

## 先読み運転支援を可能とするための歩行者認識

新村 文郷<sup>1)</sup> 久徳 遙矢<sup>1)</sup> 川西 康友<sup>1)</sup> 出口 大輔<sup>1)</sup> 福井 宏<sup>2)</sup>  
 山内 悠嗣<sup>2)</sup> 山下 隆義<sup>2)</sup> 藤吉 弘亘<sup>2)</sup> 村瀬 洋<sup>1)</sup>

### Pedestrian Recognition for Proactive Driving Support System

Fumito Shinmura Haruya Kyutoku Yasutomo Kawanishi Daisuke Deguchi Hiroshi Fukui  
 Yuji Yamauchi Takayoshi Yamashita Hironobu Fujiyoshi Hiroshi Murase

Pedestrian behavior prediction is one of fundamental technique for proactive driving support system. For their behavior prediction, not only pedestrian detection but also their attribute recognition is important key. In our COI project, we have developed those techniques and implemented them on a vehicle. This paper presents our recognition system including detection of pedestrians and recognition of their attributes for behavior prediction. To show the performance of our system, we evaluated the system on practical environments.

**KEY WORDS:** Safety, Pedestrian detection, Image processing, Attributes recognition

#### 1. ま え が き

衝突事故の低減のため、自動ブレーキシステムをはじめとした数多くの交通安全システムに関する研究開発が行われてきた。しかし、従来の自動ブレーキシステムでは前方の障害物との衝突は回避できるものの、飛び出し歩行者との衝突回避に対する効果は小さい。そのため、高齢者が苦手とするこのような素早い判断が求められるような事故の低減には有効な手段とはならない。このような問題を解決する安全運転支援システムを実現するためには、歩行者の飛び出しや先行車の急な進路変更といった危険な状況を予測する先読み運転が重要である。

先読み運転を実現するためには、対象の動きを早い段階で予測する技術が必要である。自車前方の歩行者を回避する場面においては、歩行者が数秒先にどこに移動するかを予測することが重要である。これを実現するためには、歩行者の位置を遠方から正確に把握し、予測の手掛かりとなる体向き等の属性を精度よく認識する技術が必要である。すなわち、高精度な歩行者検出技術と属性認識技術の二つが必要となる。そのため本研究では、先読み運転に必要となるこれらの歩行者認識技術の開発を目的とする。

歩行者認識のためのセンサには、カメラやLIDAR、ミリ波レーダのように様々なものが存在する。これらの中で、車載カメラは導入コストも最も小さく、撮影された画像からは多様な歩行者属性の認識が期待できる。そのため、本プロジェクトでは、前方車載カメラを主要なセンサとして利用する。

車載カメラの画像や映像からの歩行者検出や属性認識は、

- 1) 名古屋大学(464-8601 愛知県名古屋市千種区不老町)
- 2) 中部大学(487-8501 愛知県春日井市松本町 1200)

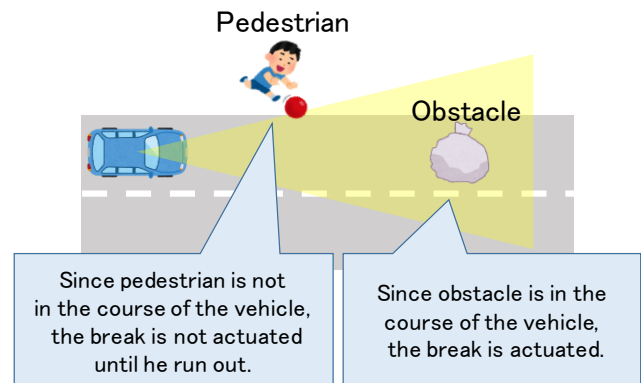


Fig.1 A situation of conventional auto braking systems.

数多くの研究機関からさまざまな手法が提案されている。代表的なものとしては、Histogram of Oriented Gradient (HOG) 特徴量と Support Vector Machine (SVM) を組み合わせた手法 (1) や、歩行者の全身と各部位の位置関係を利用する Deformable Part Model (DPM) による手法 (2) などが提案されている。近年では、深層学習を用いた検出手法 (3) も提案されており、優れた検出性能を示している。属性認識では、体向きや身長、年齢といった属性を推定する手法 (4, 5) が提案されている。本研究では、歩行者の動きの予測に特に重要と考えられる体向きの属性に注目する。歩行者検出と同様の枠組みを用いて体向きを推定する手法 (6) や、様々な属性と同時に推定する手法 (7) が提案されている。本研究では、歩行者認識技術の実車への実装を想定している。そのため、実車に搭載可能なシステム上で高精度かつリアルタイムでの歩行者認識を実現する必要がある。近年注目を集めている深層学習を用いた検出手法は高精度であるものの、処理負荷が非常に大

