

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4979033号
(P4979033)

(45) 発行日 平成24年7月18日(2012.7.18)

(24) 登録日 平成24年4月27日(2012.4.27)

(51) Int.Cl.

G06T 7/00 (2006.01)

F 1

G06T 7/00 100Z

請求項の数 10 (全 17 頁)

(21) 出願番号 特願2009-507098 (P2009-507098)
 (86) (22) 出願日 平成19年4月27日 (2007.4.27)
 (65) 公表番号 特表2009-535683 (P2009-535683A)
 (43) 公表日 平成21年10月1日 (2009.10.1)
 (86) 國際出願番号 PCT/EP2007/054195
 (87) 國際公開番号 WO2007/125115
 (87) 國際公開日 平成19年11月8日 (2007.11.8)
 審査請求日 平成22年4月20日 (2010.4.20)
 (31) 優先権主張番号 06300418.8
 (32) 優先日 平成18年4月28日 (2006.4.28)
 (33) 優先権主張国 歐州特許庁 (EP)
 (31) 優先権主張番号 06300538.3
 (32) 優先日 平成18年5月31日 (2006.5.31)
 (33) 優先権主張国 歐州特許庁 (EP)

(73) 特許権者 501263810
 トムソン ライセンシング
 Thomson Licensing
 フランス国, 92130 イッシー レ
 ムーリノー, ル ジヤンヌ ダルク,
 1-5
 1-5, rue Jeanne d' Arc, 92130 ISSY LES
 MOULINEAUX, France
 (74) 代理人 100115864
 弁理士 木越 力
 (74) 代理人 100121175
 弁理士 石井 たかし

最終頁に続く

(54) 【発明の名称】オブジェクト・ベース視覚的注意モデルの顕著性推定

(57) 【特許請求の範囲】

【請求項 1】

(1) 画像を複数の別個のオブジェクトにセグメント化するステップと、
各オブジェクトについて、
 (2) 前記オブジェクトのフィーチャ値を前記オブジェクトのピクセルの関数として抽出するステップであって、該フィーチャ値が輝度強度及び色の少なくともいずれか一方である、該ステップと、を含む、 画像の顕著性を推定する方法であって、該顕著性は顕著なオブジェクトが背景オブジェクトとどのくらい異なるかの尺度である、該方法は、

(3) 前記オブジェクトの近傍領域を選択するステップであって、該近傍領域が円形である該ステップと、

(4) 前記オブジェクトのフィーチャ値と前記近傍領域内のその他の各オブジェクトのフィーチャ値との間の平均差に、該その他の各オブジェクトのサイズを用いて重み付けし、前記オブジェクトの重み付け平均フィーチャ差を得るステップと、

(5) 前記複数のオブジェクトの前記重み付け平均フィーチャ差の関数として前記画像の顕著性を推定するステップと、をさらに含むことを特徴とする、前記方法。

【請求項 2】

オブジェクトのフィーチャ値を抽出するときに、ある範囲内に入るフィーチャ値を有する前記オブジェクトの複数のピクセルの数が、前記オブジェクト内のピクセルの総数と比較して所定の閾値を超える場合に、前記複数のピクセルのフィーチャ値の平均を、前記オブジェクトのフィーチャ値として抽出する、請求項 1 に記載の方法。

【請求項 3】

前記オブジェクトについて、ステップ(3)において複数の近傍領域を選択することができ、ステップ(4)を各近傍領域で実行し、各近傍領域内のオブジェクトの重み付け平均フィーチャ差の線形和をとることによって総重み付け平均フィーチャ差を得る、請求項1または2に記載の方法。

【請求項 4】

複数の近傍領域がある場合には、これらの複数の近傍領域が同心円状である、請求項3に記載の方法。

【請求項 5】

オブジェクトの各近傍領域が、前記オブジェクトの中心ピクセルを中心とする、請求項4に記載の方法。 10

【請求項 6】

前記重み付けステップにおいて、オブジェクトの各ピクセルが該オブジェクトの前記中心ピクセルと同等であるとされる、請求項1に記載の方法。

【請求項 7】

前記フィーチャ値は、前記オブジェクトの前記ピクセルの、値、色の変量、または強度である、請求項1に記載の方法。

【請求項 8】

その他のオブジェクトそれぞれのサイズが、その中のピクセル数である、請求項1に記載の方法。 20

【請求項 9】

追加フィーチャ値が動きである、請求項1に記載の方法。

【請求項 10】

前記円形領域が水平半径および異なる垂直半径を有する橜円である、請求項1に記載の方法。

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、画像の顕著性(salience)を推定する方法に関し、より詳細には、オブジェクト・ベース視覚注意モデルの顕著性推定方法に関する。 30

