

鉄道総研報告  
RTRI REPORT物体検知 AI とステレオカメラによる  
構内の支障物検知システム向嶋 宏記\* 影山 棕\*\*  
長峯 望\* 小室 翔嗣\*\*\*

## Stereo Camera-Based Obstacle Detection System in Railway Yards

Hiroki MUKOJIMA Ryo KAGEYAMA  
Nozomi NAGAMINE Syoji KOMURO

Several railway operators are currently advancing toward driverless operations. To achieve end-to-end autonomy, operational conditions must be investigated not only along mainlines but also within depots and station premises. Our prior work addressed a forward obstacle detection system for mainline. Because environments such as yards and stations feature distinct characteristics—including object types, operation speeds, sightlines, and operational rules—this paper redefines the performance requirements and detection targets. Focusing on these environments, this paper defines the performance requirements based on current technology, proposes a train-borne obstacle detection architecture tailored to depot/station operations, and reports evaluation results for representative scenarios.

キーワード：列車前方映像，ステレオカメラ，AI，前方監視，構内

## 1. はじめに

近年，複数の鉄道事業者においてドライバレス運転の実現に向けた取り組みが進められている．列車の運行全体をドライバレス運転とするためには，本線走行以外に車両基地や駅の構内における運転も自動化することが求められてくる．このようなドライバレス運転の取り組みの中で，列車前方の支障物検知技術は，総合的な判断による安全確保<sup>1)</sup>を実現するための要素技術の一つとして期待されている．我々は，これまでに主に本線の走行を想定した列車前方の支障物検知システムの検討を進めてきた<sup>2)</sup>．特に，本線走行中は，遠方の支障物を検出することに主眼を置いたが，構内においては要求される検知距離が短くなるなど，前提条件が一部異なる．

そこで，本報告では車両基地や駅の構内での走行にターゲットを絞り，列車前方の支障物検知システムを実現する場合の構成を提案し，システムを構成する．各手法の性能を評価した結果について述べる．

## 2. 構内における支障物検知システムの構成

## 2.1 構内における条件の整理

本稿で想定する構内での走行とは，主に入換等を目的に駅や車両基地などで車両が低速で走行している状態である．支障物検知という観点で本線と構内での走行条件を比較した時，基本的には構内での走行条件は本線での走行条件に包含されると考えられる．つまり，構内のみを対象にした場合，想定する条件の範囲が本線よりも小さくなる部分が多い．ただし，前方の支障物の検知の際の分岐器の開通方向を含めた

\* 情報通信技術研究部 画像・AI研究室  
\*\* 旧 情報通信技術研究部 画像解析研究室  
\*\*\* 元 情報通信技術研究部 画像・AI研究室

自列車の進路の認識については、本線よりも複雑な構造となる可能性がある。それ以外の条件で、本線と構内の最も大きな違いは前提とする車両の走行速度が異なる点である。本線は線区最高速度までを想定する必要があるが、入換の場合は通常 20～45km/h 程度と想定できる。目標の設定においては、この速度から支障物検知システムが支障物を検知した際に、即座に非常ブレーキを動作させれば停止可能であると考えられる 100 m を目標とした。ただし、100m はあくまでも本稿における目標の値であり、実運用を検討する際には、車両の性能や、線区ごとの勾配の影響、天候による制動距離の変動などの要素についても加味してマージンを設けるなど慎重に判断する必要がある点には注意されたい。その他、本稿における支障物検知システムの要求仕様については以下の通りである。

- ・特定の検知対象物については人物を検知すること
- ・車両改造を最低限に留めるために、センサを車両内の運転台に設置すること
- ・特急列車等の流線型の先頭形状に対応できること
- ・保守作業に使用される台車や工具箱などの事前に想定されていない支障物が原理的に検知できること

## 2.2 支障物検知システムの構成

2.1 節で整理した内容について、分岐器の開通方向や人物の検出は事前に学習させる対象が明確であることから、可視光カメラを用いた AI に対応が可能であると考えられる。一方、工具箱などの事前に想定されていない物体は、学習データとして用意して AI に学習させることはできない。そこで、不特定の物体は 3 次元的に支障物として捉えることで対応する。なお、人物についても立体物として検出が可能であるが、将来的に人物の退避状態の認識などの検討ができるように AI による物体検出を併用する。

