

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5667061号
(P5667061)

(45) 発行日 平成27年2月12日(2015.2.12)

(24) 登録日 平成26年12月19日(2014.12.19)

(51) Int.Cl.

F I

G 0 6 T 7/00 (2006.01)

G 0 6 T 7/00 1 3 0

請求項の数 2 (全 41 頁)

(21) 出願番号	特願2011-530244 (P2011-530244)	(73) 特許権者	505005049
(86) (22) 出願日	平成21年10月1日(2009.10.1)		スリーエム イノベイティブ プロパティ
(65) 公表番号	特表2012-504830 (P2012-504830A)		ズ カンパニー
(43) 公表日	平成24年2月23日(2012.2.23)		アメリカ合衆国, ミネソタ州 55133
(86) 国際出願番号	PCT/US2009/059255		-3427, セント ポール, ポスト オ
(87) 国際公開番号	W02010/039976		フィス ボックス 33427, スリーエ
(87) 国際公開日	平成22年4月8日(2010.4.8)		ム センター
審査請求日	平成24年9月28日(2012.9.28)	(74) 代理人	100099759
(31) 優先権主張番号	61/102,670		弁理士 青木 篤
(32) 優先日	平成20年10月3日(2008.10.3)	(74) 代理人	100092624
(33) 優先権主張国	米国 (US)		弁理士 鶴田 準一
		(74) 代理人	100114018
			弁理士 南山 知広
		(74) 代理人	100160716
			弁理士 遠藤 力

最終頁に続く

(54) 【発明の名称】 複数の視点のシーン解析のシステム及び方法

(57) 【特許請求の範囲】

【請求項 1】

3Dシーンの異なる視点から取得された複数の画像を受信することと、
前記複数の画像内に現れるオブジェクトの少なくともいくつかを識別する入力を受信することと、

シーン内の領域が、それらが視覚的注意を引き付ける、又は眼球動作を引き付ける尤度に関して異なる程度を決定する視覚的注意モデルを用いて、前記複数の画像を解析することと、

前記解析に基づいて、識別されたオブジェクトが前記3Dシーンの異なる視点にわたる視覚的注意を集める傾向がある程度を決定することと、

視覚的注意を集める傾向があるオブジェクトが前記3Dシーンの視覚的特性内の変化に対して堅牢である程度を決定することと、を含む、コンピュータ実行型の方法。

【請求項 2】

コンピュータシステムであって、

プロセッサ及びメモリと、

複数の視点のシーン解析モジュールであって、少なくとも

(1) 3Dシーン内から取得された複数の画像を受信する、

(2) 前記画像内に現れる少なくともいくつかのオブジェクトを識別する、

(3) シーン内の領域が、それらが視覚的注意を引き付ける、又は眼球動作を引き付ける尤度に関して異なる程度を決定する視覚的注意モジュールを前記画像に適用し、前記視

覚的注意モジュールは視覚的注意を集める傾向がある前記画像内のオブジェクトを予測するように機能する、

(4) 前記視覚的注意モジュールによって予測される前記識別されたオブジェクトのうちのどれが、前記3Dシーンの異なる視点にわたる視覚的注意を集める傾向があるかを決定するように、機能する、

(5) 視覚的注意を集める傾向があるオブジェクトが前記3Dシーンの視覚的特性内の変化に対して堅牢である程度を決定するように、機能する、複数の視点のシーン解析モジュールと、を含む、コンピュータシステム。

【発明の詳細な説明】

【技術分野】

10

【0001】

(関連出願の相互参照)

本出願は、米国特許仮出願第61/102670号(2008年10月3日出願)の利益を主張する。

【0002】

(関連出願)

本出願はまた、米国特許仮出願第61/102618号(2008年10月3日出願)及び同第61/102625号(2008年10月3日)に関連している。

【背景技術】

【0003】

20

ヒトの視覚的システムは、容量が限定されたシステムであり、そこでは、それは所与の時間で比較的少量のオブジェクトしか処理できない。いずれか所与の時間で見ることができる多くのオブジェクトが存在するという事実に対して、これは本当である。ヒトに見えるオブジェクトの配列から、ヒトの視覚的システムは、いずれか所与の時間で1つ(又は非常に少ない)オブジェクトに注意を払い、又はこれを処理する。ヒトが画像又はシーンを見ると、彼の視覚的システムは1つのオブジェクトから別のものへと注意(及び思考プロセス)を移す。

【0004】

ヒトの視覚的注意の分野において相当量の研究がされている。この研究は、多くの調査を、ヒトの視覚的注意の挙動、並びに視覚的注意の多くの計算モデルの理解に向けている。これらの計算モデル(視覚的注意モデル、視線予測モデル、注意モデル、又は顕著性モデルと呼ばれることもある)は、所与の視覚刺激(例えば、写真又はシーン)で、ある人がその視覚的注意又は視線をどこに払うかということを予測する。

30

【発明の概要】

【発明が解決しようとする課題】

【0005】

これらのモデルは、視覚的注意を引き付けるシーン内のオブジェクト又は領域についての予測を提供する。典型的な現実世界のシーンは、しかしながら、高度にダイナミックである。ヒトに投影される画像は、例えばその人の視点が変わると、シーン内のオブジェクトは、位置又は配向を変化させ、あるいは照明が変化する(異なる影がかかる)。更に、観察者自身が、不確実性を予想内に生じさせる場合がある(観察者は先入観を抱いているか、ないしは別の方法で特定の注意パターンに配置されている場合がある)。シーンから投影された画像におけるいずれかの変動、又は観察者にわたる変動、又はシーンそれ自体への小さな変化でさえも、これらのモデルによって作られる予測を著しく変える場合がある。これは、適用された設定において視覚的注意モデルを使用するときに問題となることもある。

40

【課題を解決するための手段】

【0006】

視覚的注意モデルリング技法を使用する、複数の視点からシーンを評価するシステム及び方法。代表的な実施形態では、複数の画像がシーンから取得される(ないしは別の方法

50

で生成される)。シーン内のオブジェクトは、各画像内で識別され、次いでシーンは、視覚的注意モデルを使用して解析される。視覚的注意モデルによって予測されたオブジェクトは追跡され、それによって、あるエリア内の複数の見解から、オブジェクトの顕著性についての予測をすることができる。他の代表的な実施形態では、複数の視点のシーン解析は、シーンの最適化に関連する技術及びシステムと組み合わせられてもよい。例えば、オブジェクトに対する許容可能な変更が定義されてもよく、次いで、異なる視点からとられた複数の画像がとられ、又は生成され、次いで解析され得る。これらの画像は、オブジェクトに対してされた変更を含んでもよい。更に、一部の実施形態では、堅牢性は複数の視点のシーン解析において、因数としても使用され得る。例えば、3Dシーンは解析され、堅牢であるといわれる場合があり(すなわち、算出された堅牢性の値)、すなわち3Dシーン内のオブジェクトは、堅牢であると決定され得る。

10

【0007】

一実施形態では、3Dシーン内からとられた複数の画像を受信することと、複数の画像内に現れるオブジェクトの少なくともいくつかを識別する入力を受信することと、視覚的注意モデルを用いて、複数の画像を解析することと、この解析に基づいて、複数の画像内で視覚的注意を集める傾向がある、識別されたオブジェクトを決定することと、を含む、コンピュータ実行型の方法が記載される。3Dシーン内から取られた複数の画像を有することに加えて、多くの画像がオブジェクトの共通集合の異なる視点から生成され得る。

【0008】

他の実施形態では、3Dシーンに関する少なくとも1つの視覚的目標を定義することによって、シーンは複数のオブジェクトからなり、視覚的目標は、視覚的注意モデルによって注意が払われると予測されることが望まれている、シーン内の少なくとも1つのオブジェクトを指定し、更にそこから少なくとも1つのオブジェクトが注意を払われるべきである視点を明記することと、3Dシーン内の複数の視点から、複数の画像を受信することと、シーン内の少なくともいくつかのオブジェクトに対する許容可能な変更を定義する入力を受信することと、許容可能な変更の少なくともいくつかに、コストの見積りを割り当てることと、修正された画像を作るために、定義された許容可能な変更と一致する、複数のオブジェクトのいくつかを自動的に修正することと、視覚的注意モデルを用いて、修正された画像を評価することと、評価に基づいて、シーンの少なくとも1つの視覚的目標を達成する、修正されたシーンの少なくともいくつかを決定することと、を含む、コンピュータ

20

30

【0009】

他の実施形態では、プロセッサ及びメモリ、並びに複数の視点のシーン解析モジュールであって少なくとも以下のように機能するモジュールを含むコンピュータシステムが記載される。(1)3Dシーン内から取られた複数の画像を受信する、(2)画像内に現れる少なくともいくつかのオブジェクトを識別する、(3)画像に視覚的注意モジュールを適用し、視覚的注意モジュールは視覚的注意を集める傾向がある画像内のオブジェクトを予測するように機能する、(4)識別されたオブジェクトのうちのどれが、視覚的注意モジュールによって予測されるかを決定する。

【図面の簡単な説明】

40

【0010】

【図1】代表的な視覚的注意モジュールの上位レベルの機能を図示するフローチャート。

【図2】代表的なシーンのスケッチ。

【図3A】視覚的注意モジュールからの出力の、アーティストのレンダリング。

【図3B】視覚的注意モジュールからの出力の、アーティストのレンダリング。

【図4】2つの例示のオブジェクトのプロパティを示すグラフ。

【図5】異なる背景のシーンに対して、どのように広告オブジェクトが設定され得るかを示す。

【図6】シーン又はシーン内のオブジェクトの堅牢性を評価するための、システム内の機能性モジュールのダイアグラム。

50

【図 7】例示のシーンの修正を示す図。

【図 8】視覚的注意堅牢性評価システムの他の実施形態を示し、ここでは視覚的注意堅牢性評価システムは更にウェブサーバーモジュールを含む。

【図 9】視覚的注意堅牢性評価システムが、シーン内のオブジェクト又はシーンそれ自体の堅牢性を評価するために利用することができるプロセスの 1 つの実施形態を示す上位レベルのフローチャート。

【図 10】シーン最適化の上位レベルのプロセスを示すフローチャート。

【図 11】注意サッカードの、関心のオブジェクトに対する系列位置の関数として、2 つの異なる報酬構造を示すグラフ。

【図 12】シーン最適化のための、システム内の機能性モジュールのダイアグラム。

【図 13 A】2 つの対応するシーンの、アーティストのレンダリング。

【図 13 B】2 つの対応するシーンの、アーティストのレンダリング。

【図 14 A】2 つの対応するシーンの、アーティストのレンダリング。

【図 14 B】2 つの対応するシーンの、アーティストのレンダリング。

【図 15】複数の視点のシーン解析を示す上位レベルのフローチャート。

【図 16】会議室の図。

【図 17 A】会議室の異なる視点を示す。

【図 17 B】会議室の異なる視点を示す。

【図 18】会議室内の様々なオブジェクトに実施された 3 つの異なる例示の解析を示すグラフ。

【図 19】複数の視点のシーン解析システムの図。

【発明を実施するための形態】

【0011】

視覚的注意モデル

視覚的注意モデルは、シーン内の領域が、それらが視覚的注意を引き付ける、又は眼球動作を引き付ける尤度に関して異なる程度を決定する。シーンは、視覚的注意モデルによる評価に適している入力の入力の任意の形態（例えば、グラフィック画像）であり、例えば、デジタル写真、バーチャル 3D シーン、ウェブページ、ドキュメント、又はビデオであってもよい。

【0012】

オブジェクトへの視覚的注意モデルの配置は、視覚的注意のモデルが、シーン内のオブジェクトの相対顕著性をどのように特徴付けるかということを指す。例えば、一部の視覚的注意モデルは、予測されたオブジェクトの周辺においてトレースラインを重ね合わせる。他の視覚的注意モデルは、画像の上に重ね合わされ得る、又は画像から別々に見られ得るヒートマップを生成する。更に他のものが、オブジェクトの顕著性を関連用語で示す値と共に、特定のオブジェクト及び/又は領域に値を生成し、割り当ててもよい。トレース線との関連で、オブジェクトの配置は、モデルによって「選択された」（トレースされた場合に）、又は「選択されていない」と見なされ得る。ヒートマップとの関連で、オブジェクトの配置は、アルゴリズムがオブジェクトを選択した（又はオブジェクトを選択しない）程度である。顕著性の数との関連で、オブジェクトの配置は、顕著性の数それ自体であってもよい。

【0013】

シーン内のどこに、人の視覚的注意が割り当てられるかを予測する多くの視覚的注意モデルが存在する。一般的に、これらの視覚的注意モデルは、入力として環境の単一の画像をとり、そのシーン内のどこに注意が割り当てられるかに関する予測を生成する。実証的アプローチは、ヒトをシーンに対して露出させ、彼らの眼球運動を追跡する。この実証的アプローチはリソース集中型ではあるが、しかしシーンの少なくとも一部分において、解析によって注意を予測しようと試みる多くの数学モデルが開発されてきている。この前述の実証的アプローチは、本明細書で記載されるシステム及び方法の一部として使用することができ、本明細書で使用されるとき、実証的アプローチは、視覚的注意モデルリングの

10

20

30

40

50

1つのタイプと見なされる。

【0014】

3つのモデルのうちの1つの基本的な方法論は、図1に示されており、これはItti, L. & Koch, C. (2000)の「A saliency-based search mechanism for overt and covert shifts of visual attention」(Vision Research, vol. 40, 1489~1506頁)によって提案されている。上位レベルでは、図1は、「ボトムアップ」特徴(例えば、色、動き、輝度、エッジ等)の評価によって、どのように予測が視覚的注意に対してなされるかを示し、これらはヒトの視覚の一部の視点を調節する、視覚的表示の構成要素として機能する。まず、デジタル写真の形態のシーンは、Itti & Kochモデル(ステップ10)のコンピュータ実行型版へ提供される。次いで、特徴抽出プロセスは、色、強度、配向、又は他のシーンのキュー、例えば、動き、中継、ターミネータ、又はステレオ視差、及び陰影からの形状に対するデジタル写真を解析する(ステップ11)。特徴抽出プロセスは、複数の特徴マップを生じさせ(ステップ12)、これは組み合わせられて顕著性マップを作る(ステップ13)。Itti & Kochモデルの場合では、顕著性データは、視覚的注意が次に割り当てられるとモデルが予測している「最も輝度の高い」オブジェクトを有するオリジナルのデジタル写真のレンダリングとして、ユーザーに提供される。この予測されたオブジェクトは、「競争式」タイプのアルゴリズムにおいて、視覚的に突出していると識別され(ステップ14)、そのプロセスは次いで、複数のオブジェクトがモデルによって識別されるまで繰り返す(ステップ15)。

【0015】

図2は、例えばItti & Kochの視覚的注意モデルに提供され得るシーン201の、アーティストのレンダリングである。それは説明の目的のみの本明細書に含まれる簡易化されたシーンであり、実際には、シーンは実際のデジタル写真又はビデオであることが多く、はるかに複雑である。図2は、例えば星202、花203、顔204、星205、矢印206、カップ207など、シーン内に多くのオブジェクトを含む。

【0016】

図3Aは、Itti & Kochモデルからの出力がどのように表され得るかを示す、アーティストのレンダリングである。ハイライトされた(かつ、この代表的な図においては、丸で囲まれた)オブジェクトは、モデルが、視覚的に顕著であると予測するものである。例えば、星202は、この図ではハイライトされた縁208内であり、花203は縁209内であり、顔204は縁210内であり、星205は縁211内であり、矢印206は縁212内であり、カップ207は、縁213内である。したがって、この例におけるモデルは、他のオブジェクトに対して、より視覚的に顕著である6つのオブジェクトを決定している。この特定のモデルはまた、一部の視覚的顕著性の閾値よりも上であると決定されたオブジェクト間で、どのように注意が移動するかも予測する。例えば、経路301、302、303、304、及び305は予測された視覚的注意を示す。

【0017】

図3Bは、Itti & Kochモデルからの出力が表されることもある、更なる方式を示す、第2のアーティストのレンダリングである。図3Aに示されているものに加えて、図3Bは、予測された視覚的注意のシーケンスを含む。例えば、星202は、「1」とラベル付けされ(注意シーケンス番号214)、花203は、「2」等とラベル付けされる(注意シーケンス番号215)。

【0018】

勿論、図3A及び図3Bは、視覚的注意の予測が、ユーザーに伝達され得る1つの方法に過ぎず、異なるモデルが異なる方法でこの情報(又はその一部の部分集合)を表す。例えば、全てのモデルが、予測された注意シーケンスを決定するわけではないが、そのような注意シーケンスは、視覚的顕著性の最高度を有するオブジェクトを決定し、次いでそのオブジェクトを排除し、次の上位レベルのものをを見つける等によって達することができ

る。

【0019】

Itti & Kochのモデルは、「ボトムアップ」の視覚的注意モデルの代表的なものであり、そこでは、このモデルは、シーンの特定のものの解析に基づいてその予測を行う。他のボトムアップの視覚的に顕著なモデルが、これらの参照：Gao, Mahadevan, and Vesconcelos (2008)に記載されている。

