

# 無人自動運転車における複数歩行者に向けた視線制御

銭 寅飛<sup>†1</sup> 吳 在強<sup>†1</sup> 桂 新月<sup>†1</sup> 戸田 光紀<sup>†1</sup> 張 家銘<sup>†1</sup> 五十嵐 健夫<sup>†1</sup>

**概要:** 車と歩行者のコミュニケーションモダリティとして、多くの研究が「目」を提唱している。歩行者とアイコンタクトをとる場合、合理的かつ自動的に歩行者を見ることが重要な課題となる。しかし、複数の歩行者が車に接近している場合、どのように視線を制御すればよいかは自明ではない。本研究では、複数歩行者に対応した視線制御アルゴリズムを開発している。我々は、CEATEC2022において、提案手法を実装した実車プロトタイプでデモを行った。本稿では、我々の提案手法のアルゴリズムについて説明し、多くの来場者から頂いたフィードバックについて述べる。

## 1. はじめに

無人自動運転車では、人間が運転しないため、車の意思表示は外部向けヒューマンマシンインターフェース (external Human-Machine Interface, eHMI) を介する必要がある[1][2][3]。人間のような目を用いたインタラクション手法が提案されており、車と歩行者のコミュニケーションに効果を発揮している[3][4][5][6][7][8][9]。目の概念を用いることで、歩行者はより直感的に自動車の意思を理解することができるが、当該手法では目による歩行者とのインタラクションをいかに合理的かつリアルタイムに行うかが重要な課題となっている。

本稿では、目を eHMI として利用するコンセプトをさらに追究すべく、図 1 に示す実験用車両を用いて、本 eHMI に適した自動制御用アルゴリズムの提案および開発を実施した。我々は、自動視線制御用アルゴリズムを開発し、本実験車両に実装し、先述の提案システムを実現した。本提案システムでは、インタラクションの主体を車両前方の歩行者と定義していて、歩行者とのインタラクションは、「目」が歩行者を注視と追跡することとなっている。人間の運転手に近い視線の動きをすることで、歩行者が直感的に車の意図を理解できると考え、提案システムでは複数の異なるモードを作成し組み合わせてロボットの目の動きを自動制御し、どの歩行者に注目するかの優先順位を決定する。

本提案システムの性能を検証するため、複数の人を被験者とし、異なる人数や異なる環境の場合でのパフォーマンスを記録し分析し、動作検証を実施した。また、CEATEC2022 展覧会において、来場者を被験者 (歩行者) とした車両展示を行った。来場者の観察やインタビューを通じて、来場者から多くのフィードバックや提案を得ることができ、車と歩行者のインタラクション分野における他の研究者や、我々のシステムの改善に関する今後の開発に有益な知見となった。



図 1 実験車両

## 2. 提案システム

本システムは、歩行者を注視すべく、図 2 に示すように、歩行者検出アルゴリズムと視線制御アルゴリズムで構成される。具体的には、以下の処理を経て動作する。

- (a) 車の前 (両目の中央) に固定されたフロントカメラから映像ストリームをリアルタイムで取得。
- (b) キャプチャの画角内の歩行者をアルゴリズムによって検知。
- (c-1) 検出した各歩行者をクロップ (crop) して姿勢推定 (手振り、頭の位置などを検出) を行う。
- (c-2) 各歩行者を歩行者行列 (キュー) に追加する。そして、設定したモードに基づいてキューの中の歩行者順位を並べ替える。例えば、自動車に最も近い歩行者を優先する場合は、まず最も近い人を見るように視線を制御する。また、(c-1) と (c-2) は同時並行で処理される。
- (d) 優先順位が決まった後、先頭の歩行者を選択し、頭部の位置を 2 次元座標で推定する。
- (e) 車の目を制御するために、2 次元座標を線形写像関数で回転角度に変換し、車の目に送る。