3 次元データを取得する方法は、LiDAR やミリ波が考えられるが、車内にセンサを設置する前提であるため、前面ガラスによってレーザー光やミリ波が減衰する。また、車両の先頭形状によらないことが求められるが、LiDAR は斜めの前面ガラスで反射されて距離の測定ができない場合もある。以上をふまえ、本稿では視差によって距離を推定するステレオカメラを採用する。

ステレオカメラによる支障物検知システムの構成を図 1 に示す。2 台の可視光カメラで撮影された画像を使って、ステレオ視による距離推定と立体物の検出、片方のカメラの画像を用いた検知範囲の設定と AI による人物の検出を行う。各フレームで設定された検知範囲との内外判定の結果から、過去の複数フレームにおけるフレーム単位での検知判定を行い、連続的に対象物を検知していた場合にシステムとして検知したことを確定させ、検知結果を出力する。出力された検知結果を使って、アラーム音や発光などで警報を運転士へ通知することや、将来的には車両の制御系を操作してブレーキを動作させることを想定している。次章において、処理部における各処理の詳細を述べる。

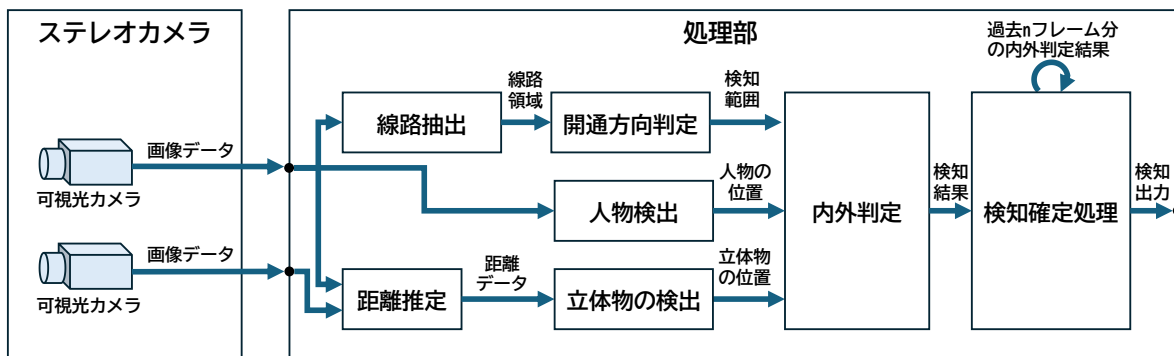


図1 ステレオカメラによる支障物検知システムの構成

### 3. ステレオカメラによる支障物検知手法

提案する支障物検知システムは、検知範囲の設定、人物の検出、立体物の検出、最終的な検知の確定処理からなる。本章では、これら各手法の詳細を述べる。

#### 3.1 検知範囲の設定と分岐器の開通方向の判定

画像から自身の進行方向のレールに沿って検知範囲を設定する。まず、画像中の特定の領域を画素単位で分類する「セマンティックセグメンテーション」と呼ばれる方式の AI モデル（以下、線路抽出 AI）を用いて「レール頭頂面」「左右レールの間の領域」を予測する。次に、隣接線の領域を除外するため、図 2 に示す通り画像下端の範囲から連続して繋がる線路領域を進行方向の線路領域として選択する。最後に左右レール頭頂面間の領域を広げて検知範囲を設定する。ここでは軌間と建築限界の幅が決まっていることを利用して、両者の比率の分だけ左右レール間の領域を広げることで、おおよそ建築限界の範囲内を検知範囲と定めている。また、線路抽出 AI は、様々なシーンの列車前方面像 3,324 枚に対しレール頭頂面、左右レール間の領域に色を塗ってラベル付けした教師データで学習した。図 3 にアノテーションの例を示す。自列車が走行しない隣接線であってもアノテーションを付与しているため、線路抽出 AI の出力は自列車の走行線と隣接線のすべてを含めた線路領域となる。

分岐の無い区間では前述した方法で進行方向の線路領域を選択することが可能である。一方で、分岐が存在する場所では、線路開通方向に沿って検知範囲を設定する必要がある。そこで我々は、分岐器の開通方向を判定する手法を提案した<sup>3)</sup>。提案手法では、左右のレールは常に分岐器の開通方向に向かって連続することに着目し、レール頭頂面の連続性によって開通方向を判定する。開通方向の判定処理を図 4 に示す。まず、線路抽出 AI で予測されたレール踏面領域に対して複数の水平線を設定し、水平線とレール領域の交点から一定の間隔を保つ点の組を左右レールのペアとして抽出する。次に、画像下端から消失点方向

