

OpenCVとDlibライブラリを用いた 瞬き検出プログラムの開発

徳雄帆¹ 伊藤哲¹ 嶋田知泰¹ 西川広記^{1,2} 孔祥博¹ 富山宏之¹

概要: 近年、自動車の運転者の状態を監視する研究が多く行われている。本論文では、運転者の目を監視し、眠気を検出するシステムを開発する。開発したプログラムは、(1) 画像の読み込み、(2) 顔の検出、(3) 目の位置の特定、(4) 瞬きの検出の4ステップから構成される。ステップ(2)と(3)において、OpenCVとDlibの2種類のライブラリを試し、瞬き検出の精度と処理速度を評価する。実験結果によると、Raspberry Pi4上で瞬き検出を行う場合、Haar-LBFの組み合わせが最も良い結果を示した。

キーワード: OpenCV, Dlib, 居眠り検出システム, IoT

1. はじめに

近年、運転者の状態を監視する研究が多く行われている。監視する方法として心拍数や血圧などの循環器系や、体温などの基礎代謝系などがある。その中でも、瞬きから運転者の状態を監視する手法が期待されている。瞬きは比較的に弱い眠気を反映し、危険を未然に防ぐことから、居眠り検出に利用されている。瞬きを検出する手法として、文献[1][2]がある。しかし、これらのセンサを用いた瞬き検出手法は、装置を長時間装着すると運転者にストレスを与える可能性がある。そこで、センサを使用せず画像から瞬きを検出する手法が検討されている。文献[3][4]では、画像から開眼・閉眼の度合いで睡眠を判定する方法が提案されている。一方で、瞬き検出技術の精度と処理速度に基づいてアルゴリズムの組み合わせを評価する研究は少ない。本論文では、運転者の目を監視し眠気を検出するシステムを開発する。開発したシステムは、(1) 画像の読み込み、(2) 顔の検出、(3) 目の位置を特定、(4) 瞬きの検出の4ステップから構成されている。本論文では、ステップ(2)と(3)において、OpenCVとDlibの2種類のライブラリを試し、瞬き検出の精度と処理速度を評価する。第2章では、瞬き検出システムの詳細を説明し、第3章では、瞬き検出技術の精度と処理速度に基づいてアルゴリズムの組み合わせを評価している。第4章では本稿の結論について述べる。

2. 瞬き検出の概要

本章では、瞬き検出システムの概要を説明する。本システムのフローチャートを図1に示す。本システムは、画像の読み込み、顔の検出、目の位置を特定、瞬きの検出で構成されている。はじめに、680×480サイズの画像を読み込み、グレースケール変換を行う。

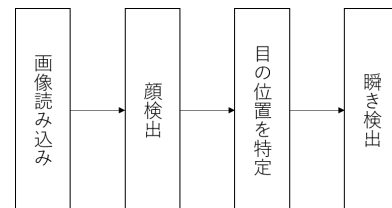


図1 瞬き検出システムのフローチャート

次に、画像内の顔の位置を検出する。顔の検出は、OpenCVのHaar特徴量を用いたカスケード分類器[5]や、DlibのHistogram of Oriented Gradients (HOG)特徴量を用いたSupport Vector Machine (SVM)分類器[6]を使用する。検出された関心領域から、OpenCVのFacemark Local Binary Features (LBF)[7]またはDlibのFacemarkKazemi (Kazemi)[8]を用いて、画像中の顔の関心領域内から68個のランドマークを検出する。顔のランドマークは、眉、鼻、口、目、顎のラインなどに一致するため、顔の部位を正確に検出できる。これにより、正確な目の位置を特定できる。ランドマークから目の位置を特定した後、目のアスペクト比(EAR: Eye Aspect Ratio)を算出する。最後に、瞬きを行っているかの判断は文献[9]を参考に行う。EARと閾値を比

¹ 立命館大学大学院理工学研究科
Ritsumeikan University, Kusatsu, Shiga, Japan

² 日本学術振興会特別研究員 DC1
JSPS Research Fellows

較し、閾値以下になった場合に瞬きを行ったと判断する。閾値は、各被験者の目の開閉時の EAR を用いて算出する。

3. 瞬き検出システムの評価実験

瞬き検出技術の精度と処理速度を評価するために、実験を行った。顔検出の比較手法として、Haar 特徴量を用いたカスケード分類器と HOG 特徴量を用いた SVM 分類器を採用した。実験では、680×480 の画像を 1778 枚使用しており、目を開けた状態と閉じた状態の画像が含まれている。また、被験者は眼鏡をかけていない人と眼鏡をかけている人を使用した。閾値は、各被験者から開眼時、閉眼時の画像をそれぞれ 10 枚選び、算出した EAR の平均値を用いた。実験の結果を表 1 と表 2 に示す。表 1 は、眼鏡をかけていないデータセットを使用した場合の結果を示し、表 2 は眼鏡をかけているデータセットを使用した場合の結果を示している。TP は目を閉じた状態の画像を用いて閉じた目を検出した場合、TN は目が開いていると判断した場合、FP は閉じた目が開いていると判断した場合、FN は開いた目が閉じていると判断して顔を検出できなかった場合を示している。