きい。そのため、深層学習のみで広範囲の歩行者検出を行うことは難しい。そこで、GPU等の演算装置を用いなくても高速な計算が可能な Aggregated Channel Features (ACF) (8) を用いた歩行者検出手法と属性(体向き)認識手法を開発した。さらに、精度向上と高速化の両立のため、Locally Decorrelated Channel Features (LDCF) (9) による追跡処理を組み合わせる。これにより、実車において高精度かつリアルタイムでの歩行者認識を実現する。

本発表では、2章にて提案する歩行者認識手法について述べ、3章にて実車を使用した評価実験について述べる。最後に、4章にて本発表をまとめる。

## 2. 歩行者認識

本手法では、まず入力された画像全体から ACF を用いて歩行者を検出する。ACF を用いることにより、複数スケールに亘る特徴量を高速に計算することが可能である。しかしながら、そのような複数スケール全体から検出窓を走査して歩行者検出を行うためには、膨大な判別処理が必要となる。そこで画像全体からの歩行者検出と同時に、一度検出された歩行者の周囲に範囲を絞って詳細な探索(追跡)を行う処理を導入する。この追跡処理は探索範囲が比較的小さく済むため、全体探索よりも高速である。ここでは、LDCF を用いることで検出の高精度化を図る。

次に、検出された各歩行者の体向きを推定する。本手法では、体向きを8方向(0°, 45°, …, 315°)に分類する。体向きの違いは歩行者の見えの違いから識別できると考えられるため、歩行者の色や輝度勾配といった見えの特徴を捉える ACF を用いる、また、これは方向をクラスとした多クラス(8クラス)分類である。ここでは、Random Forest を用いる。

歩行者認識の処理概要を図2と図3に示す。本手法は、事前に歩行者検出器や体向き推定器を統計的学習により構築する学習段階と、車載カメラより入力された画像から歩行者を検出して体向きを推定する認識段階に分かれる。

### 2.1. 歩行者検出器と体向き推定器の構築

学習段階において、歩行者検出器と体向き推定器を構築する。まず、事前に学習に用いる歩行者検出用と体向き推定用に学習画像群を用意する。歩行者検出用の学習画像群は、正例として歩行者の画像群を、負例として歩行者が写っていない背景の画像群をそれぞれ用意する。体向き推定用の学習画像群は、歩行者画像に各方向の体向きのラベルを付与した画像群を用意する。

歩行者検出器は、歩行者検出用の学習画像群を用いて、画像群から算出した特徴量をもとに歩行者と背景を識別する Boosted trees を構築する。このとき、ACF による検出と LDCF による追跡を行うために、ACF により学習した検出器と LDCF により学習した検出器の2種類を構築する。各特徴量の算出

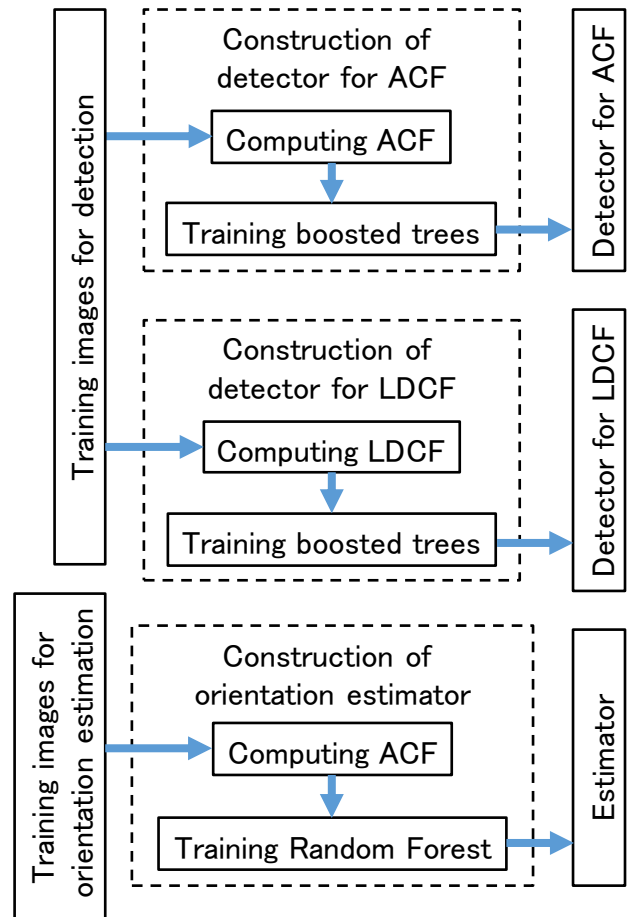


Fig.2 Process flow of the proposed method in training phase.

