

# 人パラメータ推定に基づく 移動ロボットの動作パラメータ学習に関する研究

○平山 三千昭(早大) 亀崎 允啓(早大/JST さきがけ) 濱田 太郎 斎藤 喬介 金田 太智  
菅野 重樹(早大)

## 1. はじめに

近い将来、自律移動ロボットは、パーティ会場のような比較的静的だが混雑している環境や、駅の通路のような人が動いている動的な環境など、さまざまな環境に導入されサービスを行うことが期待されている。我々は環境に合わせた自律移動制御モデルを開発し、それらを統合することにより、多様な環境条件下でスムーズな移動を実現することを目指してきた。ロボットが人への接近や声かけ、能動的な接触などの積極的な行動を選択することで、混雑環境でも効率的な移動が可能になると考え、ロボットの働きかけに対する人の反応を予測した軌道制御手法である inducible Social Force Model (iSFM) を開発した [1]。それにより、複数歩行者の移動予測、複数軌道候補生成、および移動効率と人への影響を考慮した実時間での軌道制御が可能になった。

一方、ロボットの動作パラメータ群は、導入環境に合わせて手動調整をしていたため、異なる環境での再利用や場面遷移への対応が難しく、スムーズに動作しないことがわかってきた。例えば、図 1 は静的かつ混雑しているポスター会場を想定しており、ロボットは手前のゴールまで移動しようとしている。図 1(a) では、ロボットは被験者  $\alpha$  の回避軌道を生成するが、図 1(b) のように、 $\alpha$  は我々の想定以上に体を左右に揺らしていたため、ロボットは移動するものと誤認識し、チャタリングを起こした。どれだけ事前調整をしても想定外の場面は出てきてしまうため、ロボットが環境に合わせて自動でパラメータ調整できれば、スタックや無駄な回避をせずに、様々な環境でスムーズな移動が可能になると考えられる。

ロボットが動作パラメータの自動調整をするためには、①ロボットの行動評価機能、②人の動作パラメータ推定機能、③ロボットの動作パラメータ調整機能が必要だと考えられる(図 2)。①まず、ロボットの行動が人々に許容され、かつロボットの行動目的に対して適当かどうかを評価できる必要がある。例えば人々がお互いに距離をとっている環境ではロボットも過度な接近をするべきではなく、人を優先しすぎてロボットが移動できない行動も取るべきではない。適当な行動を取るための評価指標を具備する必要がある。②次にロボットは人との相互作用や観察を通して、人の動作パラメータを推定する。センサで推定可能なパラメータの例としては、他者との距離感、接触への不快感、快適速度などが考えられる。③①と②により、人パラメータに応じたロボットの動作パラメータの自動調整が可能になると考えられる。

①ロボットの行動評価指標で代表的な例としてパーソナルスペース(PS)がある。PS は人が他者に侵入されたくない心理的な領域で、ロボットと人の距離感を決める指標として用いられる [2]。②人移動予測

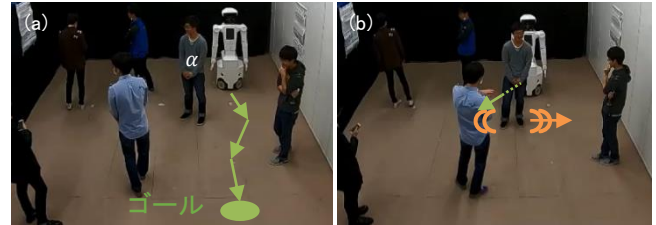
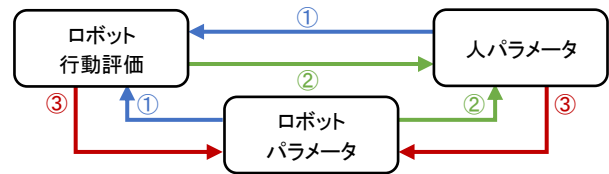


図 1 想定していない人の動作によりロボットがチャタリングを起こす場面



- ①人パラメータ×ロボットパラメータ→行動評価指標導出
- ②行動評価指標×ロボットパラメータ→人パラメータ推定
- ③人パラメータ×行動評価指標→ロボットパラメータ調整

