

# K-17 視聴覚情報を用いた移動型ロボットの指示者探索システム

## Operator Searching System of Mobile Robot using Audiovisual Information

服部雄介† 竹内義則‡ 大西昇‡  
Yusuke Hattori Yoshinori Takeuchi Noboru Ohnishi

### 1. はじめに

近年、ロボットが日常生活に多く登場し、身近なものになってきている。それに伴い、自律的に行動するロボットの研究も盛んに行われている。

こうしたロボットは、介護ロボット、ペットロボットのような主人が特定しているものと、案内ロボットのような不特定の人物にサービスするものに分けられる。不特定の人物にサービスするロボットでは、ロボットに複数の人物の中からロボットの指示者となる人物を探索することが必要不可欠である。

移動型ロボットに関して様々な研究が行われている。寺倉らは聴覚でロボットを制御する研究を行った[1]。また音声とジェスチャを用いたヒューマンロボットインターフェースに関する研究も多く行われている[2]。複数の感覚情報処理能力をもつロボットの研究として、電総研での Jijo-2 と名づけられたロボットに関するものがある[3]。オフィス内での人間のサポートを目的としており、前方位視覚センサや多方向超音波センサで外部情報を得ている。

### 2. 指示者探索システムの構成

本研究では、ロボットに命令を与える人物(指示者)をロボットが自律的に視覚情報と聴覚情報を用いて探索、特定することを目的とする。これは人間とロボットとの間での基本的かつ重要なインタラクション「人が呼んだらロボットは振り向き、その人を特定して、その人のもとにやってくる」という動作である。

システムの処理の流れを Fig.1 に示す。音声処理部分では、人からの呼びかけ音声の有無を調べるためにマイクから取り込んだ音の ONSET を常にチェックする。ONSET が見つかったら呼びかけがあったとし、それぞれのマイクから得られる ONSET 時刻から時間差を計算して音源の方位角を求める。そして音源の方向へロボットを向けるため、求めた方位角分ロボットを回転させる。こうして指示者をロボット視野内にとらえる。

次の画像処理部分では、ロボット視野内にいる人物のうち、ロボットに手招きしている人物を指示者と決定する処理を行う。カメラから画像を取り込みつつフレーム間差分処理と肌色抽出を用いて手招きしている部分を特定する。そして求めた手招き動作部分の位置をロボット視野の中心にくるようにロボットの向きを微調整し、かつロボットのソナーセンサ情報からロボットから手招き動作部分までの距離を求め、その位置まで移動する。

本研究で使った移動型ロボットは Active Media 社の Pioneer2 MOBILE ROBOTS である。

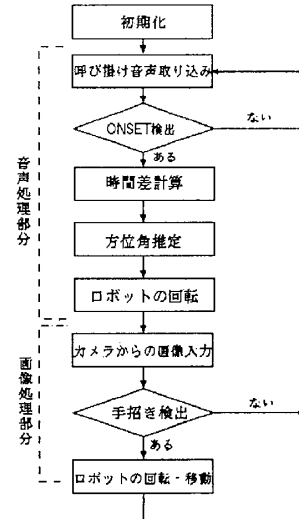


Fig.1 システムの流れ

### 3. 呼びかけ音声処理

本研究で使ったロボットの頭部を Fig.2 に示す。2組の時間差を求めるために、サウンドカードを2枚同時に使用した。マイク1とマイク2を1つのカードに、マイク3とマイク4をもうひとつのカードに接続した。まず、ロボットは音を1秒取り込んでONSETの有無を調べ、無かったらまた1秒取り込むことを繰り返す。ONSETが取り込んだ音圧を閾値処理することで求めた。ONSETが検出されたら、その時点の音データから2組の時間差を求める。マイク1とマイク2の時間差を $\Delta T_{1,2}$ 、マイク3とマイク4の時間差を $\Delta T_{3,4}$ とする。またマイク配置から幾何学的に求められるそれぞれの時間差 $\Delta t_{1,2}$ 、 $\Delta t_{3,4}$ は、

$$\Delta t_{1,2} = 0.168 (\cos(\theta - \pi/4) - \cos(\theta - 5\pi/4)) / C$$

$$\Delta t_{3,4} = 0.168 (\cos(\theta - 3\pi/4) - \cos(\theta - 7\pi/4)) / C$$

となる(Cは音速)。そして最小二乗法で

$$E(\theta) = (\Delta t_{1,2}(\theta) - \Delta T_{1,2})^2 + (\Delta t_{3,4}(\theta) - \Delta T_{3,4})^2$$

を最小にする $\theta$ を求め音源方向とする。

### 4. 手招き動作検出処理

本研究では人間のする手招き動作を、“肌色である程度

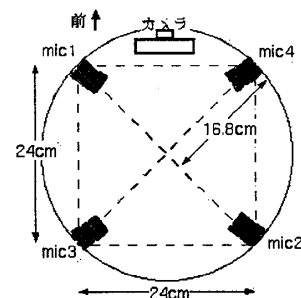


Fig.2 ロボット頭部のマイク配置

†名古屋大学大学院計算理工学専攻

‡名古屋大学大学院工学研究科情報工学専攻

同じ位置で動きのある部分”と定義する。

人間の手招き動作を検出するために、フレーム間差分処理と肌色抽出を行う。取り込む画像のフレーム間隔は約 0.054 秒で、15 フレーム分(約 0.81 秒)取り込む。まずフレーム間差分処理を行い、抽出した画素に対しさらに肌色抽出をおこなう。肌色抽出は YUV 表色系で行い、U と V が  $-20 \leq U \leq 5, -10 \leq V \leq 20$

