

一人称視点における視覚的顕著性マップモデルの性能評価

山田健太郎[†] 菅野 裕介[†] 岡部 孝弘[†] 佐藤 洋一[†] 杉本 晃宏^{††}
開 一夫^{†††}

[†] 東京大学生産技術研究所, 〒 153-8505 東京都目黒区駒場 4-6-1

^{††} 国立情報学研究所, 〒 101-8430 東京都千代田区一ツ橋 2-1-2

^{†††} 東京大学大学院総合文化研究科, 〒 153-8902 東京都目黒区駒場 3-8-1

E-mail: [†]{yamada,sugano,takahiro,ysato}@iis.u-tokyo.ac.jp, ^{††}sugimoto@nii.ac.jp,

^{†††}khiraki@idea.c.u-tokyo.ac.jp

あらまし 視覚的顕著性マップモデルは、画像・映像から視覚的注意を推定するために用いられている。しかし、一人称視点での視覚的注意推定においては、顕著性マップモデルの有効性は確認されていない。また、既存の顕著性マップモデルでは、一人称視点から撮影された映像に生じる自己運動による影響が考慮されていない。本研究では、一人称視点における顕著性マップモデルを用いた視覚的注意推定の性能評価を行うため、頭部装着型カメラにより撮影された一人称視点映像の顕著性マップと、同時に計測された視線データを比較する実験を行った。その結果、静的特徴と比較して動的特徴の性能が低く、自己運動の影響によって性能低下が生じることを確認した。

キーワード 視覚的顕著性, 視覚的注意, 視線, 一人称視点

Evaluating Conventional Saliency Map Models for Estimating Human Egocentric Visual Attention

Kentaro YAMADA[†], Yusuke SUGANO[†], Takahiro OKABE[†], Yoichi SATO[†], Akihiro
SUGIMOTO^{††}, and Kazuo HIRAKI^{†††}

[†] Institute of Industrial Science, The University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505 Japan

^{††} National Institute of Informatics

2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, 101-8430, Japan

^{†††} Graduate School of Arts and Sciences, The University of Tokyo

3-8-1 Komaba, Meguro-ku, Tokyo, 153-8902

E-mail: [†]{yamada,sugano,takahiro,ysato}@iis.u-tokyo.ac.jp, ^{††}sugimoto@nii.ac.jp,

^{†††}khiraki@idea.c.u-tokyo.ac.jp

Abstract The validity of using conventional saliency map models to predict human attention was investigated for video captured with an egocentric camera. To evaluate the validity of using saliency map models for egocentric vision, an experiment was carried out to examine the correlation between saliency maps and measured gaze points for egocentric vision. The results show that conventional models can predict visually salient regions better than chance for egocentric vision and that the accuracy decreases significantly with an increase in visual motion induced by egomotion, which is presumably compensated for in the human visual system. This latter finding indicates that a visual saliency model is needed that can better predict human visual attention from egocentric videos.

Key words Visual saliency , Visual attention , Gaze , First person vision

1. はじめに

人間の視覚的な注意は、我々の内部状態を推定するための重要な手がかりである。それゆえ、インタラクティブ広告や、Attentive User Interface といった、人間中心のインタラクティブシステムに効果的に用いられている。

人間中心のメディアの一つである一人称視点映像においても、視覚的注意を推定することは有用である。一人称視点映像とは、頭部に装着されたカメラ等の観察者と同じような視点（一人称視点）から撮影された映像のことである。一人称視点のカメラは日常的な自己の行動を観測するのに非常に適している。それゆえ、一人称視点から撮影されるダイナミックなシーンを分析することで、行動意図を理解し、支援を行うこと等を目的とした研究も行われている。このような目的のためにも、一人称視点映像中のどこに注意が向けられているのか、という情報は非常に有益である。視覚的注意と視線は密接に関係しているため、一人称視点に限らず、多くの視線計測技術が様々なアプローチに基づいて開発されてきた。しかしながら、拘束の少ない環境で視線を計測することは、現在もなお困難な課題である。

