

(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2012-123631

(P2012-123631A)

(43) 公開日 平成24年6月28日(2012.6.28)

(51) Int.Cl. F 1 テーマコード (参考)
G 0 6 T 7/00 (2006.01) G 0 6 T 7/00 2 5 0 5 L 0 9 6

審査請求 未請求 請求項の数 8 O L (全 11 頁)

(21) 出願番号	特願2010-273899 (P2010-273899)	(71) 出願人	000001007
(22) 出願日	平成22年12月8日 (2010.12.8)		キヤノン株式会社
			東京都大田区下丸子3丁目30番2号
		(74) 代理人	100126240
			弁理士 阿部 琢磨
		(74) 代理人	100124442
			弁理士 黒岩 創吾
		(72) 発明者	山中 正雄
			東京都大田区下丸子3丁目30番2号キヤ
			ノン株式会社内
		(72) 発明者	真継 優和
			東京都大田区下丸子3丁目30番2号キヤ
			ノン株式会社内
		Fターム(参考)	5L096 AA06 BA08 BA18 EA35 FA35
			FA66 GA30 JA11 KA04

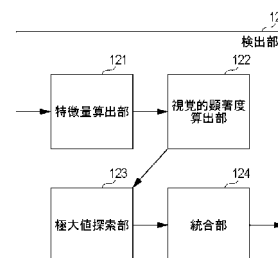
(54) 【発明の名称】 注目領域検出方法、注目領域検出装置、及びプログラム

(57) 【要約】

【課題】 入力画像中から人間にとって意味のあるひと固まりを注目領域として精度よく検出する。

【解決手段】 注目領域検出方法に、入力画像中の所定の領域において複数の低次特徴量を算出する第1の算出工程と、前記複数の低次特徴量のそれぞれの確率密度の関係に基づいて視覚的顕著度を算出する第2の算出工程と、前記視覚的顕著度の統計的分布に基づいて注目領域を検出する検出工程とを備える。

【選択図】 図2



【特許請求の範囲】**【請求項 1】**

入力画像中の所定の領域において複数の低次特徴量を算出する第 1 の算出工程と、
前記複数の低次特徴量のそれぞれの確率密度の関係に基づいて視覚的顕著度を算出する
第 2 の算出工程と、

前記視覚的顕著度の統計的分布に基づいて注目領域を検出する検出工程と
を有することを特徴とする注目領域検出方法。

【請求項 2】

前記第 1 の算出工程は、

前記所定の領域として、前記入力画像内の所定の点に対して第 1 の抽出領域と、該第 1
の抽出領域と大きさの異なる第 2 の抽出領域とを設定する抽出領域設定工程と、

前記複数の低次特徴量として、前記第 1 の抽出領域および前記第 2 の抽出領域のそれぞ
れから低次特徴量を抽出する低次特徴量抽出工程と
を有することを特徴とする請求項 1 に記載の注目領域検出方法。

【請求項 3】

前記低次特徴量抽出工程では、前記第 1 の抽出領域および前記第 2 の抽出領域のそれぞ
れから、複数種類の低次特徴量を抽出することを特徴とする請求項 2 に記載の注目領域検
出方法。

【請求項 4】

前記第 2 の算出工程では、前記視覚的顕著度を、前記第 1 の抽出領域の低次特徴量のヒ
ストグラムと前記第 2 の抽出領域の低次特徴量に対するヒストグラムの対応するヒストグ
ラムビン間の差分絶対値に基づいて算出することを特徴とする請求項 2 に記載の注目領域
検出方法。

【請求項 5】

前記第 2 の算出工程では、前記視覚的顕著度を、前記第 1 の抽出領域で得られた低次特
徴量に対する確率密度と前記第 2 の抽出領域で得られた低次特徴量に対する確率密度の比
に基づいて算出することを特徴とする請求項 2 に記載の注目領域検出方法。

