

人支援ロボットの受け渡し動作のための手認識システムの開発 ～ロボットへの提示行動の調査と手形状の分離方法の基礎的検討～

Development of a Hand Detection System for Hand Over Task of an Assistant Robot

～Survey of Presented Action to Robot and Basic Study of Hand's Shape Classification～

○佐藤弘基, カニエテ ルイス, 高橋隆行

○Kouki Sato, Canete Luis, Takayuki Takahashi,

福島大学

Fukushima University

キーワード： 人支援ロボット (human support robot), 手認識 (hand detection),
画像処理 (image processing), 距離画像 (depth image)

連絡先： 〒 960-1248 福島市金谷川 1 番地 福島大学 理工学群 共生システム理工学類 高橋隆行研究室
佐藤弘基, Tel./Fax.: (024)548-5259, E-mail: kouki@rb.sss.fukushima-u.ac.jp

1. 緒言

近年, 人の生活環境でサービスを提供する人支援ロボットの開発が盛んである. 人支援ロボットは, 人や生活環境に危害を加えない高安全性と多種多様なタスクをこなす高作業性が求められている. しかし, これらは一般にトレードオフの関係にあり, 安全性を保ちつつ作業性の高いロボットの開発は困難である. この課題に対し筆者らは, 車輪型倒立振り機構のメリットを積極的に利用し, サービスを効率よく提供することを目的とした人支援ロボット「I-PENTAR(Inverted PENDulum Type Assistant Robot)」を開発している¹⁾. I-PENTAR のハードウェアの外観を Fig. 1 に, 仕様を Table 1 に示す. I-PENTAR は, 軽量で低トルクなマニピュレータを用いるこ

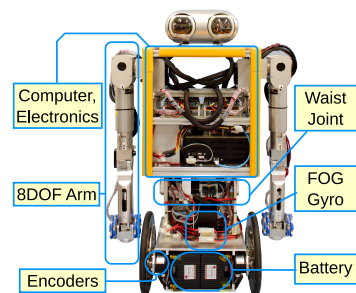


Fig.1: I-PENTAR

とで安全性を保っている. さらに, 自身が傾くことによって重心を移動させ自重を有効に利用し, 重量物の持ち上げや運搬が可能である.

筆者らは, 人支援ロボットのサービスとしてマニピュレータを用いた人との物体受け渡し動作の実現を目指している. ロボットと人との

Table 1: Hardware specification of I-PENTAR

Size	0.9 [H] × 0.3 [W] × 0.2 [D] [m]
Weight	30 [kg]
DOF	Mobile Base(2), Manipulator(8×2), Waist(1)
Sensor	Gyro sensor(2), Encoder(19)

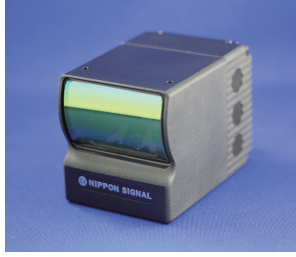


Fig.2: InfiniSoleil FX8(NIPPON SIGNAL)

物体受け渡し動作の研究には、人同士による受け渡し動作の解析とそれに基づく腕部の制御則の構築²⁾や、手の認識などの課題がある。手の認識では、予め用意された手の形状を認識し、それをロボットへの指示として用いるものが多い^{3, 4)}。本研究では、手の認識について検討する。

主な手の認識方法は、RGB センサを用いた肌色の抽出による認識と RGB-D センサを用いた認識の 2 つが挙げられる。しかし、RGB センサは色や光の影響を受けやすく、RGB-D センサは検出距離範囲が狭く、外乱光に弱いという問題点がある。そこで郎らは、外乱光に強い 3D 距離画像センサ InfiniSoleil FX8(日本信号)を用いてランダムフォレスト機械学習を行い、リアルタイムかつ距離情報のみで手の検出を行うシステムを構築し、実験によりその有用性を示した⁵⁾。しかしながら、さらに球体近似と Hough 変換による手の形状認識も試みたが期待するような結果は得られなかった。3D 距離画像センサ InfiniSoleil FX8 の外観を Fig.2 に、仕様を Table 2 に示す。