視線計測技術とは異なるアプローチで視覚的注意の推定を行う手法として、視覚的顕著性マップモデルが挙げられる。顕著性マップモデルは、人間が注意を向けやすい画像中の領域を推定するための計算モデルである。心理学分野の視覚的注意に関する研究 [1] を踏まえ、Kock と Ullman が顕著性マップの概念を提案した [2]。その後、完全な計算モデルとしての顕著性マップモデルが Itti らにより提案された [3]。それ以来、多数の顕著性マップモデルが提案されてきた [4]~[9]。たとえば、運動方向や輝度変化といった低レベルな動的特徴を導入することで、静止画像だけでなく動画にも適用可能なモデルが提案されている [4]。また、視線計測に基づいた研究 [10]~[12] では、顕著性マップは実際の人間の注意分布に適合していることが示されている。このようなモデルの拡張、有用性の確認に基づき、近年、顕著性マップモデルはセグメンテーションや画像編集等、人間の知覚・注意を利用したアプリケーションに、広く応用されている。

一人称視点においても、顕著性マップモデルを用いることで、一人称視点映像のみを入力とした視覚的注意の正確な推定が可能であれば、教育や医療、福祉における行動支援等、様々な分野での応用が期待できる。しかしながら、人間中心の視覚における視覚的注意のメカニズムは大幅に異なる。たとえば、自己運動によって生じた視覚刺激は、人間の視覚システムにおいては抑制されるが、既存の顕著性マップモデルでは、自己運動によって生じる視覚的運動刺激といった重要な観点が考慮されていないため、そのような抑制メカニズムは明示的には組み込まれていない。

また、既存の顕著性マップモデルは、これまで一人称視点映像における性能評価が行われていないため、一人称視点における視覚的注意の推定に有効であるかどうか明らかではない。顕著性マップの性能評価を行った既存の研究では、既に撮影済みの画像や映像を提示した際の注意のみが検討されてきたので

ある。具体的には、ディスプレイ等を用いて被験者に提示された画像・映像から顕著性マップが計算され、ディスプレイ上の視線座標とマップの比較により、顕著性マップの性能が評価された。つまり、既存の研究において、顕著性マップモデルによる観察者の注意の推定や視線の計測は、提示された画像・映像に限られており、より自由な環境、すなわち、実空間の物体や周辺環境を観察するような、一人称視点における視覚的注意と顕著性マップの相関は確認されていない。

そこで本研究は、一人称視点映像に対する視線計測結果とその顕著性マップの比較を行い、一人称視点における既存の顕著性マップモデルの性能を実験的に評価することを目的とする。実験では、頭部に装着したカメラを用いて一人称視点映像を撮影し、同時に、頭部装着型の視線計測装置を用いて視線計測を行った。また、複数の顕著性マップモデルと特徴の組み合わせについて定量的に性能の評価・比較を行い、一人称視点での視覚的注意の特性についての議論を行った。我々の知る限り、本研究は一人称視点における顕著性マップモデルの性能を実験的に評価した初めての研究である。

2. 関連研究

本節では、まず視覚的顕著性の背景にある理論について紹介した後、顕著性マップの計算モデルに関する先行研究について簡潔にまとめる。

人間の情報処理能力には限界があるため、大量の情報を同時に処理することは不可能である。そこで、情報の重要な部分を効率的に選択し、集中するために注意というメカニズムが必要とされる [13]。人間の視覚システムに関しても同じことが言え、視覚的注意は人間が目を通して受け取る大量の情報を扱うために必要となる。

視覚的注意のメカニズムを理解する鍵となる理論として、Treisman らによって提案された特徴統合理論 [1] がある。人間の視覚システムは初期処理において、視覚から取り入れた画像情報を、形、色、運動などの単純な特徴に分解して処理していることが知られている [14]。Treisman らは、単純な特徴や複雑な特徴による視覚刺激を被験者に探索させた実験の結果に基づき、人間の視覚的注意のメカニズムは、このような単純な特徴を統合する機能を備えていると結論付けた。

