

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-81732

(43) 公開日 平成9年(1997)3月28日

(51) Int.Cl.⁶
G 0 6 T 7/00

識別記号 庁内整理番号

F I
G 0 6 F 15/70

技術表示箇所
3 3 0 C

審査請求 未請求 請求項の数7 F D (全 32 頁)

(21) 出願番号 特願平7-266407

(22) 出願日 平成7年(1995)9月19日

(71) 出願人 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

(72) 発明者 船山 竜士

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 横矢 直和

奈良県生駒市高山町8916-5 B-301

(72) 発明者 竹村 治雄

京都府相楽郡木津町木津川台1-27-9

(72) 発明者 岩佐 英彦

奈良県生駒市高山町8916-5 C-202

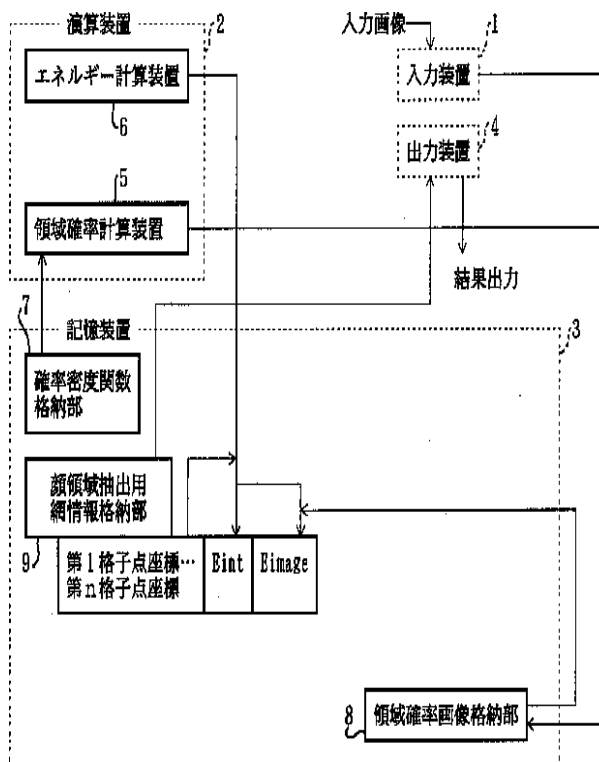
(74) 代理人 弁理士 原 謙三

(54) 【発明の名称】 領域抽出装置及びそれを用いた方向検出装置

(57) 【要約】

【課題】 照明条件や入力画像の大きさ、画像中の顔位置や大きさ、回転、背景などの変化に左右されることなく、顔領域を抽出する。

【解決手段】 顔領域を表す確率密度分布が格納された確率密度関数格納部7を参照することにより、入力装置1により入力された人物画像の顔領域確率画像を生成する領域確率計算装置5と、画像全体を覆うように設けられる仮想的な網を構成する格子点の位置関係に基づく内部エネルギーと、顔領域確率画像上の格子点が位置するピクセル色の確からしさに基づく画像エネルギーとをそれぞれ計算するエネルギー計算装置6とを有する。2つのエネルギーの合計値が小さくなる方向へ網を収縮変形させることにより、顔領域を抽出する。



【特許請求の範囲】

【請求項 1】対象物を撮像してカラーデジタル画像とする撮像手段と、

上記カラーデジタル画像において均一に同様の色を持つ所望領域を示す確率密度関数を予め記憶しておく確率密度関数格納手段と、

上記所望領域を構成する各画素の色に基づいて上記確率密度関数を計算することにより、各画素の色の確からしさを算出する領域確率計算手段と、

上記カラーデジタル画像上の任意の 3 つ以上の点を接続することにより形成され画像全体を覆うように設けられる仮想的な網の情報を記憶しておく網情報格納手段と、上記点の位置関係に基づいて定義される網の内部エネルギーと、点が位置する画素の色の確からしさに基づいて定義される網の画像エネルギーとをそれぞれ計算するエネルギー計算手段とを有し、

1 つの点を移動させたときの網の内部エネルギーと画像エネルギーとの合計値を、上記点を移動させる前の合計値と比較し、合計値が変化するときには網が収縮する方向へ点の移動を行い、合計値が変化しなくなった時点で点の移動を停止することを特徴とする領域抽出装置。

【請求項 2】上記対象物は人物であり、上記所望領域は人物の顔の領域であることを特徴とする請求項 1 記載の画像処理装置。

【請求項 3】確率密度関数格納手段には上記所望領域と異なる色を持つ部分領域を示す確率密度関数が予め記憶されると共に、網情報格納手段には上記所望領域を囲んで停止した網の情報が部分領域を抽出するための部分領域抽出用網の初期値として記憶され、

上記所望領域内に部分領域が 1 つある場合に、領域確率計算手段が上記部分領域を構成する各画素の色に基づいて上記確率密度関数を計算することにより各画素の色の確からしさを算出して、エネルギー計算手段が部分領域抽出用網の内部エネルギー及び画像エネルギーと、部分領域に収束されるように定義された移動エネルギーとをそれぞれ計算することにより、部分領域を抽出することを特徴とする請求項 1 または 2 記載の領域抽出装置。

【請求項 4】部分領域抽出用網の外周を形成する点からフレーム重心を求めると共に、画像の明るさを重みとした部分領域抽出用網が囲む領域の重心を求める重心計算手段を有し、

上記移動エネルギーを、上記フレーム重心が領域の重心の方向に引き寄せられるように定義することを特徴とする請求項 3 記載の領域抽出装置。

【請求項 5】網の外周を形成する点からフレーム重心を求めると共に、画像の明るさを重みとして網が囲む領域の重心を求める重心計算手段を有し、

確率密度関数格納手段には口唇領域を示す確率密度関数が予め記憶されると共に、網情報格納手段には上記顔領域を囲んで停止した網の情報が口唇領域及び左右の目領

域を抽出するための網の初期値として記憶され、

領域確率計算手段が、上記口唇領域を構成する各画素の色に基づいて上記確率密度関数を計算することにより各画素の色の確からしさを算出すると共に、上記左右の目領域を構成する各画素の色に基づいて顔領域あるいは口唇領域の確率密度関数を計算することにより各画素の色の確からしさを算出して、

エネルギー計算手段が、各領域の内部エネルギー及び画像エネルギーと、各々のフレーム重心が各々の領域重心の方向へ引き寄せられ、かつ、3 つの領域の重心の位置関係が予め測定され記憶された口唇及び左右の目の位置関係と等しくなるように定義されたバランスエネルギーとを計算することにより、口唇及び左右の目領域を抽出することを特徴とする請求項 2 記載の領域抽出装置。

【請求項 6】上記撮像手段が複数個設けられ、複数の撮像手段から同時に得られる複数の画像から口唇領域及び左右の目領域をそれぞれ抽出し、互いに比較することにより口唇及び目の 3 次元位置情報を導出することを特徴とする請求項 5 記載の領域抽出装置を用いた方向検出装置。

【請求項 7】抽出された目の画像と、予め測定して視線方向の情報が付加されて記憶された目の画像とを比較することにより推定視線方向を算出し、さらに推定視線方向と検出された頭部の向きとを比較することにより真の視線方向を検出することを特徴とする請求項 6 記載の領域抽出装置を用いた方向検出装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、入力された画像から、顔が存在している位置、顔の輪郭、及び顔部品などを検出する領域抽出装置、及びそれを用いて頭の方向や視線方向を検出する方向検出装置に関するものである。

【0002】

【従来の技術】顔に関する画像処理は、20 数年来という非常に長い間多くの研究がなされてきた。顔に関する画像処理は、顔の特徴点抽出による個人同定、プレートマッチングなどによる顔部品抽出、色情報を利用したデータ圧縮、モーフィングなどのアニメーション応用、位相情報の適用など、非常に広範囲に渡る。その中でも、入力された画像中で、顔がどこにあるか、そして、顔画像から、目、鼻、口といった顔部品を抽出する問題は、最も基本的、かつ非常に応用範囲の広い有用な技術である。これは、顔画像処理の最初のキーポイントであり、従来から様々な手法が提案されてきた。

【0003】

入力画像から顔位置を検出する手法は、画像の解像度を落として濃淡画像でプレートマッチングを行う手法や、色情報とセグメンテーションを利用する手法などがある。これらのいずれの手法も、光軸方向の回転やノイズなどに弱く、誤抽出を起こしやすい。

【0004】

顔部品の抽出に関しては、投影を用いる方

法が最もアルゴリズムが単純かつ高速な抽出手法であるが、この方法は照明の強度及び方向、画像中の顔の位置や大きさや回転などの変動に対して非常に弱く、照明の被写体までの距離や画質などがかなり制約された環境内で撮像された画像に対してでないとは有効でないことがわかっている。また、テンプレートマッチングを利用した顔部品の抽出法も提案されているが、これは投影を用いる方法と同様に、照明の変動などにより影が生じたり、顔の方向、回転の変更や個人差による変動などに非常に弱いということが言える。以下に投影とテンプレートマッチングについて詳しく説明する。

【0005】投影とはある画像をある方向から見た画像へ変換することである。この場合の方向とは3次元的な方向とは異なる。例えば、本実施の形態の説明に用いる図8に示す入力顔画像があるとすると、これを水平方向に微分して2値化すると図47(a)に示すようになる。水平方向に微分するとは、図48に示すように、ピクセルの明るさを横方向に順に見ていき、あるピクセルの明るさと、その前のピクセルの明るさとの差を新しいピクセルの明るさとすることである。従って、明るさが急激に変化している画像を微分すると、その変化しているところで微分値が大きくなっているため、エッジの抽出などを行うことができる。

【0006】この2値化された水平方向の微分画像を水平方向に投影すると図47(b)、垂直方向に投影すると図47(c)に示すようになる。このとき、水平方向に投影するとは、2値化した微分画像のピクセルを水平方向に見ていき、ONになっているピクセル(図47(a)では黒いピクセル)の数を数えてそれをグラフにしたものである。垂直方向への投影は、2値化した微分画像のピクセルを垂直に見ていき、ONになっているピクセルの数を数えてそれをグラフにしたものである。こ

$$\text{評価値} = \sum_{i=1}^5 \sum_{j=1}^5 \{V(I_{ij}) - V(T_{ij})\}^2$$

ここで、 $V(p)$ はあるピクセルPの明るさ、 T_{ij} はテンプレートの座標 (i, j) におけるピクセル、 I_{ij} はテンプレートに対応する領域の座標 (i, j) におけるピクセルである。

【0011】図51に示すように、入力画像左上の 5×5 領域について評価値を計算したら、次に1ピクセル水平方向にずらした 5×5 領域について同じように評価値を計算する。同様にして、右下の 5×5 領域まで、入力画像のすべての 5×5 領域とテンプレートとの間で評価値を計算する。そして、すべての評価値の中で最も点数の高いものがテンプレートと最も似ている領域として抽出される。

【0012】しかしながら、この手法は、抽出しようとしている領域がテンプレートとずれる場合、例えばテンプレートより大きかったり小さかったりした場合、また、雑音や背景などの影響でテンプレートに似た領域な

の投影画像を調べることにより、顔領域や顔部品領域を抽出することができる。例えば、図47(b)の水平方向の投影画像のピークを調べることにより、目の垂直位置がわかる。また、図47(c)の垂直方向の投影画像の左右のピークを調べることにより、顔領域の左右の端を調べることができる。

【0007】しかしながら、この手法は、照明条件が変わると影のつき方が変わるので微分画像が大きく変化すると共に、雑音や背景、顔の回転などにより投影画像が変化するという問題がある。従って、投影の手法では非常に限られた条件で撮影された画像でないとは、顔領域や顔部品の抽出ができないということがわかる。即ち、照明や画像の大きさ、画像中の顔位置や大きさ、顔の向きなどが厳密にコントロールされた環境で撮影された画像に対してでないとは、投影の手法は有効でない。

【0008】次にテンプレートマッチングについて説明する。テンプレートとは、例えば図49に示す目の画像のように、抽出すべき領域の画像を予め記憶させたものである。テンプレートマッチングは、テンプレートと入力画像の対応する領域を比べて、最も点数の高い領域をそのテンプレートにマッチする領域として抽出するというものである。

【0009】テンプレートは、入力画像より小さく適当な大きさの画像である。図50に示すように、テンプレートの大きさが 5×5 だとすると、まず、入力画像の左上の 5×5 領域と、テンプレートとを比較して評価値を付ける。比較の方法は、例えば、テンプレートとそれに対応する入力画像の領域とで、対応するピクセルの明るさの2乗和の差を点数とする。これは以下のような式となる。

【0010】

【数1】

どができた場合に誤った抽出を起しやす。この手法も投影を用いた手法と同様に、照明や画像の大きさ、画像中の顔位置や大きさ、向きなどが厳密にコントロールされた画像に対してでないとは有効でない。

【0013】上記投影やテンプレートマッチングの欠点である照明や画像の大きさ、画像中の顔位置や大きさ、向きなどが厳密にコントロールされた環境で撮像された画像という前提条件をある程度緩和するものとして、Deformable template と呼ばれる手法を用いた顔部品の抽出が提案されている。

【0014】これは、図52に示すように、抽出しようとする領域の特徴を単純な数学的な図形関数の組み合わせで表現したテンプレート(図中の点線)と、そのテンプレートに対応する画像との間でエネルギーを計算して、最もエネルギーが小さくなるようにテンプレートを変形、移動させることによって領域を抽出しようという

ものである。

【0015】テンプレートは、例えば、次のように設定されている。図53に示すように、中心 x_c ・半径 r の円が瞳に相当する。円周は瞳と白目のエッジに引き寄せられる一方、円内部は明るさの落ち込んだ領域に引き寄せられるようにエネルギーの定義を行う。

【0016】目の境界は、焦点を x_0 に持つ上下に凸の2つの放物線の一部が合わさって形成される。焦点 x_0 から目の片端までのまでの幅を b' 、即ち目の幅を $2b'$ 、上に凸の放物線の高さを a' 、下に凸の放物線の高さを c' 、焦点 x_0 を通る水平線と目の中心線（図中、焦点 x_0 を通る点線）とのなす角を回転角 θ とする。エネルギーは、目の境界がエッジに引き寄せられるように定義する。

【0017】瞳により2分割された白目領域のそれぞれの中心は、 $x_0 + p_1(\cos\theta, \sin\theta)$ 、及び $x_0 + p_2(\cos\theta, \sin\theta)$ で表される（ $p_1 > 0, p_2 < 0$ ）。2つの白目領域の中心が近傍画素の中で最も明るい領域の中心に引き寄せられるようにエネルギーの定義を行う。

【0018】目の境界と瞳の間は白目に相当し、明度の

$$E_v = - (c_1/\text{Area}) \iint_{\text{circle-Area}} \Phi_v(x) dA$$

【0023】

$$E_e = - (c_2/\text{Length}) \int_{\text{circle-Bound}} \Phi_e(x) ds$$

$$- (c_3/\text{Length}) \int_{\text{para-Bound}} \Phi_e(x) ds$$

【0024】

$$E_i = - (c_4/\text{Area}) \iint_{\text{circle-Area}} \Phi_i(x) dA$$

【0025】

$$E_i = - (c_5/\text{Area}) \iint_{\text{whites}} \Phi_i(x) dA$$

【0026】

$$E_p = c_6 \{ \Phi(x_0 + p_1 e_1) + \Phi(x_0 + p_2 e_1) \}$$

【0027】

$$E_{int} = k_1(x_0 - x_c)^2 / 2 + k_2 \{ (p_1 - (r + b')) / 2 \}^2 / 2$$

$$+ k_3 \{ (p_2 - (r + b')) / 2 \}^2 / 2 + k_4 (b' - 2r)^2 / 2$$

上記の式は、幾何的な構成が最適になったときにそれぞれのエネルギーが最小になるように定義されている。この方法は、画像の大きさ、画像面に平行な回転、頭の回転や照明の影響を受けにくいという特徴がある。

【0028】上記のように顔領域や顔部品を検出する以外に、頭部方向や視線方向を検出する技術も開示されている。

【0029】頭部方向を検出する方法としては、磁気センサなどを用いた方法がある。これは、図54に示すように、利用者が頭部に3つ以上の電磁波を放射する電磁波放射装置を装着すると共に、電磁波放射装置からの電磁波を受信する電磁波受信機を床面などに3箇所以上設置する構成である。これによれば、1つの電磁波放射装置につき3箇所以上の電磁波受信機が受信する信号の位

大きい領域に引き寄せられるようにエネルギーの定義を行う。

【0019】これらの要素には、中心 x_c と焦点 x_0 とは互いに引き寄せ合い、一致したときにエネルギーが最小となる、目の幅 $2b'$ はほぼ半径 r の4倍となる、白目領域の2つの中心点は目の中心線上に乗るような力が作用する。

【0020】以上のパラメータを少しずつ変化させ、エネルギー関数 $E_c(x_c, x_0, p_1, p_2, r, a', b', c', \theta)$ が最小となるようにテンプレートが変形することにより目にマッチングする。エネルギー関数 E_c は以下の式で表される。

【0021】

【数2】

$$E_c = E_v + E_e + E_i + E_p + E_{int}$$

ここで、 E_v は谷エネルギー、 E_e はエッジエネルギー、 E_i は画像エネルギー、 E_p はピークエネルギー、 E_{int} は内部エネルギーであり、次のように定義される。

【0022】

【数3】

$$E_v = - (c_1/\text{Area}) \iint_{\text{circle-Area}} \Phi_v(x) dA$$

【数4】

$$E_e = - (c_2/\text{Length}) \int_{\text{circle-Bound}} \Phi_e(x) ds$$

$$- (c_3/\text{Length}) \int_{\text{para-Bound}} \Phi_e(x) ds$$

【数5】

$$E_i = - (c_4/\text{Area}) \iint_{\text{circle-Area}} \Phi_i(x) dA$$

【数6】

$$E_i = - (c_5/\text{Area}) \iint_{\text{whites}} \Phi_i(x) dA$$

【数7】

$$E_p = c_6 \{ \Phi(x_0 + p_1 e_1) + \Phi(x_0 + p_2 e_1) \}$$

【数8】

$$E_{int} = k_1(x_0 - x_c)^2 / 2 + k_2 \{ (p_1 - (r + b')) / 2 \}^2 / 2$$

$$+ k_3 \{ (p_2 - (r + b')) / 2 \}^2 / 2 + k_4 (b' - 2r)^2 / 2$$

相差を利用して、各電磁波放射装置の3次元位置を検出し、電磁波放射装置が取り付けられている頭部の向きを検出することができる。

【0030】また、専用ハードウェアを用いずに頭部位置を検出する方法もある。これは、図55に示すように、利用者が頭部の適当な位置に3箇所以上、目立つ色のシールなどを貼るか、あるいはシールを貼ったヘッドギアなどをかぶると共に、2つ以上の撮像装置を備える構成である。これによれば、シールを撮影した2つ以上の画像からシールの3次元位置を測定し、頭部の方向を検出することができる。

