

ボトムアップ要因による視線分布の推移推定

Estimation of Transition of the Eye-Gaze Distribution that originates in Bottom-Up Factor

梅田 修一† 小峯 一晃† 比留間 伸行†
 Shuichi Umeda Kazuteru Komine Nobuyuki Hiruma

1. まえがき

テレビ視聴時における眼球運動を測定(視線計測)して、視聴者の興味などの内的な状態を推定することを検討している。興味ある情報を周辺視野よりも分解能の高い中心窩で知覚しようとする結果が眼球運動に現れるとするならば、視線計測は内的な状態推定のために有効であると考えられる。しかしながら、視覚的注意の研究によれば、仮に画面の中央に何か動くものがあるとしてそこに視線を向けている観測者がいるとしても、その観測者が中央の動くものに興味を持っているとは断定できない(図1)。

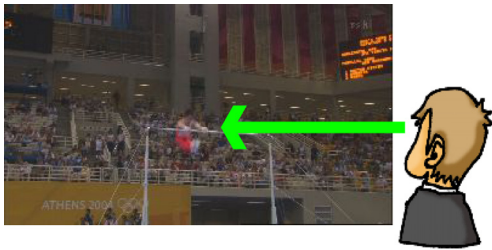


図1. 興味があるから見ているのだろうか..

どこに視線を向けるかは、興味といった観測者の内的な要因だけで決まるわけではない。視線がどこに向かうかを定める要因としては観測者の内的な状態に依存するトップダウン要因のほかに、画像を構成する色や形状等が引き起こし、観測者の興味とは無関係なボトムアップ要因がある[1]。

眼球運動そのものの測定は計測装置の性能の発達により、自然な視聴形態においても必要十分な精度の視線データを得られるようになった。それにも関わらず観測者の内的な状態を推定する手法が確立されていないのは、対象に視線を向けているときのボトムアップ要因を十分な精度で推定できておらず、今向けている視線がトップダウン要因によるものかボトムアップ要因によるものかどうかを判断できないためと推測できる。

ボトムアップ要因による視線を予測する視覚モデルの中で代表的なものに、顕著性マップと winner-take-all(WTA)を組み合わせた顕著性マップモデル[2][3]がある。トップダウン要因による影響を抑えた視線計測実験を行うことによって、最適な顕著性マップを導出することが試みられているが、導出法の確立までには至っていない。その一つの原因として、画面の中央に視線が集中する傾向を示す「センターバイアス」の影響が、計測した視線データの有効性を失わせることが指摘されている[4][5]。

本報告では、視線の動きを調べる基礎的な実験結果から明らかとなっている「グローバル効果」[6]を根拠にして顕

著性マップモデル修正を試みる。提案する修正モデルに基づくシミュレーションの結果、1枚の提示画像に対して複数の予測視線分布を導出する。視線分布の推移を表す次元が1つ増える代わりに、センターバイアスの影響を推定できる可能性を示す。

今回報告する成果は、我々の研究の目的である視聴者の興味などの推定に応用できる可能性を示唆するものである。

2. 研究の背景

2.1. ボトムアップ-トップダウン要因の視覚モデル

図2-1.に示す通り、人間の視線の位置決定にはトップダウン要因だけでなくボトムアップ要因が作用する[1]。つまり観測者の内的な状態に関係なく、画像の物理特徴量によって視線が誘導されることがある。両者が視線に与える影響度の割合は、画像提示後の時間が大きく関係すると考えられる。ボトムアップ要因のみが視線に与える影響を調べるための実験では、刺激画像の提示時間を数秒間と短くすることでトップダウン要因の影響を抑えている[4][5][7]。

