

12-1 はじめに

自動車による事故を予防し、また将来の自動運転を実現するためには道路あるいは人や他車を認識し、停止や回避などの適切な行動を取ることが求められる。これらの機能を実現するための技術は、自律移動ロボット分野において研究されてきたものとほぼ同一である。そこで、本章では予防安全・自動運転に適用可能なロボットの知能化技術として、人物の検出技術および地図生成技術について述べる。

12-2 人物検出

(1) 人物検出の重要性

自動車の安全性向上のためには歩行者を高信頼に検出することが不可欠である。画像による人物検出手法については、近年の画像の局所特微量と学習アルゴリズムの発展により、照明条件が適切であればかなりの検出精度が得られることが示されている。しかしながら、自動車で利用するためには、人の運動方向を予測するための向きと移動方向の推定や、適切な回避行動を取るための3次元情報の取得などが必要となる。本節では、人物検出・状態推定のために開発した手法について述べる。

(2) 画像情報を利用した人物の向きと移動方向の推定 [1]

図12-2-1に処理のブロック図を示す。処理は大きく2つに分かれる。検出部では、まずHOG特徴[2]とAdaboostを用いた検出器[3]を上半身検出用に学習させ、画像中より上半身を検出する。検出した上半身領域に対し、形を捉えるHOG特徴およびテクスチャを捉えるLBP特徴[4]を計算し、合わせて特徴ベクトルとする。人の向きは図12-2-2に示す8方向に量子化して推定する。

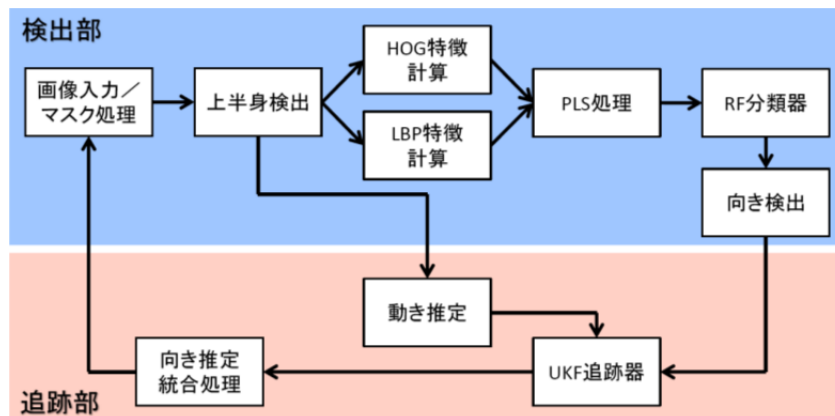


図12-2-1 人物の向き・移動方向推定処理のブロック図



図12-2-2 推定する8方向

特徴ベクトルの次元削減には部分最小2乗法 (Partial Least Squares, PLS) [5]を用いる。PLSは入力空間 (特徴ベクトル) におけるデータの分布だけでなく、出力空間 (クラスラベル、ここでは離散化された人物の向き) との相関を考慮して部分空間を作成するため、より識別性能の高い部分空間を作成することができる。PLSによって得られた特徴ベクトルを用いて向きの分類を行う。分類器には多クラス分類が可能な Random Forest[6]を用いる。

歩行中の人物は通常進行方向を向いているので、移動方向と向きとの間には関連がある。そこで、両方の情報を統合して向きを推定するために、人物の位置、速度、向きを状態変数とする Unscented Kalman Filter を利用した追跡器を構成し、情報統合を行う (図 12-2-1 の下のブロック)。また、人物追跡情報を利用して画像中の人物検出範囲を限定することによって処理時間の短縮を図る。

図 12-2-3 に TUD-Stadtmitte データベース[7]を用いて人物検出および向き推定を行った例を示す。図中、上段は人物の検出結果 (円筒) と向きの推定結果 (矢印) を示す。下段は 2 次元座標系におけるカメラ視野と検出された人物位置を示す。他のデータベースを用いた比較において、従来手法に比べて人物検出および向き推定の精度が向上していることが示されている[1]。