【0031】また、視線を検出するには、普通、装着型の視線トラッカーを用いる。視線トラッカーには、常に眼球を捉えて眼球の画像を撮影するために、眼球のすぐ

近くにCCDなどの撮像装置が配置されている。

【0032】

【発明が解決しようとする課題】しかしながら、上記Deformable template の手法は、抽出しようとしている部品を表すテンプレートの初期位置を別手法を用いて設定しなければならないという問題がある。これは、初期位置として抽出したい領域近くにテンプレートを位置させないと誤抽出が多くなると共に、計算量が莫大になるからである。別手法としては上述の投影やテンプレートマッチングなどがあげられるが、テンプレートマッチングのような従来手法の欠点を克服するための手法が、前提としてそのような従来技術を用いなければならないことが大きな矛盾となっている。

【0033】結果として、Deformable template の手法による顔輪郭・顔部品の抽出も、投影やテンプレートマッチングの手法と同じような問題が発生する。即ち、Deformable template の手法は、その抽出の精度と、照明条件や画像の大きさ、画像中の顔位置や大きさ、回転、背景などの変化に対して弱く、厳密にコントロールされた環境で撮影された画像に対してでないとは有効でない。つまり、普通の環境で人物を撮影する場合、常に同じ照明を用い、常に同じ背景・方向・向きに設定されていることはあり得ず、従って、通常の状態では撮影された画像に対してそれらの技術を適用すると、誤った領域の抽出を起しかねない。また、エネルギーの定義を部品ごとに定義しなければならないという問題も有している。

【0034】また、電磁波放射装置を用いて頭部方向を検出する方法は、電磁波放射装置を頭部に動かないようにしっかり固定する必要があり、また、そこに電源を供給する仕組みも必要となる。さらに、電磁波放射装置から一定距離以上離れた床などに電磁波受信機を固定して置かなくてはならない。この結果、頭部にハードウェアを装着しなければならないという煩わしさと、コストがかかるという欠点がある。一方、頭部にシールなどのマーカーを貼る方法は低コストで実現できるが、シールを貼ったり、あるいはヘッドギア状のものを被ったりという煩わしさを除去することはできないという問題を有している。

【0035】さらに、視線方向を検出する方法は、顔に固定して動かないような堅固な構造を持つ視線トラッカーを装着するために、装着感が非常に悪く、また、コストもかかるという問題を有している。

【0036】

【課題を解決するための手段】上記の目的を達成するために、本発明の請求項1記載の領域抽出装置は、対象物を撮像してカラーデジタル画像とする撮像手段（例えば、撮像装置）と、上記カラーデジタル画像において均一に同様の色を持つ所望領域を示す確率密度関数を予め記憶しておく確率密度関数格納手段（例えば、確率密度関数格納部）と、上記所望領域を構成する各画素（ピク

セル）の色に基づいて上記確率密度関数を計算することにより、各画素の色の確からしさを算出する領域確率計算手段（例えば、領域確率計算装置）と、上記カラーデジタル画像上の任意の3つ以上の点（例えば、格子点）を接続することにより形成され画像全体を覆うように設けられる仮想的な網の情報を記憶しておく網情報格納手段（例えば、顔領域抽出用網情報格納部）と、上記点の位置関係に基づいて定義される網の内部エネルギーと、点が位置する画素の色の確からしさに基づいて定義される網の画像エネルギーとをそれぞれ計算するエネルギー計算手段（例えば、エネルギー計算装置）とを有し、1つの点を移動させたときの網の内部エネルギーと画像エネルギーとの合計値を、上記点を移動させる前の合計値と比較し、合計値が変化するときには網が収縮する方向へ点の移動を行い、合計値が変化しなくなった時点で点の移動を停止することを特徴としている。

【0037】上記構成によれば、まず、撮像手段により対象物が撮影され、対象物のカラーデジタル画像が得られる。そして、領域確率計算手段が確率密度関数格納手段を参照し、カラーデジタル画像における抽出したい領域を構成する各画素の色の値を確率密度関数に代入して計算することにより、各画素の色の確からしさが算出される。このとき、所望領域は均一に同様の色で形成されており、かつ確率密度関数も所望領域であることを表す関数であるので、所望領域の色の確からしさが大きくなる。

【0038】次に、入力されたカラーデジタル画像全体を覆うように設けられた仮想的な網を構成する点を1つずつ移動させ、点を移動する前と後とで網のエネルギーに変化があるかどうかを調べる。ここで、網のエネルギーとは、エネルギー計算手段により求められた、網の形状を反映する内部エネルギーと網を構成する点が位置する画素の色を反映する画像エネルギーとの合計値である。

【0039】このエネルギーの合計値に変化がある場合には網が収縮する方向へ点を移動させ、合計値に変化がなくなると点の移動を停止させる。即ち、初めに画像全体を覆っていた網は、色の確からしさが大きい所望領域に向かって収縮変形し、最終的に所定領域を囲んで停止することになる。従って、所望領域を抽出することができる。

【0040】これにより、本領域抽出装置は照明条件の変化に強い色の確からしさに基づいて所望領域を抽出することができるので、撮影する場所、時間、補助照明の有無、方向などに左右されることがなくなる。また、網はどのような形状にもなり得るので、入力画像の大きさ、及び画像中の顔位置や大きさなどに制限されることがなくなる。これらの結果、撮影の自由度を増大させることができる。

【0041】請求項2記載の領域抽出装置は、請求項1

記載の構成に加えて、上記対象物は人物であり、上記所望領域は人物の顔の領域であることを特徴としている。

【0042】上記構成によれば、まず、撮像手段により人物を撮影することにより人物画像が得られる。この画像に対して仮想的な網を用いると、顔領域に網が収縮して顔領域を抽出することができる。

【0043】これにより、さまざまな家電製品などの機器に利用することができる。即ち、機器に備えられた領域抽出装置が顔領域を抽出することにより、機器と利用者とのような位置関係にあるのかを判断することができる。そして、機器の前に利用者が存在するか否か、また存在するのならばどの辺りにいるのかの情報を得ることができる。これらの情報を利用して、利用者の位置による機器の制御を行うことが可能となる。

【0044】請求項3記載の領域抽出装置は、請求項1または2記載の構成に加えて、確率密度関数格納手段には上記所望領域と異なる色を持つ部分領域を示す確率密度関数が予め記憶されると共に、網情報格納手段には上記所望領域を囲んで停止した網の情報が部分領域を抽出するための部分領域抽出用網の初期値として記憶され、上記所望領域内に部分領域が1つある場合に、領域確率計算手段が上記部分領域を構成する各画素の色に基づいて上記確率密度関数を計算することにより各画素の色の確からしさを算出して、エネルギー計算手段が部分領域抽出用網の内部エネルギー及び画像エネルギーと、部分領域に収束されるように定義された移動エネルギーとをそれぞれ計算することにより、部分領域を抽出することを特徴としている。

【0045】上記構成によれば、請求項1または2において所定領域（顔領域）を囲んで停止した網の情報が部分領域を抽出する部分領域抽出用網の初期値として網情報格納手段に記憶される。また、確率密度関数格納手段には上記所望領域と異なる色を持つ部分領域を示す確率密度関数が予め記憶されている。そして、所望領域内に部分領域が1つある場合に、部分領域の各画素の色の確からしさを求めた後、部分領域に収束されるように定義された移動エネルギーを持つ網を動かすことで部分領域を抽出する。即ち、初めに所定領域を覆っていた網は、色の確からしさが大きい部分領域に向かって収縮変形・移動し、最終的に部分領域を囲んで停止することになる。

【0046】これにより、最初の抽出領域に用いた網を、さらに領域内の別の領域の抽出に用いることができるので、領域毎に網の設定を変える必要がない。また、請求項2においては、顔領域が抽出されると、顔領域と色の異なる口唇領域や目領域が抽出できる。

【0047】請求項4記載の領域抽出装置は、請求項3記載の構成に加えて、部分領域抽出用網の外周を形成する点からフレーム重心を求めると共に、画像の明るさを重みとした部分領域抽出用網が囲む領域の重心を求める

重心計算手段（例えば、重心計算装置）を有し、上記移動エネルギーを、上記フレーム重心が領域の重心の方向に引き寄せられるように定義することを特徴としている。

【0048】上記構成によれば、部分領域抽出用が囲む領域の重心が画像の明るさを重みとしているので、領域重心の位置はフレーム重心の位置に比べて抽出したい領域近くに位置することになる。このとき、フレーム重心が領域重心の方向に引き寄せられるように移動エネルギーが設定されているので、網は部分領域に向かって移動する。これにより、所定領域内に部分領域に似た色を持つ領域があったとしても、その領域に引き寄せられることなく、部分領域を抽出することができる。この結果、処理速度が速く、高精度の部分領域の抽出が可能となる。

【0049】請求項5記載の領域抽出装置は、請求項2記載の構成に加えて、網の外周を形成する点からフレーム重心を求めると共に、画像の明るさを重みとして網が囲む領域の重心を求める重心計算手段を有し、確率密度関数格納手段には口唇領域を示す確率密度関数が予め記憶されると共に、網情報格納手段には上記顔領域を囲んで停止した網の情報が口唇領域及び左右の目領域を抽出するための網の初期値として記憶され、領域確率計算手段が、上記口唇領域を構成する各画素の色に基づいて上記確率密度関数を計算することにより各画素の色の確からしさを算出すると共に、上記左右の目領域を構成する各画素の色に基づいて顔領域あるいは口唇領域の確率密度関数を計算することにより各画素の色の確からしさを算出して、エネルギー計算手段が、各領域の内部エネルギー及び画像エネルギーと、各々のフレーム重心が各々の領域重心の方向へ引き寄せられ、かつ、3つの領域の重心の位置関係が予め測定され記憶された口唇及び左右の目の位置関係と等しくなるように定義されたバランスエネルギーとを計算することにより、口唇及び左右の目領域を抽出することを特徴としている。

【0050】上記構成によれば、請求項2において顔領域を囲んで停止した網の情報が口唇領域及び左右の目領域を抽出するための網の初期値として網情報格納手段に記憶される。また、確率密度関数格納手段には口唇領域の確からしさを示す確率密度関数が予め記憶されている。そして、口唇領域の各画素の色の確からしさと、左右の目領域の各画素の色の確からしさとを求めた後、内部エネルギー、画像エネルギー、及びバランスエネルギーを持つ3つの網を動かすことで口唇及び左右の目領域を抽出する。即ち、初めに顔領域を覆っていた網は、色の確からしさが大きい3つの領域に向かってそれぞれ収縮変形・移動し、最終的に各領域を囲んで停止することになる。

【0051】これにより、顔領域抽出用網を、口唇領域抽出用網及び左右の目領域抽出用網とすることができる。

ので、領域毎に網の設定を変える必要がない。

【0052】また、顔領域に対する顔部品の位置がわかるので、顔の回転角（顔の傾き）を認識することができる。顔の回転角がわかると、例えば、人物データベースを作る際に、写真をスキャナなどで読み込む場合に、写真の方向などを気にせずに入力することができ、手間が大いに軽減する。この結果、撮影時、及び処理時の自由度が増大する。

【0053】さらに、顔領域と口唇領域とのバランスにより、顔のおおよその方向を推定することができる。即ち、口唇領域の中心が顔領域の主軸に近い位置にあるならば、利用者は撮像手段の方向を向いており、左右のどちらかにずれていればそれぞれの方向に向いていることになる。

【0054】請求項6記載の方向検出装置は、請求項5記載の領域抽出装置を用いて、上記撮像手段が複数個設けられ、複数の撮像手段から同時に得られる複数の画像から口唇領域及び左右の目領域をそれぞれ抽出し、互いに比較することにより口唇及び目の3次元位置情報を導出することを特徴としている。

【0055】上記構成によれば、複数の撮像手段により同時に同じ人物を撮影すると、画面上で人物の位置が異なる複数の画像が得られる。そして、請求項5記載の構成の領域抽出装置により複数の画像における口唇領域及び目領域がそれぞれ抽出される。それらの複数の画像において、対応する顔部品どうしを互いに比較すると、画像どうしのずれにより口唇、左目、及び右目それぞれの3次元空間上の3点が決まる。これにより、頭部がどの方向に向いているかを検出することが可能となる。

【0056】このようにして、従来のような専用のハードウェアを用いなくとも、容易に頭部方向を検出することができる。この結果、ハードウェアなどを装着する煩わしさを除去することができると共に、装置の低コスト化が可能となる。

【0057】請求項7記載の方向検出装置は、請求項6記載の構成に加えて、抽出された目の画像と、予め測定して視線方向の情報が付加されて記憶された目の画像とを比較することにより推定視線方向を算出し、さらに推定視線方向と検出された頭部の向きとを比較することにより真の視線方向を検出することを特徴としている。

【0058】上記構成によれば、予めの測定により黒目の位置と視線方向との関係がある程度わかっているので、請求項6記載の構成にて抽出された目の画像を用いて、推定視線方向を算出することができる。推定視線方向が得られると、それをさらに頭部の向きと比較して補正を行い、真の視線方向を検出することができる。

【0059】これにより、従来のような専用のハードウェアを用いなくとも、容易に視線方向を検出することができる。この結果、ハードウェアなどを装着する煩わしさを除去することができると共に、装置の低コスト化が

可能となる。

【0060】

【発明の実施の形態】

〔実施の形態1〕本発明の実施の一形態について図1ないし図9、図11ないし図19に基づいて説明すれば、以下の通りである。

【0061】本実施の形態にかかる領域抽出装置は、カラーデジタル画像上の任意の点（ピクセル）を結ぶことによって形成される仮想的な網のモデルを用いることにより、顔領域を抽出する装置である。

【0062】上記領域抽出装置は、図1に示すように、人物などの対象物を撮像してカラーデジタル画像とするCCDなどの撮像装置（撮像手段）である入力装置1、各種計算処理を行う演算装置2、データを予め記憶したり、一時的にデータを記憶する記憶装置3、及び網が目的の領域に収束した時点の結果を出力する出力装置4により構成される。

【0063】上記演算装置2は、領域確率計算装置（領域確率計算手段）5及びエネルギー計算装置（エネルギー計算手段）6を有する。領域確率計算装置5は、入力された画像において目的の領域を抽出する前段階として、後述の確率密度関数格納部7を参照することにより、色情報としての領域の確率密度関数を導出し後述の領域確率画像を生成する装置である。エネルギー計算装置6は、後述の顔領域抽出用網情報格納部9を参照することにより網の内部エネルギー（ E_{int} ）を計算すると共に、顔領域抽出用網情報格納部9及び領域確率画像格納部8を参照することにより画像エネルギー（ E_{image} ）を計算する。

【0064】記憶装置3は、確率密度関数格納部（確率密度関数格納手段）7、領域確率画像格納部8、及び顔領域抽出用網情報格納部（網情報格納手段）9を有する。

【0065】確率密度関数格納部7には予めサンプリングすることにより求められた顔領域の肌の色の確率密度関数が格納されている。領域確率画像格納部8には上記領域確率計算装置5で生成された領域確率画像が格納される。

【0066】顔領域抽出用網情報格納部9には、入力画像上の複数の点を接続することにより形成され画像全体を覆うように設けられる仮想的な網の情報が記憶されている。即ち、仮想的な網を構成する格子点の初期座標が予め格納されている。この初期座標は、入力された画像に対して均等に位置するように自動的に計算されて設定される。また、理論的には格子点が3点あれば領域を形成することができるが、抽出する精度や抽出しようとする形、画像の大きさに合わせて格子点の数は設定される。また、顔領域抽出用網情報格納部9には網が移動することにより変化した後の格子点の座標も記憶される。さらに、顔領域抽出用網情報格納部9にはエネルギー計

算装置 6 により計算され求められた内部エネルギー (E_{int}) 及び画像エネルギー (E_{image}) が格納される。

【 0 0 6 7 】 以下に、上記構成による領域抽出装置の動作、及び網の定義の仕方などを具体的に説明する。

【 0 0 6 8 】 まず、顔領域を表す確率密度関数を算出するために、人間の顔の皮膚の色分布を調べる。この作業の工程を図 2 のフローチャートに基づいて説明する。統計をとるためにより多くの人物顔の画像を撮影し、手作業で顔の皮膚だけの画像を作成する (S 1)。例えば、図 3 (a) に示す入力人物画像に対して、図 3 (b) の A に示すように、顔の肌色部分のみを取り出す。顔領域だけを切り出したデジタル画像から適当な数のピクセル (画素) をランダムに選ぶ。そのピクセルの色を H S V 表色系で表した場合の H (色相) 及び S (彩度) の出現回数をカウントし、それらの度数分布を得る (S 2)。

【 0 0 6 9 】 色相は、赤・青・黄といった色の違いを区別する属性で、照明などによる反射や陰影の影響を受けにくいという特徴がある。また、彩度は色の鮮やかな程度を表す指標である。人間の顔は比較的彩度が高く、一方、室内などは比較的彩度が低いもので構成されていることが多い。

【 0 0 7 0 】 図 4 は、24 人の人物画像データにより、

$$N (\mu , \Sigma) = (2 \pi)^{-k/2} | \Sigma |^{-1/2} \exp [- (\mathbf{x} - \mu) \Sigma^{-1} (\mathbf{x} - \mu)^T / 2]$$

$$\mathbf{x} = (x_1 , x_2) , \mu = (\mu_1 , \mu_2) , \Sigma = (\sigma_{ij})$$

で表されるので、求められた $\mu_1 , \mu_2 , \sigma_{ij}$ を (1) 式に代入することにより顔領域を表す確率密度関数を導出することができる。この確率密度関数を予め前記確率密度関数格納部 7 に記憶させておく。図 4 及び図 5 の色度数分布に基づいて求められた確率密度関数をプロットしたものを図 6 に示す。図の高さ軸は確率密度、左側軸は色相、右側軸は彩度を表している。

【 0 0 7 4 】 領域確率計算装置 5 が、求められた確率密度関数の、 x_1 に色相値、 x_2 に彩度を代入して計算することにより、入力デジタル画像のすべてのピクセルについて、顔の色の確からしさが 0 ~ 1 の範囲で求められる。そして、そのピクセルの色相と彩度とにより求めた確からしさの値 (0 ~ 1) をピクセルの新しい値とするような画像を生成する。このようにして求められた画像を領域確率画像と称することにする。

【 0 0 7 5 】 領域確率画像の生成の過程を図 7 のフローチャートに基づいて説明する。画像が入力されると (S 1 1)、入力画像の左上から順に 1 つずつピクセルを取り出す (S 1 2)。取り出したピクセルの色相と彩度とを求め (S 1 3)、先に求めた確率密度関数の引数としてその色相と彩度を用いて、領域の確からしさを計算する。その確からしさの値をピクセルの新しい値とする (S 1 4)。そして、すべてのピクセルの新しい値が得

られたかどうかを判断し (S 1 5)、すべてのピクセルについてまだ計算されていない場合には上記 S 1 2 ~ S 1 4 の工程を繰り返し、すべてのピクセルに対して計算が行われたら処理を終了する。これにより、領域確率画像が生成される。図 9 は、図 8 に示される入力画像から生成された顔領域確率画像を示す。これは、明るいピクセルほど顔領域である可能性が高いことを表している。

【 0 0 7 1 】 顔領域における色相及び彩度の色度数分布を得ると、色相と彩度とのパラツキが正規分布に従っていると仮定して、顔領域を表す 2 次元の正規確率密度関数を導出する (S 3)。即ち、 μ_1 を色相の平均値、 μ_2 を彩度の平均値とし、 σ_{ij} を色相と彩度の分布の分散共分散行列とすると、まず、手作業で切り出してカウントした色の分布からこれらの値を計算する。

