

**科学研究費助成事業 研究成果報告書**

平成 29 年 5 月 10 日現在

機関番号：12601

研究種目：若手研究(B)

研究期間：2015～2016

課題番号：15K16018

研究課題名(和文)映像コンテンツの顕著性変動解析による特徴ベース視線推定

研究課題名(英文)Gaze Estimation based on Saliency Dynamics Analysis

研究代表者

米谷 竜 (Yonetani, Ryo)

東京大学・生産技術研究所・助教

研究者番号：40733481

交付決定額(研究期間全体)：(直接経費) 2,100,000円

研究成果の概要(和文)：映像コンテンツを閲覧する際に人間がどこに注目しているかを知る手段として、視線推定と呼ばれる技術がある。本研究では、映像コンテンツ中において人間の目を引きつける領域の動きをモデル化した顕著性変動と呼ばれる情報を活用した視線推定技術の開発を目指した。具体的には、映像中の顕著性変動と視線運動(眼球の動き)の相関関係の解析に基づく視線推定技術の開発である。映像中で動く対象を追視する際の視線運動は、その対象の動きと相関するという現象を利用することで、人物ごとの校正手続きを必要としない視線推定が可能となった。本成果は2件の国内学会で発表された。

研究成果の概要(英文)：We develop a technique to estimate where people look at when they are watching video contents. The proposed method makes use of a property of videos called 'saliency dynamics', which models how salient regions in videos dynamically move over time, for gaze estimation. Specifically, we propose to analyze the correlation between saliency dynamics and eye movements. Intuitively, when people pursue moving objects in a video, we can observe a positive correlation between object motion and eye movements. This correlation feature allows us gaze estimation without manual calibration procedures. Experimental results have demonstrated the effectiveness of the proposed approach. These results were presented at two major computer vision conferences in Japan.

研究分野：コンピュータビジョン

キーワード：視線推定

### 1. 研究開始当初の背景

「目は心の窓」と言われるように、視線はしばしばその人間の注意や意図といった内的状態を表す。視線に関する研究は視覚心理分野を中心に古くより取り組まれてきており、たとえば従事するタスクによって視線の向けられる対象が異なることや、視線の統計的解析から発達障害を判断できる可能性があることが明らかになってきている。これに加えて近年では、閲覧中の映像コンテンツと視線をそれぞれ画像処理、信号処理し、両者の時空間的な構造や確率的な依存関係を統計的機械学習によって獲得する試みがコンピュータビジョン、マルチメディア分野で盛んに行われており、映像データのみから注視されやすい箇所を予測する技術や、視線データから背後に潜む内的状態を推定する技術が研究・開発されている。

### 2. 研究の目的

上に述べたような、映像コンテンツ、それを閲覧する際の視線情報、その背後に潜む内的状態の関係性を明らかにする注視行動解析を日常生活において実現するためには、テレビやパソコン、スマートフォンといった既成端末上で簡単に駆動する視線推定システムが不可欠である。本研究では、視線推定手法の一つである特徴ベース視線推定をとりあげ、映像コンテンツ中で人間の目を引き付ける顕著性変動と呼ばれる情報を解析して用いることで、カメラ・ディスプレイのみで構成される端末上で駆動し、かつ利用者ごとの事前校正を必要としない視線推定システムを実現することを目指す。

### 3. 研究の方法

本課題では特に、視線推定で必要となる事前校正に着目し、研究に取り組んだ。

多くの視線推定は、ディスプレイをユーザが閲覧する状況を想定し、a) カメラ等によってユーザの目情報（たとえば目画像）を計測し、b) それを画像処理することで特徴（たとえば瞳孔の位置や、補助的に照射する赤外光線の反射位置）を抽出し、c) その特徴をもとにディスプレイ中の注視位置を推定するという手続きで実現される。このとき、c) の手続きを行うにあたって、ディスプレイのどの位置を注視する際にどのような特徴が目情報から抽出されるかをあらかじめ対応付けする必要がある、これを校正（キャリブレーション）と呼ぶ。校正は視線推定の前に行う必要がある（事前校正）、既存の視線推定システムの多くは、「ディスプレイ上のいくつかの位置を指定し、ユーザに注視させる」という手動での校正が行われている（図1）。これを自動化することで、テレビやパソコン、スマートフォンといった端末上で簡単に駆動する視線推定システムという目的の一部を達成することを試みた。

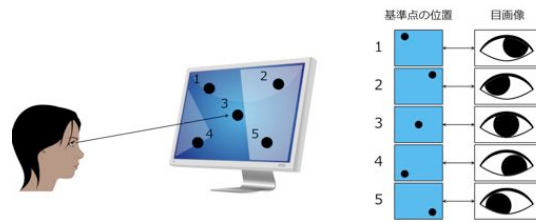


図1: 視線推定システムの手動校正

### 4. 研究成果

本研究の成果は、「滑動性眼球運動を用いた視線計測の自動校正」という新方式である。以下にその概要を述べる。

ディスプレイに提示された映像（たとえばテレビ番組やコマーシャル、あるいは町中で撮影されたサーベイランス映像）をユーザが閲覧する状況を考える。特に映像中に動いている対象があり、それをユーザが追従する場合、滑動性眼球運動と呼ばれる「動く対象を追いかける」目の動きが観測される（図2）。