表 1 眼鏡をかけていない場合の評価結果

	Haar-LBF	Haar-Kazemi	HOG-LBF	HOG-Kazemi
TP	439	455	483	464
TN	527	527	518	522
FP	7	7	16	12
FN	95	79	51	70
再現率	82.21 %	85.21 %	90.44 %	86.89 %
適合率	98.43 %	98.48 %	96.79 %	97.48 %
正解率	90.45 %	91.95 %	93.73 %	92.32 %

表 1 によると、HOG-LBF の再現率と正解率はそれぞれ 90.44% と 93.73% を達成しており、他の 3 つの手法よりも高くなっている。一方で、適合率は各手法よりも劣っている。

表 2 眼鏡をかけている場合の評価結果

	Haar-LBF	Haar-Kazemi	HOG-LBF	HOG-Kazemi
TP	289	300	155	326
TN	298	237	295	108
FP	57	118	60	247
FN	66	55	200	29
再現率	81.41 %	84.51 %	43.66 %	91.83 %
適合率	83.53 %	71.77 %	72.09 %	56.89 %
正解率	82.68 %	75.63 %	63.38 %	61.13 %

表 2 に示すように、Haar-LBF は、眼鏡をかけている被験者において、83.53% の適合率と 82.68% の正解率を達成しており、他の 3 つの手法よりも高い精度を実現している。一方で、HOG-Kazemi や Haar-Kazemi と比較すると、再現率は劣っている。また、Haar-LBF 手法以外の手法において適合率、正解率は 80% 以下を示している。これは、眼鏡の影や光の反射により画素間の差分に誤差を生じさせ、メガネをかけた被験者ではランドマークがずれてしまうためである。本論文におけるランドマークのずれは、HOG-Kazemi 手法では約 1.98 %、Haar-Kazemi 手法では

表 3 システム実行時間

	Haar-LBF	Haar-Kazemi	HOG-LBF	HOG-Kazemi
処理時間 (秒)	0.050	0.045	0.155	0.151

約 3.52 %、HOG-LBF 手法では約 47.32 % となっている。したがって、Haar-LBF 手法はすべての指標において 80% 以上を示しており、ほかの手法より優れていることがわかる。また、各手法と比較した画像 1 枚あたりの処理時間の結果を表 3 に示す。表 3 より Haar-Kazemi は、処理時間の点で最も高速な手法である。一方、HOG を用いた手法は、Haar を用いた手法に比べて非常に低速である。

4. おわりに

本論文では、OpenCV と Dlib を用いた瞬き検出の精度と処理速度を評価した。その結果、Raspberry Pi4 上で瞬き検出を行う場合、Haar-LBF の組み合わせが最も良い結果を示した。検出精度の向上が今後の課題である。

謝辞 本研究の一部は科研費 20J21208 の支援による。

参考文献

- [1] R. Chhabra, S. Verma, and C. R. Krishna, "A Survey on Driver Behavior Detection Techniques for Intelligent Transportation Systems," in *International Conference on Cloud Computing, Data Science Engineering - Confluence*, 2017.
- [2] Y. Kim, "Detection of Eye Blinking Using Doppler Sensor with Principal Component Analysis," *IEEE Antennas and Wireless Propagation Letters*, vol. 14, pp. 123–126, 2015.
- [3] A. Rahman, M. Sirshar, and A. Khan, "Real Time Drowsiness Detection Using Eye Blink Monitoring," in *National Software Engineering Conference*, 2015.
- [4] J. W. Baek, B. Han, K. Kim, Y. Chung, and S. Lee, "Real-Time Drowsiness Detection Algorithm for Driver State Monitoring Systems," in *International Conference on Ubiquitous and Future Networks*, 2018.
- [5] P. Viola and M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001.
- [6] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005.
- [7] S. Ren, X. Cao, Y. Wei, and J. Sun, "Face Alignment at 3000 FPS via Regressing Local Binary Features," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2014.
- [8] V. Kazemi and J. Sullivan, "One Millisecond Face Alignment with an Ensemble of Regression Trees," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2014.
- [9] J. Cech and T. Soukupova, "Real-Time Eye Blink Detection Using Facial Landmarks," in *Computer Vision Winter Work*, 2016.