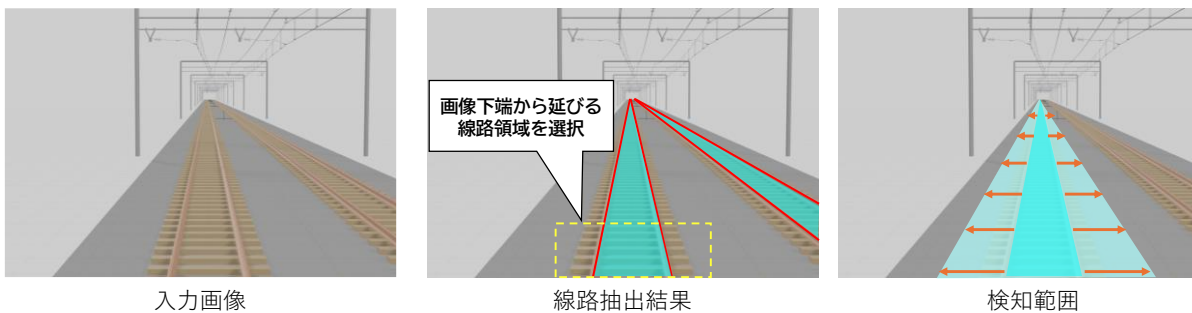


図 2 検知範囲の設定処理

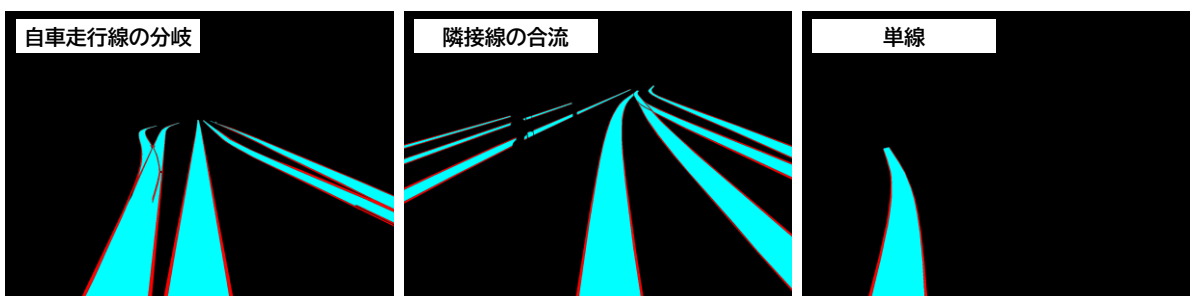


図 3 線路抽出 AI における学習データの例（アノテーションの例）

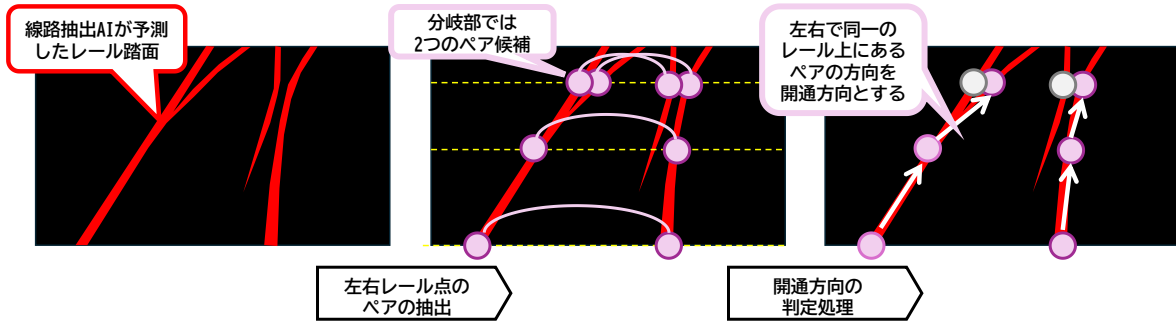


図4 分岐器開通方向の判定処理



(a) 背向方向

(b) 対向方向

図5 分岐器開通方向の判定処理結果例

に向けて、これらのペアのうち、左右それぞれで同一のレール上にあるものを順番に選択することで分岐器開通方向を判定する。最終的に選択された2本のレールを使い、検知範囲を設定する。

