

# 文化や言語の異なりを考慮した HRI(Human-Robot Interaction)

TA ZHENYU†1 坂本篤志†1 張得福†1 下川原(佐藤)英理†1 山口亨†1

**概要:** ロボットが適切に人の示す意図を理解するには文化や言語の違いを考慮したヒューマンロボットインタラクションが必要である。本研究ではロボットと人の共存のために、ジェスチャの意味の違いを考慮した人の意図理解に基づくロボット操作と、母国語の違いが感情認識に与える影響について述べる。

## 1. はじめに

HRI(Human-Robot Interaction)ではロボットが適切に人の示す意図を理解する必要がある。しかし文化や言語が異なると、動作の持つ意味が異なることや、音声に対して感じる感情が異なることがある。例えば、日本で「おいで」と手招きするジェスチャは掌を下に向けて上下に振る動作であるが、これはアメリカでは逆に「あっちに行って」というジェスチャである。また日本人が海外の人が話しているところを聞いた時、実際には怒っていないとも怒っていると認識することがある。このような文化や言語の違いを考慮したヒューマンロボットインタラクションが必要である。本研究ではロボットと人の共存のために、ジェスチャの意味の違いを考慮した人の意図理解に基づくロボット操作と、母国語の違いが感情認識に与える影響について述べる。

## 2. 文化の違いからのジェスチャの差異

### 2.1 システム構成

開発するインターフェースの入出力の流れ、構成を図1に示す。指先検出と指検出、顔の向き、体の向き認識の結果よりジェスチャ分類、動作解析を行う。そこで操縦者の意図理解をして操縦者の要求した指示を認識する。さらに安全のため障害物を検知する環境認識を行い、それらによってモビリティロボット (CHOICA, VECTOR 株式会社製) への命令決定をする。入力機器としては先行研究[1]で示された手法では指先検出にメガネ型ウェアラブルカメラ (view っとめがね, テラテクノス株式会社) を用いた。メガネ型ウェアラブルカメラは両眼の中心に装着車の前景を映すカメラが付いている。顔向き認識には深度センサ搭載カメラ (RealSense SR300, Intel Corporation) を使用し環境認識のためにレーザーレンジファインダー (URG-04LX-UG01, 北陽電気株式会社) を用いた。今回新たに示す手法では指検出にハンドトラッキングデバイス (LeapMotion, LeapMotion inc.) を使用した。