図 12-2-3 TUD-Stadtmitte データベースに対する人物検出・向き推定結果の例

(3) 画像と 2 次元距離センサデータを用いた人物の追跡と認証 [8]

画像データは人物の識別に有効であるが、見えのみの情報では誤った領域を人物としてしまう可能性がある。そこで、レーザ距離センサ (Laser Range Finder, LRF) を用いて人物を検出し、さらに両方のデータを組み合わせて人物認証を行う手法を開発した。図 12-2-4 に処理のブロック図を示す。

LRF により人物の胴体部、脚部の高さのデータを取得し、得られた領域の形状から人物を検出し、ロボットからの相対位置を計測する。さらに、ロボットの移動を考慮した足の接地位置の推定により、歩容 (歩幅、歩行周期) に関する情報を得る。一方、LRF から得られた人物位置情報と前項で述べた

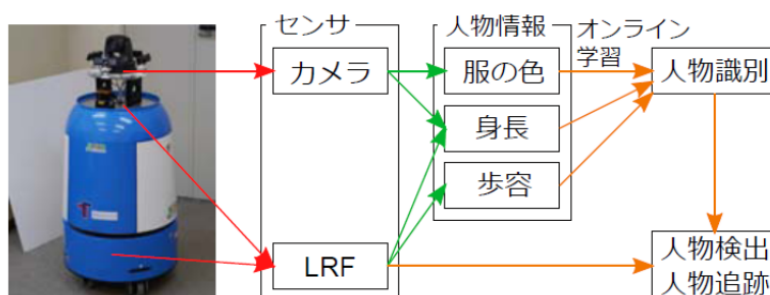


図 12-2-4 カメラとレーザ距離センサを用いた人物追跡・識別

画像による人物検出を組み合わせて画像中の人物位置を決定した後、服の色および身長を計算する。身

長の計算には、画像中から得られた頭頂部位置と LRF から得られる人物までの距離を利用する。最終的に服の色、身長、歩容の情報を統合して特定の人物を検出し、追従行動を取る。

頭頂部の検出は髪色（黒色）らしい領域と輝度勾配が強い領域の論理積を取り、人物領域内でその最上位の場所を選択することによって行う（図 12-2-5 参照）。LRF による足の接地位置の検出は、図 12-2-6（左）に示すように時間ごとの脚位置を 2 次元平面上に投票し、その投票値の高いところを接地位置とみなす。ロボットと人の位置関係によっては、片方の足でもう片方の足が隠蔽されるため、隠蔽の有無を判断し隠蔽側の接地位置の信頼性を低く設定した上で、歩幅、歩行周期を推定する（図 12-2-6（右））。

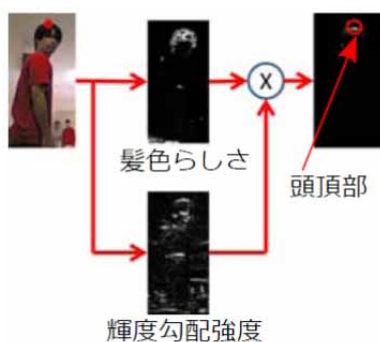
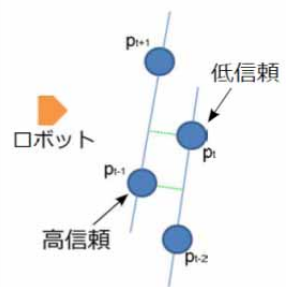
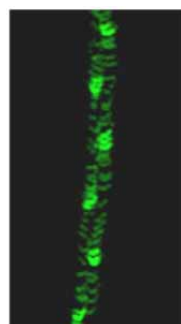


図 12-2-5 頭頂部検出処理



脚位置データの蓄積 軸足位置の信頼性
図 12-2-6 接地位置の検出

図 12-2-7 に示す状況でそれぞれの人物の追跡対象人物モデルとの類似度を計測する実験を行った。右側の人物が追跡対象人物であり、左側の人物（通行人）との比較を行った。なお、画像中の紫と緑の折れ線はそれぞれ、胴および脚の位置の LRF の距離データをカメラ画像上に射影したものである。図 12-2-8 に計測された各人物の身長、歩容データを示す。いずれも識別に十分な情報を持っていることがわかる。