図 2 統合パラメータ調整システムの要求仕様

モデルでは Social Force Model [3] など数多くの手法が存在し、それらのパラメータを推定する手法も提案されている [4]。③ロボットの動作パラメータ調整では、深層強化学習による試行錯誤的な調整手法などが提案されてきた [5]。しかし、①-③を統合し、行動評価指標をもとにパラメータ推定と調整をするような統合的なパラメータ調整の枠組みは我々の調査した限りでは提案されていない。

以上より、本研究では、様々な人共存環境でのロボットのスムーズな移動を実現するために、人の動作パラメータ推定機能とロボットの行動評価機能を有し、動的に行動調整を行う統合パラメータ調整システムを開発する。基礎検討となる本研究では、Unity を用いて人を含む様々なシミュレーション環境を模擬的に構築し、ロボットのパラメータ調整を行う。

## 2. 統合パラメータ調整システム

2.1 節では制御手法、2.2 節では調査対象のパラメータと環境条件、2.3-2.5 節で図 2 に示した提案のシステム①-③の各機能について説明する。

### 2.1 制御手法

本研究では iSFM [1] によりロボットを制御する。ロボットを  $R$ 、移動予測対象の人を  $H_i$  としたとき、 $H_i$  の動きは式 (1) でモデル化される。

$$f_{all}^{H_i} = f_{R,ph}^{H_i} + (f_{R,so}^{H_i})\delta_{aw}^{H_i} + f_{R,C}^{H_i} + f_{H_j}^{H_i} + f_{obj}^{H_i} + f_{wp}^{H_i} \quad (1)$$

$f_{all}^{H_i}$  は合力、 $f_{R,ph}^{H_i}$  は  $R$  からの物理的な衝突力、 $\delta_{aw}^{H_i}$  は認知度 ( $\delta_{aw}^{H_i} = \{0,1\}$ )、 $f_{R,so}^{H_i}$  は  $R$  からの心理的な斥力、 $f_{R,C}^{H_i}$  は  $R$  の能動接触により加わる力、 $f_{H_j}^{H_i}$  は他エージェント

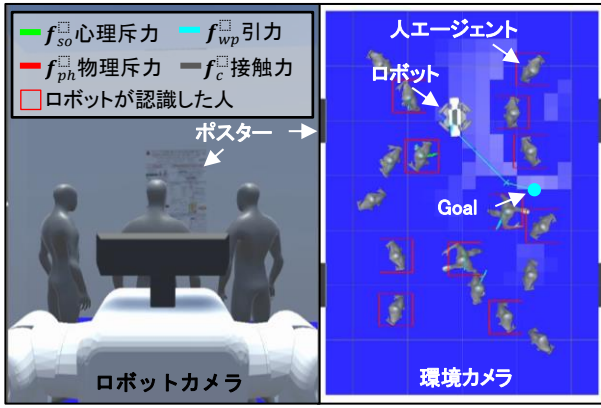


図3 シナリオ：ポスター会場

ト $H_j$ からの斥力,  $f_{obj}^{H_i}$ は壁からの斥力,  $f_{wp}^{H_i}$ は目的地からの引力を表す。ロボットの制御は式(2)の $R$ と $H_i$ を入れ替えた $f_{all}^R$ で行う。詳しくは[1]を参照されたい。

## 2.2 対象パラメータおよび環境条件

式(1)の各項にはそれぞれパラメータが含まれており、それらを環境に応じて推定および調整する必要がある。パラメータ群の中で、今回はロボットと人の距離感を決める $f_{so}$ に関するパラメータ $k_{so}$ と $S_A$ に着目し、 $k_{so}^{H_i}$ と $S_A^{H_i}$ の推定に基づく $k_{so}^R$ と $S_A^R$ の調整を行う。式(2)に $H_i$ の $f_{so}$ を示す。

$$f_{R,so}^{H_i} = k_{so}^{H_i} e^{\left(\frac{r_{R,H_i} - d_{R,H_i}}{S_A^{H_i}}\right)} v_{R,H_i} \quad (2)$$