【0020】

ボトムアップモデルに加えて、視覚的注意の「トップダウン」モデルと呼ばれる別のクラスのモデルがある。ボトムアップモデルと対照的に、これらのモデルはシーンと、明白なタスク（例えば、障害を避けること及びオブジェクトを収集すること）、又は特定の検索タスク中に、注意が割り当てられる場所に影響を与える、対象世界に関する事前知識（例えば、椅子はどちらかといえば床の上にあり、天井にはない）のいずれかで始まる。この知識（タスク及びシーンに基づく）は、ボトムアップ特徴と併せて使用され、観察されたシーン内のオブジェクトに注意を向けさせる。一部の代表的なトップダウンモデルがRothkopf, C. A., Ballard, D. H. & Hayhoe, M. M.の「Task and context Determine Where You Look」(2007, Journal of Vision 7(14):16, 1~20)に、及びTorralba, A.「Contextual Modulation of Target Saliency」(Adv. in Neural Information Processing Systems 14(NIPS)(2001)(MIT Press, 2001)にも記載されている。例えば、視覚的注意のTorralbaのモデルは、特定のタイプのオブジェクトを含む特徴に関する事前知識、及びシーン内のこれらのオブジェクトの絶対位置及び相対位置に関する情報を有する。この事前知識は、シーン内の特定の標的の検索において「トップダウン」の影響をもたらす。

【0021】

当該技術分野は、ボトムアップ及びトップダウン設計の両方の特徴を有するハイブリッドな視覚的注意モデルを含むように発達し、モデルが露出されるシーンのタイプ（例えばビデオ対静止画像、屋外画像対ウェブページ等）における相違に適合している。

【0022】

堅牢性

堅牢性とは視覚的注意モデルから以下のいずれか、あるいはこれらのいくつかの組み合わせに対して出力された予測の感度を指す。

【0023】

(a) 例えばシーン内のオブジェクトの照明、オブジェクトの色等（「外部変動」と呼ばれる）を含む、シーンの視覚的特性内の変化及び/又は変動、又は、

(b) 観察者内の変化及び/若しくは変動、又は観察のモデル（「内部変動」と呼ばれる）。

【0024】

用語「オブジェクト」は、本明細書で使用されるとき、シーン内のもの又はエリア若しくは領域を、場合によっては、あるいは、視覚的注意モデルによって解析されるシーン内の領域のいずれかを指す。用語「オブジェクト」は、場合により「エリア」又は「領域」と同じ意味で用いられる。

【0025】

2つのタイプの変化(a)及び(b)は、総称して内部又は外部、即ちIE変化と呼ばれ、かかる変化を生じさせ、次いでこれを評価する様々な方法が以下に更に記載される。

【0026】

堅牢性評価は、シーン内のどこで注意が割り当てられるかに対するIE変化の影響を測定する方法である。シーン内のオブジェクトの堅牢性は、次いで、IE変化に関わらず、視覚的注意モデルの予測した、特定のオブジェクトが変化する、又は変化しない配置の程度である。

【 0 0 2 7 】

オブジェクトのグループの堅牢性は、2つ以上のオブジェクト（オブジェクトの集合体）に対して視覚的注意モデルの配置が、I E 変化と共に変化する、又は変化しない程度である。

【 0 0 2 8 】

シーンの堅牢性は、シーン内の視覚的注意モデルの、オブジェクトの配置が、I E 変化に関わらず、変化する、又は変化しない程度の尺度である。例えば、視覚的注意モデルが、シーン内の同じ4つのオブジェクトにされる注意又は凝視を予測し、かつこれらの4つのオブジェクトがI E 変化に関わらず、視覚的注意モデルによって予測されたままの傾向にある場合、4つのオブジェクトがI E 変化に続いて変化した場合よりも、シーンは、より堅牢である傾向がある。

10

【 0 0 2 9 】

視覚的注意堅牢性評価システム

図6は、視覚的注意堅牢性評価システム403内の代表的な機能性モジュールを示す図であり、これはシーンの堅牢性、シーン内のオブジェクトの堅牢性、又は複数のシーンにわたるオブジェクトの堅牢性を評価するためのシステムである。勿論、これらの方法は手動で実施されてよい。図6に示す実施形態において、視覚的注意堅牢性評価システムは、コンピュータシステム408内である。コンピュータシステム408は、いずれかの汎用又は用途専用コンピュータ若しくはデバイスであってもよい。それはスタンドアローンのラップトップコンピュータ、又は複数のネットワーク接続コンピュータであってもよい。更にコンピュータシステム408は、携帯式コンピュータ、デジタルカメラ、若しくはタブレットPC、又は更には携帯電話であってもよい。コンピュータシステム408は、一実施形態では、オペレーティングシステムを含む様々な機能性モジュール（図6には示さず）を有する。かかるオペレーティングシステムは、コンピュータシステムのリソースに対する視覚的注意堅牢性評価システムのアクセスを促進する。コンピュータシステム408は、プロセッサ及びメモリ、並びに従来の入力／出力インターフェースを有する。

20

【 0 0 3 0 】

視覚的注意モジュール403は、任意の視覚的注意モデル又はモデルの組み合わせの任意の実施形態である。先に記載したように、異なるタイプの視覚的注意モデルが存在するが、それらは全て、ある程度、そこに視覚的注意が割り当てられる傾向にある、シーン内のオブジェクト又はエリアを予測する。視覚的注意モジュール403は、視覚的注意堅牢性評価システム402の一部として図6に示されているが、他の実施形態における視覚的注意モジュール403は、スタンドアローンコンピュータプロセス、又は更にはリモートコンピュータにおいてコンピュータネットワークのいずれかのタイプ（例えばWorld Wide Web）にわたって提供されるサービスとして動作する。

30

【 0 0 3 1 】

VAM修正モジュール404は、視覚的注意モジュール修正モジュールである。VAM修正モジュール404は、視覚的注意モジュールのパラメータ又はアーキテクチャの態様を修正する。この修正は、視覚的注意モジュール403の実施によって決まる多くのやり方で達成することができる。例えば、視覚的注意モジュール403はそれ自体が、どのように視覚的注意モジュールが働くかという態様を修正する関数コールを支持してもよい。一実施形態では、視覚的注意モジュール403は、シーンの特定の態様（例えば輝度）に与えられた重み付けを修正する関数コールを支持してもよい。他の実施形態では、視覚的注意モジュールがコマンドラインを介して呼び出された場合、様々なスイッチを利用して、視覚的注意モジュール内にある変数を変更することができる。あるいは、視覚的注意モジュール403がスクリプト又はプログラミングコード内に統合されている場合、I E 修正モジュールは、スクリプト又はプログラミングコードそれ自体を修正することができる。他の実施形態では、全体の視覚的注意モデルは、別の視覚的注意モデルによって置き換えられる。VAM修正モジュール404が視覚的注意モジュール403の下層の視覚的注意モデル（又はかかるモデルの、シーンへの適用）を修正する特別な方法が更に以下に記

40

50

載されているが、一例は、顕著性マップを生成するために使用される複数の特徴マップに関連する重み付けを修正するためのものであってもよい。これは、許容可能な値の分布に基づいて、これらの値を修正する関数コールを介して達成することができる。

【0032】

シーン修正モジュール405は、解析のために視覚的注意モジュール403に提供されるシーンを修正する。シーン変更は、環境のダイナミックな性質をシミュレーションする、外部環境に関連する変更である。このような外部変動には、例えば、シーン内のオブジェクトを動かすこと（例えば、歩行者の配置）、照明方向における変更による影の変化、又は大気条件での変化（例えば、空中の塵）を挙げることができる。この変動は、多くの方法で生成することができる。1つの方法は、シーン内にカメラを配置し、異なる時間で画像を捕捉することである。これは、実際のシーンの自然変動を捕捉する。別の方法は、シーンの単一の画像を捕捉し、個々の要素及びそのシーン内のオブジェクトにされ得る、可能性のある変更を特定することである。かかる技法の図は、図7に示される。シーン801は、最初のシーンの表示である。シーン802は、シーン変動要素、例えば影及び他のオブジェクト（例えば、車、鳥、シーン内に配置され得る任意のオブジェクト）で置換されたシーン801からのオブジェクトを示す。シーン803、804、805、及び806は、シーン変動要素を有する最初のシーン801の異なる組み合わせを示す。当業者は、例えば商標「Photoshop」でAdobe Systems (San Jose, California) から販売されているように、市販されている画像編集ソフトウェアを使用するなど、シーンを修正する、数えきれない異なる手法があるということ

10

20

【0033】

シーン修正モジュール405は、シーンに外部変動を追加し、これはひいては、複数のバージョンのシーン又はシーンの例となる。シーンの例の収集は、一実施形態では、シーン内の変動を捕捉する。オブジェクトの堅牢性又はシーンの堅牢性を測定するために、各シーンの例は、視覚的注意モジュール403に送られ、視覚的注意が各シーンの例内のどこに（すなわち、どのオブジェクトに）割り当てられるかという予測を作り出す。各ランに関連する情報はデータベース407内に保存され、次いで堅牢性評価モジュール409は、データを評価し、これらの異なる例にわたって、注意が割り当てられるであろうと、モデルが予測したオブジェクトに基づいた統計値を生成する（更なる詳細で先に記載されたとおり）。

30

【0034】

グラフィカル・ユーザー・インターフェースモジュール406は、ユーザー401との相互作用を促進する。グラフィカル・ユーザー・インターフェースモジュール406は、例えばユーザー401からの入力を求めるための、グラフィカル・ユーザー・インターフェースを構築するために、オペレーティングシステムのリソース（コンピュータシステム408からの）を求めることができる。一実施形態におけるこの入力は、シーンの場所、並びに視覚的注意堅牢性評価システムのための他の動作パラメータを含む。この入力は、一実施形態では、ユーザー401が評価に関心があるシーン内の領域及び/又は場所を特定する。かかる場所を特定することに加え、ユーザー401は、視覚的注意モジュール403によって、どのようなタイプの変動が考慮されるかということ特定することができる。これは一般的若しくは特定の内部変動、又は一般的若しくは特定の外部変動を含んでもよい。例えば、1つの特定のタイプの外部変動は、照明方向における変動により、画像を変更してもよい。光が変わると、作られた影も変化する。これは特定のタイプの外部変動であり、なぜならば、それは例えばダイナミックなオブジェクト、大気収差等の他の要因を考慮しないためである。一般的な内部変動の例は、各特徴マップのための重み付けが独立して変化するのが可能である状態である。特定の内部変動の例は、特徴マップの1つのセットの重み付けが変化するが（例えば輝度）、他のものは変化しないときである。グラフィカル・ユーザー・インターフェースモジュール406もまた、場合によってユーザーからの、どこからシーンの画像を得るべきであるかに関する入力を得るのを促進する。

40

50

可能性のある場所には、例えばデータベース又はフラットファイルが挙げられる。

【0035】

堅牢性評価モジュール409は、シーン内のオブジェクトの堅牢性又はシーンそれ自体の堅牢性を評価するために、他のモジュールの相互作用を制御する。例えば、堅牢性評価モジュール409は、不可欠なものとして、視覚的注意モジュール403、並びにVAM修正モジュール404及びシーン修正モジュール405を呼び出す。堅牢性評価モジュール409が様々なモジュールを呼び出す方法は、例えばグラフィカル・ユーザー・インターフェースモジュール406を介して堅牢性評価モジュール409に提供された、ユーザー401からの入力によって修正されてもよい。堅牢性評価モジュール409は、他のモジュールによって提供されたデータも評価し、必要に応じてレポートを生成する。

10

【0036】

データベース407は、視覚的注意堅牢性評価システム402のデータ保管への需要に対処する。特に、データベース407はシーンの画像を保持することができる。データベース407は任意のコンピュータメモリであってもよい。それは、ランダムアクセスメモリ、又はフラットファイル、又は1つ以上のデータベースサーバーを実行する1つ以上のデータベース管理システム(DBMS)であってもよい。データベース管理システムは、リレーショナル(RDBMS)、階層(HDBMS)、多次元(MDBMS)、オブジェクト指向(ODBMS若しくはOODBMS)、又はオブジェクトリレーショナル(ORDBMS)データベース管理システムであってもよい。データベース407は、例えば、Microsoft CorporationからのSQLサーバーなど単一のリレーショナルデータベースであってもよい。

20

【0037】

ユーザー401は、視覚的注意堅牢性評価システムの任意のユーザーである。一部の実施形態では、視覚的注意堅牢性評価システム402の使用は十分に優しいため、視覚的顕著性理論になじみの薄い人がシステムを使用して、オブジェクト、領域、又はシーンの堅牢性を評価できる。ユーザー401は、彼らの環境内のサイン及び非サインオブジェクトの位置決めを評価するために、企業で働く従業員によって雇用されたコンサルタント又はこの従業員であってもよい。ユーザー401は、ページの視覚的特性が変化し得るウェブページ上のデザイン及びデジタルオブジェクトの配置の評価に関心のあるコンテンツデザイナーであってもよい。

30

【0038】

図8は、視覚的注意堅牢性評価システムの他の実施形態を示し、ここでは視覚的注意堅牢性評価システムは更にウェブサーバーモジュール501を含む。ウェブサーバーモジュール501は、便宜上、視覚的注意堅牢性評価システム402の一部として示される。ウェブサーバーモジュールは、しかしながら別個のメモリ空間内でコンピュータシステム408上で動いているソフトウェアモジュールとして具体化されてもよい。あるいは、ウェブサーバーモジュール501は、ネットワークを介して視覚的注意堅牢性評価システム402に連結された別個のコンピュータシステム上にあってもよい。

【0039】

ウェブサーバーモジュール501は、インターフェースを提供し、これによってユーザー401は、クライアントコンピュータ503を介して、かつネットワーク502を介して、視覚的注意堅牢性評価システム402と通信することができる。1つの構成では、ウェブモジュール501は、Microsoft Corporation(Redmond, Washington)からのInternet Information Serverなどのウェブサーバーソフトウェアを実行する。ウェブサーバーモジュール501は、例えば、Active Server Pages(ハイパーテキスト・マークアップ・ランゲージ(HTML)で書かれたウェブページ)又はダイナミックHTML、ActiveXモジュール、Lotusスクリプト、Javaスクリプト、Java Applet Distributed Component Object Modules(DCOM)の使用を介して、遠隔ユーザー401と相互作用するためのメカニズムを提供

40

50

する。

【0040】

コンピュータシステム408によって提供されるオペレーティング環境内で実行する「サーバー側」のソフトウェアモジュールとして示されているが、視覚的注意堅牢性評価システム402を含む機能性モジュールは、ユーザー401によって使用されるとき、コンピューティングデバイス、例えばクライアントコンピュータ503上で実行する「クライアント側」ソフトウェアモジュールとして容易に実行することができる。視覚的注意堅牢性評価システム402は、例えば、クライアントコンピュータ503上で実行するウェブブラウザによって実行されるActive Xモジュールとして実行することができる。

【0041】

ネットワーク502は、任意の種類のネットワーク、公開又は非公開であってもよい。一実施形態では、ネットワーク502はインターネットである。

【0042】

図7及び8でのユーザー401はインターネットの任意のユーザーであってもよい。一実施形態では、ユーザー401は、視覚的注意堅牢性評価システムの機能性にアクセスするために、ライセンスによって事前に配置されていてもよい。

【0043】

図9は、視覚的注意堅牢性評価システム402が、シーン内のオブジェクト又はシーン自体の堅牢性を評価するために利用することができるプロセスの1つの実施形態を示す上位レベルのフローチャートである。まず、視覚的注意モジュールは、入力されたシーンを呼び出し、これが提供される(ステップ601)。シーン内から予想されるオブジェクトは次いで、視覚的注意モジュールから受信される(ステップ602)。次いで、モデルが、高度な相対顕著性を有すると予測する場所/オブジェクト/領域の一部の表示が、データベースに保存される(ステップ603)。はっきり、何がデータベースに保存されるかは、視覚的注意モジュールからステップ602で受信された出力のタイプに大きく依存する。データベースは、一実施形態では、閾値の顕著性の値に到達するオブジェクトに関する情報を保存する。他の実施形態では、データベースは、マトリックスの値(イメージ内に識別された各オブジェクトに対して1つ)を保存し、例えば、オブジェクトが閾値の顕著性の値に到達した回数を保存する。次いで、プロセスが完了したかを見るためのチェックがなされる(ステップ605)。一実施形態では、このチェックは、現在の繰り返しができる。他の実施形態では、繰り返しの数は最初に設定された繰り返しの数を超えているかを決定することができる。他の実施形態では、繰り返しの数はアルゴリズムによって決定されてもよい。更に他の実施形態では、繰り返しの数は、堅牢性の尺度を考慮して、シーン又は統計のプロパティによって決定されてもよい。例えば、特定のサイトにおいて、2つの広告のうちどちらがより良いかを決定しようとする場合、2つの広告の堅牢性の値の間に統計的に信頼性のある効果があるまで、アルゴリズムを走らせることができる。

