

# 複雑な実環境下での察するインタフェースに関する研究

代表研究者 金子正秀 電気通信大学大学院情報理工学研究科 教授

## 1 はじめに

人間共存型ロボットが我々の日常生活環境の中に入り、人間のパートナーとしての役割を果たしていくためには、様々な物が存在し、複数の人間がまちまちに往来する複雑環境下で、周囲の環境の動的変化を認識し、周囲の人間の位置や動きを的確に把握する必要がある。その上で、人間の活動に対し様々な支援を行い、人間のパートナーとして、人間の日常活動をより便利で快適なものとなるようにすることが期待される。このような、ロボットと人間との関わり合いに関する研究として、筆者らは「察するインタフェース」の概念を提唱し、研究を進めてきている。「察するインタフェース」とは、人間が明示的にロボットに指示を行わなくとも、ロボット自らが相手の人間の状態やインタラクションの状況を把握し、更には人間の内在的な行動意図或いは人間の期待を推測し(これらをまとめて、「察する」という言葉で代表している)、自身の行動を自律的に決めていくようなインタフェースである。

本報告では、複雑な実環境下での察するインタフェースに関する研究として、次の2つの研究成果を報告する。

- (1) 同行者との相対的位置関係を考慮した日常環境下でのロボットの自律的移動
- (2) 受取人の状態に応じて割込み行動を変化させる配達ロボット

## 2 同行者との相対的位置関係を考慮した日常環境下でのロボットの自律的移動

### 2-1 関連の研究動向

自律移動ロボットに関しては、ロボット単独での自律移動時における障害物回避や経路探索に関する研究が活発に行われてきている[1][2][3]。また、複数台のロボットを連携して移動させる研究も行われている。これらとは別に、人間との係り合いの中でロボットを活用していくための研究も盛んになっている[4]。一般に、人間と共存し、人間と協調して動作するようなロボットは人間共存型ロボット、或いは、人間協調型ロボットと呼ばれている。

一方、日常生活空間において人間と共存し、人間と行動を共にするようなロボットにおいては、部屋や廊下の形状や障害物を認識するだけでなく、周囲の人物との係り合いの中で自らの移動形態、移動経路を自律的に決めていく機能が必要となる。このような機能の基本として、特定人物に自動的に追従して移動する機能が挙げられる[5]。お伴ロボット ApriAttendant™ [6]も報告されている。顔、服、人物下部領域の各々を追従部位とした手法が各種提案されている[5]が、特定人物への追従機能に限られている。これに対し、人間と行動を共にして、日常生活空間を移動するロボットを考えた場合、先行する人間の動きに追従するだけでなく、人間と同行し、同行者の動きと周囲環境(障害物や他歩行者)の状況に応じて、同行者との位置関係を考慮しながら自律的に移動できる機能が不可欠である。例えば、人間が2人で廊下を歩行する際、一般的には横に並んだ「並走」の形態で歩行する。しかし、対向歩行者とすれ違う際には、道幅が狭くなるため「縦走」の形態に自然に移行する。この形態変化については、2人のどちらが先に進むのかを互いに瞬時に判断して行っている。このような動作を自律的に行うことができるロボットに関する研究については、まだ取り組みが始まった段階であり、従来研究としては花井らの研究[7]、小林らの研究[8]にとどまっている。花井ら[7]は同行するユーザの動きに合わせて、並走→縦走、縦走→並走の移動モードを予め幾つかロボットに用意しておき、これらのモードを自律的に切替える方法を提案している。この研究では、ロボットとユーザの位置関係は大きく分けて並走と縦走の2パターンであるため適用環境が限られてしまう。また、小林ら[8]は並進する介護者の進みたい方向を察知する自律移動車椅子を提案している。一人の介護者との移動を対象とし、ロボットは介護者と並走するモードと介護者に追従するモードを切替えながら移動する。しかし、実際の生活空間において、ロボットと同行者との移動形態は複雑に変化し得る。そこで本報告では、複数の人物がいる、より実際の生活空間に近い環境の中で、同行者と並走する、同行者の後方を追従する、同行者の前にも位置するなど相対的位置関係をさまざまに変化させながら、安全に自律的移動をすることができるロボットシ

システムを提案する。

## 2-2 移動ロボットの概要

### (1) 本研究での問題設定

本研究で検討する手法の前提条件は以下の通りである。

(1) ロボットの視野内に常に同行者が位置する。

(2) 人物は成人に限り、スカートではなくズボンを着用する。

(1)の条件により、ロボットは同行者を見失うことは無いとする。(2)の条件により、腰と脚部分の距離情報による人物検出を可能とする。

### (2) 想定する環境

本報告では、部屋や廊下など段差のない室内を動作環境とする。使用する移動ロボットを図1に示す。ロボットにはLaser Range Finder (LRF) を人間の腰の高さ(110[cm])と膝の高さ(30[cm])に設置する。人物検出の精度はLRFの高さに依存するため、本研究では成人を対象とした。LRFはロボットの前方向270度、半径10[m]の範囲の距離情報を取得することができる。

### (3) ロボットが行う処理の流れ

ロボットが行う処理の流れを図2に示す。まず、LRFの距離情報から人物検出を行う。そして、周辺人物と同行者の追跡を行い、同行者との相対位置関係を認識する。同行者の向きの認識を行い、同行し続けるための最適な位置関係を決定し移動する。

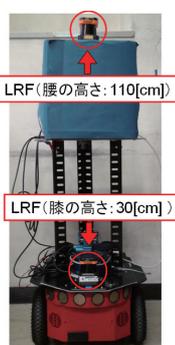


図1 移動ロボット

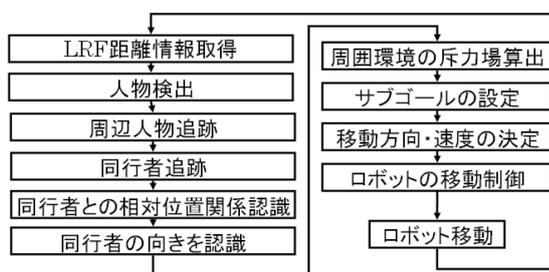
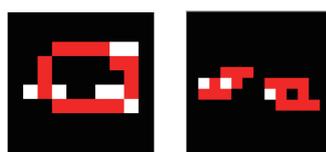