については、次節で述べる。Boosted trees の学習では、算出された特徴量の中から歩行者の識別に有効な特徴量とそのしきい値を選択することで弱識別器を構築する。これを繰り返すことで弱識別器群を構築し、それらを統合することで1個の強識別器を構築し、これを検出器として利用する。

体向き推定器は、体向き推定用の学習画像群を用いて、画像群から算出した ACF をもとに各方向と特徴量の関係を Random Forest により学習することで構築する。

### 2.2. ACF を用いた検出

認識段階では、まず画像全体に対して検出窓を走査し、各位置での ACF を算出する。ACF は、画像から色や輝度勾配強度、勾配方向ヒストグラムを計算し、それらをチャンネル化した特徴量である。色は、画像を LUV 色空間に変換することで3チャンネルの特徴量として得る。輝度勾配強度と勾配方向については微分画像から求める。ここで、輝度勾配強度は1チャンネルの特徴量、勾配方向は周辺の領域から6方向のビンのヒストグラムを計算することで6チャンネルの特徴量として得る。これにより得た10チャンネルの特徴量を連結することで ACF を得る。

一般的な歩行者検出では画像をラスタスキャンし、各検出

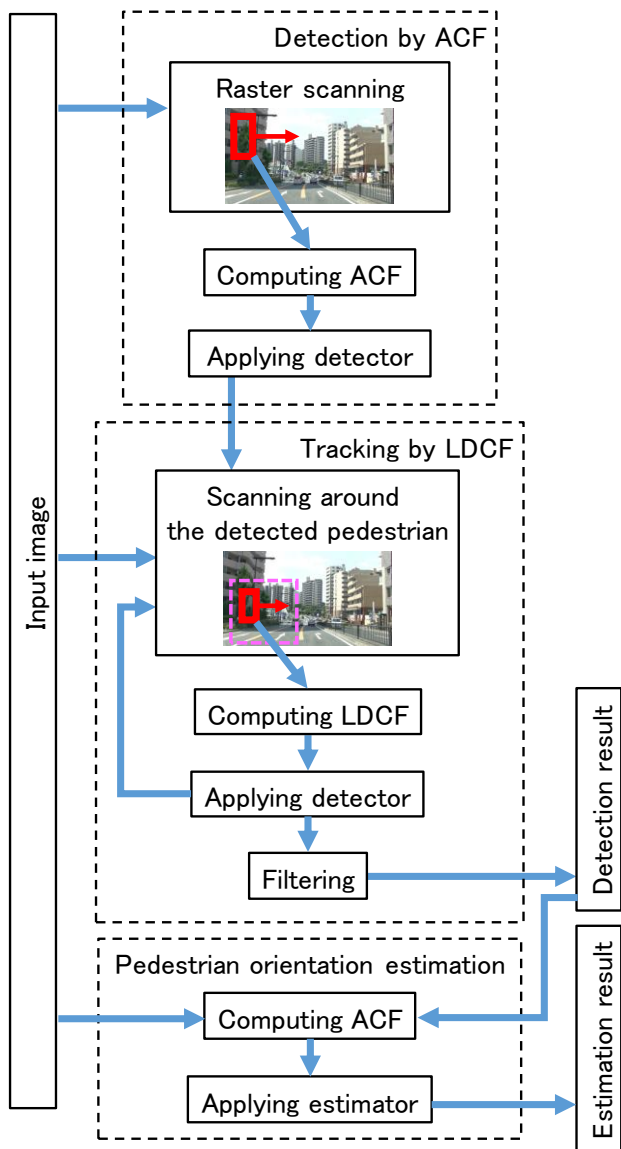


Fig.3 Process flow of the proposed method in recognition.