【0044】

堅牢性評価が完了していない場合(ステップ605において「いいえ」)、IE変化は、視覚的注意モデルの視覚的注意モデル又は入力されたシーンのいずれかに生じさせられる。IE変化は、構造化及びランダムな2つのカテゴリに分けることができる。ランダム変動は相関性のない変動である。例えば、シーン内のランダムな変動は、個々のピクセルの色及び/又は明度におけるランダムな変動を含んでもよい。そのような場合、ピクセルの変動は相関関係がない。対照的に、構造化変動は、修正される要素間で相関関係を有する。例えば、シーン内のオブジェクトの動作をシミュレーションすることによって、又はシーン内のオブジェクトを追加する若しくは除去することによって修正されるシーンは、構造化されたシーン変動を構成する。そのような場合、ピクセル修正における変化は相関関係がある。ランダムな内部変動は、視覚的注意モジュールによって使用される注意モデルにおけるランダムな変動を含んでもよい。構造化変動は、一方では、他に対して注意マップの一部における注意のためのプログラムのバイアスであり得る。変動を生成するタイプの方法が表1にまとめられている。IE変化を生じさせるサブプロセスは、以下に更に

10

20

30

40

50

詳細に記載されている。

【 0 0 4 5 】

【 表 1 】

表 1

	内部	外部
構造化	内部構造化変動を生成する多くの方法がある。このタイプの変動は、マップ又は空間的な領域にわたって共変動によって定義される。1つの方法は、顕著性マップへの、異なる特徴マップの重み付けを変化させることであろう。重み付け値は、色、配向、輝度等の個々の好み又は予測をシミュレーションしながら、1つの試行から次へとわずかに変化してもよい。他の方法は、場所の全てが、顕著性「ブースト」が与えられている顕著性マップの1つの領域（例えば下方領域）のモデルバイアスを有することである。	ピクセルの集合が共変するイメージ内の変動。この共変は、イメージ内の領域を定義し、そのような領域内の R G B 値を修正することによって生成され得る。特定の領域内のピクセルの全ては、同程度に修正される。領域から領域への、又は試行から試行への変動の量は、ガウス分布の標準偏差によって設定され得る。第2のパラメータは、共変し得る領域の「サイズ」であってもよい。
ランダム	ランダム変動は、どこに注意が割り当てられるかを視覚的注意モデルが決定する前に、顕著性マップの出力に追加される。変動は、任意の確率分布の形態であってもよいが、1つの例の実施は、ガウスの平均が、予測された顕著性の値に設定され、ガウスの幅が適切な標準偏差によって設定されるガウス分布を使用することである。	R G B 値がランダムに修正されているピクセル毎の変動である。修正の程度は、ガウスの平均が「実際の」 R G B 値に設定されているガウス分布の標準偏差によって特定されてもよい。変動に対する他の方法も使用することができる。

【 0 0 4 6 】

いったん I E 変化がなされると、プロセスはステップ 6 0 1 に戻る。

【 0 0 4 7 】

いったんプロセスが完了すると（ステップ 6 0 5 における「はい」）、視覚的注意モジュールによって予測されたオブジェクトが一連の繰り返しにおいて、どのように、そしてどの程度変化したかに対して解析がされる（ステップ 6 0 6）。この解析は、以下に更に詳細が記載される。

【 0 0 4 8 】

最終的に、堅牢性を示す出力が生成される（ステップ 6 0 7）。一実施形態では、この出力は、グラフィカル・ユーザー・インターフェースモジュール 4 0 6 を介してユーザーに提供される。しかしながら、出力は他の形態に、例えば他のプログラム又は呼び出し機能で提供されてもよい。

【 0 0 4 9 】

視覚的注意モデルへの変化

視覚的注意モデルを修正する（前述の図 9 におけるステップ 6 0 4 において生じさせることができる変化のタイプ）の1つの例は、シーン内の、又はシーン内の特定のエリア内

の特定の特徴に向けて、バイアスをプログラムで生じさせることである。視覚的注意モデルへのかかる変動は、注意が割り当てられる場所への影響を有し、観察者及び／又は次から次へとシーンを経験するヒト観察者間の変動をシミュレーションする。例として、シーンの視覚的注意堅牢性を評価するために I t t i & K o c h のボトムアップの視覚的顕著性モデルを使用してもよい。そのようなモデルでは、顕著性「マップ」は、異なる視覚的特徴マップからの入力を組み合わせることによって生成される。例えば、いくつかの具体例では、3つの特徴マップがあり、第1は色に対して感度が高く、もう1つは配向に対して、第3は輝度に関連するものである。顕著性マップへの入力、これらの3つのマップの重み付けされた組み合わせである。典型的に、これらの3つのマップは顕著性マップに同一の重み付けを有し、もう一方に対して1つのタイプの特徴に対するバイアスはないということを示し、かつ重み付けベクトルとして示され得る（例えば、3つの機能マップの等しい重み付けに対して $[1 \quad 1 \quad 1]$ ）。観察者のバイアスにおける変動を生成する1つの方法は、もう一方に対して1つの特徴の方にバイアスされる可能性のある視聴者をシミュレーションしながら、これらの重み付けを修正することである。例えば、ある人は環境内の品目を、 $[0.5 \quad 0.5 \quad 2]$ に重み付けを設定することによって、より明るいものに偏倚させ得る。内部の変動を生成する（又は視聴者の変動をシミュレーションする）1つの方法は、異なる重み付け値を有するモデルの集合体を生成することである。これは、各モデルに対して値をランダムに設定することによって実施され得る。

【0050】

別のアプローチは、シーンのモデルの内部の表示をランダムに表示することである。例えば、ヒト視覚的システムにおけるニューロンは、幾分ノイズが多いため、それらの活性化は同じ画像の表示でさえも変化する。この変動をシミュレーションすることは、視覚的注意モデルがシーンに対応して発達する、個々の特徴マップ（例えば、色、形状、配向、輝度）に関連する値の内部の表示を混乱させることによって実施され得る。

【0051】

他のモデルは、画像のどの領域が最も顕著であるかを識別するのに使用される顕著性マップにおける値を混乱させることである。例えば、それぞれの凝視計算に続いて、内部の表示を混乱させる、又はシーンが表示される度に値を独立して混乱させてもよい。

【0052】

シーンに対する変更

シーンは、シーンのプロパティが変化される、既に存在しているシーンに基づいて、グラフィックエディタ（例えば P h o t o s h o p ）、シーンの3D表示（例えば、パーティクルリアリティモデル）、又はビデオ（また、本明細書において全て総称して「シーン」と呼ばれる）を使用して新しい画像を作ることによって変えられる。例えば、そのような変化したプロパティは、照明においてシミュレーションされた（又は実際の）変更、又は新しく追加された、又は取り除かれたオブジェクト（実際の、又はパーティクルの）を含むことができ、すなわちピクセルの色をランダムに変化させる。

【0053】

シーンを変化させるのに多くの異なる方法がある。どのような外部変化を作るべきかの決定は、状況から状況へと変化してもよい。シーンにどのようなタイプの変化をなすべきかを決定する1つの方法は、シーン内に典型的に発生するシーン変動のタイプを決定し、解析される画像にこれらの変化を生じさせることである。例えば、ホテルのロビーの人々は、シーンにわたって動いている。時には、ゲストはカウンターに立っており、時にはエレベータで立っており、又はあるいはある場所からある場所へ歩いている。これらの状況をシミュレーションすることは、実際のシーンにおいて変動を捕捉するための方法を提供する。

【0054】

外部変動がシーンに追加される程度を変化させたい、及び／又は測定したい場合がある。外部変動の程度を定量化するための1つの方法は、シーン内に操作されるオブジェクトの数を変化させることである。例えば、歩行者を有するホテルロビーの例に戻りながら、

シーン内の歩行者の数を変化させてもよい。歩行者が少ないのは、小さな量の変動に変換し、多くの歩行者は多くの変動に変換する。変動を測定する他の方法は、生成されるイメージの全てに対してピクセル変動を測定することである。ピクセル変動の1つの例示の尺度は、生成されるイメージのセットに対する各ピクセルの平均分散変動を測定することである。いったん各ピクセルの変動が算出されると、ピクセルの全ての平均分散量を算出することによって、単一の数字を出すことができる。また、イメージ内の変動が増加するとき、この値もまた増加する。

【0055】

変更を作る1つの方法は、例えば商標「Photoshop」でSan JoseのAdobe Systems Inc.によって販売されているような、市販の画像編集ソフトウェアを使用して、デジタルで画像を修正することによって、画像を変更させることである。この方法を使用しながら、デジタルでオブジェクトを配置する、オブジェクトを除去する、又は照明変化をシミュレーションすることによって、シーンを修正することができる。これらの変化は次いで、モデルによって読み込まれ、解析される画像の集合体に変換される。これらの変化はまた、レイヤーを画像にアルゴリズム的に重ね合わせることで、自動的にイメージに適用することができる。

10

【0056】

外部の構造的変動を生成するための他の方法は、シーンのバーチャルな表示を生成することである。このアプローチを使用して、オブジェクトの配置、オブジェクトのプロパティ、及び照明条件を容易に修正することができる。これらのバーチャル表示から、画像は、オブジェクト、それらのプロパティ、それらのポーズ及び照明条件が変化された状態で生成され得る。

20

【0057】

これまで、シーンを変化させるアプローチは「合成」であるといえ、すなわち、それらはシーンそれ自体というよりはむしろ、シーンの表現に対する変化である。シーンを修正するための別のアプローチは実際のシーンを修正することである。例えば、一連の時間差の画像（例えば、定義された期間で300秒毎の写真）を撮ってもよい。この一連の写真は次いで、視覚的注意堅牢性評価システムの、連続する繰り返しに使用することができる。そのような解析（時間にわたるマルチ画像）は、外部の構造変動の尺度を提供する。更に、そのような変動を与えられて（例えば、太陽の位置における変化により、照明方向における変化と共に移動する人々及びオブジェクト）を与えられて、環境内のどこに注意が一般的に割り当てられるかということを解析することができる。

30

【0058】

外部及び内部の変動の追加に加えて、無差別に、シーン内の異なるオブジェクトの視覚的顕著性による変動を生じさせることもできる。例えば、注意が割り当てられる変動の程度を減少させ、かつ注意が割り当てられないこれらの領域に対して注意を増加させたい場合がある。

【0059】

堅牢性の評価

例えば、「ヒートマップ」グラフィック表示、内部エントロピー値、堅牢性係数、又は堅牢性の値などを使用する、シーンの堅牢性を特徴付ける多くの方法がある。視覚的注意モデルをシーンに繰り返し適用し、その一方で、各繰り返しに対して、あるタイプのIE変化を提供することから統合された結果を要約する1つの基本的なアプローチは、シーンの表示上に重ねられる相対値又は領域のグラフィック表示を生成することである。そのようなアプローチは、ヒートマップに類似もののものであり、クエスチョンにおける変動に対し、より堅牢性である、より「明るい」オブジェクトを備える。

40

【0060】

視覚的注意モデルの予測を阻害するノイズの影響の評価に加えて、堅牢性を評価する他の方法は、シーンの、モデルの内部表示における変動性を測定することである。例えば、Itti & Kochモデルは、内部の「顕著性マップ」表示を使用して、視覚的注意

50

が配置される場所を決定する。I E 変化の関数として、顕著性マップのこの内部表示における変化の量又はエントロピーを、堅牢性の測定基準として測定することができる（方法だけでなく）。個々の特徴マップの、又は視覚的注意モデルの他の内部の表示の応答を見て、I E 変化の影響を測定することもできる。

【 0 0 6 1 】

堅牢性の測定基準を生成する他の方法は、I E 変化がパラメータ化できるときである。これは、例えばランダムな輝度のノイズをシーンに追加する（したがって、ノイズ量を、イメージ内の輝度エネルギーに対して、ノイズの輝度エネルギーを修正することによって、変化させる）場合である。又は、視覚的注意モデルの接続の重み付けの変動を増加させ、エントロピーの測定基準として、変化の程度を特徴付けることができる。オブジェクト、オブジェクトの群、又はシーンの堅牢性を特徴付ける他の方法は、これらのオブジェクトが、それらの視覚的顕著性の階層を保持しながら、耐えることができるノイズの量を決定することである。この測定基準の方法は、生成された変動の量の関数として、オブジェクトが割り当てられる尤度を測定することである。高度に堅牢であるこれらのオブジェクト及び／又はシーンは、注意がノイズ要素に引き付けられ、オブジェクト／領域には引き付けられない前の大容量のノイズにも耐性がある。得られる測定基準は、オブジェクトからノイズに向けて注意が引き付けられる前に、オブジェクトが耐え得るノイズのレベルである。

【 0 0 6 2 】

図 4 は、オブジェクトに対する注意の確率対 2 つのオブジェクト（低ノイズ耐性を有する第 1 オブジェクト 4 5 1（低堅牢性の値）及び高ノイズ耐性（高堅牢性の値）を有する第 2 オブジェクト 4 5 2）に対する変動を示すグラフである。グラフは、オブジェクト 4 5 1 及びオブジェクト 4 5 2 に注意が割り当てられるかどうかにおける変動の効果を示し、それぞれは異なる対応の堅牢性係数を有する。この例では、少量の変動を追加することは、注意がオブジェクト 4 5 1 に割り当てられるということをモデルが予測するという確率に著しく影響を及ぼす。これは、少量の変動のみを有する確率における降下によって、図 4 に示される。対照的に、オブジェクト 4 5 2 は、少量の変動によって影響を受けない。これは、より大きなレベルの変動と共に発生する「減少」ポイントによって示される。これらの「減少」ポイントを測定することにより、堅牢性係数に関して、シーン内の異なるオブジェクトを定量化することができる。

【 0 0 6 3 】

ちょうど記載されたヒートマップのアプローチよりはむしろ、又はこれに加えてシーン内の特定のオブジェクト（又は複数のオブジェクト）の堅牢性を定量化することは時には有用であり得る。そのような定量化は、シーン内の特定のオブジェクトの顕著性の程度を低減するために始まる I E 変化の程度の決定することなど、後に続く解析を可能にし得る。

【 0 0 6 4 】

また、シーン自体の堅牢性を評価する、及び／又は定量化することは、時には有用であり得る。オブジェクトとシーン堅牢性との間の区分は、多くのタイプの、興味深い後に続く解析に対して門戸を開く。例えば、シーンが低い堅牢性の値を有するということが（すなわちモデルが、視覚的注意を集めると予測するオブジェクトのセットは、I E 変化の関数として変化する）が可能であり、同時にそのシーン内のオブジェクトは高い堅牢性の値を有する。「高い」オブジェクトの堅牢性の値及び「低い」シーンの堅牢性の値の例は、例えば、標的オブジェクトが、上位 3 つの注視状態に常にあるが、他のオブジェクトは著しく変化すると見なされるものである。

【 0 0 6 5 】

表 2 は、標的オブジェクトが、低い及び高い堅牢性の値を有するシーンにわたって高い及び低い堅牢性の値を有する例示のデータを提供する。この表において、1 つの、可能性のある堅牢性測定基準が示され、ここでは「堅牢性の値」と以下に呼ばれる正規化された尤度が、各オブジェクトに対して算出される。各シーンに対して、左の列は、I E 変化が

あるときに、オブジェクトが最初の3つのサッカーカードに現れる時間のパーセントを示す。表2に示される例において、シーン内には14のオブジェクトがあり、1つは標的オブジェクトであり、13ディストラクタであり、ディストラクタは、顕著性を有する非標的オブジェクト又は領域である。モデルがこれらのオブジェクトをランダムに選択する場合、各オブジェクトは 21.42% の選択される可能性を有する($3 \times 1 / 14 = 0.2142$)。これは、非常に少ないオブジェクト及び/又はシーン堅牢性があるという状況であり、したがって、それは重要な基準として働き、これによってシーンの堅牢性を比較する。これは、確率に機会のもに確率が近づくにつれて、堅牢性が減少するためである。この例では、堅牢性の値は、選択される可能性%によって分られた、数学的に注意が払われる%であり、したがって最初の堅牢性の値 $4.6667 = 100\% / 21.42\%$ である。シーンのための堅牢性の値を計算する1つの方法は、上位Kオブジェクト(注意選択の数)の平均堅牢性の値を算出することである。表2の下は、上位3つのオブジェクトの平均堅牢性の値である。

10

【0066】

【表 2】

堅牢性	標的オブジェクト=高 シーン=低		標的オブジェクト=高 シーン=高		標的オブジェクト=低 シーン=高		標的オブジェクト=低 シーン=低	
	割り当てられた	堅牢性の値	割り当てられた	堅牢性の値	割り当てられた	堅牢性の値	割り当てられた	堅牢性の値
オブジェクトID								
標的オブジェクト	100%	4.6667	100%	4.6667	5%	0.2333	9%	0.4200
ディストラクタ_1	15%	0.7000	90%	4.2000	97%	4.5267	30%	1.4000
ディストラクタ_2	15%	0.7000	5%	0.2333	5%	0.2333	26%	1.2133
ディストラクタ_3	17%	0.7933	2%	0.0933	2%	0.0933	15%	0.7000
ディストラクタ_4	14%	0.6533	80%	3.7333	88%	4.1067	32%	1.4933
ディストラクタ_5	13%	0.6067	3%	0.1400	91%	4.2467	28%	1.3067
ディストラクタ_6	21%	0.9800	2%	0.0933	2%	0.0933	22%	1.0267
ディストラクタ_7	18%	0.8400	1%	0.0467	1%	0.0467	32%	1.4933
ディストラクタ_8	19%	0.8867	4%	0.1867	4%	0.1867	33%	1.5400
ディストラクタ_9	16%	0.7467	2%	0.0933	2%	0.0933	38%	1.7733
ディストラクタ_10	20%	0.9333	1%	0.0467	1%	0.0467	18%	0.8400
ディストラクタ_11	14%	0.6533	2%	0.0933	2%	0.0933	17%	0.7933
ディストラクタ_12	18%	0.8400	8%	0.3733	8%	0.3733	9%	0.4200
ディストラクタ_13	19%	0.8867	0%	0.0000	0%	0.0000	30%	1.4000
シーンの堅牢性の値	2.1933		4.2000		4.2933		1.6022	