図2 処理概要

## 2-3 人物検出・追跡・同行者の向き

### (1) 人物検出

壁や障害物などが存在する室内環境からロボットが人物のみを区別できるようにするために人物検出を行う。LRFより得られる距離情報を2次元画像上にマッピングしX-Z平面画像を作成する。X-Z平面画像における1画素は、実環境において5[cm]×5[cm]の大きさに相当する。腰の高さに設置したLRFから作成したX-Z平面画像からは、人間の腰周りに特徴的な三日月型の形状を観察することができる。腰回りの形状は本来楕円状であるが、LRFではLRFに面している部分のみが検出されるので、三日月型の形状となる。X-Z平面画像に対して楕円フィッティングを行うことにより人間の腰部分の位置を検出する。また、膝の高さに設置したLRFからも、脚の形状の特徴らしき三日月型の形状を2ヶ所観察できる。腰部の検出と同様にして楕円フィッティングを行うことにより脚の検出を行う。図3に腰部と脚部の楕円フィッティングの結果例を示す。



(a)腰部 (b)脚部

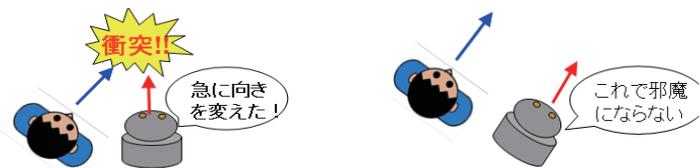
図3 楕円フィッティング

図 3(a)は  $X-Z$  平面画像（腰部分）に楕円フィッティング処理を行った結果である．図 3(b)は脚検出用の LRF より得た  $X-Z$  平面画像に楕円フィッティングを行い，脚を検出した結果画像である．フィッティングによる検出精度を向上するため，長径が 60[cm]以下の楕円を人物の腰候補とし，30[cm]以下の楕円を脚の候補とする．腰部と脚は  $X-Z$  平面画像上ではだいたい同じ位置に存在すると考えられるため， $X-Z$  平面画像(腰部分)から検出された人物候補領域の中心から 35[cm]以内の範囲に， $X-Z$  平面画像(脚部分)より脚候補が検出された場合，ロボットはそこに人間が位置すると認識する．

### (2) 同行者の追跡

複数人物がいる環境において，頑健に同行者のみを追跡するために人物追跡を行う．追跡処理にはパーティクルフィルタ[9]を使用する．これにより少々のおクルージョンや人物検出の失敗にも頑健な追跡を行うことが出来る．複数の人物を検出した場合はロボットに一番近い領域を同行者と設定し，現フレームとの時間差に応じ重み付けをする．過去 11 フレーム(約 600[msec])の累積した位置情報を使用することにより，周辺人物とのすれ違いによって起きるおクルージョンへの対応を可能とする．人物の中心位置の計算には現フレームと過去 2 フレームの情報を使用することにより，実際の人物の位置との誤差を少なく観測できる．

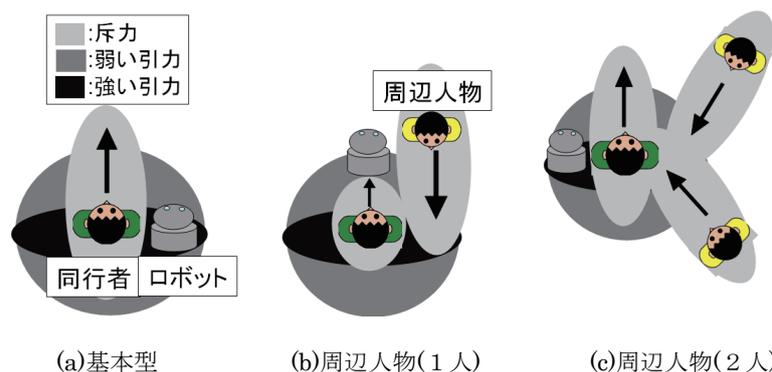
### (3) 同行者の向き



(a) 向きを考慮せずに移動 (b) 向きを考慮した移動

図 4 同行者の向きを考慮した移動判断例

現フレーム情報のみで同行者の進行方向を予測するために，身体の向きを検出を行う．これにより，同行者の進行方向にロボットが位置した場合に，瞬時に(約 0.5[msec])邪魔になるかどうかを認識し移動することができる．具体的には，現フレームの情報から同行者がロボット側を向いていることを検出することにより，同行者のロボット側に移動しようとする意思を予測し，図 4(b)のように相対的位置関係を変化させ衝突を回避する．検出方法としては，腰部分にフィッティングさせた楕円について短径方向を求め，ロボットの正面側と同じ側にある短径がロボットの正面方向となす角度をロボットと同行者との向きの差として認識する．なお，同行者の腰部分に対する楕円フィッティングの精度を向上させるため，追跡した同行者の中心付近の距離情報に基づいて，1 画素を 3[cm]×3[cm]に設定した，より精細な  $X-Z$  平面画像を作成し (2-3 (1) 人物検出における  $X-Z$  平面画像においては，1 画素は 5[cm]×5[cm] の大きさに相当する)，楕円フィッティングを改めて行うようにしている．



(a)基本型 (b)周辺人物(1人) (c)周辺人物(2人)

図 5 ロボットと同行者の相対的位置関係の例

## 2-4 ロボットの移動判断

ロボットが周囲環境を認識することは，安全に自律移動する上で重要である．周囲の障害物や歩行者の状態からロボットの移動方向を決定するために，ポテンシャル法[10]を用いる．ロボットの目的地となる引力ポテンシャルを同行者の周囲に設定し，「並走」形態となる位置には，より強い引力ポテンシャルを設定する．