【請求項 6】

前記検出工程は、

前記視覚的顕著度に基づいて複数の局所的注目領域を設定する局所的注目領域設定工程
と、

前記複数の局所的注目領域をそれらの距離の大きさに基づいて統合して局所的注目領域
群を生成する統合工程と、

前記局所的注目領域群ごとに視覚的顕著度の合計値を算出し、その最大値を与える局所
的注目領域群を特定する特定工程と、

特定された前記局所的注目領域群を包含する領域を前記注目領域として設定する注目領
域設定工程と
を有することを特徴とする請求項 1 乃至 5 のいずれか 1 項に記載の注目領域検出方法。

【請求項 7】

入力画像中の所定の領域において複数の低次特徴量を算出する第 1 の算出手段と、
前記複数の低次特徴量のそれぞれの確率密度の関係に基づいて視覚的顕著度を算出する
第 2 の算出手段と、

前記視覚的顕著度の統計的分布に基づいて注目領域を検出する検出手段と
を有することを特徴とする注目領域検出装置。

【請求項 8】

請求項 1 に記載の注目領域検出方法の各工程をコンピュータに実行させるためのプログ
ラム。

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、画像中の注目領域検出に関し、特に注目領域検出における視覚的顕著度の記述方法ならびに当該視覚的顕著度を用いた注目領域検出に関するものである。

【背景技術】

【0002】

入力画像中から人間にとって意味のあるひと固まりの領域を検出する方法として、例えば、特許文献1では、以下のような方法が提案されている。すなわち、入力画像中から基礎特徴画像を複数種類抽出し、その多重解像度表現である多重解像度画像を抽出する。また、多重解像度画像の各種類について解像度の異なる画像の間の差分である解像度差分画像を複数抽出し、解像度差分画像の各種類について、解像度の異なる解像度差分画像を統合することにより視覚的顕著度画像を抽出する。さらに、得られた視覚的顕著度画像において顕著度がある閾値以上の領域として注目領域が検出できる。

10

【先行技術文献】

【特許文献】

【0003】

【特許文献1】特開2009-3615号公報

【非特許文献】

【0004】

【非特許文献1】Hido, S., Tsuboi, Y., Kashima, H., Sugiyama, M., & Kanamori, T. Statistical outlier detection using direct density ratio estimation. Knowledge and Information Systems, to appear.

20

【発明の概要】

【発明が解決しようとする課題】

【0005】

しかし、このような入力画像中からの基礎特徴画像を直接用いて視覚的顕著度を算出する場合、環境的、観測的要因に依るノイズの影響を受けやすく、注目領域の検出精度が低下するという問題があった。

【課題を解決するための手段】

【0006】

30

上記課題を解決するために、本発明に係る注目領域検出方法は、入力画像中の所定の領域において複数の低次特徴量を算出する第1の算出工程と、前記複数の低次特徴量のそれぞれの確率密度の関係に基づいて視覚的顕著度を算出する第2の算出工程と、前記視覚的顕著度の統計的分布に基づいて注目領域を検出する検出工程とを備える。

【発明の効果】

【0007】

本発明によれば、注目領域の検出精度を向上させることができる。

【図面の簡単な説明】

【0008】

【図1】第1実施形態に係る注目領域検出装置の機能構成図である。

40

【図2】検出部の機能を説明する図である。

【図3】データ抽出領域を説明する図である。

【図4】視覚的顕著度の極大値が検出された検出点を説明する図である。

【図5】局所的注目領域群を説明する図である。

【図6】設定される注目領域を説明する図である。

【図7】学習部の機能を説明する図である。

【図8】パラメータ候補を説明する図である。

【図9】係数算出部の処理を説明する図である。

【図10】パラメータの決定を説明する図である。

【図11】パラメータ候補を説明する図である。

50

【図 1 2】注目領域検出装置を実現する情報処理装置のハードウェア構成を示すブロック図である。

【発明を実施するための形態】

【0009】

以下、図面を参照しながら、本発明の各実施形態について説明する。

【0010】

[第1の実施形態]

図 1 は、本発明の第 1 の実施形態に係る注目領域検出装置 1 の機能構成を示すブロック図である。本実施形態に係わる物体識別装置は、半導体集積回路 (LSI) を用いて実現されるものとする。図 1 に示すように注目領域検出装置 1 は、学習部 1 1、検出部 1 2、制御部 1 3 を有する。これらの構成要素は注目領域検出装置 1 が果たす機能にそれぞれ対応している。