図 12-2-7 実験状況

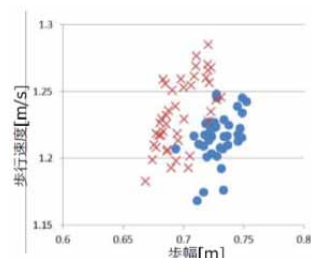
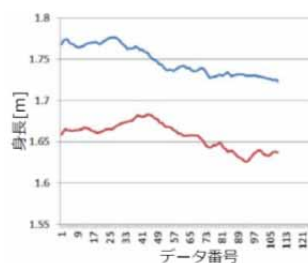
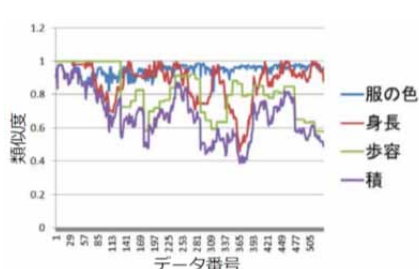
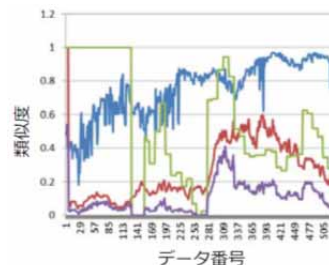


図 12-2-8 追跡対象（赤）と通行人（青）のデータ

図 12-2-9 に各特徴から得られる追跡対象人物らしさ（尤度）とすべての情報を統合した尤度の時間変化を示す。特定の情報で区別が付きにくいときでも、情報統合により正しい人物に対し高い尤度が計算されている。



(a) 追跡対象



(b) 通行人

図 12-2-9 追跡対象人物らしさ（尤度）の変化

（４）３次元距離センサを用いた人物の発見と追跡 [9]

前項で用いた距離センサ（LRF）は 2 次元平面上を走査するものであったが、3 次元空間を走査す

るタイプの距離センサを用いると人物の形状を認識することが可能である。ここでは、高品位 3 次元 LRF を用いて人物検出を行う城殿らの手法 [10] を、より解像度の低い 3 次元 LRF に適用する。

人物検出の手順を図 12-2-10 に示す。取得した 3 次元データ(a)から路面部分のデータを除去し(b)、物体ごとにクラスタリングした後その大きさから人物候補を検出し(c)、さらに識別器によって人物を検出する(d)。識別には、従来研究で用いられてきた 3 次元的な点の分布などに加え、特に遠方でデータ点数が少ない場合に対処するために考案した、大まかな形状を捉えるためのスライス特徴および衣服や金属ポール等材質の違いを捉えるための反射強度ヒストグラムを用いる。識別器には RealAdaBoost [11] を用いた。図 12-2-11 に検出例を示す。800 枚の距離データに対して検出実験を行い、再現率 0.90、適合率 0.94 であった。なお、図に示す移動ロボットシステムは、対象人物が保持する発信機の電波到来方向をエスパアンテナ[12]を用いて推定し、一時的に対象人物が視野から外れても再発見を行った後、追従行動を再開・継続することができる [9]。

