

ステレオカメラ搭載移動ロボットへの指の本数による動作指示

Movement Directive Using the Number of Fingers for a Mobile Robot Equipped with Stereo Camera

○学 近藤 正人 (静岡大) 正 山下 淳 (東京大)
正 金子 透 (静岡大) 正 浅間 一 (東京大)

Masato KONDO, Shizuoka University
Atsushi YAMASHITA, The University of Tokyo, yamashita@robot.t.u-tokyo.ac.jp
Toru KANEKO, Shizuoka University, tmtkane@ipc.shizuoka.ac.jp
Hajime ASAMA, The University of Tokyo, asama@robot.t.u-tokyo.ac.jp

In recent years, introduction of autonomous mobile robots which support humans in environments close to our life is expected. Examples are shopping support robots and porter robots carrying the luggage while accompanying the person. To utilize these robots, real time movement directive to the robots is important. In this paper, we propose a mobile robot system which performs the actions by recognizing the number of fingers. The robot recognizes the number of fingers from images obtained with a stereo camera system. Also, as one of the actions, the mobile robot tracks human who is attaching the mark. We verified the validity of the system by using a wheel mobile robot in indoor environment.

Key Words: Mobile Robot, Man-Machine Interface, Gesture Recognition, Human Following, Stereo Camera

1. 序論

本研究では、利用者の提示した指の本数を認識して動作を実行する移動ロボットシステムの構築を目的とする。

近年のロボット産業の発展に伴い、人間の生活に近い環境を自律的に移動し、人を支援するロボットの導入が期待されている。そのようなロボットの例として、大型店舗で人に同行しながら商品を運ぶ買い物支援ロボット、ホテルや空港で荷物を運搬するポーターロボットが挙げられる。これらのロボットを人が利用するためには、リアルタイムでの人からロボットへの動作指示が重要となる。

リアルタイムでの人からロボットへの動作指示には、人にとって直感的で使いやすい手ジェスチャを用いるのが有効である[1-3]。ここで、手ジェスチャとは手の動きと形を意味する。手ジェスチャを用いることで、利用者はリモコン等を携帯する必要がなく、より自然に動作指示を行うことが可能となる。

そこで本研究では、手ジェスチャとして、指の本数によりロボットへの動作指示を行う。ロボットには視覚としてステレオカメラを搭載し、ロボットが移動中であっても安定して指示を認識することが可能なロボットシステムの構築を行う。また、ロボットの動作の1つとして、利用者にマークを装着することで人物追跡を行う。

2. 処理の概要

本研究では、車輪移動ロボットを用いる(図1)。ロボットには平行ステレオとして2台のカメラを搭載し、これによりロボット正面の画像を取得する。また、ロボットの動作環境は段差のない屋内とする。利用者は肌色に近くない色の服の上から背中にマークを装着し、動作を指示する際にロボットの正面に手を向けることとする。

処理の流れとして、まず、ステレオカメラにより正面の画像を取得する。次に、取得した画像から肌色情報を用いて指の本数を認識する。最後に、認識結果に対応した動作を実行する。これらの処理をロボットが移動している間も繰り返し

行うことで、移動中でも動作を変更することが可能となる。

人が提示する指の本数と動作の対応を表1に示す。本研究では買い物支援ロボットを想定し、ロボットが人物を追跡しながら荷物の運搬を行う際に必要な動作として、表1のような動作を設定した。また、人物追跡では、取得画像に対してテンプレートマッチングを行うことで利用者に装着したマークを検出し、追跡を行う。

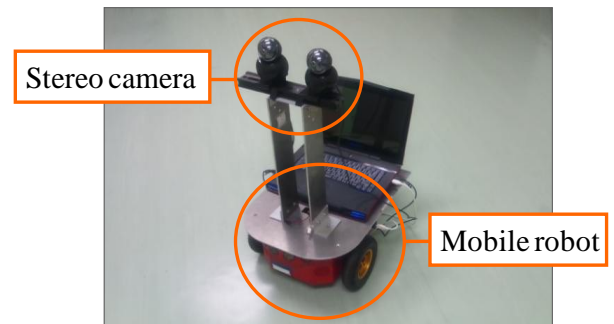


Fig. 1 Mobile robot equipped with stereo camera.

Table 1 Number of fingers and actions.

Number of fingers	Actions
0	Stop
1	Human following
3	180 degrees turn
5	Back

3. 指の本数の認識

3.1 指の本数の認識方法

指の本数の認識では、まず、ステレオカメラで取得した左右画像のうち左画像から肌色領域を抽出する。次に、抽出した肌色領域から手領域を抽出し、その特徴量を求める。最後に、求めた特徴量を用いて指の本数を認識する。



Fig. 2 Acquired image.

Fig. 3 Extracted skin region.

