

広範囲空間における頭部姿勢変動に頑健な
キャリブレーションフリー視線推定

2017年7月

田村 仁優

主 論 文 要 旨

報告番号	甲 乙 第	号	氏 名	田村 仁優
主 論 文 題 目： 広範囲空間における頭部姿勢変動に頑健なキャリブレーションフリー視線推定				
(内容の要旨) コンピュータの性能向上とイメージセンサの低価格化に伴い、カメラが様々なデバイスに内蔵され、日常生活における応用が広がっている。ノート PC やタブレット、ショッピングモールや駅構内にあるデジタルサイネージでは、利用者の顔を捉えるカメラが機器に内蔵され、テレビ通話やジェスチャ認識、インタラクションやマーケティングに利用されている。このようにカメラの搭載されたデバイスは今後とも増える見込みである。これらの機器に搭載されたカメラに、キャリブレーション作業や使用時の頭部位置拘束のない視線推定機能が搭載されれば、様々な商業応用価値を生むと考えられており、その実現が期待されている。 本論文では、社会での実利用に即した視線推定手法を目的とし、キャリブレーションフリーかつカメラに対して広範囲な頭部位置で利用可能な注視点推定手法を提案する。従来の手法は、赤外線 LED や距離センサなど特別な装置を要求するものや、使用前のキャリブレーションを強いるものが多く、一般的な用途には適していなかった。近年、単眼カメラのみでキャリブレーションフリーを実現した手法が多数報告されているが、いずれも使用者の頭部位置のカメラ近傍への拘束や、目領域解像度が低い場合に大きな精度低下が発生するといった問題があり、実利用上の課題となっていた。これに対し本論文では、解像度に依存しない頑健な虹彩追跡手法を軸とし、顔特徴点検出、虹彩追跡、注視点推定から構成される広範囲空間内での視線推定を実現する手法を提案する。 第1章では、研究背景として、社会における視線推定の重要性や応用例を紹介し、その現状について述べる。また、関連する手法を挙げて、本研究の位置付けや研究目的を明らかにする。 第2章では、顔画像からの顔特徴点検出手法について説明する。顔特徴点検出では、自然環境下で撮影された様々な人種、顔向き画像を含むデータセットで学習した畳み込みニューラルネットワークにより、環境光変化や頭部姿勢変動に頑健かつ高精度な顔特徴点検出を実現する。さらに、視線推定のための前処理として、目領域のみを学習したネットワークにより詳細な目形状の推定について述べ、顔特徴点検出のために一般的に用いられるデータセットを用いて精度比較実験を行う。 第3章では、虹彩追跡について説明する。従来手法では外乱エッジや低解像度環境が課題であったが、提案手法では、3次元眼球モデルをもとに、虹彩エッジの強度および勾配方向に着目した尤度評価と密なサンプリングにより頑健な虹彩追跡を実現する。このように低解像度に強い追跡を実現する手法について述べる。 第4章では、注視点推定について解説する。広い空間内で非拘束視線推定を実現するため、独自の視線データセットを作成した。本データセットにより学習した回帰モデルによって、画面上の注視点を推定する方法について詳細に述べる。 第5章では、提案手法の有効性を検証する。第4章で作成した視線データセットによる検証を通して、従来のデータセットより低解像度の条件における注視点推定精度の比較を行う。 第6章では、本論文をまとめ、応用例とともに今後の課題と展望について述べる。				

以上

SUMMARY OF Ph.D. DISSERTATION

School Integrated Design Engineering	Student Identification Number	SURNAME, First name TAMURA, Kimimasa
Title Calibration-free gaze estimation in wide area robust to head posture variation		
Abstract <p>As the price of cameras has decreased and the performance of computers has increased, more consumer-grade devices have been equipped with cameras. Such cameras built in laptop PCs, tablets, digital signage in shopping malls and stations capture the user's faces, that are used for video calls, natural user interface, human computer interactions and marketing analysis. The number of built-in cameras is expected to increase. If a gaze estimation function, which does not require individual calibration processes nor constrained head pose, could be integrated into such cameras, they could have many applications. Thus, the realization of easy-to-use gaze estimation method is awaited.</p> <p>In this thesis, an unconstrained head pose and calibration free gaze estimation method in a wide space is proposed. The aim of this method is to be used in real world. The existing methods have required special equipment such as infrared LEDs, depth sensors or individual calibration prior to tracking, which are not suitable for general use. In recent years, some calibration free methods using a monocular camera have been reported, however, they are only applicable to the area near to the camera. In addition, the requirement of eye size in the image is also the crucial problem on gaze estimation, since the resolution affects the accuracy. On the contrast, the proposed method realizes a gaze estimation method robust to head pose changes in a wide space, which consists of the facial feature point detection robust to low quality of image, the iris tracking robust to low resolution, and the user-independent point of gaze regressor.</p> <p>In chapter 1, the importance and applied examples of gaze estimation are introduced as the background for this study, and then its current state and possibilities are described. In addition, related gaze tracking methods are cited, and the objectives of this study are clarified.</p> <p>In chapter 2, the facial feature points detection method is explained in detail. Facial feature point detection robust and highly accurate for ambient light change and head posture change is realized by convolutional neural networks learned with large scale data set including various race and face oriented images taken under natural environment. Furthermore, as preprocessing for gaze estimation, detailed refinement of eye shape by another network learned only with eye region is described. The proposed method of facial feature points detection is compared with conventional one in a commonly used dataset for face.</p> <p>Chapter 3 explains about iris tracking. In conventional method, outlier edges and low resolution caused problem. However, in the proposed method, the iris tracking is performed by a novel likelihood evaluation focusing on the intensity and gradient direction of the iris edges with dense sampling, which achieves robustness to low-resolution and low-quality of eye images.</p> <p>Chapter 4 describes about point of gaze estimation. In order to achieve unconstrained gaze estimation in a wide space, a unique gaze dataset were collected. Furthermore, this section explains about the method to estimate the point of gaze on a display by a regression model learned in the dataset.</p> <p>Chapter 5 describes the evaluation of a proposed method. Using the gaze dataset, comparison about point of gaze estimation accuracy under lower resolution than an existing dataset.</p> <p>In chapter 6, conclusion and future works of our study are shown.</p>		