【 0 0 7 2 】 2 次元の正規確率密度関数は、

【 0 0 7 3 】

【 数 9 】

... (1)
られたかどうかを判断し (S 1 5)、すべてのピクセルについてまだ計算されていない場合には上記 S 1 2 ~ S 1 4 の工程を繰り返し、すべてのピクセルに対して計算が行われたら処理を終了する。これにより、領域確率画像が生成される。図 9 は、図 8 に示される入力画像から生成された顔領域確率画像を示す。これは、明るいピクセルほど顔領域である可能性が高いことを表している。

【 0 0 7 6 】 次に網のモデルについて説明する。網は、図 1 1 に示すように、デジタル画像上のピクセルに相当する黒点を結ぶことにより形成される。この網を形作る黒点を格子点と呼ぶ。この格子点の位置関係によって網のエネルギーが定義される。網のエネルギーは、網自身の形状から生じる内部エネルギーと、網とそれが覆う画像とで決まる画像エネルギーとの 2 つのエネルギーを持っている。この 2 つのエネルギーの合計値が小さくなるように格子点を移動させて網を変形させることにより目的の領域を抽出する。

【 0 0 7 7 】 ある格子点は隣り合う別の格子点と接続関係を持っており、この点どうしの位置関係によって網の内部エネルギーが定義される。網の内部エネルギーは、格子点の距離が小さければ小さいほどエネルギーも小さく、また、格子点どうしを結ぶ網を構成するラインが滑らかであればあるほどエネルギーが小さくなるように、

格子点どうしを結んで構成される格子の向かい合う辺が平行になるように定義する。また、格子点でのピクセルの明るさにより画像エネルギーを定義する。つまり、格子点上にあるピクセルの状態をエネルギーに反映することにより領域の抽出を行うのである。

【0078】領域の抽出は、上記のような網のモデルを前記領域確率計算装置5によって生成された領域確率画像に対して適用することにより行われる。図12(a)~(d)は、網のモデルを用いてある領域を抽出するようすを示している。図12(a)に示す入力画像から領域確率画像を算出し、同図(b)に示すように、画像全体を覆うような網をかぶせる。すると、網は目的の領域に向かって収縮し(同図(c)参照)、最終的にその領域を囲んで停止する(同図(d)参照)。顔領域を抽出する場合には、図13(a)~(d)に示すようになる。即ち、顔領域確率画像に対して、画像全体を覆うような網をまずかぶせる。そして、この網のエネルギーが小さくなるように格子点を移動させていき、エネルギーの変化がなくなったところで、格子点の移動を停止させる。すると、網は顔領域を囲むような形に変形している。

【0079】(1)網の形状

上記網の形状は図14に示すように定義される。ある点を中心に半径を次第に大きくしていき複数の同心円を作る。そして、その中心点から放射状に線分を伸ばし、最も外側の同心円まで伸ばす。同心円の円周と、その放射

$$E_{int}(t,s) = E_t(t,s) + E_{tt}(t,s) + E_s(t,s) + E_{ss}(t,s) + E_{ts}(t,s) \dots \quad (2)$$

【0083】

【数11】

$$E_t(t,s) = |p(t,s) - p(t+1,s)|$$

【0084】

$$E_{tt}(t,s) = |p(t-1,s) - 2 \times p(t,s) + p(t+1,s)| \quad (t > 1)$$

$$= 0 \quad (t = 0)$$

【0086】

$$E_{ss}(t,s) = |p(t,s-1) - 2 \times p(t,s) + p(t,s+1)|$$

【0087】

$$E_{ts}(t,s) = |p(t,s) - p(t,s-1) - p(t-1,s) + p(t-1,s-1)|$$

上記 E_t は、図16(a)に示すように、同一スポークs上にある隣り合う格子点 $p(t,s)$ と $p(t+1,s)$ との距離を表す。距離が小さいほど E_t は小さくなり、 $p(t,s)$ と $p(t+1,s)$ とが一致すると $E_t = 0$ となる。

【0088】 E_s は、同一レイヤーt上にある隣り合う格子点 $p(t,s)$ と $p(t,s+1)$ との距離を表す。距離が小さいほど E_s は小さくなり、 $p(t,s)$ と $p(t,s+1)$ とが一致すると $E_s = 0$ となる。

【0089】 E_{tt} は、図16(b)に示すように、 $p(t,s)$ と、同一スポークs上で $p(t,s)$ の前後の格子点

状に伸びた線分との交点を格子点とする。従って、格子点の数は、同心円の数をT、放射状線分の数をSとすると、 $T \times S + 1$ 個になる。ここで、同心円をレイヤー、放射状線分をスポークと呼ぶことにする。また、一番外側のレイヤーをレイヤー0とし、内側に向かってレイヤー1、レイヤー2、...とする。一番外側のレイヤー0と各スポークとの交点を最外郭格子点と呼び、それ以外の格子点を内部格子点と呼ぶ。なお、網の形状は円形としているが、図15に示すように、長方形に変形しても格子点どうしの接続関係は変わらず、同じように扱うことができる。

【0080】(2)網の内部エネルギー

あるレイヤーtとスポークsとにより形成される格子点を $p(t,s)$ と表すことにする。あるレイヤーtより1つ内側のレイヤーをレイヤー(t+1)、1つ外側のレイヤーをレイヤー(t-1)とする。もちろん、最外郭格子点においては、レイヤー(t-1)は存在しない。同様に、あるスポークsより時計回り方向の次のスポークをスポーク(s+1)、反時計回り方向の次のスポークをスポーク(s-1)とする。

【0081】格子点 $p(t,s)$ の内部エネルギー $E_{int}(t,s)$ は、次のように表される。なお、本文中及び

図16のpはベクトルを表す。

【0082】

【数10】

【数12】

$$E_s(t,s) = |p(t,s) - p(t,s+1)|$$

【0085】

【数13】

【数14】

【数15】

$p(t-1,s)$ 及び $p(t+1,s)$ とが滑らかに接続されているかどうかを表す。3つの格子点が直線となる場合、 $E_{ts} = 0$ となる。

【0090】 E_{ss} は、 $p(t,s)$ と、同一レイヤーt上で $p(t,s)$ の左右の格子点 $p(t,s-1)$ 及び $p(t,s+1)$ とが滑らかに接続されているかどうかを表す。3つの格子点が直線となる場合、 $E_{ss} = 0$ となる。

【0091】 E_{ts} は、 $p(t,s-1) - p(t,s)$ と、 $p(t-1,s-1) - p(t-1,s)$ とが同じベクトルである、即ち互いに平行で大きさが同じであるときに $E_{ts} = 0$ となる。

【0092】ある格子点 $p(t,s)$ に注目し、その周囲の格子点の位置から、上で求められる $E_{int}(t,s)$ を計算する。これがその格子点の持つ内部エネルギーである。これをすべての格子点について計算しその和を求めると、網の内部エネルギー E_{int} が算出される。なお、(2)式における $E_{int}(t,s)$ を構成する各項について、画像の大きさと格子点の数との比や色の確からしさ

$$E_{image}(t,s) = -FRPI(t,s) \quad (t > 1)$$

$$= FRPI(t,s) \quad (t = 0)$$

ここで、 $FRPI(t,s)$ は、領域確率画像における格子点 $p(t,s)$ 上のピクセルの明るさ(領域の確からしさ)を示している。即ち、内部格子点 ($t > 1$) においては領域確率画像における格子点 $p(t,s)$ 上のピクセルの明るさにマイナス符号を付け、最外郭格子点 ($t = 0$) においては符号を付けずにそのまま格子点 $p(t,s)$ の画像エネルギーとする。

【0095】ある格子点 $p(t,s)$ に注目し、上で求められる $E_{image}(t,s)$ を計算する。これがその格子点の持つ画像エネルギーである。これをすべての格子点について

$$E_{net}(t,s) = E_{int}(t,s) + E_{image}(t,s) \quad \dots (4)$$

網全体のエネルギー E_{net} は、(4)式をすべての格子点について計算し、合計したものであり、次のように表される。

【0098】

【数18】

$$E_{net} = \sum_i \sum_s E_{net}(t,s)$$

このようにして、定義した網全体のエネルギー E_{net} を計算し、このエネルギーの大きさが小さくなるように格子点を移動させる。エネルギーの変化がなくなった時点で格子点の動きを止めると、網は顔領域を囲んで収束している(図13参照)。

【0099】(4)格子点の移動

次に、格子点をどのように移動させていくかを図17のフローチャートに基づいて説明する。

【0100】まず、上述のように、領域確率計算装置5により顔領域確率画像を生成する(S21)。次に、初期状態の網を設定する。この設定は、図13に示すように、入力画像全体を覆うように設定する(S22)。最初に最外郭格子点の中から1つの格子点を選び、その格子点のエネルギー $E_{net}(t,s)$ を計算する(S23)。そして、図18に示すように、この格子点が位置するピクセルの周りに配置されている8つのピクセル(以下、8近傍と称する)のいずれかに格子点を移動させる。このとき、格子点を8近傍のそれぞれに移動したときのエネルギーを計算し(S24)、その中で最もエネルギーが小さくなるピクセルに格子点を移動させる(S25)。最初の格子点の位置でのエネルギーが最も小さい

の範囲によっては、適当に重みを決めてやってもよい。

【0093】(3)網の画像エネルギー

ある格子点 $p(t,s)$ の画像エネルギー $E_{image}(t,s)$ は、その格子点上のピクセルの状態として、以下のように定義される。

【0094】

【数16】

$$\dots (3)$$

て計算しその和を求めると、網の画像エネルギー E_{image} が算出される。なお、(3)式における $E_{image}(t,s)$ を構成する各項については、(2)式の場合と同様に、適当に重みを決めてやってもよい。

【0096】任意の格子点 $p(t,s)$ のエネルギー $E_{net}(t,s)$ は、格子点の持つ内部エネルギー $E_{int}(t,s)$ と画像エネルギー $E_{image}(t,s)$ との和で表される。

【0097】

【数17】

場合は、格子点を移動させる必要はない。

【0101】最外郭格子点から内部格子点に順に格子点を1つずつ、すべての格子点について8近傍のどこに移動させればエネルギーが最も小さくなるかを調べて、そこに格子点を移動させる。すべての格子点についてピクセル1つ分移動させる移動処理(以下、1単位変形と称する)が終了したかどうかを調べ(S26)、処理が終了していない場合にはS23の工程に戻り、終了した場合にはS27の工程へ進む。

【0102】S27では、各格子点でのエネルギーを合計した網全体のエネルギー E_{net} が、格子点を移動する前のエネルギー E_{net} と比較してその大きさに変化が生じているかどうかを調べる。S27で網全体のエネルギー E_{net} に変化がある場合にはS23の工程に戻って再び上S23~S26の処理を行い、S27で網全体のエネルギー E_{net} に変化がない場合には格子点の移動を終了する。格子点の移動が終了したときには、網は顔領域全体を囲んでいる(図13参照)。

【0103】以上により、本実施の形態における領域抽出装置は、格子点の位置関係に基づく網の内部エネルギーと、格子点が位置するピクセルの色の確からしさに基づく画像エネルギーとの合計値が変化しなくなるまで網を収縮変形させ、合計値が変化しなくなった時点で格子点の移動を停止する構成である。

【0104】従って、入力画像全体を覆っていた網は、色の確からしさが大きい顔領域に向かって収縮し、最終的に顔領域を囲んで停止する。この結果、顔領域を抽出することができる。

【0105】このとき、顔領域の色の確からしさは照明条件に左右されない色相に基づいて求められるので、本領域抽出装置では常に同じ照明条件で撮影する必要がない。即ち、本領域抽出装置は照明条件の変化に強いので、撮影する場所、時間、補助照明の有無、方向などを気にしなくてもよい。なお、従来では採光窓のない室内の同じ位置で、同じ照明器具を用い、同じ方向から人物を撮影することによって常に同じ照明条件を設定しなければならなかった。

【0106】また、本領域抽出装置は、網はどのような形状にもなり得るので、入力画像の大きさがどんなものでも対応することができる。一般に領域抽出の処理を行うためには、画像をデジタルで扱う必要がある。1次的に入力される画像は、デジタルイメージであっても通常のフィルム写真であってもどのような形態でもかまわないのだが、抽出処理を行う段階ではデジタルイメージである必要がある。従って、入力された画像の形態が異なるものであれ、領域抽出処理をする前段階として、デジタルイメージへの変換が必要である。

【0107】このとき、画像の大きさによらないということは、例えば1枚のデジタル画像が1000×1000ドットでも512×512ドットでも全く関係ないということである。従来のテンプレートマッチングのような手法を用いると、抽出しようとしている部品大きさに合わせてテンプレートを設定しなければならないので、画像の大きさにおのずと制限がついてしまう。また、Deformable template による手法でも設定されたテンプレートに対応する部品のみしか抽出できない。

【0108】一方、本領域抽出装置は画像の大きさによらないので、どのような大きさの画像に対してでも対応することができる。従って、どのような解像度の撮像装置を用いてもよく、また、アナログ画像をサンプリングするときどのようなレートでもよいので、撮影の自由度が増える。

【0109】同様に、本領域抽出装置は画像中の顔位置や大きさによらないので、顔が映っていれば、人物が画面いっぱい映っていたとしても、あるいは、画面の隅に小さく映っていても顔を抽出することができる。この結果、人物を好きな位置から好きな構図で撮影することができるので、撮影の自由度が向上する。なお、従来のテンプレートマッチングの手法では、前述したように、抽出しようとしている領域に合わせてテンプレートを設定するため、入力画像の大きさだけでなく、画像中の顔の大きさにも制限がついてしまう。また、投影を用いる方法においても、画像中の顔の大きさがある程度わかっていないと、部品の抽出の精度は極端に落ちてしまう。

【0110】また、顔領域が抽出されることにより、これを情報圧縮に利用することが可能となる。例えば、顔領域以外の領域の解像度を下げたり、また、色数を落とすなどして情報を圧縮し、少ない記憶容量で多くの人物

画像を保存することができる。これは、人物データベースなどを作成する際に大きなメリットとなる。また、単位時間当たりの電送容量が少ない回線で人物中心の画像を転送する際、人物以外の情報を荒くすることによって情報量を減らし、狭い回線でもコマ数の多い滑らかな画像を送ったり、その他の多くの情報を送ることが可能となる。これは、テレビ会議、テレビ電話、電子秘書、及びナビゲーションシステムなどで人物による説明などが入るようなアプリケーションに適用可能である。

【0111】また、本領域抽出装置は、家電製品などのさまざまな機器の制御などに利用することができる。図19は、本領域抽出装置11を機器13の制御に適用した場合の基本的なブロック図である。機器13に取り付けられた撮像装置12が利用者の姿を捉え、そのデジタル画像を領域抽出装置11に送る。領域抽出装置11は、画像から利用者の顔領域を抽出し、その情報を制御装置14に送る。制御装置14は機器13に接続されており、領域抽出装置11から送られてきた情報をもとに機器13を制御する。この一連の作業を短時間で繰り返すことにより、実時間で機器13の制御を行うことのできるシステムを構築することができる。

【0112】このようにして、機器13に備えられた領域抽出装置11が顔領域を抽出することにより、機器13と利用者とはどのような位置関係にあるのかを判断することができる。そして、機器13の前に利用者が存在するか否か、また存在するのならどの辺りにいるのかの情報を得ることができる。

【0113】これにより、家電製品などを人物の向いている方向に向かって制御することが可能となる。例えば、エアコンに領域抽出装置を取り付けることにより、部屋のどの位置に人間がいて、どの位置にいないかを検出することができるため、人間に向かって風を送ったり、逆に人間を避けて風を送ったりといった制御が可能となる。また、人間がいないことを検知して、送風をストップさせることも可能である。

【0114】さらに、テレビやオーディオ製品などに適用した場合、画面の方向や音の方向を利用者の方向に向けて制御したり、音量やステレオの場合の左右の音量のバランスを調整したり、人がいないときにスイッチを切ると共に人が現れたらスイッチを入れたりすることも可能となる。

【0115】このように、利用者が機器の近くに来ることによって機器のスイッチをON/OFFしたり、利用者の位置によって機器を制御することは、ほとんどの電気製品に適用することが可能である。これにより、スイッチの入切の手間や、スイッチの切り忘れ、機器の調整の手間を低減することができる。

【0116】なお、電気製品などの人のいる/いないにあわせて制御するようなアプリケーションにおいては、従来では人間がスイッチをいちいちオン/オフしたり、

赤外線画像を用いて人のいる / いないを判断したりしていたが、前者の場合は操作の煩わしさがあり、後者の場合は認識の精度をあげようとする、高価な赤外線撮像装置などを用いなければならず高コストになってしまう問題があった。本領域抽出装置を用いれば、低コストで人のいる / いないに関する電気製品の制御が可能となる。

【0117】〔実施の形態2〕本発明の他の実施の形態について図3、図8ないし図10、図20ないし図31に基づいて説明すれば、以下の通りである。なお説明の便宜上、前記の実施の形態の図面に示した部材と同一の部材には同一の符号を付記し、その説明を省略する。

【0118】本実施の形態にかかる領域抽出装置は、入力された人物画像により顔領域、及び口唇領域を抽出するものである。ここでは、顔領域を抽出した後、その抽出された顔領域を初期状態とした口唇領域抽出用網を設定し、それを移動させて口唇領域を抽出する。

【0119】上記領域抽出装置は、図20に示すように、実施の形態1の構成に加えて、演算装置2に重心計算装置（重心計算手段）21を、記憶装置3に口唇領域抽出用網情報格納部（網情報格納手段）22を有している。

【0120】重心計算装置21は、最外郭格子点の座標を平均することにより求められる最外郭フレームの重心（ g ）と、網内部のピクセルの座標を口唇領域の画像の明るさを重みとして平均することにより求められる網が囲む領域の重心（ G ）とを算出する装置である。

【0121】顔領域抽出用網情報格納部9には重心計算装置21により求められた顔領域抽出用網の2つの重心と、実施の形態1と同様に顔領域抽出用網の内部エネルギー（ E_{int} ）及び画像エネルギー（ E_{image} ）とが格納される。また、口唇領域抽出用網情報格納部22には、初期状態として顔領域に収縮した網の座標が格納される。さらに、口唇領域抽出用網の2つの重心と、口唇領域抽出用網の内部エネルギー（ E_{int} ）及び画像エネルギー（ E_{image} ）と、後述の形状エネルギー（ E_{form} ）と、上記2つの重心から求められる移動エネルギー（ E_{move} ）とが格納される。また、確率密度関数格納部7には、顔領域を表す確率密度関数に加えて、口唇領域を表す確率密度関数が予め格納されている。

【0122】上記構成によれば、領域確率計算装置5が領域確率画像格納部8を参照して、顔領域及び口唇領域を構成する各ピクセルの色に基づいて確率密度関数を計算することにより各ピクセルの色の確からしさを算出し、顔領域確率画像及び口唇領域確率画像を生成する。次に、実施の形態1と同様に、エネルギー計算装置6が顔領域抽出用網の内部エネルギー（ E_{int} ）と画像エネルギー（ E_{image} ）を計算し、それを顔領域抽出用網情報格納部9に格納する。これらのエネルギーの合計が小さくなるように網を収縮変形・移動させ、まず顔領域を