<sup>†1</sup> 東京大学 大学院 情報理工学系研究科

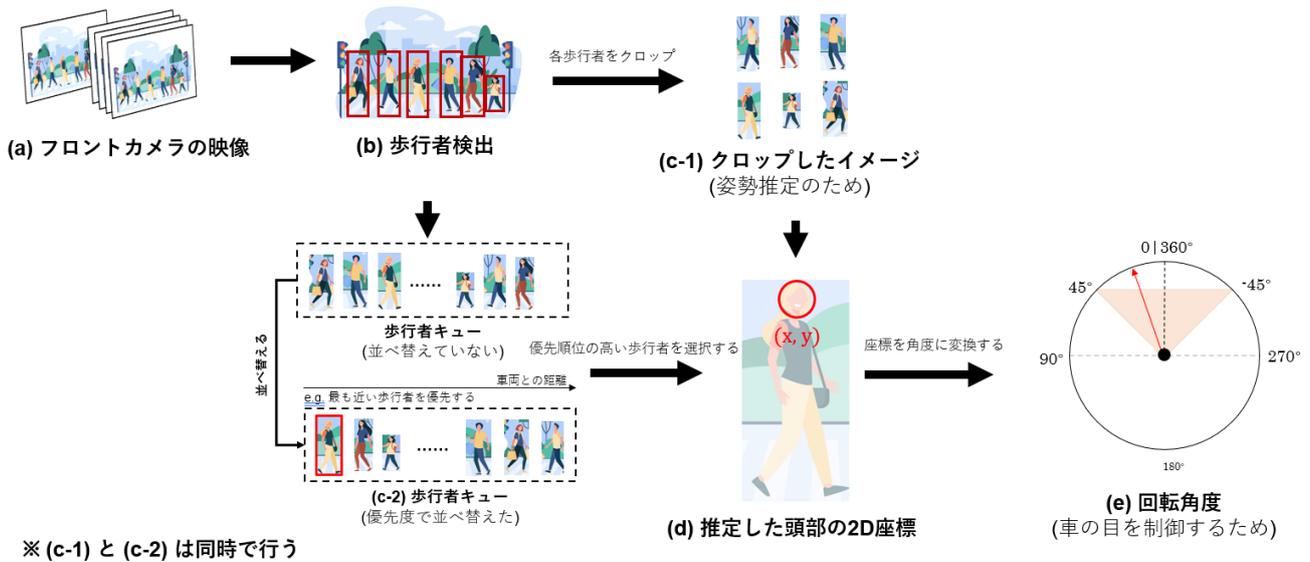


図 2 提案システムのワークフロー

## 2.1 歩行者検知アルゴリズム

まず、車に固定したフロントカメラで撮影した映像から歩行者を検出する。その後、得た映像の各フレームに対して既存の物体検出アルゴリズムで歩行者を検出する。歩行者の検出には、以下のルールを設けている。

1. 車に最も近い歩行者を優先して、一度に 5 秒まで視線を送る。
2. 車に向かって歩いている歩行者、移動速度の速い歩行者、車に向かって手を振っている歩行者を優先的に見る (図 3, 4)。

また、本システムは、以下のアルゴリズムによりルールを実現する。

1. 歩行者の車両からの距離、移動速度、顔の向きや動きなどの状態を推定し、どの歩行者を優先的に注視すべきかを計算。
2. 歩行者検出を本研究用に改良した Yolov4[10]を用いて行う。歩行者のデータセットは、Microsoft COCO dataset[11]を用いる。
3. MediaPipe[12]を用いて歩行者の姿勢を推定することで、手振りを検出。

歩行者検出のパイプラインは Praveen et al.[13]の研究と同様だが、GPU での並列計算を可能にしたのと、この後説明する車の目の制御と別のスレッドで動かさせたため、より高速に検出処理が実行される。

## 2.2 歩行者検知アルゴリズム

本研究用のプロトタイプとして、黒いプラスチック製の眼球と透明な半球状のアクリルカバーで機械式の「目」一対を作製し、本実験車両に搭載した。各眼球は 2 つのアクチュエータ(RC サーボモーター)で駆動され、回転について 2 自由度を有する。

