

FIR-V ハイブリッドカメラを使った歩行者検知システム*

長谷川 辰雄**、菊池 貴**、宇都宮 弘純**、鎌田 智也***、加藤 浩之****

クルマの安全性を高める技術は急速に高性能化しており、車両に対する自動ブレーキなど、交通事故を未然に防ぐ予防安全技術の実用化が進んでいる。しかし、歩行者のカメラ検知では、歩行者の形が一様でないことによる未検知の削減と、背景と混同する誤検知の削減を両立することが難しく検知率の向上が課題であった。そこで、車載の遠赤外線カメラと可視光カメラの波長の異なった画像を組み合わせ、さらにオプティカルフローによる動体検出とパターン認識による機械学習の画像処理により、独自の評価データで歩行者検知率 95.2% を達成できた。

キーワード：歩行者検知、遠赤外線（FIR）カメラ、可視光（V）カメラ

Pedestrian-Detection System with Far-Infrared-Visible Hybrid Camera

**Tatsuo Hasegawa, Takashi Kikuchi, Kojoyun Utsunomiya,
Tomoya Kamata and Hiroyuki Kato**

In recent years, the performance of car-safety technology has rapidly improved and has led to the practical use of safety devices such as automatic brakes. However, detecting pedestrians by using cameras remains difficult because of their ill-defined shape, which leads to confusion between pedestrians and the background. In this report, we present an improved pedestrian-detection system. The approach is based on using images captured by cameras sensitive to two different wavelengths: far-infrared and visible light. The detection algorithm includes motion detection based on optical flow, and the image processing uses pattern recognition-based machine learning. The proposed algorithm achieves a pedestrian-detection rate of 95.2% with our evaluation datasets.

key words : pedestrian-detection, far-infrared camera, visible range camera

1 はじめに

交通事故防止の観点から、自動車の衝突被害軽減（プリクラッシュセーフティ）の有効性が世界的に認められている。車両周辺監視が衝突被害軽減に有効であり、その中でも歩行者検知の難易度は高く各国の技術開発が進んでいる。車両周辺監視のセンサはミリ波レーダとカメラが代表的であり、30～150mの遠方検知にはミリ波レーダを用い、30m以内の近距離検知にはカメラを用いる方法が有効とされている。より高度な衝突被害軽減のためには、歩行者の認識とその速度、進行方向を実時間に検知し、衝突を事前に予測する必要がある。しかし、ミリ波レーダでは歩行者自体の認識やその移動方向、複数の物体を分離する等の検出が困難であるため、現在ではカメラ画像を用いて、歩行者の特徴形状と移動方向を画像処理で求める方法で研究開発が行われている。

本研究は、遠赤外線カメラ（FIR：Far-Infrared camera）と可視光カメラ（V：Visible camera）のハイブリッドカメラ（FIR-V）を使って、歩行者を実時間

で検知可能な実用装置を開発し、衝突被害軽減などの高度な安全装置を開発する国内メーカーに技術（製品）を提供することを目的とした（図1）。開発目標は、車速50km/h以下で車両から30m以内にいる歩行者を検出率95%以上で検出することとした。