この理論に基づいて、Koch と Ullman は顕著度を符号化した 2 次元のマップとして、視覚的顕著性マップの概念を提案した [2]。この顕著度は単純な視覚特徴を統合することで生成され、その領域がどれくらい人間の注意を引きつけるかを意味している。

これに続き、Itti ら [3] は、完全な計算モデルとしての顕著性マップモデルを提案した。彼らは、画像から単純な視覚特徴を抽出する処理を導入し、人間の視覚における受容野の働きを模倣した処理により、顕著度の計算を行った。その後、顕著性マップモデルは運動方向と輝度変化という 2 つの動的特徴を追加することで拡張され、静止画像だけでなく、動画を扱うことが可能になった [4]。

また、異なるアプローチにより、顕著性マップの高精度化を

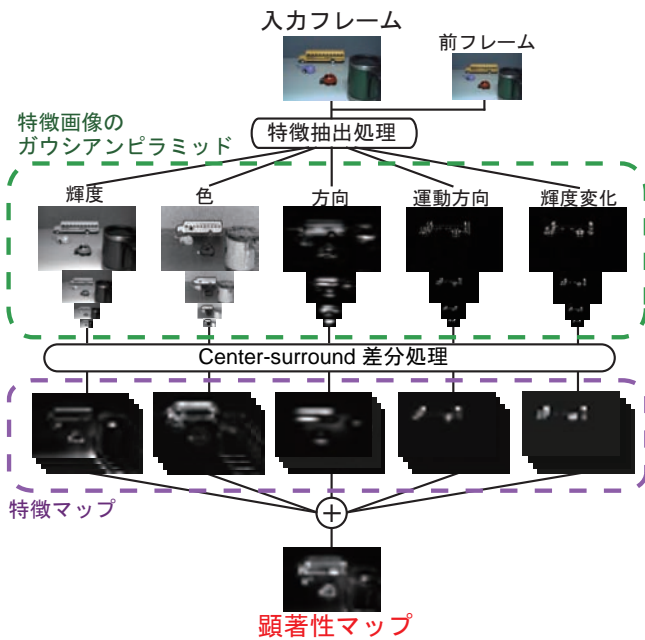


図 1 動画の顕著性マップ生成処理。

目指したモデルも提案されている．たとえば，近年グラフベースのアプローチによるモデルがいくつか提案されている [7] ~ [9]．入力画像のグラフはノード間の相違度関数と距離関数によって定義され，顕著度はグラフの定常状態を解析することで計算される．このようなグラフベースのアプローチをとった研究では，用いられる画像特徴の性能の向上よりも，単純な画像特徴を用いて，より高精度な顕著度を計算することに重点が置かれていた．

3. 動画の顕著性マップ生成処理

本研究では，一人称視点において顕著性マップモデルを適用することの有効性を，2種類の代表的な顕著性マップモデルを用いて評価実験を行った．1つは center-surround メカニズムに基づいた Itti らのモデル [4] であり，もう1つは Harel らのグラフベースモデル [7] である．本節では，まず Itti らのモデルの計算処理について説明した後に，それを改良した Harel らのモデルの処理について説明する．

主に3つのステージで構成される顕著性マップの計算処理を図1に示す．第1ステージでは，特徴抽出処理により，入力フレームから特徴画像のガウシアンピラミッドを生成する．第2ステージでは，center-surround メカニズムの処理により，特徴画像から，特徴マップを生成する．第3ステージでは，全ての特徴マップを正規化し，統合することで最終的な単一の顕著性マップを生成する．

第1ステージでは，単純な線形フィルタにより，入力画像を5種類の視覚的特徴画像に分解する．動画からの顕著性マップ生成においては，輝度，色，方向が静的特徴として，運動方向，輝度変化が動的特徴として，一般に用いられる．輝度特徴画像は，入力画像の RGB (赤，緑，青) チャンネルの平均として得られる．Itti らのモデルでは，赤と緑，青と黄色，という補色の関係にある2つの色チャンネルのセットから生成され