【0067】

表2は、注意が割り当てられる場所を、オブジェクトによって分類する。しかし、前述のように、用語「オブジェクト」は画像（即ち、シーン）内の領域又はエリアとして大まかに定義される。本明細書に記載される方法及びシステムは、オブジェクトベースのアプローチに限定されないが、他の同様なアプローチもまた機能する。例えば、等しいサイズ領域であり得るグリッドは、ヒト視覚的システム（例えば、視聴者の中心窩のサイズ）の

10

20

30

40

50

プロパティに基づいて、シーン又はシーンの領域にわたって定義される。

【 0 0 6 8 】

シーンにわたる堅牢性の評価

上記の方法がオブジェクト又はシーンの堅牢性を評価すると仮定すると、次に、他の方法で堅牢性の評価を広げることは可能である。例えば、「シーンにわたる堅牢性」とは、オブジェクト（又はオブジェクトのグループ）が、異なるシーンにわたってどのように変化するかという尺度である。オブジェクトに対する堅牢性測定基準を提供することにより（上記のように）、シーンから別個に、異なるシーンにおける特定のオブジェクトの堅牢性を評価することができる。例えば、広告主が、3つの異なるシーンにある、3つの異なる看板に配置される予定の2点の異なる広告コンテンツのどちらかに決めようとしている状況を考える。図5は、このシナリオの一例の表示を提供する。シーン703は看板700を含む。看板700は、シーン704及び705において繰り返される。同様に、看板706はシーン707、708及び709に見られる。

10

【 0 0 6 9 】

先に記載された方法（例えば、シーンのグラフィック表示に広告コンテンツを重ね合わせることを）を使用して、3つのシーンのそれぞれにおいて、各看板の堅牢性の値を決定することができる。すなわち6つの異なるオブジェクト堅牢性の値を生成する。顧客は次いで、一番高い平均の堅牢性の値（又は、堅牢性の値の集合に基づく他の測定基準）を有する広告（オブジェクト）を選択することができる。

【 0 0 7 0 】

20

1セットのオブジェクトを有するが、オブジェクトを配置するシーンのセットからも選択する必要があるときに、同様な解析を使用することもできる。この看板広告の例で続けると、顧客は、彼らが2つの異なるシーンに配置したい単一の広告を有してもよい。上記の技法を採用して、3つの異なるシーンにおいて各広告の堅牢性の値を生成するために使用され得るデジタル画像を生成することができる。シーン選択を促進するために、最も高い堅牢性の値を有する2つのシーン（広告がシーン内に重ね合わされた状態で）を顧客は選択することができる。

【 0 0 7 1 】

シーンの固定されたセットのためのオブジェクトの集合から、単一のオブジェクト（例えば、広告看板）を、又は単一のオブジェクトのためのシーンのセットを選択することに加えて、オブジェクトを堅牢性の値に基づいて特定のシーンに割り当てることができる。例えば、顧客は、彼らが開発した3つの異なる広告（看板上に配置されているコンテンツ）及びこれらの表示がその上に配置され得る20の異なる看板を有することができる。20のシーンに配置された3つの広告のそれぞれに対して堅牢性の値を生成することによって、顧客は次いで、各個々のシーンに対して一番高い堅牢性の値を有する広告の使用を選択することができる。更に、最も良い広告を、最も高い堅牢性の値を備える10のシーンに割り当てることができる。

30

【 0 0 7 2 】

これまで例示は、どのように堅牢性の値の測定基準が、看板及び広告に関してなされる選択を改善するために使用することができるかを示しているが、堅牢性の測定基準の効果は、この定義域に限定されず、店又はホテル内のデジタル表示、静止表示、製品表示、製品パッケージ構成、又はウェブサイト間など、コンテキストの集合体及び視覚刺激（オブジェクト）の集合体を有する任意の定義域において使用することができる。

40

【 0 0 7 3 】

シーン内のオブジェクトの堅牢性を評価するための、本明細書に記載のシステム及び方法は、視覚的注意を決定するためのいずれか特定の測定方法に依存しない。むしろ、それらは一般的に、視覚的注意を評価する任意のモデルと共に使用することができ、一部の実施形態では、複数の異なる視覚的注意モデルは、オブジェクト又はシーンの堅牢性を評価するために使用され得る。

【 0 0 7 4 】

50

堅牢性及び注視シーケンス

ここまでの考察は、大部分はシーン内のオブジェクトの堅牢性、又はシーン若しくはシーン内のオブジェの集合体の堅牢性に集中しており、オブジェクトが割り当てられる順序の考察がなかった。例えば、ここまでの解析は、注意がオブジェクトに割り当てられたか否かということモデルが実際に予測したかどうかを評価するのみであった。しかしながら、要素が実施に問題になる、順序における状況が存在する。例えば、マルチ - オブジェクトのシーンに対して、他のオブジェクトの前にどのくらいの頻度で1つのオブジェクトが現われるかということ、及び2つの別個のオブジェクトに対する注意の間に随行する介在オブジェクト（いわゆる介在オブジェクト）があるかどうかを知りたい場合がある。

【 0 0 7 5 】

10

変動をシーンに生じさせることに関して、上記の方法を使用し、注視シーケンスの集合体を有効にさせる。このデータを使用して、標的順序（例えば、オブジェクト B の前にオブジェクト A に注意を払う）が得られるシーケンスの全てを識別することができる。シーケンス堅牢性解析は、シーケンス標的堅牢性の尺度として標的シーケンスを取得する確率を使用することができる。シーケンス標的堅牢性を測定する1つの方法は、ランダムにオブジェクトを選択する場合、標的順序（すなわち、オブジェクト B の前にオブジェクト A）が発生するという尤度を算定することである。シーケンス標的堅牢性の値は、標的シーケンスが、それが偶然発生する尤度により割ることによって得られた確率である。この測定基準の背後の原則は、上記の堅牢性の値の背後の原則と同様である。

【 0 0 7 6 】

20

注意の相対順序が問題とはならないが、介在オブジェクトが問題となる第1の場合を考える。例えば、会社が、レストラン自体の近くで、看板上にレストランのディナースペシャルを広告している場合であるとき、目的は、オブジェクト B の前にオブジェクト A に引き付けられた注意を有することである。会社は、ディナースペシャルの2つの控えめな広告を検討している場合がある。目標は、近くを通る人々をまずディナースペシャルの広告に注意を払わせ、続いてレストラン自体に注意を払わせることである。2つの異なる広告（すなわち、デジタル画像に広告を挿入し、よって、それらが看板上有るように見える）を評価するために上記の方法を使用して、シーン内に配置し、次いで I E 変化を適用し、レストランの前に、どのくらいの頻度で看板が視覚的注意を集めるかを計算することができる。表3及び4は、それらが全く注意を集めていないかどうか（オブジェクト - # N o t A t t . ）と共に、オブジェクト A 及びオブジェクト B の相対順序を示す、可能性のある尤度を提供する。これらの表から、オブジェクト B の前にオブジェクト A が、広告コンテンツ1と共に時間の65%、しかし広告コンテンツ2と共に40%のみ発生したということが分かる。したがって、このシーケンスが重要である場合、顧客は広告コンテンツ - 1を選択する傾向にあり得る。

30

【 0 0 7 7 】

【表3】

表3 広告コンテンツー1

	オブジェクトーA P o s 1	オブジェクトーB P o s . 2	オブジェクトーA N o t A t t .
オブジェクトーB P o s . 1		2 5 %	7 %
2. オブジェクトーB P o s 2.	6 5 %		
オブジェクトーB N o t A t t .	2 %		1 %

40

【 0 0 7 8 】

【表 4】

表 4 広告コンテンツ-2

	オブジェクト-A P o s 1	オブジェクト-B P o s . 2	オブジェクト-A N o t A t t .
オブジェクト-B P o s . 1		2 0 %	2 0 %
2. オブジェクト-B P o s 2.	4 0 %		
オブジェクト-B N o t A t t .	5 %		1 5 %

10

【 0 0 7 9 】

堅牢性及びマルチサンプル状況

本記載はこれまで、単一の場所からオブジェクトを観察するときに焦点を当ててきた。しかしながら、世界は高度にダイナミックであり、観察者は空間を通過して移動することが多い。観察者が空間を通過して移動するとき、観察者は、特定のオブジェクトを「処理」するための複数の機会を有する場合がある。しかしながら、この動作は、オブジェクトが見られる時間及び凝視の回数であり、変化し得る堅牢性解析に対する他の重要な態様を生成する。本発明者らは、これを「視認性持続時間」と呼ぶ。すなわち、どのくらいの長さで特定のシーン又はオブジェクトが、観察者によって見られるかということである。視認性持続時間を捕捉するための1つの方法は、視覚的注意の場所の予測されたシーケンスをリポートするモデルを使用することによる。これらのモデルを有して、視認性持続時間は、シーンに対する堅牢性評価のためにモデルが考慮する凝視の数を限定することによって特定され得る（これはまた、以下に記載されるシーン最適化にも適用される）。より長い視認性持続時間は、より多くの凝視に対応し、同時に、より短い視認性持続時間は、より少ない凝視に対応する。視認性持続時間における変化は、堅牢性解析に有意な影響を与える場合がある。

20

【 0 0 8 0 】

シーンを処理し得る凝視の数が限定されているときに、オブジェクトが注意を集めるか否かは、それが視覚的注意シーケンス内にあるかどうかによって決まる。例えば、顧客はホテル内にデジタル表示を配置してもよい。デジタル表示は次から次へと2つのコンテンツを表示している。1つのコンテンツは3秒間、2番目のものは6秒間表示される。その限定された視認性持続時間を前提として、3秒のコンテンツは、6秒のコンテンツよりも注意シーケンスにおいてより早く現われる必要がある。例えば、人々が毎秒2回の凝視をする場合、モデルは、最初の6回の凝視で、注意がオブジェクトに引き付けられるということモデル予測しなければならない。一方、6秒のコンテンツは、最初の12回の凝視で視覚的注意を集めなければならない。これらのダイナミクスを前提として、それらと同様に他のものは、視認性持続時間を考慮しないと、人々がシーン内に注意を払うオブジェクトに関して不正確な予測につながる恐れがある。

30

【 0 0 8 1 】

これを示すために、本発明者らは、看板の例を拡大する（しかし、この同じコンセプトはダイナミックに変化する任意のディスプレイに適用する）。道路に沿って複数の表示がある長い道路を検討する。更に、この道路は平坦であり、看板以外の他のオブジェクトはない。この条件下で、看板の全ては同じ視認性持続時間を有する。すなわち、視覚的システムがいずれか特定の看板を処理する時間は同じであり、車の速度、表示のサイズ、及びシーン内のいずれかの大気の収差によって決定される。

40

【 0 0 8 2 】

ここで、道路上の最後の看板の1つの前側に500フィート（152.4メートル）の木列がある、同じ道路を検討する。他の看板の全ては同じ視認性持続時間を有するが、この最終看板は、より短い視認性持続時間を有する。視認性持続時間におけるこの変化を

50

前提として、堅牢性評価のためにモデルが考慮する凝視の回数を限定したい。この第1のセットの看板は、多数の凝視を有し、その一方で、最終看板の持続時間を考慮するとき、画像内に動画又はより少ない凝視から、より少ないフレームを考慮することができる。

【0083】

視認性持続時間のこの態様は、シーン内に多数のオブジェクトがある状況を考慮するとき、重要な役割を果たし得る。一部の状況下では、オブジェクトが処理されるのに十分な顕著性（又は注意を捕捉）であるが、しかし、シーン内の他のオブジェクトの妨げになるほど顕著性ではないように、注意捕捉要素を分散させたい場合がある。（この概念は、シーン最適化にも関連し、これは特定の目標を達成する方法でシーンを修正することを考慮する。シーン最適化は、以下で更に記載される。）これを仮定すると、オブジェクト、オブジェクト又はシーンの集合体の堅牢性を解析するとき、視認性持続時間を考慮したいであろう。例えば、パス上で短い視認性持続時間を有するオブジェクトに対して、顕著性要素（例えば、動き、輝度、色、コントラスト等）を増加させて、特定のオブジェクトが、その短い視認性持続時間中に処理される尤度を増加させる。しかしながら、その同じパスに沿って見える他のオブジェクトは、より長い視認性持続時間を有し、そのオブジェクトを処理する機会がより多くあるという事実を利用することができ、したがって、そのオブジェクトに顕著性要素をより少なく、又はより弱い顕著性要素を割り当ててもよい。

10

【0084】

視認性持続時間は、オブジェクトがパスに沿ったどこかで視覚的注意を集めたかどうかを考慮することによって、堅牢性解析に組み入れられる。このように、他よりも長い視認性持続時間を有する特定のオブジェクトは、それらに関連する、より多くの注視を有し、注意が、そのシーケンスに沿ったどこかでオブジェクトに割り当てられた場合、それは「ヒット」と見なされる。したがって、より長い視認性持続時間を有するオブジェクトは、より多くのサンプルを有し、処理されることのより高い尤度を有し、したがって、処理されるためにはより低いレベルの顕著性を必要とする場合がある。より短い視認性持続時間を有するオブジェクトは、より少ないサンプルを有し、したがって、注意が払われる可能性はより少なくなり、したがって、その短いシーケンス中に、検出されるべき、より上位な顕著性を必要とする場合がある。

20

【0085】

シーン最適化

30

ここまでは、本開示は堅牢性に焦点を当ててきた。ここで、本発明者らは、一部の実施形態で、前述の堅牢性関連方法及びシステムから利益を享受することができるが、堅牢性の評価を必ずしも必要としない他の視覚的注意モデルに戻る。1つのそのような概念は、シーン最適化のものである。

【0086】

先に記載したように、被検体が、シーン内でどこに、その視覚的注意を割り当てるかを予測することができる視覚的注意モデルが存在する。しかしながら、そのようなモデルは、特定の視覚的目標を得るために、どのようにシーンが修正され得るかということを識別するためのいずれのメカニズムも提供しない。なぜならば、ヒトの視覚的システムは、シーンの見えるエリア全体を実際には処理せず、代わりに、注意が引き付けられたこれらの領域のみを処理するからであり、多くの実際の状況では、人々に、シーン内の特定のオブジェクトを「見える」ようにするだけでなく、彼らに、特定のオブジェクトに「注意を払わせる」ようにすることが望ましい。

40

【0087】

視覚的目標は、次いで、被験者がシーン内のオブジェクトに注意を払う方法を指す。例えば、視覚的目標は、単に特定のオブジェクトが注意を払われるということを望むものであってもよい（すなわち、重要ではない、又は有害とさえ見なされるオブジェクトの集合体に加えて、ある人が視覚的注意の観点から重要であると判定するシーン内のオブジェクトの集合体）。あるいは、それは、特定のシーケンス、又は特定の時間に注意が払われる特定のオブジェクトを有することを望むものであってもよく、又はそれは特定の視点から

50

注意が払われるが、しかし必ずしも他からは注意が払われない特定のオブジェクトを有することを望むものであり得る。このセクションの残りは、視覚的注意の計算モデルを利用して、視覚的目標を達成するためにシーンを最適化する方法に関して記載する。記載される方法の一部は、先に記載したように、堅牢性の評価 (assessment or robustness) を利用するが、利用しないものもある。

【 0 0 8 8 】

図 10 を参照して、いったん視覚的目標が定義されると (ステップ 1 2 0 1)、明確なコスト / 報酬構造をシーン内のオブジェクトに割り当てることによってシーン最適化が開始する (ステップ 1 2 0 5)。そのような割り当ては、定量的に視覚的目標を定義する。コスト / 報酬構造は、シーン内のどのオブジェクトが高い値のオブジェクト、低い値のオブジェクト、及び更には、気をそらす又は視覚的目標に弊害をもたらすものであるとしてユーザーが見るオブジェクトであるかということ定義する。ユーザーは、視覚的目標 (有益な報酬) 又は有害である (負のコスト) の一部であると見なすオブジェクトのそれぞれに「注意有用性」を配置する。又は、ユーザーは、どの要素が、他よりも視覚的目標に「より」価値があるかを特定しながら重点を置くことができる。

【 0 0 8 9 】

次いで、多くの、シーンに対して可能性のある変更が定義される (ステップ 1 2 1 0)。これは、例えば照明、オブジェクトの色、オブジェクトの位置決め等の単純な変更、又は、例えばロビー内のどこに受付が建てられるべきかというような、より複雑なデザインの変更であり得る。勿論、理想的に、受付の配置など基本的な何かを評価するときは、ホテルのロビーを建てる前にされるのが一番であり、よって、本明細書に記載されるシーン最適化方法の 1 つの利用は、そのようなシーン内の、合成又は部分的に合成のシーン / デザインレイアウトの選択であると予測される。