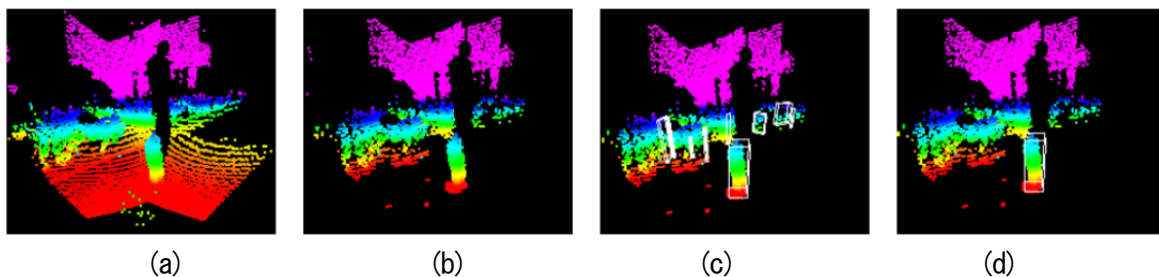


図 12-2-10 3 次元距離センサデータからの人物検出

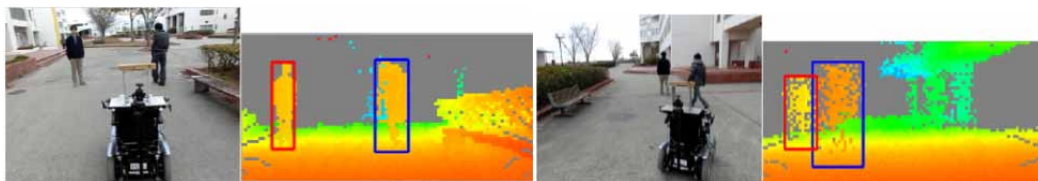


図 12-2-11 人物検出例

同様の手法を電気自動車に設置した距離センサに適用した。2 次元 LRF を 1 軸回転機構で回転させることにより、3 次元距離データを取得する (図 12-2-12 参照)。147 枚の距離データに対して検出実験を行い、再現率 0.87 であった。また、自動化タクシーへの適用を想定して、合図 (手を上にあげる (手上げ)、手を横に出す (手差出し)) の検出を試みた (図 12-2-13 参照)。現在、自転車など歩行者以外の移動体も含めた検出システムを開発している。



図 12-2-12 3 次元距離データ取得システム

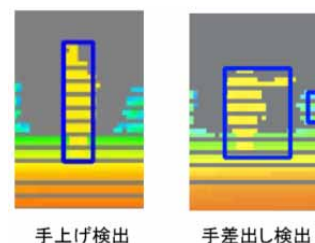


図 12-2-13 合図検出結果

12-3 屋外地図生成 [13]

不整地や障害物の存在する道路上を移動する際には、通行可能領域の検出が不可欠である。ここでは、そのための技術として屋外の3次元地図生成手法を開発している。屋内環境でのSLAM（地図と移動量の同時推定）技術はかなり成熟した段階にあるが、屋外では移動量推定に有用な構造物が常にあるとは限らない。GPSは有用な位置情報を提供するが、逆に構造物の近くでは十分な位置推定情報が得られない。そこで、場所に応じて適切な移動量推定手法を選択するとともに、閉ループ検出（以前通った場所に戻ってきたことの検出）やGPSによる相対位置情報に基づいて移動量修正を行うことにより、移動経路を推定して3次元地図を生成する手法を開発した。センサとしては3次元レーザ距離センサを主に使用し、2つ並べて使うことにより約120度の視野角を確保している。図12-3-1は取得した点群データの例である。

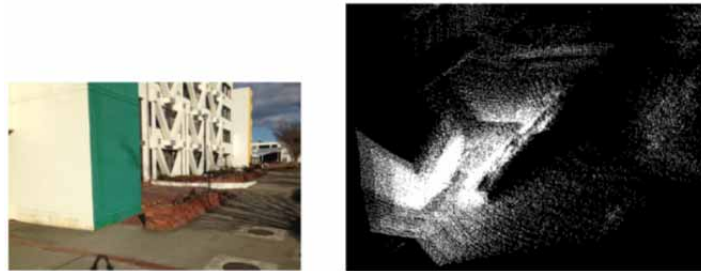


図12-3-1 点群データの例：（左）計測環境 （右）取得した点群

提案手法では、移動量推定の方法として、平面の対応付けに基づく手法、画像特徴点の対応付けとその3次元位置推定に基づく方法、ステレオ視覚オドメトリ、車輪オドメトリの4種を選択的に利用している。図12-3-2に提案する移動量推定手法を用いた地図生成結果を示す。地図の青色は移動可能な平面、赤色は建物や木などの障害物を示す。

