

令和 5 年 6 月 2 日現在

機関番号：13801

研究種目：基盤研究(C)（一般）

研究期間：2020～2022

課題番号：20K11828

研究課題名（和文）車載カメラ映像を対象とした注視領域の推定法の提案とデータセットの構築

研究課題名（英文）Estimation method and dataset of visual attention for dashcam images

研究代表者

大橋 剛介（Ohashi, Gosuke）

静岡大学・工学部・教授

研究者番号：80293603

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：静止画と映像（動画）に対する顕著性を比較したところ、対象が移動するシーンでは視線が対象の前方を注視しており、対象の運動が変化したシーンでは対象を注視できないことがわかった。また、注視領域は対象の移動方向に偏ること、注視領域は運動が変化する前の運動に影響されることがわかった。Carlaシミュレータと実際の画像に対して、画像変換モデルを用いて、昼間の車載カメラ映像のデータセットから夜間の車載カメラ映像のデータセットを作成した。オプティカルフローから算出したFOEを用いた注視領域推定モデルを提案し、車載カメラ映像のデータセットを用いて有効性を検証したところ、高精度に注視領域が推定できた。

研究成果の学術的意義や社会的意義

本研究では、車載カメラ映像に対して、運転者の注視領域の推定モデルを開発することとデータセットを構築することを目的としている。注視情報付きデータセットで公開されている一般的な動画（映像）を対象としたものは、車載カメラ映像における注視領域とは異なる。したがって、本研究の学術的意義は、1）静止画と映像（動画）との相違点、2）車載カメラ映像のデータセット、3）運転者の注視特性、を明らかにしたことである。本研究の社会的意義は、本研究の成果により、既存の膨大な映像データを有効に活用でき、運転者の注視領域を明らかにすることが可能になり、運転支援システムの開発に貢献できる。

研究成果の概要（英文）：The comparison of saliency maps between the static image and moving image proved that the gaze is fixated ahead on the moving target object, and the target object can not be gazed at for objects that change motion in a scene. Moreover, it was found that visual attention is biased toward the moving direction of the target object and affected by the target object's motion before the motion changes. The nighttime dataset of the dashcam image is created from the existing dataset of the daytime dashcam image using the image transformation model for the Carla simulator and real images. A visual attention model was proposed incorporating FOE from the optical flow and verified the model's effectiveness by predicting visual attention with high accuracy using the dashcam image dataset.

研究分野：画像処理

キーワード：データセット 車載カメラ映像 注視領域 顕著性マップ 深層学習 ドライビング・シミュレーター

## 1. 研究開始当初の背景

車載カメラ映像から、運転者がどこを注視するかを知ることは、運転支援技術の開発に有用である。注視領域は、視線追跡装置を用いて知ることができるが、視線追跡装置を用いずに、画像から注視領域を計算によって求める顕著性マップの研究が、Itti らによる提案 (L. Itti et. al., IEEE Trans. on Pattern Analysis and Machine Intelligence, 1998) 以来、盛んに研究されている。また、近年、画像認識分野で著しい成果をあげている深層学習を用いた手法が、静止画を対象とした顕著性マップのベンチマークにおいて、上位を占めている (SAM-ResNet : Rita Cucchiara et. al, IEEE Trans. on Image Processing, 2018, DSCLRCN : IEEE Trans. on Image Processing, 2018)。代表者は、U-net 構造に基づく深層学習モデルを提案し、顕著性マップのベンチマークである SALICON Saliency Prediction Challenge において、7つの評価指標のうちの3つ (CC、SIM、AUC) で世界最高精度を得ている。しかしながら、動画像 (映像) に対しては、2017年に最初の深層学習モデルが提案された (C. Bak, et. al., IEEE Trans. on Multimedia, 2017) が、未だ顕著性マップを高精度に推定するには至っていない。車載カメラ映像に対しても、注視情報 (Ground Truth) が付与されたデータセットは限定的であり、未だ顕著性マップを高精度に推定するには至っていない。車載カメラ映像を対象とした Ground Truth の注視情報付きデータセットおよび顕著性マップモデルは限定的であり、動画像 (映像) を対象とした Ground Truth の注視情報付きデータセットで公開されているものは、テレビ番組のゴルフなどのスポーツシーンや映画のワンシーンなどである。そのため、スポーツ映像では人の体に、映画のシーンでは顔に、顕著性マップは出力する傾向があり、車載カメラ映像における注視領域とは異なる。したがって、本研究課題の核心をなす学術的「問い」は、1) 静止画と映像 (動画像) との相違点は何か、2) 車載カメラ映像の特性は何か、3) 運転者の注視特性は何か、を明らかにし、推定モデルにいかに関与するか、である。