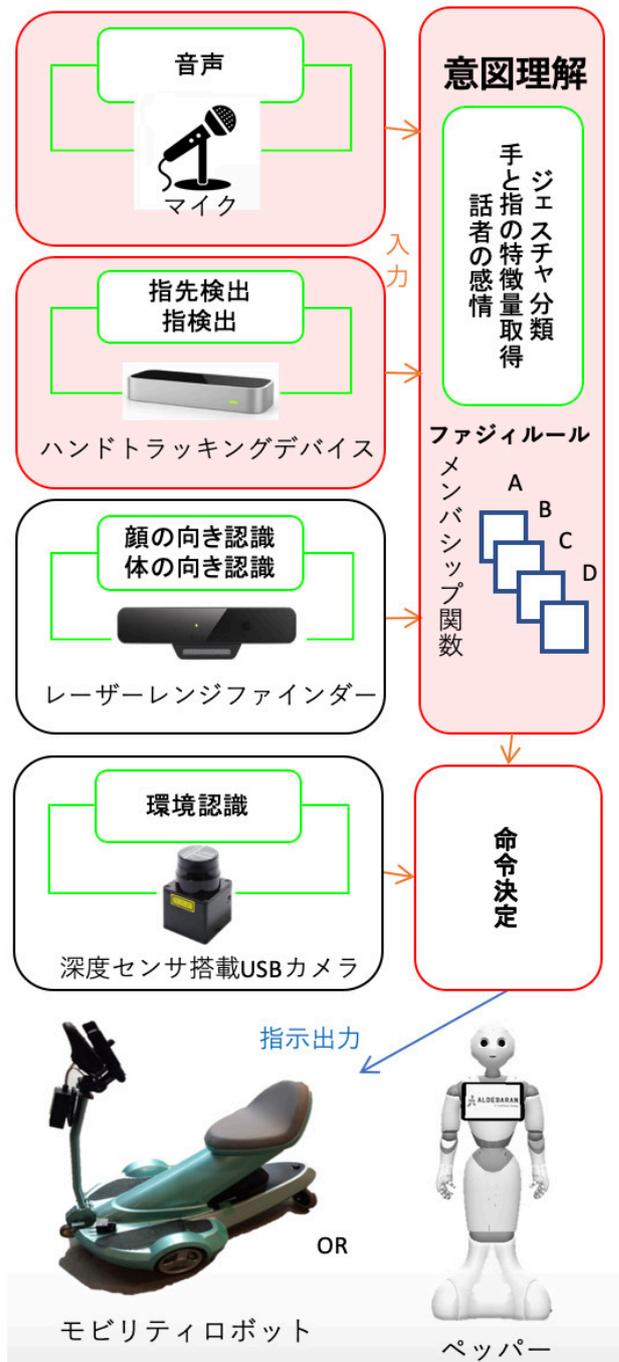


図1 HRCで扱うシステム構成

†1 首都大学東京

## 2.2 操縦者の指示の種類

意図理解のためにジェスチャ分類を行った。操縦者が行うことのできる指示を以下に示す。

### (1) 前方指差しジェスチャ

操縦者がモビリティロボットの前方を向きながら指差しを行なった時はモビリティロボットを前進させる。

### (2) 左右指差しジェスチャ

操縦者がモビリティロボットの正面から左右に寄っている方向に指差しする動作。メガネ型ウェアラブルカメラの場合は顔の向きまで旋回させた後前進させる。ハンドトラッキングデバイスの場合は指の指す方向まで旋回させた後前進させる。

### (3) スクロールジェスチャ

指を差しながら左右どちらかに水平移動させる動作。移動させた方向に 90 度旋回させる。

### (4) カメラ隠しジェスチャ

カメラを隠すことでモビリティロボットを停止させる動作。直感的ではないが、安全面に考慮し確実に停止させることができる。

### (5) 指を差して回すジェスチャ

操縦者が指を差して回した時、モビリティロボットをその差している方向まで旋回させる。旋回中に停止命令があれば旋回を止める。

### (6) 手を広げて前方に出すジェスチャ

図 5 に示すように、手を広げることでモビリティロボットの移動、旋回を停止させる。

### (7) 拳を握るジェスチャ

図 4 に示すように、手を握ることでモビリティロボットの移動、旋回を停止させる。

### (8) 手を交差させるジェスチャ

図 3 に示すように、手を交差させ、ばつ印を作ることでモビリティロボットの移動、旋回を停止させる。

(6), (7), (8) はどれモビリティロボットを停止させるジェスチャである。

## 2.3 文化の差異を考慮した緊急ブレーキ実装

モビリティロボットの操作においてブレーキは緊急性を要するために操縦者の安全を確保する上で重要である。よってジェスチャで操作する場合、ジェスチャの文化的差異

を考慮した実装をする必要がある。図 2 に示すように、複数のジェスチャを OR 回路で繋いでその出力をブレーキにした。このようにすることで誰が操縦者になっても安全に止まることができた。停止命令のジェスチャを中国人 10 人に調査し解答された三つのジェスチャを実装した。