また、周辺人物、障害物や壁には斥力ポテンシャルを設定する。図 5 にロボットと同行者の相対的位置関係の例を示す。引力を■と■の色、斥力を■の色で表す。ロボットは図 5(a)を基本型として移動する。そこに他の人物が近づいてくるとロボットは同行者と周辺人物に合わせ、図 5(b)のように位置を変化させる。周辺人物の位置や速度によって、同行者との位置関係を変化させるタイミングが変わり、サブゴールを設置し、そこを経由して並走にもどる。サブゴールは、ロボットと引力ポテンシャル（同行者）との間に周辺人物等による斥力場が生じた時に設置される。ロボットは、サブゴールを経由して衝突を回避し移動する。斥力場がロボットと引力ポテンシャルの間にあり、その時、同行者がロボットの前方に位置した場合には、同行者が先に移動したいという意図であると認識し、サブゴールを決める探索範囲を同行者の後ろに設置する。一方、同行者がロボットの後方に位置した場合は、同行者がロボットの後に移動したい（ロボットに先に移動してほしい）という意図であると認識し、サブゴールの探索範囲を同行者の前に設置する。探索範囲の中で一番引力が強い部分をサブゴールとする。認識処理は毎フレーム（約 60[msec]）行う。但し、サブゴールを安定して設置するために、過去 5 フレームにおける平均値から位置を定める。また、周辺人物が複数人近づいてきて、強めの引力（■部分）が消滅し、同行者に対して右側の位置では衝突してしまうと判断すると、サブゴールを設置しそこを経由することで、図 5(c)のように自律的に左側へと位置関係を変化させる。従来研究 [7] , [8]では、2-1 で触れたように予め設定された幾つかの動作モードの切替えによって移動形態が決まるため、移動形態のパターンに限られ、周辺環境の変化への柔軟な対応は難しい。これに対し、本研究のポテンシャル法では、同行者の周りに環状に引力ポテンシャルを設置することが特徴であり、様々な周辺環境に対応してロボットは同行者に対する相対的位置関係を臨機応変に変化させ自律的に移動することができる。

図 6 に、室内環境においての引力と斥力の関係をグラフにしたものを示す。実環境において、斥力と引力が同行者の周りに環状に生成されていることが確認できる。

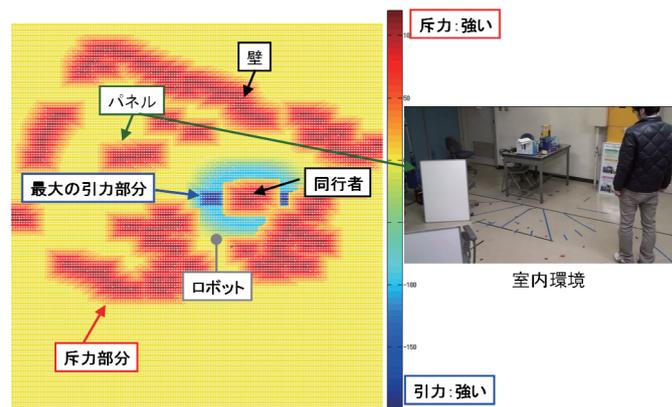


図 6 室内環境での斥力と引力の分布(同行者は停止)

## 2-5 廊下での実験

図 7 に廊下においての実験結果を示す。丸印がロボットの位置、輪印が同行者、同行者の左側の丸印が引力ポテンシャルの位置、矩形が対向人物、線分がカルマンフィルタの結果であり、速度が速いほど線分は長くなる。図 7(a1), (a2)において、ロボットは同行者の隣りに設置した引力ポテンシャルの力を受けて移動を開始する。ロボットと引力ポテンシャルの間にカルマンフィルタの結果によって生成された斥力場が存在すると、図 7(b1), (b2)に示すようにロボットはサブゴール(濃い丸印)を設置し、そこを経由する。このとき、同行者がロボットの前方に位置するため、ロボットは同行者がロボットより先に移動したいという意図を察し、図 7 (c1), (c2)に示すようにサブゴールを同行者の後方に設置した。図 7(d1), (d2)では、ロボットは同行者に合わせて停止する。図 7(a3)～(d3)には引力と斥力の関係を図で表現した。また、図 8 にロボットが同行者の前方を経由した場合の実験結果を示す。図 8(b1), (b2)に示すように、斥力がロボットと引力ポテンシャルの間に検出された時、同行者はロボットの後方に位置する。この場合には、ロボットは同行者はロボットに道を譲っていると認識し、サブゴールを同行者の前方に設置して、図 8(c1), (c2)のように、サブゴールを経由することにより衝突回避ができた。図 8(d1), (d2)では、同行者に合わせて停止する。1 フレーム当たりの処理速度は 50～70 [msec]であり、実時間でロボットの移動が可能であった。他の移動パターンとして、図 9 に、廊下から入口を通過し部屋内へ移動する場合についての実験結果を示す。