た，2つの異なる画像が色特徴画像として用いられている．一方，本研究ではこの RGB 色空間による色特徴画像の代わりに，Derrington-Krauskopf-Lennie (DKL) 色空間 [15] を用いて生成された色特徴画像を用いた．DKL 色空間は，網膜と視床における生理学的な色処理の3つのチャンネルの軸によって表現された色空間で，人間の脳内の色処理を扱うのに適している．方向特徴画像は，4方向 (0°, 45°, 90°, 135°) のガボールフィルタを用いて，輝度画像から計算される．

運動方向と輝度変化の特徴画像を得るためには，2つの入力フレームが必要となる．運動方向特徴画像は，現在のフレームと，前フレームの間で各4方向の方向特徴画像を空間的にシフトした差分画像から生成される．輝度変化特徴画像は，現在のフレームと，前フレームの間の輝度特徴画像の差分画像として計算される．結果として，輝度が1種類，色が2種類，方向が4種類，運動方向が4種類，輝度変化が1種類で，合計12種類の特徴画像が生成される．続いて，各特徴画像ごとに，9つのスケール (スケール 0 = 1:1 ~ スケール 8 = 1:256) ガウシアンピラミッド [16] を生成する．

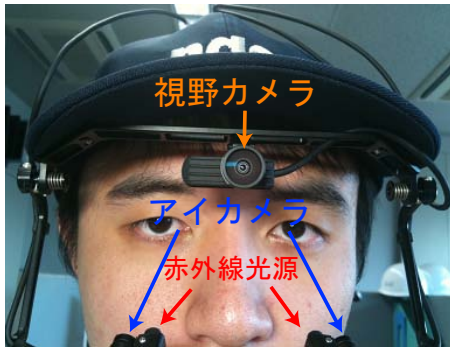
第2ステージでは，この特徴画像のガウシアンピラミッドから，center-surround メカニズムを用いて特徴マップが計算される．まず，2種類の異なるスケールサイズのガウシアンピラミッドを組み合わせ，6つのセットを作成する．続いて，各セットにおけるスケール間の補間と各ピクセルごとの減算により得られた，スケール間の差分画像を用いて，各特徴画像毎に6種類の特徴マップが計算される．

第3ステージでは，輝度が6種類，色が12種類，方向が24種類，運動方向が24種類，輝度変化が6種類の，合計72種類の正規化された特徴マップを統合することで最終的な顕著性マップを生成する．統合前には，各特徴マップにおいて， $(M - \bar{m})$ を全体に乗ずることで，正規化処理を行う．ここで， M はそのマップ全体の最大値であり， \bar{m} は全体の最大値を除いた局所最大値の平均である．この正規化処理には，少数の強いピークが存在するマップを全体的に増幅し，多数の同程度のピークが存在するマップを全体的に抑制する効果がある．

Harel らのモデル [7] は，用いる特徴画像は Itti らのモデルと同一であるが，第2ステージと第3ステージにおいてグラフベースのアプローチを採用している．特徴マップと最終的な顕著性マップはマルコフ連鎖グラフの平衡分布を計算することによって生成される．第2ステージにおいては，ノード間の相違度関数と距離関数を定義し，この2つの関数を互いに乗ずることで各ノードの重みを計算する．第3ステージにおいては，特徴マップ上の対応する位置の値と距離関数を乗ずることで各ノードの重みを得る．

4. 実験

上述のように，既存の顕著性マップモデルは単純な低レベルの特徴画像を用いて顕著性マップを計算している．これらのモデルは既に撮影済みの画像や映像を人間に提示した場合の視覚的顕著性を計算するために設計されており，自己運動によって生じた視覚刺激の取り扱いには考慮されていない．一人称視点に



(a)



(b)

図 2 (a) EMR-9 [17], ナックイメージテクノロジー製のモバイル型視線計測装置。EMR-9 は 2 つのアイカメラと、2 つの赤外線光源を備え、眼球運動を 240[Hz] で計測することができる。また、それと同時に、頭部に装着された視野カメラにより、一人称視点映像を 30[fps] で撮影を行う。視野カメラの水平画角は 121° であり、撮影される映像の解像度は 640 × 480 である。(b) 実験中に視野カメラから撮影された映像のフレームの例。