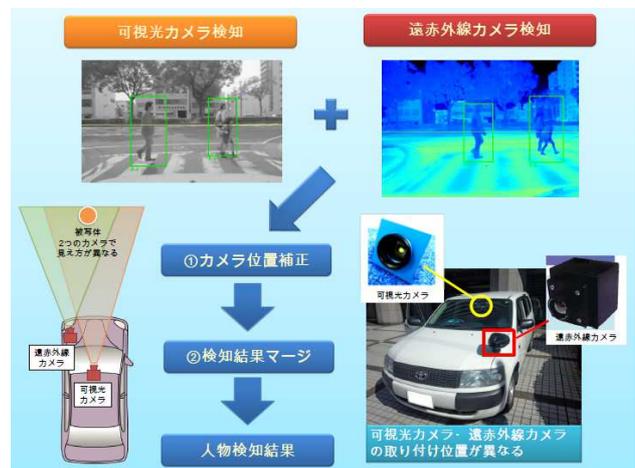


図1 ハイブリッド検知概要図

* 平成 24～26 年度 経済産業省 戦略的基盤技術高度化支援事業

** 電子情報技術部 *** アイエスエス株式会社 **** 萩原電気株式会社

目標達成の方法として、最大事後確率法とオプティカルフローを基盤とした動体検出及び、パターン認識手法を基盤とした学習による識別アルゴリズムを開発した。また、歩行者検出器に必要な遠赤外線(FIR)カメラと可視光(V)カメラで学習及び評価用の画像を録画し、それぞれから画像特徴量を抽出した歩行者検出器を構築した。歩行者画像を入力に HoughForest¹⁾処理を行った後、その Hough 投票のスコアが閾値以上の場合に歩行者と認識する手法を構築し評価を行った。

2 実験方法

2-1 歩行者検知の方法

近年では、計算機性能の向上に伴い、シーン映像からの人物検知を行う手法として、機械学習²⁾によるものが大きなウェイトを占めてきている。本研究においても最終的な人物領域の検出は学習器を用いたものとなるが、一般に学習器による検出計算は膨大な特徴量計算を要するために、実時間性の面で課題がある。本アルゴリズムの狙いは、初めに撮影シーンを解析することによって、歩行者が存在する可能性のある領域を特定し、次に、限定された領域にのみ人物検出処理を行うことで処理全体の計算量を大きく軽減することにある。シーン理解のための前段処理アルゴリズムは、計算時間を要する精度を期待するよりもむしろ、計算量の少ない単純な処理であることが望ましい。そこで、画像領域全体の密なステレオ計算を行うのではなく、疎な特徴点群のオプティカルフローに基づき3次元シーン推定を行う手法の開発を行った。図2は本研究の歩行者検知アルゴリズムの流れ図である。カメラ入力映像に対し、①レンズ歪み補正とカメラ取付け姿勢補正を行い、この補正画像に対し、自車両の進行方向を推測する運動推定³⁾を行った後、②立体物の領域を判定する。最後に Hough Forest 法による歩行者検知を実施する。



図2 歩行者検知アルゴリズム

2-2 自車両運動の推定

運動する車両に設置した単眼カメラにて、車両前方の路上シーンを撮影する場合、停車中の歩行者検出は監視

カメラ分野の背景差分法を応用することにより比較的容易に検出可能である。一方、移動車両の場合、背景を含む車両シーン全体が動くため背景差分法を使うことができない。このため、歩行者を検出する処理では、画面全体にわたって歩行者画像を探索する必要があり、組み込みシステムにおいては計算負荷が大きな課題となる。

本研究は、計算資源が制限される組み込みシステム上で歩行者検知処理のリアルタイム処理を実現するため、車両の運動とカメラの間で成立する幾何学条件を用いて、歩行者を含む立体物の領域を予め抽出し、その領域にのみ歩行者検出処理^{4, 5)}を適応する。これによって、HoughForest 法による歩行者検出処理を適応すべき領域面積を削減し、歩行者検出に要する計算コストを低減できるアルゴリズムを開発した。図3に自車両運動推定アルゴリズムと立体物判定アルゴリズムの処理の流れを示す。