【 0 0 9 0 】

実際の現実世界の「コスト」(すなわち見積)は、次いで各可能性のあるシーン変更に関連する (ステップ 1 2 1 5)。例えば、常連が、特定の表示に注意を払うことが目標であるホテルロビーの場合、表示の色を変更することは、比較的安く (\$ 2 0 0 の見積もられた変化値として割り当てられる)、その一方で花崗岩の床の色を変えることは高価である (\$ 2 0 , 0 0 0 見積もられた変化値として割り当てられる)。現実世界のコスト見積りを割り当てることで、複数の変化に伴う価格数字を関連付けることが可能である。例えば、シーンに対して可能性のある変化の範囲を考慮することができ、そのうちの一部は全ての目標に達し、その内の一部は大半の目標に達し、そのうちの一部は目標に達し、かつ最も安く達成し、定義された目標の 9 0 % を達成し、目標の残りの 1 0 % を達成するために、不均衡な投資資本をとるということを見出すことができる。換言すれば、現実世界のコストを一部の実施形態で可能性のある変化と関連付けることは、オプションの、より有意な評価を可能にする。結局、本方法は、注意の有用性を最大化しながら、同時に、オブジェクトの特徴構成と関連付けられたコストを最小化するシーン構成を提供する (ステップ 1 2 2 0)。

【 0 0 9 1 】

現実世界のコストは単に、特定の変更の相対コストがどのように共相関され得るか、すなわち他のシステム、例えば特定の変更に関してより高いコストと共相関する高い点を有するポイントシステム、及びより安価な変更である、より低いポイントはちょうど容易に使用され得るということに注意されたい。

【 0 0 9 2 】

別個のオブジェクトを追跡し、操作することができるグラフィックエディタは、シーンに対して可能性のある変化を定義するのに有用であり得る。例えば、そのようなグラフィックエディタ内で見られるシーンにおける表示は、識別 (例えば、マウスで右クリック) されることができ、ユーザーはそのオブジェクトの変更可能なプロパティを選択することができる。これらの変更可能なプロパティは、色、照明、レイヤー内の位置決め (例えば、オブジェクトは一部の他のオブジェクトの前又は背後に置かれてもよい)、コントラスト

ト、影、サイズ等を挙げることができる。変更され得る個々のプロパティを選択することに加えて、ユーザーは、許可された範囲の変更又は他の関連パラメータを定義する能力を有してもよい。例えば、色に関しては、シーン内にある特定の壁に対して、クライアント又はユーザーが許容可能であると見なし得る色は茶褐色の陰影であってもよい。したがって、色の属性は、特定のスペクトル範囲内で変化するのみであると定義される。同様に、属性がサイズである場合、オブジェクトのサイズがクライアント又はユーザーにとって許容可能でない前に、特定のオブジェクトが拡大され得る（又は、縮小され得る）サイズに対する明らかな制限がある。先に記載したとおり、実際の現実世界の「コスト」は、各可能性のある変更に関連する。可能性のある変更の範囲があるところでは、ユーザーは、その範囲内の例示のポイントのコストを定義することができ、支持するコンピュータシステム（後述する）は、例示のポイントに対してベストフィットカーブを外挿する。ユーザーは、多くの可能性のあるベストフィットアルゴリズムと共に表示され得、したがって、どれを使うべきか選択するよう依頼される。例えば、ユーザーは、表示のサイズに関して、最小の許容可能な表示コストは \$ 100 であり、中間表示のコストは、2つのコストのポイント間で線状に上昇する（サイズに基づいて）と定義したいだけである場合がある。

【0093】

定義された目標（又は複数の目標）、目標のための注意の有用性、可能性のあるシーンの変更、及び定義された可能性のあるシーン変更のコストと共に、次のステップはシーンに対する可能性のある変更の効果を評価し、ステップ1201で定義される視覚性目標を達成するシーン構成を見つけようとするものである。これは、一実施形態では、予測された報酬を最大化するシーンのプロパティ、又はコストを最小限にする一方で、報酬を最大にするシーンの構成、又は最小コストに対して定義された目標を単に達成するシーン構成をアルゴリズムで修正することによって実施される。この期待された報酬は、特定のオブジェクトが注意を集め、及びこれらのオブジェクトに注意が払われるための報酬/コストを受容する尤度として算出される。シーン内の場所に注意を払うための報酬/コストに加えて、最適化ルーチンもまた、シーンに特定のタイプの変更をする（例えば、シーン内の色及び/又は位置を変更すること）ためのコスト/報酬における要因である。注目を集めるオブジェクトの尤度を算出する1つの方法は、前のセクションに記載された堅牢性の解析を使用して測定され得る。

【0094】

等式1は、視覚的注意の計算モデルを使用して、ユーティリティ関数を公式化する1つの例を提供する。

【0095】

【数1】

$$ER(F) = \sum_{a_F \in A_F} [a_F(xy)R_a(O_f(xy))] + \sum_{o \in O} R(o_f)$$

等式1

【0096】

Fはシーン内のオブジェクトに関する変更可能な属性のセットである。例えば、それは、シーン内のオブジェクトのそれぞれに対する色、テクスチャー、又は位置を含む。Aは、モデルが所与のこの特徴構成Fを予測する注視のセットである。R(Oxy)は、特徴セットfと共に位置(xy)においてオブジェクトに注意を払うための報酬（正及び負の両方）である。a(xy)は、モデルが場所xyに割り当てられる注意を予測する尤度である。視覚的注意の一部のモデル及びアプローチに対して、a(xy)はバイナリ値（注意が場所に対して割り当てられる、又は割り当てられないかどうかを示す0又は1）であってもよく、一方で他のモデル及びアプローチに対して、これはこの位置(0...1)に割り当てられる尤度であり得る。どちらにしても、等式のこの部分は、現在のオブジェ

クト特徴セットに対する注意の割り当てに対して報酬を特定する。

【0097】

機能の他の部分は、オブジェクト o に対して特徴セット f を使用することに対するコストを特定する。 $R(O_f)$ はオブジェクト o において特徴セット f を使用することに対する報酬/コストを特定する。場合によっては、特定の特徴に対するコストは0.0であってもよい（例えば、これはオブジェクト o に対する現在の特徴セットに対するコストであってもよい）。しかしながら、モデルが、同じように難しい全ての可能性のある特徴の組み合わせを考慮することを望む場合がある。この場合では、全ての特徴に対する報酬は同等である（又は最も容易に0.0）。しかしながら、一部の状況では（例えば、シーン内のカーペットの色を変化させることに対して、花瓶を動かすこと）、許容可能な特徴構成における制約がある。そのような状況では、オブジェクト（ o ）及び特徴（ f ）上のコスト/報酬に関してこれらの値を特定することができるか、又はそれらは単に許容可能な変化として非許容可能な特徴構成を定義することを単に控えることができる。

10

【0098】

この報酬関数を使用して、解空間は、述べられた報酬関数を最適化させる特徴構成に対して展開される。いったん報酬関数が特定されると、最適の解を達成するための多くの方法がある。これらの方法には、閉形式の等式、Monte Carloシミュレーション、Simulated Annealing、Genetic Algorithms、及びStochastic Gradient Descentが挙げられるがこれらに限定されない。これらの近似値アプローチに加えて、一部の視覚的注意モデルに対しては、閉形式の解析を実施することができる。

20

【0099】

目標及び関連の付随するコスト情報を達成する解空間からの解は、次いで評価されるために利用できる。

【0100】

多くの例はこれまで、そのシーン内のオブジェクトの視覚的顕著性を増加させることによってシーンを最適化することに焦点を合わせていることに注意されたい。しかしながら、一部の視覚的目標は、他のオブジェクトからの視覚的注意（又は乱れ）を低減することによって達成され得るということに注意されたい。どのようにコストモデルが設定されるかということによって決定される、本明細書に記載されるシーン最適化方法は、シーンの様子をミュートすることになり得る（常にオブジェクトを、より視覚的に顕著性があるようにしていない）。

30

【0101】

シーン最適化及び注意シーケンス

上記のとおり、一部の目標はオブジェクトに注意が払われる順序を考慮しない場合がある。そのような条件下では、オブジェクトに注意を払うための報酬/コストは、シーケンス内のその位置及び/又は既存の注視の前若しくは後に注意が払われたオブジェクトによって影響を受けない。しかしながら、シーン内のオブジェクトの注視順序は重要であり得る状況が存在する。例えば、被検体が、指示の特定のシーケンスに従うとき、順序が問題になる傾向がある。

40

【0102】

この、より複雑な視覚的目標に対処するために、上記（等式1）の期待された報酬関数は、期待された報酬が順序依存であるように、特徴構成を最適化しよう拡大されてもよい。これは、サッカーの数の関数として、可変報酬構造を利用することによって実施され得る。等式1では、オブジェクトに注意を払うことに対する報酬は、ここの注意サッカー（ R_a ）によってインデックスが付けられる。サッカーのシーケンシャル位置（ a ）によって異なる報酬を特定することによって、それによってシーンが予測されたサッカー順序によって最適化される方法を作ることができる。図11は、2つの報酬構造を示すグラフである。1つの報酬構造は、サッカー位置（列1225）と恒常的に示され、2番目はサッカー位置（列1230）に依存する。サッカー依存の報酬構造は

50

、この特定のオブジェクトに注意を払うことに対する期待された報酬は、それが早く発生した場合には非常に高いが、シーケンスにおいて、注意が後で割り当てられるときは減少するということを特定する。このタイプの報酬構造は、たとえば構造体区域のシーンにおける歩行者など、「高価値」のオブジェクトに関連されてもよい。

【0103】

注意シーケンスにおいてどれくらい早くオブジェクトが注意を払われるかという報酬に基点を置くこと（図11が例示するように）よりはむしろ、シーケンスを基準とした目標における報酬を基点とし、オブジェクトの報酬は、その前及びその後に注意を集めるオブジェクトに基づく。例えば、ホテルのオーナーは、2つの屋外の表示を有している場合がある。1つは、彼らのレストランでスペシャルを広告し、2番目は彼らのホテルの名前及びロゴを表示する。特別な表示広告は、ホテルの表示の前に見られるべきであることが重要であるとホテルのオーナーは確信する。この視覚的目標を前提として、解析は、ホテル名の前に「レストランスペシャル」の表示がいつ注意を払われるかということに非常に高い報酬を設定する。更に、低い報酬は、ホテル名が「スペシャル広告」表示の前に注意が払われたときに与えられ、もしどちらか1つが他のものがない状態で注意が払われた場合は、報酬が与えられない場合がある。シーケンスは相対シーケンス（前対後）であってもよく、又は絶対シーケンス（視覚的注意を集める第1オブジェクトとしてオブジェクトAが発生し、オブジェクトBは視覚的注意を集める第2のオブジェクトであるとき、オブジェクトAが発生するまでオブジェクトA及びBに注意を払うことに対して報酬はない）であってもよい。勿論、報酬構造が当業者によって理解されるように、定式化され得る多くの他の方法がある。

【0104】

これまで、このセクションは、サッカードの位置又はサッカードの順序が報酬に影響を与える2つの可能性のある報酬関数を記載している。当業者は、いずれかの任意のシーケンシャルな報酬関数を、オブジェクトのセット及びサッカードのセットにわたって定義してもよいということを理解するであろう。より一般的に、クエスチョンにおけるサッカードシーケンスの長さである、M次元の空間（各オブジェクトに対して一次元）を定義することによって、報酬構造の、このシーケンシャルな構成要素を定義することができる。

【0105】

この最適化ルーチンが設定されて、所与の報酬構造へのシーンに対する最良の構成を提供し得るが、得られるシーンは「推薦」であって、適切ではないということが見出すことができる。例えば、それをより顕著に、又はより顕著にしないようにするために、1つのオブジェクトの色を修正しようとした場合、どのように注意シーケンスが変化するか疑問に思う場合がある。1つのオブジェクトをより顕著にすることは、非自明な、非線形の効果予想されるシーケンスに有する場合がある。これは、注意の割り当てが、ゼロサムゲームの定義によるという事実起因しているためであり、1つのオブジェクトに注意を割り当てることは、必ずしも注意が他のオブジェクトから割り当てられるということであることを必ずしも意味しない。したがって、1つのオブジェクトの特徴を、その顕著性を増加させるために修正することは、このオブジェクトが注意を集める（かつそれが注意集める注意シーケンスにおいて）尤度を変化させるだけでなく、他のオブジェクトが注意を集めるという尤度及びこれらのオブジェクトが注意を集める注意シーケンスにおいても影響を及ぼす。

【0106】

最適化ルーチンへの1つのアプローチは、複数の特徴構成を自動的に考慮するため、解空間を定義する多くの（又は全ての）可能性のある構成は展開されているであろう。非最適な解は、ユーザー又はクライアントの大きな関心であり得る。例えば、どの色が表示を、その位置、例えばサッカードシーケンスの25番目の位置からトップ5に移動させるかを知りたいと思う場合がある。システムは保存された注意シーケンスを検索し、そこでは、他のオブジェクトの全ての特徴が一定に保たれており、感心のオブジェクトはトップ5のサッカードにある。

【0107】

堅牢性は、上記のとおりシーン最適化でも使用され得る。例えば、その画像のプロパティに対してのみではなく、一日の又は異なる視聴者に対してシーンが見える場合があるシーンに対して、シーンを最適化することができる。すなわち、最適化ルーチンは、シーンが経験し得る変動に対して堅牢であるシーンを推奨する。これまで、本発明者らは、入力されたシーン及びモデルに対する内部及び外部変動を捕捉し、生成する方法を記載した。これらのアプローチの目的は、実際のシーンの予測変動をシミュレーション（又は捕捉）することであった。例えば、観察者の変動を捕捉する1つの方法は、視覚的注意モデルのパラメータを変化させることであり（例えば、顕著性を計算するための個々の機能マップの重み付け）及びこれらのモデルを同じイメージ上に走らせることである。異なるモデルパラメータ（内部変動）の多数のインスタンスによって、各シーン構成を走らせることは、特定のシーン構成に対して複数のスコア（各モデルに対して1つの注意シーケンス）を与える。各モデルに対する平均スコアをとることによって、所与の変動を有するシーン構成に対する期待されたスコアを生成することができる。平均において最良のスコアを提供するシーン構成を推奨する場合がある。

10

【0108】

また、堅牢性に関して視覚的目標を定義し、特定の堅牢性の値にシーンのオブジェクトを最適化する場合がある。

【0109】

視認性持続時間

20

堅牢性の記載に関して言及したように、異なるオブジェクト又は異なるシーンは異なる視認性持続時間を有する場合がある。備忘として、視認性持続時間は、オブジェクト及び/又はシーンが見られる時間を指す。時間に関して特定され得る視認性持続時間は、典型的には、最適化ルーチンにおいて考慮される、多くの予測された注視（又は顕著性領域）に変換される。視認性持続時間は、様々なシーンの解析に使用される凝視のセットを限定するのに使用される。より具体的には、それは視覚的目標解析において使用される。

【0110】

視認性持続時間は、多数の視覚的目標を考慮するときに効果を有することができる。例として、異なる報酬（例えば報酬値は1、2、3、4、5、6）を備える、6つの視覚的目標を有するドキュメント又はコンテンツを有する場合がある。コンテンツがデジタル表示に5秒間表示された場合、人々は毎秒約2つの凝視をして、これは10の凝視の視認性持続時間に変換する。視認性持続時間が10の凝視とすると、視認性持続時間は、視覚的目標の全てを捕捉するのに十分長い。この条件下で、モデルは比較的等しい6つの品目の顕著性を作る（シーン内には他のオブジェクトはなく、変化を作るためには等しいコストであると仮定する）。視認性持続時間と共に最適化ルーチンがおおよそ等しい視覚的目標を作る理由は、モデルが、この状況下でターゲットの全てに注意を払う傾向があるということである。標的（例えば、最も高い報酬を備えるオブジェクト）のうちの1つの顕著性が、他のオブジェクトのうちの1つの顕著性よりも著しく高い場合、大半の視覚的注意モデルは、最も顕著なオブジェクトに、次いで2番目に最も顕著なものへ注意を払うが、それは典型的には最も顕著性のあるオブジェクトに再び戻る。前のオブジェクトに戻ることに
40
関して追加的な報酬がない場合、この注視は全体的な報酬を増加させない。しかしながら、オブジェクトを顕著性において比較的等しくすることにより、モデルは、標的オブジェクトの全てに注意を払う傾向があり、したがって、視認性持続時間を与えられた、より多くの視覚的目標（全部ではないが）を達成する。最適化ルーチンのこの態様及び視認性持続時間は、標的オブジェクトの報酬又は関連性を相関付けられたオブジェクトの顕著性を単に作ることは非常に異なる。

【0111】

図12は、シーン最適化システム1255の上位レベルのダイアグラムである。この図において、類似の名前の構成要素は、先に記したこれらのモデルの機能性及び能力に類似である。シーン最適化システム1255は、データベースグラフィカル・ユーザー・イン