において既存の顕著性マップモデルを用いる有効性について評価するため、我々は以下に説明するような実験を行った。

4.1 実験手順

一人称視点における顕著性マップモデルの性能を評価するため、頭部に装着されたカメラから撮影された映像に対する視線計測結果と、顕著性マップの関係性を分析する実験を行った。

実験では、ナックイメージテクノロジー製のモバイル型視線計測装置 EMR-9 [17] を用いて視線位置の計測と一人称視点映像の撮影を行った。図 2(a) に示すように、EMR-9 は 2 つのアイカメラと、2 つの赤外線光源を備え、眼球運動を 240[Hz] で計測することができる。また、それと同時に、頭部に装着された視野カメラにより、一人称視点映像を 30[fps] で撮影する。視野カメラの水平画角は 121° であり、撮影される映像の解像度は 640 × 480 である。

前節で述べた Itti らのモデル [4] と、Harel らのモデル [7] を代表的な顕著性マップモデルとして評価に用いた。実験は、室内で着席した状態、室内で自由に動き回る状態、野外で自由に動き回る状態の、3 種類の実験設定で行った。5 人の被験者が実験に参加し、各実験設定につき、約 1 分間の一人称視点映像を 3 種類撮影し、合計 15 種類のデータを作成した。図 2(b) に、実験中に視野カメラから撮影された映像のフレームの例を示す。瞬目等によるエラーと高速な視線移動中の視線位置を取り除いた上で、各一人称視点映像につき 8,000 点の視線位置を実験に

表 1 特徴を統合した顕著性マップの AUC。

	静的特徴	動的特徴	全特徴
Itti らのモデル [4]	0.763	0.586	0.740
Harel らのモデル [7]	0.809	0.672	0.764

用いた。

人間の注意は実行中のタスクによって影響されるが、既存の顕著性マップモデルは、このような高レベルの注意のメカニズムを効果的に扱うことは出来ない。我々の研究の目的は、一人称視点において顕著性マップモデルを用いることの有効性を検討することであるため、着席の指示と自由な移動以外の特別なタスクを被験者に与えなかった。

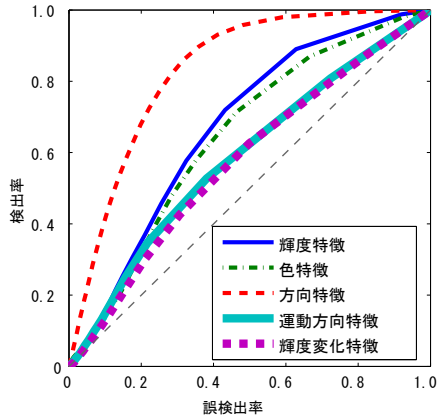
4.2 実験結果

視覚的注意の推定精度に各特徴がどの程度寄与しているかを評価するため、5 つの各特徴について、単一の特徴のみを用いて計算された顕著性マップ、及び、それらを組み合わせた顕著性マップと、視線計測結果の関係を分析した。図 3 に示す曲線は、実験に用いた全データセットを用い、顕著度の閾値を最小から最大まで変化させることで算出された ROC(Receiver Operating Characteristic) 曲線である。縦軸は検出率であり、顕著性マップにおいて顕著度が閾値よりも高い値を持つ領域に含まれた実際の視線位置の割合に対応する。横軸は誤検出率であり、顕著性マップ中で閾値よりも高い値を持ち、視線が存在しない領域の割合を示している。