分岐器開通方向の判定処理の結果例を図5に示す。図5の例では(a)が背向方向、(b)が対向方向の分岐器であり、提案したアルゴリズムでどちらの分岐器も開通方向が判定でき、さらに奥にある分岐器についても開通方向が判定できていることが分かる。開通方向の判定精度はセマンティックセグメンテーションによるレールの頭頂面の領域抽出の精度に依存する。

### 3.2 人物検出

人物の検出にはYOLOX<sup>4)</sup>を用いる。COCOデータセット<sup>5)</sup>での事前学習済みモデルでは、転倒している人物やヘルメットと安全チョッキを着用した作業員に対する性能が低下する。これは、COCOデータセットに含まれる画像には、転倒している人物や、安全チョッキを着用した作業員がほとんど含まれていないためである。そこで、鉄道環境における人物の学習データでファインチューニングする。

線路内で転倒した人物を撮影できるのは、試験線路などの場所に限られる。しかし、AIの汎化性能を向上させるためには様々な鉄道環境で撮影した人物で学習させることが望ましい。そこで、人物が写っていない列車前方画像に対して、生成AIによって人物を合成して学習データを作成し、YOLOXの学習に用いる<sup>6)</sup>。図6に生成AIによって線路内に転倒した作業員を生成した例を示す。

### 3.3 ステレオカメラによる立体物検出

#### 3.3.1 カメラキャリブレーションの省力化

ステレオカメラでは、2カメラ間で視差を推定する際の処理を簡便にするため、2カメラの画像が完全に平行となるように配置されることが望ましい。しかし、実際には完全に平行に設置することは難しいため、

キャリブレーションにより、2 カメラ間で画像が平行になるように変換を行う。一般的なキャリブレーションは、図 7 のように特定の模様を描かれた板をカメラの前にかざし、画像から検出した模様の位置を基準に板とカメラ間の位置を推定し、2 カメラの位置関係を間接的に計算する。適切なキャリブレーション結果を得るためにはカメラの視野範囲をカバーするように複数の箇所でもボードを撮影する必要がある。

しかし、鉄道環境でキャリブレーションを実施する場合、作業員が触車する恐れがあるため、車両の前方の線路内で板を掲げての撮影が容易ではない。他の車両が確実に線路内に入って来ない措置を取るか、カメラを車両から取り外してから行う必要があり、保守の際の負担が大きいことが想定される。

そこでカメラを鉄道車両に設置したままで、2 カメラに共通で映るレールの位置を基準に平行化を行う手法を提案した<sup>7)</sup>。具体的には、画像のまくらぎ方向の傾きおよび上下方向のずれを補正したのち、画像に映るレールを線路抽出 AI によって検出した結果から仮想的な無限遠点である左右レールの消失点を求め、2 カメラ間で消失点の位置における視差がなくなるように2つの画像を変換する。変換前後の左右カメラ画像を重ねたアナグリフ画像と、変換後の2カメラの画像から距離を推定した結果を図 8 に示す。図 8(a)からは2つの画像が平行になっていることから、レール位置を基準とした平行化が適切に動作していることが分かる。また、図 8(b)はカメラから物体までの距離が色に変換されており、手前の設備から奥の設備まで距離が推定されていることが確認できる。



図 6 生成 AI によって生成した線路周辺で転倒した作業員

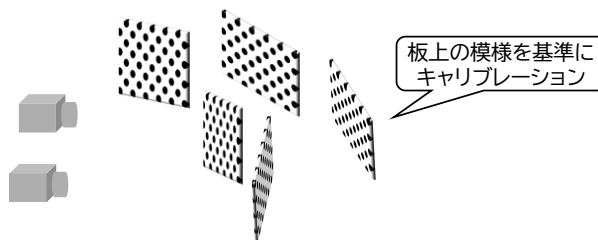
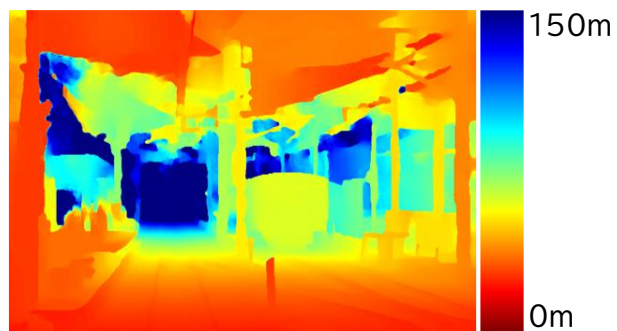