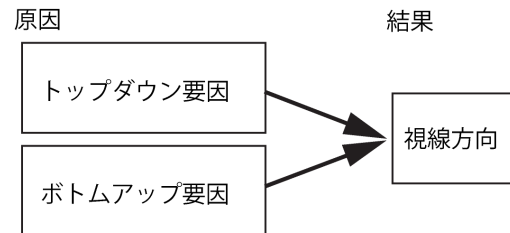


図2-1. ボトムアップ-トップダウン要因 視覚モデル

2.2. 顕著性マップについて

本報告ではボトムアップ要因による視線分布の予測のために顕著性マップを用いる。顕著性マップは視覚環境内に存在する顕著性をコードするモジュールであり、種々の初期視覚特徴をコードしている特徴マップの重み付き線形和により構成される[8]。ただし、重みの最適な値を求める手法はまだ確立されていない。特徴マップの種類には色・方位・輝度・運動・点滅等が考えられており、多くの研究では全ての特徴マップの重みを等しくして顕著性マップとしている。顕著性マップは提示画像における顕著性の高さを、マップ上の画素値として表したものである。視線を予測するためには、この顕著性マップを元に、顕著性の高い部分に視線を向けていくとする winner-take-all(WTA)のモジュールが組み合わされた顕著性マップモデルが提案されている(図2-2)。WTAには顕著性の高い部分から順番に視線を向ける決定論的 WTA[1]と、確率的に顕著性の高い部分を選択する確率論的 WTA[8]があるが、顕著性の高い部分のみに視線を分布させることになる点は共通する。

† NHK 放送技術研究所 人間・情報科学研究部
 Japan Broadcasting Corporation Science & Technology
 Research Laboratories Human & Information Science
 Reserch Division

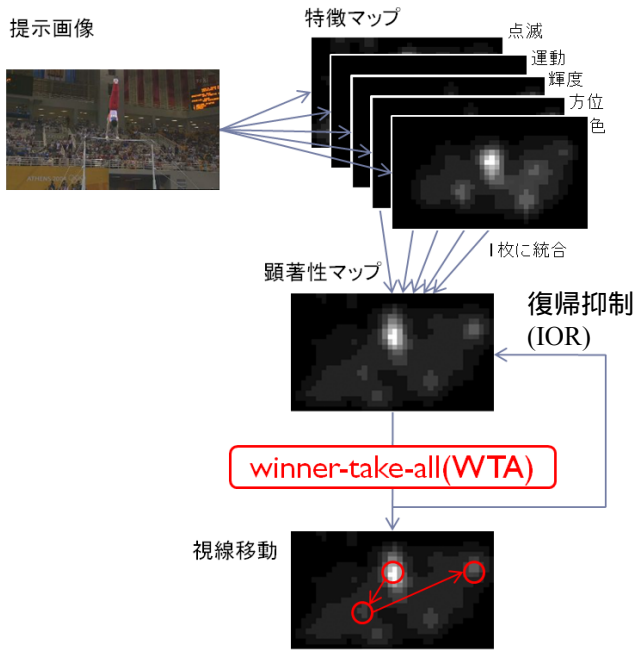


図 2-2. 顕著性マップモデル

2.3 グローバル効果について

視線測定実験を行うと、提示画像中の複数の注視対象の中間点に向かう視線移動が見られる。この現象はグローバル効果(global effect)として知られている[6]。

文献[6]が示す実験結果は、WTA で予測される視線とは異なり、顕著性が低い部分にも視線が分布することを示す。視線移動に影響を与えるこの効果は、視覚経路における比較的後期の段階の働きであることが示唆されている[9]。

2.4 センターバイアスについて

強制的に画面中央から外側に視線をそらす誘導を行わない限り、画像を見る際にはセンターバイアスが作用する。番組から抽出した映像を見ている時の視聴者の視線を測定すると[7]、多くの場合において視線は画面の中央に集中する(図 2-3)。予測した視線と実際の視線の一致度を比較した研究[4]によると、色や輝度、運動といった特徴量から予測した視線よりも、センターバイアスのみから予測した視線の方が、計測した視線との一致度が高いことが多い。



図 2-3. 実験に用いた番組(左)と計測した視線(右) [7]

3. 修正視覚モデルの提案

これまで考えられてきたボトムアップ要因を生じる段階よりも高次であるが、トップダウン要因とするまでには至らない影響を追加した視覚モデルを提案する。これまでの顕著性マップモデルでは WTA により選択的に、ボトムアップ要因の視線を予測してきた(図 2-2)。しかしながら 2.3.で指摘したグローバル効果が示すとおり、顕著性が高