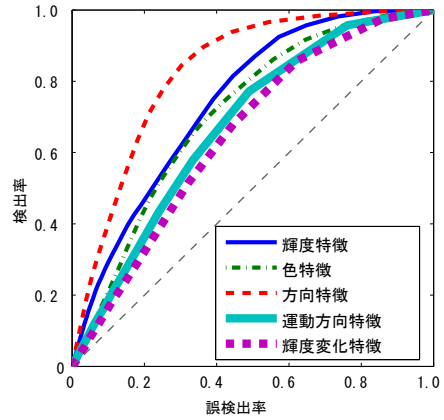
図 3 (a) 及び (b) は前節で説明した 2 つの顕著性マップモデルを用いて生成した、5 つの各特徴ごとの顕著性マップの性能比較である。図 3 (c) 及び (d) は静的顕著性マップ、動的顕著性マップ、標準顕著性マップの、3 つの顕著性マップの性能比較である。静的顕著性マップは輝度、色、方向の静的特徴のみを用いて計算された顕著性マップであり、動的顕著性マップは運動方向、輝度変化の動的特徴のみを用いて計算された顕著性マップである。標準顕著性マップは全ての特徴を統合した、通常の顕著性マップである。図 3 (a) 及び (c) は Itti らのモデル [4] を用いた結果であり、図 3 (b) 及び (d) は Harel らのモデル [7] を用いた結果である。表 1 に示す AUC (Area under the curve, ROC 曲線の曲線下面積) は、図 3 (c), (d) における 3 本の ROC 曲線の推定性能を表している。これらの結果から、これらの既存の顕著性マップモデルは、一人称視点における視覚的注意を偶然に比ると十分良く推定できているが、両モデルにおいて動的特徴は性能に寄与しておらず、それどころか精度を低下させてしまっていることがわかる。

4.3 考察

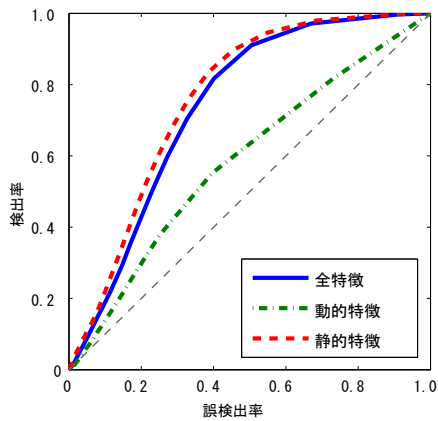
実験結果は、一人称視点において、運動方向と輝度変化の動的特徴が明らかに予測性能を低下させていることを示している。しかしながら、実験において、動的な視覚的刺激が被験者の注意を引いた場面が複数回観測された。図 4 は被験者が移動物体 (歩行者) を見ていたシーンの視線軌跡の例である。また、図 5 は輝度が急激に変化した物体 (ノート PC のモニター) を含むシーンの視線軌跡の例である。これらのケースでは、被



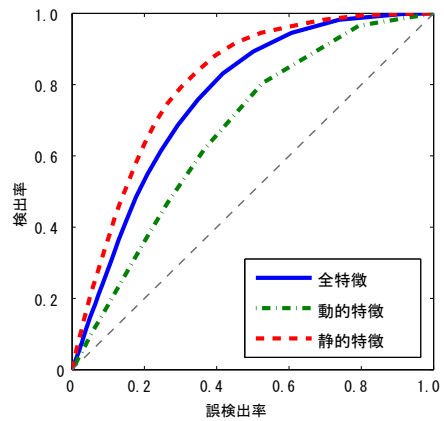
(a) 各特徴 / Itti らのモデル [4]



(b) 各特徴 / Harel らのモデル [7]



(c) 特徴の組み合わせ / Itti らのモデル [4]



(d) 特徴の組み合わせ / Harel らのモデル [7]

図 3 各特徴ごとのマップの ROC 曲線 ((a) Itti らのモデル [4], (b) Harel らのモデル [7]) 及び、静的特徴, 動的特徴, 全ての特徴を用いたマップの ROC 曲線 ((c) Itti らのモデル [4], (d) Harel らのモデル [7]). ROC 曲線は, 実験に用いた全データを用い, 顕著度の閾値を最小から最大まで変化させることで算出した. 縦軸は検出率であり, 各マップにおいて顕著度が閾値よりも高い値を持つ領域に含まれた実際の視線位置の割合に対応する. 横軸は誤検出率であり, 各マップ中で閾値よりも高い値を持ち, 視線が存在しない領域の割合を示している.



