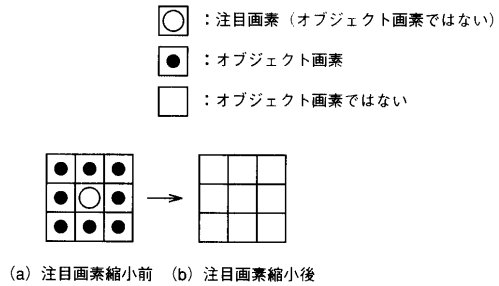


1: 動き画像
0: 静止画像

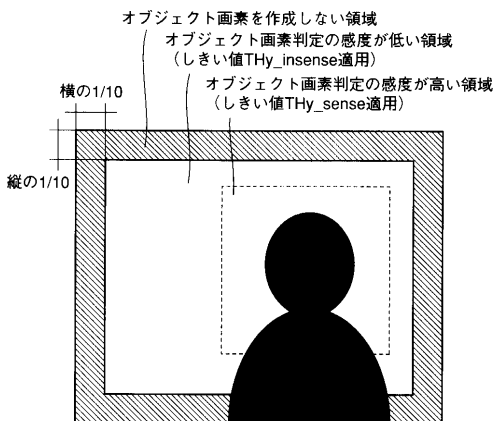
【 図 9 】



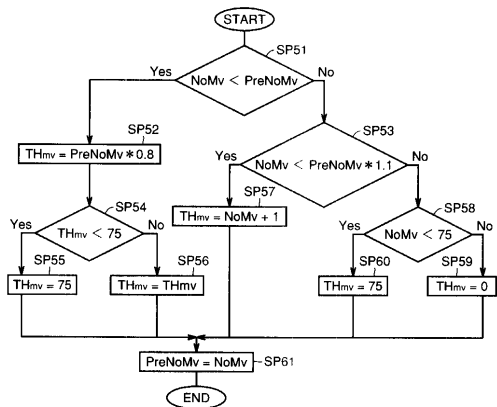
【 図 10 】



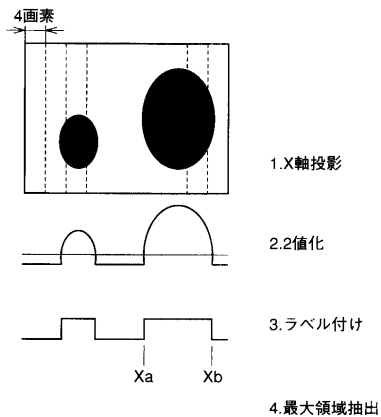
【 図 8 】



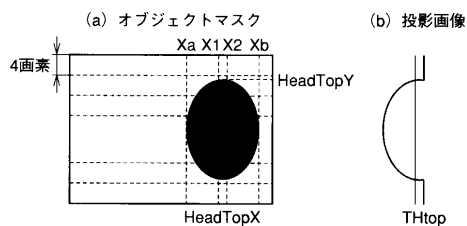
【 図 1 1 】



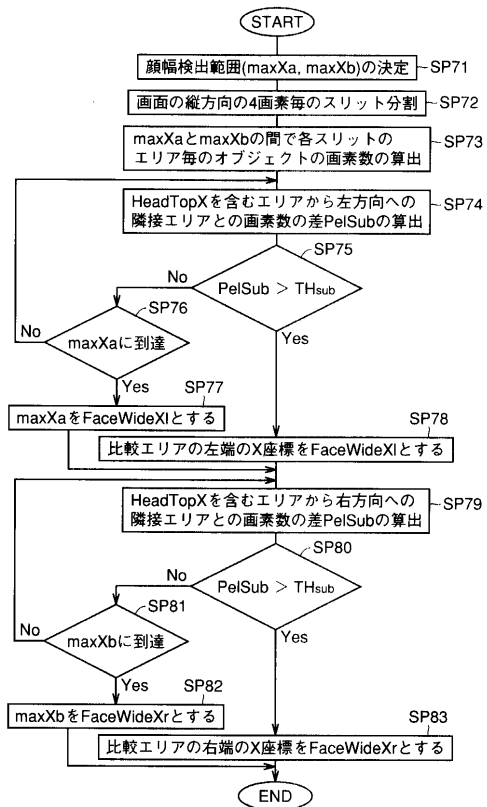
【 図 1 2 】



