

## 視線追跡と Saliency Map による視覚的注意の解析

江川陽<sup>†</sup> 白山晋<sup>††</sup>

<sup>†</sup> 東京大学工学系研究科環境海洋工学専攻 〒277-8568 千葉県柏市柏の葉 5-1-5

<sup>‡</sup> 東京大学人工物工学研究センター 〒277-8568 千葉県柏市柏の葉 5-1-5

E-mail: <sup>†</sup> egawa@race.u-tokyo.ac.jp, <sup>‡</sup> sirayama@race.u-tokyo.ac.jp

あらまし 本稿では、視覚情報の視認性を向上させるために、視覚的注意の高度な解析手法を提案する。はじめに、視覚的情報に対するトップダウンのプロセスを、視線分析で測定し、それを可視化する。次に、視覚的注意のボトムアップのプロセスのモデルである Saliency Map モデルを視覚情報に適用する。そして、これらの結果を比較することで、トップダウンのプロセスとボトムアップのプロセスの関係性を分析する。分析結果を利用して、視覚情報における重要領域を抽出し、その領域と見方を提示することによって視認性の向上を図る。

## Analysis of Visual Attention using Eye-tracking and Saliency Map

AKIRA EGAWA<sup>†</sup> SUSUMU SHIRAYAMA<sup>††</sup>

<sup>†</sup> Department of Environmental and Ocean Engineering, School of Engineering, The University of Tokyo 5-1-5

Kashiwanoha, Kashiwa-shi, Chiba, 277-8568 Japan

<sup>‡</sup> Research into Artifacts, Center for Engineering, The University of Tokyo 5-1-5 Kashiwanoha, Kashiwa-shi, Chiba, 277-8568 Japan

E-mail: <sup>†</sup> egawa@race.u-tokyo.ac.jp, <sup>‡</sup> sirayama@race.u-tokyo.ac.jp

Abstract We propose a method for advanced analysis of visual attention to improve visibility of visual information. First, gaze to the visual information as top-down process of visual attention is visualized. Secondly, saliency map model as bottom-up process of visual attention is applied to the visual information. Finally, we analyze relations between top-down process and bottom-up process of visual attention by drawing a comparison between visualized gaze and saliency map. We extract important area on the visual information using results of analysis. We show a method to improve visibility of the visual information by presenting the important area and way of gaze.

### 1. はじめに

近年のデータ取得技術やシミュレーション技術の発展により、膨大なデータが生じ、蓄積されている。そのようなデータを効果的、効率的に処理するために、視覚情報化が行われることが多く、中でも可視化の役割が重要になっている。この可視化結果を効果的に処理するためには、視覚情報の視認性を向上させることが必要となる。視覚情報の視認性の向上のためには、見方や情報自体の特徴を考える必要があり、主として、

- 「見る」という行為の背後にあるルールを抽出する
  - 視覚情報自体に注意を集めやすい特徴領域が存在する
  - 学習や経験によって決まる見方が存在する
- の観点から研究が行われている。a)における「見る」という行為は、人間の経験や能力、または視覚情報の内容や与えられた課題の内容に基づく、能動的なトップダウンのプロセスの一つである<sup>1,2)</sup>。a)の代表である視線分析は、このトップダウンのプロセスにおける注視点の変化を分析する方法である。b)は視覚的な刺激により引き起こされる受動的な注意であり、視覚情報に大きく依存する。これはボトムアップのプロセスと呼べ、Saliency map による分析<sup>3)</sup>といった固有の手法がある。c)には、マンガのような慣習的な見方が存在するものを視線追跡によって分析する研究がある<sup>4)</sup>。

また、これらすべてを包括するものに視覚的注意に関する研究がある。視線追跡を利用して人間が視覚情報のどのような部分に注意(注視)を向けるかなどの、視覚的注意を解析する研究などがある<sup>5)</sup>。本研究では、この視覚的注意に着目して、視覚情報の

見るべき部分と見方を限定することによって視認性を向上させ、視覚情報の理解を支援することを考える。

このために、トップダウンのプロセスとボトムアップのプロセスのモデルを組み合わせる。既存研究では、ボトムアップのプロセスに Saliency map を生成する数理モデルを用い、トップダウンのプロセスには視線追跡などの実験的手法を組み合わせたものが多い<sup>6)</sup>。しかしながら、トップダウン側の視線追跡の数理モデル化は難しく、補助的に使われているのが現状である。この一つの理由は、Saliency map は領域を示すが、注視点は点を示すことから生じる。注視点を領域に拡張するような処理が施されるが、Saliency map と整合性のある処理となっていない。これが第一の課題となる。また、Saliency map は静的な表現のため、視線追跡によって得られる注視点の時間的な移動情報が活用されないことも問題である。これが第二の課題である。