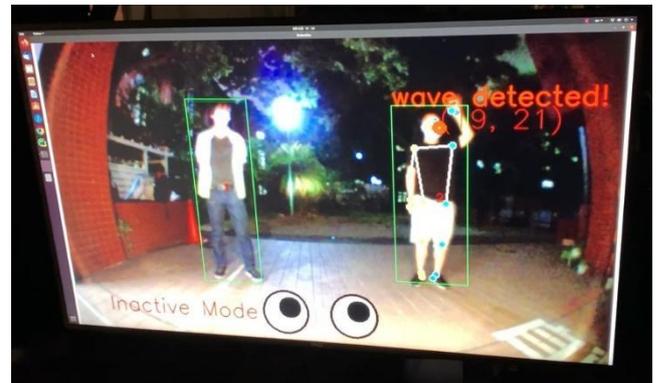


図 3 提案システムの UI

見るべき歩行者を検出した後、歩行者顔の 2D 座標を線形写像関数により車の目が回転する角度を計算し、アクチュエータに送信して眼球を動かす。ただし、機構上の制約から、各軸の最大回転角度は 45 度である。そこで、ハードウェアを保護するため、モーターの動作速度の上限に合わせて、適切なコマンドの送信頻度を推定した。

加えて、モーター側との通信処理で歩行者の検出などに遅延が生じたため、モーター制御の処理を、歩行者検知と対象の選定処理とは別のスレッドで実行している。

## 2.3 従来システムからの改善

本システムは、Praveen et.al [13]の研究を基に、検出速度の高速化、送信速度の高速化、インタラクションアルゴリズムの改良を行い、本実験車両にて動作検証を実施した。