そこで本研究では、まず、人の提示行動を調査し、手の形状をパターンに分類する。本研究

Table 2: Specification of Laser Ranging Image Sensor InfiniSoleil FX8

方式	光パルス飛行時間計測法	
検出距離範囲	0 ～ 15 [m]	
画角	水平	60 [deg]
	垂直	50 [deg]
応答速度	4 [f/s]	
測距点数	約 100×60	
距離分解能	Min.4 [mm]	
距離精度	0.3 ～ 5 [m] (Ta:25 [°C])	
耐外乱光	200,000 [lx] 以上	
外形寸法	W 62 × H 66 × D 89 [mm]	
質量	約 0.5 [kg]	
動作温度	-10 [°C] ～ +50 [°C]	

では、人がロボットから物体を受け取りたいと思った時にとる行動を提示行動と呼ぶ。得られた形状パターンを認識するシステムを構築することで、人が自然にロボットに差し出した手とロボットとの受け渡し動作が可能となる。本研究では、その認識システムについても検討する。

2. 手認識システムの概要

手認識システムには距離情報のみを用いたランダムフォレスト機械学習を用いている⁵⁾。教師あり学習であるランダムフォレスト機械学習は、手認識の準備として大量の教師データによる学習を行い、決定木を複数構築し識別器を生成しておく必要があるが、予め識別器を構築するだけで高速な検出が行える。手認識は、構築しておいた識別器に、認識したい距離画像を与え、ピクセルごとに手の識別をさせる。

2.1 教師データの取得方法

機械学習に必要な教師データは、距離画像中の手部分をラベリングしたものである。本研究で用いる FX8 は、光パルス飛行時間計測法によりセンサからの距離を計測している。そのた

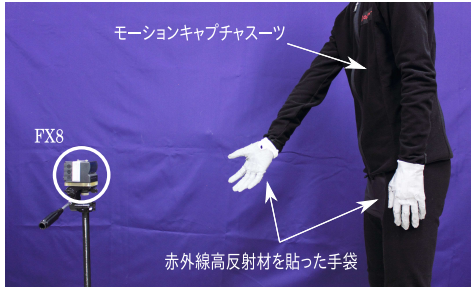


Fig.3: Training data acquisition

めこのセンサは、距離情報の他に、赤外線反射率も取得することができる。この赤外線反射率の情報を用いて手のラベリングを行った。教師データの取得風景を Fig.3 に示す。まず、3次元モーションキャプチャーのマーカースーツとして用いられる赤外線高反射材を貼った手袋を着用し、距離画像を取得した。そして、赤外線反射率の高いピクセルを手として教師データを作成した。この時、服装によっては赤外線反射率の高い部分が出てしまうので距離画像取得時には、赤外線反射率を抑える為に、モーションキャプチャスーツ (OptiTrack) を着用した。

2.2 Depth image features

ランダムフォレスト機械学習では、決定木を構築するために、feature と呼ばれる「特徴量」を用いる。本研究では、以下で説明する Depth image features を用いた。Depth image features $f_\theta(I, x)$ は式 (1) のように定義され、決定木におけるノード分岐に利用する。本 feature はノード分岐に用いられる特徴量としては、シンプルで非常に弱い識別能力しか持たない。しかし、1つでは識別能力の弱い特徴量も 20 階層の決定木を 3つ用いると $2^{20} \times 3$ 個集まることとなり、距離画像を認識する有効性と高い精度が示されている⁶⁾。

$$f_\theta(I, x) = d_I \left(x + \frac{u}{d_I(x)} \right) - d_I \left(x + \frac{v}{d_I(x)} \right) \quad (1)$$

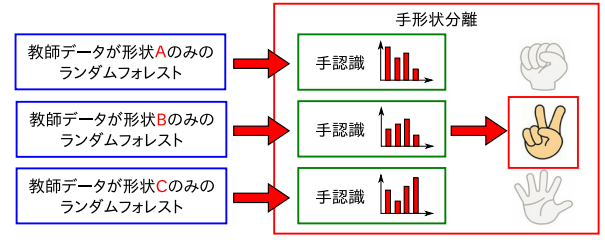


Fig.4: proposed hand shape detection system

Depth image features は、注目ピクセル x から $\theta = (u, v)$ だけオフセットした 2 点の距離の差分である。 $d_I(x)$ は距離画像 I 中のピクセル x の距離である。

3. 手形状分離手法の提案

ロボットと人との自然な物体受け渡し動作を行う為には、手の位置情報のみでは不十分であり、手形状も認識する必要がある。

3.1 提案手法の概要

本研究で検討する手法の模式図を Fig. 4 に示す。教師データが異なるランダムフォレストを複数構築し、それらの認識結果の差異から用意した手形状パターンに分離する。しかし、本研究で開発しているシステムは受け渡し動作の為のものであるが、今まで受け渡し動作においてどのような手形状となるのか検討されていなかった。そのためこの手法を用いるには