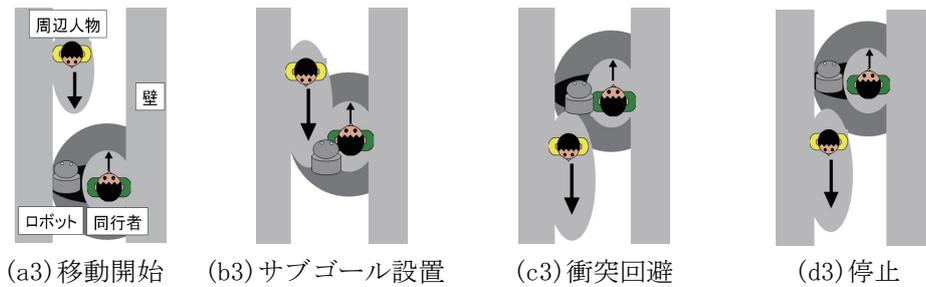
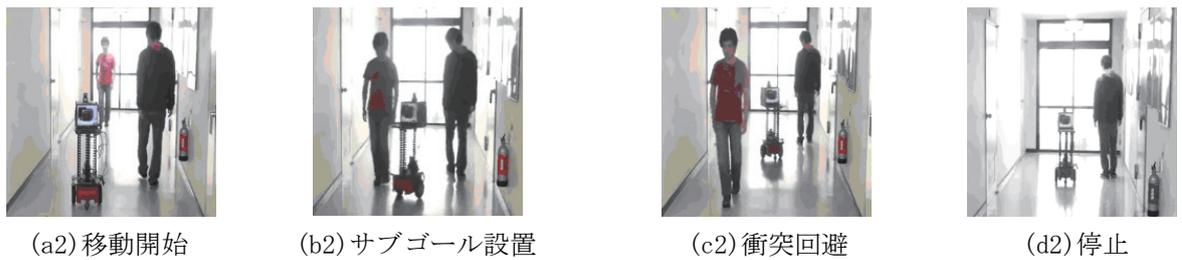
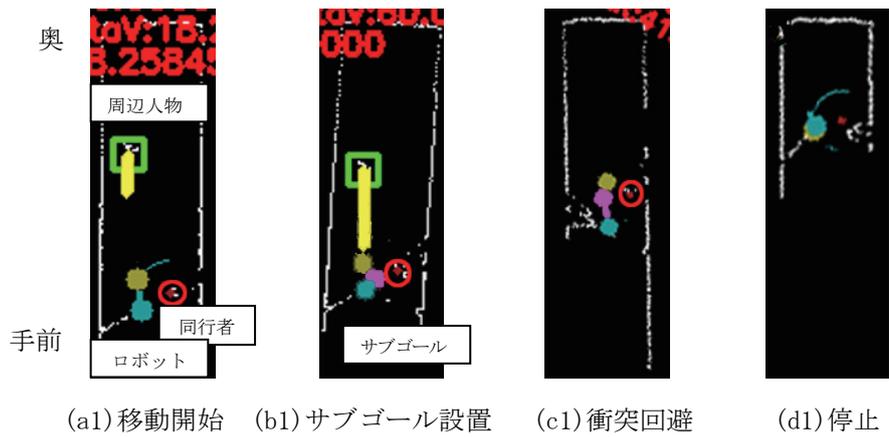
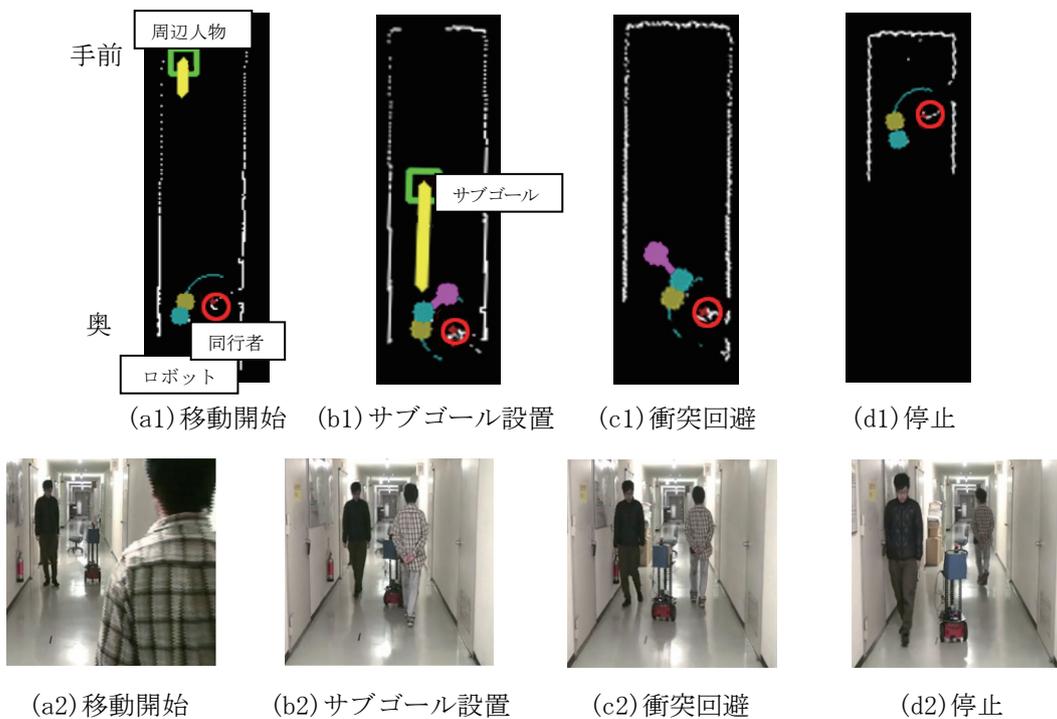


図7 実験結果 (同行者の後方を經由)



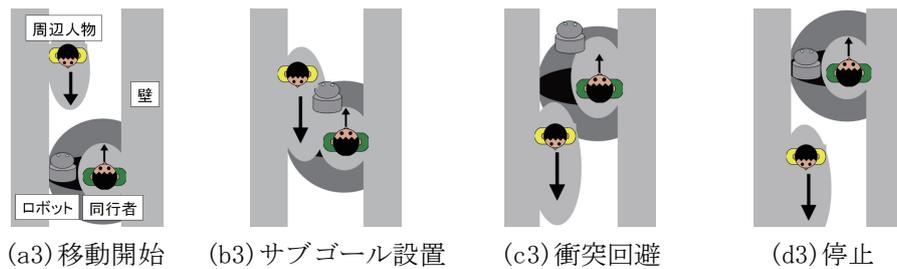


図 8 実験結果（同行者の前方を經由）

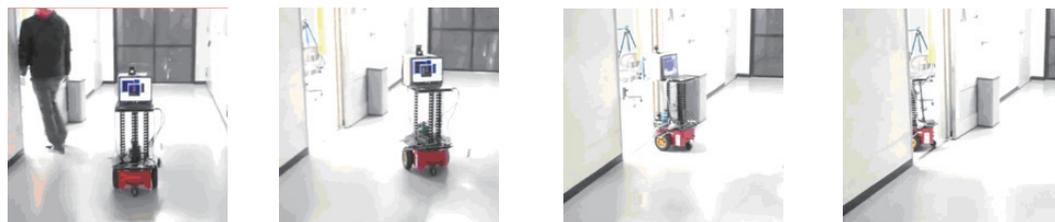
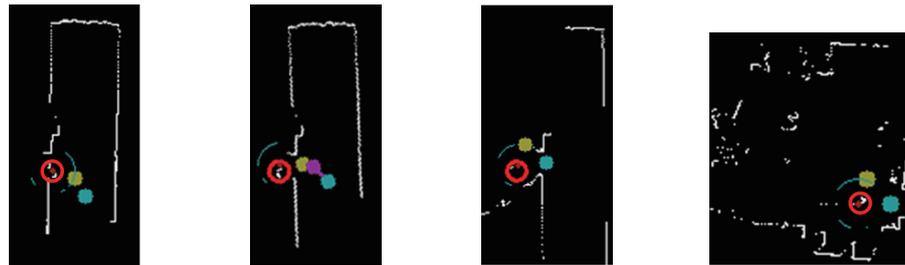