抽出する。この結果が出力装置4に出力される。

【0123】重心計算装置21が、顔領域抽出用網情報格納部9に格納されている網の座標を参照して最外郭フレームの重心（ g ）を、上記網の座標と領域確率画像格納部8に格納された顔領域確率画像とを参照して顔領域抽出用網の囲む領域の重心（ G ）を計算し、それぞれ顔領域抽出用網情報格納部9に格納する。同様に、重心計算装置21が、口唇領域抽出用網情報格納部22に格納されている口唇領域抽出用網の座標を参照して最外郭フレームの重心（ g ）を、上記網の座標と領域確率画像格納部8に格納された口唇領域確率画像とを参照して口唇領域抽出用網の囲む領域の重心（ G ）を計算し、それぞれ口唇領域抽出用網情報格納部22に格納する。

【0124】また、エネルギー計算装置6が口唇領域抽出用網の内部エネルギー（ E_{int} ）、画像エネルギー（ E_{image} ）、形状エネルギー（ E_{form} ）、及び移動エネルギー（ E_{move} ）を計算し、それらを口唇領域抽出用網情報格納部22に格納する。これらの口唇領域抽出用網のエネルギーの合計値が小さくなるように口唇領域抽出用網を変形収縮・移動し、口唇領域を抽出する。その結果が出力装置4に出力される。

【0125】以下に、上記構成による領域抽出装置の動作、及び網の定義の仕方などを具体的に説明する。

【0126】まず、口唇領域を表す確率密度関数の算出するために、実施の形態1と同様に、例えば、図3（a）に示す入力人物画像に対して図3（b）のBに示すように口唇のみを取り出して、口唇の色分布を調べる。図21は、24人の人物画像データにより、1画像につき100ピクセルの点をサンプリングしたときの口唇領域の色相の色度数分布を示す。図22は、同様にサンプリングしたときの口唇領域の彩度の色度数分布を示す。なお、各図の縦軸及び横軸は、実施の形態1の図4及び図5と同じである。

【0127】そして、実施の形態1と同様に求めた口唇領域を表す2次元の正規確率密度関数をプロットしたものを図23に示す。図の高さ軸は確率密度、左側軸は色相、右側軸は彩度を表している。

【0128】前記領域確率計算装置5が、求められた確率密度関数に色相値及び彩度を代入して計算することにより、入力デジタル画像のすべてのピクセルについて、口唇の色の確からしさが0～1の範囲で求められる。そして、そのピクセルの色相と彩度とにより求めた確からしさの値（0～1）をピクセルの新しい値とする口唇領域確率画像を生成する。図10は、図8に示される入力画像から生成された口唇領域確率画像であり、明るいピクセルほど口唇領域である可能性が高いことを示している。

【0129】次に、（1）最外郭フレーム、（2）最外郭フレームの重心、及び（3）網が囲む領域の重心を定義する。

【0130】最外郭フレームとは、図24に示すように、最外郭格子点を順に結んでできる輪郭のことであり、レイヤー0のことである。即ち、網の外形を表している。

【0131】最外郭フレームの重心とは、最外郭格子点の座標を平均したものである。即ち、最外郭格子点の数をn、i番目の最外郭格子点を表すベクトルをp_iとすると、重心を表すベクトルgは、次のように表される。

【0132】

【数19】

$$g = \frac{\sum_{i=1}^n p_i}{n}$$

網が囲む領域の重心とは、図25に示すように、網内部のピクセルの座標をその明るさの重みとして平均したものである。即ち、網内部にあるピクセルの数をN、網内部のピクセルjを表すベクトルをp_j、ピクセルjにおける明るさをI(p_j)とすると、網が囲む領域の重心Gは、次のように定義される。なお、Iは網内部のすべてのピクセルについて合計している。

【0133】

【数20】

$$G = \frac{\sum_{j=1}^N \{p_j \times I(p_j)\}}{N}$$

(1) 網の内部エネルギー

口唇領域を抽出するための網の内部エネルギーの定義は以下の通りである。本実施の形態の内部エネルギーは、実施の形態1の内部エネルギーに、顔幅から推定される口唇の平均的な形状と大きさとに網が収縮するようなエネルギーを加える。このエネルギーを形状エネルギーE_{form}(t,s)と呼ぶことにする。

【0134】具体的には、まず、図26に示すように、顔領域抽出用網の最外郭フレームF_fの重心g_fと、口

$$I_{mf}(x,y) = I_m(x,y) - I_f(x,y) \quad (I_m(x,y) > I_f(x,y))$$

$$= 0 \quad (I_m(x,y) \leq I_f(x,y))$$

実際の画像では、図9に示す顔領域確率画像と図10に示す口唇領域確率画像との差分をとると、図29に示す口唇領域確率差分画像が得られる。このようにして生成した口唇領域確率差分画像のピクセルの明るさに-1を掛けたものを内部格子点でのエネルギーとする。最外郭格子点のエネルギーは、顔領域確率画像のピクセルの明

唇領域抽出用網が囲む領域A_mの重心G_mを結ぶ線に直角で、重心g_fを通る直線を引く。これを顔幅直線と呼ぶ。顔幅直線と最外郭フレームF_fとの2つの交点F_a・F_bを結ぶ線分の長さを顔幅の推定値F_{ab}とする。顔幅の推定値F_{ab}に一定の比率を掛けることによって、顔幅に対する口唇の平均的な大きさ、即ち口唇の幅及び高さを推定する。

【0135】そして、図27に示すように、口唇の形状を楕円に近いものとして、重心G_mを中心とし、推定した口唇の幅を長径、推定した口唇の高さを短径とし、長径が顔幅直線と同じ傾きをもつ楕円e_mを考える。

【0136】最外郭格子点の形状エネルギーE_{form}(t,s)は、この楕円e_mの中心(重心G_m)から口唇領域抽出用網の最外郭格子点m₁の一つに向かって直線を引いたとき、その直線と楕円e_mとの交点と、最外郭格子点m₁との距離として計算される。また、内部格子点m₂についても最外郭格子点m₁と同様に、形状エネルギーE_{form}(t,s)が計算されるが、内部格子点m₂が楕円e_mの内側にあるときにはE_{form}(t,s) = 0とする。このようにして求められた形状エネルギーE_{form}(t,s)を実施の形態1の(1)式に加えたものを口唇領域抽出用網の内部エネルギーとする。

【0137】(2) 網の画像エネルギー

画像エネルギーは、口唇領域確率差分画像のピクセルの明るさを用いる。口唇領域確率差分画像とは、領域確率計算装置5で生成された口唇領域確率画像と顔領域確率画像との対応するピクセルの明るさの差をとったものである。口唇領域確率画像のピクセルが顔領域確率画像のピクセルより明るい場合には、その明るさの差をとり、そうでない場合には0とする。即ち、図28に示すように、座標(x,y)における口唇領域確率画像のピクセルの明るさをI_m(x,y)、顔領域確率画像のピクセルの明るさをI_f(x,y)、口唇領域確率差分画像のピクセルの明るさをI_{mf}(x,y)とすると、次のようになる。

【0138】

【数21】

るさに-1を掛けたものを用いる。即ち、口唇領域抽出用網の格子点p(t,s)での画像エネルギーは、以下のようになる。

【0139】

【数22】

$$E_{\text{image}}(t, s) = -\text{MFDRPI}(t, s) \quad (t > 1)$$

$$= -\text{FRPI}(t, s) \quad (t = 0)$$

ここで、MFDRPI(t,s) は口唇領域確率差分画像での格子点 p(t,s) での明るさ、FRPI(t,s) は顔領域確率画像での格子点 p(t,s) での明るさである。

【0140】(3) 移動エネルギー

口唇領域抽出用網を動かすために、網の内部エネルギー及び画像エネルギーの他に網の移動エネルギーを定義する。これは、口唇領域を抽出するための網が、精度良く口唇領域を抽出するために、網全体を口唇領域の方へ動かすためのエネルギーである。この移動エネルギー E_{move} は、以下のように定義する。

【0141】

【数23】

$$E_{\text{move}} = |G_{\text{net}} - g_{\text{net}}|^2$$

ここで、 G_{net} は口唇領域抽出用網が囲む領域の重心を表すベクトル、 g_{net} は口唇領域抽出用網の最外郭フレームの重心を表すベクトルである。即ち、 E_{move} は、2つの重心が重なるように網を移動させる力をもたらす。なお、口唇領域抽出用網の最外郭フレームの重心は初期状態では、顔領域抽出用網の最外郭フレームの重心と一致している。

【0142】次に、上記のように定義された口唇領域抽出用網が口唇領域を抽出するプロセスを図30のフローチャートに基づいて説明する。

【0143】まず、実施の形態1と同様にして顔領域抽出用網により顔領域を抽出する(S31)。従って、仮想的な網は、図31の上段に示すように、顔領域を囲むような形・大きさになっている。口唇領域は、顔領域の内部にあるのだから、顔領域が抽出されたときの網を口唇領域抽出用網の初期値として、口唇領域抽出用網を動かす始める(S32)。顔領域抽出用網のときと同じようにして、すべての格子点について8近傍を移動させ、網を1単位変形させる(S33)。次に、先に挙げた口唇領域抽出用網が囲む領域の重心 G_{net} と、口唇領域抽出用網の最外郭フレームの重心 g_{net} を計算し、それらの距離が小さくなる方向に網を移動させる(S34)。網全体のエネルギーに変化があるかどうかを調べ(S35)、S35でエネルギー変化がある場合にはS33の工程へ戻り、S35でエネルギーに変化がなくなったら網の変形を停止させる。網が停止したときに網は図31の下段に示すように、口唇領域を囲っている。

【0144】以上により、本実施の形態における領域抽出装置は、顔領域を囲む網を初期値とした口唇領域抽出用網を設定し、格子点の位置関係に基づく網の内部エネルギーと、格子点が位置するピクセルの色の確からしさに基づく画像エネルギーと、口唇領域抽出用網の最外郭フレームの重心と網が囲む領域の重心とが一致するように定義された移動エネルギーとの合計値が変化しなくな

るまで網を収縮変形・移動させ、合計値が変化しなくなった時点で格子点の移動を停止する構成である。

【0145】従って、顔領域を覆っていた網は、色の確からしさが大きい口唇領域に向かって収縮し、最終的に口唇領域を囲んで停止する。この結果、口唇領域を抽出することができる。

【0146】また、実施の形態1の効果に加えて、顔領域の最外郭フレームの重心と口唇領域抽出用網の囲む領域の重心とを結ぶ直線と、垂直軸とのなす角を求めることにより、顔の回転角を認識することができる。ここで、顔の回転とは顔の傾きのことであり、即ち、撮像装置から撮影方向に延びた軸の回りの回転を示す。

【0147】顔の回転角がわかると、例えば、人物データベースを作る際に、写真をスキャナなどで読み込む場合に、写真の方向などをきこせずに入力することができ、手間が大いに軽減する。また、同様にして、バラになった写真の方向を揃えるような装置に適用することが可能である。さらに、撮影時のブレやフィルムイメージをデジタル化するときなどの方向の間違いを気にしなくてもよくなる。この結果、撮影時、及び処理時の自由度が増大する。

【0148】さらに、顔領域と口唇領域とのバランスにより、顔のおおよその方向を推定することができる。即ち、口唇領域の中心が顔領域の軸に近い位置にあるならば、利用者は撮像装置の方向を向いており、左右のどちらかにずれていればそれぞれの方向に向いていることになる。

【0149】また、動画像から口唇領域を連続的に抽出することにより口唇の動きを検出することができ、マイクなどと組み合わせると音声認識の認識精度をあげたり、マイクが拾った音が利用者が喋っているのか別のところから聞こえてきたものかを判断したりすることが可能となる。

【0150】さらに、本領域抽出装置は、実施の形態1と同様にさまざまな機器の制御などに利用することができる。

【0151】例えば、テレビに本領域抽出装置を取り付けて、利用者の顔の向きを検出することにより、利用者がテレビの方を向いていないときだけ画面を暗くして表示面を保護したり、音センサと組み合わせ、単純な発生によってスイッチのON/OFFを行ったり、単純な機器の制御を行ったりすることができる。さらに、マイクで拾った声が利用者のものかどうかを判断できるので、誤った検出で制御されることを防ぐことができる。

【0152】〔実施の形態3〕本発明のその他の実施の形態について図32ないし図38に基づいて説明すれば、以下の通りである。なお説明の便宜上、前記の実施

の形態の図面に示した部材と同一の部材には同一の符号を付記し、その説明を省略する。

【0153】本実施の形態にかかる領域抽出装置は、入力された人物画像により SDO BIJ><DP N=0001><RTI ID 目領域を抽出するものである。ここでは、顔領域を抽出した後、その抽出された顔領域を初期状態とした口唇領域抽出用網、左目領域抽出用網、及び右目領域抽出用網を設定し、それぞれの網の位置関係を考慮に入れながらそれらの網を同時に移動させて口唇領域、及び左右の目領域を同時に抽出する。

【0154】上記領域抽出装置は、図32に示すように、実施の形態2の構成に加えて、記憶装置3に左目領域抽出用網情報格納部23、右目領域抽出用網情報格納部24、バランスエネルギー($E_{balance}$)格納部25、及び顔部品位置関係格納部26を有している。

【0155】左・右目領域抽出用網情報格納部23・24には初期状態として顔領域に収縮した網の座標がそれぞれ格納される。また、左・右目領域抽出用網の2つの重心と、網の内部エネルギー(E_{int})、画像エネルギー(E_{image})、及び形状エネルギー(E_{form})とがそれぞれ格納される。バランスエネルギー格納部25には3つの網の位置関係の制御を行うバランスエネルギー($E_{balance}$)が格納される。顔部品位置関係格納部26にはサンプリングにより予め求められた顔部品の位置関係が予め格納されている。

【0156】上記構成によれば、実施の形態2と同様に、領域確率計算装置5が領域確率画像格納部8を参照して、顔領域及び口唇領域を構成する各ピクセルの色に基づいて確率密度関数を計算することにより各ピクセルの色の確からしさを算出し、顔領域確率画像及び口唇領域確率画像を生成する。次に、エネルギー計算装置6が顔領域抽出用網の内部エネルギー(E_{int})と画像エネルギー(E_{image})を計算し、それを顔領域抽出用網情報格納部9に格納する。これらのエネルギーの合計が小さくなるように網を収縮変形・移動させ、まず顔領域を抽出する。この結果が出力装置4に出力される。

【0157】重心計算装置21が、顔領域抽出用網情報格納部9・口唇領域抽出用網情報格納部22・左目領域抽出用網情報格納部23・右目領域抽出用網情報格納部24にそれぞれ格納されている網の座標を参照して最外郭フレームの重心(g)を計算し、それぞれ格納する。また、各々の網の座標と領域確率画像格納部8に格納された顔領域確率画像とを参照して顔領域抽出用網・口唇領域抽出用網・左右の目領域抽出用網の囲む領域の重心

$$I_e(x, y) = \text{Max} (I_m(x, y), I_f(x, y))$$

このようにして生成された目領域抽出画像のピクセルの明るさを内部格子点での画像エネルギーとする。最外郭格子点の画像エネルギーは、顔領域確率画像のピクセルの明るさに-1を掛けたものを用いる。即ち、目領域抽出用網のある格子点 $p(t, s)$ での画像エネルギーは、以

(G)をそれぞれ計算し、顔領域抽出用網情報格納部9・口唇領域抽出用網情報格納部22・左目領域抽出用網情報格納部23・右目領域抽出用網情報格納部24に各々格納する。

【0158】また、エネルギー計算装置6が口唇領域抽出用網・左右の目領域抽出用網の内部エネルギー(E_{int})、画像エネルギー(E_{image})、及び形状エネルギー(E_{form})を計算し、それらを口唇領域抽出用網情報格納部22・左目領域抽出用網情報格納部23・右目領域抽出用網情報格納部24に格納する。さらに、エネルギー計算装置6が顔部品位置関係格納部26と4つの網情報格納部に格納されている重心($g \cdot G$)とを比較することによりバランスエネルギー($E_{balance}$)を計算し、格納する。

【0159】口唇領域抽出用網・左右の目領域抽出用網のエネルギーの合計値がそれぞれ小さくなるように3つの網を変形収縮・移動し、口唇領域及び左右の目領域を抽出する。その結果が出力装置4に出力される。

【0160】次に、口唇領域抽出用網、左目領域抽出用網、及び右目領域抽出用網のエネルギーについて説明する。口唇領域抽出用網のエネルギーは実施の形態2と同様である。

【0161】(1)目領域抽出用網の内部エネルギー
左右の目領域抽出用網のエネルギーの内部エネルギーについては、口唇領域抽出用網と同様に、顔幅から推定される目の平均的な形状と大きさに網が収縮するようなエネルギーを加える。即ち、顔幅に対して一定の比率を乗ずることによって得られる長径及び短径を持つ楕円を想定することによって得られる形状エネルギー E_{form} が実施の形態1の(1)式に加えられることにより目領域抽出用網の内部エネルギーが定義される。

【0162】(2)目領域抽出用網の画像エネルギー
左右の目領域抽出用網のエネルギーの画像エネルギーは、目領域抽出画像のピクセルの明るさを用いる。目領域抽出画像とは、領域確率計算装置5で生成された口唇領域確率画像と顔領域確率画像との対応するピクセルの明るさの大きいほうをとったものである。即ち、図33に示すように、座標(x, y)における口唇領域確率画像のピクセルの明るさを $I_m(x, y)$ 、顔領域確率画像のピクセルの明るさを $I_f(x, y)$ 、目領域抽出画像のピクセルの明るさを $I_e(x, y)$ とすると、次のようになる。

【0163】

【数24】

下のようになる。

【0164】

【数25】

$$E_{\text{image}}(t, s) = ERPI(t, s) \quad (t > 1)$$

$$= -FRPI(t, s) \quad (t = 0)$$

ここで、 $ERPI(t, s)$ は、目領域抽出画像での格子点 $p(t, s)$ での明るさ、 $FRPI(t, s)$ は顔領域確率画像での格子点 $p(t, s)$ での明るさである。

【0165】以上のようにして、目領域抽出画像の内部エネルギーと画像エネルギーとが定義される。

【0166】(3) 網のバランスエネルギー
これら左右の目領域抽出用網及び口唇領域抽出用網の3つの網が協調しながら動くように定義した網のバランスエネルギーについて説明する。

【0167】通常、人間の左右の目と口唇との位置関係は同じであり、個人差はあるものの例えば、両目の間の

距離や、目及び口唇の高さなどは顔幅に対して一定の比率を持っている。従って、これらの顔部品の位置関係に関する先見的知識を用いて、それぞれの領域を抽出する網の位置関係を制御する。

【0168】表1は24人のサンプルデータより測定した顔部品の大きさ及び位置関係に関するデータである。これは、図34(a)及び(b)に示すように、顔幅 h を1としたときの各測定幅 $a \sim l$ の値の平均、分散、及び標準偏差を示している。各々のデータは、図34に示す箇所を測定している。これらのデータにより算出された顔部品の位置関係は、前記顔部品位置関係格納部26に格納される。