- 1) 認識、分離すべき手形状
- 2) 認識結果に差異が生じ易い教師データ
- 3) 分離に効果的なパラメータ

を検討する必要がある。

3.2 予備実験

まず筆者らは、教師データの異なるランダムフォレストで、認識結果に形状を特定できるよ

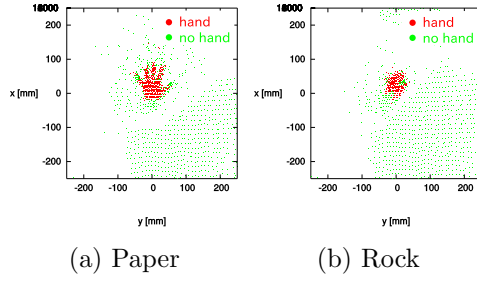


Fig.5: Random Forest outputs of paper only training data

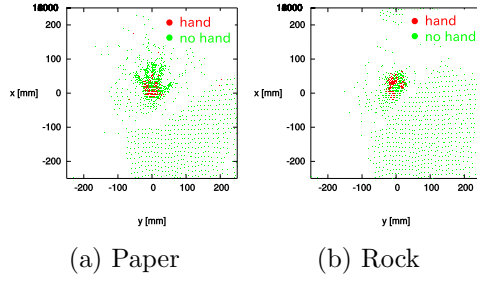


Fig.6: Random Forest outputs of rock only training data

うな有意差があるか検討した。そこでまず、郎らが課題としていたグーとパーの形状判別を試み、パーのみを教師データとしたランダムフォレストとグーのみを教師データとしたランダムフォレストの構築を行なった。2つのランダムフォレストにそれぞれグーとパーを認識させた結果をFig. 5とFig. 6に示す。教師データがパーのみのランダムフォレストでグーを認識してしまっている。これは、パーの掌部とグーでDepth image featuresが近くなってしまった為と考えられる。しかし、Fig. 5(a)とFig. 6(a)を比較するとFig. 6(a)の認識ピクセル数が非常に少なく、指部が認識できていないことがわかる。

教師データに近い形状のみを認識することはできなかったが、認識ピクセル数が大きく異なる形状があることを確認した。特に指部の認識に差が生まれ、形状特定をする上で有意差となる可能性があり、提案手法による形状分離が行える可能性が示唆された。

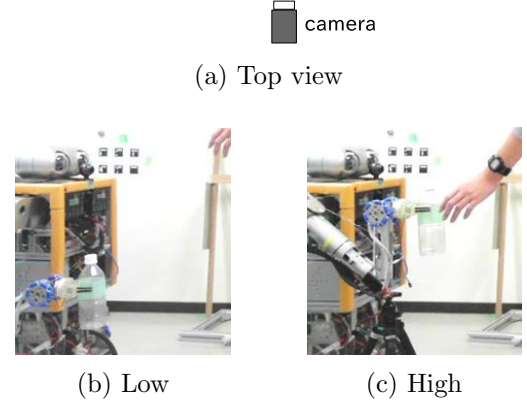
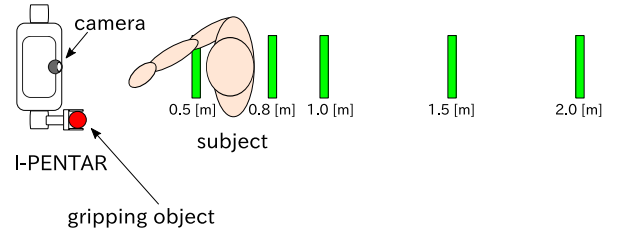


Fig.7: Survey method of presented action by human to robot



Fig.8: Results of survey

4. ロボットへの提示行動の調査

筆者らが開発している I-PENTAR の大きさに近いロボットに対する人の提示行動を調査した研究はない。そこでまず、提示行動そのものを調査すると共にその時の手形状を確認し、認識すべき手形状を検討した。また、I-PENTAR 頭部に FX8 を搭載した場合に、提示行動を撮影することができるのか確認した。