図 9 実験結果（廊下から室内へ移動）

## 2-6 むすび

LRFから得られる  $X-Z$  平面画像に基づいて人物検出，追跡，障害物の検出を行うことにより，同行者と並走し，また，同行者の向きや同行者との位置関係，周辺人物の位置や歩行速度を考慮した移動を，廊下や室内などの日常環境下においてロボットに実時間で行わせることができた．今後の課題として，同行者を見失ったときへの対応が挙げられる．

## 3 受取人の状態に応じて割込み行動を変化させる配達ロボット

### 3-1 はじめに

お手伝いロボットのように，人間に対して働きかけを行うロボットが世の中に登場し始めている．人間への働きかけの場面では，人間の作業を邪魔して不快感を与えないよう，割込むタイミングを計る必要がある．従来研究として，パソコンでの作業中における忙しさをパソコンの操作状態から調べる試み[11]や，不快感の低下を狙って動作と発話のタイミングをずらす方法[12]などが報告されている．しかし，これらの研究ではロボットと人間が1対1の関係にあり，実際の生活場面で見られるような人間同士の会話などの関わり合いが考慮されていない．

そこで本報告では，受取人に荷物を届ける配達ロボットを例として，受取人が他者と会話している可能性を考慮して，割込みのタイミングを計る手法を提案する．会話状態の判断には，受取人と会話相手の関係を考慮したマルコフモデルを利用する．ロボットは，視覚情報による個人識別及び聴覚情報による話者推定の結果を用いてこのモデルを更新し，会話に十分な「間」があると判断したタイミングで割込みを行う．

### 3-2 ロボットのセンシング・認識機能の実装

#### (1) ロボットの構成

ロボットにはMobileRobots社製のPioneer3DXを用い、受取人とその周囲の環境をセンシングするために、距離画像の取得が可能なTOF (Time of Flight) カメラ、カラー単眼カメラ、及び5本のマイクから構成されるマイクロホンアレイを搭載した。TOFカメラとカラーカメラは並べて同一方向に設置しているため、光軸がほぼ一致していると見なせる。

このロボットは図10のようなステートを持つ。まず、「視覚探索」ステートではTOFカメラによる距離情報とカラーカメラによる色情報を用いて、ロボットの視野内に人間がいるか、いるならばそれが誰かを識別する。次に、「聴覚探索」ステートではマイクロホンアレイを用いて音源位置推定を行い、視覚探索の結果と合わせて、誰が発話しているかを認識する。この話者推定の結果を用いて会話モデルの状態の確率を更新し、割込みのタイミングを決定する。

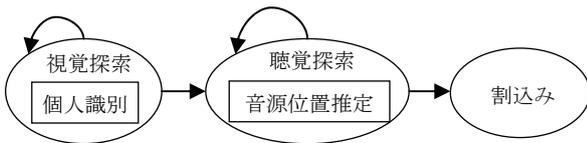
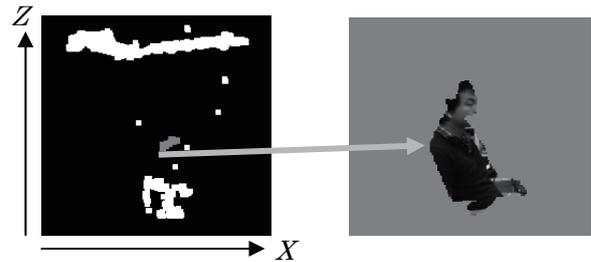


図10 ロボットのセンシング・認識の流れ



(a) X-Z平面 (b) 抽出された人物

図11 人物領域の検出

#### (2) 視覚探索

##### (a) 人物検出

TOFカメラによる距離画像から人物領域を検出する手法を述べる。まず、距離情報の各点を、横軸が水平(X)方向に、縦軸が奥行き(Z)方向に対応したX-Z平面画像上に投票していく。なお、床から1m以下の高さについては椅子や机などの人物以外の物体による距離情報が大半を占めるため、投票対象は1m以上の高さに限定する。次に、投票画像を2値化处理し、物体の存在候補領域を得る。この時に、遠方に行くほど距離情報のサンプルが疎になることを考慮し、奥行き $z$  [m]に応じて、式(1)のように閾値 $T_z$ を変化させる。

$$T_z = 50 / z^2 \quad (1)$$

ここで、この物体の存在候補領域にラベリングを行い、人間らしい大きさを持つ領域を人物候補領域とする。これに対応する領域をカラーカメラ画像から抽出し、HSV表色系を用いた肌色検出を行う。この肌色領域が一定の大きさ以上を持つ場合は、これを人物領域とする。図11(a)では中心付近の領域が人物領域である。図11(b)に示すように、この人物領域に対応するカラーカメラ画像の領域を人物抽出画像とする。

##### (b) 個人識別

(2)(a)で得た人物抽出画像について、色情報と勾配情報を用いて個人識別を行う。

色情報を表すため、人物抽出画像をHSV表色系で表し、H成分とS成分を各々15値、4値に量子化して計60次元のヒストグラムを得る。このヒストグラムについては、事前に用意した1人につき5枚のデータベース画像(以下DB画像)から得られたヒストグラムと比較し、最も距離が近い画像の人物IDを $R_{hs}$ とする。

勾配情報を表すため、人物抽出画像をグレースケール化し、128次元の特徴量を持つSURF特徴[13]を計算する。この特徴のそれぞれがどの人物に所属するかを、各人物用のSVM(Support Vector Machine[14])を用いて分類する。このSVMは事前にDB画像により学習したものである。分類されたSURF特徴の数が最も多い人間のIDを $R_{surf}$ とする。

SURF特徴は、識別問題で高い性能を示すことが知られているが、模様の少ない服等では信頼度が落ちる。このことを考慮して、最終的な識別結果は、SURF特徴の信頼度が高い時には $R_{surf}$ 、そうでなければ $R_{hs}$ とする。式(2)を満たせば、SURF特徴は信頼できるとする。ここで $v_1 \sim v_n$ は各人物へのSURF特徴の投票数を多い順に並べたものである。

$$v_1 \geq 0.5 \sum_{i=1}^n v_i \quad (2)$$

個人識別のための処理は計算量が膨大となる．そこで，過去の個人識別結果をできる限り利用し，全体の処理の高速化を図る．個人識別の結果は  $X-Z$  平面と同じ画素数の個人位置マップに格納しておくことにする．フレームごとに得られた人物検出の結果を個人位置マップに対応させ，過去の情報がある場合にはそれを活用し，ない場合にのみ個人識別を新たに行う．