い部分を選択するだけでは予測不可能な視線の存在が明らかになっている。そこでグローバル効果の影響を視覚モデルに組み込むことにする(図 3.)。

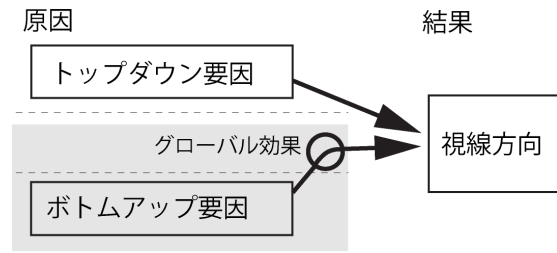


図 3. グローバル効果を追加した修正視覚モデル

グローバル効果はしばしば重心効果とも呼ばれる[9]。重心という用語から連想される力学系における重心においては、構成要素の剛体の質量と位置が全て計測可能ならば、重心位置の推定が可能である。そこで力学系における構成要素を、視覚モデルの構成要素に置き換えることによってモデルを提案する。

力学系において系に複数の剛体が分布するとして、 i 番目の剛体の質量を M_i 、位置を \mathbf{r}_i とする。このとき、

$$\mathbf{r}_c = \frac{\sum_i M_i \mathbf{r}_i}{\sum_i M_i} \quad (1)$$

により重心位置 \mathbf{r}_c が求められる。ここで力学系における構成要素を以下の通り、修正視覚モデルの構成要素に置き換える。

力学系	修正視覚モデル
剛体	注意要素 (仮想的な要素)
質量 M_i	視線決定に及ぼす重み
重心位置 \mathbf{r}_c	顕在化する視線の位置

注意要素は、視線として顕在化することはないが、視線位置の決定に影響を及ぼす仮想的な要素である。顕在化されないという点では視覚的注意研究における潜在的注意[9]に類似する。

提案した視覚モデルによれば、画面上において、固有の重みをもつ複数の注意要素を一つの物体と見たときに、一点でちょうど釣り合うところに顕在化した視線が現れる、と解釈することができる。

4. 修正視覚モデルの実装

4.1. 顕著性マップを注意要素位置決定のための確率分布として使用するシミュレーション

3章で提案した修正視覚モデルにおいて、1枚の提示画像の中に注意要素が同時に n 個存在すると仮定する。この仮定のもとで、視線分布を以下のアルゴリズムでシミュレーションにより予測した。

(a) 顕著性マップを求める。

最適な顕著性マップを導出する方法は確立されていないが、理想的な重みによる線形加算で統合された顕著性マップが得られるものとする。顕著性マップの導出には[10]を用いた。ただし、いくら顕著性が低い部分でも視線が分布する可能性は0ではないことから、顕著性マップの各画素の値は最低でも0より大きい値を持つものとする。

(b) n個の注意要素の位置を決定する。

顕著性マップが示す確率分布関数に従う乱数を発生させて、提示画像内の2次元座標平面上の点をn個求める。この点を注意要素の位置とする。つまり顕著性マップが高い値を示す位置には、注意要素が存在する可能性が高い。ここでは近傍にある注意要素は同一の要素であるとみなして、視野角の5°以内に2個以上の注意要素を存在させない。