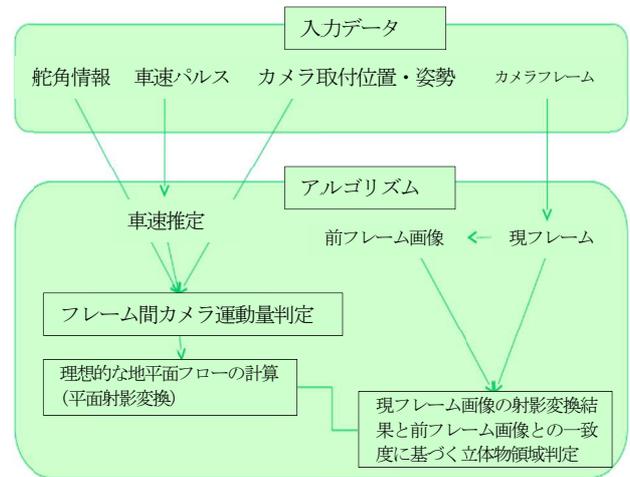


図3 自車両運動推定アルゴリズムと立体物判定

2-3 ハイブリッド方法

遠赤外線(FIR)カメラと可視光(VR)カメラのハイブリッド方法は、それぞれの Hough Forest 法の投票値を組合せて歩行者を認識する⁶⁾。投票値とはエッジ等の特徴量により人の中心座標に投票値を設定するものである。これまでは各投票値を加算した後、閾値より高い投票値を歩行者と定義していたが、この方法では片方のカメラが未検出の場合にハイブリッド法でも未検出となり検知率が低下する課題があった。そこで、図4に示す通り、片方のカメラが検出していれば正検知となる論理 OR 方式を検討し評価した。

高投票値の箇所に枠付

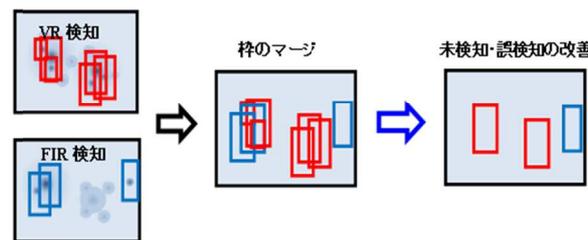


図4 ハイブリッド方法概念図

2-4 歩行者画像 DB の構築

実験車両により内部で利用する車両走行シーンの録画撮影を実施した。撮影した画像の中から良質の学習サンプルを選定し評価した。過去2年間の検知結果では、可視光カメラの検知性能が遠赤外線カメラよりも優れていた。そこで、遠赤外線カメラの検知率向上のため、遠赤外線画像 DB から良質な歩行者サンプルを学習成果の中から比較するために、これまでに使用した学習サンプル群から選別した。選別方針として不適格サンプル対象を以下の3項目により、目視により不適格な画像を削除した。

不適格サンプル対象

- (A) 歩行者の背後に直線状オブジェクトがある画像
- (B) 歩行者の中心がずれている画像
- (C) 歩行者が重複している画像

学習サンプルとして不適格と判断した画像例を図5に示す。遠赤外線画像は視認性がおもしろくないので、画像例はコントラストを明瞭化するイコライズ処理を施して選別した。



図5 不適格画像サンプルの例

2-5 リアルタイム検知

歩行者検知アルゴリズムの組込コード開発は、Xilinx 製 Zynq (以下 Zynq) を使用した。また、ユーザ企業が本研究成果を評価できるように、専用の評価ボードを試作した。一方、夏季の気温が高い中部地方の特徴を捉えるため、遠赤外線カメラと可視光カメラで学習用及び評価用の路上映像の録画を内部で実施した。過去2年間で、Zynq 上での開発環境構築、専用評価ボードの製作、歩行者検知コードのハードウェア化を実施した。しかし、リアルタイム (33ms) で動作させるには、メモリ構造やタイミングの設計など難易度の高い課題が残った。そこで、高速処理化のためにベアメタル (OS レス) 化や領域マスクの工夫を行い、Zynq に歩行者検知コードを実装し評価した。同ボードに実装した歩行者検知コードの構成図を図6に示す。

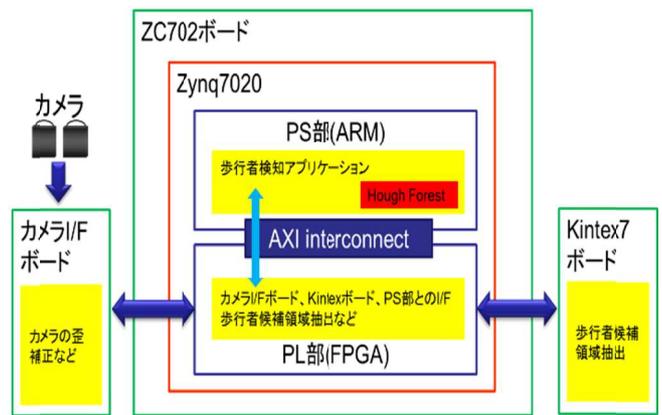


図6 ハードウェアブロック図

2-6 ハイブリッド位置合わせ

可視光カメラと遠赤外線カメラのそれぞれの特性を考慮し、可視光カメラはバックミラー付近、遠赤外線カメラは車外に設置した。ハイブリッドは両カメラ画像の位置合わせが必要であるが、両カメラの間隔は約1mと距離が離れ、また画角も多少異なるレンズを用いていることから、位置合わせの誤差が大きくなる課題があった。これに対し、両カメラ画像の較正⁷⁾ (キャリブレーション) の工夫と、ルックアップテーブル (LUT) を使った位置合わせにより、ハイブリッド化の精度向上を狙った。