図2: 滑動性眼球運動。赤点がユーザの注視位置を示す

もし映像から動く対象を自動抽出でき、またその動く対象を追従する眼球運動を視線データから同定できれば、それらの組は、視線推定システムの校正において必要となる「ディスプレイのどの位置を注視する際にどのような特徴が目情報から抽出されるか」という情報を知るための大きな手がかりとなる。これが提案手法の基本的なアイデアであり、新規性である。具体的な、以下のようなアプローチで提案手法の開発を実施した。

#### 4.1 映像中からの動対象の検出

コンピュータビジョン分野において映像の動き推定は基本的な課題であり、さまざまな手法が提案されている。特に本研究では、ディスプレイに提示される映像およびその中で動く対象の種類（たとえば人物、車）が多様かつ未知であることから、対象カテゴリに依存せず動対象を検出する手法が必要となる。また、事前校正という目的を考慮し、「ユーザが注視しうる」という条件を満たす領域を検出する必要がある。

そこで本研究では、近年動作認識分野で提案された dense trajectory と呼ばれる技術と、視覚心理分野で古くから研究されている視覚的顕著性 (visual saliency) と呼ばれるモデルを組み合わせたアプローチを採用した。Dense trajectory は、映像中から特徴的な点（たとえばエッジやコーナー）を多数抽出し、それを時間にわたって追跡するアプローチである。また、視覚的顕著性は、映像フレーム中で「色や明度が周囲と異なる部分」を人間が注意を向けやすい部分としてハ

イライトするモデルである (図3).

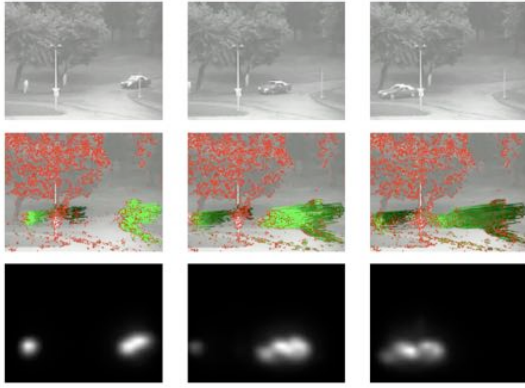


図3: 動対象の検出 上: 入力映像 中: Dense trajectory 下: 顕著性マップ

具体的には、多数生成された dense trajectory のうち、特に視覚的顕著性の高い部分のもののみを、映像中においてユーザが滑動性眼球運動によって注視する可能性の高い動対象として検出する。

#### 4.2 滑動性眼球運動の検出

映像を閲覧する際の眼球運動は、これまでに述べた「滑動性眼球運動」に加え、静止する対象を注視する際に発生する「固視」および注視対象を切り替える際に発生する「サッカド」によってしばしば特徴づけられる。これら3種類の眼球運動は、多くの場合、眼球運動の速度に閾値処理を施すことで分類することが可能である。しかしながら、どのような閾値をもって滑動性眼球運動を同定するかは、同運動がどの程度の速度で移動する物体を追従する際に発生するか大きく依存し、またその速度は映像において多様である。結果として、あらかじめ定めた単一の速度閾値をもって安定して滑動性眼球運動を同定することは難しい。

そこで本研究では、多重解像度解析の考え方にもとづき、さまざまなスケール(解像度)で表現される視線の階層構造を用いて、複数の閾値で同定された滑動性眼球運動を得る。図4にその概要図を示す。

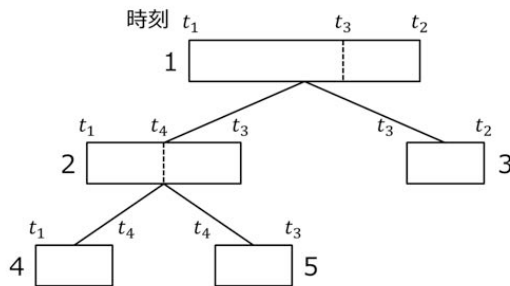


図4: 視線の階層構造. それぞれの分節が滑動性眼球運動の候補区間に相当する。

詳細は発表文献 1, 2 に譲るが、直感的には速度閾値を大きな値から徐々に小さくしつつ、閾値以上の速度を持つ眼球運動を検出す

ることにより、さまざまな速度で行われる滑動性眼球運動を検出することが可能である。

#### 4.3 動対象と滑動性眼球運動のマッチング

4.1 において映像中の動対象が、4.2 において眼球運動中の滑動性眼球運動が、それぞれ得られた。最終的に、どの動対象がどの滑動性眼球運動によって注視されているかというマッチングを計算することにより、事前校正を達成することができる。

