ガイドロボットのための複数移動人物の検出手法の検討

A Study on a Detection and Tracking Method of Multiple Moving Persons for a Guide Robot

○佐々井 拓也 高橋 陽 小谷 充範 正 中村 明生 (東京電機大)

Takuya SASAI, Tokyo Denki University, sasai@is.fr.dendai.ac.jp
You TAKAHASHI, Mitsunori KOTANI and Akio NAKAMURA, Tokyo Denki University,

One of necessary functions for a guide robot is to detect persons around it. We propose the method that enables a robot to detect multiple moving persons even if the robot is also moving. The robot has an omni-directional camera and a laser range finder. First, skin-color and hair-color extraction based on image processing technique is done and the multiple candidate regions of persons are obtained. Then, template matching is applied to range data corresponding to the candidate regions. Thus, the robot can determine the exact person regions. Experimental result basically shows the effectiveness of the proposed method.

Key Words: Image Processing, Guide Robot, Laser Range Finder, Omni directional Camera, Human Detection

1. 緒言

近年, 我々の生活する環境内でロボットを日にする機会が増えている. 施設案内や人物誘導などを行うロボットは“ガイドロボット”と呼ばれており, 施設での実証実験を行う例や既に施設で運用されているものもある[1][2]. ガイドロボットに関する研究は様々な企業や研究機関により進められており, 今後発展が見込まれる分野である[3]. このような人間を相手として活動するガイドロボットに必要な機能を「周辺環境の認識機能」, 「人物の検出機能」, 「インタラクション機能」の3点に分類する.

ガイドロボットが施設案内や人物誘導を行う際には対象となる人物を発見できなければならない. このことから「人物の検出機能」が分類の中では最も重要であると考え, 本研究で検討を行う. ガイドロボットが活動する環境には, ロボット周囲の様々な方向に複数の人物が存在する. これらの人物がガイドロボットと何らかのインタラクションを行う際には, Hall が提唱する Personal space[7]より, Social space 内 (約 1.2-3.5[m]) で行われると考えられる. そこで, ロボットは 3.5[m]以内の全ての方向の人物を検出できなければならない. また, 円滑な施設案内・人物誘導のためにはロボットが停止することなく移動しながら人物検出を行う能力が必要である. これらの条件をふまえ, 本研究では以下を要件とした人物検出システムの検討を行う.

- 広範囲の観測が可能であること
- 移動中に人物検出が可能であること
- 複数の人物を検出可能であること

2. 提案システム

2.1 使用するセンサ種別

まずシステムの要件より, 検出に用いるセンサは全方向を観測可能でなければならない. 移動ロボットに関する研究の多くでは, 広範囲を観測する際に周辺の照明条件の変化に対して頑強な LRF (Laser Range Finder)を用いている. LRF を用いて人物を検出する手法として, 人物の足のスキャンデータをテンプレートとし, センサデータに対してテンプレートマッチングを行ことで人物を検出する方法がある[4]. しかしながら, 一般環境中に類似した形状の物体が多数存在した場合, LRF 単体では判別が困難である. このような問題は, 壁など移動が発生しない物体の領域を除外することで軽減することが可能である. 移動のない領域の取得には, センサデータを

数フレーム間で統合し, 変化の無い領域を除外する背景差分法の利用や, 事前にロボットに環境中の地図情報を与えるなどの手法がある[4]. しかし背景差分法はセンサが固定されていなければならない, 移動するロボットからでは静止した背景情報を得ることが困難である. エンコーダなどにより得たロボットの移動データを基に背景情報を除外することも可能であるが, タイヤのスリップなどのセンサで観測不可能な外部要因により誤差が発生する. カメラからの画像データから背景情報を取得する際にはこの誤差の影響を強く受け, 正常に背景情報を取得することが不可能である. また, 事前にロボットに環境地図を与えておいた場合, 壁などの背景情報は容易に除去することが可能であるが, 静的なオブジェクト以外を除外することが不可能であることから, 常に状況の変化する一般環境中においては根本的解決になりえない. そこで, 本システム人物以外の領域の除外に対して画像情報を併用する.

一般的な画像センサで, 広範囲を観測可能な LRF と同等の観測範囲を得るためには, 単眼カメラを複数台使用するか, 単眼カメラに首振り機構を搭載したパンチルトカメラを使用する. しかし, 単眼カメラの複数台使用の場合には, カメラの台数に応じて画像間のマッチング処理がシステムの負荷となる. また, パンチルトカメラには特定方向を観測している間には他方向が死角となるといった問題が存在する.