10

【0011】

注目領域検出装置 1 が果たす機能を大別すると 2 つある。その 1 つは学習機能であり、学習部 1 1 で実行される。もう 1 つは検出機能であり、検出部 1 2 で実行される。ここで、学習部 1 1 では、検出部 1 2 で用いられるパラメータ α を算出する。また、検出部 1 2 は、学習部 1 1 で算出したパラメータ α を用いて入力画像内の注目領域を検出する。さらに、制御部 1 3 は、物体識別装置 1 の各構成要素の制御を行っている。

【0012】

一方、注目領域の検出結果は、注目領域検出装置 1 の上位階層の装置に伝達され、様々なアプリケーション (デジタルスチルカメラのオートフォーカス機能やセキュリティカメラの異常検出機能など) に適用される。

20

【0013】

図 2 は検出部 1 2 の機能構成を示すブロック図である。図 2 に示すように検出部 1 2 は、第 1 の算出部である特徴量算出部 1 2 1、第 2 の算出部である視覚的顕著度算出部 1 2 2、極大値探索部 1 2 3、統合部 1 2 4 から構成される。

【0014】

図 3 は、データ抽出領域を説明する図である。特徴量算出部 1 2 1 は、図 3 のように、注目領域検出装置 1 の外部から入力される入力画像中に、訓練データ抽出領域と検証データ抽出領域を設定する抽出領域設定を実行し、各々の領域から無作為に所定の個数の複数種類の低次特徴量抽出を行う。

30

【0015】

ここで、訓練データ抽出領域と検証データ抽出領域は、大きさの異なる 2 つの円形領域で与えられる。このうちの半径の大きい方の円形領域を第 1 の抽出領域である検証データ抽出領域とし、半径の小さい方の円形領域を第 2 の抽出領域である訓練データ抽出領域とする。

【0016】

また、訓練データ抽出領域と検証データ抽出領域の半径、抽出する低次特徴量の個数、およびその種類 (輝度値、エッジ強度、テクスチャなど) は、あらかじめ学習部 1 1 により決定され、制御部 1 3 により伝達される。

40

【0017】

なお、ここで得られた低次特徴量は視覚的顕著度算出部 1 2 2 に出力される。

【0018】

視覚的顕著度算出部 1 2 2 は、特徴量算出部 1 2 1 で得られた低次特徴量に基づいて、入力画像内の任意の点における視覚的顕著度 S を算出する。具体的には、訓練データ抽出領域から得られた低次特徴量と、検証データ抽出領域から得られた低次特徴量とのそれぞれを用いて推定される確率密度の比に基づいて、視覚的顕著度 S を算出する。

【0019】

ここで、訓練データ抽出領域における低次特徴量の確率密度 p_{train} と、検証データ抽出領域における低次特徴量の確率密度 p_{test} との密度比 (p_{test} / p_{train})

50

i_n) は、たとえば、非特許文献 1 の密度比推定手法を用いて算出できる。

【 0 0 2 0 】

これから、視覚的顕著度 S は、特徴量算出部 1 2 1 で抽出される低次特徴量が単一種の場合（たとえば、輝度値 Y のみの場合）は、その低次特徴量による確率密度の比の標準偏差 σ の逆数 $1 / \sigma$ で与えられる。

【 0 0 2 1 】

また、視覚的顕著度 S は、特徴量算出部 1 2 1 で抽出される低次特徴量が複数種の場合（たとえば、輝度値 Y 、エッジ強度 E 、テクスチャ T の 3 種類の場合）は、以下のようになる。すなわち、学習部 1 1 より入力されるパラメータ $\alpha = (\alpha_Y, \alpha_E, \alpha_T)$ と、各々の低次特徴量による確率密度の比の標準偏差 $(\sigma_Y, \sigma_E, \sigma_T,)$ の逆数 $(1 / \sigma_Y, 1 / \sigma_E, 1 / \sigma_T)$ による線形和を用いて、式 (1) のように与えられる。