図 7 キャリブレーションのイメージ図



(a) アナグリフ画像



(b) 距離画像

図 8 ステレオカメラによる距離推定

### 3.3.2 距離画像からの立体物の検知

人物を含めた不特定の立体物を、線路抽出 AI によって推定したレール頭頂面の情報と、ステレオカメラからの距離画像を用いて検知する<sup>8)</sup>。立体物検知の処理を図9に示す。まず、距離画像を、カメラを中心とした3次元座標系の点群に変換する。線路抽出 AI で推定したレール頭頂面の領域が対応する点群から、レールレベルの平面を Random Sample Consensus (RANSAC)によって推定する。レールレベルからの高さが一定以上となる点のみを立体物として抽出し、距離画像に戻す。

このようにして得られた立体物のみの距離画像には、実際には地面付近などにおける距離推定の誤差に起因するノイズなども含まれる。このようなノイズを除去する手法としては、3次元空間でのクラスタリングにより物体を抽出する方法が考えられる。しかし、これらの手法は計算負荷が比較的大きいという課題がある。

そこで、立体物の距離分布に基づく軽量のノイズ除去手法を提案する。まず、図10に示すように距離画像を列方向に一定幅の短冊状領域へ分割する。次に、各領域において距離ごとの画素分布をヒストグラムとして算出する。実際の物体に由来する距離値は分散が小さい一方で、距離推定誤差に起因するノイズは分散が大きい、あるいは領域が極端に小さい傾向をもつ。このため、物体が存在する距離に対応するヒストグラムはスパイク状のピークとして観測される。最後に、得られたヒストグラムに対してしきい値処理を適用し、特定の距離に集中して分布する成分のみを抽出することで、ノイズを効果的に除去する。