### (3) 聴覚探索

#### (a) 尤度マップを用いた音源位置推定

ここでは，多数のマイクを用いて取得された音声信号の時間差を元に音源位置推定を行う．まず，周波数領域での位相差を手がかりにする CSP (Cross-power Spectrum Phase) 法によって，時間差の確からしさ (CSP 係数) を求めた．次に，尤度マップ法[15]を用いてこの時間差を位置に対応させ，音源位置を推定する．

尤度マップ法の詳細を述べる．ロボット座標の原点から見て，ある座標  $p$  の位置に音源があると仮定する．マイクロホンペア MicA, MicB の 3 次元座標を  $m_a$ ,  $m_b$  とする．音速を  $c$ ，マイクロホンペアにおける時間差を  $\Delta t$  とすると，以下の式が成り立つ．

$$\Delta t = \frac{(|p - m_a| - |p - m_b|)}{c} \quad (3)$$

この関係を満たす座標  $p$  は双曲線を描く．この双曲線上に， $\Delta t$  に対応する CSP 係数を割り当てる．こうすることで，そのマイクロホンペアに対して音源位置の確からしさを表すマップを得ることができる．この音源位置の確からしさを，水平垂直方向の角度の空間にマッピングしたものを尤度マップとする．ここでは，それぞれの角度に対応する直線上で，CSP 係数の最大値を求め，これをその角度に対する尤度とする．この尤度マップを複数のマイクロホンペアで平均することにより音源位置を精度よく絞り込むことができる．図 12 は 5 個のマイクロホン (10 ペア) について平均を取った尤度マップであり，垂直方向  $\pm 45^\circ$ ，水平方向  $\pm 180^\circ$  (全周) の尤度を表している．また，色が白いほど尤度が高いことを示す．この場合，ロボット正面方向に音源があることが分かる．

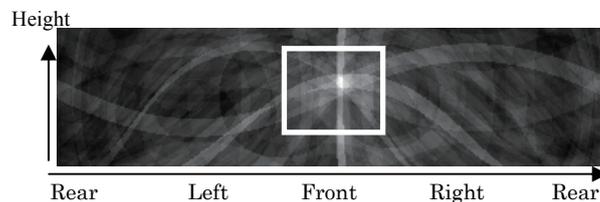


図 12 尤度マップと TOF カメラの視野



図 13 話者推定法

#### (b) 話者推定

図 12 のロボットの正面に当たる部分 (図 12 中の白枠) を切り抜き，閾値処理を行うことで TOF カメラの視野と対応した音源位置画像を得る．この音源位置画像に対して 2 値化処理を行った画像と，3-2 (2) (a) の人物抽出画像の背景を黒にした画像の AND 演算を行ったものが図 13 となる．この AND 演算の結果画像に黒以外の画素が残っていた場合，その人物抽出画像の人物が発話したと推定する．

### 3-3 割り込み判断のための会話状態推定

ロボットが荷物配達のための割り込みのタイミングを決定するため、受取人が他の人間と会話している状態を、マルコフモデルを用いて図 14 のようにモデル化する。ただし、ここでは受取人に加え人物 A, B の計 3 人がいる場合を考える。「Silent」(以下「S」と表記)は会話が行われていない状態、「Listening」(以下「L」と表記)は、受取人本人は発話していないものの、A か B の発話を聞いている状態、「Talking」(以下「T」と表記)は受取人が A または B に向けて発話している状態を示している。

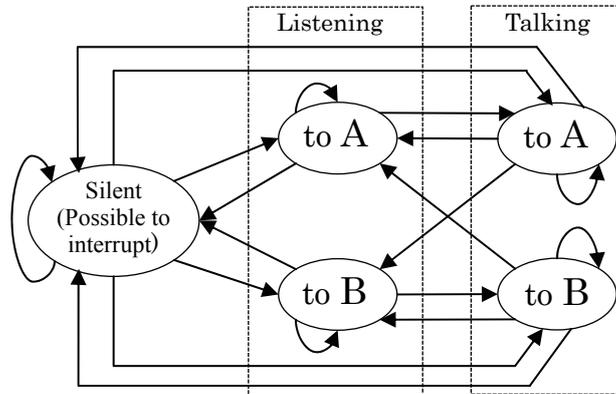


図 14 会話状態の状態遷移図

表 1 会話状態モデルの遷移確率例(受取人 R 発話時)

		To				
		S	L <sub>A</sub>	L <sub>B</sub>	T <sub>A</sub>	T <sub>B</sub>
From	S	0.2	0	0	*1	*2
	L <sub>A</sub>	*3	*4	N/A	0.3	N/A
	L <sub>B</sub>	*5	N/A	*6	N/A	0.1
	T <sub>A</sub>	0.1	0.1	0	0.8	N/A
	T <sub>B</sub>	0.1	0	0.1	N/A	0.8

$$\begin{aligned}
 &0.4 * L_A / (L_A + L_B) \dots *1 & 0.4 * L_B / (L_A + L_B) \dots *2 \\
 &0.2 - 0.2C_A \dots *3 & 0.5 + 0.2C_A \dots *4 \\
 &0.2 - 0.2C_B \dots *5 & 0.5 + 0.2C_B \dots *6
 \end{aligned}$$

表 2 会話状態モデルの遷移確率例(人物 A 発話時)

		To				
		S	L <sub>A</sub>	L <sub>B</sub>	T <sub>A</sub>	T <sub>B</sub>
From	S	0.8	0.05	0	0.15	0
	L <sub>A</sub>	*1	*2	N/A	0.1	N/A
	L <sub>B</sub>	*3	N/A	*4	N/A	0.1
	T <sub>A</sub>	0.05	0.4	0.05	0.5	N/A
	T <sub>B</sub>	0.15	0.3	0.05	N/A	0.5

$$\begin{aligned}
 &0.4 - 0.4C_A \dots *1 & 0.5 + 0.4C_A \dots *2 \\
 &0.4 - 0.4C_B \dots *3 & 0.5 + 0.4C_B \dots *4
 \end{aligned}$$