10

【 0 0 2 2 】

さらに、視覚的顕著度 S は、特徴量算出部 1 2 1 で抽出される低次特徴量が複数種の場合（たとえば、 N 種類の場合）は、式 (1) を容易に拡張でき、視覚的顕著度 S は式 (2) のように与えられる。あるいは、より一般的に、視覚的顕著度 S は式 (3) のように σ_n ($n = 0 \sim N$) に関する非線形関数としてもよい。

【 0 0 2 3 】

なお、得られた視覚的顕著度 S は、極大値探索部 1 2 3 に出力される。

【 0 0 2 4 】

【 数 1 】

$$S = \frac{\alpha_Y}{\sigma_Y} + \frac{\alpha_E}{\sigma_E} + \frac{\alpha_T}{\sigma_T} \quad (1)$$

20

【 0 0 2 5 】

【 数 2 】

$$S = \frac{\alpha_0}{\sigma_0} + \frac{\alpha_1}{\sigma_1} + \dots + \frac{\alpha_n}{\sigma_n} + \dots + \frac{\alpha_N}{\sigma_N} \quad (2)$$

【 0 0 2 6 】

【 数 3 】

$$S = f(\sigma_0, \sigma_1, \dots, \sigma_n, \dots, \sigma_N) \quad (3)$$

【 0 0 2 7 】

30

極大値探索部 1 2 3 は、特徴量算出部 1 2 1 と視覚的顕著度算出部 1 2 2 を用いて、入力画像内のさまざまな点 (x, y) における視覚的顕著度 $S = S(x, y)$ を算出し、視覚的顕著度 S の統計的分布を求めるために、極大値（または、所定閾値以上の値）を与える点を検出する。

【 0 0 2 8 】

なお、極大値（または、所定閾値以上の値）の検出された点 p_k ($k = 0 \sim K$) は、統合部 1 2 4 に出力される。ただし、 K は、視覚的顕著度 S の極大値（または、所定閾値以上の値）が検出された点 (x, y) の個数を表す。

【 0 0 2 9 】

統合部 1 2 4 は、極大値探索部 1 2 3 で得られた、視覚的顕著度 S の極大値（または、所定閾値以上の値）が検出された検出点 p_k ($k = 0 \sim K$) に対応する領域を局所的注目領域として設定する局所的注目領域設定を実行する。そして、それら局所的注目領域を 2 つの検出点の間の距離 d に基づいて統合する。

40

【 0 0 3 0 】

例えば、図 4 のように、入力画像内の 3 点 (p_i, p_j, p_k) において視覚的顕著度 S の極大値（または、所定閾値以上の値）が得られたとする。図 4 において、 p_i, p_j, p_k は、各々の円形領域の中心であり、それぞれの半径の大きさは訓練データ抽出領域の半径の大きさに相当する。まず、点 p_i と p_j に着目し、これらの点の間の距離 d_{ij} が所定閾値 d_{th} より小さい場合、この 2 つの点 (p_i, p_j) を同一グループとして統合して局所的注目領域群を生成する。次に、点 p_j と p_k に着目し、これらの点の間の距

50

離 d_{jk} が所定閾値 d_{th} より小さい場合、この2つの点 (p_j, p_k) を同一のグループとして統合する。同様に、点 p_i と p_k に着目し、これらの点の間の距離 d_{ik} が所定閾値 d_{th} より小さい場合、この2つの点を同一のグループとして統合する。ただし、上記所定閾値 d_{th} は、学習部 12 により決定され、制御部 13 により伝達される。

【0031】

これを視覚的顕著度 S の極大値（または、所定閾値以上の値）が検出された点 p_k ($k = 0 \sim K$) の、すべての組に対して実行する。これにより、点 p_k ($k = 0 \sim K$) に対応する局所的注目領域を複数のグループ（局所的注目領域群） g_m ($m = 0 \sim M$) に統合する（図5）。さらに、グループ g_m ($m = 0 \sim M$) 毎に視覚的顕著度 S の合計値 S_m ($m = 0 \sim M$) を算出し、その最大値を与えるグループ g_m を包含する矩形領域を設定し（図6）、これを最終的な注目領域とする注目領域設定を実行する。ただし、上記 M は、グループ個数を表す（図5の例では $M = 3$ ）。