【0169】

【表1】

測定位置	平均	分散	標準偏差
a	0.23	0.00035	0.019
b	0.07	0.00021	0.015
c	0.22	0.00038	0.020
d	0.07	0.00021	0.014
e	0.40	0.00118	0.034
f	0.15	0.00135	0.037
g	0.47	0.00043	0.021
h	1	0	0
i	0.90	0.00359	0.060
j	0.51	0.00204	0.045
k	0.86	0.00253	0.050
l	1.77	0.01192	0.109

バランスエネルギー E_{balance} は、内部エネルギー及び画像エネルギーに加えてそれぞれの網の位置関係を特定の制約条件に拘束するように定義される。バランスエネ

$$E_{\text{balance}} = E_{\text{mg}} + E_{\text{leg}} + E_{\text{reg}} + E_{\text{iod}} + E_{\text{emh}} + E_{\text{ia}} + E_{\text{ema}}$$

図35に示すように、 E_{reg} は、口唇領域抽出用網の最外郭格子点の重心 g_{r} が、その網が囲む領域の重心 G_{r} に引き寄せられるようなエネルギーである。

【0171】 E_{leg} は、左の目領域抽出用網の最外郭格子点の重心 g_{le} が、その網が囲む領域の重心 G_{le} に引き寄せられるようなエネルギーである。

【0172】 E_{reg} は、右の目領域抽出用網の最外郭格子点の重心 g_{re} が、その網が囲む領域の重心 G_{re} に引き寄せられるようなエネルギーである。なお、上記重心 g_{r} 、重心 g_{le} 、及び重心 g_{re} は、初期状態では顔領域抽出用網の最外郭格子点の重心 g_{r} に一致している。

【0173】 E_{iod} は、左右の目領域抽出用網が囲む領域の重心 G_{le} 、 G_{re} 間の距離が、顔幅に対する特定の比になろうとするエネルギーである。

【0174】 E_{emh} は、左右の目領域抽出用網が囲む領

域の重心 G_{le} と G_{re} とを結ぶ線分の中点と口唇領域抽出用網が囲む領域の重心 G_{r} との距離が顔幅に対する特定の比率になろうとするエネルギーである。

【0170】

【数26】

E_{ia} は、左右の目領域抽出用網が囲む領域の重心 G_{le} と G_{re} とを結ぶ線分の中点と口唇領域抽出用網が囲む領域の重心 G_{r} とを結ぶ軸と、顔の傾きを表す軸とのなす角が0になろうとするエネルギーである。

【0175】 E_{ema} は、左右の目領域抽出用網が囲む領域の重心 G_{le} と G_{re} とを結ぶ線分と、その線分の中点と口唇領域抽出用網が囲む領域の重心 G_{r} とを結ぶ軸とのなす角が直角になろうとするエネルギーである。

【0176】 E_{ema} は、左右の目領域抽出用網が囲む領域の重心 G_{le} と G_{re} とを結ぶ線分と、その線分の中点と口唇領域抽出用網が囲む領域の重心 G_{r} とを結ぶ軸とのなす角が直角になろうとするエネルギーである。

【0177】以上のように、口唇・左目・右目領域抽出用網の内部エネルギー及び画像エネルギーは網を収縮させ、かつ滑らかに保とうとしながら、各領域の画像の特徴にあわせて各領域に縮んで生き、バランスエネルギー

E_{balance} はそれぞれの網のバランスを一定の平行状態に保つ。3つの網がそれぞれの領域に向かって変形・収縮し、網の位置関係が一定の平衡状態を保つように移動することにより、それぞれの網は最終的に個々の領域を囲むように収束する。図36に、入力画像に対して本領域抽出装置を適用した結果を示す。同図(a)~(c)の順に、口唇、左目、及び右目領域抽出用網が各領域を抽出する様子を示している。

【0178】次に、口唇・左目・右目領域抽出用網がそれぞれの領域を抽出するプロセスを図37のフローチャートに基づいて説明する。

【0179】まず、実施の形態1と同様にして顔領域抽出用網により顔領域を抽出する(S41)。従って、図38の上段に示すように、仮想的な網は、顔領域を囲むような形・大きさになっている。口唇領域及び左右の目領域は、顔領域の内部にあるのだから、顔領域が抽出されたときの網を口唇・左目・右目領域抽出用網の初期値として、各網を動かし始める(S42)。顔領域抽出用網のときと同じようにして、口唇領域抽出用網のすべての格子点について8近傍を移動させ、網が収束しているかどうかを調べ(S43)、S43で網が収束していなければ網を1単位変形させ(S44)、S43で収束していれば次に左右の目領域抽出用網について同様に調べる(S45~S48)。

【0180】各網が1単位変形変形した後、バランスエネルギーを計算して各網を移動させる(S49)。そして、口唇領域抽出用網のエネルギーに変化があるかどうかを調べ(S50)、変化がない場合には口唇領域抽出用網の収束済フラグをONにする(S51)。同様にして、左右の目領域抽出用網についても調べる(S52~S55)。その後、すべての網のエネルギーに変化があるかどうかを調べ(S56)、エネルギー変化がある場合にはS43の工程へ戻り、すべての網のエネルギーに変化がなくなったら網の変形を停止させる。すべての網が停止したときには、図38の下3段に示すように、各網は各領域を囲っている。

【0181】以上により、本実施の形態における領域抽出装置は、顔領域を囲む網を初期値とした口唇・左右の目領域抽出用網を設定し、格子点の位置関係に基づく網の内部エネルギーと、格子点が位置するピクセルの色の確からしさに基づく画像エネルギーと、3つの領域の重心の位置関係が予め測定され記憶された口唇及び左右の目の位置関係と等しくなるように定義されたバランスエネルギーとの合計値が変化しなくなるまで網を収縮変形・移動させ、合計値が変化しなくなった時点で格子点の移動を停止する構成である。

【0182】従って、顔領域を覆っていた網は、色の確からしさが大きい口唇・左右の目領域に向かってそれぞれ収縮し、最終的に各領域を囲んで停止する。この結果、口唇及び左右の目領域を抽出することができる。

【0183】これにより、実施の形態1及び2の効果に加えて、口唇・目の形を認識することができ、これらは個人によって特有な値を持つので、これを利用して個人同定を行うことができる。即ち、予め複数の人物の画像を入力しておき、領域抽出装置により各領域を抽出、それぞれの部品の相対位置関係や大きさ、形状を数値化して、各人ごとにデータベースに記憶しておく。データベース検索装置に撮像装置と領域抽出装置を取り付けることにより、撮像装置に映された人物の顔領域、口唇領域、目領域を抽出して、同じように各部品の位置関係や大きさを数値化する。得られた数値とデータベースの数値とでマッチングを取り、入力画像がデータベース中のどの人物であるかを判断する。これにより、来客の多い企業の受付やホテルなどで予めデータベースを作っておき、来客人物を撮影してデータベースからその人物のプロフィールなどを自動検索し、それによってスムーズな対応をするアプリケーションを想定することができる。

【0184】また、顔輪郭だけでなく、顔部品の抽出まで行えるようになると、情報の圧縮率はさらに高まることになる。実施の形態1で述べたように、顔領域を抽出し、顔以外の解像度や色数を落としたりすることで情報を圧縮できるが、さらに、顔領域の中で、顔部品の領域を最も高い解像度・色数にし、次いで顔領域、顔領域以外、という順序で解像度・色数を落とすことにより、画像の顔領域の劣化を少なく保ちつつ情報を圧縮することが可能となる。さらに、顔部品及び顔領域で利用される色数は限られており、それぞれ独自の色コードを持つことにより、さらなる情報の圧縮が可能となる。これにより、実施の形態1や2よりも少ない容量でより多くの画像を保存でき、あるいは、狭い帯域幅でよりコマ数の多い滑らかな画像を電送することが可能となる。

【0185】また、人物データベースなどのように、画像を多くファイリングする必要があるアプリケーションにおいては、ひとつひとつの画像の大きさがその装置に記憶することのできる画像の数に大きく関わってくる。上述のように、1枚の人物画像をより少ない容量で保存することができるなら、低コストで多くの画像を記憶することのできるデータベースを作成することができる。また、帯域幅を広くしなくとも人物画像を電送できるようになる。従って、テレビ会議、テレビ電話、及び電子秘書などの低コスト化が実現できる。

【0186】また、人物データベースを作成する際に、今までは画像と一緒に検索のためのキーとなるその個人を表すデータを人間が手作業で単語として登録し、検索する際もそのキーワードを手で入力したり、あるいは、データベースを最初から順番に見つかるまで方端から見なければならなかったが、本領域抽出装置では画像から顔領域及び顔部品が自動的に抽出できるので、データベースに登録する際に顔部品を抽出して、装置自身がデータベース検索装置に接続された撮像装置に映すだけでよ

くなる。この結果、来客や久しぶりに会った人が誰であるかを瞬時に見つけ出すといったアプリケーションが可能となる。

【0187】さらに、目や口唇の領域が抽出されると、表情の認識を行うことが可能となる。人間の表情は、目や口などの顔部品に特徴が現れるため、目や口の形を領域抽出することにより調べることによって、その人がどのような表情をしているのかを判定できるようになる。これは、人の感情にあわせて対応を変化させるロボットや精神分析などの自動化などを行うことが可能にする。

【0188】同様に、表情の認識ができるということは、表情の生成をもできるということになる。目や口の領域が抽出されると、その領域を感情特有の形に変形させることにより表情を作り出せる。例えば、人が笑っているときの典型的な目と口の形状を予め測定してデータベース化しておき、入力画像から抽出された目や口の領域をそのデータベースにしたがって変形させれば、その画像中の表情を笑わせることができる。さまざまな表情を予めデータベース化しておけば、撮像した画像の顔を自由自在に笑わせたり、怒らせたりといったことが可能となる。

【0189】これは、人間的なやさしいインターフェイスを持った機器を開発することを可能にする。従って、機器の操作説明などで現れる顔の表情を自在に変化させ、利用者にとって親しみ易い機器の提供などを行うことが可能となる。また、ある人間を別の人間、あるいは、他の動物などに連続的に変化させるアニメーションなど、映像作品における特殊効果的な利用もこの技術を用いれば全自動で行うことができる。来るべきマルチメディア時代のエンターテインメントコンテンツとしての利用価値は大きい。

【0190】また、表情生成、似顔絵作成などが自動的にできるようになることは、エンターテインメント分野でのマルチメディアコンテンツ供給の低コスト化を実現できる。従来では手作業と独特のセンス、技術を要求されるものであったものが、自動的にできるようになるため、将来的に増大するであろう需要に対しても対応することが可能となる。

【0191】また、実施の形態1及び2と同様に、リアルタイムに顔領域及び顔部品を抽出することにより、より細かい機器の制御を行うことが可能となる。例えば、顔部品の位置関係により顔の傾きを精度良く検出することが可能となる。顔の傾きの大きさによって、オーディオ装置の音量をコントロールしたり、照明の明るさを制御したりすることが可能となる。あるいは、瞬きなどを検出して両手が使えない状態、例えば、車の運転中にラジオのチューニングを変えるということが可能となる。

【0192】あるいは、顔部品を抽出し個人同定を行うことにより、個人の好みに合わせたエアコンの強弱やテレビ、オーディオなどの音量などの自動制御や、パソコ

ンのログインや電子ロックといったセキュリティ方面への応用も可能となる。これにより、そのつど機器を調整しなければならない手間やパスワード入力をする手間、入力間違いなどが削減される。

【0193】なお、上記実施の形態1ないし3では、領域確率計算装置5により一旦領域確率画像を生成し、この領域確率画像を用いて画像エネルギーを計算しているが、これは、処理が簡単になると共に処理が高速化できるという効果がある。しかしながら、画像エネルギーを計算する際に、格子点上にあるピクセルの色から各領域の確からしさをそのつど計算することによっても画像エネルギーを計算することができる。この場合は、領域確率画像を生成しなくてもよい必要なメモリを減らすことができるというメリットがある反面、格子点上のピクセルの領域の確からしさをそのつど計算する必要があるので処理速度が低下する。

【0194】また、上記実施の形態の網のエネルギーは網の形が顔領域及び顔部品を囲むようになったときにエネルギーが最も小さくなるように定義されているが、網の形が顔領域及び顔部品を囲むようになったときにエネルギーが最も大きくなるように定義しても同様の効果が得られる。

【0195】また、本実施の形態の網の形状は放射状としているが、エネルギーの設定次第で格子状など他の形状とすることもできる。即ち、点と点との接続関係の情報があればどのような形状とすることもできる。ただし、本実施の形態のように楕円に近い形状を持つ顔、口唇、目領域などを抽出する場合にはエネルギー設定が行いやすいので放射状が好ましい。

【0196】また、本実施の形態では、格子点を移動させる際に最外郭格子点から移動させているが、これに限られることはなく、どの格子点から移動させてもよい。ただし、すべての格子点が1単位移動するまで、同じ格子点が二重に動かないようにする必要がある。

【0197】また、本実施の形態では格子点の移動を8近傍としているが、これに限られることはなく、例えば4近傍としてもよい。さらに、隣接するピクセルではなく、離れたピクセルに移動してもよいが、この場合は、抽出したい領域の境界部分を飛び越えてしまわない程度の移動距離にする必要がある。

【0198】なお、実施の形態2及び3における内部エネルギーには、楕円を想定することにより得られる形状エネルギーが設定されているが、抽出しようとする物の形状がわからない場合には設定しなくともよい。ただし、本実施の形態のように、抽出対象の形状がわかっている場合は、形状エネルギーを設定した方が精度があがるので望ましい。

【0199】また、実施の形態2における移動エネルギーについても同様に設定しなくともよいが、設定したほうが精度が良くなるので望ましい。

【0200】〔実施の形態4〕本発明のその他の実施の形態について図39ないし図43に基づいて説明すれば、以下の通りである。なお説明の便宜上、前記の実施の形態の図面に示した部材と同一の部材には同一の符号を付記し、その説明を省略する。

【0201】本実施の形態に係る方向検出装置は、実施の形態3の領域抽出装置を利用すると共に2つの撮像装置を用いて、各撮像装置に映った人物画像よりそれぞれ顔部品を抽出した後、各画像に映っている顔部品の位置のずれを調べることにより、頭部の3次元方向を検出するものである。

【0202】図39に示すように、機器43に取り付けられた撮像装置（撮像手段）42L・42Rが利用者の姿を捉え、そのデジタル画像を領域抽出装置41に送る。領域抽出装置41は、それぞれの画像から利用者の顔領域を抽出し、その情報を頭部方向検出装置45に送る。頭部方向検出装置45は2つの画像から顔部品の3次元情報検出することにより頭部の方向を検出し、その情報を制御装置44に送る。制御装置44は機器43に接続されており、頭部方向検出装置45から送られてきた情報をもとに機器43を制御する。

【0203】上記2つの撮像装置42L・42Rを用いて頭部方向を検出する方法は以下の通りである。まず、撮像装置42L・42Rの配置はエピソード拘束に従う。エピソード拘束とは、図40に示すように、一方の撮像装置42Lのデジタル画像Lに映っているものが、他方の撮像装置42Rのデジタル画像Rにて、走査線に平行に映っているような制約のことである。即ち、撮像装置を2台設置する際には、同様の特性を持つ撮像装置を光軸が同一平面上にのるように設置しなければならない。

【0204】まず、図41に示すように、2つの人物画像である画像L・Rからそれぞれ口唇領域、左右の目領域を抽出する。次に、画像Lにおける3つの顔部品が画像Rにおいてどの位置にあるかを調べる。画像Lにおける口唇位置の、画像Rにおける口唇位置の水平方向のずれを d_m 、画像Lにおける左目位置の、画像Rにおける左目位置の水平方向のずれを d_{le} 、画像Lにおける右目位置の、画像Rにおける右目位置の水平方向のずれを d_{re} とする。これらの2つの画像L・Rでの顔部品のずれの逆数が、画像面からの顔部品の相対的な距離に相当する。

【0205】そこで、図42に示すように、画像面の各顔部品から画像面奥の方へ垂直に向かってそれぞれ $1/d_m$ 、 $1/d_{le}$ 、 $1/d_{re}$ だけ離れた距離にある点をそれぞれ p_m 、 p_{le} 、 p_{re} とすると、頭部の3次元向きは p_{le} と p_{re} とを結ぶ線分の中点と、 p_m とを通る直線で p_m から中点に向かう方向となる。

【0206】上記方向検出装置のプロセスを図43のフローチャートに基づいて説明する。まず、2つの撮像装

置42L・42Rにより利用者の画像を撮影する（S61）。そして、領域抽出装置41により各画像から口唇・左右目領域を抽出する（S62）。2つの画像間での口唇・左目・右目それぞれの顔部品の水平位置のずれを算出する（S63）。各顔部品の画像面からの距離を計算する（S64）。そして、頭部の3次元方向を計算する（S65）。このプロセスを繰り返すことにより、撮像装置前の利用者の頭部方向をリアルタイムに検出することが可能となる。

【0207】即ち、2つの撮像装置42L・42Rにより顔画像を同時に撮影し、それぞれの画像の顔領域及び顔部品を抽出することにより、それぞれの画像の視差により口唇、左目、右目、それぞれの3次元空間上の3点が決まるため、頭部がどの方向を向いているかを検出することができる。

【0208】これにより、専用の装置を頭部に装着しなれば頭部の方向を検出できなかった従来のものと比較して、本方向検出装置は専用のハードウェアを用いなくてもよいので、低コストで製品が実現できると共に装着する際の煩わしさなどを除去することができる。また、専用のハードウェアを用いなくても済むので応用範囲が広がる。

【0209】また、顔の方向を検知して利用者が自分の方を向いているときだけ作動するような仕組みを提供したり、例えば、顔の向きでチャンネルや音量が変更されるようなテレビやオーディオ製品が可能となる。

【0210】また、バーチャルリアリティに応用することも可能である。即ち、撮像装置の前の人物の頭の位置・方向に合わせて、コンピュータ上で表現された仮想世界での仮想的な利用者の位置・方向を決めることができ、実際の利用者の頭部の動きと、仮想世界での利用者の頭部の動きを、特殊なヘッドギアなどを用いずに同調させることが可能となる。また、3次元CADにおいて、特定の3次元位置の点を指定したり、複数の属性を持つものを同時に調節、例えば、テレビの色合い、コントラスト、明るさなどを同時に調整することも可能となる。

【0211】なお、撮像装置は2つに限られることはなく、3つ以上とすることも可能である。この場合、撮像装置の配置はエピソード拘束に従う必要はない。

【0212】〔実施の形態5〕本発明のその他の実施の形態について図44ないし図46に基づいて説明すれば、以下の通りである。なお説明の便宜上、前記の実施の形態の図面に示した部材と同一の部材には同一の符号を付記し、その説明を省略する。

【0213】本実施の形態にかかる方向検出装置は、図44に示すように、実施の形態4の構成に加えて、頭部方向検出装置45により検出された頭部方向の情報が入力される視線方向検出装置46を有する。