また、A, B と受取人との社会的人間関係を事前に定め  $C_A$ ,  $C_B$  と表す。 $C_A$ ,  $C_B$  は社会的立場の差や親しさを反映し、会話の重要性が高いほど大きな値になるものとする。これにより、会話相手によって割り込みまでのタイミングを変化させることが出来る。

図 14 のマルコフモデルに対して、「無音」、「受取人が発話」、「A が発話」、及び「B が発話」のイベントに対応する計 4 個の遷移確率テーブルを作成する。表 1 は「受取人発話」時のテーブル、表 2 は A が発話した時のテーブルである。受取人が実際には会話に参加していないのに、人物 A または B が発話し続けることにより、 $L_A$  及び  $L_B$  の確率が高くなることを防ぐために、表 2 の「 $T_A \rightarrow L_A$ 」を表 2 の「 $S \rightarrow L_A$ 」に比べて大きくし、また表 1 の「 $S \rightarrow T_A$ 」が大きな値になるようにする。これにより、受取人 R が A の話を聞いているという確率

「 $L_A$ 」が上昇するためには、RがAに対して発話する「 $T_A$ 」を通過することが条件となっている。この会話状態モデルを用いて、本研究では「S」の状態確率が0.7を超えた場合に、ロボットは割込みを行うことにした。

### 3-4 実験

人物認識、個人識別、及び話者推定をロボット実機に実装し、実験を行った。実験では、受取人Rと会話相手A、会話に参加しないBの計3人がいる会話環境を記録し、その結果に対して3-3の会話状態モデルを当てはめて、各状態の確率をオフラインで計算した。実験1回目は、RとAが会話を行い、 $C_A$ の変化による確率の変化について検証した。実験2回目はAのみが発話し、Rが元々会話に参加していなかった場合を検証した。

図15及び図16に、実験1回目の $(C_A, C_B) = (0.3, 0.5)$ 及び $(0.6, 0.5)$ の時の会話状態モデルの各状態の確率を示す。なお、図中の上部の点は、話者推定を行った結果である。図15より、時刻25秒付近の発話非検知区間で「S」の確率が上昇し、0.7を超えていることがわかる。また図16より、 $C_A$ が大きくなると、Aの話を受けている確率「 $L_A$ 」が高くなり、「S」の確率の上昇が鈍くなる。この場合、割込みが可能になったのは時刻65秒となり、会話相手の重要度が反映された結果になっている。

図17に、実験2回目の結果を示す。図17の0~3秒付近までのグラフより、Rが発話をせずにAのみ発話をしている状態においては、「 $L_A$ 」は一定以上上昇せず、割込みに対して大きな影響を及ぼさないことがわかる。これは3-3で示したように、「 $L_A$ 」が上昇するためには、Rが発話することが必要条件になっているためである。

### 3-5 むすび

本報告では、受取人に荷物を届ける配達ロボットを例として、受取人が他者と会話している可能性を考慮して、割込みのタイミングを計る手法を提案した。会話状態の判断には、受取人と会話相手の関係を考慮したマルコフモデルを利用する。ロボットは、視覚情報による個人識別及び聴覚情報による話者推定の結果を用いてこのモデルを更新し、会話に十分な「間」があると判断したタイミングで割込みを行う。

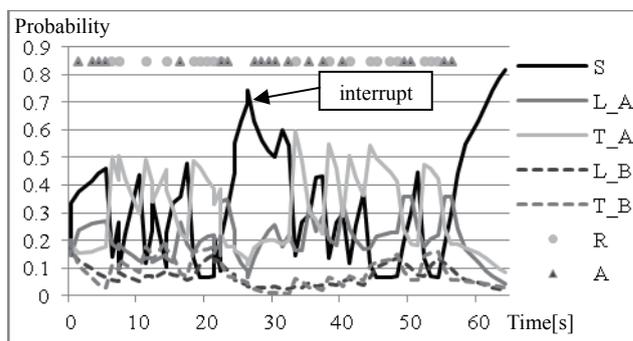


図15  $C_A = 0.3, C_B = 0.5$ の時の各状態の確率

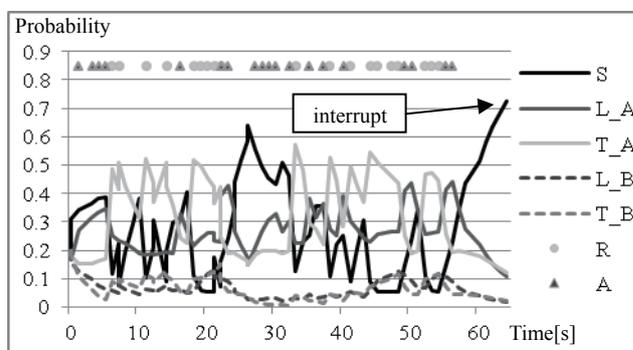


図16  $C_A = 0.6, C_B = 0.5$ の時の各状態の確率

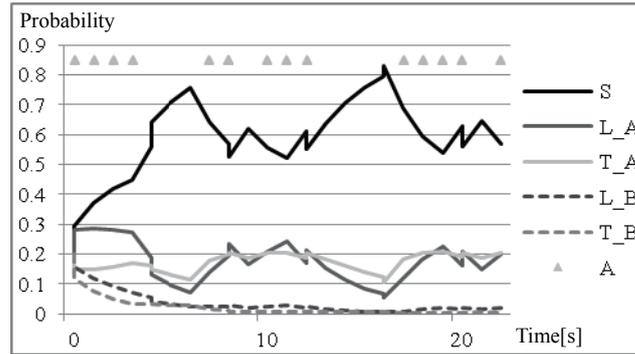


図 17 受取人 R が発話しない時の各状態の確率

今後の課題として、事前に与えられた地図情報を用いることで、観測対象がすべて TOF カメラの視野内に取まらないケースにも対応することが挙げられる。また、今回は割込みの閾値を 0.7 としたが、この値を荷物の緊急度によって可変にすることで、より臨機応変な行動を可能にすることが挙げられる。

