

# 視野外の状況も考慮したロボットの自律的行動選択

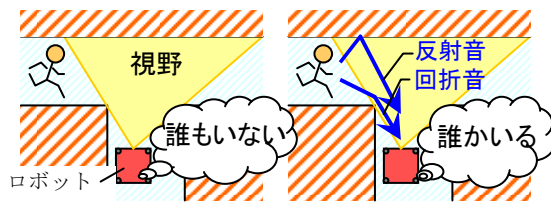
## ～ 音情報による物陰の人物の認識 ～

渡部 翔      金子 正秀

電気通信大学 大学院電気通信学研究科

### 1. はじめに

ロボットがオフィスなどで自律的に行動する際、安全かつ効率よく動くためには素早く正確な環境認識が必要となる。特に人物については位置や形状が時々刻々変化するので、リアルタイムで認識する必要がある。従来手法では、レーザーレンジファインダ(LRF)を用いた手法[1]や、LRFとカメラを組み合わせる手法[2]などがある。しかし、これらの手法で認識できるのは可視範囲のみであり、曲がり角の先など、物陰にいる人物を考慮した例はない。そこで本論文では、図1に示したように、自律移動ロボットが物陰の人物にも対応できるよう、音を利用した環境認識手法を提案する。音は物陰から発せられても反射や回折によりロボットまで届く。このため、足音などの物音を利用することで物陰の情報のある程度取得することが可能である。最終的には、ロボットが視野外の人物の位置や行動を聴覚を用いて認識し、視野外の人物との衝突を回避しながら自律的に移動できるようにすることを目的とする。



(a) 視覚のみの場合    (b) 音を利用する場合  
図1. 音による物陰の人物の認識

### 2. 音源位置推定

#### 2.1. CSP法[3]

音源の方向を推定する手段として、複数のマイクロホンへの音の到来時間差を利用する。1組のマイクロホンペアにおける到来時間差はCSP法により計算する。2つのマイクロホンの受信音を用いてCSP係数は式(1)により計算される。

$$CSP(k) = FFT^{-1} \left( \frac{FFT(s_i(t)) \cdot FFT^*(s_j(t))}{|FFT(s_i(t))| \cdot |FFT(s_j(t))|} \right) \quad (1)$$

ただし  $s_i, s_j$  は各マイクロホンにおける受信音信号、 $t$  は時間、 $k$  は遅延サンプル数である。

あるマイクロホンペアのCSP係数は、そのマイクロホンペアにおける受信音の遅延時間ごとの尤度を表している。CSP係数の大きな部分の遅延サンプル数から、2つの信号の時間差を求めることができる。

#### 2.2. 音源尤度マップによる音源位置推定

通常CSP法による音源位置推定では、音源とマイクロホンアレーの距離が十分に遠いと仮定し、CSP法によって得られたCSP係数から近似的に音源の方向を角度として算出する。ところがこの方法では、音源が近い場合に近似が破綻して正確な推定ができなくなる欠点がある。

そこで本論文では、予め各マイクロホンペアについてさまざまな方向・距離に音源を仮定した場合の遅延時間差をシミュレートしたデータを用意した。それを受信音から得られるCSP係数と比較することで、音源の存在する可能性(尤度)を2次元マップとして取得する。このマップを音源尤度マップと呼ぶ。この方法では近似を使用しないので、距離によらず正確な音源位置推定が可能となる。また、マップ状のデータなので視覚画像との統合も容易である。

#### 3. 人物の移動方向推定

マイクロホンに直接音が届く範囲に音源がある場合には、前述のように遅延時間差を利用して音源位置を推定することができる。ところが、音源が物陰にある場合には、音が直接届かないため、この方法で正確な位置を推定することは難しい。