窓が歩行者か否かを判定する。これを画像のスケールを様々なに変化させて行うため、特徴量の計算回数が膨大な数になる。この問題を解決するため、ACF ではまずいくつかのスケールで特徴量を計算し、それらの間の補間処理により複数スケールの特徴量を高速に求める。

最後に、算出した特徴量をもとに、事前に構築した検出器を用いて歩行者か否かを識別する。

### 2.3. LDCF を用いた追跡

前フレームで検出された歩行者の周囲を探索し、各位置の検出窓について LDCF を算出する。前フレームの検出結果には、直前に ACF の検出結果が得られた場合は ACF の検出結果を用い、得られなかった場合は前フレームの LDCF の検出結果を用いる。LDCF (9) は、ACF の各チャンネルの特徴量に対して decorrelating filters を適用して特徴量の無相関化を行う

手法である。まず、ACF と同様の方法で画像から 10 チャンネルの特徴量を抽出する。各チャンネルの特徴量について、 $5 \times 5$  画素の局所領域毎に decorrelating filters を適用し、4 チャンネルの特徴量に変換する。具体的には、特徴量に対して主成分分析を行い、第 4 主成分までを decorrelating filters として取り出す。これらの処理により、40 チャンネルの特徴量が得られる。

LDCF は、ACF の特徴量にフィルタの畳み込み処理を行うため、ACF よりも若干処理時間が増加する。その反面、より識別に有効な特徴量が得られるため検出精度が向上する。

最後に、算出した特徴量をもとに、事前に構築した検出器を用いて歩行者か否かを識別する。このとき、歩行者が現実にはあり得ない位置に検出された場合、それは誤検出と考えられる。そのため、画像内の道路面と消失点を推定し、検出位置や大きさが適切か否かの判定を行い、最終的な検出結果を求める。

### 2.4. 体向き推定

検出された各歩行者の画像からその体向きを 8 方向に分類する。まず、検出時と同様の方法で、歩行者画像から ACF を算出する。算出した ACF をもとに、事前に構築した推定器を適用して体向きを求める。Random Forest による推定器は処理を並列化することで高速に推定が可能であるため、処理を並列化して体向きの推定を行う。

## 3. 評価実験

### 3.1. 実験目的

提案手法の有効性を示すため、車載カメラから得られた映像を用いて歩行者認識の実験を行った。本実験では、提案手法による歩行者検出と体向き推定が有効であることを示す。また、歩行者認識の処理速度を計測し、提案する検出と追跡の組み合わせによる検出速度の高速化を示す。

### 3.2. 実験環境

前方車載カメラとして、Grasshopper3 (Point Grey Research 社) を使用した。カメラから得られる画像の解像度は  $1,920 \times 1,440$  画素、フレームレートは 26fps である。カメラの設置位置は図 4 に示す。また、カメラから得た画像の例を図 5 に示す。

歩行者検出の学習用データとして、歩行者検出用の公開データセットである Caltech Pedestrian Dataset (10) の学習用画像に加え、独自に収集した画像を用いた。体向き推定では、独自に収集した画像に体向きのアノテーションを付与して用いた。

実験に使用した計算機は、Intel (R) Core i7-4930K, 3.40 GHz の CPU と実装メモリを 64 GB を搭載した汎用 PC である。





Fig.4 Installing position of the in-vehicle camera.



Fig.5 Samples of input images.

### 3.3. 実験結果

車載カメラ画像に対し、本手法を適用して歩行者を検出し、体向きを推定した結果を図 6 に示す。図中の青枠は歩行者の検出結果である。また、推定された体向きは赤の楕円と扇形の組み合わせにより示した。扇形が下向きときは手前を向いていることを、上向きときは奥を向いていることを示している。

図に示した結果より、歩行者を検出して体向きを推定できていることが確認できる。しかし、棒状の人工物では誤検出が、2 人の歩行者が重なっている場合には体向きの誤推定がそれぞれみられた。検出や体向き推定は歩行者の見えをもとに行うため、歩行者と見えの似た人工物の誤検出や、遮蔽等により見えが変化した歩行者の体向きの誤推定が生じたと考えられる。

歩行者検出および体向き推定を合わせた処理時間は、前述した計算機上にて、30fps であった。したがって、提案する歩行者認識がリアルタイムにて動作することを確認した。ACF は高速に計算可能だが、 $1,920 \times 1,440$  画素の大きな画像の探索には時間が掛かり、ACF を用いた検出のみでは 10fps 程度の速度であった。これに追跡を加えることで 30fps の速度で検出が可能となり、検出と追跡を組み合わせた提案手法が有効であることが確認できた。