図 4 被験者が移動物体 (歩行者) を見ていたシーンの視線軌跡例. 一人称視点映像に運動方向特徴のみによる顕著性マップを重ね合わせてある. 白い十字は視線位置を示す.

験者は動的な視覚刺激の変化に対して注意を向けているが, 物体の移動, 輝度の変化に対応する運動方向特徴, 輝度変化特徴による顕著性マップでは, 動的変化の無い他の位置に, 高い顕著度が与えられていた. このことは, 既存のモデルで提案された特徴が, 一人称視点においては動的な視覚刺激を適切に扱うことができないことを示唆している.

図 6 に, 自己運動を伴うシーンにおける, 動的特徴による顕

著性マップの例を示す. (a) は運動方向特徴のみを用いて生成された顕著性マップであり (b) は輝度変化特徴のみを用いて生成された顕著性マップである. 上段が入力画像, 中段が各動的特徴による顕著性マップ, 下段が入力画像に顕著性マップを重ね合わせたものである.

既に撮影済みの画像や映像を人間に提示するケースと異なり, 我々の想定しているケースでは, 自己運動の影響が生じる. 人

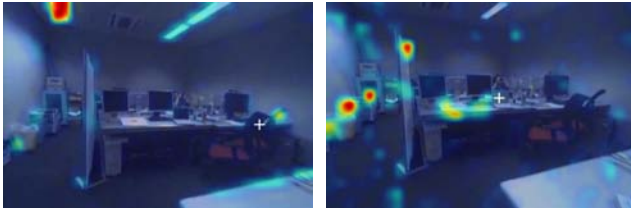
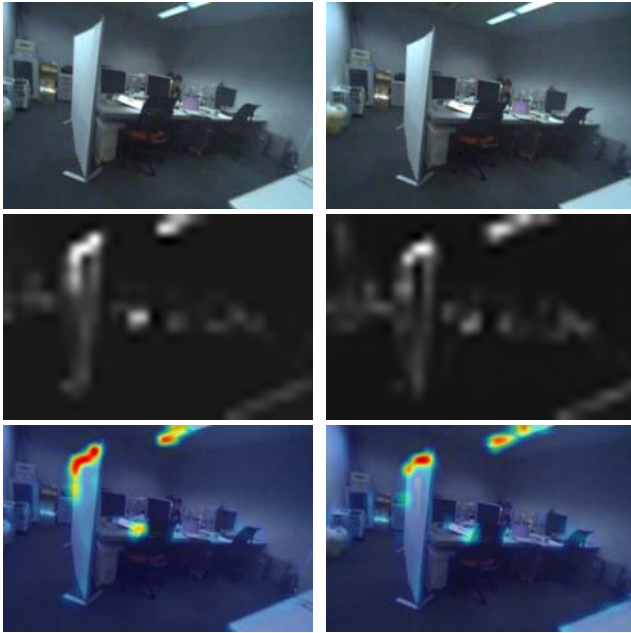


図 5 輝度が急激に変化した物体（ノート PC のモニター）を含むシーンの視線軌跡例．一人称視点映像に輝度変化特徴のみによる顕著性マップを重ねてある．白い十字は視線位置を示す．



(a) 運動方向特徴 (b) 輝度変化特徴

図 6 自己運動を伴うシーンにおける，動的特徴による顕著性マップの例．(a) は運動方向特徴のみを用いて生成された顕著性マップであり (b) は輝度変化特徴のみを用いて生成された顕著性マップである．上段が入力画像，中段が各動的特徴による顕著性マップ，下段が入力画像に顕著性マップを重ね合わせたものである．

間の視覚システムが自己運動を補償する機能を有している [18] 一方で，既存の顕著性マップモデルはこのような補償のためのメカニズムを考慮していない．それゆえ，動的特徴による顕著性マップでは，実世界では動的な変化が無いにもかかわらず，自己運動の影響により，視覚的な運動方向や輝度変化が周囲と異なる領域に，高い顕著度が現れている．しかしながら，このような視覚刺激に対する注意は，自己運動の補償機能により抑制されるため，顕著性マップの性能低下の原因となっている．