【0214】視線方向を検出する方法は以下の通りであ

る。まず、実験として、視線の検出精度を設定する。その精度から検出しようとしている方向の数を決め、例えば図45(a)~(f)に示すように、その数だけ被験者にその方向を注視してもらおう。そして、その注視している時の目の画像を撮影する。この目の画像は、人間がある方向を見た時に、瞳がどの位置にくるかを示している視線データベースとなる。領域抽出装置によって検出された目の画像を、この視線データベースと比較することにより、利用者の視線の方向を推定することができる。そして、実施の形態4の頭部方向検出装置45で検出された頭部の3次元方向の情報と、この視線方向の情報とを組み合わせることで、真の視線方向を検出することが可能となる。

【0215】上記方向検出装置のプロセスを図46のフローチャートに基づいて説明する。まず、頭部検出装置により、頭部の3次元方向を検出する(S71)。この時点で目の領域はすでに検出されているため、抽出された目の領域を視線データベースにある画像にあわせてサイズを変更する。そして、変更された目の領域の画像と、視線データベースの画像との間でマッチングをとる(S72)。このマッチングにはテンプレートマッチングなどの手法を用いる。そして、2つの画像の差を取り、差が最も小さくなるものを選択することにより、視線方向を推定する(S73)。すでに検出されている頭部の3次元方向と、推定された視線方向とを合わせることで、真の視線方向を検出する(S74)。

【0216】これにより、本方向検出装置は、特別な装置を用いずに視線方向を検出することができる。この結果、ハードウェアなどを装着する煩わしさを除去することができると共に、装置の低コスト化が可能となる。また、情報機器などのポインティングデバイスの代用、例えば、コンピュータディスプレイの任意の一点をマウスやペン、あるいは、現存するアイトラッカーのような大がかりな装置を使用することなく指定することができ、さまざまな電子機器の制御を行うことが可能となる。また、キーボードを操作しながら注目点を目で追跡したりすることができ、機器の作業の効率をあげることができる。さらに、両手が塞がった状態でも他の操作を行うことが可能となる。また、体の不自由な人への適用も考えられる。

【0217】

【発明の効果】以上のように、本発明の請求項1記載の領域抽出装置は、対象物を撮像してカラーデジタル画像とする撮像手段と、上記カラーデジタル画像において均一に同様の色を持つ所望領域を示す確率密度関数を予め記憶しておく確率密度関数格納手段と、上記所望領域を構成する各画素の色に基づいて上記確率密度関数を計算することにより、各画素の色の確からしさを算出する領域確率計算手段と、上記カラーデジタル画像上の任意の3つ以上の点を接続することにより形成され画像全体を

覆うように設けられる仮想的な網の情報を記憶しておく網情報格納手段と、上記点の位置関係に基づいて定義される網の内部エネルギーと、点が位置する画素の色の確からしさに基づいて定義される網の画像エネルギーとをそれぞれ計算するエネルギー計算手段とを有し、1つの点を移動させたときの網の内部エネルギーと画像エネルギーとの合計値を、上記点を移動させる前の合計値と比較し、合計値が変化するときは網が収縮する方向へ点の移動を行い、合計値が変化しなくなった時点で点の移動を停止する構成である。

【0218】これにより、本領域抽出装置は照明条件の変化に強い色の確からしさに基づいて所望領域を抽出することができるので、撮影する場所、時間、補助照明の有無、方向などに左右されることがなくなる。また、網はどのような形状にもなり得るので、入力画像の大きさ、及び画像中の顔位置や大きさなどに制限されることがなくなる。これらの結果、撮影の自由度を増大させることができるという効果を奏する。

【0219】請求項2記載の領域抽出装置は、請求項1記載の構成に加えて、上記対象物は人物であり、上記所望領域は人物の顔の領域である構成である。

【0220】これにより、顔領域を抽出することができるので、さまざまな家電製品などの機器の制御に利用することができる。即ち、利用者の位置による機器の制御を行うことが可能となるという効果を奏する。

【0221】請求項3記載の領域抽出装置は、請求項1または2記載の構成に加えて、確率密度関数格納手段には上記所望領域と異なる色を持つ部分領域を示す確率密度関数が予め記憶されると共に、網情報格納手段には上記所望領域を囲んで停止した網の情報が部分領域を抽出するための部分領域抽出用網の初期値として記憶され、上記所望領域内に部分領域が1つある場合に、領域確率計算手段が上記部分領域を構成する各画素の色に基づいて上記確率密度関数を計算することにより各画素の色の確からしさを算出して、エネルギー計算手段が部分領域抽出用網の内部エネルギー及び画像エネルギーと、部分領域に収束されるように定義された移動エネルギーとをそれぞれ計算することにより、部分領域を抽出する構成である。

【0222】これにより、最初の領域抽出に用いた網を、さらに領域内の別の領域の抽出に用いることができるので、領域毎に網の設定を変える必要がないという効果を奏する。

【0223】請求項4記載の領域抽出装置は、請求項3記載の構成に加えて、部分領域抽出用網の外周を形成する点からフレーム重心を求めると共に、画像の明るさを重みとした部分領域抽出用網が囲む領域の重心を求める重心計算手段を有し、上記移動エネルギーを、上記フレーム重心が領域の重心の方向に引き寄せられるように定義する構成である。

【0224】これにより、フレーム重心が領域重心の方向に引き寄せられるように移動エネルギーが設定されているので、部分領域に似た色を持つ領域があったとしても、その領域に引き寄せられることなく、部分領域を抽出することができる。この結果、処理速度が速く、高精度の部分領域の抽出が可能となるという効果を奏する。

【0225】請求項5記載の領域抽出装置は、請求項2記載の構成に加えて、網の外周を形成する点からフレーム重心を求めると共に、画像の明るさを重みとして網が囲む領域の重心を求める重心計算手段を有し、確率密度関数格納手段には口唇領域を示す確率密度関数が予め記憶されると共に、網情報格納手段には上記顔領域を囲んで停止した網の情報が口唇領域及び左右の目領域を抽出するための網の初期値として記憶され、領域確率計算手段が、上記口唇領域を構成する各画素の色に基づいて上記確率密度関数を計算することにより各画素の色の確からしさを算出すると共に、上記左右の目領域を構成する各画素の色に基づいて顔領域あるいは口唇領域の確率密度関数を計算することにより各画素の色の確からしさを算出して、エネルギー計算手段が、各領域の内部エネルギー及び画像エネルギーと、各々のフレーム重心が各々の領域重心の方向へ引き寄せられ、かつ、3つの領域の重心の位置関係が予め測定され記憶された口唇及び左右の目の位置関係と等しくなるように定義されたバランスエネルギーとを計算することにより、口唇及び左右の目領域を抽出する構成である。

【0226】これにより、顔領域抽出用網を、口唇領域抽出用網及び左右の目領域抽出用網とすることができるので、領域毎に網の設定を変える必要がない。また、顔領域に対する顔部品の位置がわかるので、顔の回転角（顔の傾き）を認識することができ、撮影時、及び処理時の自由度が増大する。さらに、顔のおおよその方向を推定することができるという効果を奏する。

【0227】請求項6記載の方向検出装置は、請求項5記載の領域抽出装置を用いて、上記撮像手段が複数個設けられ、複数の撮像手段から同時に得られる複数の画像から口唇領域及び左右の目領域をそれぞれ抽出し、互いに比較することにより口唇及び目の3次元位置情報を導出する構成である。

【0228】これにより、従来のような専用のハードウェアを用いなくとも、容易に頭部方向を検出することができる。この結果、ハードウェアなどを装着する煩わしさを除去することができると共に、装置の低コスト化が可能となるという効果を奏する。

【0229】請求項7記載の方向検出装置は、請求項6記載の構成に加えて、抽出された目の画像と、予め測定して視線方向の情報が付加されて記憶された目の画像とを比較することにより推定視線方向を算出し、さらに推定視線方向と検出された頭部の向きとを比較することにより真の視線方向を検出する構成である。

【0230】これにより、従来のような専用のハードウェアを用いなくとも、容易に視線方向を検出することができる。この結果、ハードウェアなどを装着する煩わしさを除去することができると共に、装置の低コスト化が可能となるという効果を奏する。

【図面の簡単な説明】

【図1】本発明の実施の一形態にかかる領域抽出装置の構成を示すブロック図である。

【図2】上記領域抽出装置における領域確率計算装置の動作を示すフローチャートである。

【図3】(a)は入力画像を示し、(b)は入力画像から顔の肌色部分と口唇部分とを取り出した状態を示す説明図である。

【図4】顔領域の色相の色度数分布を示すグラフである。

【図5】顔領域の彩度の色度数分布を示すグラフである。

【図6】顔領域を表す2次元の正規確率密度関数を示すグラフである。

【図7】領域確率画像の生成の過程を表すフローチャートである。

【図8】入力装置に入力される画像を表し、ディスプレイ上に表示された中間調画像を示す図面代用写真である。

【図9】図8に示された入力画像から生成された顔領域確率画像を表し、ディスプレイ上に表示された中間調画像を示す図面代用写真である。

【図10】図8に示された入力画像から生成された口唇領域確率画像を表し、ディスプレイ上に表示された中間調画像を示す図面代用写真である。

【図11】網を形成する格子点を示す説明図である。

【図12】網のモデルを用いて領域を抽出する様子を示す説明図であり、(a)は入力画像、(b)は網の初期状態、(c)は収束中の網の状態、(d)は収束後の網の状態をそれぞれ示す。

【図13】網のモデルを用いて顔領域を抽出する様子を示す説明図であり、(a)は入力画像、(b)は網の初期状態、(c)は収束中の網の状態、(d)は収束後の網の状態をそれぞれ示す。

【図14】網の形状を示す説明図である。

【図15】他の網の形状を示す説明図である。

【図16】(a)(b)は、網の内部エネルギーを構成する各エネルギーの定義を示す説明図である。

【図17】格子点の移動の仕方を示すフローチャートである。

【図18】あるピクセルの8近傍を示す説明図である。

【図19】上記領域抽出装置を機器に利用したときの構成を示すブロック図である。

【図20】本発明の他の実施の形態にかかる領域抽出装置の構成を示すブロック図である。

【図 2 1】口唇領域の色相の色度数分布を示すグラフである。

【図 2 2】口唇領域の彩度の色度数分布を示すグラフである。

【図 2 3】口唇領域を表す 2 次元の正規確率密度関数を示すグラフである。

【図 2 4】最外郭フレーム及び最外郭フレームの重心を示す説明図である。

【図 2 5】網が囲む領域の重心を示す説明図である。

【図 2 6】形状エネルギーの定義に必要な各パラメータを示す説明図である。

【図 2 7】形状エネルギーの定義を示す説明図である。

【図 2 8】口唇領域確率差分画像の求め方を示す説明図である。

【図 2 9】図 9 に示す顔領域確率画像と図 10 に示す口唇領域確率画像とにより算出された口唇領域確率差分画像を表し、ディスプレイ上に表示された中間調画像を示す図面代用写真である。

【図 3 0】口唇領域抽出用網が口唇領域を抽出する過程を示すフローチャートである。

【図 3 1】口唇領域抽出用網が口唇領域を抽出する過程を示す模式図である。

【図 3 2】本発明のその他の実施の形態にかかる領域抽出装置の構成を示すブロック図である。

【図 3 3】目領域抽出画像の求め方を示す説明図である。

【図 3 4】顔部品の高さ及び位置関係に関するデータを測定したときのデータ測定箇所を示す説明図である。

【図 3 5】バランスエネルギーの定義に必要な各パラメータを示す説明図である。

【図 3 6】(a) は口唇領域抽出用網、(b) は左目領域抽出用網、及び(c) は右目領域抽出用網がそれぞれ収縮する様子を示す説明図である。

【図 3 7】口唇及び左右の目領域抽出用網が各領域を抽出する過程を示すフローチャートである。

【図 3 8】口唇及び左右の目領域抽出用網が各領域を抽出する過程を示す模式図である。

【図 3 9】本発明のその他の実施の形態にかかる、領域抽出装置を利用した方向検出装置の構成を示すブロック図である。

【図 4 0】エビポーラ拘束の状態を示す説明図である。

【図 4 1】2 つの画像における顔部品のずれを算出する過程を示す説明図である。

【図 4 2】上記ずれの値から算出される顔部品の 3 次元位置情報を示す説明図である。

【図 4 3】上記方向検出装置の頭部検出動作を示すフローチャートである。

【図 4 4】本発明のその他の実施の形態にかかる、領域抽出装置を利用した方向検出装置の構成を示すブロック図である。

【図 4 5】(a) ~ (f) は、視線データベースの一例を示す説明図である。

【図 4 6】上記方向検出装置の視線検出動作を示すフローチャートである。

【図 4 7】従来の投影による抽出手法を示すものであり、(a) は図 8 に示す入力画像を水平方向に微分した状態を示す説明図、(b) は(a) の画像を水平方向へ投影したときのピクセル数をカウントしたグラフ、(c) は(a) の画像を垂直方向へ投影したときのピクセル数をカウントしたグラフをそれぞれ示す。

【図 4 8】上記水平方向の微分処理を示す説明図である。

【図 4 9】従来のテンプレートによる抽出方法に用いられる入力画像とテンプレートとを示す説明図である。

【図 5 0】5 × 5 の大きさのテンプレートと入力画像との比較の状態を示す説明図である。

【図 5 1】テンプレートと入力画像との比較の順序を示す説明図である。

【図 5 2】従来の Deformable template による抽出方法を用いて目を抽出する様子を示す説明図である。

【図 5 3】上記抽出方法におけるテンプレートの定義に必要な各パラメータを示す説明図である。

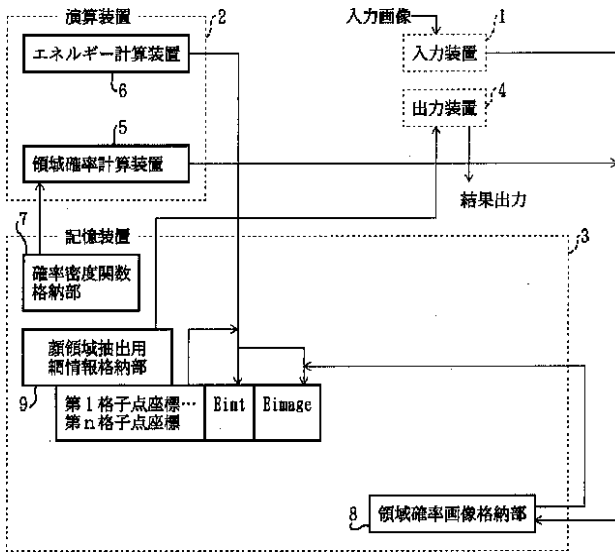
【図 5 4】従来の頭部方向の検出方法を示す説明図である。

【図 5 5】従来の他の頭部方向の検出方法を示す説明図である。

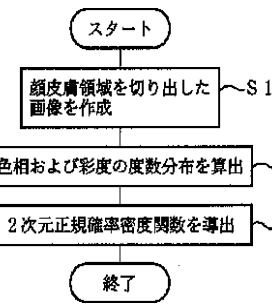
【符号の説明】

- 1 入力装置 (撮像手段)
- 5 領域確率計算装置 (領域確率計算手段)
- 6 エネルギー計算装置 (エネルギー計算手段)
- 7 確率密度関数格納部 (確率密度関数格納手段)
- 9 顔領域抽出用網情報格納部 (網情報格納手段)
- 2 1 重心計算装置 (重心計算手段)
- 2 2 口唇領域抽出用網情報格納部 (網情報格納手段)
- 2 3 左目領域抽出用網情報格納部 (網情報格納手段)
- 2 4 右目領域抽出用網情報格納部 (網情報格納手段)