そこで, 単一の画像センサで広範囲の情報を取得することにより負荷問題と死角問題を解決する. 広範囲の情報取得が可能なレンズとして代表的なものとしては, 全周魚眼レンズと全方位レンズがある. 全周魚眼レンズは, 画像中心からの距離と角度が比例するようなレンズで, 180 度の画角を得られる. 全方位レンズは全方位ミラーにより水平全方位の視野を得ることが可能なレンズである. 魚眼レンズは特性上, レンズ端では収差・歪曲が大きく, レンズを上方に向けて水平視野を得ようとした際には中央部に対して画像の劣化が問題となる. そこで, 本システムで使用するカメラには, 水平視野を得るのに最適な形状のミラーを用いている全方位レンズを使用する.

2.2 人物検出処理

画像情報を用いた人物検出手法は, 現在までに様々なものが提案されている[6]. 通常の動物体抽出手法は LRF の場合と同様にロボットが移動するため使用できない. フレーム間で画像内に存在する特徴点の位置をマッチングすることでロボ

ットの動きを得る手法も考えられるが、複数の特徴点抽出処理は計算に時間が掛かり、平坦な画像中からは特徴点を取ることが困難であるといった問題が存在する。そこで画像中に人物が存在した場合、頭部に肌が露出している顔領域が存在することを利用し、肌の色で抽出処理を行うことで簡便に人物の検出を行う。しかし、一般環境中には人物の肌の色と類似色の物体が多数存在することが考えられるため、単純に肌の色の領域を抽出するだけでは誤検出が予想される。この類似色の物体は「人物の頭部には髪領域が存在する」と仮定し、肌色領域と髪領域が存在するかを判別することで除去が可能である。これらの処理によって得られた人物位置の候補領域のみを LRF で走査することにより、LRF の走査範囲を制限しなかった場合と比較してマッチング処理に要する時間や、誤検出のリスクを抑えることが可能である。これらの処理を以下にまとめる。

- Step1. 全方位画像に対して肌色抽出
- Step2. 髪領域を推定し髪が存在するか判定
- Step3. 人物と判定された領域のみを LRF でスキャン
- Step4. LRF からのセンサデータに対してマッチング

次節では、この一連の流れを実装するためのアルゴリズムについて述べる。

3. 画像センサによる人物検出

3.1 肌色領域抽出

コンピュータなどに使用されている RGB 表色系は各パラメータが色に関するもので、輝度に関するパラメータが分離されていない。そのため、照明変化により周辺の輝度の変化するような環境で特定の色を抽出するには不向きである。輝度成分が分離されている表色系としては HSV 表色系、YCrCb 表色系などがある。HSV は色相・彩度・輝度、YCrCb は輝度・赤の色差・青の色差により色表現を行っている。特定の色を抽出する際には、色を表現するパラメータが多い表色系の方が有利なため、本システムでは YCrCb 表色系の固定閾値処理により肌色抽出を行う。また、閾値は実験的に求めた値を使用する。色抽出後は適切にノイズ・微小領域の除去を行っている。図 1 に色抽出処理とノイズ処理を行った結果を示す。

Fig. 1 Extract skin color and noise removal

3.2 顔及び髪領域の推定

人物の顔の肌露出領域に対する髪の領域のサイズは、カメラに対する顔の方向などにより変化したが、肌領域の縦横幅から推定することが可能である。肌色領域について、式(1)で定義する縦横比 F_{AR} 、肌色領域の横幅 F_w 、縦幅 F_h を様々な方向の場合で測定し結果を Fig. 22 に示す。

$$F_{AR} = F_h / F_w \quad (1)$$

図 2 より、縦幅 F_h が大きい場合は人物の顔が上向きであり、カメラには髪の領域が小さく映っていることが分かる。逆に縦幅 F_h が小さい場合は人物の顔が下向きで、髪の領域は大き

くなっている。

Fig. 2 Change of the region size of the skin

横幅 F_w の変化が少ないことから、縦横比 F_{AR} は縦幅によって変化する係数と取れる。この縦横比 F_{AR} を利用して髪領域を推定することで、顔の向きに対して頑強に髪領域を得ることが可能である。歩行中の人物の顔の向きに対する F_{AR} は、基礎実験より 0.8-1.8 の間となるため、この値を閾値とする。 F_{AR} が閾値内の場合には F_{AR} を係数として Fig. 33 に示す計算式で髪領域の推定方法を行う。

Fig. 3 Estimate to hair color region from skin color region.