(c) 各注意要素について、視線決定に及ぼす重みを決める。目立つ部分に常に視線が向かうとは限らない、との考えから重みは一樣乱数により求めた。

(d) 式(1)から r_c を求めて、顕在化する視線位置とする。

(e) 予測視線分布のマップ上で、(d)で求めた位置の画素に1を加算した後、画素値の総和が一定となるように正規化する。

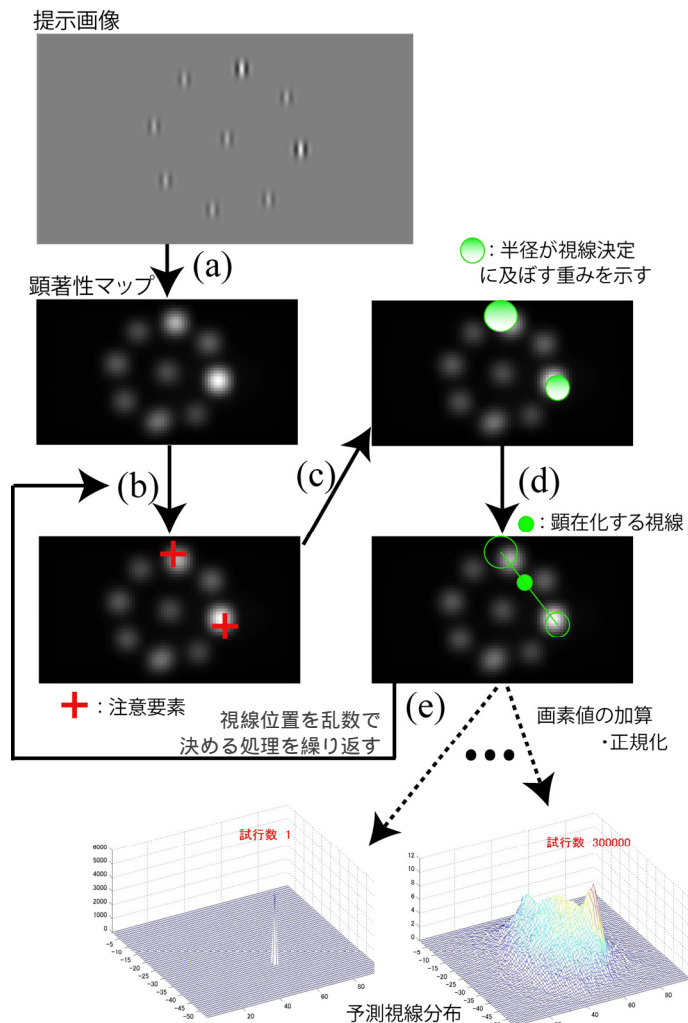


図 4-1. 注意要素数 n = 2 のときの予測視線分布

(b)~(e)の処理を繰り返して視線分布を求める流れを図 4-1.に示す(試行数 30 万回)。(b)~(d)の処理は図 2-2.の顕著性マップモデルにおける WTA を置き換えるモジュールであり、図 4-1.は全体として顕著性マップモデルを修正する。

4.2. 注意要素数を変化させた予測視線分布

提示画像に図 4-1.と同じものを用いて、n を変化させてシミュレーションした結果を図 4-2.に示す(マップサイズ 96×54 に対して試行数 100 万回)。

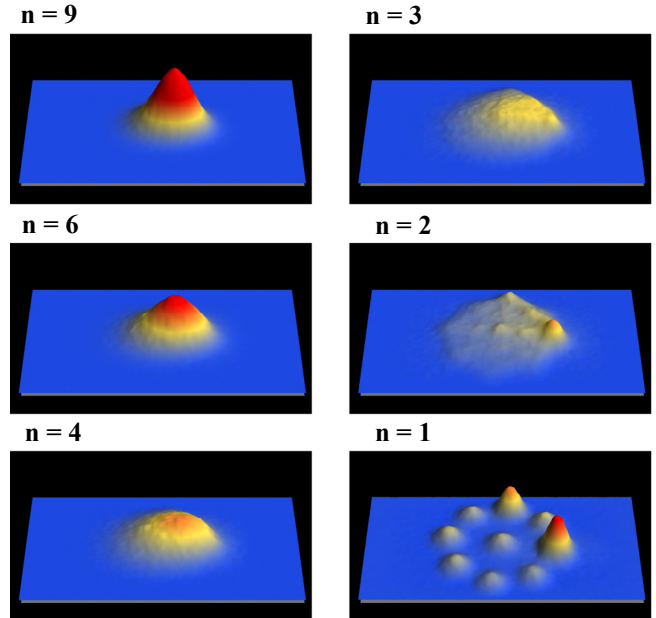


図 4-2. 注意要素数 n を変えたときの予測視線分布

提案した視覚モデルに基づく視線分布のシミュレーション結果は、十分な回数試行した場合、注意要素数 n = 1 では予測視線分布と顕著性マップは同等となることを示す。また、n が大きいほど予測視線分布は画面中央に集中する。