本稿では、数理モデル化の難しい視線追跡に Saliency map を加えることで、視覚的注意の解析精度の向上を試みる。また、提案手法によってどのように視覚情報の視認性が向上するかを考察する。具体的には、Saliency map を利用した対象領域の分割を行い、視線追跡を利用し、その領域での注目度を算出し、第一の問題を解決する。第二の問題は、Saliency map での形式的な特徴領域間の移動情報と注視点の移動情報からネットワークを生成し、ネットワークの比較によってその解決を試みる。このように視覚情報における重要領域を抽出し、その領域と見方を提示することによって視認性の向上を図る。なお、本稿では視覚情報として静止画像を扱う。また、特に断らない限りはc)については考慮しない。

## 2. 提案手法

### 2.1 注視点の計測

はじめに、視線追跡を利用して、視覚情報に対する被験者の注視点の時系列の移動データを取得する。注視点計測の対象領域は、視覚情報の左上の座標を $(x_0, y_0)$ 、右下を $(x_n, y_n)$ とする矩形領域とする。注視点はこの矩形領域内の点で表される。ある注視点を $a_m$ とし、 $a_m(m, t_m, x_m, y_m)$ と定義する。ここで、 $m$ は注視点の指標、 $t_m$ は測定された時刻、 $(x_m, y_m)$ は、対象領域における注視点の座標である。

視線追跡は、ナックイメーজテクノロジーのEMR-8<sup>9)</sup>を利用して行う。視線追跡の結果は、指標、時刻、座標からなる移動軌跡データとして表すことができる。

### 2.2 注目度マップの作成

対象となる視覚情報を機械的に格子状に分割する。 $x$ 方向を $I$ 分割、 $y$ 方向を $J$ 分割する(図1)。分割した小領域をセルと呼ぶことにする。セルの幅を $\Delta w$ 、高さを $\Delta h$ とするとその大きさは、

$$\Delta w = \frac{(x_s - x_e)}{I}, \Delta h = \frac{(y_s - y_e)}{J} \quad (1)$$

となる。次に、注視点をサンプリングするための空間を用意する(以降、サンプリング空間と呼ぶ)。全体もセルも対象領域のものと同じ大きさとする。また、サンプリング空間の格子点を $(i, j)$ で表す。ここで、 $i=1, \dots, I, j=1, \dots, J$ である。セルを $c_j$ で識別する( $i=1, \dots, I, j=1, \dots, J$ )。サンプリング空間上のセルは注目度を算出する最小の単位である。

次に、注視点の移動軌跡データを用いて、各セルに存在する注視点の数を算出する。一定時間に得られた注視点の数を $M$ とする。移動軌跡データは、 $a_1, a_2, \dots, a_m, \dots, a_M$ である。例えば、注視点 $a_m$ がセル $c_j$ 内に存在する場合、

$$\begin{aligned} p_s + i\Delta w \leq x_m \leq p_s + (i+1)\Delta w \\ q_s + j\Delta h \leq y_m \leq q_s + (j+1)\Delta h \end{aligned} \quad (2)$$

が成り立つ。

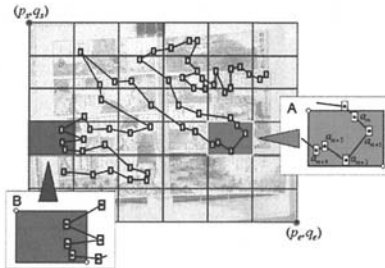


図1 注視されたセルの計算

図1右のAに示すように、セルに対して連続した注視点が複数入っている場合、そのセルが注視されたとする。左下のBのように連続していない場合、そのセルは注視されたとみなさない。本稿では、Aのようなセルに対する注目度を、セルの中にある注視点の停留時間の和として定義する。ただし、注目度は、セルの大きさに依存する。セルの大きさに関しては、後述するSaliency mapの一点が代表している入力画像の領域の大きさと同じものにする。計算されたセルの注目度の値に従って当該セルを色付けること