ここで、 $d_{R,H_i}$ は $R$ と $H_i$ 間の距離、 $r_{R,H_i}$ は $R$ と $H_i$ の体幹半径の合計、 $v_{R,H_i}$ は $R$ から $H_i$ への単位ベクトルを表す。なお、 $k_{so}$ は $f_{so}$ の大きさ、 $S_A$ は $f_{so}$ の及ぶ範囲を決めるパラメータと理解できる。

本研究では、人口密度と移動者割合を操作することで環境条件を変化させる。場所、日時、文脈、人々の行動目的など様々な要因が距離感に影響すると考えられるが、本研究では簡単かつ一般化可能な環境条件として、上記の2つに着目する。

- 人口密度：一般的に公園のように広い空間で人数が少ないときは他者と距離を取ることができるが、エレベータのように狭い空間では他者の接近を許容せざるをえなくなる。そのため人口密度は他者との距離感を決める条件の一つだといえる。
- 移動者割合：多くの人が移動している動的な環境では衝突による転倒など危険が伴うため、満員電車ほど他者とは接近できない。そのため、環境の動的度合いを表す移動者割合によって、他者との距離感が変化すると考えられる。

以上より、環境条件として人口密度と移動者割合に着目する。なお、実験では部屋の大きさを固定し人数を変化させることで人口密度を変化させる。

実験環境はロボットを導入可能かつ人の行動原理をモデル化可能と考えられる、図3に示す閉空間のポスター会場とする。右図の赤枠はロボットが視認可能な人を表し、体幹中心から出ている線が式(1)の各力を表している。人もロボットもポスターの前で一定時間停止して閲覧し、その後ランダムに別のポスターに移動することを繰り返す。図のシーンでは

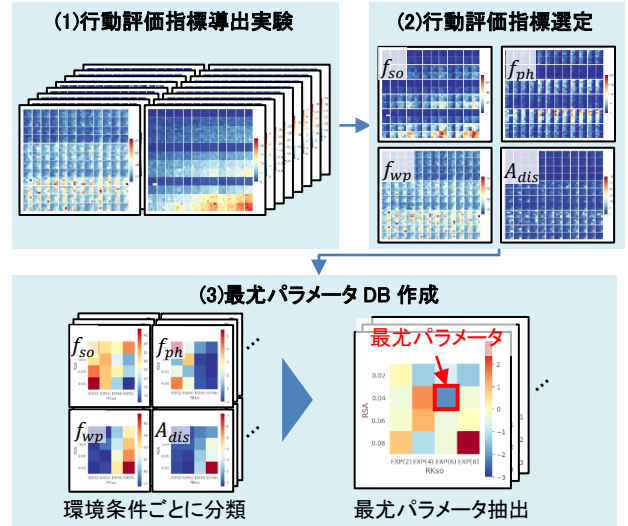


図4 行動評価指標導出フロー

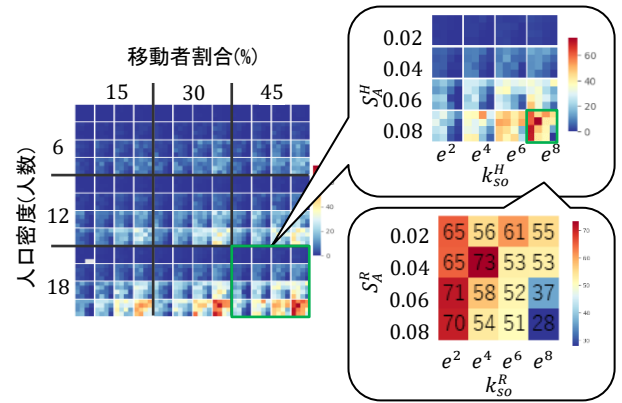


図5 行動評価指標の例： $f_{so}^{RH}$

ロボットが左上のポスターの鑑賞を終え、右上のポスターが鑑賞可能な位置に移動しようとしている。

## 2.3 行動評価指標導出

ロボットの行動評価指標導出の手順を図4(1)～(3)に示す。まず、(1)行動評価指標導出実験を行う。次に、(2)実験で得られた複数の評価指標候補の中から、ロボットの用途に応じて評価指標を選定する。(3)その後、選定した評価指標と環境条件から、それぞれの環境での適切なロボットのパラメータを抽出することで、最尤パラメータDBを作成する。