5. 考察

5.1. グローバル効果の再現

文献[6]が示す実験結果のグローバル効果は、WTA を用いた顕著性マップモデルでは再現できないが、提案した視覚モデルによる予測視線分布では再現する(図 5-1.)。

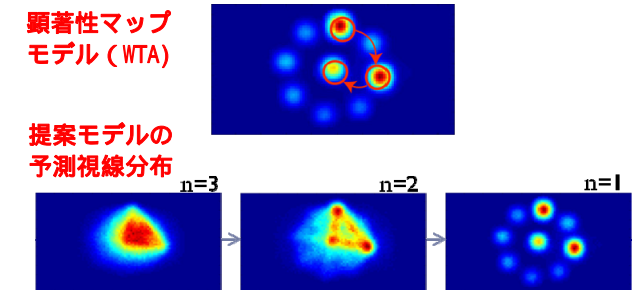


図 5-1. グローバル効果の再現

5.2. センターバイアスの再現

注意要素数 n を大きくすることでセンターバイアスの影響が大きい視線分布を予測できるとしたら、画面隅に顕著性が高い対象が集中しているとしても、n を大きくすると、予測視線分布のピークは画面隅から離れて画面中央に近づ

くと考えられる。図 5-2.のシミュレーション例は、この仮説に矛盾しない結果を示す。

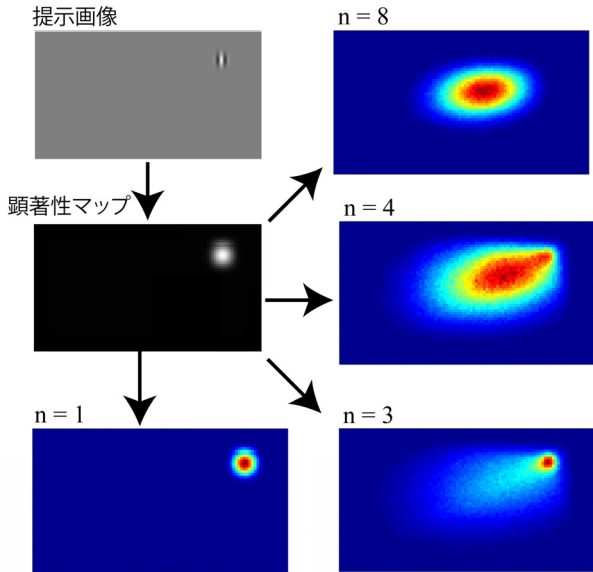


図 5-2. センターバイアスの再現

顕在化することがない仮想的な注意要素の数 n が推定できない場合には、1 枚の提示画像に対して複数枚の予測視線分布を導出する必要がある。しかしながら、センターバイアスを再現するシミュレーション結果は、注意要素数 n とセンターバイアスの影響による視線分布とを関連づけて、注意要素数 n を推定することができる可能性を示唆する。

ここで仮に注意要素数を推定できたとして、ある観測者から計測した視線と予測した視線の一致度を求めるとする。このとき注意要素数が変化すると一致度も変化する(図 5-3.)。このことから、計測した視線と予測視線の一致度を分析する際には、注意要素数 n を一致度以外の基準で推定した上で評価することが重要だと考えられる。

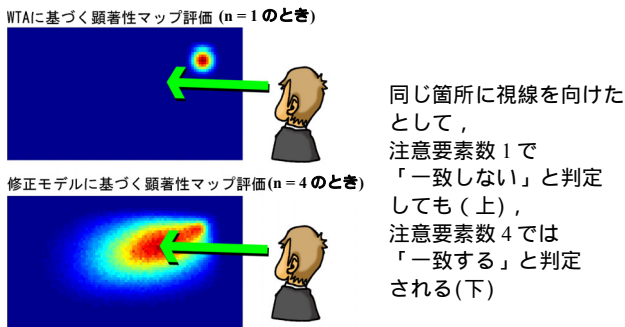


図 5-3. 正しい注意要素数推定の重要性

6. まとめ

ボトムアップ要因を拡張した視覚モデルを提案し、これに基づいて視線分布を予測した。その結果、従来の顕著性マップモデルでは再現できなかったグローバル効果や、センターバイアスの影響を再現する可能性を示した。