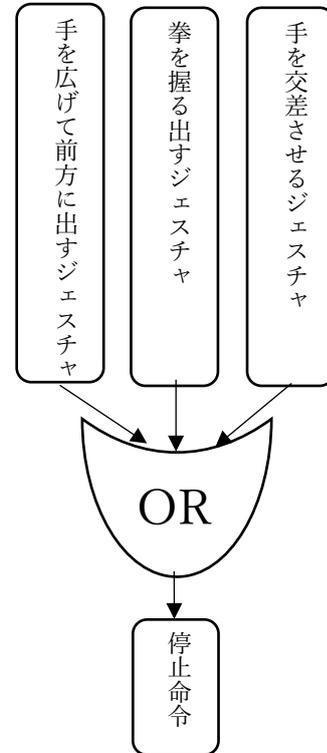


図 2 文化の差異を考慮したジェスチャ分類構成

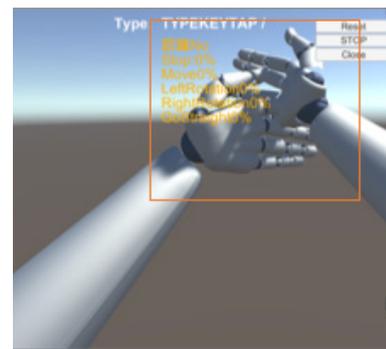


図 3 手を交差させるジェスチャ

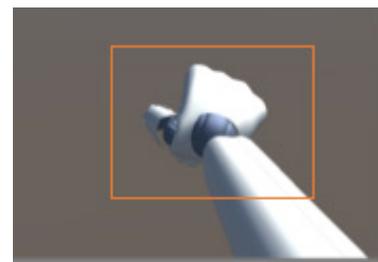


図 4 拳を握る出すジェスチャ

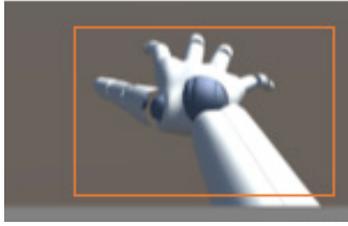


図5 手を広げて前方に出すジェスチャ

## 2.4 手ジェスチャ認識

操縦者が指示を出しているか否かを認識することをここでは指示中認識とする。メガネ型ウェアラブルカメラではGMM 前景分離を活用し指示中認識とジェスチャ認識を行っている。LeapMotionは2台の赤外線カメラと赤外線照射器で構成されている。その赤外線カメラで手の形状を推測する。赤外線は太陽光に含まれている。しかし直射日光を受ける場所であっても認識精度は変わらないことが確認できた。これはLeapMotionが照射している赤外線の周波数でフィルタをかけているためだと考えられる。指先の位置をサンプリング周波数10 Hz、窓幅3データで取得しLeapMotion空間におけるユークリッド距離を計算した。その結果からファジィ推論で認識した。ファジィ推論にすることで操縦者に合わせてメンバシップ関数を用意し、個人差を考慮した認識を行うことができる。ジェスチャ認識では伸ばしている指の本数が一本の時、前方指差しか左右指差しかを指先の向く方向からファジィ推論で認識させた。操縦者の前方を0度として左右に-90度から90度を-1から1とする。

## 3. 文化の違いからの感情認識の差異

### 3.1 クロスカルチャーに基づく感情の認識

クロスカルチャーに基づく感情を認識する先行研究[2][3][4]として、クロスカルチャーに基づく顔感情認識という研究がある。この研究は顔データベースを使用し、事前にサンプルで「これは楽しい顔」、「これは怒っている顔」と学習させておいて認識させている。しかしながら、クロスカルチャーに基づく音声による感情の認識はまだ未知見が不足している。そこで、私は異なる文化的背景の実験者による感情認識率を比較することで、文化の違いが感情の認識に与える影響について考察する。

### 3.2 実験データの収集

中国の学生2名(男性:1名, 女性:1名)に録音してもらい、5つの感情を含む中国語音声データを得た。国立情報学研究所の感情音声を聞き、手作業で5つの感情を判定し、日本語音声データを得た。学生12名(日本人:6名, 中国人:6名)に録音した60個(中国語:30個, 日本語:30個)の音声を聞いてもらい、感情を判定させた。表1は被験者の回答の個数を表している。その結果、実験者はラ

ンダムな推測レベルを超えて他の言語の発話感情を認識することができ、母語の発話感情に対する実験者の認識率は他の言語よりも高いことが分かった。

表1 日本人と中国人の解答個数

emotion	Chinese(%)					Japanese(%)				
	A	F	H	S	P	A	F	H	S	P
Chinese	A	100				8.3	66.7	41.7	8.3	8.3
	F	66.7	8.3	25		8.3	33.3	16.7		16.7
	H		66.7		50	16.7		25	16.7	25
	S		33.3	16.7	75		16.7	16.7	75	8.3
	P			8.3		41.7	8.3		33.3	41.7
Japanese	A	58.3		25		8.3	75		8.3	8.3
	F	41.7	16.7	16.7		8.3		50		8.3
	H	25		41.7	16.7	41.7	8.3	8.3	66.7	16.7
	S		58.3		66.7		25	16.7	91.7	16.7
	P	16.7		16.7		41.7	16.7	16.7	8.3	58.3

### 3.3 実験データの解析

実験データを統計解析ソフト「SPSS」で解析する。カイ二乗分析により、性別認識率のマッピングを実行し、図6のように、中国語と日本語の音声認識における感情認識の割合の折れ線グラフで一貫した傾向があるため、性別は認識に影響を与えないことが分かった。次に、実験者の認識結果についてMDS分析を行い、図7と図8、図9にこの結果を示す。日本語話者の感情を認識するとき、日本語を学習する中国人は日本語を学習しない中国人よりも日本人に近いことを発見した。