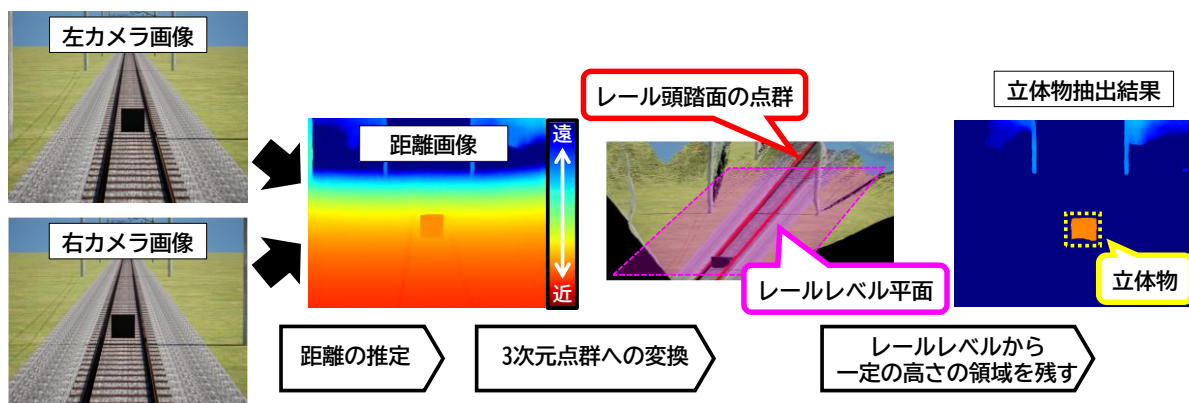


図9 立体物の検知処理

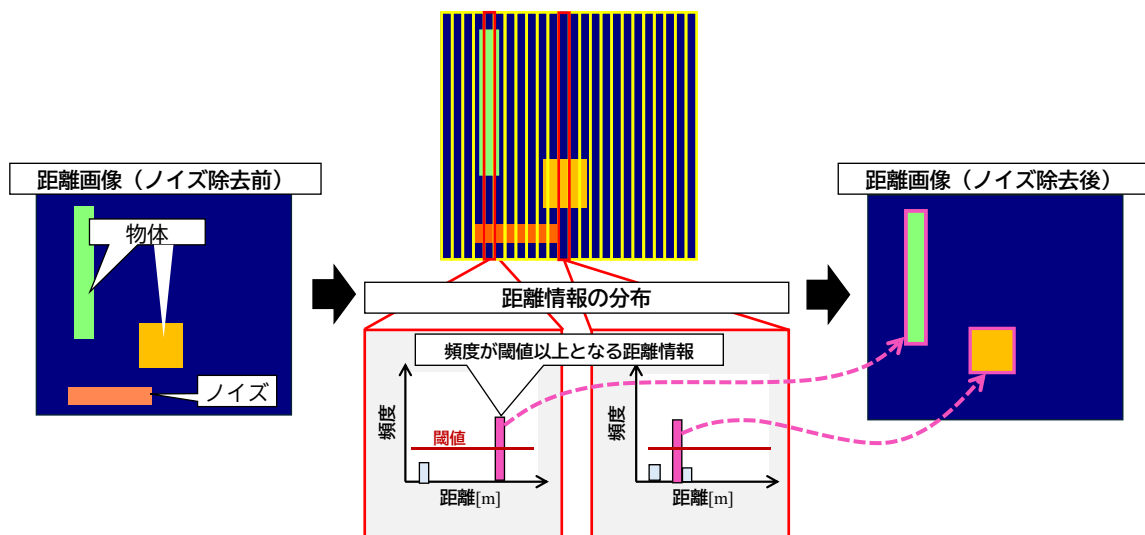


図10 立体物の距離画像からのノイズ除去処理

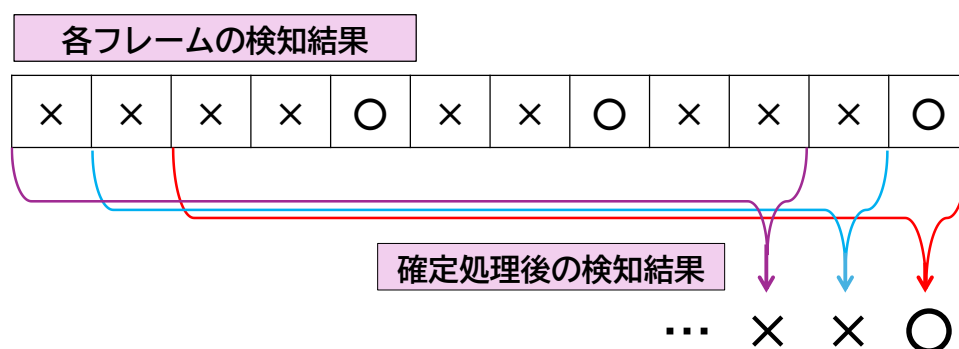


図11 検知結果の確定処理

### 3.4 検知の確定処理

フレーム単位での検知処理の結果の中には瞬間的な誤検知や未検知も含まれる。最終的な出力におけるこれらの影響を除くため、過去の連続するフレームのうち、一定数以上のフレームで検知出力がされている場合に検知を確定させる処理を施す。検知結果の確定処理の例を図11に示す。図11では、検知が10フレーム中3フレーム以上存在する場合に確定処理において検知と出力する様子を示している。以降の評価試験においても、検知確定処理のパラメータは10フレーム中3フレーム以上と設定した。

## 4. 検知性能評価

### 4.1 評価用データの撮影

対象との距離ごとの検知性能を評価するため、実際の車両基地構内においてデータ取得試験を実施した。本試験では、実際の車両にステレオカメラを仮設し、車両を被写体から遠ざける方向に走行させてデータを取得した。この1回の走行を1試番として、2023年6月～8月までの8日間に、昼間43試番、夜間36試番の合計79試番を実施した。被写体は図12に示す通り、服装や姿勢の異なる人物、および人物以外に20cm<sup>2</sup>～30cm<sup>2</sup>程度の工具箱・工具バッグを対象とした。また、撮影に用いたカメラの諸元を表1に示す。カメラは、レールレベルからの高さは約2.7mで俯角は約4.0°、基線長は25cmでまくらぎ方向に平行に配置した。

### 4.2 評価方法

検知性能評価にあたり、実際のシステムの適用場面に近い条件で評価するため、撮影時に車両が被写体から遠ざかる向きに画像を取得し、評価時に逆再生することにより、被写体に近づく方向を再現した。なお、安全のために、被写体となる人物は車両が離れていく方向に動き出したことを確認した後、線路内に立ち入った。検知距離については、隣接フレーム間での軌道面の画素値の移動量を推定し、走行試番の画像群全体で積算したものを実際の走行距離と対応付けることで、フレームごとの対象との距離を算出した。