## 2. 研究の目的

本研究では、車載カメラ映像に対して、運転者がどこを注視しやすいかを推定できる顕著性マップ推定モデルを開発することを目的としている。さらに、顕著性マップを既存の膨大な車載カメラ映像データに適用し、データセットを構築する。

## 3. 研究の方法

### 1) 動画を対象とした視覚的顕著性マップ推定のための視線解析

画像から人間の注視領域を予測する顕著性マップ推定が盛んに研究されている。顕著性マップ推定には静止画を対象としたものと、動画を対象としたものがあるが、動画には静止画にない動きという特徴が存在するため、動画を対象とした顕著性マップ推定は困難な課題となっている。そこで、本研究では、動画を対象とした顕著性マップ推定モデルの精度向上のための知見を得るため、移動する対象が存在する動画を対象として、視線追跡装置によって計測した視線による顕著性マップと、動画を対象とした顕著性マップ推定モデルの中で、現在最も精度が高いモデルの一つである STSANet より推定された顕著性マップを比較し、解析した。

### 2) 車載カメラ映像のデータセットの構築

大規模なデータセットは、ディープラーニングモデルの性能を向上させる。車載カメラ画像データセットには、昼間のシーンが多く、夜間のシーンは少ない。そこで夜間のデータセットを増やすために、視覚的な外観や内容を保持したまま昼間の道路シーンを夜間の道路シーンに変換する画像間変換モデルを提案した。敵対的生成ネットワークの一種で、画像のスタイル変換が可能な CycleGAN を用いて、昼間の車載カメラ映像を夜間車載カメラ映像に変換する。このとき、データセットの構築を目的としたとき、視認性の向上のみではなく、ラベル情報の保持も必要となる。そこで、画像をスタイルとコンテンツに分離し、昼間の車載カメラ映像のコンテンツは保持して、スタイルのみ参照する夜間車載カメラ映像に置き換える CycleGAN に基づく手法を提案した。

### 3) 車載カメラ映像に対する注視領域の推定

運転者の注視特性は何か、を明らかにするため、視線追跡装置を用いて取得し、注視領域と FOE の関係を明らかにする。また、オプティカルフローの湧出点である FOE (Focus Of Expansion) は運転時の注視において、重要な要素とされている。そこで、本研究では車載カメラ画像における運転時の注視領域の推定精度を向上させるため、一般的な視線を推定する顕著性マップモデルにオプティカルフローの湧出点である FOE を組み込む手法を提案した。さらに、FOE に関しては、オプティカルフローのノルムに着目したロバスト性の高い手法を提案した。

## 4. 研究成果

### 1) 動画を対象とした視覚的顕著性マップ推定のための視線解析

動画を対象とした注視領域推定モデルの精度向上のために、動的顕著性と静的顕著性を計測して比較したところ、対象が大きく移動するシーンでは動的顕著性と静的顕著性が類似しないことを確認した。対象が滑らかに移動する動画の、視線追跡装置によって計測した視線により作

成した顕著性マップ（計測結果）と、STSA Net の推定による顕著性マップ（推定結果）を図 1 に示す。図 1 の (a), (b), (c) の計測結果では対象、対象の前方への注視が存在するが、推定結果では、対象と画像中央を中心に広い範囲に顕著性が現れている。対象の運動が途中で変化する動画の運動が変化した直後のフレームの、計測した視線により作成した顕著性マップ（計測結果）と STSA Net の推定による顕著性マップ（推定結果）を図 2 に、計測した視線の平均座標を図 3 に示す。背景が停止している場合 (fixed background)、背景が移動する場合 (moving background) に関わらず、計測結果では対象の移動先である左側に顕著性が出現しているのに対して、推定結果では対象と画像中央を中心に広い範囲に顕著性が出現している。このことから注視領域は対象の移動方向に偏ること、対象の運動が変化したとき、注視領域は運動が変化する前の運動に影響されることがわかった。また、現在の顕著性マップ推定モデルでは、これらの動画を適切に推定できないことがわかった。