図 4.では、最も代表的な視線移動パターンの一例を示す。図の下に詳細を記載した。

先行研究[13]の問題と本研究での対応、あるいは構想を以下に述べる。

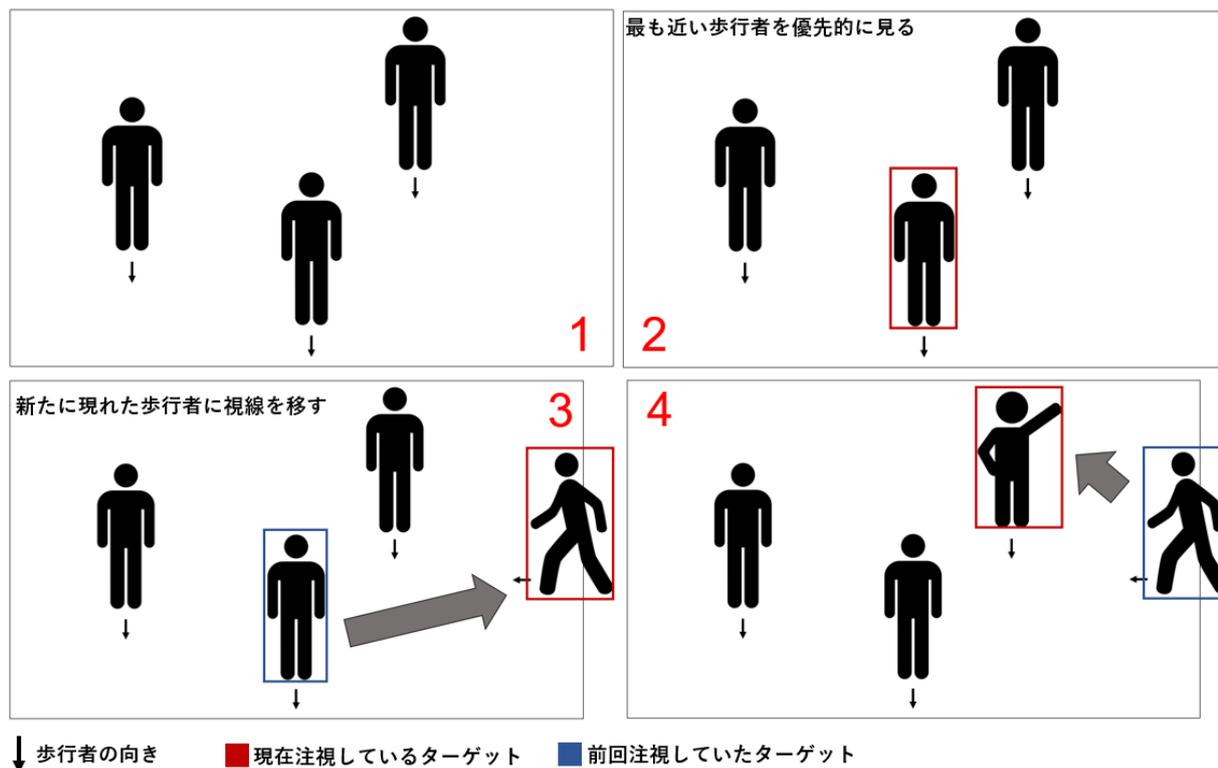


図 4 優先度ルールに基づく視線の動き

図 4 の説明：設定したルールの一つである。1 と 2 では、3 人の歩行者が車と向き合っているため、最も車に近い人を優先的に見るようにしている。3 では、元のターゲットから新たに現れた歩行者に視線を移す。4 では、手を振っている歩行者がいるため、最優先のターゲットとなり視線を移す。2 人以上の場合は、1 人の歩行者を 2 秒間見てから、歩行者列の中で優先順位の高い次の歩行者を見るように設定している。

(a) 対象を見失ったときや、検出環境（照明環境など）が不安定な際、眼球運動が乱れる。

現在、我々の手法では、視覚情報を基に歩行者を検知している。従来のシステムでは、カメラが撮影したフレームが明るすぎたり暗すぎたりすると、歩行者の位置を正確に特定することが難しかった。一度、対象を見失うと、新たな角度指令を受けることができず、角度の値が 0 になるため、目が元の位置に戻るなどの障害があった。また、検出が不安定になると、同じ対象の座標が変化し、目が小刻みに震える。そこで、歩行者に特化したニューラルネットワークを利用し[14]、精度と安定性を向上させた。さらに、撮影した元画像を調整して、露出過多や露出不足にならないようにした。また、対象を見失った場合、角度の値は 0 にならず、数フレームの間、元の位置を維持してから追跡を再開するルールを追加した。将来は、車両のレーダーや深度カメラ、赤外線カメラなどを組み合わせ、夜間などの環境下での歩行者認識性能の向上を図る予定である。

(b) 実験車両での性能テストでは、遅延時間が長い。従来のシステムでは、モーターの設計上、1 つのスレッドであまりに速い指令を送ると、未処理の指令が積み重なり、歩行者の検知から目の移動までの全処理に最大 3~5 秒の遅延が発生する。遅延を軽減させるため、歩行者検知とコ

マンド送信を複数のスレッドやプロセスに分割することで、車の目が常に一番新しいコマンドを受信できるように対応した。実験では、平均 92% の待ち時間短縮を実現した。良好な照明条件下では、待ち時間は約 0.3 秒から 0.6 秒となった。視線移動が多い状況（例：歩行者が多く、大きな視線移動が必要）では、遅延は 0.8 から 1 秒程度となる。今後は、モーター側との通信の遅延の軽減を進めつつ、より高出力かつ応答性の高いモーターで目を操作することで、よりリアルタイムに近い制御を実現する予定である。

(c) 目の焦点が一点に合っていない

当初のシステムでは、両目のモーターが同じ回転角度を受け取っていた。そのため、眼球が平行になり、一点に焦点を合わせていなかった。このような状況では、近い位置にいる 2 人の歩行者が、眼球が自分を見ていると勘違いする恐れがある。そこで、アルゴリズムで歩行者のおおよその距離を判断し、それぞれの眼球に個別に回転角度を送り、歩行者の顔の一点にフォーカスするよう工夫を施した。例えば、歩行者が車両の真正面に近づくと、目は内側を向くようになる。歩行者が車から遠ざかるほど、目は平行になる。実験では、従来のシステムと比較して、明らかに見られている感覚があると被験者から報告された。