【背景技術】**【0002】**

神経生物学的概念では、注意するということは、緻密または入念な観察によってあるオブジェクトに精神力を集中することを意味する。注意領域とは、人間の注意を多く引きつける傾向があるピクチャ内の領域である。ピクチャの注意領域を自動的に検出するよう設計されたシステムを、注意モデルと呼ぶ。検出された注意領域は、限られた資源がある注意領域に集中する、検索/探索の方向付けをする、解析を単純化するなど、多様な応用分野で広く利用される。

【0003】

図1は、最もよく使用される注意モデルの大まかな構成を示す図である。最初に、推定対象の画像が注意モデルに入力される。次いで、特徴抽出ステップの後で、強度、色、向きなどの特徴を得る。第3のステップで、上記の特徴の顕著性を推定する。フュージョン・スキーム(fusion scheme)ステップおよび後処理ステップの後で、最終的に注意領域が得られる。 40

【0004】

マクロブロック(MB)を基本単位とする空間位置認識仮説に基づいて注意を駆動するほとんどの従来のマシン・ビジョン・システムで用いられる注意モデルと異なる、他の視覚注意を方向付けするモデルは、オブジェクト主導であり、オブジェクト・ベース視覚注意モデルと呼ばれる。

【0005】

10

20

30

40

50

MB (マクロブロック) 空間ベース視覚注意に関する研究は数多くなされており、例えば、L. Itti他の「A Model of Saliency-Based Visual Attention for Rapid Scene Analysis」、IEEE Transactions on Pattern Analysis and Machine Intelligence、Volume 20、No. 11、1998年11月、や、Y. F. Ma他の「A User Attention Model for Video Summarization」、ACM Multimedia '02、pp. 533~542、2002年12月、などがある。しかし、オブジェクト・ベース視覚注意は、本質的に難解であるために、それほど広く研究されているわけではない。Y. Sun他は、「Object-Based Visual Attention for Computer Vision」、Artificial Intelligence、pp. 77~123、2003年5月、において、オブジェクト・ベース視覚注意の構造を提案している。別のオブジェクト・ベース視覚注意モデルが、F. O'raabona他の「Object-Based Visual Attention: a Model for a Behaving Robot」、3rd International Workshop on Attention and Performance in Computational Vision、2005年6月、に提案されている。どちらのオブジェクト・ベース視覚注意方式も、依然として、図1に大まかに記載した注意モデルの構成を基本としたものである。「顕著性推定」以外の全てのプロセスは、IttiのMB空間ベース視覚注意モデルから直接生まれたものである。 10 20

【0006】

MB空間ベース視覚注意モデルでもオブジェクト・ベース視覚注意モデルでも、低レベルの空間的 / 時間的特徴を最初に抽出し、その後、ピクチャ全体にわたって、各単位毎の各顕著性（異なる、周囲と比べて顕著な、すなわちより魅力的な）特徴マップを推定し、その後、全ての特徴マップを単純にボトムアップ式に送ることによって、マスタ「顕著性マップ」を生成する。

【0007】

オブジェクト・ベース視覚注意モデルと比較して、MB空間ベース視覚注意モデルは、はるかに容易且つ速く作成することができる。しかし、MB空間ベース視覚注意モデルには、本質的な欠点がいくつかある。 30

1) 注意領域が自然オブジェクト境界を破壊する。

2) 各マクロブロックが、多数の自然オブジェクトをカバーすることがある。

【0008】

従って、マクロブロックの抽出された特徴は、これら全ての自然オブジェクトの混合したプロパティであり、そのため、注意領域の検出精度を低下させることになる。

【0009】

オブジェクト・ベース視覚注意モデルの重要な問題は、2つの特徴にある。1つは、特徴を抽出する前のオブジェクトのグループ化であり、もう1つは、画像内の全てのオブジェクトにおける各オブジェクトの個々の効率的な顕著性推定である。現在使用されている顕著性推定方式の中心にある概念は、Y. Sun他が提案しているように、ガウス距離測定 (Gauss distance measure) に基づく。 40

【0010】

顕著性推定の対象となるオブジェクトを x 、全ての背景オブジェクトを y_i ($i = 1, 2, \dots, n$)、入力画像の幅および高さの最大値を w 、 x と y_i の間の物理的距離を $||x - y_i||$ と表すと、ガウス距離は、以下の数式 (1) のように定義される。

【数1】

$$d_{gauss}(x, y_i) = \left(1 - \frac{\|x - y_i\|}{w-1} \right)^{e^{-\frac{1}{2\sigma^2} \|x - y_i\|^2}} \quad (1)$$

ここで、スケール w / w に設定する。 w は正の整数であり、一般に $1 / w$ は、2%、4%、5%、または20%、25%、50%など、 w の百分率に設定することができる。 10