50

ターフェース 1 2 4 5 を含み、これは、ネットワーク 5 0 2 にわたってユーザー 4 0 1 からの、コンピュータ 5 0 3 を介した入力を受信し、シーン及びシーンに対して可能性のある変更のための 1 つ以上の視覚的目標を定義する。シーン内のオブジェクトに対するこれらの可能性のある変化はデータベース 1 2 3 5 に保存される。シーン最適化モジュール 1 2 6 0 は、視覚的目標を達成する、可能性のあるシーンの分野を繰り返し適用して検索し、シーン修正モジュール 1 2 4 0 を呼び出して、ユーザー 4 0 1 によって定義された可能性のある変化と一致する方法で入力されたシーンを修正する。これは修正されたシーンを作り出し、修正されたシーンは視覚的注意モジュール 4 0 3 に提供され、視覚的注意モジュール 4 0 3 は視覚的注意に関する出力を提供し、視覚的注意はデータベース 1 2 3 5 に保存される。シーン最適化システム 1 2 5 5 は、ユーザー 4 0 1 が遠隔である場合（かつ図 1 2 に示されているように）ウェブサーバーモジュール 1 2 5 0 を含んでもよい。

10

【 0 1 1 2 】

複数の視点のシーン解析

この点まで、本記載は、単一の静止画像又はムービーをを利用して、画像内のどこに注意が割り当てられるかということをも予測する、視覚的注意モデリングに主に焦点を当ててきた。これらの 2 つのアプローチはしばしば多くの状況に対して有用であるが、それらは、実際には複雑な 3 次元であるものの単一の 2 次元ビューで動作するということに制限される。

【 0 1 1 3 】

我々が生きる 3 D の世界では、全体像における小さな変化は、観察者の網膜上に投影されている画像上に著しい変化を有する場合がある。1 つの全体像から見えるオブジェクトは、他においては全く見えない場合がある。更に、オブジェクト間の空間的な関係（すなわち、網膜上の、第 2 オブジェクトに対する 1 つのオブジェクトの投影された画像の位置）は、1 つの全体像から他のものへと著しく変化し得る。全体像における変化は、網膜に投影される画像において大きな変動を生成することができ、それらはまた、ヒトの視覚的注意がシーン内のどこに割り当てられるかということに有意な著しい影響を有する。例えば、図 1 3 A は、デジタル表示 1 2 9 0 をデスクの背後に有する受付デスクを含む、ホテルロビーのシーンのアーティストのレンダリングである。シーンは、図 1 3 A において最も注意を引き付ける 2 つのオブジェクトが、デジタル表示 1 2 9 0 及び広告カード 1 2 8 5 であると予測する、視覚的注意モデルによって解析されている。図 1 3 B は、同じホテルロビー内の同じ受付エリアであるが、デジタル表示 1 2 9 0 は、視覚的注意モデルによって特定されたいくつかのオブジェクトの中にはない。視覚的目標が、常連にデジタル表示 1 2 9 0 に注意を払わせることであり、図 1 3 A に示されるシーンのみが解析される場合、視覚的目標が一貫して達成されているという、間違った安心感が存在するであろう。しがたって、3 D シーン内のシーンコンテンツを測定し、評価する効果的な方法を有する必要がある。

20

30

【 0 1 1 4 】

3 D シーン評価（本明細書では、複数の視点のシーン解析と呼ぶ）は、ムービーシーケンスなどのものに対する 2 D の視覚的注意の単なる拡大と同じではないことに注意されたい。そのようなアプローチは多くの全体像からのデータを提供するが、最終的には、それが空間を通る特定の動作に対する注意の割り当てを解析する能力を有するものを提供することができるという点で、単一の画像アプローチのものと似た限定を有するが、空間を通る他のパス及び / 又は動きを考慮する能力を提供することはできない。静止シーンの視覚的注意又は連続する複数の静止シーン（ビデオ）の視覚的注意の解析のいずれも、画像が 3 D 設定から生じているという事実に対処していない。

40

【 0 1 1 5 】

よって、観察者が、3 次元である地理的空間横断することができる 3 次元及び無数の方法を提供する複数の視点のモデリングプロセスを有することは有用である。

【 0 1 1 6 】

図 1 5 は、複数の視点の視覚的注意モデリングプロセスを示すフォローチャートである

50

。このプロセスはサイト計画（ステップ1329）で始まり、これは3Dシーン内のどの場所及びオブジェクトを解析したいかを決定することから構成される。実際には、これは、解析される3Dシーンの平面図を取得する、又は開発すること、並びにユーザーの観察視点の代表するであろう平面図における場所を決定することを意味する。勿論、3Dシーンの平面図がないと、厳密性の少ないアプローチが使用される可能性があり、そこではユーザーは単にサイトに行き、どの場所が関心のものであるかという決定を行い、これらの場所から写真を撮る。更に、ユーザーは、画像が撮られた位置及び向き記録する場合があり、これは報告目的には有用であり得る。

【0117】

写真よりはむしろ、フレームがビデオ又はビデオシーケンスから捕捉されるビデオ技術もまた使用することができる。フレーム（ビデオからの画像）又は画像（例えば（or example）デジタルカメラからの）が使用されるとき、ユーザーは、2つの異なる画像間にあるビューを生成するためのビュー補間技法を使用することもできる。先述のように、画像は実際の環境から生成されるものに限定されず、それらは合成（バーチャル）環境からも生成され得る。しかしながら、2つの場合では、画像が撮られる環境における場所、及びカメラの視点（どのような方法でカメラが向けられたか）を記録若しくは事前に特定しなければならない。場所を特定するための単純なアプローチは、次いで、北、東、南、西軸の周囲を90度を進展する（以下の例で実施されるように）全体像からそれぞれ連続する画像を有することである。しかし、場所の事前特定及びカメラ照準プロトコルがない場合は、カメラは、代わりに、光学追跡技術と組み合わせで、GPS追跡技術を使用して追跡することができる。例えば、画像が撮られる度に、機器はカメラに取り付けられ、3つの位置寸法（X、Y、Z）、及び3つの方向寸法（ヨー、ピッチ及びロール）を記録し、画像が捕捉される場所からの視点の明確な表示を提供する。これらの6つの値はメモリに保存され、その時に捕捉された画像に関連付けられる。勿論、カメラ位置及び向きは手動で記録されてもよい。

【0118】

測定された3Dシーンと共に、次に、3D環境からの次の多数の画像が受信され、多数の画像は、観察者が3D環境を通してやりとりし、ナビゲートするとき、観察者が経験し得るビューのセットを表示する（ステップ1330）。一実施形態では、これは、3Dシーン内（再び、ホテルのロビーなど）の複数の位置及び向きから撮られた複数の写真によって達成される。他の実施形態では、観察者がいるであろうと予期され得る多数の代表的なエリアから撮られたショット共に、ビデオは3Dシーンで作成される。更に他の実施形態では、バーチャル3Dモデルが使用されて、ビューは、バーチャル3Dモデルを通してバーチャルカメラを移動させることによって生成される。どれくらい生成されたとしても、結果は、3Dシーン内の様々な位置からの複数の2Dイメージとなる。事前に指定されていない場合、3D環境内の場所及びカメラの向きを示すデータも収集される。このデータは、観察者が3D空間を通じて、とる可能性がある多くの異なるパスの評価と共に、多くの視点からシーンを評価できるようにする。

【0119】

いったん画像が収集されると、対象のオブジェクトに対応する2次元の画像の領域が選択され、特定される（ステップ1335）。これは、これらのオブジェクトを自動で抽出する方法、領域の手動選択、又は自動及び手動のタグ付け及びラベル付け技法の両方を使用する混合の技法を含む（しかしこれらに限定されない）多くの異なる方法を使用して達成することができる。このプロセスを達成するための1つの方法（手動の）のグラフィック結果の表示が、図14A及び14Bに示されている。壁画1310、デジタル表示1315、及び花瓶1320などのオブジェクトは、マニュアル選択プロセスによって特定されている。

【0120】

次に、関心のオブジェクトを含む画像は、視覚的注意データを作るために視覚的注意モデルを使用して処理される（ステップ1340）。先に言及したように、1つのそのよう

10

20

30

40

50

なモデルは、Itti & Koch (2001) によって記載されているが、任意の視覚的注意モデルが使用されてもよい。観察者が各視点にいるときに、視覚的注意が割り当てられるであろうとモデルが予測する2次元の場所が、次いで例えばデータベースに記録される。モデルが注意を集めると予測する、見えるエリア内のオブジェクトを決定するために、これらの場所は次いで、各視点に対してタグ付けされ、ラベル付けされた領域と比較される。

【0121】

それぞれ個々の視点を解析し、それぞれ個々の視点に対して視覚的注意を集めると予測するオブジェクトの算定の後、どの場所からどのオブジェクトに注意が払われるかを検討するデータが生成され、保存される。このデータは次いで、例えば、特定のオブジェクトが潜在的な観察者によって、その者が3Dシーンを先回するときに見られる（全く見られない）尤度；実際に特定のオブジェクトを3Dシーン内で観察する潜在的な観察者のパーセント；特定のビューのサブセットに対して注意が払われるオブジェクトの尤度（ロビーに入ることにに対してロビーを出るなど、特定の移動パスに関する情報の解析に有用であり得る）、又はオブジェクトが見えるとき（一部のオブジェクトは、環境内の可能性のある視点の大半から目立たない必要がある得るが、オブジェクトが見える視点に関しては、それに対して注意が引き付けられる、高度な確実性を有したい場合がある）、又はオブジェクトが3Dシーン内で見えるとき（又はオブジェクトに注意が払われる）オブジェクトに注意が払われる尤度を決定するために使用することができる。

【0122】

この3D視覚的注意モデリングは、シーン最適化に関して上記のシステム及び方法と組み合わせることができる。3D視覚的注意モデリングに関して定義され得る視覚的目標は、しかしながら異なってもよい。例えば、90%の潜在的な観察者が、実際にある点で特定のデジタル表示を観察し、同時にホテルのロビーを横切るように、視覚的目標が定義されてもよい。

【0123】

シーン最適化と組み合わせられた3D視覚的注意モデリングを適用する例として、一部の視点から観察されたときはオブジェクトが目立たないままであり、オブジェクトが関連性のあるときは、他の視点から目立つようにしたいケースを検討する。例えば、その顧客に見て欲しく、かつ彼らの滞在中に注意を払って欲しい、3つの異なる形態の広告コンテンツを有するホテル。第1のコンテンツは、部屋のアップグレードに関するスペシャルを広告することであり、第2のコンテンツは、ルームサービスに関する広告であり、第3のコンテンツはホテルのダンスホールで行われる劇のチケットを広告することである。これらの異なる形態のコンテンツは、顧客の滞在中に異なる時間に関連する。部屋のアップグレードは、顧客がホテルにチェックインするときに関連があるが、他の時間では関連がない。ルームサービスは、顧客が彼らの部屋に行くときに関連があるが、顧客がホテルを去るときには必要ではない。劇のチケットは、反対に、ほぼどんな時でも顧客に関連する。シーン最適化技法と組み合わせられた3D視覚的注意モデリング（両方とも上で記載された）を使用して、以下の方法でこれらの広告材料の配置及びコンテンツを最適化することができる。第1に、情報が最も関連するホテルのロビーにおける場所を決定することができる。例えば、部屋のアップグレードは、顧客がホテルにチェックインするときに関連があり、これは通常、顧客がチェックインカウンタの前に立っているときに発生する。エレベータの隣のエリアは、一般的な広告（芝居のチケットなど）に最適であり得る。ちょうどホテルが、特定のタスクに対応する視点から特定の表示が目立つように望む場合があるように、ホテルはまた、関連がない表示には目立たないようにしたい（例えば、エレベータを待つときに、部屋のアップグレード）。チェックインカウンタにおいて立っているときに、その状態であろう位置及び向きのセットに基づいて、広告材料の位置及びコンテンツを解析することができる。

【0124】

3D視覚的注意モデリング及びシーン最適化方法の一実施形態を試験するために、本発

明者らは、彼らの試験 3 D シーンとして標準的な会議室で行い、このダイアグラムは図 1 6 に見ることができる。会議室はテーブル 1 3 4 5、黄色の表示 1 3 5 0、緑色のバスケット 1 3 6 0、電話 1 3 5 5、紫色の表示 1 3 6 5、並びに会議室に見出すことが予測される他の典型的なもの（椅子、ゴミ箱、スクリーン）を含む。観察者が部屋を見ることを期待され得るものからの代表的な点は、マニュアルで決定され、8 つの代表的な観察場所が生じた（観察場所 1 3 6 6、1 3 7 7、1 3 7 8、1 3 7 9、1 3 8 0、1 3 8 1、1 3 8 2、及び 1 3 8 3）。この試験に関して、観察場所は、部屋の邪魔されない（家具のない）領域全体で約 1 . 2 1 メートル（4 フィート）離間していた。デジタルカメラを使用して 4 つの画像が撮られ、8 つの観察場所のそれぞれから、3 2 の画像が生じた。観察場所から外側に延びる矢印は、デジタルカメラが各写真に対して向けられた一般的な方向を示し、各観察場所において、各写真に対して約 9 0 度の向きである。3 2 の異なる画像のそれぞれに対する場所及び向きが記録された。

10

【 0 1 2 5 】

本発明者らは次いで、3 2 の画像の少なくとも 1 つに見出された 1 2 の異なるオブジェクトに関連するピクセルを特定し、タグを付けた。これは、3 2 の画像のそれぞれ上で、関心のオブジェクトを包む 2 D 領域を画定する多角形領域をユーザーに選択させることによって実施された。図 1 7 A 及び 1 7 B は、3 2 の画像のうちの 2 つの、アーティストのレンダリングを示し、多角形は、関心のオブジェクト、例えば緑色のバスケット 1 3 6 0 及び黄色の表示 1 3 5 0 を囲む。「タグ付け」は単に、オブジェクト（例えば「黄色の表示」）を含む領域の名前を付けることという。オブジェクトを囲む多角形は、タグ付けソフトウェアの目的の代表であり、下線のある画像は、実際には多角形で修正されてはならず、識別された領域は、ユーザーによっていったん特定されると、最初の画像上でどんな方法でも重要ではない。

20

【 0 1 2 6 】

画像をタグ付けし、ラベル付けした後、画像は視覚的注意モデルに提出され、視覚的注意が割り当てられると、モデルが予測する画像において場所を収集する。画像の全てを送信した後、各画像に対してコンピュータは、視覚的注意が割り当てられると、モデルが予測した x、y 座標を記録する。コンピュータはまた、これらの注意のそれぞれに対して、ユーザーによってタグ付け及びラベル付けされた画像の領域内にそれが入るかどうか、算出した。コンピュータはまた、注意を集めると予測されなかった（「ミス」）タグ付き領域を含む各画像を記録した。全てのデータはデータベースに保存され、これは次いで、シーン内のオブジェクトの顕著性に関する一連のサマリーを生成するために使用された。

30

【 0 1 2 7 】

図 1 8 は、会議室内でタグ付けされ、ラベル付けされた 1 2 のオブジェクト 1 4 0 0 に実施された、3 つの異なる実施例解析を示すグラフ 1 3 9 5 を示す。第 1 の解析は、オブジェクトが可視である、即ち見える（ $p(\text{Visible})$ ）尤度である。これは、オブジェクトが、画像が、画像の全数によって割られた画像内にある画像の比率である。 $p(\text{可視})$ は次いで、関心のオブジェクトが、どれくらいうまく設定内に配置されるかという、いくつかのの表示を与える測定基準である。決定される測定基準は、それが可視である（ $p(\text{Attended} | \text{Visible})$ ）ということを仮定されて、オブジェクトに注意が払われた尤度であり、これは、オブジェクトが可視であり、その凝視を予測した視覚的注意モデルが、特定のオブジェクトを定義する領域内に発生するかどうかを特定する可能性である。計算された測定基準は、特定のオブジェクトが全てに注意が払われる尤度（ $p(\text{Attended})$ ）であり、これはモデルが、少なくとも 1 回のオブジェクトへの注意の割り当てを予測し、次いでその値を画像の全部の総数によってわっている。

40

【 0 1 2 8 】

複数の視点のシーン解析は、標的オブジェクトが多くの異なる距離から見ることもできるという事実を示す。例えば、上記の看板の例を検討する。長い平坦な道を運転しているとき、看板に関心を向け、したがって、これを処理する多くの機会が存在する。一方、他の看板に関して、極めて最後の瞬間まで表示を遮蔽する丘又は木の群がある場合がある。

50

これらの異なる視点を考慮し、オブジェクトが、それが見られ得る異なる視点から視覚的注意を集める尤度をより正確に解析できる。複数の視点なしで、かつ単一のビューのみを使用して、それが注意を払われる可能性のある多くの視点を仮定すると、オブジェクトが視覚的注意を集める、又は集めないということを誤って予測する場合がある。

【0129】

図18は、上記の解析から得られるデータで実施することができる、可能性のある評価のサブセットを示す。このデータから引き出せる多くの結論がある。第1に、紫色の表示オブジェクトは、スクリーンオブジェクトよりも見えないことが多いということに注意されたい(即ち、それは、環境内で撮られた少ない画像であった)。しかしながら、これらの2つのオブジェクトの黒いバーを見ると、紫色の表示が見えるとき(すなわち、画像内にそれが存在する)、それは常に注意が払われており($p(\text{Attended} | \text{Visible}) = 1.0$)、しかしスクリーンが見えるときは、注意が払われていない。紫色の表示は、極めて見られないことが多い定置に配置されているが、それがビュー内にあるときは、モデルはそれに注意が割り当てられているということを予測する。