4.1 調査方法

調査の模式図を Fig. 7 に示す。被験者には、物体を把持しているロボットから定められた5つの距離で「物体が欲しいというアピールをして欲しい」とだけ指示し、2台のカメラで撮影した。今回の調査は、ロボットから0.5[m], 0.8[m], 1.0[m], 1.5[m], 2.0[m]の距離で行った。Fig. 8 に示すような画像を取得するために、1台のカメラは、ロボット頭部中央に設置し被験者を正面から撮影した。他方は、ロボットと被験者の側面を撮影できるように設置した。比較のため、物体は、500[ml]のペットボトルと $\phi 65$ [mm]のラテックス系ゴムボール、長さ約60[mm]の鍵の3種類用意し、ロボットが物体を把持している高さも変え調査を行った。また、Fig. 8 (a) から写っている手がFX8の垂直画角50[deg]の範囲内か確認した。これは、Fig. 8 (a) に写っている camera から頂角50[deg]の二等辺三角形を描画し、その三角形内部に手がある場合は、FX8でも提示行動が撮影可能とした。被験者は18歳から22歳の男女37名である。

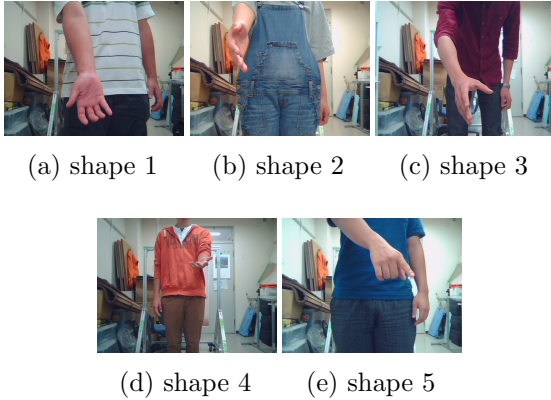


Fig.9: Five of hand's shape by presented action

4.2 結果

4.2.1 手形状

全被験者の Fig. 8 (b) を例とする調査結果画像を比較すると、提示行動における手形状は5つに分類できることがわかった。提示行動における手形状の代表例を Fig. 9 に示す。Fig. 9 (a)

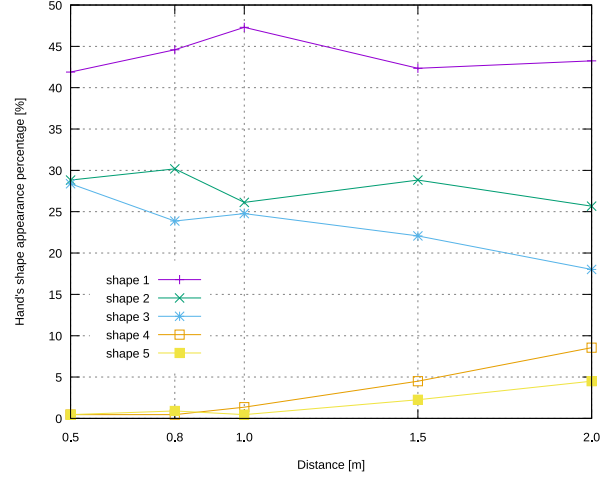


Fig.10: Percentage of presented hand shape vs. distance

Table 3: Valuation basis of hand position change

Point	Description
5	higher
4	slightly higher
3	nochange
2	slightly lower
1	lower

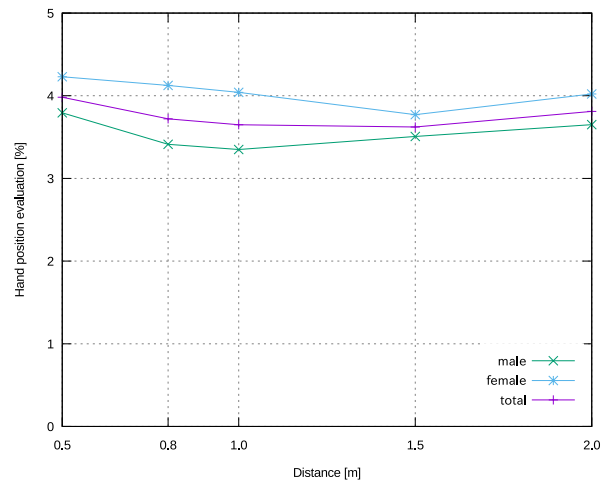


Fig.11: Average of hand position change valuation

は、掌が完全に写るものであり、Fig. 9 (b) は、掌が写るが指で一部分が隠れてしまうものである。Fig. 9 (c) は、手の甲が写るものである。Fig. 9 (d) は、手を完全に開いた状態で I-PENTAR から手が水平に見える。Fig. 9 (e) は、指差しである。本研究で開発する形状分離システムでは、この5つの手形状を認識する必要がある。なお、5つの手形状の出現割合を Fig. 10 に示す。結果より、人の提示行動の 80[%] 以上は shape 1～3 であり、ロボットの人との距離が離れるにつれて shape 4, 5 の割合が増加する傾向がみられた。