3.3 髪領域による人物の判定

3.2 節で取得した髪領域に対して髪色での色抽出を行い、髪領域に対する抽出領域の割合で人物か否かを判定する。

まず人物の髪色が常に黒色であると仮定し色抽出を行う。HSV 表色系で黒色を表現しようとした場合、色相成分は使用せずに彩度成分・輝度成分の 2 つのパラメータで行うことになる。このとき十分に彩度・輝度が低い領域ならば、基の色に依らず反応するため、黒色のみの抽出には不適當と判断した。また、YCrCb 表色系は色の表現を赤の色差・青の色差で行っているため黒色を表現することが難しく、HSV 表色系と同様に不適切と判断した。よって、全パラメータが色に関係しており、また標準で使用されている表色系のため変換の必要が無い RGB 表色系を使用する。3.2 節の方法により、肌色領域から推定された髪領域に対して髪色による抽出処理を行った際の結果を図 4 に示す。

Fig. 4 Determination in a hair region

図 4 より手前側の人物は、髪の領域の割合が 97[%] と高い値となっていることが分かる。後方の人物は顔の肌色部分の影

響を受け、抽出率が 72[%]と手前側の人物に対して落ちている。しかしながら、人物を含まない右下部の領域では抽出率が 18[%]と低くなっていることから、十分に判別が可能である。この結果から、髪領域を用いて人物を判別することが可能であるといえる。

3.4 全方位カメラへの実装

前節までの処理を、全方位レンズを使用した全方位カメラからの画像に対して適用する。全方位カメラから取得した画像は、図 5(a)に示すように、水平方向の視野が同心円状に画素が配置されたような画像となっている。通常、画像処理アルゴリズムの多くは画素が矩形に配置されている画像について行うことを想定しているため、視覚的な扱い易さの面でも展開処理が必要となる。展開処理は、極座標で表現されている全方位画像の画素を式(2)によって極座標から直交座標に変換することで行える。

$$re^{j\theta} = \begin{cases} r \cos \theta \\ r \sin \theta \end{cases} \quad (2)$$

図 5(a)を展開したものを、図 5(b)に示す。以降この展開処理をパノラマ展開とし、展開された画像を θ° パノラマ画像とする。

Fig. 5 About omni directional camera

パノラマ展開を行った画像に対し、前節までに述べた髪領域を用いて人物を推定する手法を適用する。図 6 は観測範囲中に人物が 3 名存在している環境で、適度に物が配置され、人物の活動する環境内である。領域が強調され、枠から画像下部まで直線が伸びている部分が人物と検出された領域で、枠のみが表示されている領域は色抽出に反応した候補領域である。図 6 より、様々な距離に存在する 3 名の人物を適切に検出されている。色検出の時点では図 6(a)-(c) のように色の近い物体に誤反応しているが、髪領域の色抽出結果により正常に除外されていることがわかる。

Fig. 6 Example which applied panorama image to detection system

4. 距離センサによる人物検出

4.1 LRF からのセンサデータの制限

3 章で肌色領域と髪色領域により、人物が検出されている。このとき、図 5 よりパノラマ画像中の座標とロボットに対する角度は常に対応していることを利用し、式(3)でパノラマ中の座標から角度に変換することが可能である。

$$\theta = x \times \frac{360}{n} - 180 \quad (3)$$

$$\begin{cases} \theta : \text{angle of person's position [deg]} \\ x : \text{coordinate of x axis of person's position [pix]} \\ n : \text{width of panorama image [pix]} \end{cases}$$

この角度情報を基に LRF の走査範囲を限定することで、人物が存在するとされた領域のみをスキャンすることが可能である。図 7(c)は、(b)の LRF のセンサデータに対して、図 7(a)のパノラマ画像に対する人物検出によって得られた角度位置 θ [deg]を利用して走査範囲を制限した様子である。なお、制限範囲は ± 30 [deg]としている。

Fig. 7 LRF scan image and limited range

4.2 LRF センサデータからの人物検出

足の位置を検出する場合、地面に設置されたオブジェクトに対する誤検出の影響が大きい。また、人物の頭部位置には個人差が大きく高さ位置の決定が困難である。そこで、本稿では胴体部の形状を利用し、人物の位置を検出する。