10

【0032】

なお、ここで得られた注目領域の入力画像上における位置とサイズは、制御部 12 に出力される。

【0033】

なお、上記手段が注目領域検出装置 1 における検出部 12 の一例に相当する。

【0034】

図7は学習部 11 の機能構成を示すブロック図である。図7に示すように学習部 11 は、画像データベース 111、係数算出部 112 から構成される。学習部 11 は、検出部 12 で用いられるパラメータ α を、画像データベース 111 に格納された GT （グランドトゥールズ）画像に基づいて決定する。ここで、 GT 画像とは、入力画像内の予め意味のあるまとまった物体領域として、注目領域を図6の点線で示す矩形枠で定義し、その位置とサイズがあらかじめ設定された画像である。係数算出部 112 は、検出部 12 を用いて同様の注目領域検出結果が得られるように、パラメータ α を学習する。

20

【0035】

具体的には、パラメータ α の候補 $\alpha_{0,m}$ ($m = 0 \sim M$) を N 次元空間のランダムな点の座標で与える。ただし、 N は低次特徴量の種類数を表し、本実施例では、輝度値 Y 、エッジ強度 E 、テクスチャ T の3種類 ($N = 3$) を考えることにする。また、 M はパラメータ α の候補数を表す。

30

【0036】

すると、パラメータ α の候補 $\alpha_{0,m}$ ($m = 0 \sim M$) は、図8のように、3次元空間 ($\alpha_Y, \alpha_E, \alpha_T$) における半径1の球体内のランダムな点の座標で与えられる。ここで、パラメータ α の候補数 M は、学習時間を短縮したい場合は、比較的小さめ（例えば10～50）に、注目領域の検出精度を重視したい場合は、比較的大きめ（例えば100～500）に設定する。

【0037】

また、係数算出部 112 は、検出部 12 を用いて、パラメータ α の候補 $\alpha_{0,m}$ ($m = 0 \sim M$) の各々を用いた場合の検出精度 $R_{0,m}$ ($m = 0 \sim M$) を測定する。ここで、係数算出部 112 は、図9のように、画像データベース 111 に格納された GT 画像と、検出部 12 を用いて得られた注目領域検出結果を各々照らし合わせる。検出精度 $R_{0,m}$ は、それらの重複する面積 s と GT 画像における注目領域の面積 s' との面積比 (s/s') の平均値で与えられる。

40

【0038】

この結果、図10のように、最も良好な検出精度 R_m を与えるパラメータ α の候補 $\alpha_{0,m}$ が特定できたとする。次に、係数算出部 112 は、図11のように、新たなパラメータ α の候補 $\alpha_{1,m}$ ($m = 0 \sim M$) を、3次元空間 ($\alpha_Y, \alpha_E, \alpha_T$) における点 $\alpha_{0,m}$ を中心とする、半径 $1 * \gamma$ ($0 < \gamma < 1$) の球体内のランダムな点の座標で与える。ここで同様に、係数算出部 112 は、検出部 12 を用いて、パラメータ α の候補 $\alpha_{1,m}$ ($m = 0 \sim M$) の各々を用いた場合の検出精度 $R_{1,m}$ ($m = 0 \sim M$) を算出する

50

。

【 0 0 3 9 】

係数算出部 1 1 2 は、以下同様に上記の操作を繰り返し、球体の半径が所定閾値 y_{th} より小さくなった場合、処理を打ち切り、その時点で得られたパラメータ α を制御部 1 3 に出力する。ただし、上記所定閾値 y_{th} は、非負の実数として、試行錯誤的に決定される。以上が注目領域検出装置 1 における学習部 1 1 の一例に相当する。

【 0 0 4 0 】

このようにして得られた注目領域検出結果は、注目領域検出装置 1 のさらに上位階層に伝達される。例えば、注目領域検出後にロボットアームにより注目対象物体をピックアップするような状況においては、ロボットアームと注目領域検出装置 1 とを制御するための装置、プログラム等に伝達され、様々なアプリケーションに適用される。