被写体が最も遠い場所から順に検知処理を行うと、一定の距離まで近づいたタイミングで対象を検知できるようになり、被写体が一時的に画角から外れない限りは被写体の手前まで検知状態が持続する。このような検知・未検知の境界となる距離を「最遠検知距離」と定義し、条件別に集計した。なお、撮影時においてカメラと被写体が最も近接した状態において対象を検知できなかった場合、実際にどの程度の近傍距離で検知が可能であるかを評価することが困難である。このため、本報告では最遠検知距離を0mとして集計した。



図 12 被写体の例

表 1 使用したカメラの諸元

|         |                        |
|---------|------------------------|
| 解像度     | 水平 1,920px, 垂直 1,536px |
| 素子サイズ   | 3 $\mu\text{m}$        |
| レンズ焦点距離 | 24 mm                  |
| 感度※     | 1,953 mV               |

※露光時間 1/30 秒での値

### 4.3 評価試験の結果

人物について、時間帯および姿勢ごとの最遠検知距離の平均値と標準偏差を箱ひげ図で示したグラフを図 13 に示す。公開データセットのみで学習した AI モデル①と、AI によって生成した転倒している作業員を学習させた AI モデル②のそれぞれで評価している。AI モデル①では、ステレオカメラによる立体物の検出でも、転倒した状態では、レールレベルとの差が小さく転倒した人物を検出することができなかった。一方で、AI モデル②では転倒状態を含めた人物について、最遠検知距離は 100 m 以上となった。これは人物認識の AI の性能向上によるものである。

夜間においては、対象が撮像される距離が短くなり、最遠検知距離は 50m 程度であった。これは、画像全体が暗くなることでステレオカメラによる視差の推定ができず、AI による認識では安全チョッキによる光の反射によって検出性能が低下したためであると考えられる。対策としては、明るさを補正する処理を追加し、ステレオカメラによる視差の計算をしやすくする方法と、安全チョッキによって反射した状態の学習データを追加することで AI による認識性能を向上させることが考えられる。

最後に、人物以外のバッグや工具箱については、今回のシステムでは検出ができなかった。これは、RANSAC によるレールレベル平面の推定精度が不十分であるためと考えられる。具体的には、レール頭頂面の抽出誤差や、ステレオカメラによる距離推定誤差の影響により、推定される平面が実際のレールレベ

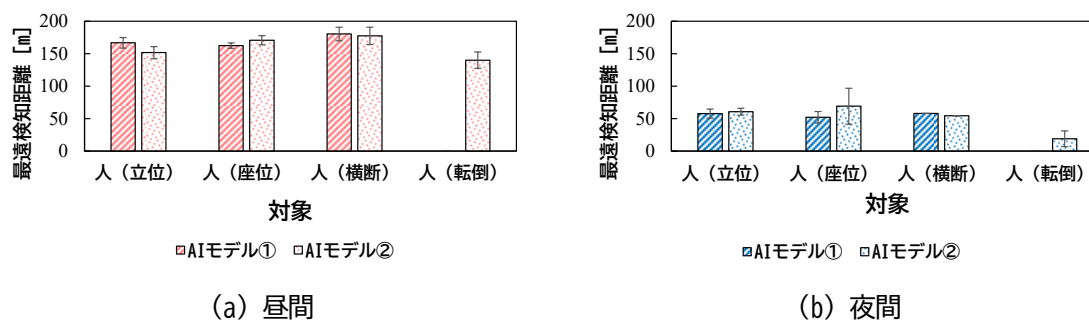


図 13 AI モデルごとの各対象の最遠検知距離

表 2 処理装置の諸元

|       |                                                            |
|-------|------------------------------------------------------------|
| CPU   | Intel® Core i9-10900TE (1.8 GHz/4.5 GHz 10 コア 20 MB キャッシュ) |
| GPU   | NVIDIA RTX A4500                                           |
| メモリ   | 64 GB                                                      |
| ストレージ | 2 TB                                                       |
| 備考    | ボックス PC 鉄道規格 EN50155 認証                                    |

表 3 リアルタイム動作における条件別の最遠検知距離

| No. | 対象物    | 時間帯 | 最遠検知距離 (m) |
|-----|--------|-----|------------|
| 1   | 人 (立位) | 昼   | 104        |
| 2   | 人 (立位) | 昼   | 184        |
| 3   | 人 (立位) | 昼   | 162        |
| 4   | 人 (立位) | 昼   | 108        |
| 5   | 人 (横断) | 昼   | 181        |
| 6   | 人 (横断) | 昼   | 130        |
| 7   | 人 (立位) | 夜   | 23         |
| 8   | 人 (立位) | 夜   | 23         |