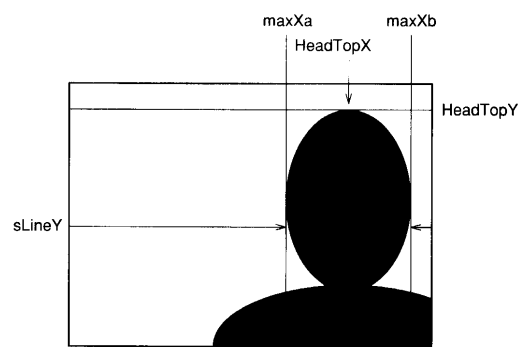
【 図 1 3 】



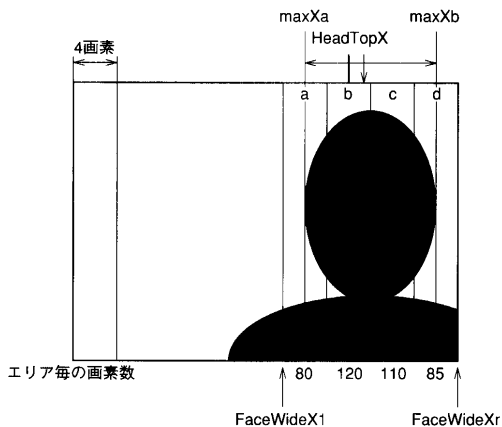
【 図 1 4 】



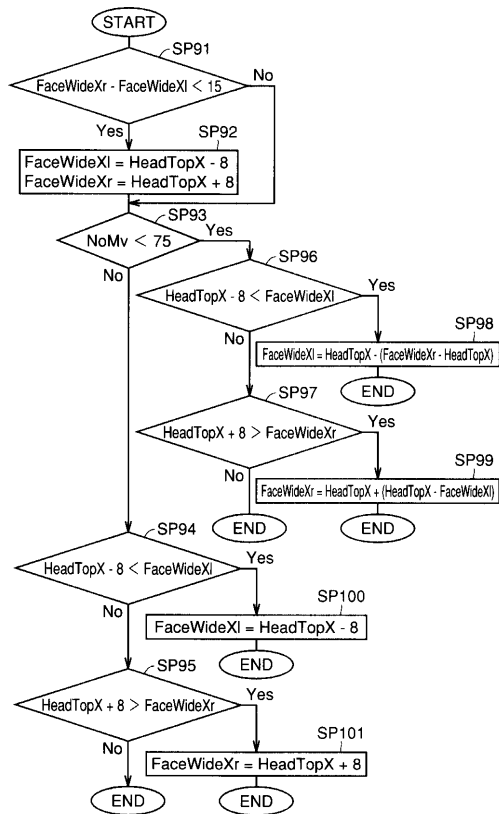
【 図 1 5 】



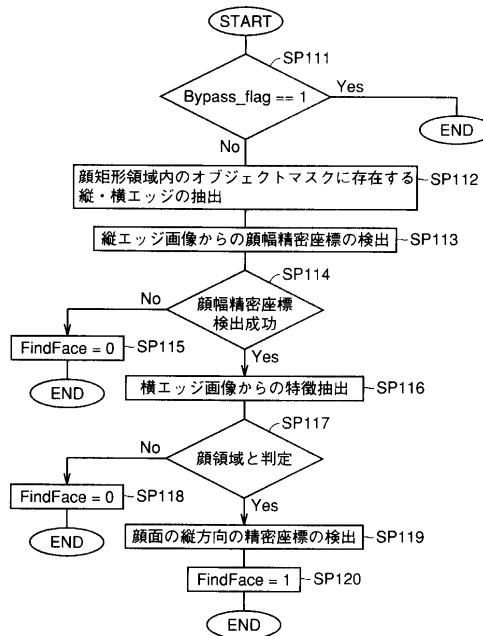
【 図 1 6 】



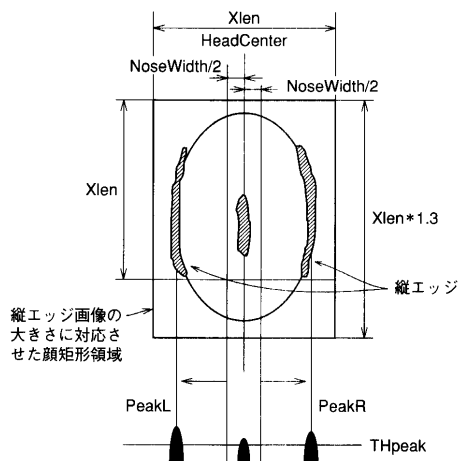
【 図 1 7 】



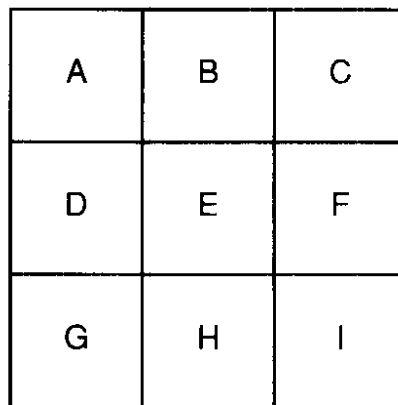
【 図 1 8 】



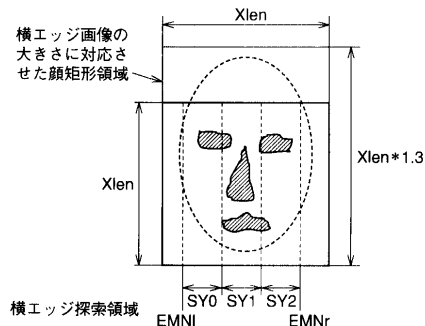