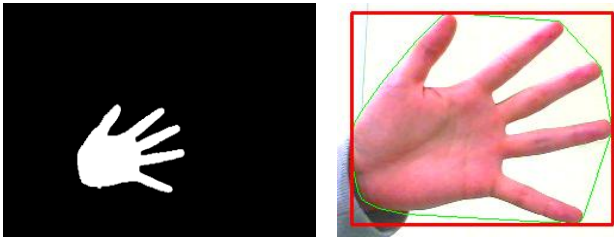


Fig. 4 Extracted hand region.

Fig. 5 Areas.



(a) Hand.

(b) Without fingers.

(c) Fingers only.

Fig. 6 Process regions.

3.2 肌色領域の抽出

取得した左画像（図2）をRGB表色系からHSV表色系に変換する．HSV表色系では色が色相H，彩度S，明度Vにより表現されるため，RGB表色系よりも明るさの変化等にロバストに対応することができる．左画像をHSV表色系に変換した後，各画素のH，S，Vのそれぞれの値について設定した範囲内にあるかどうかを判定する．このようにして抽出した肌色領域を図3に示す．図3では肌色領域を白，それ以外を黒で示している．

3.3 手領域の抽出

抽出した肌色領域（図3）は小さな穴や点といったノイズを多く含んでいる．そこで，このようなノイズを除去するために，肌色領域に対して同回数の膨張収縮処理を行う．

次に，抽出した肌色領域から手領域を抽出する．肌色領域（白画素）に対してラベリングを行い，連結領域を分ける．そして，連結領域のうち面積が最大となる領域を手領域として抽出する．ここで，連結領域の最大面積が閾値以下ならば処理を打ち切る．このようにして抽出した手領域にはまだ領域内部に小さな穴がある場合がある．そこで，手領域の輪郭を検出して内部を穴埋めする．このようにして抽出した手領域を図4に示す．

3.4 指の本数の検出

抽出した手領域の大きさは，ロボットまでの距離や指の本数に応じて変わるため，指の領域の面積は変化する．そこで，手領域と手領域を囲む矩形および凸包領域の面積比を求める（図5）．これらの比は距離に依らずほぼ一定となるため，距離に不変な特徴量とすることができる．求めた特徴量から収縮膨張処理の回数と指の面積の閾値を決定する．

最後に，手領域（図6(a)）に対して決定回数の収縮膨張処理

を行い，指のない領域（図6(b)）を求める．この際，膨張処理を収縮処理よりも多く行うことで，指のみの領域を確実に求められるようにする．続いて，図6(a)から図6(b)の差を取り，指のみの領域（図6(c)）を求める．最後に，求めた指のみの領域を数えることで指の本数を認識する．

4. ロボットの動作

移動ロボットは指の本数の認識結果が表1に示した本数であったとき，その本数に対応した動作を実行する．ただし，後退の動作は安全を考慮して5本を認識している間のみ実行する．そして，動作が変更した際には利用者に分かるように動作に対応した音声を出力する．例えば，人物追跡に切り替わった際には「追跡します」，停止する際には「停止します」といった音声を出力する．また，動作変更後は次の指示を誤認識しないように次の認識まで少し時間を空ける．

5. 人物追跡

5.1 人物追跡の方法

本研究では，図7のような4色の正方形からなるマークを利用者に装着することで人物追跡を行う．ここで選択した色は明るさの影響を受けにくい4色であり，また実空間に存在しにくい色の組み合わせであることから，取得画像中で安定して検出することが出来る．予め，図7と同じテンプレート画像を用意しておく．

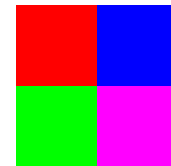


Fig. 7 Mark.

人物追跡では，まず，ステレオカメラで取得した左右画像にテンプレートマッチングを行い，マークの位置を検出する．次に，左右のマークの位置から人物までの距離と角度を計算する．最後に，距離と角度に応じてロボットを移動させることで人物の追跡を行う．

5.2 マークの検出

まず，ステレオカメラで取得した左画像上でテンプレート画像をラスタスキャンし，相違度を計算する．相違度の計算にはSAD（Sum of Absolute Difference）を用いる．テンプレート画像の画像サイズを $M \times N$ ，取得画像上での座標を (x, y) ，テンプレート画像上での座標を (i, j) とすると，テンプレートの画素値 $T(i, j)$ とテンプレートと重ね合わせた取得画像の画素値 $I(x+i, y+j)$ から相違度 R_{SAD} は式(1)で計算される．

$$R_{SAD} = \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} |I(x+i, y+j) - T(i, j)| \quad (1)$$