【0011】

特徴 F 内のオブジェクト x および y_i の絶対差を $S_F(x, y_i)$ と表すと、特徴 F 内のオブジェクト x の総合的顕著度である顕著性推定 $S_F(x)$ は、以下の式で表すことができる。

【数2】

$$S_F(x) = \frac{\sum_{i=1}^n S_F(x, y_i) \cdot d_{gauss}(x, y_i)}{\sum_{i=1}^n d_{gauss}(x, y_i)} \quad (2)$$

20

【0012】

顕著性推定をこのように定義することにより、以下のように結論づけることができる。

1. オブジェクトとその周囲との間の差が大きくなるほど、そのオブジェクトは顕著になる。

2. オブジェクトと特徴の異なる周囲とが接近するほど、そのオブジェクトは顕著になる。すなわち、人間の視覚は、距離に応じて違いを見分ける能力が低下する。その減衰係数は、この視覚生理学のテーマと整合する d_{gauss} で測定される。

【0013】

これにより、 $S_F(x)$ は、確実に特徴 F における有用な顕著性推定となる。ただし、残念ながら、いくつかの重要な人間の知覚の性質が、 $S_F(x)$ では考慮されていない。 30

【0014】

図2aは、推定対象であるスケートのオリジナル画像であり、図3aは、従来のオブジェクト・ベース視覚注意モデルを用いて行った、図2aの顕著性推定の結果である。

【0015】

図2bは、推定対象である沿岸警備隊のオリジナル画像であり、図3bは、従来のオブジェクト・ベース視覚注意モデルを用いて行った、図2bの顕著性推定の結果である。

【0016】

図3aおよび図3bの何れにおいても、白色は非常に顕著なオブジェクトであることを意味し、黒い色は顕著ではないオブジェクトであることを意味し、白と黒の間のグレーレベルが、顕著度を表している。 40

【0017】

図3aから、観衆は、その色がその近傍の色と大きく異なるために顕著であるとみなされているが、実際には、観衆の部分は詳細を全く含んでいないことが分かる。この画像を見る者は、通常、観衆には注目せず、観衆を「ビデオ・テクスチャ(video texture)」とみなす。

【0018】

さらに図3aには、女性スケーターの頭部の左側に小さなグレーのブロックがある。このブロックは、男性の黒い衣装と女性の肌で囲まれた白いスケート・リンクの一部であり、これはこの局所領域において顕著である。しかし、結局のところ、このブロックは、広 50

いスケート・リンクの一部であって、見る者の注意を引きつけることはない。これは、「局所効果」と呼ばれる。この局所効果によって、オブジェクトとその近傍の間の差は蓄積されて大きくなり、その結果、「顕著性」と認識される。

【発明の開示】

【発明が解決しようとする課題】

【0019】

以上のことから、従来のオブジェクト・ベース視覚注意モデルは、効率が不十分であり、下記のように、考慮されていない人間の視覚特性が数多く存在する。

【0020】

1. オブジェクトのサイズ オブジェクトのサイズが顕著度に与える影響の推定は、複雑な問題である。例えば、(a) 全ての近傍のオブジェクト y_i が同じサイズ s であり、オブジェクト x のサイズが s から 0 まで減少する場合には、その結果として、 x の顕著度 ($S_F(x)$) は、徐々に減少する。(b) 全ての近傍のオブジェクト y_i が同じサイズ s であり、オブジェクト x のサイズが s_1 から s_2 ($s_1 > > s$ 、且つ $s_1 > s_2 > s$) まで減少する場合には、 $S_F(x)$ は、徐々に増大する。従って、オブジェクトのサイズと顕著度の関係が単調でないことが分かる。また、この問題は、各オブジェクトが任意のサイズを有する可能性がある場合には、さらにずっと複雑になる。

10

【0021】

2. 局所効果 あるオブジェクトがそのごく近い近傍（局所領域）の中で顕著でなく、それより少し離れた近傍がそのオブジェクトと大きく異なる場合には、2つの結果が生じる可能性がある。すなわち、(a) 画像全体の中で、そのオブジェクトが全く顕著でない場合と、(b) その画像内で、その局所領域が全体として顕著になり、そのオブジェクトはその局所領域の一部として含まれる。何れの場合にも、オブジェクトの顕著度は、上記で定義したものとは一致しない。

20

【0022】

3. ビデオ・テクスチャ ある画像のオブジェクト・特徴が無作為に不均一である場合には、人間は、通常、画像全体の細部を無視し、その画像のいかなるオブジェクトも顕著ではないが、上記で定義した $S_F(x)$ は、画像内の何れかのオブジェクトに対して大きな数になる。