人物の胴体部をスキャンした時、センサデータは図 8 のように半楕円形状となる。また、環境中に存在する類似した形状の誤検出は走査範囲の制限により、最小限に収めることが可能である。この半楕円形状をテンプレートとし、センサデータに対し相関係数を利用したテンプレートマッチングを行う。相関係数は、同形状同士であってもアフィン変換がなされている場合は一致しないため、式(3)で求めた人物角度に応じてテンプレートを回転することで 1 枚のテンプレートにより複数方向に対応することが可能である。これにより LRF でテンプレートマッチングにより画像処理の段階では検出されてしまうが、ポスターなどの形状が人とかけ離れている物体に対しては相関係数が一定以下となり、除外することが可能である。図 9 のように、適切な走査範囲制限とテンプレート回転により人物を検出することが可能である。

Fig. 8 The feature of the person when scanned it in LRF

Fig. 9 Limits a scanning range of LRF

5. 実験

5.1 検出能力の評価実験

LRF として北陽電機製"URG-04LX-UG01"を、全方位カメラとしてヴイストン製"VS-C42U-200-TK"を使用した。また、計算機として PC/AT 互換機(CPU:Core2DuoE8400 MEM:3GB OS:WindowsXP)を用いた。ロボットには LRF と全方位カメラを図 10(a)の配置で設置した全高 150[mm]の三脚を移動台車に搭載したものを使用する。2名の人物(Person 1, Person 2)がロボット左前方から図 10(b)に示す経路で歩行し、ロボットが前進動作を行っている際のそれぞれの人物の検出率を評価した。

(a) Sensor's alignment (b) Experiment environment

Fig. 10 About of experiment

5.2 実験結果

ロボット、複数の人物が共に移動している場合の本システムの有効性を検証した。実験により得られたデータについて、3.5[m]以内での検出成否を評価した。システムはセンサで周辺を毎秒 8 回観測し、検出数を観測毎に記録している。観測時のデータを先頭に 1 秒分のデータを纏めて評価する。

(1) 両名が検出範囲内の場合

Person1, 2 がロボット検出範囲内の場合のグラフを図 11 に、結果を表 1 に示す。検出率は共に 92.0[%]で、安定した複数名の検出が可能となった。

Fig. 11 Graph of detection result (both persons in detection range)

Table 1 Result of both persons in detection range

	Person 1	Person 2
検出数	162	162
総観測数	176	176
検出率[%]	92.0	92.0

(2) それぞれの人物ごとの検出率

Person1, 2 がそれぞれ検出範囲に入ってから抜けるまでのグラフを図 12, 13 に、結果を表 2 に示す。それぞれの検出率は Person1 は 78.9[%], Person2 は 93.6[%]となった。Person1 の検出率が Person2 に対して低くなっているのは、照明のちらつきにより色抽出が正常に行えなかったのが原因と考えられる。

Fig. 12 Graph of detection result (Person1)

Fig. 13 Graph of detection result (Person2)

Table 2 Each detection result

	Person 1	Person 2
検出数	265	277
総観測数	336	296
検出率[%]	78.9	93.6

6. 結言

全方位カメラで取得した画像中の人物の顔領域と髪領域に着目することで背景情報を利用することなく人物を検出し、この結果を利用して LRF の走査範囲を制限することでデータ中の類似物体を除外し、LRF で物体の形状を確認することで、全方位カメラでの検出結果の校正を行うことが可能となった。

ロボットに搭載し行った実験結果より、提案手法において人物・ロボットの移動に影響を受けずに、複数の人物を検出することが可能であることを確認した。

今後の展望として、人物の特徴量を利用して髪色・肌色に影響されない検出系の校正と、検出した対象に追跡を行うための手法の導入を行う。

謝辞 本研究の一部は東京電機大学総合研究所研究(Q10J-5)による。

文 献

- [1] 安川電機, "RobPorter", 2009.
- [2] Forth Valley Royal Hospital, "use robot 'worker'", 2010.
- [3] JETRO, "日本のロボット産業の動向", 2006.
- [4] 奥迫伸一, 坂根茂幸, "レーザレンジファインダを用いた移動ロボットによる人の追跡", 日本ロボット学会, pp.605-613, 2006.
- [5] 糟谷智樹, 小林貴訓, 久野義徳, "ミュージアムガイドロボットへの遠隔対話指示", IEICE, 95, 2008.
- [6] 岩井儀雄, 勞世竝, 山口修, 平山高嗣, "画像処理による顔検出と顔認識", CVIM, pp.149, 2005.
- [7] Hall, "The Hidden Dimension" Doubleday & Company, Inc., E.T. New York, 1966.