

運転手注視点予測用動的顕著性マップを用いた
対象物体に対する誘目性の定量評価法の検討

Quantification Method for Visual Attention of Target Objects
via Dynamic Saliency Maps for Predicting Driver's Gazing Points

稲川 巧巳[†] 水野 倫太郎[‡] 中田 洋平^{†,‡}
Takumi Inagawa Rintaro Mizuno Yohei Nakada

1. はじめに

現在、自動車交通事故は人間のミスに起因するものが多く、そのような事故を減らすべく様々な取り組みが行われている。完全自動運転技術が家用車に搭載されるようになるためには、まだ多くの技術的・法的な困難が存在し、先んじて安全運転を促す運転支援技術が必要とされている。また、そのような運転支援技術の一つにフロントガラスに注意喚起を表示する技術がある。その発展に寄与することを最終的な目標として、著者らのグループでは、これまで車載カメラ動画画像から、運転手の注視点を予測するための動的顕著性マップの構築法の研究[1][2]を進めてきた。しかし、実際に前述のような運転支援技術で利用するためには、運転手の注視点を予測するだけでなく、歩行者や対向車などの特定の物体に対して、運転手がどの程度目が行きやすいかを定量評価する手段が必要となる。そこで、本稿では、そのような定量評価を可能とする手法を検討する。具体的には、本定量評価手法では、正規化後の動的顕著性マップの対象物体の映る画像領域内における積分値を算出し、それを定量評価値とする。また、著者らのグループで試作してきた車載カメラ動画画像シミュレータ[3]により再現した動画画像を用いて、本手法の初期的検証を行う。

2. 用いる動的顕著性マップ構築法の概要

本稿では、動的顕著性マップの算出に、前述した運転手注視点予測用の動的顕著性マップ構築法[1][2]を用いる。この構築法では車載カメラ画像から10種類の特徴マップを算出し、特徴マップを組み合わせることで動的顕著性マップを構築する。より具体的には、4つの静的な特徴マップと、それら静的特徴マップと入力画像から算出したオペイカルフローの大きさに基づく5つの動的な特徴マップ、一様特徴マップの計10個の特徴マップを用いる。また、それらの組み合わせ比率を決定するパラメータをEMアルゴリズムで学習する。なお、後述の検証実験では、文献[3]で用いているのと同様の構成とパラメータにより、動的顕著性マップを算出している。

3. 誘目性の定量評価法の検討

特定物体への誘目性を定量評価値として、その物体への目の行きやすさを予測する確率値として算出することは、自然な発想と考えられる。また、適切な手段で算出された顕著性マップを正規化したものは、しばしば各画素に対する人の目の行きやすさを予測する離散化された2次元確率分布と見做されてきている。更に、評価したい領域を D とすると、その領域 D の確率値は、2次元確率分布 $p(x, y)$ に



図1 使用する人物モデル (左から A, B, C, D とする)

関する以下のような積分式で表される。

$$P(D) = \iint_D p(x, y) dx dy \quad (1)$$

以上のことを踏まれば、正規化された顕著性マップの画素値 $p_{ij} \in [0, 1]$ と領域 D を指定するマスク画像の画素値 $m_{ij} \in \{0, 1\}$ が与えられれば、確率値 $P(D)$ は次式のように近似できる。

$$P(D) \approx \sum_{i=1}^N \sum_{j=1}^M p_{ij} m_{ij} \quad (2)$$

ここで、自然数 i と j は各々水平画素番号と垂直画素番号であり、 N と M は水平画素数と垂直画素数を表す。

すなわち、本稿で検討する定量評価法では、別途、歩行者や対向車などの対象物体の画像内での領域を指定するマスク画像を算出し、それと動的顕著性マップ構築法[1][2]による正規化後動的顕著性マップを用いて、対象物体への誘目性の定量評価値として確率値(2)を算出する。なお、これと類似する定量評価法は、顕著性マップを利用して特定物体への誘目性を高める画像加工を施すような手法[4][5]などでも用いられている。ただし、著者らの調査の限り、本稿のように、運転手の視線予測に関する顕著性マップに対して用いた例は見当たらない。