過去の実験では、遠赤外線カメラと可視光カメラのハイブリッド化のためのキャリブレーションと、両カメラの位置合わせを線形的な幾何学的手法によって検証したが、本研究の各カメラは、画角の違いや、取り付け3次元座標 (x, y, z) の差が大きく、画面の位置で変換誤差が大きくなる課題があった。具体的にはキャリブレーション精度及び位置合わせの精度の向上である。この解決には、従来の線形変換ではなく画面の位置によって変化量を変えるような非線形の変換が必要となる。そこで、両カメラの位置合わせ手法は、図7に示す通り人手で設定した可視光カメラの歩行者座標と遠赤外線カメラの歩行者座標を1対1に対応させたルックアップテーブル (LUT) により行った。対応座標が設定出来ない箇所は非線形方程式による補完により対応座標を推定した。



図7 可視光カメラの歩行者座標 (左図の赤枠) と遠赤外線カメラの歩行者座標 (右図の赤枠)

3 実験結果

3-1 ハイブリッド歩行者検知結果

ハイブリッド手法による歩行者検知は、学習パラメー

タの最適化、学習サンプルの厳選、フィードバック学習の各技術を発展させ、効果的に活用したことで誤検知・未検知の低減に成功し、論理 OR 方式のハイブリッド手法により目標の検知率 95.2% (評価指標: FPPI=1) を達成することができた (図 8)。FPPI(False Positive Per Image)=1 とは、1 フレームに 1 個の誤検知を許容した場合の歩行者検知率を意味する。可視光カメラ単体では 94.8%、遠赤外線カメラ単体では 96% の検知結果となった。評価方法は、内部で撮影した評価映像の 17 種類 (3,946 フレーム) に対し、正解画像と検出画像を照合する自作検証ツールを使用して評価した。夜間 (遠赤外線カメラのみ、20 種類: 4,043 フレーム) を含む、昼夜間 (24H) の評価では、全 37 種類 (7,989 フレーム) に対し、95.9% (FPPI=1) の成果が得られた。また、このときの結果画像の一部を図 9 に示す。