本研究ではさらに高速にテンプレートマッチングを行うために，まずテンプレート画像の4隅で画素の色相について相違度 R_{SADj} を計算する．そして，計算した相違度 R_{SADj} を閾値と比較し，閾値以下の場合にのみテンプレート画像全体で画素の色相について相違度 R_{SAD} を計算する．この R_{SAD} が最小となる位置を左画像上でのマークの位置として検出する．左画像上でテンプレートマッチングを行った後，エッジ線に沿って右画像についても同様にテンプレートマッチングを行い，マークの位置を検出する．

5.3 追跡時のロボットの動作

ロボット座標系を左右カメラ間の中心とする。検出した左右画像の位置よりロボット座標系から見た追跡人物までの距離と角度を計算して追跡を行う。追跡時のロボットの移動速度は距離と角度に応じて変化させる。

6. 実験

6.1 実験装置

本研究では、MobileRobots社製の車輪移動ロボットPioneer2を用いた。このロボットは、2駆動輪1キャスタの機構である。カメラはLogicool社製のQcamOrbitAFを用いた。計算機の仕様はCPUがIntel Core i7 X920 2.0GHz、メモリが8.0GBである。

6.2 実験条件

本実験では、指の本数の認識精度の評価と指の本数による移動ロボットへの動作指示を行った。実験環境の様子を図8に示す。実験は床面に段差のない屋内環境で行った。取得画像のサイズは640×480pixel、テンプレート画像のサイズは11×11pixelとした。また、装着するマークは5×5cmとした。マークを装着した様子を図9に示す。



Fig. 8 Environment.



Fig. 9 Mark attaching.

6.3 実験結果

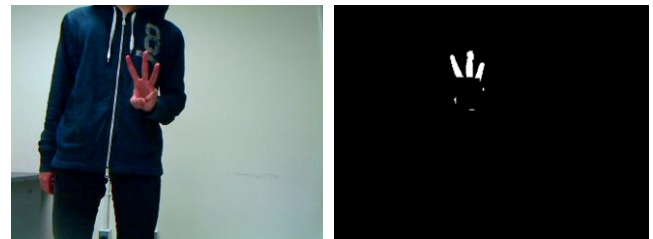
まず、指の本数の認識精度の評価を行った。単純背景下にて、人が待機中のロボットに対して3種類の指示を提示し、その様子を撮影した。撮影距離は実際のロボットの利用を想定し、手からロボットまでの距離を50, 100, 150cmとした。また、撮影回数は指示と撮影距離ごとに200回ずつとした。そして、取得した画像に対して指の本数の認識を行い、認識結果が提示した指の本数と合致しているかを確認した。

各撮影距離における取得画像と抽出した指のみの領域を図10から図12に示す。各距離において、本手法によって指のみの領域を抽出することが出来た。次に、指の本数の認識結果を表2に示す。表2より、指の本数を安定して認識できていることが確認できた。誤認識の原因としては、画像のブレや照明の影響による色の変化が挙げられる。色の変化によって肌色領域の画素値が設定した範囲内に収まらず、上手く抽出することが出来ないことがあった。また、距離が50cmにおいて

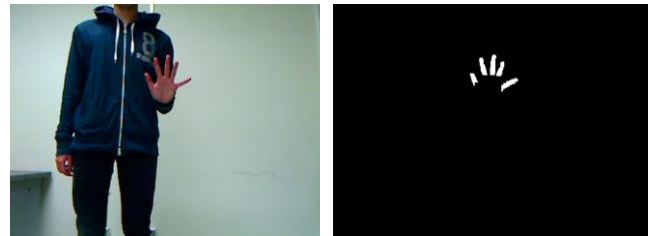
認識精度が低下している。これは撮影距離が近いことで手領域の面積の変化が大きくなり、設定した収縮膨張の回数では不十分である場合が生じたためである。しかし、ロボットは1m以上離れて追跡を行うため、距離の近さによる認識精度の低下は問題ないといえる。この認識精度の評価を踏まえ、実際に動作指示を行う際には誤動作を防ぐために3フレーム連続して認識した動作を実行することとした。



(a) Acquired image. (b) Finger region.
Fig. 10 Appearance at a distance of 50cm.



(a) Acquired image. (b) Finger region.
Fig. 11 Appearance at a distance of 100cm.



(a) Acquired image. (b) Finger region.
Fig. 12 Appearance at a distance of 150cm.

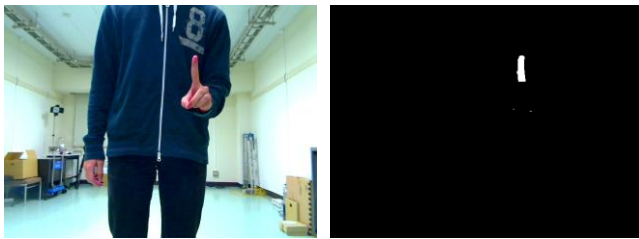
Table 2 Recognition accuracy of the number of fingers.