で対象領域に対する被験者の注目度合いを図2のように可視化する。この可視化図を注目度マップと呼ぶことにする。素材は文献8からとった。

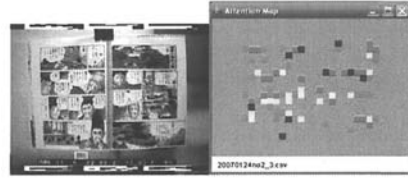


図2 注目度マップ

### 2.3 注視点チャートの作成

注目度マップからは分析できない、どのような順番で対象領域を見ているのかという情報を表すために、注視点位置の時間変化の特徴を示す方法を提案する。

ある基準に基づいて対象領域を小領域に分割する。ここでは説明のため矩形領域を用いて分割する。その小領域を $R_k$ によって識別する。次に、移動軌跡データ： $a_1, a_2, \dots, a_m, \dots, a_M$ を、注視点が存在する小領域の指標 $R_k$ によって符号化する。

横軸に経過時間を、縦軸に指標 $R_k$ をとり、符号化された移動軌跡データを表すと、図3の右図に示すようなチャートができる。これを注視点チャートと呼ぶことにする。

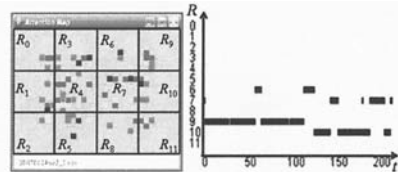


図3 注目度マップ(左)と注視点チャート(右)

### 2.4 注視点ネットワーク

注目度マップと注視点チャートを合わせて、注視点の大まかな移動パターンをネットワークとして求める。はじめに、小領域 $R_k$ 中の注目度の重心を求める。次に、注視点チャートを利用して、小領域間の注目点の移動を重心間の矢印で示す。図4に図3の注目度マップと注目点チャートに基づく、注目点の移動パターンを示す。これを注視点ネットワークと呼ぶことにする。

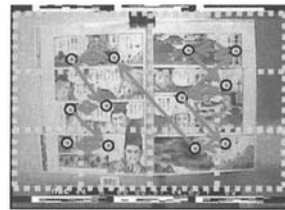


図4 注視点ネットワーク

### 2.5 Saliency mapの作成

Saliency mapモデル<sup>2)</sup>は、視覚的注意におけるボトムアップのプロセスのモデルの一つである。入力画像における不連続部分や目立った部分など、人間が注意を向けそうな特徴的な部分を抽出することができる。Ittiら<sup>2)</sup>に従って作成する。図5に作成した

Saliency map の例を示す。

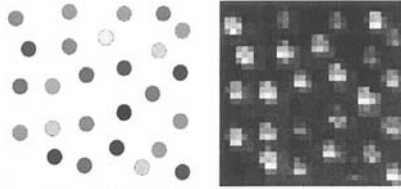


図5 視覚情報 (左) と Saliency map (右)

本稿では, Saliency map から特に saliency が高い領域のみを抽出し, クラスタリングを行う。クラスタリングされた Saliency map を, ここでは Saliency Cluster Map (以下 SCM) と呼ぶ。また, 総クラスタ数を  $M$  とする。

最後に SCM を基準として注意を集めやすい部分を Focus Of Attention (FOA) として抽出する。ここでは Iti ら<sup>2)</sup>を参考に saliency が高いクラスタ順に FOA が移動するとする。図 6 には, 図 5 の Saliency map から作成した SCM と, この SCM から推定した FOA の移動を示す。SCM ではクラスタを識別するためにクラスタを色付けして示している。また, 右図には元画像を重ねている。

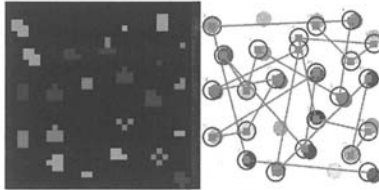


図6 SCM (左) と FOA の移動 (右)