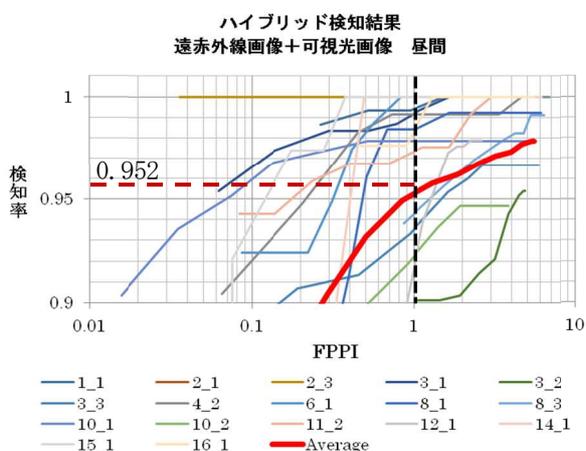


図 8 ハイブリッド歩行者検知結果

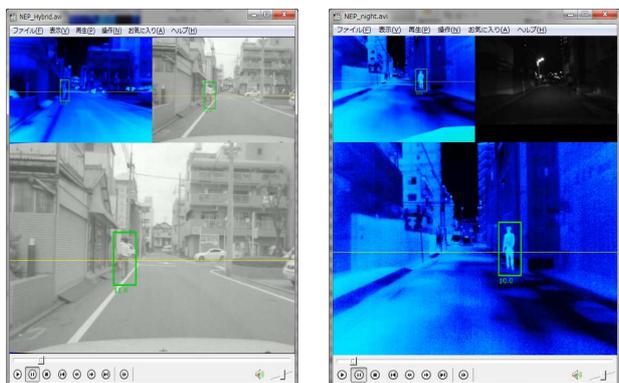


図 9 ハイブリッド検知結果画像

3-2 リアルタイム検知

歩行者検知アルゴリズム(Hough Forest)を Zynq 内の ARM へ実装した。表 1 に示すように Zynq ボードの動作環境を構築した上で実装した。Zynq ボードの動作環境は、カメラ入力部と演算負荷を低減させる FPGA ボードをケーブルで接続した構成となっている (図 10)。しかし、検知率を向上させるためにコード量が増加したことが原

因で、目標であった処理速度の 0.033 秒(30fps)に対し、0.177(5.6fps)の結果にとどまった。今後、高速化アルゴリズム検討と高速 CPU の利用など、継続して目標の 30fps を目指す予定である。

表 1 Zynq ボード動作環境

基板/開発ソフトウェア	バージョン
ZC702 Evaluation Kit	—
Xilinx SDK	2014.2
OpenCV	2.3.1



図 10 Zynq ボードの動作環境図

3-2-1 高速化の検討

Linux 等の OS 処理を行わない (OS レス) と、画像の拡大縮小処理を 1 段階とする高速化を実施したが、大きな成果が得られなかった。そのため、高速化の対策を行うために、歩行者検知器のパラメータを変化させながら、実際の処理時間を計測した。この結果を表 2 に示す。

表 2 歩行者検知の処理速度結果

検出処理スケール倍率	学習木データファイル数	処理時間 (秒)	プラットフォーム
1.0~3.0 7段階	18	30.4	Zynq(Linux)
2.0~2.0 1段階	18	3.8	Zynq(Linux)
2.0~2.0 1段階	6	1.4	Zynq(Linux)
1.0~3.0 7段階	18	26.7	Zynq(ベアメタル)
2.0~2.0 1段階	18	3.3	Zynq(ベアメタル)
2.0~2.0 1段階	6	1.3	Zynq(ベアメタル)
1.0~3.0 7段階	18	0.7	Windows(※1)

この OS レス化により約 10~15% 程の高速化ができたが、目標のリアルタイム化 (33ms) には不十分であった。また、同時に検出スケールと学習木の木数で調整を試みたが、学習木の数を減らした場合、検知率が低下することが分かったため検出スケールを固定化する等のいくつかの対策を実施し、さらなる高速処理化に取り組んだ。

3-2-2 高速化の結果

検知率の結果に影響を与えず、且つ、高速化できるかという観点で検討し、次の 2 項目を実施した。

- ① 1 画素ごとの処理を 4 画素間引く
- ② 不要な左右の領域を 50 画素省く。

この結果、表 3 に示す通り、ソフトウェア (Zynq) とハ

ードウェア (FPGA) の統合処理で、処理時間 0.177 秒と高速化ができた。

表 3 処理速度比較

条件	プラットフォーム	処理時間(秒)
PC版	Windows	0.7
Zynqのみ実施	Zynq(Linux)	30.4
ソフト高速化1	Zynq(ベアメタル)	3.3
ソフト高速化2	Zynq(ベアメタル)	0.295
ソフト・ハード統合高速化	Zynq(ベアメタル)+FPGA	0.177

3-3 ハイブリッド位置合わせ

ハイブリッド検知は FIR カメラと VR カメラの位置合わせ精度が重要であるが、これまでの平面射影行列による歩行者足元の3次元座標変換では、誤差が大きくなる課題があった。これに対し、ステレオ平行化による位置合わせを実施したが、両カメラの画角の違いと視差の大きさが影響し誤差縮小はできなかつた。そこで、遠赤外線画像と可視光画像の対応する歩行者の座標を活用した LUT を新たに構築することで、この問題の解決を図った。両カメラ間で座標を変換した際の誤差について図 11 に示す。