(1) 行動評価指標導出実験：環境条件(人数、移動者割合)と人およびロボットパラメータ( $k_{so}^{H_i}$ ,  $S_A^{H_i}$ ,  $k_{so}^R$ ,  $S_A^R$ )を総当たりで変化させ、様々な環境でのロボットのふるまいデータを収集することで、行動評価指標を導出する。なお、ここではロボットの行動評価に使用する可能性のあるデータを行動評価項目候補と呼んでおり、移動効率や移動距離、接触回数、心理斥力など様々なものが考えられる。

実験は人が圧迫感を感じない程度の広さの5m × 7mの環境で行い、ポスター4枚を左右の壁に3mおきに配置した。人の数は6人(低密度)・12人(中密度)・18人(高密度)として、15%、30%、45%を移動させることで移動者割合を変化させた。人とロボットについて $k_{so}=e^8$ ,  $S_A=0.08$ をデフォルト値として $k_{so}=\{e^2, e^4, e^6, e^8\}$ ,  $S_A=\{0.02, 0.04, 0.06, 0.08\}$ を調査した。

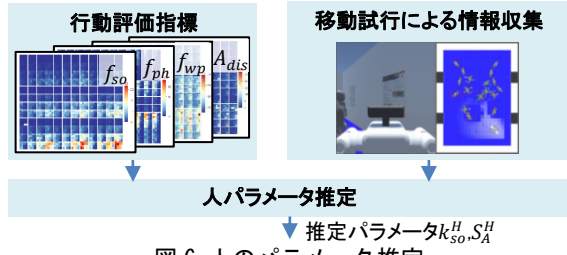


図 6 人のパラメータ推定

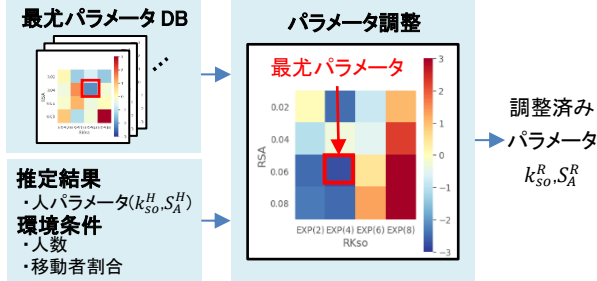


図 7 動的パラメータ調整

[3]では、 $k_{so}=2000 \approx e^{7.6}$ 、 $S_A=0.08$ と設定していたため、これを包含するキリのいい値をデフォルトとし、等間隔でそれより小さい4点をとることで、接近を許容していったときの動作を評価することにした。以上より、 $3 \times 3 \times 4 \times 4 \times 4 \times 4 = 2304$ パターン、実験時間は5分間として合計192時間分のデータを取得した。行動評価項目候補は、ポスター間の移動時間、移動距離、回避量  $A_{dis}^R$ 、 $f_{so}$ 、 $f_{wp}$ 、 $f_c$ 、 $f_{ph}$ など考える16種類のデータとし、それらを積分することで指標候補を算出した。また、 $r_{R,H_i}=0.6$ とし、その他パラメータは[1]と同じ値にした。

図4(1)が実験で得られた16種類の行動評価指標候補である。さらに図5に、ロボットが人に与えた  $f_{so}^{RH}$ の結果を例示する。最も大きいグラフはX軸が移動者割合、Y軸が人口密度である。それぞれの移動者割合・人口密度での人パラメータ  $k_{so}^H$ 、 $S_A^H$ がその中に格納されており、さらにその中にロボットの  $k_{so}^R$ 、 $S_A^R$ を変化させたときの  $f_{so}^{RH}$ の積分値が格納されている。 $f_{so}^{RH}$ の値が小さいほど色が青く、大きいほど赤い。