### 3.4. 考察

提案手法は歩行者検出と体向き推定を組み合わせた認識を行った。体向き推定は前段の歩行者検出の結果を用いて行うため、検出時に枠が歩行者の体からずれたときに体向きの推定を誤る場合がみられた。この問題への対処として、体向き推

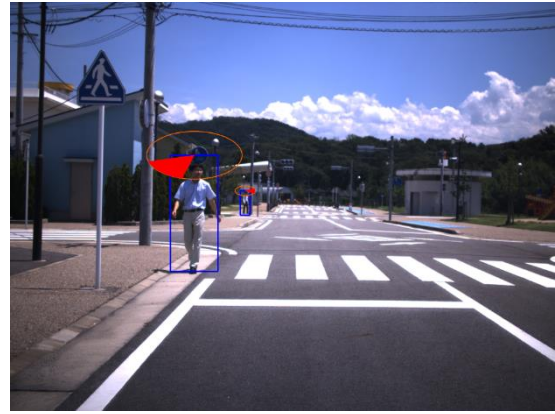


Fig.6 The results of pedestrian recognition.

定器の学習時に、実際の検出結果の画像を用いる方法や枠をずらした画像を生成して追加する方法も検討した。しかし、これらの対処方法では十分な解決にはならないと考えられる。この問題を解決するためには、検出された枠が歩行者を正確に囲うよう、枠の位置や大きさを補正する手法が有効と考えられる。

## 4. まとめ

本発表では、実車での動作を想定した歩行者検出および体向き推定手法を提案した。リアルタイムでの歩行者検出と体向き推定を実現するため、高速に計算が可能な ACF を特徴量として用いた。また、歩行者検出ではさらに高速化するために

追跡処理を組み合わせる手法を提案した。また、歩行者の動きの予測を行うため、検出された歩行者の体向きを推定した。これから起こり得る危険の予測という点から、属性認識は非常に重要である。

今後の展望として、前段の歩行者検出においてより正確な枠を取得する手法の検討が挙げられる。また、より詳細に歩行者の動きを予測するため、体向き以外の多様な属性の認識を行う計画である。

#### 謝 辞

本研究の一部は国立研究開発法人科学技術振興機構（JST）の研究成果展開事業「センター・オブ・イノベーションプログラム（名古屋 COI：高齢者が元気になるモビリティ社会）」の支援及び、科学研究費補助金の援助によって行われた。

#### 参 考 文 献

- (1) N. Dalal: Histograms of Oriented Gradients for Human Detection, Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol.1, p. 886-893 (2005)
- (2) P. Felzenszwalb, R. Girshick, D. McAllester, and D. Ramanan: Object Detection with Discriminatively Trained Part Based Models, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.32, No.9, p.1627-1645 (2010)
- (3) J. Hosang, M. Omran, R. Benenson, and B. Schiele: Taking a Deeper Look at Pedestrians, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, p.4073-4082 (2015)
- (4) S.X.M. Yang, P.K. Larsen, T. Alkjær, T.B. Juul-Kristensen, E.B. Simonsen, and N. Lynnerup: Height estimations based on eye measurements throughout a gait cycle, Forensic Science International, Vol.236, p.170-174 (2015)
- (5) Y.H.K.Y.H. Kwon and N. D.V. Lobo: Age classification from facial images, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (1994)
- (6) D. Baltieri, R. Vezzani, and R. Cucchiara: People orientation recognition by mixtures of wrapped distributions on random trees, Proceedings of the 12th European Conference on Computer Vision, p.270-283 (2012)
- (7) H. Fukui, T. Yamashita, Y. Yamauchi, H. Fujiyoshi, and H. Murase: Robust Pedestrian Attribute Recognition for an Unbalanced Dataset using Mini-batch Training with Rarity Rate, Proceedings of IEEE Intelligent Vehicle Symposium, MoPosterAT3.3 (2016)

- (8) P. Dollár, R. Appel, S. Belongie, and P. Perona: Fast feature pyramids for object detection, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.36, No.8, p.1532-1545 (2014)
- (9) W. Nam, P. Dollár, and J.H. Han: Local Decorrelation for Improved Detection, Proceedings of Neural Information Processing Systems Conference (2014)
- (10) P. Dollár, C. Wojek, B. Schiele and P. Perona: Pedestrian Detection: An Evaluation of the State of the Art, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.34, No.4, p.743-761 (2012)