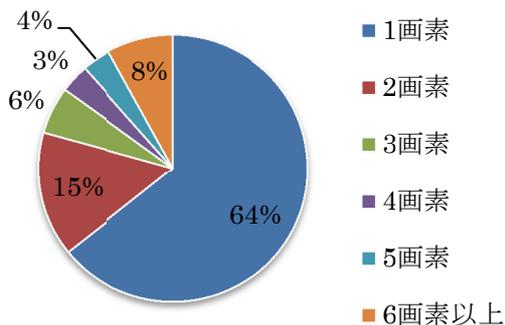
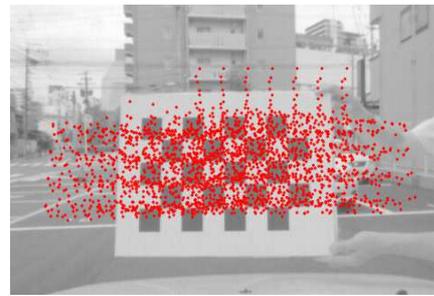
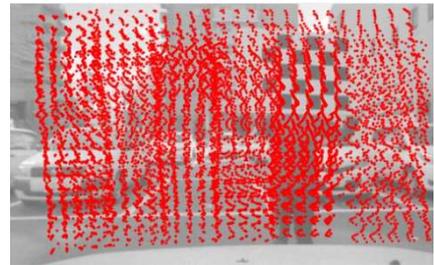


図 11 位置合わせ誤差の割合

これにより、従来では 50 画素あった位置合わせの最大誤差を 27 画素に縮小することができた。また評価データ全体の 97.8%を±10 画素以内、92.6%を±5 画素以内に収めることができた。この位置合わせ誤差の縮小がハイブリッド検知率の向上に寄与したと考える。また、従来の加算式ハイブリッド法に対し、論理 OR 方式にすることでハイブリッド検知率 95.2%の達成に貢献することができた。初回実験のキャリブレーションは、平面射影モデルの Zhang の手法を用いて、遠赤外線画像 45 枚、可視光画像 81 枚のキャリブレーション画像を用いてカメラ内部パラメータ (焦点距離、中心座標、歪補正係数) を求めた。しかし、図 12(a) の通りカバーしきれない領域が存在し、画像中心と歪補正の誤差が大きくなり、結果的にハイブリッド化の検知率が向上とはならなかつた。そこで、図 12 (b) の通り画面全体をカバーするようにキャリブレーション座標を取得することで、中心座標と



(a) 81 枚



(b) 558 枚

図 12 可視光カメラのキャリブレーション

レンズ歪補正の誤差を縮小化した。

図 13 は FIR 画像のキャリブレーションの原画像と歪補正画像で、VGA (640 × 480) の中心座標を基準点 (320, 240) としたとき、キャリブレーションで求めた中心座標 (329, 225) との距離を示している。446 枚の計算によって、距離誤差 44pixel から 17pixel へ精度向上することができた。遠赤外線カメラは画角が広くレンズ歪も大きかったことから、歪補正の効果が大きかったと考察する。可視光カメラは 558 枚のデータ枚数のキャリブレーションで、46pixel から 28pixel へ誤差の改善ができた。

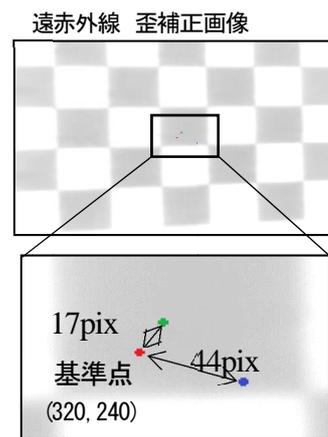


図 13 位置合わせ座標変換の誤差

4 考 察

遠赤外線カメラは、物体が放射する温度を画像化するため、可視光 (環境光) の影響を受けない画像が得られる。可視光線と遠赤外線の異なる波長の画像を比較すると、可視光カメラで確認できる歩行者の服の模様や色が、遠赤外線カメラでは確認できない事など、遠赤外線カメ