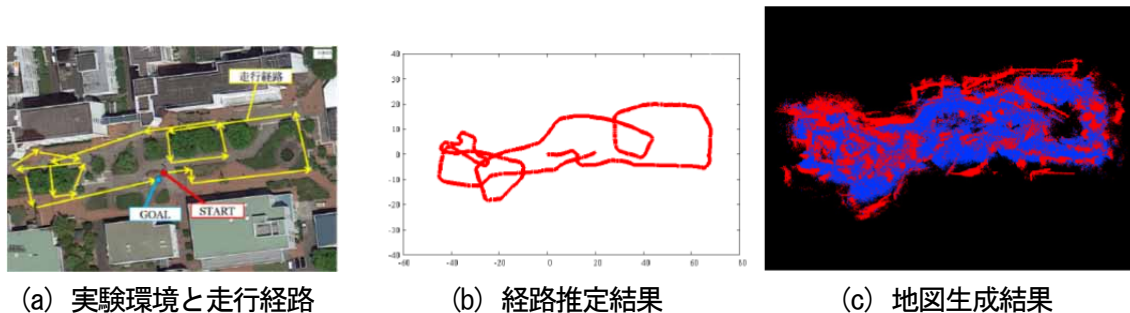
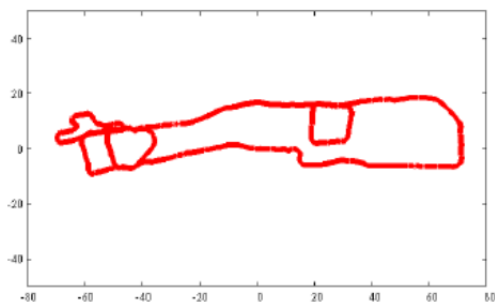
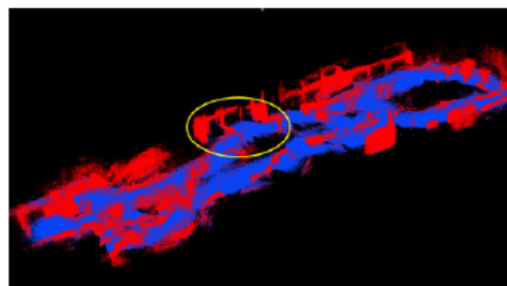


図12-3-2 選択的移動量推定に基づく経路推定と地図生成

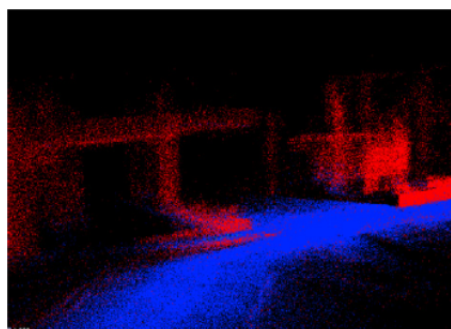
図12-3-2の結果ではおおよその経路と地図は復元されているものの細部が正確ではない。そこで、SURF特徴を利用したBag-of-visual-words法[14]により時間的に異なる2つの計測位置がほぼ同じ位置にあることを検出し、その2点間の相対移動量を上記移動量推定によって求めて得た情報、およびGPSデータから異なる2地点間の相対移動量を求めた情報の二種類の情報を追加することによって、経路推定および地図の改善を図る。改善には計測位置をノード、位置間の関係をエッジとするpose graphを生成した後に最適化を行う、pose graph SLAM[15]を適用する。図12-3-3に修正した経路推定と地図生成の結果を示す。