ルよりも高い位置、あるいは実際とは異なる傾きを有する平面として推定される場合がある。その結果、レールレベル平面以上の領域から立体物を抽出する段階で、本来検出されるべき地面近傍の対象物が平面下に埋もれてしまい、立体物として正しく抽出されないという問題が生じた。

この問題に対する対策としては、レール抽出の精度を向上させることや、ステレオカメラの基線長を拡大することで距離分解能を向上させ、距離推定精度の改善を図ることが有効であると考えられる。

#### 4.4 リアルタイム動作の検証

提案手法について、2 カメラからの画像の取得からモニタへの検知結果の出力までの一連の処理を、実車環境でリアルタイムに動作させる試験を 8 試番実施した。車両基地構内において、実車の運転台に、ステレオカメラおよび処理装置、モニタを仮設した。処理装置の諸元は表 2 の通りである。処理装置は実際の車載装置に近い仕様を有するものとして、鉄道車両に搭載することを想定して設計されたボックス PC を用いた。被写体から車両は遠ざかる方向に走行し、フレーム単位の検知結果を保存した。列車が被写体に近づく方向で走行した場合と同等になるように、逆再生して評価することでシステムの最遠検知距離を求めた。

人物の昼夜別の最遠検知距離を算出し、条件別にまとめた結果を表 3 に示す。昼間は、100 m 以上遠方で人物を検知できていることが分かる。処理の性能に関しては、1 秒あたり 2~3 フレーム程度であったため、今後高速化などの検討を進める。

## 5. まとめ

本報告書では、構内をターゲットにセンサを用いた列車前方支障物検知システムの構成と検知アルゴリズムを提案し、実際に車内に設置したステレオカメラを用いた評価試験の結果を示した。

線路の分岐が多い構内において線路の開通方向に検知範囲を設定するため、AI が予測した左右レールが画像下端から連続する方向に判定することで、開通方向を判定する手法を提案した。提案手法では、昼間の人物について姿勢によらず約 100 m から検知できることを示した。一方で、夜間における検知性能、および人物以外の軌道面付近の物体の検知性能に課題があることを示した。

今後は軌道面付近での物体検知性能の課題への対応と、時間帯以外に天候状態が変化した際の提案手法における性能評価を行う予定である。

## 文 献

- 1) 国土交通省：鉄道における自動運転技術検討会とりまとめ, <https://www.mlit.go.jp/tetudo/content/001351026.pdf>, 2022 (参照日: 2024 年 12 月 26 日)
- 2) 影山 椋, 長峯 望: 列車前方監視のためのセンサフュージョンによる支障物検知手法の検討, 電気学会論文誌 D, Vol.144, No.3, pp.70-78, 2024
- 3) 影山 椋, 長峯 望, 山下 光太郎, 丸山 量志: 構内における列車前方監視のための検知範囲設定手法, 第 30 回鉄道技術・政策連合シンポジウム (J-RAIL2023) 講演論文集, S6-1-6, pp.410-413, 2023
- 4) Y Zheng Ge and, Songtao Liu, Feng Wang, Zeming Li and Jian Sun, “YOLOX: Exceeding YOLO Series in 2021,” CoRR, vol. abs/2107.08430, arXiv preprint arXiv:2107.08430, 2021. URL: <https://arxiv.org/abs/2107.08430>.
- 5) L. Tsung-Yi, et. al., “Microsoft coco: Common objects in context,” In Proc. of the 13th European Conference on Computer Vision, pp.740-755, 2014.
- 6) 清水 義輝, 影山 椋, 長峯 望, 山下 光太郎, 丸山 量志: 学習データの改良による車上型支障物検知手法の検知性能向上, 電気学会交通・電気鉄道研究会, TER-23-067, 2023
- 7) 影山 椋, 長峯 望: レール位置を用いた車載ステレオカメラの効率的な平行化手法, 電気学会交通・電気鉄道研究会, TER-24-094, 2024
- 8) 影山 椋, 長峯 望: 列車前方監視のためのステレオカメラによる線路内の立体物検知手法, 第 31 回鉄道技術・政策連合シンポジウム (J-RAIL2024) 講演論文集, SS6-3-3, 2024