### 3. 動作検証

我々は複数の被験者を歩行者とし、本提案システムの動作を検証した。その結果、歩行者が車から離れすぎていると、視線を向けても気づきにくいことが判明した。また、車から近すぎると、機械的な理由で注視や認識ができなくなる。そこで、対象範囲を車両前方から 0.75m~3.5m 以内に設計した。この範囲内で、着衣した歩行者を日中（良好な照度の場合）と夕方（低照度の場合）に自動車の前に立たせて、あるいは歩かせて実験を行った。その結果、低照度下では歩行者検知性能が低下し、検知対象を見落とす可能性が高くなることが確認された。また、1人から8人までの検出の効果を検証した。その結果、検出人数が規定範囲内の6人を超えると画面が混雑し、歩行者検出のエラー率が上昇した。本システムは、日中や明るい場所であれば、歩行者が6人以下の場合はリアルタイムに近い動作することが可能であることがわかった。

### 4. CEATEC2022 への出展

展覧会である CEATEC2022 では、本システムを搭載したプロトタイプ車両を展示し、来場者に本システムを実演・説明した（図 1,5）。来場者からの意見とフィードバックを3点に絞って記載する。

#### (1) 歩行者が多いときにシステムが故障しないかどうか

将来的には、より精度の高い歩行者検知を行い、頭部・顔面認識を追加して、各歩行者の検知をより適切に行うことと予定がある。ただし、我々の eHMI の利用シーンは群衆に特化したものではないので、多数の歩行者や群衆に遭遇した場合は、現状のシステムを無効化するか、群衆に適した別の eHMI に置き換える予定である。

#### (2) 歩行者の反応に対する目の動き

現在、私たちのシステムでは、歩行者の顔の向きから車に気づいたかどうかを推定している。また、歩行者が車に向かって手を振った後、歩行者に視線を送ることで、気づいたことを示すような応答も行っている。ただし、歩行者は車に対して普通は反応することが少ないことを考慮し、今後は歩行者の年齢層や状態特性（歩行速度、方向など）を推定して、視線を向けるタイミングなどを調整する予定である。また、車の目の動き（アニメーション）の設計を行い、アニメーション表現による歩行者とのコミュニケーションを目指す。

#### (3) その他のフィードバック

眼球型 eHMI は、一般的な無人自動運転車だけでなく、例えば無人配達車や重機、自動運転車イスなど、特に信号や交通標識がない場所や人ごみの多い場所でも使用することが可能ではないかとのフィードバックを頂いた。



図 5 CEATEC 2022 でのデモンストレーション

### 5. おわりに

本稿では、車の目の視線制御アルゴリズムを開発することで、複数の歩行者が車に接近している場合、視線移動の効果について研究した。よりリアルタイムに近い動作をさせるために、アルゴリズムの最適化を行った。同時に、本システムのパフォーマンスと効果を検証するための動作検証を実施した。その結果、理想な光環境とある程度以下の人数の場合において、本システムが設計通りに動作できることを確認した。また、CEATEC2022 で頂いた多くのフィードバックを参考に開発を継続し、システムの性能を向上させていく予定である。実際の自動運転における認識エンジンと結合することも重要な課題として残されている。その後、より詳細なユーザー調査を行い、定量的・定性的なユーザースタディを通じて本システムの有効性を分析することで、自動車と歩行者のコミュニケーション分野へ貢献していきたいと考えている。

**謝辞** 本研究は JST CREST JPMJCR17A1 の支援を受けたものである。また、車両は株式会社ティアフォーから提供されたものである。加えて、「目」のハードウェア開発を株式会社カラクリプロダクツに委託した。

### 参考文献

- [1] Yeti Li, Murat Dikmen, Thana G. Hussein, Yahui Wang, and Catherine Burns. 2018. To Cross or Not to Cross: Urgency-Based External Warning Displays on Autonomous Vehicles to Improve Pedestrian Crossing Safety. In Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '18). ACM, New York, NY, USA, 188–197. DOI:<https://doi.org/10.1145/3239060.3239082>
- [2] Stefanie M. Faas, Johannes Kraus, Alexander Schoenhals, and Martin Baumann. 2021. Calibrating Pedestrians' Trust in Automated Vehicles: Does an Intent Display in an External HMI Support Trust Calibration and Safe Crossing Behavior?. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (pp. 1-17). DOI:<https://doi.org/10.1145/3411764.3445738>