30

【0023】

これら全ての制限のために、従来のオブジェクト・ベース視覚注意モデルは、実用的とは言えない状態にある。従って、オブジェクト・ベース視覚注意モデルを改良することが望まれている。

【課題を解決するための手段】

【0024】

本発明は、計算の複雑さを低下させることができ、応用可能性を高めることができるマルチレベル同心円型方式を利用した、オブジェクト・ベース視覚注意モデルの顕著性推定方式を提供する。

【0025】

一態様では、本発明は、画像の顕著性を推定する方法を提供する。この方法は、画像を複数の推定対象オブジェクトにセグメント化するステップと、各セグメント化オブジェクトについて特徴マップを抽出するステップと、抽出された特徴マップに基づいて、各セグメント化されたオブジェクトの顕著性を、当該オブジェクトの中心ピクセルの周囲に規定された1組の円内で計算するステップと、全ての円内の各セグメント化されたオブジェクトの顕著性を統合して、各セグメント化オブジェクトの総合的顕著性推定を得るステップとを含む。好ましい一実施例によれば、特徴マップを抽出するステップは、画像の色変化の測定に基づく。別の好ましい実施例によれば、各セグメント化されたオブジェクトの顕著性を計算するステップは、推定対象オブジェクトの色特徴を、当該推定対象オブジェクトの周囲に規定された各円内の任意のその他のオブジェクトの色特徴と比較するサブステップを含む。

40

50

【0026】

本発明のマルチレベル同心円型顕著性推定方式に基づくオブジェクト・ベース視覚注意モデルは、計算の複雑さが低く、人間の視覚とより連絡しやすい、オブジェクト・ベース視覚注意モデルを構築するための効率的なフレームワークを提供得きるという効果を有する。

【0027】

本発明のその他の特徴および利点は、添付の図面を援用しながら以下で説明する本発明の非限定的な実施例の説明から、明らかになるであろう。

【発明を実施するための最良の形態】

【0028】

本発明の技術的特徴について、実施例を参照してさらに説明する。これらの実施例は、単なる好ましい例であって、本発明を限定するものではない。以下の詳細な説明を添付の図面とあわせて読めば、本発明をよく理解できるであろう。

【0029】

前記の内容から、顕著性推定プロセスは、次のように説明することができるところが分かる。

入力：幅 w および高さ h を有する画像 $I = \{ p_i \mid i = 1 \dots w * h \}$ 。

出力：顕著性マップ $sal [1 \dots w * h]$ 。 $sal [i]$ は、画像内のピクセル p_i の顕著度。

【0030】

本発明の方法は、主に以下に述べる 3 つのステップを含む。

【0031】

ステップ1 前処理（画像のセグメント化）

画像 I は、最初にこのステップで、1 組のオブジェクト $I = \{ o_1, o_2, \dots, o_n \}$ に分解される。画像ベースのセグメント化およびグループ化は、人間の視覚的認識において大きな役割を果たし、この領域の研究は数多くなされている。本発明では、発明者等は、画像の色変化の測定に基づく、P. F. Felzenszwalb 他による「Image Segmentation Using Local Variation」、IEEE Computer Society on Computer Vision and Pattern Recognition、1998年6月、に提案されているオブジェクト・セグメント化方式を採用する。以下、この方式について簡単に説明する。

【0032】

処理の前に、無向グラフ $H = (V, E)$ を、画像 I に基づいて定義する。ここで、 I の各ピクセル p_i は、対応する頂点 $v_i \in V$ を有し、エッジ $(v_i, v_j) \in E$ は、頂点 v_i と v_j とを結びつけるエッジである。どのピクセルが E 中のエッジによって結びつけられるかの精密な定義は、以下の式 (1-1) によって決まる。

【数3】

$$E = \{(v_i, v_j) \mid \|p_i - p_j\| \leq d\} \quad (1-1)$$

【0033】

所与の距離 d に対して、エッジの重み関数 $weight(\cdot)$ は、個々の頂点 v_i と v_j の間の類似性（または差異）の非負測度をいくつか与える。 $weight(\cdot)$ は、以下の式 (1-2) のように定義する。

【数4】

$$weight((v_i, v_j)) = \begin{cases} |Color(v_i) - Color(v_j)| & (v_i, v_j) \in E \\ \infty & \text{その他の場合} \end{cases} \quad (1-2)$$

ここで、color(v_i)は、画像内のピクセル p_i の色である。

【0034】

$S = \{C_i\}$ は、 V のセグメント化を表し、各 C_i は、セグメント化されたオブジェクトに対応する。 C の内部変化は、以下の数式(1-3)のように定義する。 10