4.2.2 手の差し出し高さ

前節の調査において、一度でも画角の範囲外に手を差し出した被験者の割合は、全被験者の約 60[%] であった。そのため、手の差し出し高さについてまず検討を行なった。I-PENTAR の物体把持位置の低い Fig. 7 (b) と物体把持位置の高い Fig. 7 (c) を比較し、手の位置に I-PENTAR 物体把持位置の影響があるか確認した。Fig. 7 (b) の手の位置を基準とし、Fig. 7 (c) の手の位置の変化を Table 3 に示すような 5 段階評価を行い、その平均を Fig. 11 に示す。Fig. 11 の通り平均が「3 nochage」を上回った。したがって、手の位置は I-PENTAR 物体把持位置の影響を受け、把持位置が高くなれば、提示行動の手の差し出される高さも高くなる。これを利用し、FX8 の範囲外に手を差し出した場合は、I-PENTAR の物体把持位置を高くし、範囲内に誘導することが可能である。

5. 手形状分離実験

5.1 実験方法

第 3.1 節で提案した手法による Fig. 9 に示した手形状への分離実験を行う。教師データが異なると認識ピクセル数が異なる手形状があることを第 3.2 節で確認した。よって、本研究では

複数のランダムフォレストの認識ピクセル数割合の比をパラメータとして分離に用いる。

まず、複数のランダムフォレストを構築し、それぞれ認識ピクセル数の割合を 2 つの手法を用いて導出した。これらの割合の比を Fig. 9 に示した手形状ごとにパターン分けを行い、そのパターンが連続的な手形状変化においても出現するか確認する。評価に用いる連続的な手形状変化は、I-PENTAR から 0.8[m] の位置で手を差し出し、Fig. 9 の shape 1, 2, 3, 4, 5, グーの順に連続的に変化させた距離画像とする。

Table 4: Parameters for learning and learning time

	本研究	先行研究
教師データの数	1000	3000
フィーチャー数	100	100
閾値数	30	30
木の深さ	18	18
木の数	1	3
学習時間 [h]	2～3	16
認識率 [%]	60～80	80～95