従来、顕著性マップと計測した視線とを直接比較することによって、顕著性マップ自体の正当性の評価を行ってきた[4][5][7]。しかしながら、センターバイアスの影響力を示す先行研究[4]と本報告が示す結果は、センターバイアスの影響を正しく推定した上で視線と比較することによ

て有効な評価となる可能性を示唆する(図 6.)。特に時々刻々と変化するテレビ映像の場合はセンターバイアスの影響は無視できないほど大きく、センターバイアス推定の重要性は大きいと考えられる。

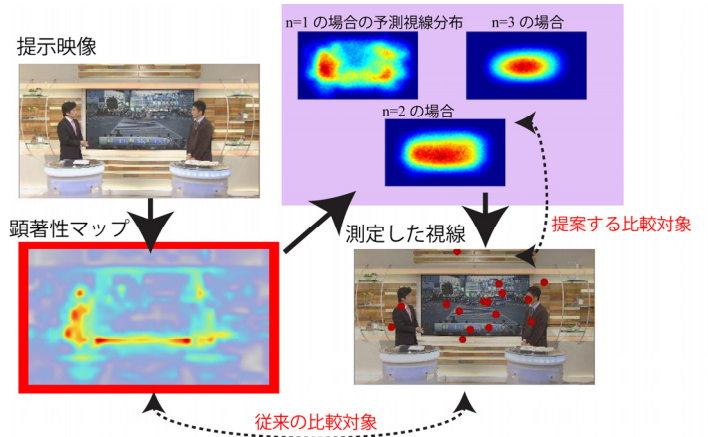


図 6. 顕著性マップの有効な評価法の提案

本報告では顕著性マップを用いた視線のボトムアップ要因推定のために、注意要素数という仮想的な要素の個数を仮定した。今後は注意要素数を実数に拡張して、より詳細に視線分布を予測する。さらには拡張したモデルが、ボトムアップ要因に由来する視覚的注意のメカニズムを解明する手掛かりとなることを期待している。

参考文献

- [1] L.Itti, C.Koch,"Computational modeling of visual attention",Nature Reviews Neuroscience,2(3),pp.194-203(mar.2001)
- [2] C.Koch, S.Ullman,"Shifts in selective visual attention :towards the underlying neural circuitry",Hum Neurobiol,Vol.4,pp.219-227(1985).
- [3] 原口健,岡嶋克典, "視覚探索における誘目性の定量化", the journal of the VISION society of japan,vol.23,No.1,pp.1-18, 2011
- [4] O. L. Meur,P. L. Callet, and D.Barba,,"Predicting visual fixations on video based on low-level visual features", Vision Research,Vol.47,issue 19,pp.2483-2498 (Sep.2007).
- [5] R.Carmi, L.Itti, "Causal Saliency Effects During Natural Vision",Proceedings of the 2006 symposium on Eye tracking reserch applications ETRA 06,Vol.1,issue March,ACM Press, pp.11 (2006).
- [6] J.M.Findlay, "Saccade Target Selection During Visual Search",Vision Research,Vol.37,issue 5, pp.617-631 (Mar.1997).
- [7] 上向俊晃,小峯一晃,森田寿哉,"動画コンテンツにおける注視点マップと顕著性マップとの関係性に関する考察",情報科学技術フォーラム講演論文集,Vol.8(3),pp.37-42 (Aug.2009).
- [8] 小池耕彦,斎木潤,"確率的誘導探索モデルによる視覚的注意の移動メカニズムの研究", Technical Report on Attention and Cognition,No.23 (2003)
- [9] J.M.フィンドレイ,I.D.ギルクリスト,"アクティブ・ビジョン眼球運動の心理・神経科学". 北大路書房, 2006
- [10] J.Harel, C.Koch, and P.Perona, "Graph-Based Visual Saliency", Proceedings of Neural Information Processing Systems (NIPS)(2006)