【数5】

$$Int(C) = \max_{e \in MST(C, E)} weight(e) \quad (1-3)$$

ここで、MST(C, E)は、 E の集合に対する C の最小スパンニング・ツリーである。

【0035】

2つのオブジェクト C_1 および C_2 の外部変化は、以下の数式(1-4)のように定義する。 20

【数6】

$$Ext(C_1, C_2) = \min_{v_i \in C_1, v_j \in C_2} weight((v_i, v_j)) \quad (1-4)$$

【0036】

セグメント化のプロセスは、セグメント化されたオブジェクトのうちの任意の2つについて以下の数式(1-5)を満たすようにするものである。 30

【数7】

$$Ext(C_1, C_2) \leq \min(Ext(C_1) + k / |C_1|, Ext(C_2) + k / |C_2|) \quad (1-5)$$

ここで、 k は定数であり、発明者等の実施態様では、 k を100に設定する。オブジェクト C 内部の拡張内部変化を、 $Int(C) + k / |C|$ で表す。

【0037】

セグメント化を実施するためには、まず、非減少エッジ重みによって E を $= (e_1, e_2, \dots, e_m)$ に分類し、最初に画像を $w * h$ 個の個々のピクセル・オブジェクトにセグメント化し、次いで、各 $e_q = (v_i, v_j)$ ($q = 1, 2, \dots, m$)に対して、次のプロセスを繰り返す。すなわち、 v_i および v_j が異なるオブジェクトに属し、且つ $weight(v_i, v_j)$ が、 v_i および v_j が属する2つのオブジェクトの拡張内部変化($Int(C) + k / |C|$)以下である場合には、これら2つのオブジェクトがマージされて、新たな1つのオブジェクトを形成する。 40

【0038】

これは、コンピューティング資源をそれほど多く費やさない効率的なオブジェクト・セグメント化方式を与えることが分かる。実施に際しては、ここでは、8連結近傍を用いて

E を構成する。すなわち、 $d = 1$ である。図 6 a および図 6 b はそれぞれ、図 2 a および図 2 b のセグメント化の結果を示している。

【0039】

ステップ2 前処理（特徴抽出）

y_i 、 r_i 、 g_i および b_i がそれぞれピクセル p_i の輝度チャネル、赤チャネル、緑チャネル、青チャネルを表すものとして、次に、ステップ1でセグメント化した各オブジェクトの特徴を抽出する。

【0040】

拡張内部変化 $I_{nt}(C) + k / |C|$ の定義を考えると、 $k / |C|$ は、内部変化に対する加算項である。これは、内部変化を過小評価することが、要素の発達を防止するのに望ましくないからである。その結果、小さいオブジェクトの方が、その内部の内部変化にかかわらず発達しやすい。例えば、図 6 a のスケートの例では、黒い衣装の大部分は、主に黒いピクセルで構成されるが、依然としていくつかの白いピクセルが例外的に存在している。

10

【0041】

上述の状況における特徴抽出の問題を解決するために、オブジェクト $o = \{v_1, v_2, \dots, v_t\}$ の特徴マップ $F = f(v_i)$ において、演算子 $\text{Major}(f, o)$ を定義する。 $\text{Major}(f, o)$ の戻り値は、以下を満たすように定義されたオブジェクト o の代表特徴である (d_1, d_2 および γ は定数であり、発明者等の実施態様ではそれぞれ 2, 64 および 95 % に設定される)。

20

(1) 数式 $\max - \min + 1 - d_1$ を満たす範囲 $[\min, \max]$ があり、その値が特徴マップ $F = \{f(v_1), f(v_2), \dots, f(v_t)\}$ 全体にわたって範囲 $[\min, \max]$ 内に収まる特徴の比率が γ 以上である場合には、 $\text{Major}(f, o)$ を、範囲 $[\min, \max]$ 内に収まる値を有するそれらの特徴の平均値として定義する。

(2) そうでない場合で、オブジェクトの大きさが定数 d_2 超である場合には、ステップ1と同じプロセスでそのオブジェクトを2つのサブオブジェクトに分割し、次いで、上記ステップを各サブオブジェクトに対して繰り返し実行する。そうでない場合で、オブジェクトが小さすぎる場合には、 $\text{Major}(f, o)$ を全ての特徴の平均値として定義する。

30

【0042】

演算子 $\text{Major}(f, o)$ を用いて、 Y_i 、 R_i 、 G_i 、 B_i および Ye_i を、以下のようにオブジェクト o_i の輝度チャネル、赤チャネル、緑チャネル、青チャネルおよび黄チャネルとして定義する。

【数8】

$$Y_i = \text{Major}(y, o_i)$$

$$R_i = \text{Major}(r, o_i) - (\text{Major}(g, o_i) + \text{Major}(b, o_i)) / 2$$