(2) 評価指標選定：ロボットの行動評価指標に正解はないため、目的に応じて選択する。今回は指標候補から、人とロボット双方の自然な振る舞いに着目し、人に与えた心理的・身体的影響である  $f_{ph}^{RH}$ 、 $f_{so}^{RH}$ 、ロボットの移動効率を表す  $f_{wp}^R$ 、 $A_{dis}^R$ の4つを選択した(図4(2))。これらはトレードオフの関係にあるため、適切にバランスすることで人・ロボット双方を考慮した評価指標が導出可能だと考えられる。

(3) 最尤パラメータ DB 作成：トレードオフの関係にある  $f_{ph}^{RH}$ 、 $f_{so}^{RH}$ と  $f_{wp}^R$ 、 $A_{dis}^R$ をバランスした最尤パラメータを環境ごとに導出し、DBを作成する。まず行動評価指標のテーブルを環境条件と  $k_{so}^H$ 、 $S_A^H$ ごとに分類する。その後、式(3)に示す最尤パラメータ導出式を用いて、分類した条件ごとのロボットの最尤パラメータを計算し、DB化する。

$$k_{so}^R, S_A^R = \underset{k_{so}^R, S_A^R}{\operatorname{argmin}} \sum \alpha \cdot Z(s^*) \quad (3)$$

$$s^* \ni f_{so}, f_{ph}, f_{wp}, A_{dis}$$

式(3)ではそれぞれの条件での評価指標の値  $s^*$  からそれらのZスコアを求め、係数  $\alpha$  を掛け合わせた値を足し合わせる。その値が最小のときのパラメータを最尤パラメータ  $k_{so}^R, S_A^R$  として導出する。なお、今回は重み  $\alpha$  を  $\alpha_{f_{so}}=0.5$ 、 $\alpha_{f_{ph}}=1.0$ 、 $\alpha_{f_{wp}}=1.0$ 、 $\alpha_{A_{dis}}=1.0$  とした。 $\alpha_{f_{so}}$  のみ0.5としたのは  $f_{so}$  の寄与度が他の項より大きかったためである。計算の結果、図4(3)の条件では  $k_{so}^R=e^6$ 、 $S_A^R=0.04$  が最尤パラメータだとわかった。

## 2.4 人パラメータ推定

概要を図6に示す。今回、環境内の人数と移動者割合は既知として、人パラメータのみ推定する。ロボットは適当なパラメータ  $k_{so}^R$ 、 $S_A^R$  で実際に人とインタラクションを繰り返しながら情報収集を行う。収集データと2.3節の行動評価指標から、適当なタイミングで式(4)による人パラメータ推定を行う。

$$k_{so}^H, S_A^H = \underset{k_{so}^H, S_A^H}{\operatorname{argmin}} \sum [Z\{\operatorname{Error}(s^* - s)\}] \quad (4)$$

$$s, s^* \ni f_{so}, f_{ph}, f_{wp}, A_{dis}$$

$s^*$  は設定した  $k_{so}^R, S_A^R$  と既知の環境条件(人数と移動者割合)における、各  $k_{so}^H, S_A^H$  候補での評価指標のスコア、 $s$  は移動試行で収集したデータを表す。 $s^*$  と  $s$  の誤差 ( $\operatorname{Error}$ ) を標準化したZスコアの合計が最小のときの  $k_{so}^H, S_A^H$  を人パラメータとして推定する。

## 2.5 ロボットの動的パラメータ調整

最後に、ロボットは推定した  $k_{so}^H, S_A^H$  と環境条件をキーとして最尤パラメータ DB から適切な  $k_{so}^R, S_A^R$  を探索し、最尤パラメータとして選択する(図7)。