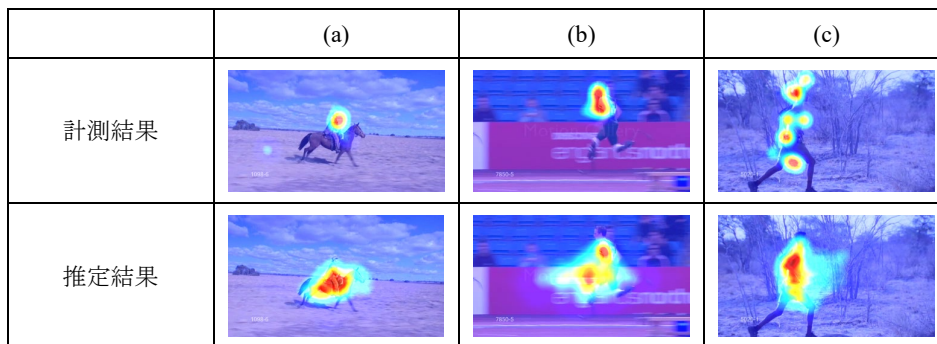


図 1 対象が滑らかに移動する動画の顕著性マップ

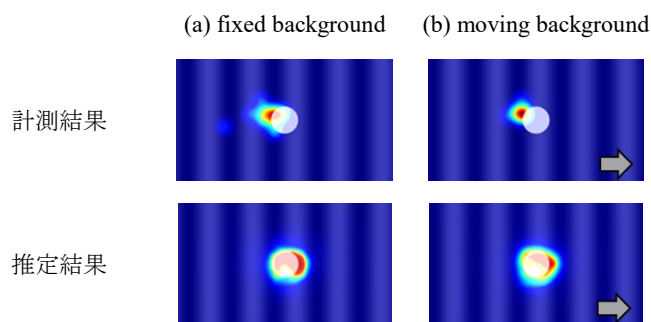


図 2 対象が左に移動してから停止する動画の顕著性マップ

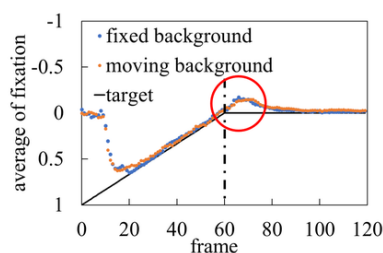
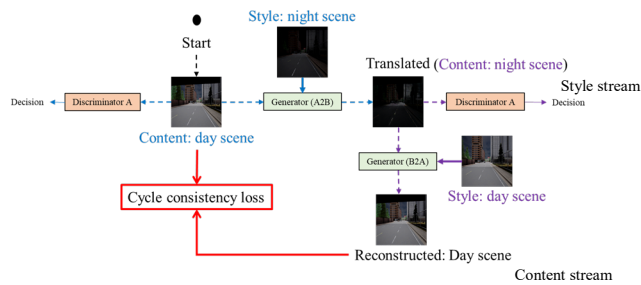


図 3 対象が左に移動してから停止する動画の視線の平均座標

## 2) 車載カメラ映像のデータセットの構築

サイクルの一貫性と内容スタイルの分離概念からなる提案モデルの概略図を図 4 に示す。ドライビングシミュレータと実画像データセットを用いて提案モデルを訓練し、シミュレータデータセット中の同じシーンの昼間と夜間の画像により、モデル性能を調べた。実験結果を図 5 に示す。提案モデルの有効性を、定性的、定量的に最新モデルと比較し、提案モデルの有効性を確認した。



(a) Cycle consistency

(b) Content-style separation in generator (A2B)

図 4 提案モデル

Example	Daytime	Nighttime	CycleGAN	TSIT	DRIT	MNUIT	Proposed
Carla (Simulator image)							
INIT (Real image)		-					

(a) Qualitative result

Model	Original (day)	Original (night)	Translated (night)	FID (↓)	
				Carla dataset	INIT dataset
-	✓	✓	×	102.8	114.7
CycleGAN	×	✓	✓	65.8	56.7
TSIT	×	✓	✓	84.6	72.9
DRIT	×	✓	✓	176.7	96.4
MUNIT	×	✓	✓	147.7	58.3
Proposed	×	✓	✓	<b>62.6</b>	<b>56.5</b>

(b) Quantitative result. (↓) indicates that the lower the score is, the better the model performance is.

図 5 実験結果

### 3) 車載カメラ映像に対する注視領域の推定

視線追跡装置を用いて取得し、注視領域と FOE の相関を算出した。表 1 に示す算出結果から FOE と注視領域の結果から高い相関を確認した。

本研究では車載カメラ画像における運転時の注視領域の推定精度を向上させるため、一般的な視線を推定する顕著性マップモデル (HD<sup>2</sup>S) にオプティカルフローの湧出点である FOE を組み込む手法を提案した。FOE をベースラインの HD<sup>2</sup>S に組み込んだ提案モデルを図 6 に示す。FOE に関しては、オプティカルフローのノルムに着目したロバスト性の高い手法を提案した。提案モデルに対する定性的評価(図 7)と定量的評価(表 2)より、ベースラインである HD<sup>2</sup>S と比較して高精度に注視領域が推定できていることを確認した。

表 1 注視領域と FOE の相関

		FOE		
		直進	左カーブ	右カーブ
注視領域	直進	<b>0.99</b>	0.55	0.34
	左カーブ	0.65	<b>0.98</b>	0.02
	右カーブ	0.45	0.03	<b>0.97</b>