(a) 修正された移動軌跡



(b) 生成された3次元地図



(c) (b)の黄色枠部分の拡大図



(d) 各大部分の実際の環境

図 12-3-3 修正した経路推定と地図生成の結果

12-4 おわりに

本稿では、予防安全あるいは自動運転に不可欠な環境認識技術として、人物の検出および地図の生成について述べた。そこでは、複数種類のセンサデータや時系列データなどを統合して解釈すること（センサフュージョンと呼ぶ）により、認識の信頼性を向上させるアプローチが有効である。移動ロボットは自動車に比べより多様な環境を認識する必要があるが、移動速度が遅いため処理にある程度時間をかけられること、人物個別の情報を電波等によって得ることも可能であることなど有利な点もある。一方、自動運転では可能な環境・状況は一般に道路上に限定できるが、移動速度が大きいため高速処理や遠方のデータの処理が必要となる。これらの違いを十分に認識しつつ、技術の共通化と個別化を行っていく必要があると考えられる。

なお、本稿で紹介した研究を筆者と共同で行った豊橋技術科学大学行動知能システム学研究室の Igi Ardiyanto, 谷地中宏基, 三栖一城, 小出健司, 波平勇気の各君に感謝する。

参考文献

- [1] I. Ardiyanto and J. Miura, "Partial Least Square-based Human Upper Body Orientation Estimation with Combined Detection and Tracking", Image and Vision Computing, Vol. 32, No. 11, pp. 904-915, 2014.
- [2] N. Dalal and B. Briggs, "Histograms of Oriented Gradients for Human Detection", Proc. 2005 IEEE Conf. on Computer Vision and Pattern Recognition, pp. 886-893, 2005.
- [3] Q. Zhu et al., "Fast Human Detection using a Cascade of Histograms of Oriented Gradients", Proc. 2006 IEEE Conf. on Computer Vision and Pattern Recognition, pp. 1491-1498, 2006.
- [4] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 24, no. 8, pp. 981-987, 2002.
- [5] W.R. Schwartz et al., "Human Detection using Partial Least Squares Analysis", Proc. 12th

- IEEE Int. Conf. on Computer Vision, pp. 24-31, 2009.
- [6] L. Breiman, "Random Forest", Machine Learning, Vol. 45, pp. 5-32, 2001.
 - [7] M. Andriluka, S. Roth, and B. Schiele, "Monocular 3D Pose Estimation and Tracking by Detection", 2010 IEEE Conf. on Computer Vision and Pattern Recognition, pp. 623-630, 2010.
 - [8] 小出健司, 三浦純, "服の色・身長・歩容特徴を用いた付き添いロボットのための人物識別", 第 32 回日本ロボット学会学術講演会, 2014.
 - [9] 三栖一城, 三浦純, "3 次元レーザ距離センサとエスパアンテナを用いた移動ロボットによる特定人物発見・追跡", 第 32 回日本ロボット学会学術講演会, 2014.
 - [10] 城殿清澄, 渡邊章弘, 内藤貴志, 三浦純, "高解像度レーザレーダによる歩行者識別", 日本ロボット学会誌, Vol. 29, No. 10, pp. 963-970, 2011.
 - [11] R.E. Schapire and Y. Singer, "Improved Boosting Algorithms Using Confidence-Rated Projections", Machine Learning, Vol. 37, No. 3, pp. 297-336, 1999.
 - [12] H. Kawakami and T. Ohira, "Electronically Steerable Passive Array Radiator (ESPAR) Antennas, IEEE Antennas and Propagation Magazine, Vol. 47, No. 2, pp. 43-50, 2005.
 - [13] 谷地中宏基, 三浦純, "移動ロボットによる屋外 3 次元地図生成と通行可能領域の分類", 第 32 回日本ロボット学会学術講演会, 2014.
 - [14] C. Csurka et al., "Visual Categorization with Bags of Keypoints", Proc. ECCV-2004 Workshop on Statistical Learning in Computer Vision, pp. 1-22, 2004.
 - [15] E. Olson, J. Leonard, and S. Teller, "Fast Iterative Alignment of Pose Graphs with Poor Initial Estimates", Proc. 2006 IEEE Int. Conf. on Robotics and Automation, pp. 2262-2269, 2006.