の範囲にあるかどうかを調べる。範囲内にあったらその画素は手招き動作候補であるとしその xy 座標を記憶する。

この処理を 15 フレーム分繰り返し、手招き動作候補として最も検出頻度が多かった部分を最終的に手招き動作部分と決定する。しかし、手招き以外の動作の検出を防ぐため、検出回数が 4 回以上でない手招き動作とは判定しないことにした。この数値は経験的に決めた。

### 5. 実験・考察

実装した指示者探索システムの性能テストを行った。指示者はロボットに向かって「おーい」という呼びかけを行った後、手招きをした。

まず、呼びかけ音声の方位角推定の結果を示す(Fig.3)。各方向から 10 回ずつ試行した。平均誤差は 22.44°、最大誤差が 140° でロボットの視野外にでてしまったのが 80 回中 13 回であった。失敗した原因は呼びかけ音声の立ち上がり音が不安定で ONSET の誤検出が起きてしまった。ゼロクロス法や相関法でより正確に時間差を求める必要がある。

次に手招き動作検出結果を示す(Fig.4)。ロボット視野内に指示者しかいない場合、複数人物がいる場合どちらでも手招きは正確に検出できているといえる。しかし今回の手招き動作の定義に近いもの(拍手やジャンケンなど)は誤検出してしまった。これに対してはオプティカルフローを用いて手招きの周期性を利用するなどして、手招き動作をより厳密に定義する必要がある。

呼びかけ音声処理から手招き動作検出まで一連のプロセスでの実験の結果は次の 4 パターンに分類できた。

- ① 1 回で音源定位に成功し、指示者をロボット視野内に収め、次の手招き動作検出で手招き動作を検出し指示者を特定した。
- ② 1 回目の音源定位に失敗し、指示者をロボット視野内に収めることができなかった。しかし手招き動作は検出せず音源推定部分に戻った。そして、2 回目の音源定位、手招き動作検出に成功して指示者を特定した。
- ③ 1 回目の音源定位に失敗し、指示者をロボット視野内に収めることができなかったが、ロボット視野内の他の人物を誤検出してしまった。

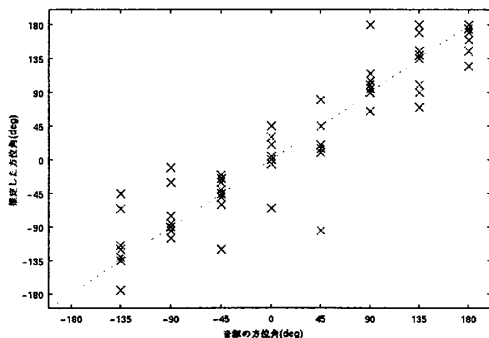


Fig.3 呼びかけ音声の方向推定結果



処理前

手招き抽出後



処理前

手招き抽出後

Fig.4 手招き抽出結果

- ④ 1 回で音源定位に成功し、指示者をロボット視野内に収めたが、手招き動作検出に失敗し、音源定位部分に戻ってしまった。

全 10 回の試行の内、①は 5 回、②は 3 回、③は 1 回、④は 1 回であった。上記のように呼びかけ音声処理、手招き動作検出処理の改善が必要であるが、加えて、ロボットが誤ってしまった場合の指示者からの「違うよ」といったリアクションや、ロボットから指示者に尋ねるといった対話処理が必要である。

### 6. まとめ・今後の課題

本研究では視聴覚情報を用いた指示者探索ロボットについて述べた。ロボットは呼びかけ音声を音源定位し、その方向に回転して指示者となる人物をロボット視野内に収める。その後、手招き動作検出で指示者を特定する。実験の結果、指示者の周囲に複数の人物が存在しても指示者だけを精度よく的確に特定できることを示した。より人間同士のコミュニケーションに近い形でかつ効率よく指示者の探索が可能であることがわかった。

本システムをより良いヒューマンインタフェースにするには、視覚処理部分と聴覚処理部分をプロセス間通信で処理し、それぞれのタスクを常に独立して実行する方法が考えられる。加えて、音声対話により、人間とロボットが自然言語で情報をやり取りすることが不可欠である。

### 参考文献

- [1] 寺倉生剛, 大西昇, 杉江昇, “音源定位機能を持った移動型ロボット”, 計測自動制御学会論文誌, Vol. 30, No. 5, pp. 600-602, 1994
- [2] 高橋拓弥, 中西知, 久野義徳, 白井良明, “音声とジェスチャによる対話に基づくヒューマンインタフェース”, 情報処理学会ヒューマンインタフェース研究会, インタラクション'98 論文集, pp. 161-168, 1998
- [3] 村井俊浩, 麻生英樹, John Fly, 浅野太, 本村陽一, 原功, 栗田多喜夫, 遠見悟, 山崎信行, “オフィス移動型ロボット Jijo-2 の音声対話システム”, 日本ロボット学会誌, Vol. 18, No. 2, pp. 300-307, 2000