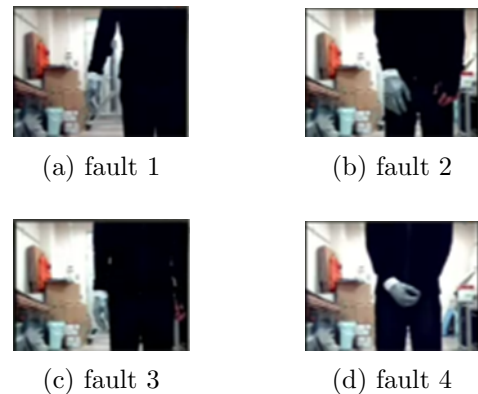


Fig.12: Fault patterns

5.1.1 ランダムフォレストの構築

本研究と先行研究の学習条件と学習に要した時間を Table 4 に示す。先行研究における手の認識精度は約 80～95[%] であったが、教師デー

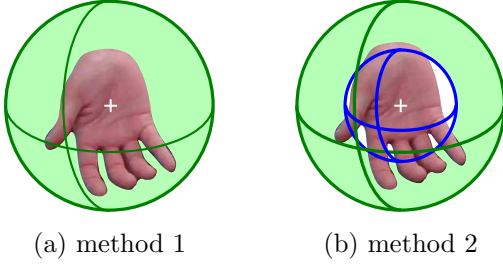


Fig.13: Search area

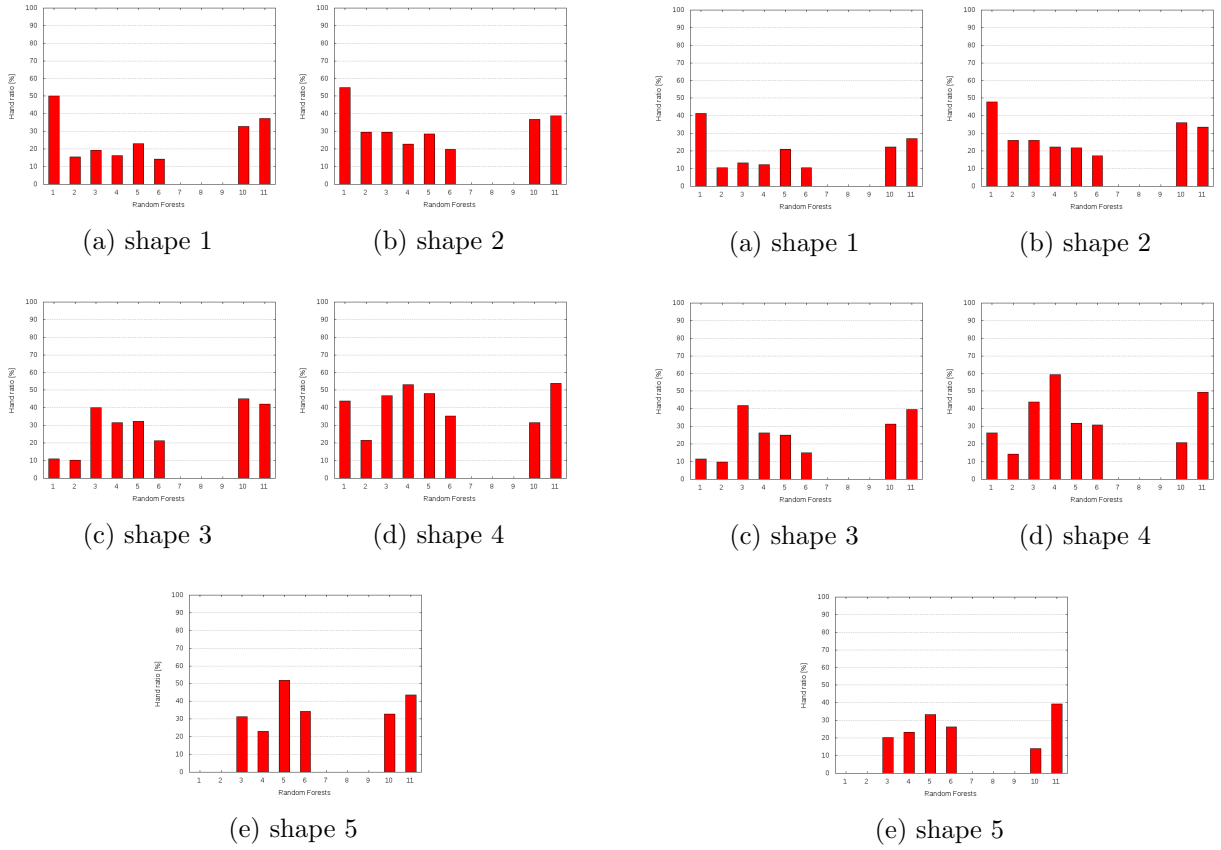


Fig.14: Recognition pattern by method 1

Fig.15: Recognition pattern by method 2

タを 1000 に減らしたことで約 60～80[%] に低下してしまった。しかしながら、基礎的検討段階にある本研究では、様々なデータを調査する必要がある。よって、認識精度よりも学習時間の短縮を優先し、以下で行う学習では教師データ数を 1000 とすることとした。

学習に必要な教師データを手形状ごとに作成した。まず、提示行動調査と同様な環境を用意し、認識対象である提示行動中の 5 つの手形状の距離画像をそれぞれ取得し、教師データを作成した。shape 1～5 の教師データで学習を行い作成されたランダムフォレストをそれぞれ Random forest 1～5 と呼ぶ。但し、I-PENTAR から 1.5[m] と 2.0[m] の位置では、手の部分の 픽セル数が少なくなってしまう為、教師データは 0.5[m], 0.8[m], 1.0[m] の位置で提示行動を行ったものとした。また、比較の為に以下 6 個の教師データを追加的に取得し学習を行なった。まず、shape 5 に近い形状であるグーのみを教師データとし学習したものを Random forest 6 とする。Fig. 12 に示す提示行動とは完全に異なる手形状又は状態の教師データを 4 つ作成し、順に Random forest 7～10 とする。そして、shape 1～3 を全て含む形状を教師データとし、学習したものは Random forest 11 とした。本研究では、これらの 11 個のランダムフォレストを用いて形状分離を試みた。