【0130】

紫色の表示のオブジェクトは次いで、ホテルのロビーとの関連で、望ましいと先に記載された、プロパティのタイプを生成する。すなわち、多くの場所から見えない表示(それは目立たない)が、それが見える場所に人々がいるときに(チェックインカウンタの側)、彼らはそのオブジェクトに常に関心を向ける。これは、紫色の表示のオブジェクトが見える確立(白いバー)は約15%であるという事実によって示される。しかし、それが見えるときに、注意は紫色の表示によって捕捉される確率(黒いバー)は1.0である。

【0131】

図19は、複数の視点のシーン解析システム1415を含む上位レベルの機能性モジュールを示すブロック図であり、これは、様々な実施形態において、図15に関して記載された複数の視点の視覚的注意モデリングプロセスを実行することができる。それは、一実施形態では、コンピュータシステム408内に存在し、これは多くの他の機能性モジュール(例えばオペレーティングシステム)及びハードウェア、例えばメモリ又はプロセッサ(それらのいずれも図19には示されていない)を含む。単一のコンピュータとして図19に示されているが、実際には、機能性の様々な部分が、ネットワーク構成でいくつか又は多くのコンピュータ間に広がっている。複数の視点のシーン解析システム1415は、視覚的注意モジュール403(上記)を含む。それは、必要に応じて、関心の解析(図15のステップ1340に関して先に記載されている)をするために視覚的注意モジュール403を呼び出す複数の視点のシーン解析モジュール1425も含み、解析から結果(どの定義されたオブジェクトが、どの画像内の視覚的注意モジュールによって特定されたかを含む)を受理し、これらの結果又は結果の要約をデータベース1410に保存する。データベース1410は、コンピュータ炭層ファイル、コンピュータメモリ、又はデータベースなど、いずれかのデータ保存装置又はシステムである。複数の視点のシーン解析システム1415は、グラフィカル・ユーザー・インターフェースモジュール1420も含み、これは多数の画像(図15のステップ1330で得られた)の入力促進し、次いで、この実施形態では、画像内の関心のオブジェクトの識別及びタグ付け(図15におけるステップ1335)を促進する。

【0132】

ユーザー1405は、複数の視点のシーン解析システムと相互作用することに興味のあるいずれかの人間又は他のコンピュータシステムである。一実施形態では、ユーザー1405は、会社に雇用されたコンサルタントであって、会社に所有された又は管理された3Dシーンへの構成変化を解析及び推薦する。

【0133】

視聴者は十分な時間、同じ場所に残ることが多い。例えば、ある人が食料品店、ホームセンター、又はホテルから、支払いをするために列で待っている場合がある。この間、個人は「視覚的採餌」作業を行う場合がある。視覚的採餌とは、観察者が特に何かを探して

10

20

30

40

50

はいないが、興味のある何かを、環境周辺で単に見回していることである。この視覚的採餌中に、人はその目を動かすことによってことある情報に関心を向け、その人の目がそれらの回転軸の縁部に到達したとき、その人は頭を動かす。典型的に、その人は関心の品目が凝視の中心にあるように頭と目を動かす。現在の最先端技術は、画像上への、このタイプの再センタリング作用をシミュレーションしない。かわりに、画像が解析されるとき、画像の中心は常に凝視されたままである。凝視ポイントが、画像（又は視界）の縁部にあるときでさえ、これは本当である。再センタリングすることなく、現在の最先端技術のアプローチは、画像の縁部に固定できるのみであり、その点を超えては決して固定しない。一方、ヒトは、彼らの目が凝視のセンターに整列されるように、彼らの視界の縁部に関心を向け彼らの頭を回転する。これは視覚的システムが、同じ方向において他の凝視を作るのを可能にする。単一の画像と共に、画像の縁部から出る情報はもはやないという事実により、同じ方向で凝視を作ることができない。

10

【 0 1 3 4 】

本明細書に記載のこのシステム及び方法は、一部の実施形態では、シーンの多数のビュー又は単一のパノラマビューを使用して、視覚的採餌中に目の再センタリングをシミュレーションする。これは以下のとおり実施され得る。

【 0 1 3 5 】

1. 単一の視点から多数の画像を生成する。画像は、360度のパノラマカメラ、又は視点軸（又は別の方法で垂直）の周りで回転された多数の単一画像によって生成される。多数の画像が撮られてもよく、ここでビューは互いに「重なり合う」。ビューの向きはまた、各ビューに割り当てられる。

20

【 0 1 3 6 】

2. 最初のビュー（「開始ビュー」）は、視覚的注意モデルに与えられる。ビューは、誰かがそれらの視覚的採餌（例えばお店の列において、それはキャッシャーの方へ「前方」を見ている場合がある）を開始する典型的な視認包囲によって測定され得る。ランダムに選択された向きで開始してもよい。パノラマビューの場合では、パノラマビューの「スライス」が、「開始」ビューの中心に置かれて使用され得る。多数のビュー／画像の場合では、開始場所の中心に最も近い画像が使用される。

【 0 1 3 7 】

3. 「開始ビュー」は、視覚的注意モデルで解析される。初期の凝視は、モデルによって予測される。この凝視の方向が計算される（これは三角法を使用して実施され得る）。パノラマビューが使用される場合、パノラマビューの新しい「スライス」が、この新しい凝視の中心にされる。多数の画像が使用される場合、新しい凝視に最も近く中心にされる画像が使用される。

30

【 0 1 3 8 】

4. この新しいビューを備えて、システムは次の顕著性のある領域を解析する。

【 0 1 3 9 】

a. このプロセスは次いで、繰り返す（凝視を決定し、視点を中心に置く）。

【 実施例 】**【 0 1 4 0 】**

いくつかの非限定的な例が以下に提供され、これらはどのように前述のシステム及び方法が実用化され得るかということを示す。

40

【 0 1 4 1 】

実施例 1：外部変動を使用する、堅牢性の計算

背景：ホテルのオーナーは2つのデジタル表示をホテルのロビーに設置したいと思っている。彼女は、顧客がそれらに気付いて欲しく、3つの潜在的な場所のいずれかにそれらを置くことができ、2つのデジタル表示の、3つの可能性のある構成となった（すなわち、場所1～2、1～3、又は2～3における表示）。堅牢性は、気付かれるために最も良い表示の場所を推薦するために算出される。

【 0 1 4 2 】

50

1. デジタルカメラを使用してホテルのロビーの画像を捕捉し、この画像を、注意モデルを走らせることができるコンピュータにダウンロードする。コンピュータは、Matlab (商標) (The MathWorks (Natick, Massachusetts) から入手可能) と共に、インストールされた視覚的注意モデルソフトウェア (例えば Koch & Itti) を有する。

【0143】

2. 画像を修正してシミュレーションされたデザイン表示を含む。3つの修正された画像を生成し、それぞれは3つの潜在的な場所において2つの表示の全ての組み合わせが作られるように、それぞれ2つのデジタル表示をシミュレーションする。Photoshop (商標) (Adobe Co. (San Jose, CA) から入手可能) など、標準のデジタル写真操作プログラムを使用する。各シミュレーションされたデザイン表示は、正しく拡大縮小され、コンテンツ、例えばホテルのロゴのグラフィックなどコンテンツをシミュレーションした。ピクセルのアドレスによって定義されたように、コンピュータ上のファイル内の3つのデジタル表示の場所のそれぞれと関連して、画像領域を保存する。

【0144】

3. 注意モデルを介して修正された画像を走らせる。出力は修正された画像の予測された顕著性のある領域を含む。それぞれ顕著性のある領域は、ステップ2で保存されたデジタル表示ピクセルアドレスと比較される。顕著性のある領域が保存されたピクセルのアドレス内にある、又はそれと重なり合う場合、期待された注意は望ましい場所に行く。3つの修正された画像のそれぞれは、上位10の凝視状態にあるデジタル表示を示し、したがって3つの場所のいずれかがを良好な候補であるということを確認する。

【0145】

4. 一連の静止写真を使用して、又はビデオ及びビデオストリームからのサンプル画像のいずれかを使用して、同じシーンの複数の画像を捕捉する。画像は16時間にわたって毎5分とられ、したがって、シーンの外部変動を捕捉し、照明状態及び歩行者の動きの変動となる。目標は、これらのタイプの変動 (照明及び歩行者の動作) に、このタイプの変動に対して耐性のある表示場所を有する。コンピュータにこれらの画像をロードし、ステップ2で記載されたようにそれらをシミュレーションされたデジタル表示と共に修正する。

【0146】

5. ステップ4からのそれぞれ修正された画像は、注意モデルによって解析され、ステップ3に記載されたように、保存されたピクセルアドレスを比較される。表示場所1及び2に関連する修正された画像のシリーズは、画像の20%において、予測された凝視が両方のデジタル表示場所に行くということを示した。同様に、場所1及び3は、両方の表示場所に行く凝視の35%を有し、一方で、場所2は、両方の表示場所に行く凝視の85%を有する。場所2及び3において設置された表示を有することは、最も耐性のある構成となり、ホテルにとって最良の解を提供する。この解をホテルのオーナーに推薦した。

【0147】

実施例2：内部変動を使用する、堅牢性の計算

背景：ホテルのオーナーは2つのデジタル表示をホテルのロビーに設置したいと思っている。彼女は、顧客がそれらに気付いて欲しく、3つの潜在的な場所の1つにそれらを置くことができ、2つのデジタル表示の、3つの可能性のある構成となった (すなわち、場所1~2、1~3、又は2~3における表示)。堅牢性は、気付かれるために最も良い表示の場所を推薦するために算出される。

【0148】

1. デジタルカメラを使用してホテルのロビーの画像を捕捉し、この画像を、注意モデルを走らせることができる汎用コンピュータにダウンロードする。コンピュータは、Matlab (商標) (The MathWorks (Natick, Massachusetts) から入手可能) と共に、インストールされた視覚的注意モデルソフトウェア (例えば Koch & Itti) を有する。

【 0 1 4 9 】

2. 画像を修正してシミュレーションされたデザイン表示を含む。3つの修正された画像を生成し、それぞれは3つの潜在的な場所において2つの表示の全ての組み合わせが作られるように、それぞれ2つのデジタル表示をシミュレーションする。Photoshop (商標) (Adobe Co. (San Jose, CA) から入手可能) など、標準のデジタル写真操作プログラムを使用する。各シミュレーションされたデザイン表示は、正しく拡大縮小され、コンテンツ、例えばホテルのロゴのグラフィックなどコンテンツをシミュレーションした。ピクセルのアドレスによって定義されたように、コンピュータ上のファイル内の3つのデジタル表示の場所のそれぞれと関連して、画像領域を保存する。

【 0 1 5 0 】

3. 注意モデルを介して修正された画像を走らせる。出力は修正された画像の予測された顕著性のある領域を含む。それぞれ顕著性のある領域は、ステップ2で保存されたデジタル表示ピクセルアドレスと比較される。顕著性のある領域が保存されたピクセルのアドレス内にある、又はそれと重なり合う場合、期待された注意は望ましい場所に行く。3つの修正された画像のそれぞれは、上位10の凝視状態にあるデジタル表示を示し、したがって3つの場所のいずれかがを良好な候補であるということを確認する。

【 0 1 5 1 】

4. ステップ1に指定されたように、Koch & Ittiの基本的な視覚的注意モデルと共に開始する。修正された画像 (例えば100のモデル変動) の解析に利用するためのモデル変動の数を指定する。各視覚的注意モデルは3つの異なる特徴マップ (色、向き、及び輝度) を有し、顕著性マップはこれらのマップのそれぞれの、重み付けされた組み合わせとして算出される。基本的な視覚的注意モデルは、各マップに対する重み付けパラメータを等しいように (1, 1, 1) 設定する。100のモデル変動を作るために、各モデルの重み付けベクトルをランダムに設定する。これは、各重み付けをランダムに設定し、重み付けの合計を3に標準化するアルゴリズムによって完了される ($3 * (RandWeights / sum(RandWeights))$)。

【 0 1 5 2 】

5. 100の視覚的注意モデルの変動 (100のランダムに重み付けしている値によって定義される) によって各画像を解析し、ステップ2及び3に記載されているように、これらの結果を保存されたピクセルアドレスと比較する。デジタル表示場所1及び2に関連する修正された画像のシリーズは、予測された凝視が両方のデジタル表示場所に、画像の20%に行くことを示す。同様に、場所1及び3は、両方の表示場所に行く凝視の35%を有し、一方で、場所2及び3は、両方の表示場所に行く凝視の85%を有する。場所2及び3において設置された表示を有することは、ホテルにとって最も耐性のある構成となる。この推薦をホテルのオーナーに提供する。

【 0 1 5 3 】

実施例3：シーン最適化

背景：ホテルのオーナーは、彼女のロビー及び2つのデジタル表示に表示されるコンテンツを視覚的に最適化したいと思う。彼女の具体的な視覚目標は、顧客が4つの標的オブジェクトに気付くことである：第1及び第2デジタル表示、ホテルのレストランを広告する静止グラフィック表示、及びチェックインカウンタの後ろにいるスタッフ。

【 0 1 5 4 】

1. 最適化オブションのためのスコアを生成するために、標的オブジェクトに注意を引いた変化には報酬が与えられ、現実世界のコストは、許容可能な変化に関連付けられる。労働力及び供給コストに関連する、ドルで見積もられたコストは、以下のように考慮される潜在的な変化に割り当てられる。

【 0 1 5 5 】

チェックインカウンタの背後の配置された絵を移動させること：\$ 100、

レストランの表示の背後の照明を変更すること：\$ 2500、及び

2つのデジタル表示に表示されるコンテンツを再度デザインすること：各\$ 250。

【0156】

視覚的目標を達成するために割り当てられた報酬値は以下のとおり。

【0157】

2つのデジタル表示に注意を引き付けること：それぞれ\$500、
レストランの表示に注意を引き付けること：\$250、
チェックインカウンタの背後のスタッフに注意を引き付けること：\$150。

【0158】

2. デジタルカメラを使用して既存のロビーの画像を捕捉し、この画像を、注意モデルを走らせることができるコンピュータにダウンロードする。コンピュータは、Matlab (商標) (The MathWorks, Natick, Massachusetts) と共に、視覚的注意モデリングソフトウェア (例えばKoch & Itti) を有する。

10

【0159】

3. 可能性のある変化の全ての可能性のある組み合わせと関連付けられた複数の画像を作るために、検討される変更を反映するために画像を修正する。Photoshop (Adobe (San Jose, CA) から入手可能) など、標準のデジタル写真操作プログラムを使用する。顧客の視覚的目標に関連する標的オブジェクトのピクセルのアドレスもまた特定され、メモリに保存される。

【0160】

4. 注意モデルを使用して、ステップ3からの各画像を解析し、モデルによって予測された顕著なオブジェクト、標的オブジェクトのための保存されたピクセルのアドレスと比較する。標的オブジェクトのためのピクセルのアドレスと共に、予測された視覚的注意の重なり合いによって示された、修正された画像において視覚的目標を達成するために、報酬値からの変化に対するコストを引くことによって、各シミュレーションされた構成に対してスコアが算定される。例えば、注意がレストランの表示に割り当てられるとき、絵の移動の変化を使用して、スコアは\$250 - \$100 = \$150である。シミュレーションされた画像の全てを解析した後、見出された、最もコスト的に有効な解決策は、\$100のコストで絵を動かし、\$250のコストでコンテンツのうちの1つの色を修正することである (総コスト\$350)。これらの変化は、視覚的目標の全てが、得られるのを可能にし、\$1400の報酬スコア及び\$1050の合計スコアを生み出す。

20

30

【0161】

実施例4：複数の視点のシーン解析

背景：実施例3から続いて、推奨された変化がなされている。ホテルのオーナーは、ホテルのロビーにおいて多数の視点から見られたときの各標的オブジェクトの視覚的顕著性を知りたいと思っている。

【0162】

1. ロビー全体に分散された4つの関心の場所が特定され、4つのデジタル写真が、カメラを90度インクリメントにおいて回転させることによって撮られ、合計16画像となり、各画像は1つの視点を表している。画像は、デジタルカメラを使用して撮られた写真である。画像を、注意モデルを走らせることができるコンピュータにダウンロードする。コンピュータは、Matlab (商標) (The MathWorks, Natick, Massachusetts) と共に、視覚的注意モデルソフトウェア、例えばKoch & Itti を有する。各視点に対して、標的オブジェクトのためのピクセルのアドレスが特定され、コンピュータのメモリに保存され、標的オブジェクトは識別子でタグ付けされた。

40

【0163】

2. ステップ1からの16画像のそれぞれは、注意モデルを使用して解析される。各画像に関して、どの標的オブジェクトが見ることができ、どのオブジェクトが、注意を引き付けると予想されるかは、視覚的注意モデルによって決められる。

【0164】

50

３．各標的オブジェクトが画像の全てにわたって見ることができる確率が算定され、それが注意を払われる確率もまた算定される。このデータは、ホテルのオーナーへの報告に示され、様々な視点からのロビーにおける視覚的特性のより良い理解をもたらす。

【 0 1 6 5 】

本明細書に記載の方法及びシステムのための例示の用途は、記載された特定の用途を超えて広い用途を有するということに注意されたい。例えば、これらの用途には小売りの環境が挙げられる。