40

$$G_i = \text{Major}(g, o_i) - (\text{Major}(r, o_i) + \text{Major}(b, o_i)) / 2$$

$$B_i = \text{Major}(b, o_i) - (\text{Major}(r, o_i) + \text{Major}(g, o_i)) / 2$$

$$Ye_i = (\text{Major}(r, o_i) + \text{Major}(g, o_i)) / 2 - |\text{Major}(r, o_i) - \text{Major}(g, o_i)| / 2 - \text{Major}(b, o_i)$$

【0043】

強度特徴は、以下の数式 (2 - 1) として抽出される。

【数9】

$$I_i = Y_i \quad (2-1)$$

【0044】

「二重反対色 (c o l o u r d o u b l e - o p p o n e n t)」ベースの色特徴は、以下の数式 (2-2) および (2-3) として抽出される。

10

【数10】

$$RG_i = R_i - G_i \quad (2-2)$$

$$BYe_i = B_i - Ye_i \quad (2-3)$$

【0045】

向きは、オブジェクト・ベース視覚注意モデルにおいて複雑な特徴となる。全てのオブジェクトが色変化によってセグメント化されるので、オブジェクト自体は、オブジェクトの境界線を除けば、いかなる向き情報も含まないことになる。セグメント化されたオブジェクトにはこの特殊な性質があるので、発明者等は、本実施態様では向きを考慮しない。

20

【0046】

向きと比較すると、オブティカル・フロー技術が現在ますます進歩していることを考えれば、動きの方が、付加的特徴として、より可能性がある。

【0047】

しかし、簡潔にするために、発明者等は、本発明では、3つの特徴マップ I_i 、 RG_i および BY_i のみを考慮している。

【0048】

30

ステップ3 顕著性推定

上記の2つのステップの後で、画像 I をオブジェクト $I = \{o_1, o_2, \dots, o_n\}$ にセグメント化し、3つの特徴マップ I_i 、 RG_i および BY_i ($i = 1 \dots n$) を抽出する。残る問題は、 $SalF(o_i)$ で表される各特徴マップ F ($F \in \{I, RG, BY\}$) の顕著性マップをどのように推定するかである。

【0049】

画像の任意のオブジェクト o_i について、サイズ (オブジェクト内のピクセル数) を s_i で表し、オブジェクトの中心ピクセルを $c_i = (X_i, Y_i)$ で表す。 X_i および Y_i は、以下の数式 (3-1) で表される。

【数11】

40

$$X_i = \frac{1}{s_i} \sum_{j=1}^{s_i} x_j \quad ; \quad Y_i = \frac{1}{s_i} \sum_{j=1}^{s_i} y_j \quad (3-1)$$

【0050】

顕著性推定プロセス中には、オブジェクト o_i の各ピクセルは、区別されずに中心ピクセル c_i に等しいとみなされるので、そのオブジェクトは図4に示すように、中心ピクセルを s_i 個複製したものとみなされる。

50

【0051】

この仮定に基づくと、マルチレベル同心円方式のオブジェクト o_i の顕著性推定が得られる。この方式の第1のステップでは、オブジェクトの中心ピクセル c_i を中心とした1組の同心円が定義され、 $C_1 \dots C_t$ (C_j は水平半径 r_x 、垂直半径 r_y の橙円であり、 j レベル円と呼ばれる)が、中心ピクセル c_i に近い近傍領域から遠い近傍領域に向かって分布する。各 j レベル円について、 $SalC_F(o_i, C_j)$ で表される C_j 内のオブジェクト o_i の顕著性を推定し、次いで、オブジェクト o_i の総合的推定顕著性を、数式(3-2)のように定義する。ただし、 k_t は線形積分の定数である。

【数12】

10

$$Sal_F(o_i) = \sum_{j=1}^t k_t \cdot SalC_F(o_i, C_j) \quad (3-2)$$

【0052】

この場合、領域 C_j 、および領域 C_j 全体から抽出した特徴 F を有する領域 C_j 内のオブジェクト o_i について、人間の視覚特性を考慮して領域 C_j 内のオブジェクト o_i の顕著性を推定するにはどうすればよいか。ここで、発明者等は、まず演算 $SalC_F$ を以下のように定義する。

20

【0053】

(1) 中心ピクセルを領域 C_j 内に有するオブジェクトの集合を S とする。

(2) 集合 S 内の各オブジェクト o_u について、 $F'_{u} = abs(F_u - F_i)$ を定義する。

【0054】

次いで、以下の数式を使用して、集合 S 内の F'_{u} の重み付け平均を計算する。

【数13】

$$avgF' = \frac{\sum_{\forall u, c_u \in S} F'_{u} \cdot s_u}{\sum_{\forall u, c_u \in S} s_u} \quad (3-3)$$