- [3] Stefanie M Faas, Lesley-Ann Mathis, and Martin Baumann. 2020. External HMI for self-driving vehicles: which information shall be displayed? *Transportation research part F: traffic psychology and behaviour* 68 (2020), 171–186.
- [4] Yoichi Ochiai and Keisuke Toyoshima. 2011. Homunculus: The Vehicle as Augmented Clothes. In *Proceedings of the 2nd Augmented Human*. 3–6. DOI:<https://doi.org/10.1145/1959826.1959829>
- [5] Chia-Ming Chang, Koki Toda, Daisuke Sakamoto, and Takeo Igarashi. 2017. Eyes on a Car: an Interface Design for Communication between an Autonomous Car and a Pedestrian. In *Proceedings of the 9th international conference on automotive user interfaces and interactive vehicular applications (AutomotiveUI '17)*, pp. 65-73. DOI: <https://doi.org/10.1145/3122986.3122989>
- [6] Jaguar Land Rover Automotive Plc., 2018. The Virtual Eyes Have it. Retrieved June 19, 2022 from <https://www.jaguarlandrover.com/2018/virtual-eyes-have-it>
- [7] Chia-Ming, Chang, Koki Toda, Takeo Igarashi, Masahiro Miyata, and Yasuhiro Kobayashi, 2018, September. A Video-based Study Comparing Communication Modalities between an Autonomous Car and a Pedestrian. In *Adjunct Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '18)*, pp. 104-109. DOI: <https://doi.org/10.1145/3239092.3265950>
- [8] Xinyue Gui, Koki Toda, Stela H. Seo, Chia-Ming Chang and Takeo Igarashi. 2022. "I am going this way": Gazing eyes on a self-driving car show multiple moving directions. In *Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '22)*.
- [9] Chia-Ming Chang, Koki Toda, Xinyue Gui, Stela H. Seo and Takeo Igarashi, 2022, Can Eyes on a Car Reduce Traffic Accidents?. *The 14th ACM International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2022)*, Seoul, South Korea, 17-20 September 2022
- [10] Zicong Jiang, Liqun Zhao, Shuaiyang Li and Yanfei jia. 2020. Real-time object detection based on improved YOLOv4-tiny DOI: <https://doi.org/10.48550/arXiv.2011.04244>
- [11] Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, Piotr Dollár. 2014. Microsoft COCO: Common Objects in Context DOI: [arxiv.org/abs/1405.0312](https://arxiv.org/abs/1405.0312)
- [12] Camillo Lugaresi, Jiuqiang Tang, Hadon Nash, Chris McClanahan, Esha Uboweja, Michael Hays, Fan Zhang, Chuo Ling Chang, Ming Guang Yong, Juhyun Lee, Wan-The Chang, Wei Hua, Manfred Georg, Matthias Grundmann. 2019. MediaPipe: A Framework for Building Perception Pipelines DOI: [arxiv.org/abs/1906.08172](https://arxiv.org/abs/1906.08172)
- [13] Praveen Singh, Chia-Ming Chang, and Takeo Igarashi. 2022. I See You: Eye Control Mechanisms for Robotic Eyes on an Autonomous Car. In *14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '22 Adjunct)*, September 17–20, 2022, Seoul, Republic of Korea. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3544999.3552320>
- [14] P. Fan et al., "Improved YOLOv4-tiny network for pedestrian detection," *2022 4th International Conference on Intelligent Control, Measurement and Signal Processing (ICMSP)*, 2022, pp. 1130-1133, doi: 10.1109/ICMSP55950.2022.9859096.