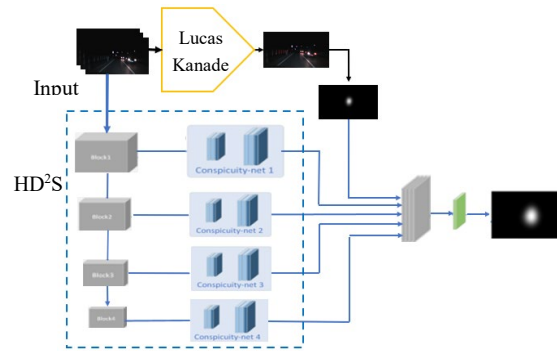


図 6 HD<sup>2</sup>S に基づく提案モデル

注視領域			
ベースモデル (HD <sup>2</sup> S)			
本手法 (HD <sup>2</sup> S +FOE)			

図 7 HD<sup>2</sup>S に基づく算出結果例

表 2 HD<sup>2</sup>S に基づく定量的評価

手法	CC (↑)	SIM (↑)	KLD (↓)
HD <sup>2</sup> S	0.585	0.499	1.22
本手法 (HD <sup>2</sup> S+FOE)	<b>0.658</b>	<b>0.558</b>	<b>0.956</b>

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件／うち国際共著 0件／うちオープンアクセス 0件）

1. 著者名 R. Sultana, G. Ohashi	4. 巻 E106-A
2. 論文標題 GAN-based Image Translation Model with Self-attention for Nighttime Dashcam Data Augmentation	5. 発行年 2023年
3. 雑誌名 The IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 池野谷玲太、大橋剛介	4. 巻 143
2. 論文標題 動画を対象とした顕著性マップ推定のための視線解析	5. 発行年 2023年
3. 雑誌名 電気学会論文誌C（電子・情報・システム部門誌）	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 R. Sultana, G. Ohashi	4. 巻 E106-D
2. 論文標題 Prediction of driver's visual attention in critical moment using optical flow	5. 発行年 2023年
3. 雑誌名 The IEICE Transactions on Information and Systems	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transinf.2022EDP7146	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計9件（うち招待講演 0件／うち国際学会 3件）

1. 発表者名 Rebeka Sultana, Yuki Hikosaka, Gosuke Ohashi
2. 発表標題 GAN Based Image-to-Image Translation Model for Nighttime Road Scene Dataset
3. 学会等名 The 29th International Display Workshops (IDW '22) (国際学会)
4. 発表年 2022年

1. 発表者名 池野谷玲太, 大橋剛介
2. 発表標題 動的顕著性の特徴に着目した視線解析
3. 学会等名 第27回知能メカトロニクスワークショップ
4. 発表年 2022年

1. 発表者名 池野谷玲太, 大橋剛介
2. 発表標題 動的顕著性マップ予測モデルのための視線解析
3. 学会等名 電気学会 電子・情報・システム部門大会
4. 発表年 2022年

1. 発表者名 丸山裕太, 大橋剛介
2. 発表標題 FOE に着目した車載カメラ画像の注視予測モデル
3. 学会等名 第28回画像センシングシンポジウム
4. 発表年 2022年

1. 発表者名 金澤陽, 平賀晃介, 池野谷玲太, 大橋剛介
2. 発表標題 ドライビングシミュレータを用いたNバック課題による水平有効視野の測定
3. 学会等名 第29回画像センシングシンポジウム
4. 発表年 2023年

1. 発表者名 Rebeka Sultana, Gosuke Ohashi
2. 発表標題 Predicting Driver's Visual Attention In BDD-A Dataset Using Optical Flow
3. 学会等名 The Eleventh International Workshop on Image Media Quality and its Applications (IMQA 2022) (国際学会)
4. 発表年 2022年

1. 発表者名 Tatsuya Oyabu, Rebeka Sultana, Gosuke Ohashi
2. 発表標題 Pre-training Dataset Generation for Classifying Beams of Vehicle Headlights from Nighttime Camera Image
3. 学会等名 Quality Control by Artificial Vision (QCAV2021) (国際学会)
4. 発表年 2021年

1. 発表者名 伊熊宗介, 大橋剛介
2. 発表標題 ドライビングシミュレータを用いた注視領域, 顕著性マップ, 深層学習の判断根拠の比較・解析
3. 学会等名 電気学会 電子・情報・システム部門大会
4. 発表年 2021年

1. 発表者名 スルタナレベッカ, 大橋剛介
2. 発表標題 Predicting driver's attention in BDD-A dataset using Multi-Level Feature-Based attention map and ConvLSTM
3. 学会等名 電気学会知覚情報/次世代産業システム合同研究会
4. 発表年 2021年



〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------