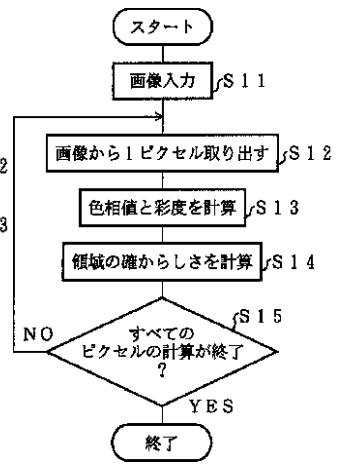
【図1】



【図2】



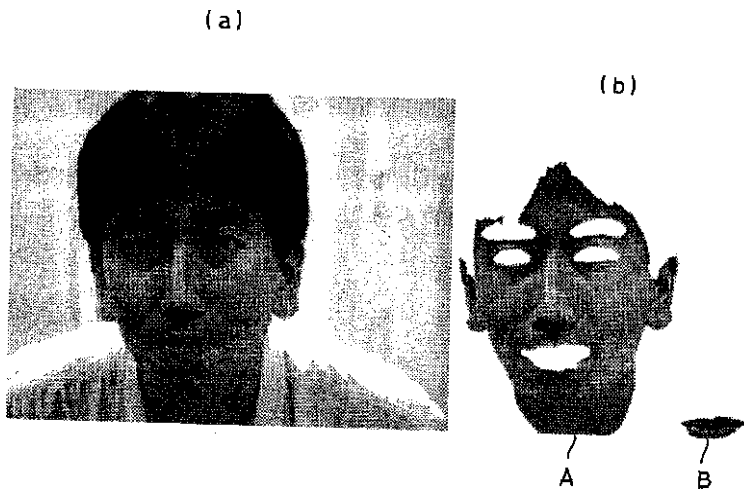
【図7】



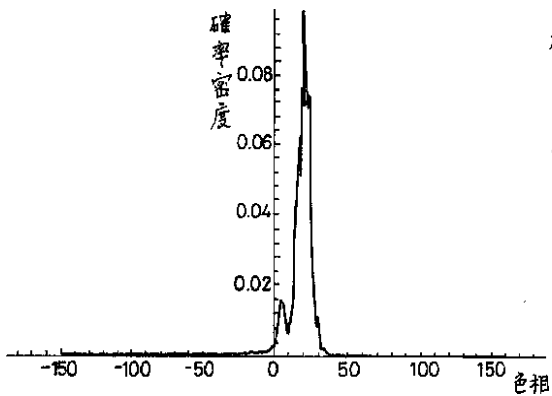
【図11】

【図18】

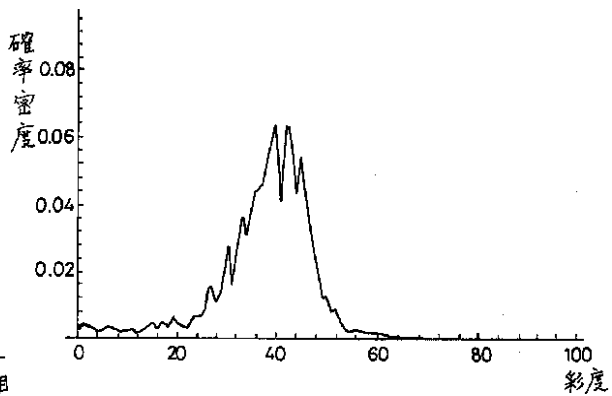
【図3】



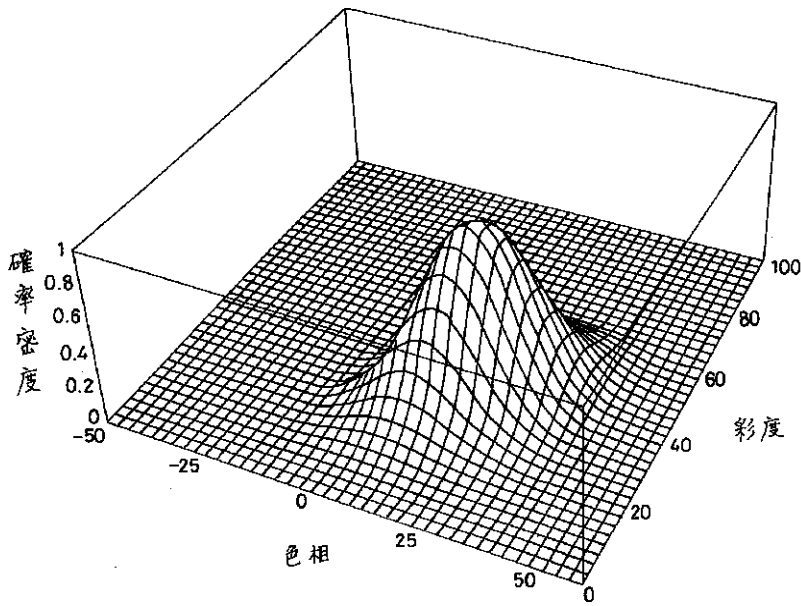
【図4】



【図5】

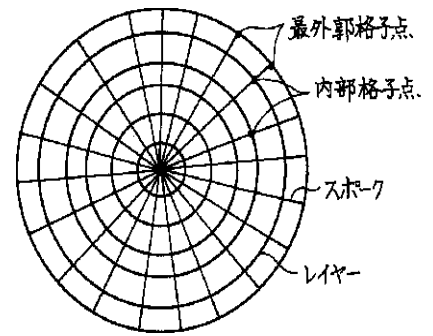


【図6】

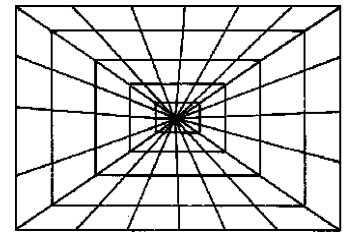


【図8】

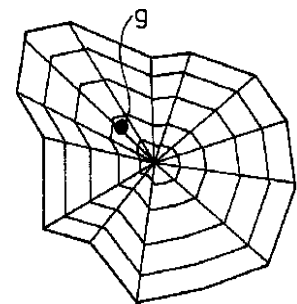
【図14】



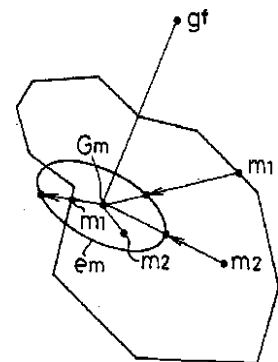
【図15】



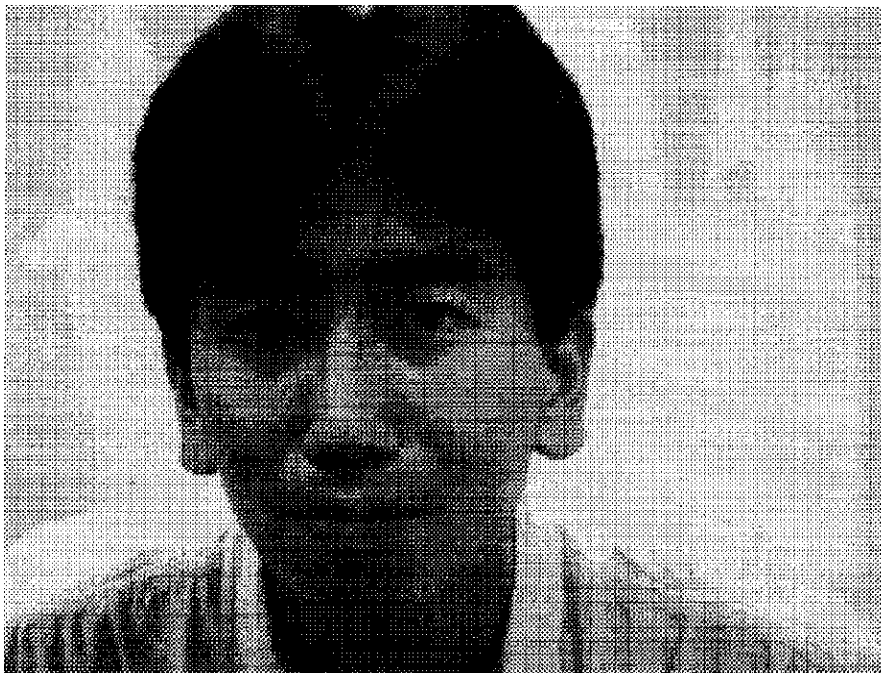
【図24】



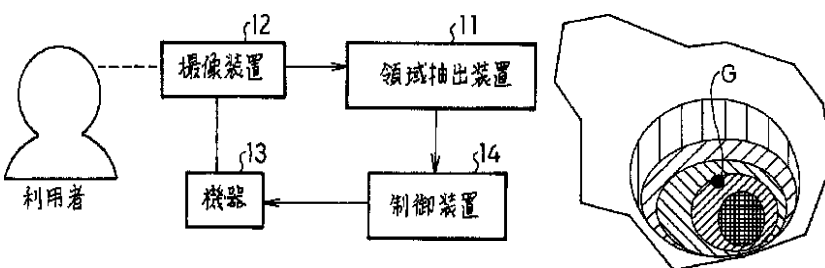
【図27】



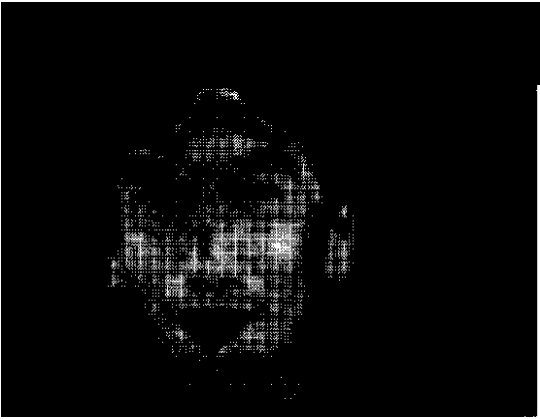
【図25】



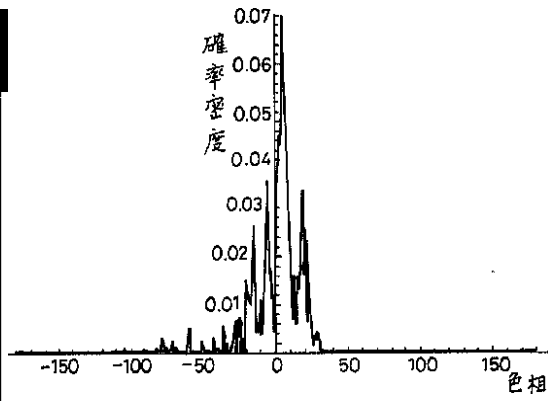
【図19】



【図 9】



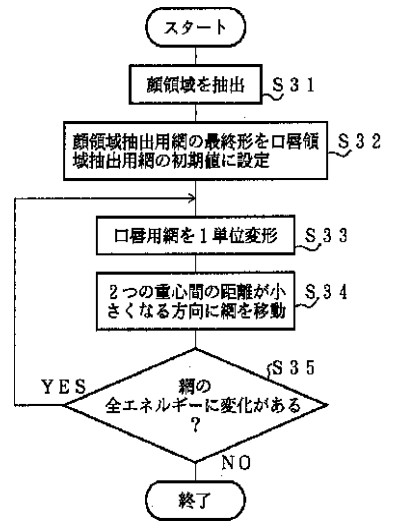
【図 2 1】



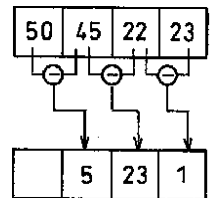
【図 1 0】



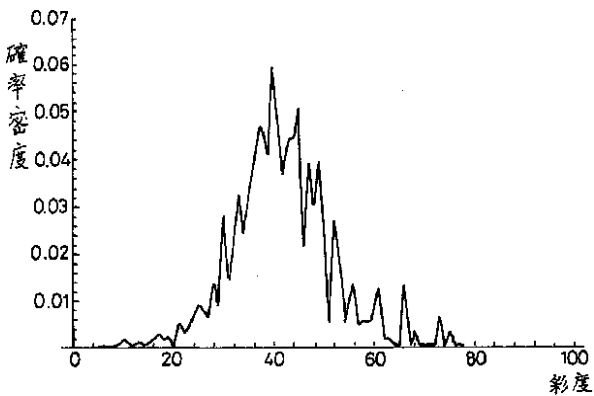
【図 3 0】



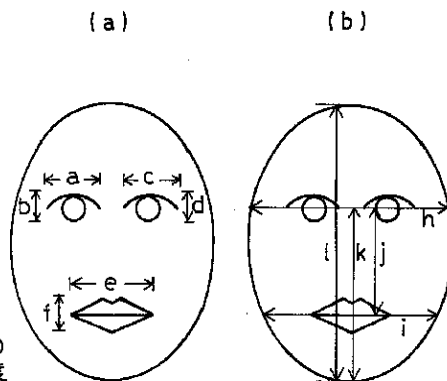
【図 4 8】



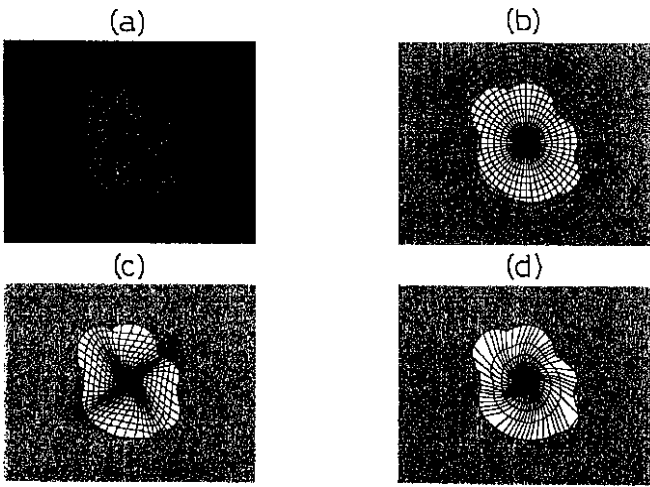
【図 2 2】



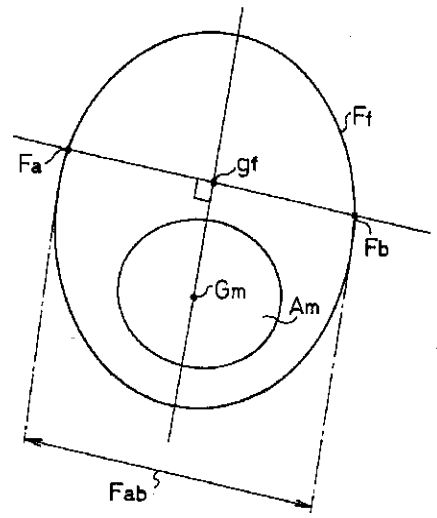
【図 3 4】



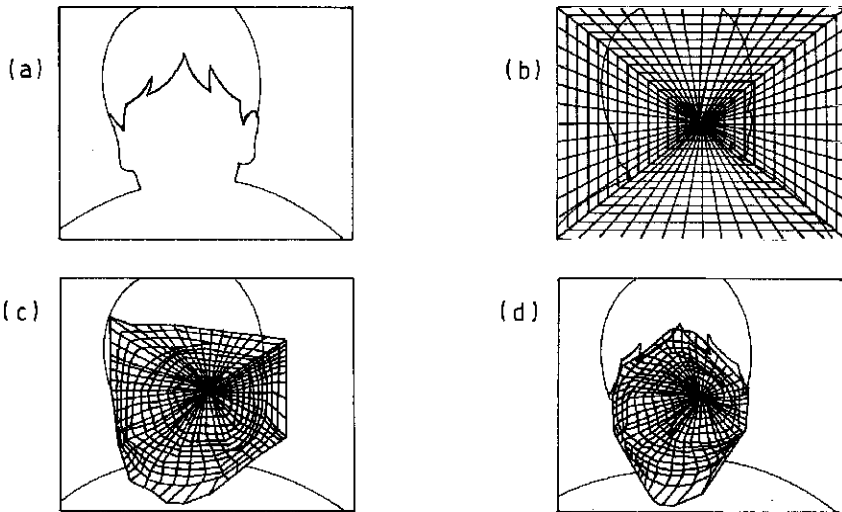
【图 1 2】



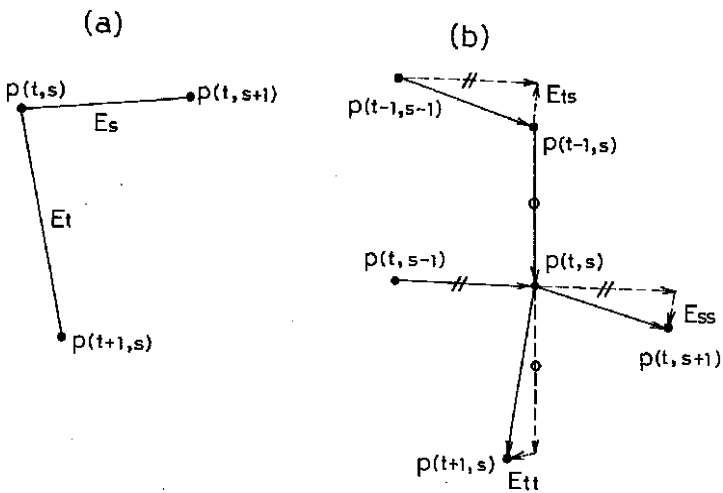
【图 2 6】