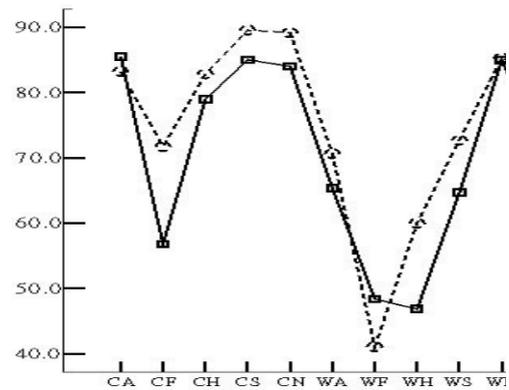


図6 性別による認識率

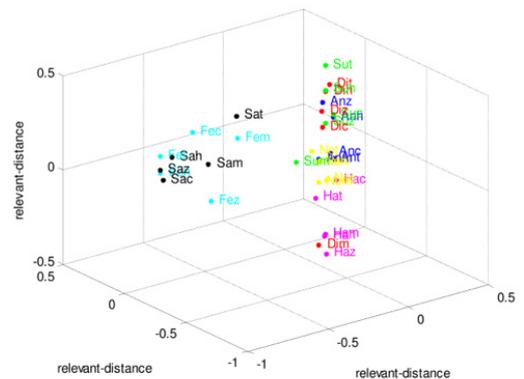


図7 日本語を学習していない中国人

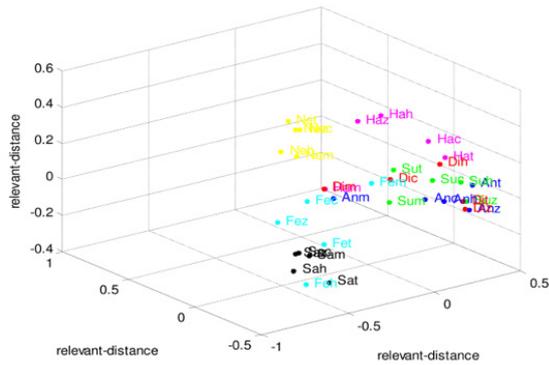


図 8 日本語を学習している中国人

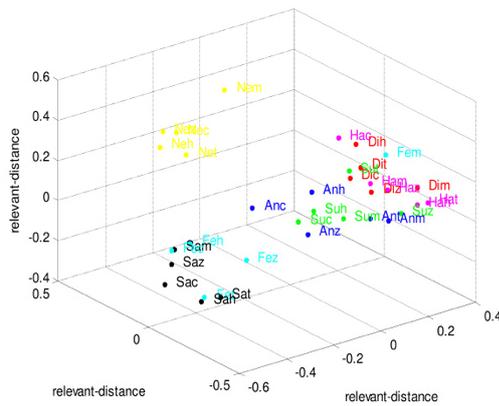


図 9 日本人

#### 4. おわりに

本論文では、ジェスチャの文化の違いを考慮した人の意図理解に基づくロボット操作と、母国語の違いが感情認識に与える影響について述べた。ジェスチャ認識では停止命令を複数認識することで文化の違いに対応させた。感情認識では日本人と中国人で差異があり、日本語を学習している中国人では日本人に近くなることが分かった。今後の研究において、これらの知見を統合することによって、文化の違いによって起こる差異を考慮したコミュニケーションロボットを構築する。

#### 参考文献

- [1] Eichi Tamura, Yoshihiro Yamashita, Taisei Yamashita, Eri Sato-Shimokawara, and Toru Yamaguchi. Movement operation interaction system for mobility robot using fingerpointing recognition. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 2017, Vol.21, No.4, pp.709-715.
- [2] Ekman, P., & Friesen, W. V.. Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 1971, 17(2), 124-129.
- [3] Anne Wichmann. The Attitudinal Effects of Prosody and How They Relate to Emotion [C], ISCA Workshop on Speech and Emotion, 2000.
- [4] Michael Biehl, David Matsumoto, Paul Ekman, Valerie Hearn, Karl Heider, Tsutomu Kudoh, Veronica Ton. Matsumoto and Ekman's Japanese and Caucasian Facial Expressions of Emotion