## 2.6 視線追跡と Saliency Map の統合

トップダウンのプロセスを計測する視線追跡と, ボトムアップのプロセスのモデルである Saliency map を統合する。

第 1 に, 3.3 節で述べた注目度マップを本稿で用いる Saliency map の一点の大きさ(16x16 ピクセル)に合わせて生成する。

第 2 に, Saliency map から SCM を作成する。

第 3 に, SCM を利用して 3.4 節で述べた注視点チャートを生成する。SCM が対象領域のすべてを覆うわけではないので, SCM とそれ以外の小領域 (地の部分) を  $R_k$  に分ける。  $k=1, \dots, M$  が SCM の領域,  $k=0$  がそれ以外の領域である。SCM 以外の領域が大きい場合, 格子状の矩形領域を併用する。各小領域で SCM 以外の部分に識別番号を付ける。

第 4 に, 移動軌跡データ:  $a_1, a_2, \dots, a_m, \dots, a_M$  を, 注視点が存在する小領域の指標  $R_k$  によって符号化する。3.3 節で述べた方法に従って注視点チャートを作成し, 3.4 節で述べた方法に従って注視点ネットワークを生成する。

## 2.7 視覚的注意の解析のためのフレームワーク

はじめに, 空間的な観点から, 視線追跡結果と Saliency map モデルの関係を, 以下の 4 つのパターンに整理する。

- A. saliency が高く, かつ注視された部分
- B. saliency が高いが注視されなかった部分
- C. saliency は低いが注視された部分
- D. saliency が低く, かつ注視されない部分

前節で述べた Saliency map を取り入れた注視点ネットワークは, 3.5 節の SMC と FOA から作成したネットワークと多くの共通のノード点を持つ。両者のネットワークを 4 つのパターンから比較, 分析することによって視覚的注意の解析を行う。

この際, 空間的な視点のみならず, 時間的な観点から考慮する必要がある。注視点ネットワークにおける注視点の時系列変化と, FOA の時系列変化が同じであるか, あるいは異なっているかを部分的に分析する。

## 3. 視覚的注意の解析

対象とした視覚情報は, 情報の可視化から得られた図 7 で示すネットワーク図とする。赤棒と S, G の文字は説明のために付したもので被験者には示していない。

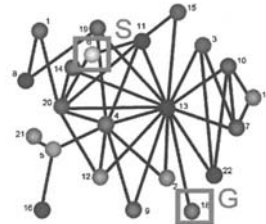


図7 対象とした視覚情報

視線追跡実験は, 二名の被験者で行った。実験をする際に, 被験者に対して課題を与えた。今回は, 図 7 中の S と G のノードをスタートとゴールのノードとして, その間の経路をさがすものである。課題の指示は, 視覚情報を被験者に提示する前に行った。よってこの設定においては, 被験者は視線追跡実験の初めからスタートとゴールのノードの場所と, そのノード間の経路を探索するというトップダウンのプロセスを働かせると考えられる。

はじめに, Saliency map の結果を示す。図 8 には, 図 7 の視覚情報を素材とした Saliency map と SCM を示す。ノードだけでなくリンクについても特徴領域として抽出されていることがわかる。従って, Saliency map モデルによれば, ネットワークのリンクに対しても注意が向けられることが推定される。

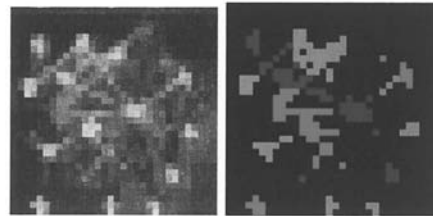


図8 Saliency map (左) と SCM (右)

図 9 に, 図 8 で示された SCM から作成した FOA の移動推定図を示す。緑の円の部分が, 初めの FOA であり, その後赤い線に沿って FOA が移動すると推定される。基本的にはノードとその近傍が FOA として抽出されるが, リンク部も FOA として推定されている。

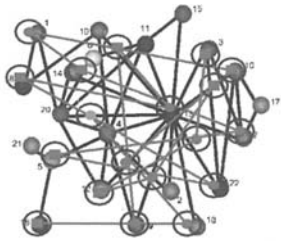


図9 FOAの移動推定図

次に視線追跡実験の結果を示す。

ある被験者に対する計測結果から得られた注目度マップを図10に、注視点チャートを図11に示す。注視点チャートは図8右のSMCに基づき、SMC以外の領域を0として作成した。また、図11に基づき、注視点ネットワークを作成する(図12)。