10

【 0 0 4 1 】

以上が注目領域検出装置 1 の一例に相当する。

【 0 0 4 2 】

〔 他の実施形態 〕

第 1 の実施形態では、注目領域検出装置は学習機能および識別機能の両者を実行するものとしたが、学習機能および検出機能のいずれか一方のみを実行するようにしてもよい。

【 0 0 4 3 】

第 1 の実施形態では、特徴量算出部 1 2 1 において、訓練データ抽出領域および検証データ抽出領域は 2 つの円形領域としていたが、円形状のみならず、楕円形状、矩形形状など、その他の任意形状を用いてもよい。また、最終的な注目領域も矩形領域に限らず、他の形状でもよい。

20

【 0 0 4 4 】

第 1 の実施形態では、視覚的顕著度算出部 1 2 2 は式 (2) を用いて視覚的顕著度 S を計算していたが、式 (2) の代わりに、訓練データ抽出領域の面積 s を用いて、式 (4) を用いてもよい。また、訓練データ抽出領域の面積 s の任意関数 $f(s)$ を用いて、式 (5) を用いてもよい。

【 0 0 4 5 】

【 数 4 】

$$S = \frac{s^2}{2s+1} \cdot \left(\frac{\alpha_0}{\sigma_0} + \frac{\alpha_1}{\sigma_1} + \dots + \frac{\alpha_n}{\sigma_n} + \dots + \frac{\alpha_N}{\sigma_N} \right) \quad (4)$$

30

【 0 0 4 6 】

【 数 5 】

$$S = f(s) \cdot \left(\frac{\alpha_0}{\sigma_0} + \frac{\alpha_1}{\sigma_1} + \dots + \frac{\alpha_n}{\sigma_n} + \dots + \frac{\alpha_N}{\sigma_N} \right) \quad (5)$$

【 0 0 4 7 】

また、第 1 の実施形態では、視覚的顕著度算出部 1 2 2 は、視覚的顕著度 S を、訓練データ抽出領域における低次特徴量の確率密度と、検証データ抽出領域における低次特徴量の確率密度との関係として、両者の密度比に基づいて算出していた。これに限らず、それぞれの抽出領域の低次特徴量のヒストグラムを求め、それぞれのヒストグラムの対応するヒストグラムビン間の差分絶対値に基づいて、視覚的顕著度 S を算出するようにしてもよい。

40

【 0 0 4 8 】

なお、本発明は、複数の機器（例えばホストコンピュータ、インタフェース機器、リーダ、プリンタなど）から構成されるシステムに適用しても、一つの機器からなる装置（例えば、複写機、ファクシミリ装置など）に適用してもよい。

【 0 0 4 9 】

また、本発明は、以下の処理を実施することによっても実現される。即ち、上述した実施形態の機能を実現するソフトウェア（プログラム）を、ネットワーク又は各種記憶媒体を介してシステム或いは装置に供給し、そのシステム或いは装置のコンピュータ（または

50

CPUやMPU)がプログラムを読み出して実行する処理である。

【0050】

図12は、プログラムを実行することで上述した注目領域検出装置を実現する情報処理装置のハードウェア構成を示すブロック図である。

【0051】

CPU201は、各種プログラムを実行し、装置各部の制御を行う。ROM202は、不揮発性のメモリであり、情報処理装置を初期動作させる際に必要なプログラムなどを記憶する。RAM203は、CPU201のワークエリアを提供し、2次記憶装置204から読み出されたプログラムなどを一時記憶する。2次記憶装置204は、CPU201が使用するプログラム210を記録し、画像データベースで利用する画像を格納する。なお、プログラム210は、OS211、アプリケーション212、モジュール213、およびデータ214から構成される。