5.1.2 認識 픽セル数による手形状分離方法

手形状分離には、探索領域内全ての 픽セルを分母とし、手と認識した 픽セル数の割合を用いた。探索領域決定には、手の認識 픽セル数の最も多いランダムフォレストの出力における手の重心を利用した。探索領域の模式図を Fig. 13 に示した。手法 1 は、手の重心位置を中心とする球を探索領域とした。手法 2 は、指に着目⁷⁾し、手の重心位置を中心とする 2 つの大きさの異なる球を用意し、大きな球と小さな球の間をくり抜いてできる球体のシェルを探索領域とし

た。球の半径は、手法 1 では 100[mm]、手法 2 では 100[mm] と 30[mm] とした。

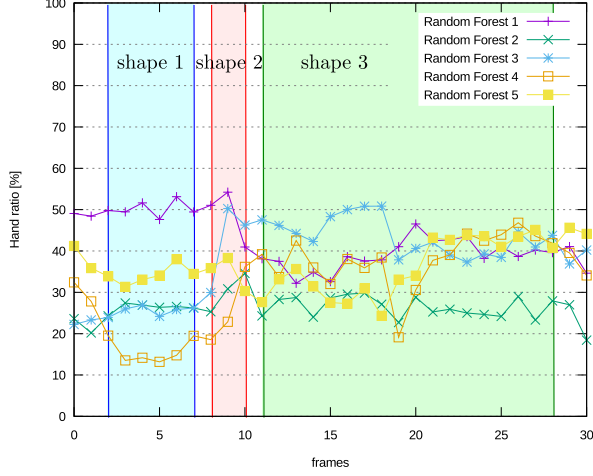
そして、提示行動を I-PENTAR から 0.8[m] の位置で行い、FX8 で距離画像を取得した。取得した距離画像から手法 1, 2 を用いて認識 픽セル数の割合をそれぞれ計算した。5 つの手形状ごとのパターンをそれぞれ Fig. 14, 15 に示す。これを形状分離基準パターンとし、近いパターンを取れば形状分離が可能と判断した。

5.2 結果

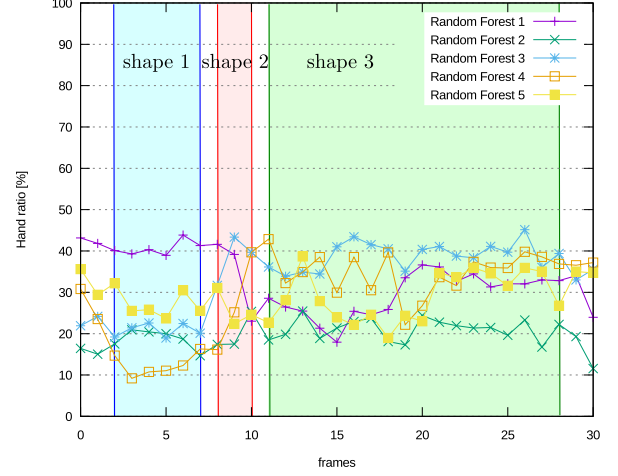
まず、連続的な手形状変化内の shape 1～3 までの手法 1, 2 における認識 픽セル数の割合の変化をそれぞれ Fig. 16, 17 に示す。連続的な手形状変化においても Fig. 14, 15 のパターンが出現し、本手法による手形状分離の可能性を確認した。shape 1 はどちらの手法においても安定して同一なパターンが得られた。しかし、shape 2, 3 では、基準パターンとは異なるパターンが出現することがあり、特に手法 1 でのパターンが乱れてしまった。同様に shape 4, 5 でもパターンの乱れが起きた。これらは教師データの数を増やすことにより認識精度が向上し改善できると考えている。