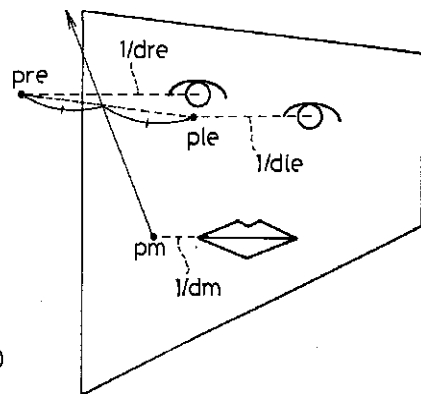
【图 1 3】



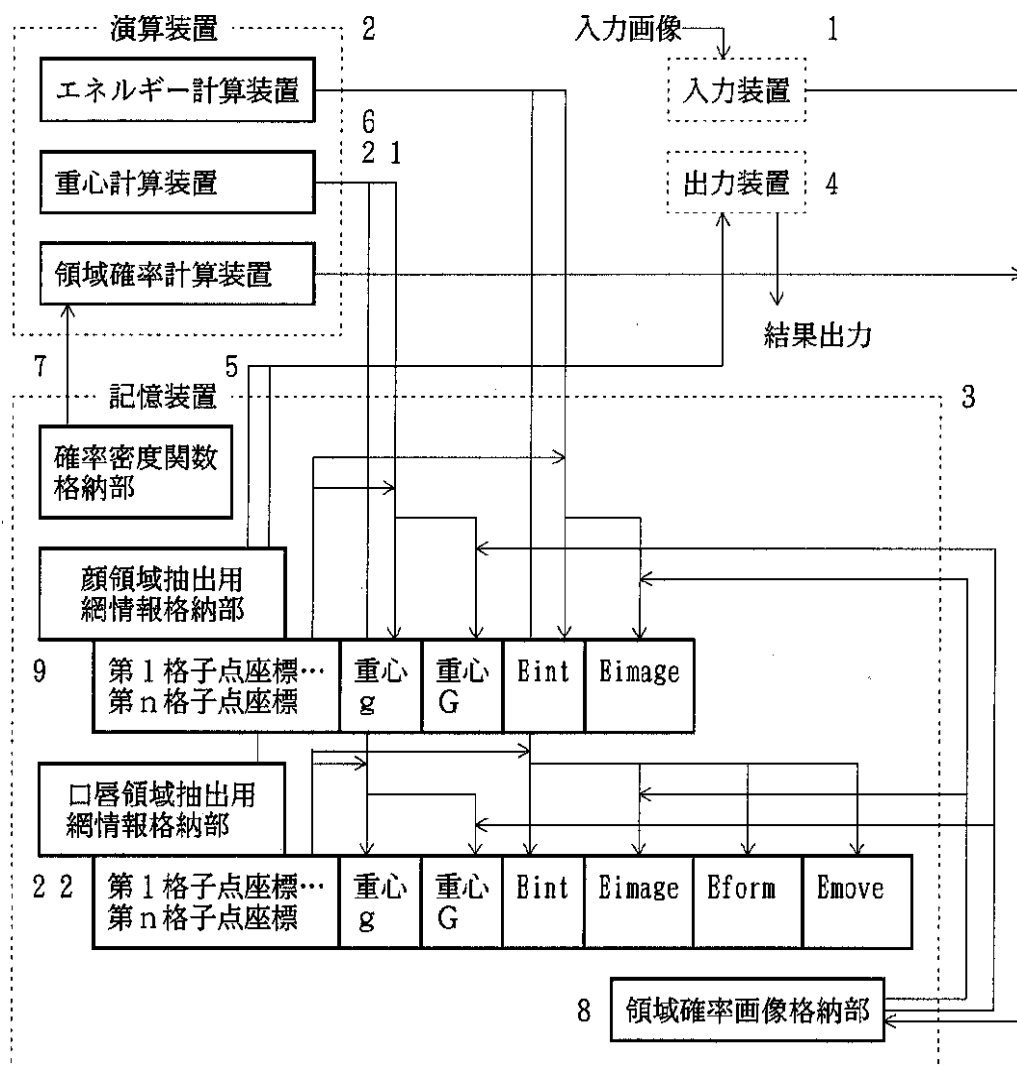
【图 1 6】



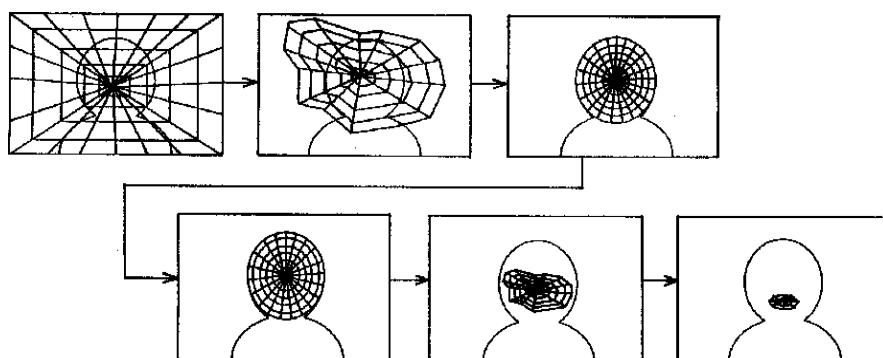
【图 4 2】



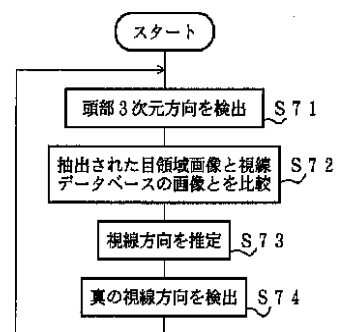
【図20】



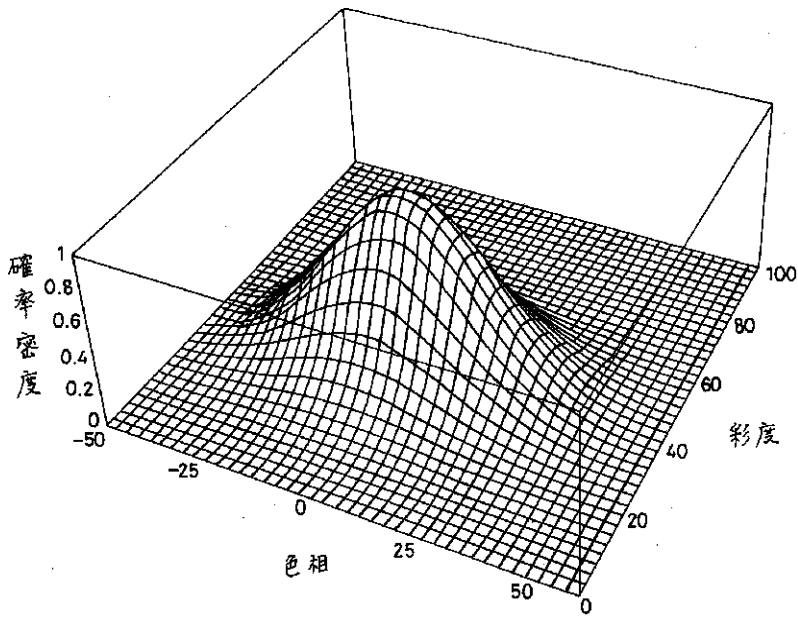
【図31】



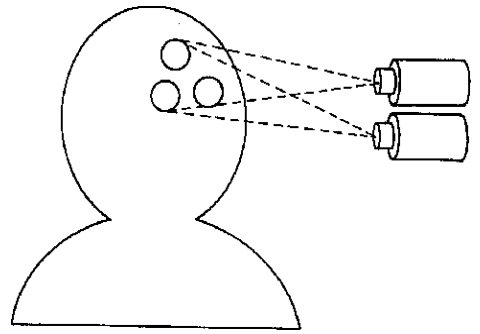
【図46】



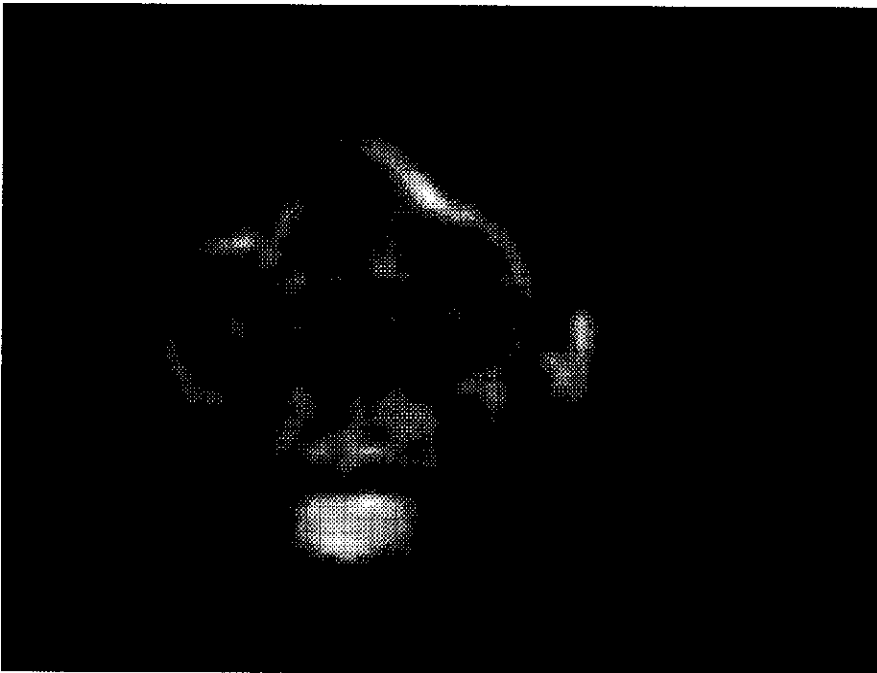
【図23】



【図55】

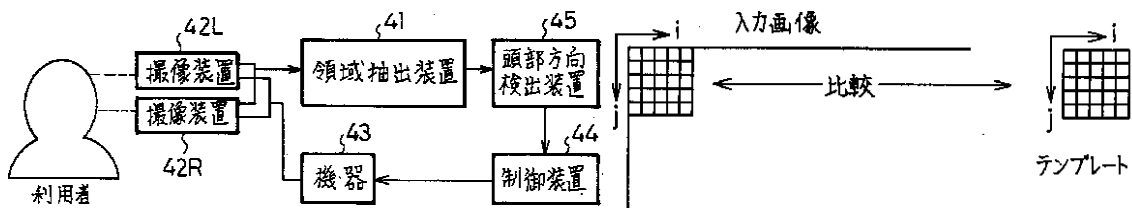


【図29】

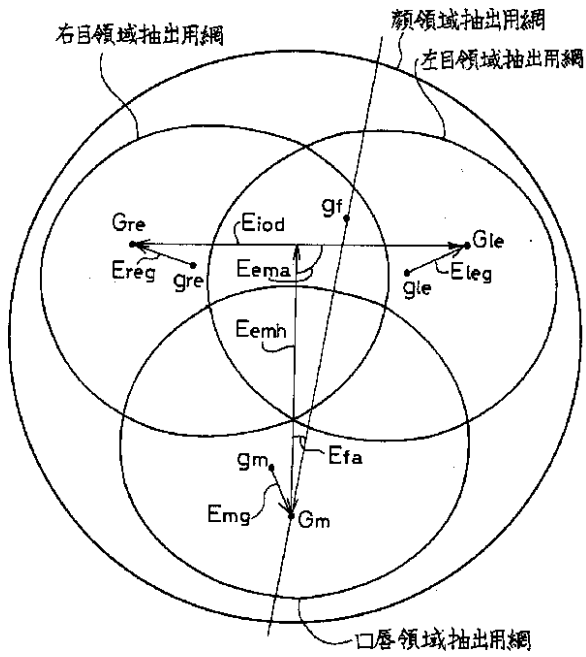


【図39】

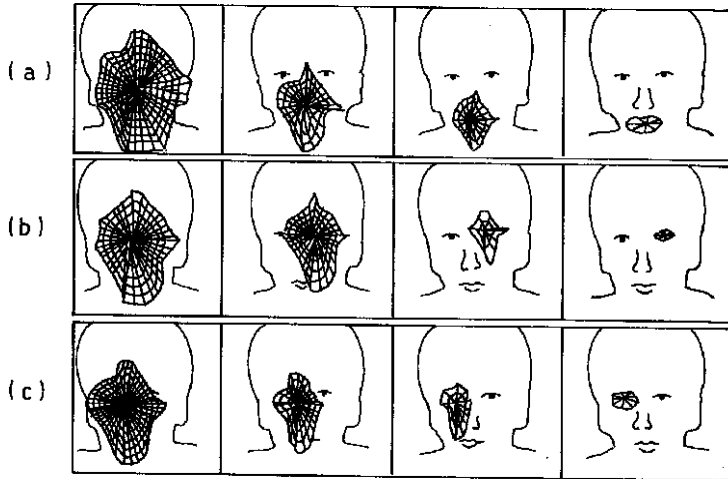
【図50】



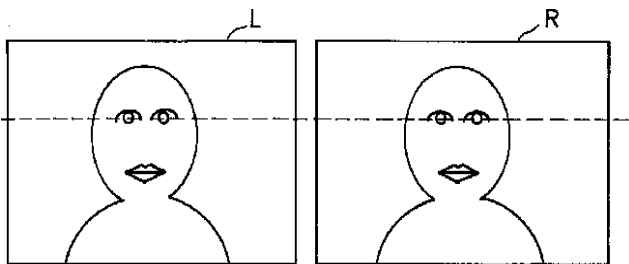
【図35】



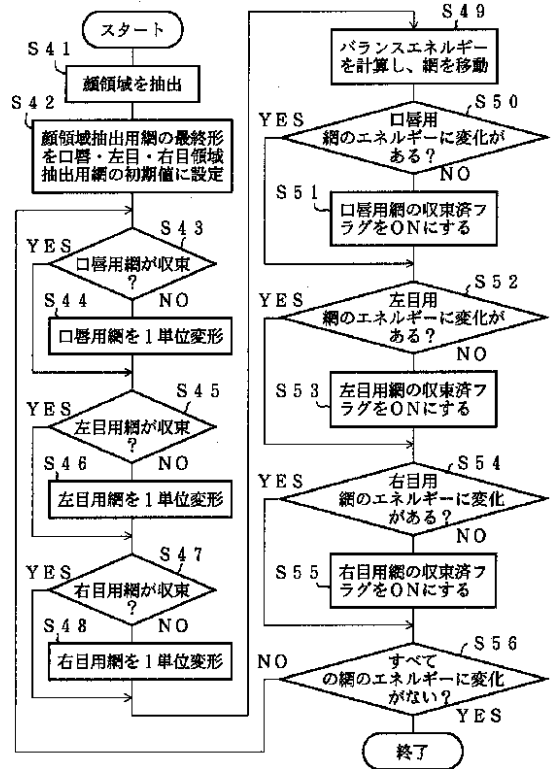
【図36】



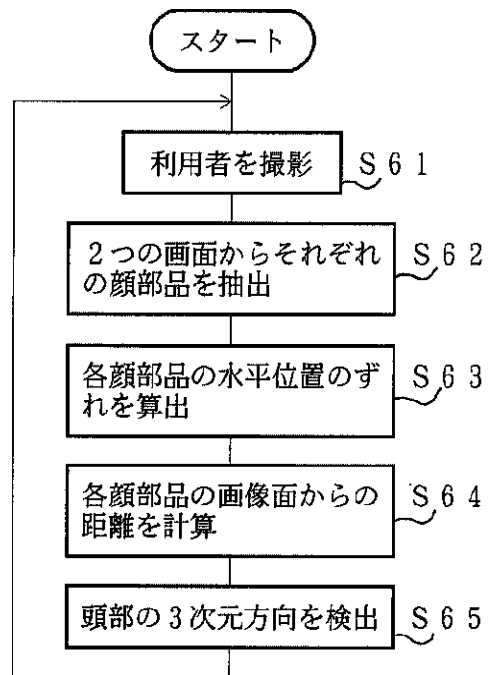
【図40】



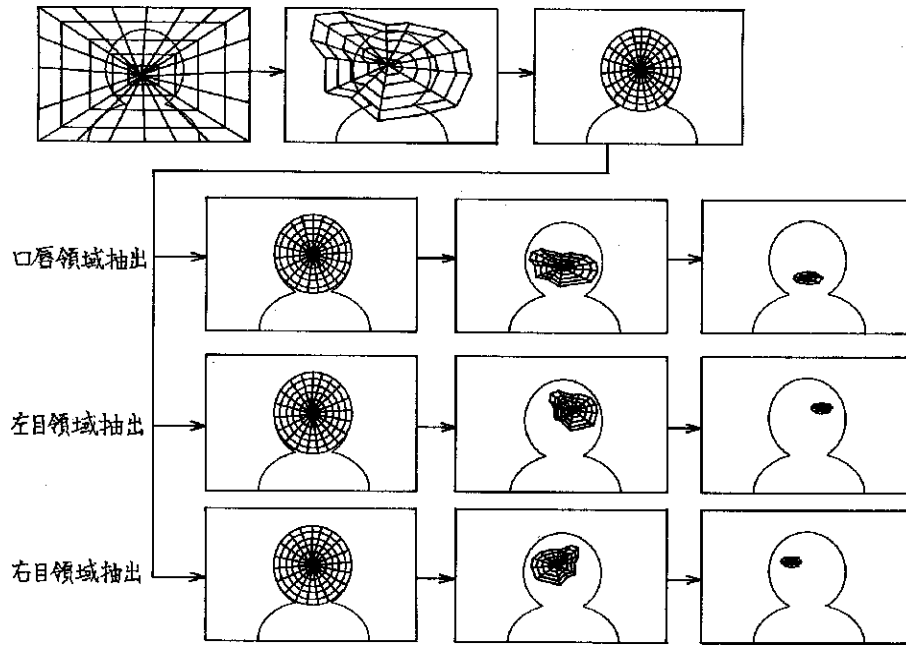
【図37】



【図43】

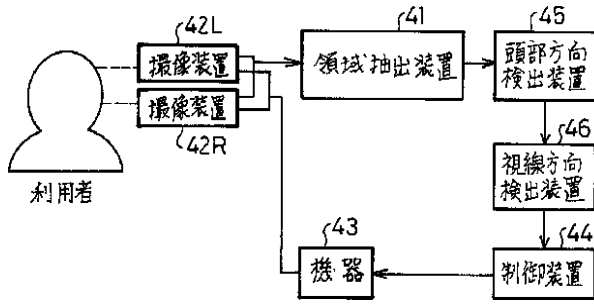


【図38】

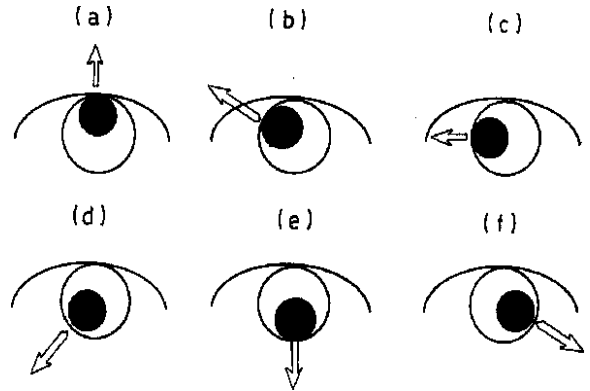


【図44】

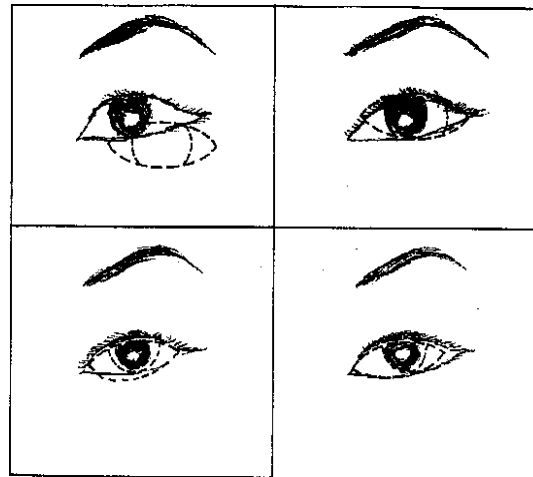
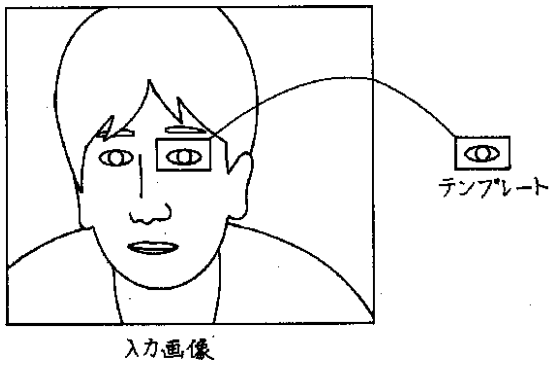
【図45】



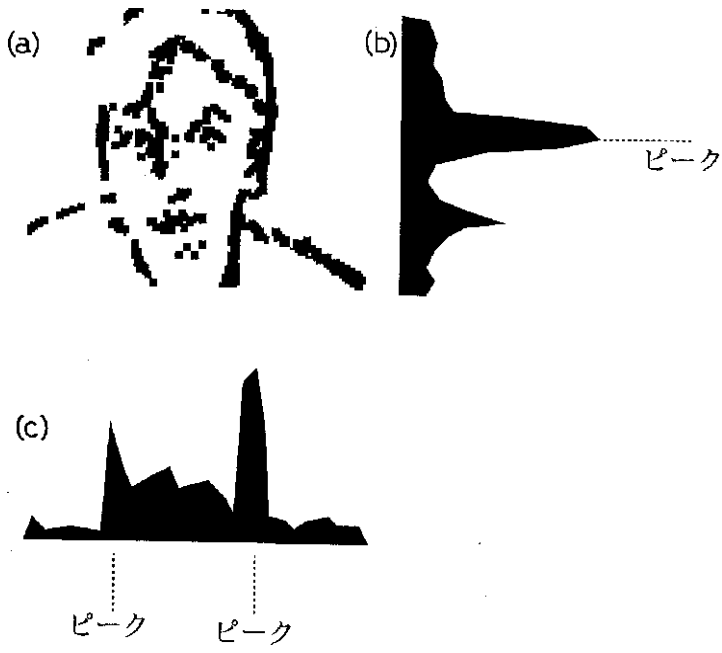
【図49】



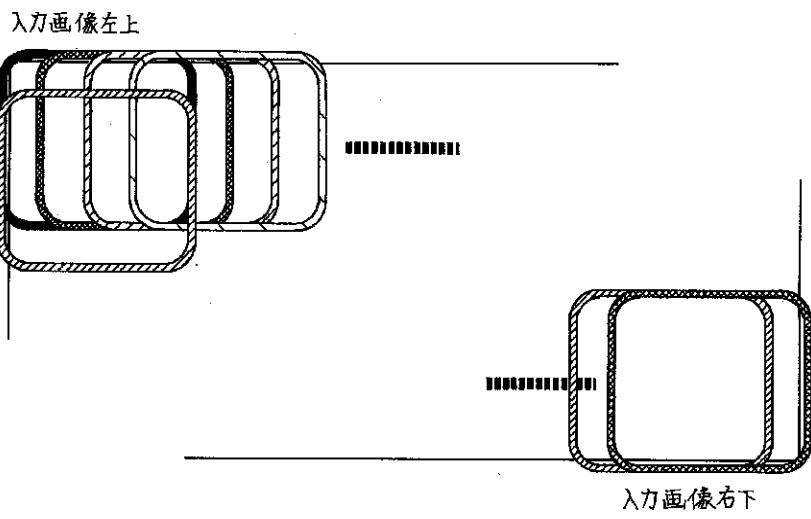
【図52】



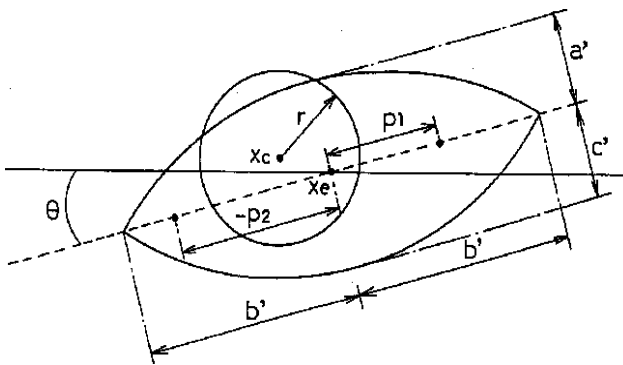
【図47】



【図51】



【図53】



【図54】