4. 初期的検証実験

本検証実験では著者らの研究室で試作を進める車載カメラ動画画像シミュレータ[3]を用いる。なお、文献[3]にも示される通り、本シミュレータは、Unity [6]を用いて作成されており、様々な状況を車載カメラ目線から撮影し再現できるものである。また、よりリアルに運転状況を再現するため、株式会社ゼンリンが公開している3D都市モデル[7]や、RenderPeople [8]の3D人体モデルとアニメーションを使用している。本検証実験では、このようなシミュレータにより、異なる人物モデルで特定運転状況の車載カメラ動画を再現し、各場面での評価値を算出することで定量評価法の動作確認を行う。より具体的には、本運転状況では、自車が直進し、図1に示されるように、A. 女性、B. ベビーカーを押す男性、C. ジョギングをしている男性、D. 少年の

[†] 明治大学 総合数理学部
School of Interdisciplinary Mathematical Sciences, Meiji University

[‡] 明治大学大学院 先端数理科学研究科
Graduate School of Advanced Mathematical Sciences, Meiji University

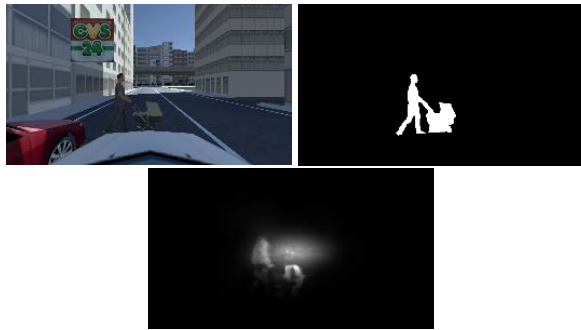


図2 元画像 (上段左), マスク画像 (上段右), 動的顕著性マップ (下段中央) の例

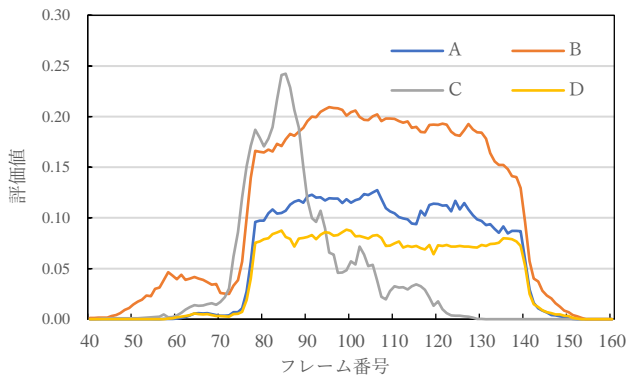


図3 各人物モデルにおける評価値の推移

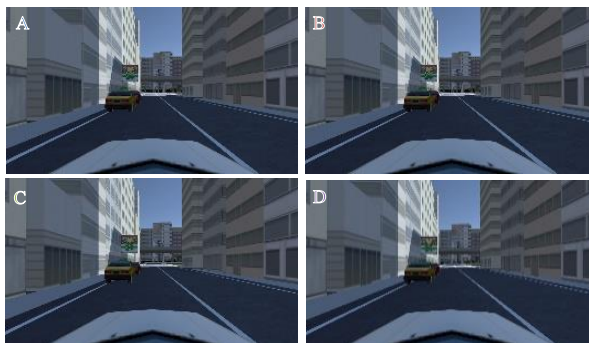


図4 各人物モデルにおける40フレーム目の画像

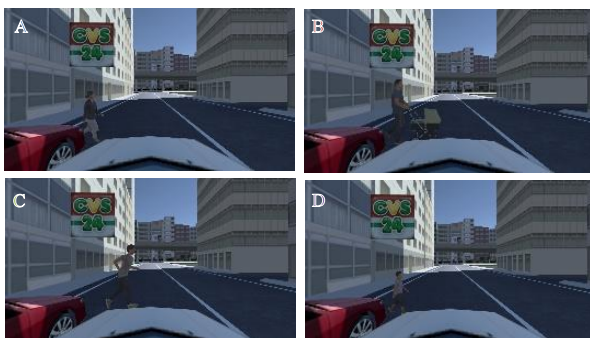


図5 各人物モデルにおける76フレーム目の画像

4種類の人物モデルのいずれかが飛び出してくる。そして、自車は、人物モデルの直前で停車してから再度発進する。その間の120フレームの評価値(2)を算出する。なお、評価値(2)を算出するにあたり、対象の範囲を示すマスク画像の