6. 結言

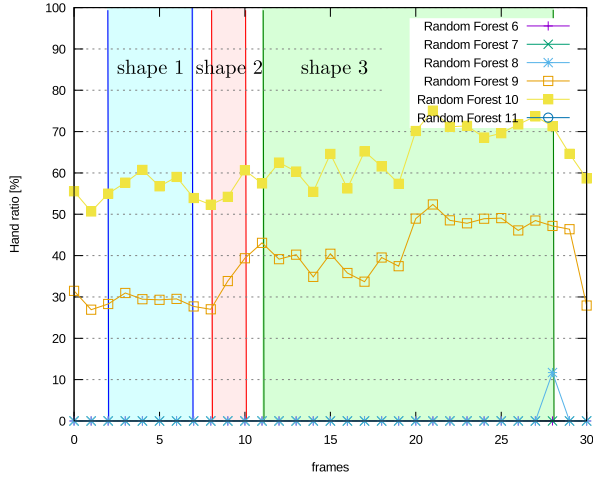
本論文では、人支援ロボットの受け渡し動作のための手認識システムの開発について述べた。まず、人とロボットとの間で行われる提示行動中の手形状について調査し、ロボットに差し出される手形状が 5 つあり、人は手を物体方向に差し出すことを確認した。そして、調査結果に基づき、物体受け渡し動作時に認識すべき手形状を決定し、複数のランダムフォレストの認識結果の差異からそれらの手形状への分離を試みた。認識 픽セル数の比を 2 つの手法で導出し、認識パターンを確認した。手法 1 ではパターンの乱れが大きく、本論文では手法 2 のパターン



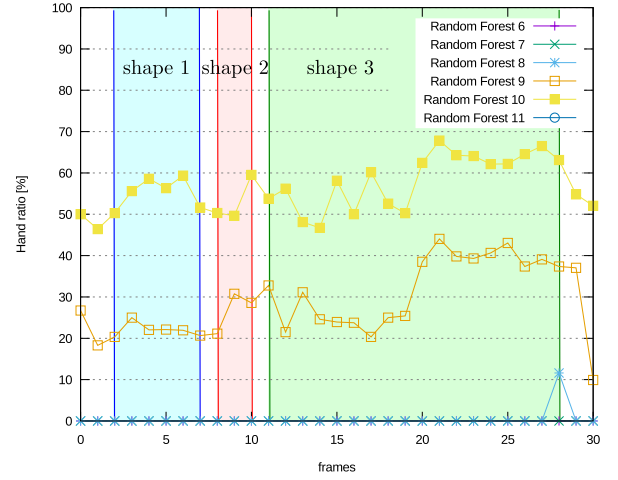
(a) Random forests 1~5 variation



(a) Random forests 1~5 variation



(b) Random forests 6~11 variation



(b) Random forests 6~11 variation

Fig.16: Recognition pattern variation by method 1 for shape 1 to 3

Fig.17: Recognition pattern variation by method 2 for shape 1 to 3

が安定していると言える。これらから複数のランダムフォレストの認識結果の差異から手形状を分離できる可能性を示した。

今後は、教師データ数を増やし、各ランダムフォレストの認識精度向上を行う。これにより形状分離精度が向上し、パラメータが安定すると考えられる。また、同時に形状分離に有効な教示データとパラメータの検討も行っていく。そして、実機での物体受け渡し動作の実験を行う。

参考文献

- 1) 近藤修平, Luis CANETE, 高橋隆行: マニピュレータを搭載した車輪型倒立振子ロボットの制御—拡張状態オブザーバを用いたマニピュレータによる物体受け渡し制御, ロボティクスメカトロニクス講演会 2015, 1P2-G03 (2015)
- 2) 小壁正義, 柴田論, 山本智規: 動作開始遅れと受け取り位置を考慮したロボットから人への物体差し出し運動生成, 日本機械学会論文集, 75-754, 1680/1689 (2009)
- 3) 神代充, 柴田論, 山本智規, 清水顯: 人の好みを考慮した手渡しロボットに関する研究—手部のサインを用いた動的特性の調整—, 人間工学, 41-2, 57/64 (2005)
- 4) 小林太, 丸野由貴, 小島史男: 人間の手姿勢を考慮したハンド/アームロボットによる手渡し動作, ロボティクスメカトロニクス講演会 2013, 1A1-J07 (2013)
- 5) 郎詩萌, Canete.Luis, 高橋隆行: 人支援ロボットの受け渡し動作のための手認識システムの開発—ランダムフォレスト機械学習の適用及び手形状認識の検討—, ロボティクスメカトロニクス講演会 2016, 2A1-20a5 (2016)
- 6) Shotton J., Fitzgibbon A., Cook M., Sharp T., Finocchio M., Moore R., Kipman A., Blake A.: Real-time human pose recognition in parts from single depth images, IEEE Eng. in Computer Vision and Pattern Recognition (2011)
- 7) 根本祐也, 中島克人: 単眼カメラによるハンドジェスチャー認識のための高速手指検出, 情報処理学会第 76 回全国大会講演論文集, 2D-3, 77/78 (2014)