Number of fingers	Success rate[%]		
	50cm	100cm	150cm
1	94.0	95.5	97.0
3	92.0	98.5	94.0
5	95.0	97.0	98.0

次に、実験環境において指の本数による移動ロボットへの動作指示を行った。実験では、停止しているロボットに対してまず人物追跡を指示し、その後、180° 旋回、人物追跡、後退、停止という順番で動作を指示した。

実験の結果、利用者が提示した指の本数を移動ロボットが安定して認識し、各動作を行うことが出来た。3フレーム連続して認識した動作を実行することで誤動作はなかった。図13から図17に動作変更時に取得した画像と抽出した指のみの領域を示す。このように、ロボットが移動中であっても正確に指のみの領域を抽出することが出来た。

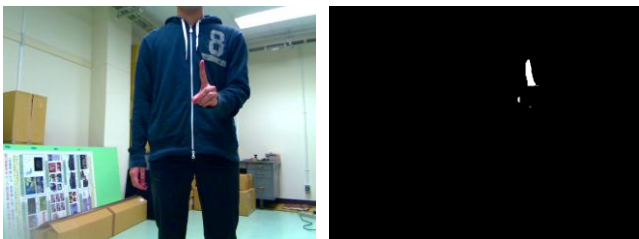
また、人物追跡時に取得した画像を図18に示す。図18より、各左右画像においてマークの位置を検出することが出来た。よって、本手法により利用者の追跡が可能であることが確認できた。



(a) Acquired image. (b) Finger region.
Fig. 13 1st movement directive: Human following.



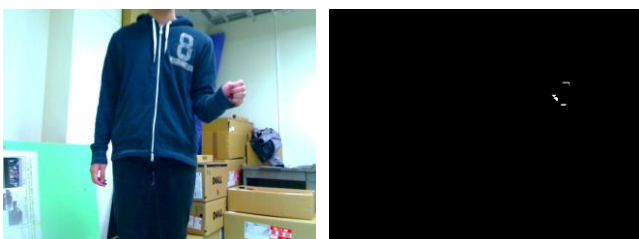
(a) Acquired image. (b) Finger region.
Fig. 14 2nd movement directive: 180 degrees turn.



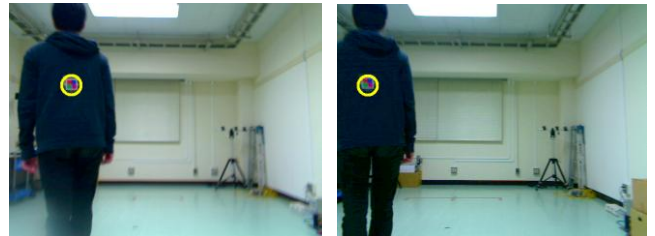
(a) Acquired image. (b) Finger region.
Fig. 15 3rd movement directive: Human following.



(a) Acquired image. (b) Finger region.
Fig. 16 4th movement directive: Back.



(a) Acquired image. (b) Finger region.
Fig. 17 5th movement directive: Stop.



(a) Left image. (b) Right image.
Fig. 18 Appearances in human following.

以上の結果より、実験環境において、本手法により移動ロボットへの指の本数による動作指示と人物追跡が達成できることが確認できた。

7. 結論

本研究では、利用者の提示した指の本数を認識して人物追跡等の動作を行うロボットシステムの構築を行った。ロボットへの動作指示に指の本数を用いることで移動中であっても正確に動作指示することが出来た。また、3フレーム連続して認識した結果を用いることで、誤動作をなくした。さらに、マークを人物に装着することで安定して人物追跡を行うことが出来た。

しかし、本手法では肌色の範囲が固定であるため、利用者によっては上手く肌色領域を抽出出来ない場合がある。そこで今後の課題として、利用者の肌色の値を取得することで、利用者に応じた範囲を設定する必要がある。また、本手法では片方の手のみを認識しているため、指示可能な動作は限られる。そこで、左右の手を認識することによる動作指示の拡張が考えられる。

文 献

- [1] Matthew M. Loper, Nathan P. Koenig, Sonia H. Chernova, Chris V. Jones, and Odest C. Jenkins: "Mobile Human-robot Teaming with Environmental Tolerance", Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction, pp.157-164, 2009.
- [2] 丸谷誠慶, 西川敦, 宮崎文夫: "2次元画像からの指先位置検出に基づく実時間手形状認識と移動ロボット遠隔操縦タスクへの応用", 情報処理学会研究報告コンピュータビジョンとイメージメディア, Vol.2002, No.34, pp.145-152, 2002.
- [3] 田中信也, 梅田和昇: "ジェスチャ認識を用いた移動ロボットの操作システム", 電気学会論文誌 C, Vol.121, No.9, pp.1457-1463, 2001.