【 図 1 】

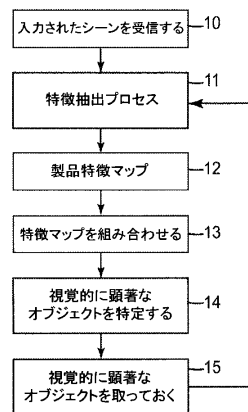


FIG. 1

【 図 2 】

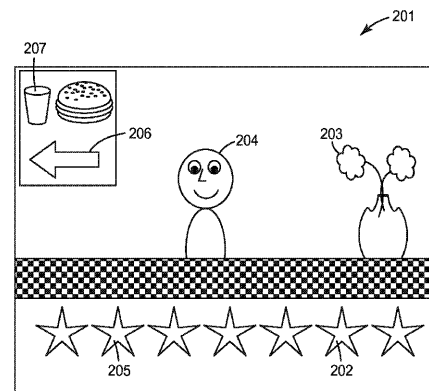


FIG. 2

【図 3 A】

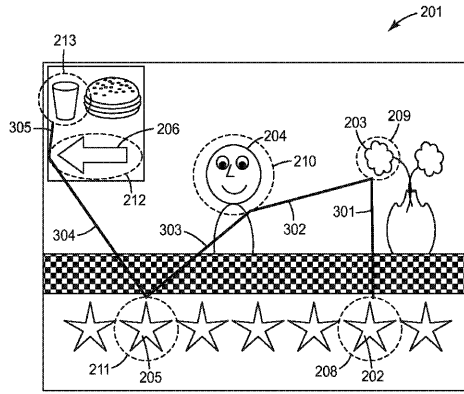


FIG. 3A

【図 3 B】

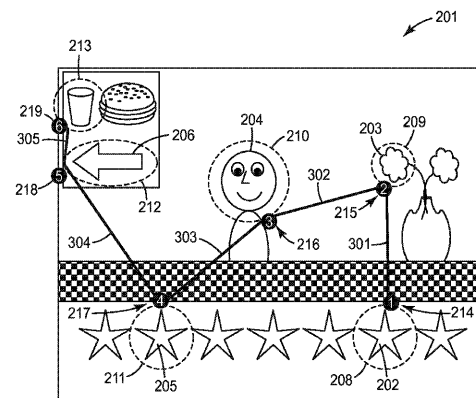


FIG. 3B

【図 4】

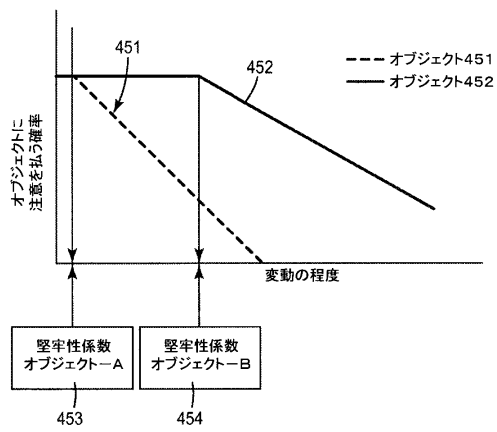


FIG. 4

【図 5】

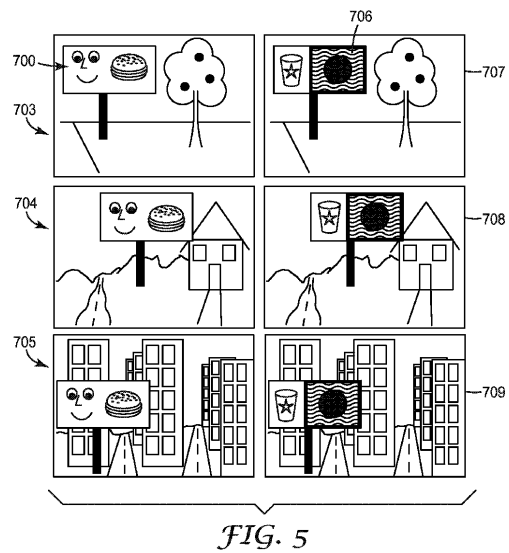


FIG. 5

【図 6】

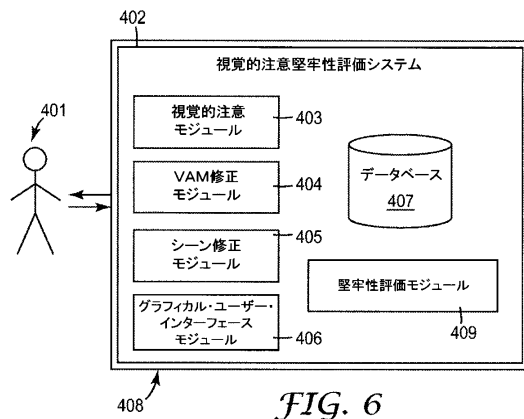


FIG. 6

【図 7】

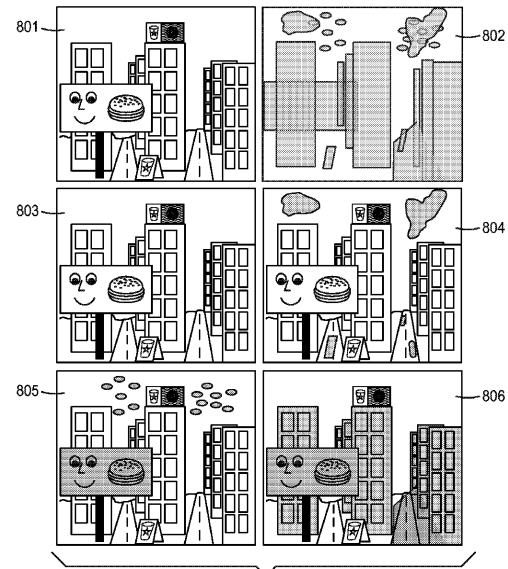


FIG. 7

【図 8】

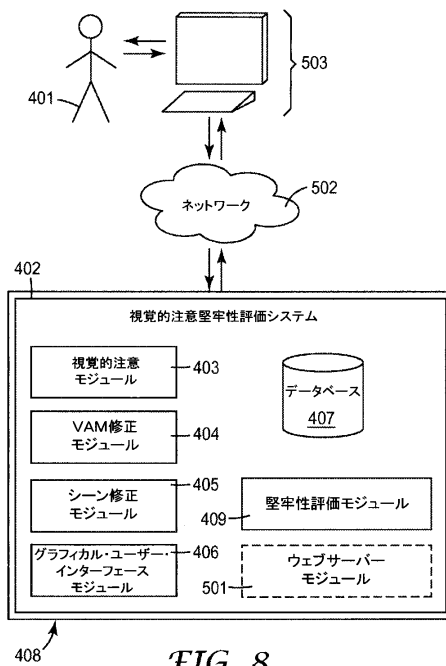


FIG. 8

【図 9】

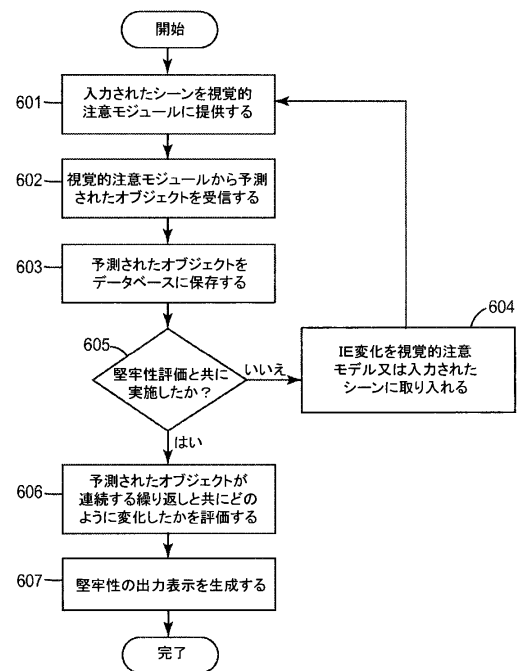


FIG. 9

【図 10】

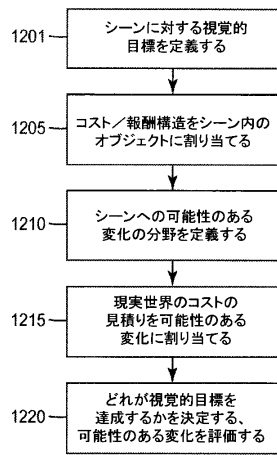


FIG. 10

【図 11】

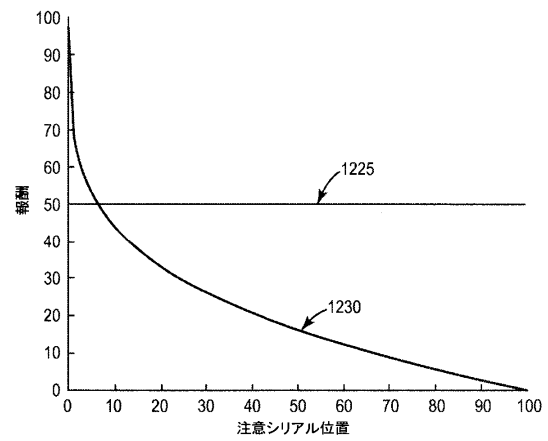


FIG. 11

【図 12】

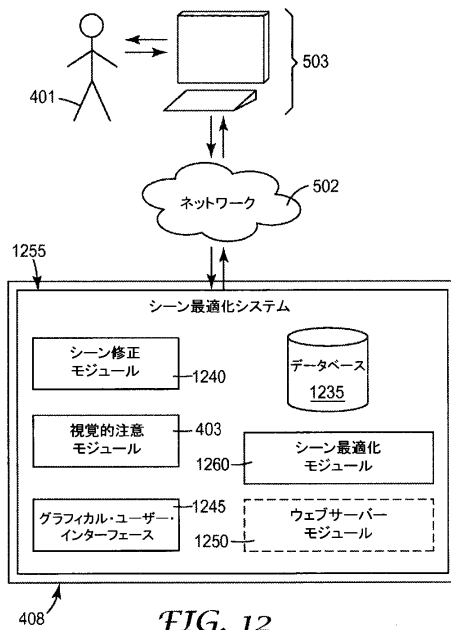


FIG. 12

【図 13 A】

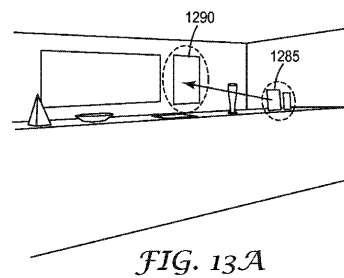


FIG. 13A

【図 13 B】

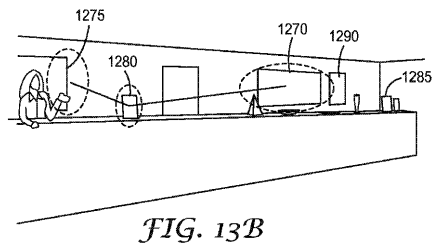


FIG. 13B

【図 14 A】

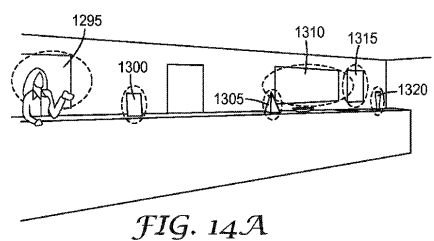
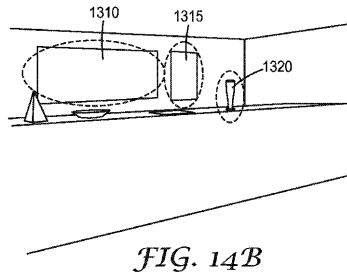
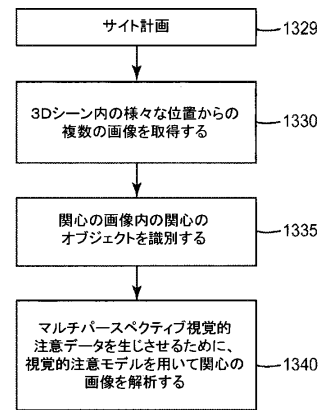


FIG. 14A

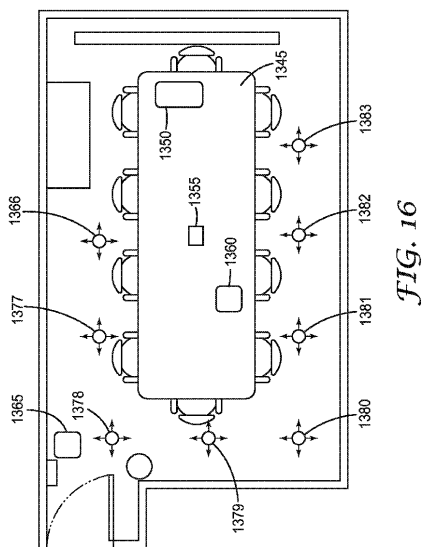
【図 14 B】



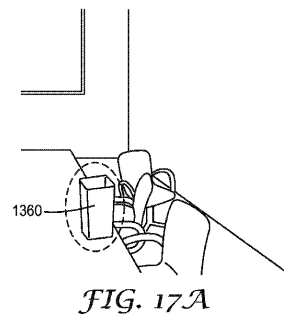
【図 15】



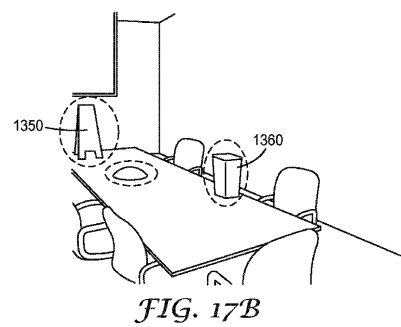
【図 16】



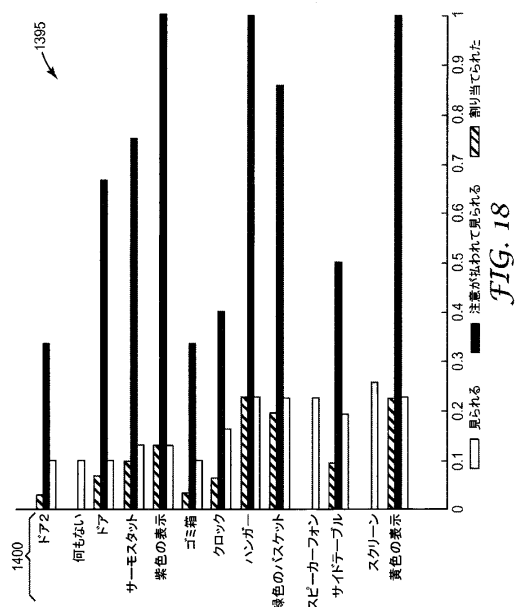
【図 17 A】



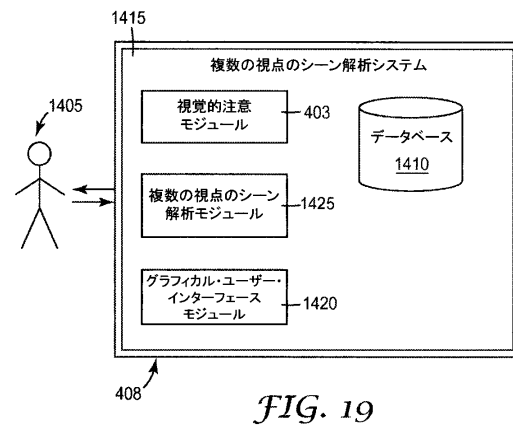
【図 17 B】



【図 18】



【図 19】



フロントページの続き

(74)代理人 100165191

弁理士 河合 章

(72)発明者 スタンキーウィクス, ブライアン ジェイ.

アメリカ合衆国, ミネソタ 55133-3427, セント ポール, ポスト オフィス ボックス 33427, スリーエム センター

(72)発明者 ブルックス, ブライアン イー.

アメリカ合衆国, ミネソタ 55133-3427, セント ポール, ポスト オフィス ボックス 33427, スリーエム センター

(72)発明者 リンジー, ブライアン エル.

アメリカ合衆国, ミネソタ 55133-3427, セント ポール, ポスト オフィス ボックス 33427, スリーエム センター

(72)発明者 アンダーソン, ネイサン ジェイ.

アメリカ合衆国, ミネソタ 55133-3427, セント ポール, ポスト オフィス ボックス 33427, スリーエム センター

(72)発明者 カナバン, マイケル ケリー

アメリカ合衆国, ミネソタ 55133-3427, セント ポール, ポスト オフィス ボックス 33427, スリーエム センター

(72)発明者 キャスナー, グレン イー.

アメリカ合衆国, ミネソタ 55133-3427, セント ポール, ポスト オフィス ボックス 33427, スリーエム センター

(72)発明者 ガードナー, ティモシー ジェイ.

アメリカ合衆国, ミネソタ 55133-3427, セント ポール, ポスト オフィス ボックス 33427, スリーエム センター

(72)発明者 ミスマー, デイビッド ケー.

アメリカ合衆国, ミネソタ 55133-3427, セント ポール, ポスト オフィス ボックス 33427, スリーエム センター

審査官 新井 則和

(56)参考文献 特表2004-520735(JP, A)

Heidmann, G., "Region saliency as a for colour segmentation stability.", Image and Vision Computing, 2007年 6月, Vol.26 No.2, pp.211-227

(58)調査した分野(Int.Cl., DB名)

G06T 1/00-7/60