作成は、人物モデルがない同様の運転状況を再現した動画との背景差分法を用いて作成している。また生成されるマスク画像には、ノイズを除去するべく、モルフォロジー変換を行っている。図2は、人物モデルBを用いた運転状況におけるあるフレームでの元画像、マスク画像、動的顕著性マップを表している。なお、このときの評価値(2)は約0.1116と算出されている。

図3は、実際に算出した各人物モデルの場合の評価値の推移を表している。ただし、人物モデルが登場する直前の40フレーム目からの推移を示している。また、図4と図5は、各々、40フレーム目と76フレーム目の画像を示している。図2を見ると、他の人物モデルと比べ、Cは他のモデルに比べ移動速度が速いため、評価値の変動が最も大きく、評価値がピークに達するのも早いことが分かる。Aは、人物の前にベビーカーがあるため、最も反応が早く、45フレーム目辺りから増加が始まっているが、76フレーム目辺りからは、A、B、Dには速度に大きな差がないため、対象物の画像中の面積の大きさに対応したような変動を見せている。なお、モデル別の最も評価値が高くなっている場面は、それぞれフロント中央部に到達しているときであり、最も評価値に変動が大きかったのはフロント中心部から一定の距離に到達したフレームであった。これは、動的顕著性マップを構築する際に用いているセンターバイアス処理が大きく影響していると考えられる。以上のことから、目標物の大きさ、速さ、人間の視覚的特性であるセンターバイアスも踏まえ、妥当と思われる結果を示していると言える。

5. 今後の課題

今後は、本定量評価法で得られた評価値が、実際の目の行きやすさに対応しているかを、アイマークカメラを用いた計測実験の結果などと比較検証していく。

謝辞

本研究を進めるにあたり、同研究室の中澤空知氏には貴重なご意見ご助言を頂きました。深謝いたします。

参考文献

- [1] S. Nakazawa, S. Ushijima, and Y. Nakada, "An EM Algorithm based Method for Constructing Dynamic Saliency Maps considering Characteristics while Driving", Proc. of the 10th International Symposium on Information and Communication Technology, pp. 314-321 (2019).
- [2] S. Nakazawa and Y. Nakada, "Improvement of Mixture-of-Experts-Type Model to Construct Dynamic Saliency Maps for Predicting Drivers' Attention", Proc. of 2020 IEEE Symposium Series on Computational Intelligence, pp. 2008-2015 (2020).
- [3] 水野 倫太郎, 中澤 空知, 中田 洋平, "車載カメラ動画シミュレータを用いた運転手注視点予測用動的顕著性マップ構築法の検証", 第20回情報科学技術フォーラム (2021).
- [4] 鈴木 菜摘, 中田 洋平, "視覚的顕著性マップを活用したエフェクト列選択手法", 画像電子学会誌, Vol. 46, No. 4, pp. 498-509 (2017).
- [5] N. Suzuki and Y. Nakada, "Effects Selection Technique for Improving Visual Attraction via Visual Saliency Map" Proc. of 2017 IEEE Symposium Series on Computational Intelligence, pp. 1030-1037 (2017).
- [6] Unity Technologies, Unity: <https://unity.com/ja>
- [7] 株式会社ゼンリン, Unity向け3D都市モデルデータ: ZENRIN City Asset Series, <https://www.zenrin.co.jp/contents/product/service/3d/asset/index.html>
- [8] Renderpeople, レンダリング用3D人体モデル, <https://renderpeople.com/jp/>