#### 4 むすび

本報告では、複雑な実環境下での察するインタフェースに関する研究として、次の 2 つの研究成果を報告した。

##### (1) 同行者との相対的位置関係を考慮した日常環境下でのロボットの自律的移動

複数の人物がいる、より実際の生活空間に近い環境の中で、同行者と並走する、同行者の後方を追従する、同行者の前にも位置するなど相対的位置関係をさまざまに変化させながら、安全に自律的移動をすることができるロボットシステムを提案した。LRF から得られる X-Z 平面画像に基づいて人物検出、追跡、障害物検出を行う。また、周囲の障害物や歩行者の状態からロボットの移動方向を自律的に決定するために、ポテンシャル法を導入した。障害物、壁、周辺人物には斥力ポテンシャルを、同行者の周囲には環状に引力ポテンシャルを設定する。廊下や室内などの日常環境下において、同行者と並走し、また、同行者の向きや同行者との位置関係、周辺人物の位置や歩行速度を考慮した移動を、ロボットに実時間で行わせることができることを確認した。

##### (2) 受取人の状態に応じて割込み行動を変化させる配達ロボット

受取人が他者と会話をしている可能性を考慮して、割込みのタイミングを計ることが可能な、受取人への荷物配達ロボットについて提案した。会話状態の判断には、受取人と会話相手の関係を考慮したマルコフモデルを利用する。ロボットは視覚情報による個人識別及び聴覚情報による話者推定の結果を用いて、このモデルを更新し、会話に十分な「間」があると判断したタイミングで割込みを行う。人物認識、個人識別、話者推定の各手法をロボット実機に実装して実験を行い、受取人の状態に応じた割込み行動を確認した。

今後ロボットが様々な形で我々の日常生活環境の中に入り、人間のパートナーとしての役割を果たしていくことが考えられる。この様な中で、通常的生活環境の中で起こり得る様々な状況をロボット自らが把握し、人間と自身（ロボット）との関係を察して自律的に行動できることは極めて重要である。

今後の課題として、察するインタフェースとしての様々な機能を統合化して実装し、実環境内でストレスなくかつ安全に動作する様なロボットの実現を図っていきたい。

#### 【参考文献】

- [1] Roland Siegwart, Illah R. Nourbakhsh and Davide Scaramuzza : “Introduction to Autonomous Mobile Robots (2nd Ed.),” The MIT Press, 2011.
- [2] 松原仁, 野田五十樹, 松野文俊, 稲見昌彦, 大須賀公一編 : “ロボット情報学ハンドブック,” 近代科学社, 2010.
- [3] 日立のロボット EMIEW : [http://www.hitachi.co.jp/rd/research/robotics/emiew1\\_01.html](http://www.hitachi.co.jp/rd/research/robotics/emiew1_01.html)
- [4] 特集「人間と共存し発展するロボットのためのビジョン」: 日本ロボット学会誌, vol. 27, no.6, 2009.7.

- [5] 渡邊睦：“人物自動追従・状態自動認識に関するビジョン技術,” 日本ロボット学会誌, vol. 27, no.6, pp.626-629, 2009.7.
- [6] 生活支援ロボットの開発について : [http://www.toshiba.co.jp/about/press/2005\\_05/pr\\_j2001.htm](http://www.toshiba.co.jp/about/press/2005_05/pr_j2001.htm)
- [7] 花井久美江, 今井順一, 金子正秀：“画像情報に基づき自律的にユーザとの並走・縦走の切替えが可能なロボットの実現,” 電子情報通信学会技術研究報告, 画像工学研究会, IE2006-285, 2007.
- [8] 小林貴訓, 金原悠貴, 久野義徳：“周辺状況を考慮して介護者と協調移動するロボット車椅子,” 第16回画像センシングシンポジウム, IS-1-12, 2010.
- [9] Michael Isard and Andrew Blake：“ICondensation: Unifying low-level and high-level tracking in a stochastic framework,” Proceedings of the 5th European Conference on Computer Vision, vol. 1, pp.893-908, 1998.
- [10] P. Vadakkepat, K. C. Tan, and M. Wang：“Evolutionary artificial potential fields and their application in real time robot path planning,” Congress of Evolutionary Computation, pp. 256-263, 2000.
- [11] 田中貴紘, 藤田欣也：“ユーザの割り込み拒否度を考慮した円滑な会話開始支援エージェント,” 電子情報通信学会論文誌 A, vol. J92-A, no.11, pp. 852-863, 2009.
- [12] Michiya Yamamoto and Tomio Watanabe：“Timing control effects of utterance to communicative actions on embodied interaction with a robot,” Proceedings of the IEEE 2004 International Workshop on Robot and Human Interactive Communication, pp.467-472, 2004.9.
- [13] H. Bay, T. Tuytelaars and L. V. Gool：“SURF: Speeded up robust features,” European Conference on Computer Vision, pp. 404-417, 2006.
- [14] Vladimir Vapnik：“The Nature of Statistical Learning Theory,” Springer-Verlag New York, 1995.
- [15] 高橋哲史, 今井順一, 金子正秀：“全周方向での音源位置推定及び肌色情報との統合による話者位置推定への適用,” 電子情報通信学会論文誌 D, vol.J90-D, no.8, pp.2232-2241, 2007.8.

#### 〈 発 表 資 料 〉

題 名	掲載誌・学会名等	発表年月
同行者との相対的位置関係を考慮した日常環境下でのロボットの自律的移動	映像情報メディア学会誌, vol.65, no.10, pp.1452-1457	2011.10.
受取人の状態に応じて割り込み行動を変化させる配達ロボット	映像情報メディア学会技術報告, メディア工学研究会, ME2012-36	2012.2.
Testing GMM algorithm for skin color segmentation in outdoor environments using a stereo camera	Proceedings of the 26th UEC International Mini-Conference for International Students (JUSST Program), pp.92-97	2011.8.
Study on particle filtering and fusion of stereo-camera and mono-camera for visual object tracking	Proceedings of the 26th UEC International Mini-Conference for International Students (JUSST Program), pp.104-105	2011.8.
動領域・静止領域の確率的マッピングを用いた動的環境の認識—自律移動ロボットにおける自己位置推定への応用—	情報処理学会第74回全国大会, 6C-2	2012.3.
受取人の状態に応じて割り込みのタイミングを変化させる配達ロボット	情報処理学会第74回全国大会, 5T-3	2012.3.
Skin color extraction in outdoor environments based on GMM using Bumblebee stereo camera	Proceedings of the 27th UEC International Mini-Conference for International Students (JUSST Program), pp.15-22	2012.3.
Research on visual object tracking based on combination of TOF camera and single optical camera	Proceedings of the 27th UEC International Mini-Conference for International Students (JUSST Program), pp.39-44	2012.3.