【0052】

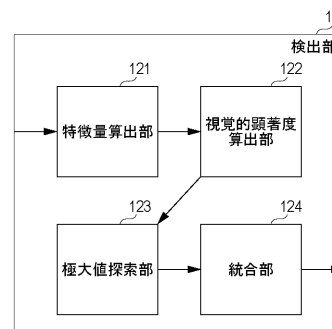
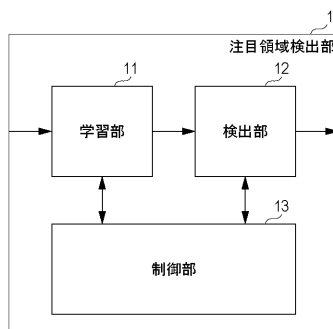
各デバイス201～204は、バス205を通じて情報をやり取りする。情報処理装置は、バス205を介して、ディスプレイ206、キーボード207、マウス208、I/Oデバイス209とつながっている。

【0053】

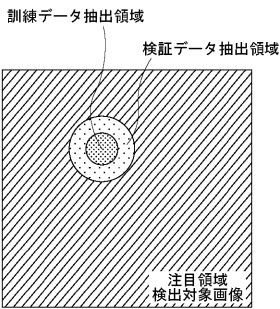
ディスプレイ206は、ユーザに処理結果や処理の途中経過等の情報を表示するのに用いられる。キーボード207とマウス208は、ユーザからの指示を入力するのに用いられ、特にマウス208は表示上の位置を入力するのに用いられる。I/Oデバイス209は、処理対象の画像を取り込むために用いられる。例えば、I/Oデバイス209は、対象物体を撮影する撮影装置から入力画像を取り込む。また、I/Oデバイス209は、情報処理結果として得られた注目領域を撮影装置や画像処理装置など他の情報処理装置へ出力することもできる。

【図1】

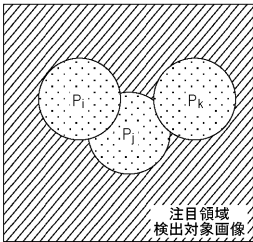
【図2】



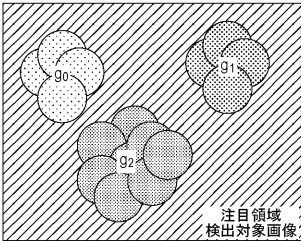
【 図 3 】



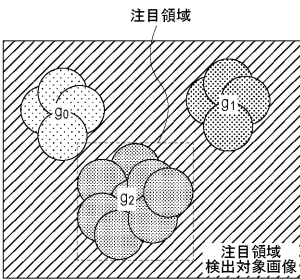
【 図 4 】



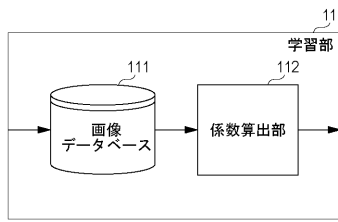
【 図 5 】



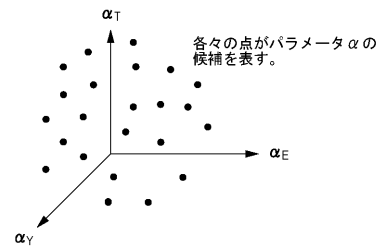
【 図 6 】



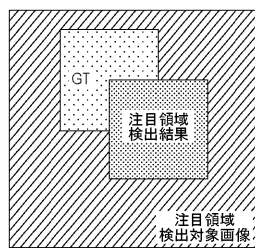
【図 7】



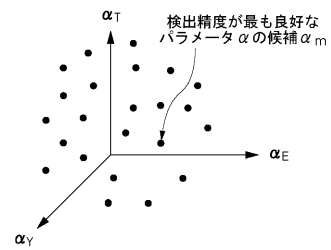
【図 8】



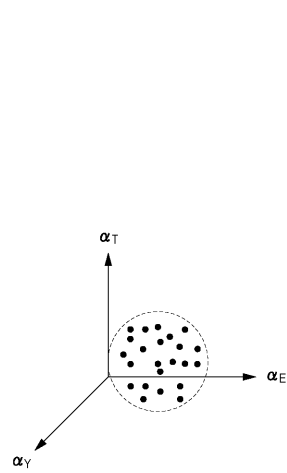
【図 9】



【図 10】



【図 1 1】



【図 1 2】