ラの情報量は少ない。しかし、可視光カメラでは捉えることの出来ない温度の情報を把握できる特徴がある。それぞれのカメラの特徴を組み合わせることで、歩行者検知の性能を上げることが出来た。図14は左上が遠赤外線カメラの結果、右上が可視光カメラの結果、中央画面が両カメラを使ったハイブリッドの結果である。可視光カメラでは看板が誤検知(赤枠)となっているが、遠赤外線カメラでは誤検知は無く、両カメラのハイブリッドの結果は歩行者のみを検知している。これは両カメラの特徴を効果的に利用できた結果と考える。一方で、遠赤外線カメラは熱画像であるため、外気温が高い夏場など、背景と歩行者が同じ温度となる「とけこみ」現象や、歩行者よりも背景温度が高くなる逆転現象などの課題がある。この課題の対策として、外気温が高い場合は可視光カメラを優先させることが有効と考えられる。図15は夜間の遠赤外線カメラのみの検知結果である。このように夏場でも、夜間は外気温が低下するため「とけこみ」現象は少なくなり検知率は向上するため、夜間は遠赤外線カメラを優先させるなど、環境温度の違いによって優先するカメラを選択することが検知率向上に効果的であると考える。

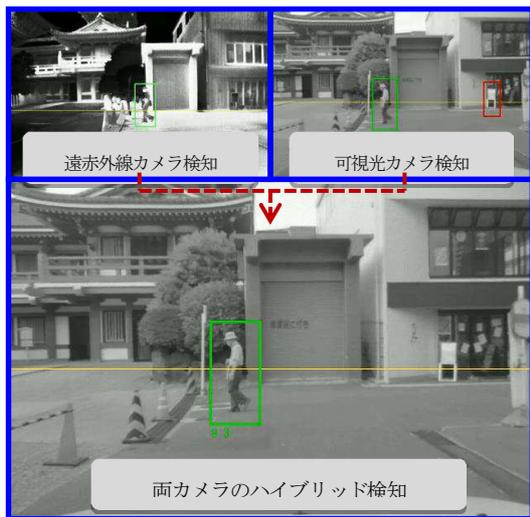


図14 誤検知が除去できた結果



図15 夜間の遠赤外線カメラ検知結果

5 結 言

FIRカメラとVRカメラによるハイブリッド検知は、波長の異なる画像を組み合わせることで相乗効果を発揮し検知率を向上させた。17種類(3,946フレーム)の評価用映像に対し検知率95.2%(目標95%以上、条件FPPI=1)を達成した。

しかし、電子基板(Zynq)での目標30fps稼働は、実績値5.6fpsの結果にとどまった。これは、歩行者検知率95%以上達成のため、プログラム量とデータ量が増加したことが原因であった。今後は、高速化の手法を再検討し、継続して30fpsの実現へ取り組む。これには、第一にアルゴリズム自体の高速化が必須と考え、最新の高速化アルゴリズムを検討しながら高速化に取り組んでいく予定である。事業化へ向けては、FIRカメラによる夜間の歩行者検知が有力と考えており、検知率も高い水準であることから実用化のアピール点である。車載歩行者検知は実用化されているが、未だ十分な性能に至っていない。このことから、車載センサによる歩行者認識の難易度が高く、解決すべきキーテクノロジーであることは明らかであり、本研究をさらに推進させることで、国内外でも競争力の高い製品(技術)へ展開が十分期待できると考える。

文 献

- 1) J Gall, V Lempitsky : Class-specific Hough forests for object detection, In CVPR2009, p1022-1029 (2009)
- 2) David Geronimo, Antonio M. Lopez : Vision-based Pedestrian Protection Systems for Intelligent Vehicles, SpringerBriefs in Computer Science (2014)
- 3) Hernan Badino, Uwe Franke, David Pfeiffer : The Stixel World ? A Compact Medium Level Representation of the 3D-World, 31st DAGM Symposium on Pattern Recognition (2009)
- 4) Olga Barinova, Victor Lempitsky, Pushmeet Kohli : On Detection of Multiple Object Instances using Hough Transforms, IEEE Conference on Computer Vision and Pattern Recognition (2010)
- 5) Paul Viola, Michael J. Jones : Robust Real-Time Face Detection, International Journal of Computer Vision, p137-154 (2004)
- 6) 岩元、菅谷、金谷 : 3次元復元のためのバンドル調整の実装と評価, IPSJ SIG Technical Report (2011)
- 7) Zhengyou Zhang : A flexible new technique for camera calibration, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11), p1330-1334 (2000)