## 3. 実証実験および考察

### 3.1 実証実験

人パラメータ推定とロボットのパラメータ調整を実時間で実行することで、提案システムの動作検証を行う。2.3(1)節と同じ環境で人数は18人、移動者割合は45%、人パラメータは他者の接近をある程度許容した  $k_{so}^H=e^4$ 、 $S_A^H=0.04$  に設定する。ロボットの初期パラメータは未知環境での移動であるため安全側に振り、デフォルト値  $k_{so}^R=e^8$ 、 $S_A^R=0.08$  とする。実験時間は5分間として、ロボットは直前の1分間のデータを使用して調整を4回繰り返す。以上の実験を12セッション行う。実験結果は収集したデータ  $s$  のZスコアとロボットのポスター間の移動回数で評価する。

### 3.2 実験結果

図8と図9にZスコアと移動回数の時系列推移を示す。調整回数0は調整前、1-4は調整を実施した回数を表す。図8のスコアは高いほど評価が悪く、低いほどよい。パラメータ調整前は安全のため人と距離を取っており、 $f_{wp}^R$  と  $A_{dis}^R$  のスコアは悪い。一方で人に心理的・物理的に干渉しないため  $f_{ph}^{RH}$ 、 $f_{so}^{RH}$  は良くなっている。1回目の調整以降はそれぞれ0周辺を推移しており、各スコアをバランスしたパラメータに調整できたことがわかる。図9の移動回数は多いほどよい。パラメータ調整前は1分間で平均1.6回移動しているが、調整後は平均3回以上移動している。パ



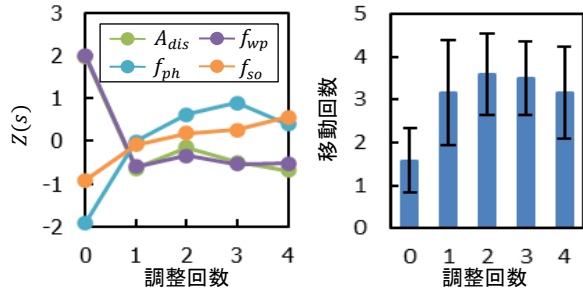


図8 Zスコア

図9 移動回数

表1 推定結果と調整結果

調整回数	人パラメータ $k_{so}^H, S_A^H$	ロボットパラメータ $k_{so}^R, S_A^R$
調整前	(正解値) $e^4, 0.04$	(初期値) $e^8, 0.08$
1回目	(推定値) $e^6, 0.02$	(調整値) $e^2, 0.02$
2回目	(推定値) $e^6, 0.04$	(調整値) $e^2, 0.06$
3回目	(推定値) $e^4, 0.04$	(調整値) $e^8, 0.02$
4回目	(推定値) $e^4, 0.06$	(調整値) $e^6, 0.04$

※人が  $e^4, 0.04$  のときの最適なロボットパラメータは  $e^8, 0.02$

ラメータ調整によって適切な距離感を学習し、スムーズに移動可能になったといえる。

あるセッションでの  $k_{so}^H, S_A^H$  推定結果と  $k_{so}^R, S_A^R$  調整結果の推移を表1に、調整前後のキャプチャを図10に示す。表1の2行目は人の正解パラメータとロボットの初期パラメータである。調整1回目で  $k_{so}^H=e^6, S_A^H=0.02$  と推定し、 $k_{so}^R=e^2, S_A^R=0.02$  と調整、2回目以降も同様に推定と調整を繰り返した。図10では、 $t=47$  でロボットは右上のポスターを閲覧するために軌道を生成し、移動を開始する。 $t=57$  で左方から移動してきた人  $\alpha$  に進路を阻まれ、初期値 ( $k_{so}=e^8, S_A=0.08$ ) では接近できないため、移動不能になった。ここで1回目のパラメータ調整が入り  $k_{so}=e^2, S_A=0.02$  に調整されたため、 $t=63$  で  $\alpha$  と壁の間をすり抜けることができ、 $t=69$  でゴールに到達した。

### 3.3 考察

実験の結果、提案システムが正常に動作することが確認できた。さらに、パラメータ調整によって取得データのZスコアもバランスがとれ、移動回数が改善することがわかった。一方、安全側に設定した初期パラメータでは大きな力を受けて急な加減速することがあったため、特に混雑環境では危険であった。iSFMのように仮想的な力に基づき制御する手法では、適切なパラメータ調整が重要なことが確認された。