30

【0055】

(3) F'_{u} が $avgF'$ を超えない集合 S 内のピクセルの比率を ρ と定義する。

【数14】

$$\rho = \frac{\sum_{\forall u, c_u \in S} bool(F'_{u} \leq avgF') \cdot s_u}{\sum_{\forall u, c_u \in S} s_u} \quad (3-4)$$

40

ここで、 $bool(exp)$ は、 exp が真のデータミナントである場合には1を戻し、そうでない場合には0を戻す。

【0056】

(4) 図5に示す検出関数 $textture(.)$ の定義を用いて、以下のように $SalC_F$ を設定する。

【数15】

$$SalC_F(o_i, C_j) = avgF' \times texture(\rho) \quad (3-5)$$

ここで、`texture(.)`は、「オーディエンス・エリア(`audience area`)」、すなわち注意すべき領域として認識されないことが期待される観衆などの無作為な特徴を有する領域を検出するための の経験的関数である。検出関数`texture()`は、 の値が小さくなるほど、`texture()`の値が大きくなる、従ってこの領域が「オーディエンス・エリア」すなわち画像のビデオ・テクスチャとして認識される可能性が高くなるという条件を満たす。この検出関数`texture(.)`を使用することにより、領域内の注意する必要のないオブジェクトが注意すべきオブジェクトとして認識される可能性は低下する。 10

【0057】

上記のことから、この顕著性推定方式を以下のように結論づけることができる。

(a) 画像の各オブジェクト o_i について、同心円 C_j ($j = 1 \dots t$) の集合を定義する。

(b) 上記の`SalC_F`の定義に従って`SalC_F(o_i, C_j)`を計算する。

(c) 数式(3-2)に従って全ての C_j について顕著性推定を積分して、総合的推定顕著性を得る。 20

【0058】

図7aおよび図7bはそれぞれ、本発明による顕著性推定方式を用いた図2aおよび図2bの顕著性推定実験結果を示している。本発明の顕著性推定方法を用いることによって、図2aの観衆および図2bの背景は、顕著ではないとみなされ、図3aの女性スケーターの頭部の左側の小さなブロックも図7aでは除去されている。本発明は、局所効果およびビデオ・テクスチャを処理することができ、応用可能性が高められている。

【0059】

以上、本発明の好ましい実施例および態様について説明したが、本発明を逸脱することなく、多くの改変を設計または構造の細部に加えることができることを、当業者なら理解するであろう。本発明は、個々に開示したものも、可能な配列および組合せで開示したものも含めて、全ての特徴を包含する。 30

【0060】

本発明のマルチレベル同心円型顕著性推定方式に基づくオブジェクト・ベース視覚注意モデルは、画像の理解の精度を向上させ、計算効率を大幅に向上させ、以下のいくつかの利点を有する。

1. 本発明は、オブジェクト・ベース視覚注意モデルを構築するための効率的なフレームワークを提供する。その計算の複雑さは低い。

2. 提供されるフレームワークは、人間の視覚とはるかに連絡しやすい。従来の方式で考慮されていなかった人間の視覚の特性(オブジェクトのサイズ、局所効果、およびビデオ・テクスチャなど)がよく考慮されている。 40

3. このフレームワークは拡張可能である。

【図面の簡単な説明】

【0061】

【図1】最もよく使用される注意モデルの大まかな構造を示す図である。

【図2a】顕著性推定対象であるスケートのオリジナル画像を示す図である。

【図2b】顕著性推定対象の沿岸警備隊のオリジナル画像を示す図である。

【図3a】従来のオブジェクト・ベース視覚注意モデルを用いて行った、図2aの顕著性推定の結果を示す図である。

【図3b】従来のオブジェクト・ベース視覚注意モデルを用いて行った、図2bの顕著性 50

推定の結果を示す図である。

【図4】本発明の好ましい実施例による顕著性推定のマルチレベル同心円型方式を示す図である。

【図5】本発明の *texture(·)* の例示的な定義を示す図である。

【図6a】本発明の好ましい実施例による図2aのセグメント化の結果の一例を示す図である。

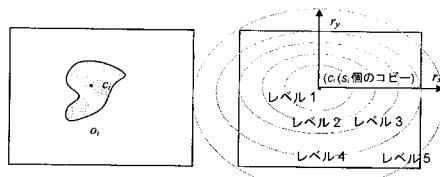
【図6b】本発明の好ましい実施例による図2bのセグメント化の結果の別の例を示す図である。

【図7a】本発明の好ましい実施例による顕著性推定方式を用いて行った、図2aの顕著性推定の結果を示す図である。

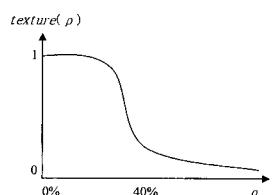
【図7b】本発明の好ましい実施例による顕著性推定方式を用いて行った、図2bの顕著性推定の結果を示す図である。