5. 結論と今後の課題

本研究では，一人称視点から撮影された映像から計算した顕著性マップと視線位置の関係を実験的に分析することで，一人称視点における視覚的顕著性マップモデルを用いた視覚的注意推定の性能評価を行った．実験の結果，既存の顕著性マップモデルは，一人称視点においても視覚的注意を偶然よりよく予測できることを確認した．しかしながら，既存のモデルは人間の視覚が持つ，自己運動による影響を補償するメカニズムをモデ

ル化できておらず，一人称視点においては，動的特徴が性能を低下させていた．それゆえ，既存のモデルは一人称視点映像を効率よく取り扱うためには，改良の余地が十分残されているといえる．

今後の課題としては，自己運動の影響を考慮することで，動的特徴を改良し，一人称視点に対応した顕著性マップモデルを構築することが挙げられる．

文 献

- [1] A.M. Treisman and G. Gelade, "A feature-integration theory of attention," *Cognitive Psychology*, vol.12, no.1, pp.97–136, 1980.
- [2] C. Koch and S. Ullman, "Shifts in selective visual attention: towards the underlying neural circuitry.," *Human neurobiology*, vol.4, no.4, pp.219–227, 1985.
- [3] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.20, no.11, pp.1254–1259, 1998.
- [4] L. Itti, N. Dhavale, F. Pighin, et al., "Realistic avatar eye and head animation using a neurobiological model of visual attention," *SPIE 48th Annual International Symposium on Optical Science and Technology*, vol.5200, pp.64–78, 2003.
- [5] T. Avraham and M. Lindenbaum, "Esaliency (extended saliency): Meaningful attention using stochastic image modeling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.32, no.4, pp.693–708, 2010.
- [6] M. Cerf, W. Harel, J. and Einhäuser, and C. Koch, "Predicting human gaze using low-level saliency combined with face detection," *Advances in Neural Information Processing Systems*, vol.20, pp.241–248, 2008.
- [7] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," *Advances in Neural Information Processing Systems*, vol.19, pp.545–552, 2006.
- [8] L.F. Costa, "Visual saliency and attention as random walks on complex networks," *ArXiv Physics e-prints*, arXiv:physics/0603025, pp.1–6, 2006.
- [9] W. Wang, Y. Wang, Q. Huang, and W. Gao, "Measuring visual saliency by site entropy rate," *Computer Vision and Pattern Recognition IEEE*, pp.2368–2375 2010.
- [10] T. Foulsham and G. Underwood, "What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition," *Journal of Vision*, vol.8, no.2:6, pp.1–17, 2008.
- [11] L. Itti, "Quantitative modelling of perceptual salience at human eye position," *Visual Cognition*, vol.14, no.4, pp.959–984, 2006.
- [12] D. Parkhurst, K. Law, and E. Niebur, "Modeling the role of salience in the allocation of overt visual attention," *Vision Research*, vol.42, no.1, pp.107–123, 2002.
- [13] L.M. Ward, "Attention," *Scholarpedia*, vol.3, no.10, p.1538, 2008.
- [14] D.E. Broadbent, *Perception and communication*, Pergamon Press, 1958.
- [15] A.M. Derrington, J. Krauskopf, and P. Lennie, "Chromatic mechanisms in lateral geniculate nucleus of macaque.," *The Journal of Physiology*, vol.357, no.1, pp.241–265, 1984.
- [16] H. Greenspan, S. Belongie, R. Goodman, P. Perona, S. Rakshit, and C. Anderson, "Overcomplete steerable pyramid filters and rotation invariance," *Computer Vision and Pattern Recognition IEEE*, pp.222–228 1994.
- [17] "nac image technology inc. emr-9". <http://www.nacinc.com/products/Eye-Tracking-Products/EMR-9/>
- [18] I.P. Howard, "The optokinetic system," *The Vestibulo-ocular Reflex and Vertigo*, pp.163–184, 1993.