基本的なアプローチは、動対象と滑動性眼球運動それぞれの持つ動きパターンの相関を計算し、その値の高いものほど対応の可能性が高い組として扱う、というものである。ただし、4.2 で述べたように滑動性眼球運動はさまざまな閾値をもって階層的に得られているため、ある階層で滑動性眼球運動がマッチングされると、その滑動性眼球運動が起こった区間と時間的にオーバーラップする別の滑動性眼球運動はマッチングされ得ない、という制約を考慮する必要がある(図5)。

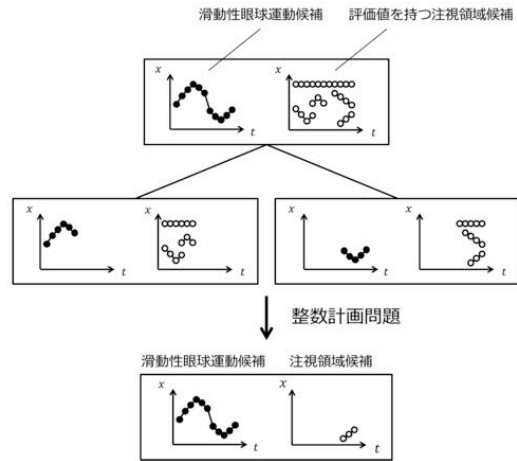


図5: 動対象と滑動性眼球運動のマッチング

そこで提案手法では、このような制約つきマッチングを整数計画問題で定式化するアプローチを導入し、その有効性を確認した。

#### 4.4 視線推定システムの校正

4.3 でマッチングが達成されることで、ある滑動性眼球運動が行われた際に、映像中のどの位置を注視していたかという情報が、動対象の位置として得られることになる。最終的に、これらの関係を表す線形モデルを最小二乗法等でフィッティングすることにより、事前構成を達成することができる。

重要な点は、以上の手続きを行うにあたって、ユーザはただ映像を閲覧しているだけでよいということである。これにより、視線推定システムの自動校正という技術が実現されている。

#### 4.5 実験

本研究では、視覚的顕著性モデルの評価で利用されている公開データを用いて評価実験を行った。実験の様子を図6に、また結果



の一例を図7に示す。



図6: 実験の様子

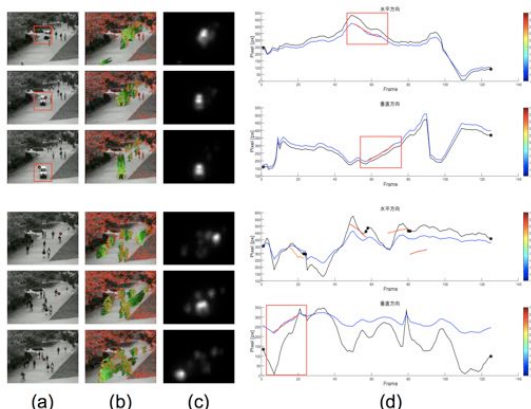


図7: 実験結果. (a) 入力映像 (b) dense trajectory (c) 視覚的顕著性 (d) 黒: 真の注視位置 赤: 事前校正に利用された動対象の位置 青: 校正された注視位置

実験の結果, a) 単一の閾値のみを用いて滑動性眼球運動を検出する手法 や b) 視覚的顕著性を利用せずすべての動対象を検出する手法 と比較し, 提案手法はより誤差の小さい校正を実現している. 詳細な結果は発表文献 1, 2 を参照にされたい.

#### 4.6 今後の展望

本成果は, たとえば人々がリビングでテレビ番組を閲覧している際に, 人々が余計な手続きを行うことなく視線推定を行うことができることを示唆している. これにより, たとえば視線から人々の興味を推定し適切な番組を推薦するといった実応用につながるものが期待される.

技術的には, 本研究の過程で, 視線推定の自動校正に適した映像とそうでない映像があることが示唆された. これをもとに, 今後は「自動校正を行うことが容易な映像デザイン」といった新たな問題に取り組むことが考えられる.

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 0 件)

[学会発表] (計 2 件)

- ① 松本 大輝, 米谷 竜, 佐藤 洋一, ”滑動性眼球運動を用いた視線計測の自動校正”, コンピュータビジョンとイメージメディア研究会, 大阪, 大阪大学吹田キャンパス 銀杏会館, 2016年1月
- ② 松本 大輝, 米谷 竜, 佐藤 洋一, ”滑動性眼球運動を用いた視線計測の自動校正”, 画像の認識・理解シンポジウム, 大阪 ホテル阪急エキスポパーク, 2015年7月

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

○取得状況 (計 0 件)

[その他]

特になし

#### 6. 研究組織

(1) 研究代表者

米谷 竜 (YONETANI, Ryo)

東京大学 生産技術研究所 助教

研究者番号: 40733481

(2) 研究分担者 なし

(3) 連携研究者 なし

(4) 研究協力者 なし