10

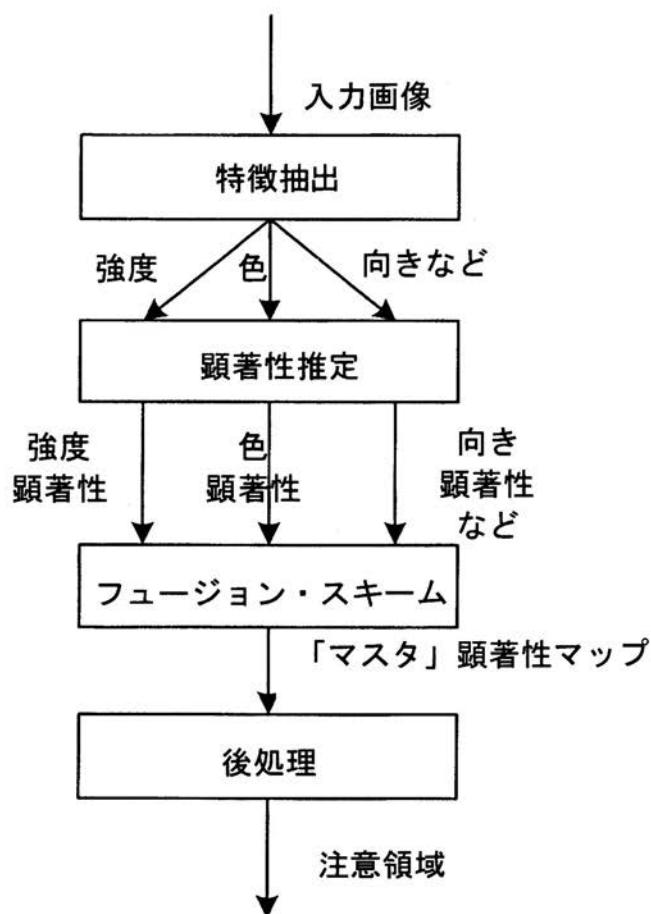
【図4】



【図5】



【図1】



【図2a】



【図2b】



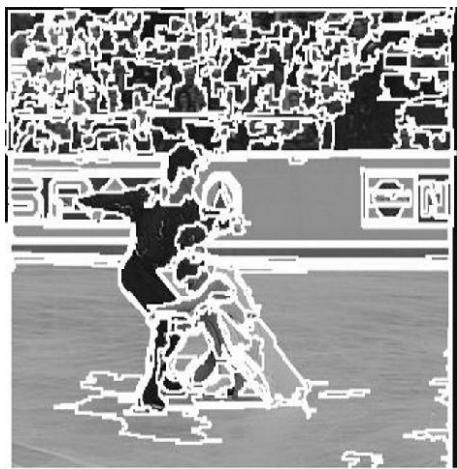
【図3a】



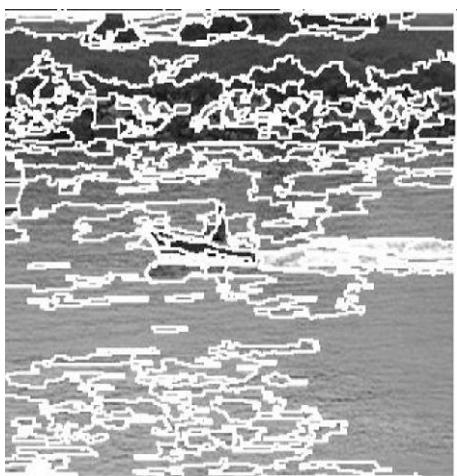
【図3b】



【図 6 a】



【図 6 b】



【図 7 a】



【図 7 b】



フロントページの続き

(72)発明者 グ , シヤオ ドン

中華人民共和国 ベイジン ハイ・ディアン・ディストリクト クー・キン・ロード 8 テクノロジー・フォーチュン・センター ビルディング エー 8エフ ルーム 03-09

(72)発明者 チエン , ジ ボ

中華人民共和国 ベイジン ハイ・ディアン・ディストリクト クー・キン・ロード 8 テクノロジー・フォーチュン・センター ビルディング エー 8エフ ルーム 03-09

(72)発明者 チエン , チュウ チン

中華人民共和国 ベイジン ハイ・ディアン・ディストリクト クー・キン・ロード 8 テクノロジー・フォーチュン・センター ビルディング エー 8エフ ルーム 03-09

審査官 佐藤 実

(56)参考文献 特開2000-207564(JP,A)

特開2003-248825(JP,A)

(58)調査した分野(Int.Cl. , DB名)

G06T 7/00~7/60