一方で、提案システムにはいくつかの改善点が見られた。まず、提案システムでは直前1分の結果のみを用いてパラメータ調整を繰り返すため、パラメータの継続的な改善ができない。また、行動評価指標と直前1分で取得したデータの比較のみで人パラメータを推定するため、推定精度があまり高くない(表1)。今回導出した行動評価指標は離散的なデータしか扱えないが、線形補間等により連続的な数値が扱えるようになれば、さらに高精度な推定や調整が可能になると考えられる。また、人同士のインタラクションの観察や外部カメラと連携することで、推定精度のさらなる向上が可能だと考えられる。

また、低密度の環境設定であっても局所的に高密

(a)調整前  $t=47$  (b)調整前  $t=57$  (c)調整後  $t=63$  (d)調整後  $t=69$

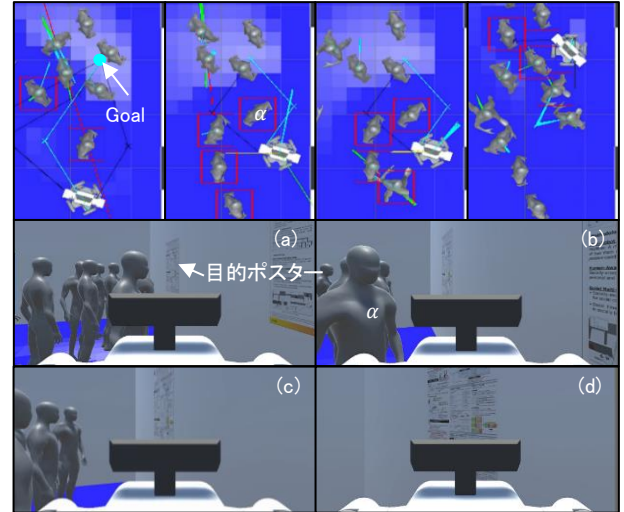


図10 実験シーン

度になり、ロボットが移動不能になることや、1分おきに調整したためロボットが人と壁に挟まれて30秒間動けなくなる場面があった。そのため、提案した人口密度と移動者割合に応じた大局的な環境条件でのパラメータ調整だけでなく、局所的な場面に応じた適切なタイミングでの調整が必要なが分かった。

### 4. まとめ

本研究では、様々な人共存環境でのロボットのスムーズな移動を実現するため、実験による行動評価指標の導出と、それを用いた人の動作パラメータ推定機能とロボット自身の行動評価機能を有する、動的パラメータ調整システムを開発した。実証実験の結果、パラメータ調整により取得データのZスコアがバランスがとれ、移動回数が向上したため、本提案の有効性が確認された。今後は、行動評価指標の使い方に関するさらなる検討、局所的な状況に応じたパラメータ調整、調整タイミングの調整、 $f_{so}$ のパラメータ以外のパラメータへの本提案の応用とともに、実環境への導入を目指す。

### 謝辞

本研究の一部は、NEDO「次世代ロボット中核技術開発」、JST さきがけ(JPMJPR1754)、ならびに、早稲田大学理工学術院総合研究所の支援を受けて行われました。ここに謝意を表します。

### 参考文献

- [1] 金田太智, 他: "Inducible Social Force Modelを用いた複数人移動予測に基づく混雑環境下でのロボットの接近・接触移動", Robomech'20, pp. 2A1-C01, 2020-5.
- [2] L. Takayama, et al, "Influences on proxemic behaviors in human-robot interaction," 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2009.
- [3] D. Helbing, et al, "Social Force Model for Pedestrian Dynamics," 1998.
- [4] Z. Wan, et al, "A Learning based Approach for Social Force Model Parameter Estimation," IEEE, 2017.
- [5] Y. F. Chen, et al, "Socially Aware Motion Planning with Deep Reinforcement Learning," 2018.