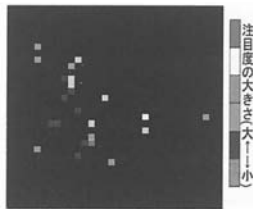


図10 注目度マップ

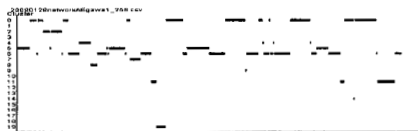


図11 注視点チャート

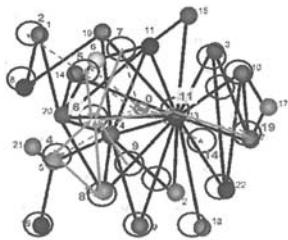


図12 注視点ネットワーク

視線追跡実験の結果と、Saliency map モデルを比較すると、注視領域と Saliency が高い領域が一致しないことがわかる(図8左図と図10)。Saliency が高い領域は図全体に広がっているが、注目度マップからわかる注目度の高い領域はスタートとゴールの間に分布している。課題に関係ない領域は、課題に依存するトップダウンのプロセスにおいては、注目されづらいことを示している。

また、図9と図12を比較すると、時系列での注視点の移動とFOAの移動が大きく異なることがわかる。前述のように、FOAの移動から形成されるネットワークは道であり、ノードとリンク

の重複はない。一方、注視点ネットワークにおいては、過去に注視した領域をもう一度注視するという動きが頻繁に見られる。特に、スタート及びその周辺部においてこの傾向が強い。

Saliency map に現れるボトムアップのプロセスは、画像が表示された瞬間からスタートとなるノードを注視するまでの間に現れと考えられる。注視点チャートと注視点ネットワークからは画像全体を探索するのではなく、あるまとまった領域を注視して、比較的早い段階でスタートとなるノードを注視している。

このように、提案手法を利用することで、ボトムアップのプロセスとトップダウンのプロセスの双方に対して、視覚情報における同じ領域での注目度の差、および注視点の移動パターンの差を顕在化できる。これらの情報に基づき、元画像において視覚的に重要ではない領域を隠蔽し、注釈付けなどによって見方を付与することで視認性が向上するものと考えられる。

#### 4. まとめ

本稿では、視覚情報の選択機能である視覚的注意に着目し、視覚情報の見るべき部分と見方を限定して示すことによって視覚情報の視認性を向上させ、視覚情報の理解を支援するという方法を提案した。具体的には、トップダウンのプロセスである視線追跡に、ボトムアップのプロセスの Saliency map の要素を加え、視線追跡側から両者を統合する方法である。また、ボトムアップのプロセスとトップダウンのプロセスを統合した形で視覚情報の見るべき部分と見られ方を示せる方法でもある。

提案手法を用いた被験者実験により、視線追跡と Saliency map のそれぞれ単独では発見が難しかった、視覚的注意においてトップダウンのプロセスのみが働いた箇所、ボトムアップのプロセスのみの箇所、双方が強め合う箇所を分離できる可能性を示した。

#### 参考文献

- 1) Igarashi,H., Suzuki, S., Sugita,T., Kurisu, M. and Kakiura, M.: Extraction of Visual Attention with Gaze Duration and Saliency Map, Proceedings of the 2006 IEEE International Conference on Control Applications, pp.562-567, (2006).
- 2) Itti, L., Koch, C., Niebur, E.,: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).
- 3) 金野武司, 橋本敬:共同注視における意図理解の構成論的モデル, MPS シンポジウム 2005 「計算科学シンポジウム」講演論文集, pp.179-186,(2005).
- 4) 石井拓, 井垣竹晴, 倉田敦子, 大森貴秀, 増田直衛:マンガにおける視線誘導要因の推定と眼球運動の制御一心の解明に向けての統合的方法論構築一, 慶應義塾大学 21 世紀 COE プログラム, 平成 15 年度成果報告書, (2004).
- 5) 三浦利章, 原田悦子:編著:事故と安全の心理学:リスクとヒューマンエラー, 東京大学出版, pp.129-157, (2007) .
- 6) 伊藤隆文, 富樫裕基, 斎藤元生, 水谷健太郎, 大森隆司, 山内康一郎:自動車運転者の運転プランを考慮した視線移動決定過程のモデル化, デンソーテクニカルレビュー, vol.12, No.1, pp.120-129, (2007).
- 7) <http://www.cycemark.jp/lineup/EMR-8/EMR-8b.html>
- 8) 入間田宣夫監修, 森藤よしひろ漫画:集英社版・学習漫画日本の歴史7鎌倉幕府の成立, 集英社, (1998).