【 図 1 9 】



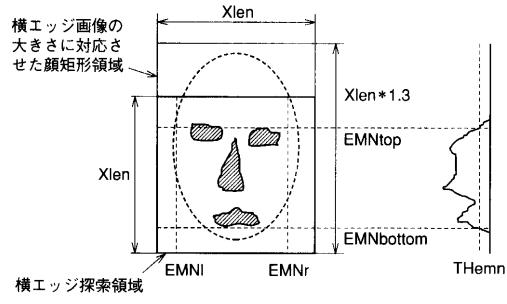
【 図 2 0 】



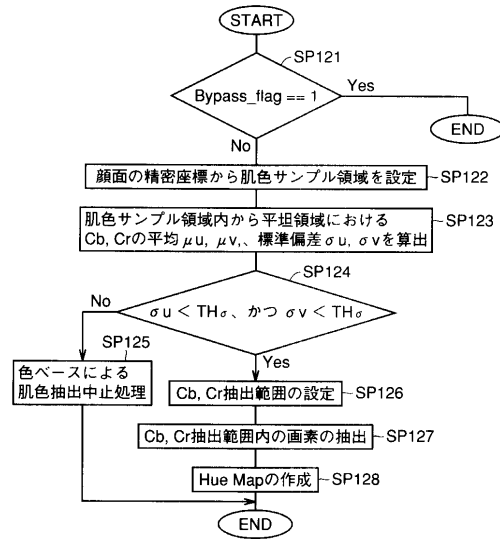
【 図 2 1 】



【 図 2 2 】



【 図 2 3 】



フロントページの続き

- (56)参考文献 特開平05 - 304662 (JP, A)
特開昭64 - 080185 (JP, A)
特開平06 - 133303 (JP, A)

- (58)調査した分野(Int.Cl.⁷, DB名)

H04N7/137

H04N7/15

H04N11/00~11/22

G06T7/00~7/60