そこで本論文ではマイクロホン間の音圧比を用いて音源までのおよその距離を測定し、その時間的变化から人物の移動方向を推定することとした。具体的には式(2)で表されるマイクロホン間音圧差(Interaural Level Difference: ILD)を用いる。音源として人物が移動時に発する足音を考える。マイクロホンペアを図2のように垂直に配置すると、音源までの距離に応じて音源から各マイクロホンまでの距離の差が変化するので、ILDの値が変化する。距離とILDの関係を予め測定しておくことで、ILD値から距離を推定することができる。ILDと距離の関係を測定した結果が図3である。

$$ILD = 10 \log_{10} \left( \frac{\sum_{n=0}^{N-1} s_i^2(n)}{\sum_{n=0}^{N-1} s_j^2(n)} \right) \quad (2)$$

Mobile Robot Autonomously Selecting Actions  
by Estimating Conditions Outside Its View  
～ Recognition of People in Occluded Area Using  
Sound Information ～

Sho WATANABE and Masahide KANEKO

The University of Electro-Communications  
1-5-1 Chofugaoka, Chofu-shi, Tokyo 182-8585, Japan  
{s\_watanabe, kaneko}@radish.ee.uec.ac.jp

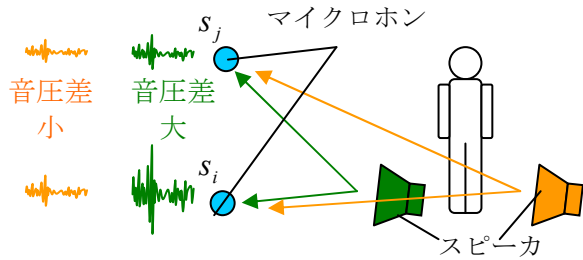


図2. マイクロホンペアにおける音圧差の発生

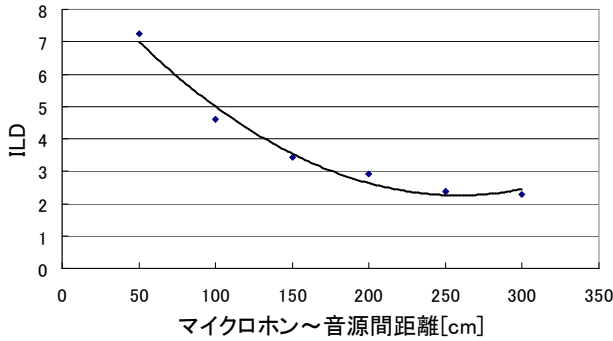


図3. 距離とILDの関係

#### 4. ロボットの動作判断

音源位置は音源尤度マップ，人物の移動方向はILDによりそれぞれ個々に推定し，両者から必要があると判断した場合にのみ，ロボットが衝突回避行動をとるようにする．前方の音源に関しては視覚情報も利用して人物検出を行い，音源尤度の高い方向に人物が検出されない場合，遮蔽があると判断する．

動作のパターンとしては，

- ① 後方から人物が接近してきた場合，道を譲る．
- ② 前方の物陰から人物が接近してきた場合，停止するか，もしくは，よける．
- ③ 横から人物が接近してきた場合，加速または減速して衝突を回避する．

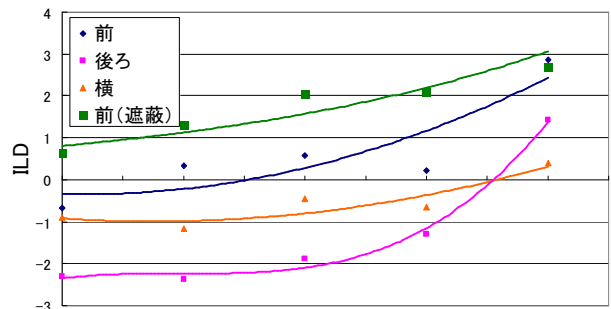
とする．どれにも該当しない場合には必要のない回避行動はとらず，効率のよい自律移動を行う．

#### 5. 評価実験

以上の方法をロボットに実装し，動作実験を行った．ここでは，動作の際にロボットが取得した音源尤度マップおよびILDのデータから，正しい環境認識ができていくかどうかを客観的に評価する．

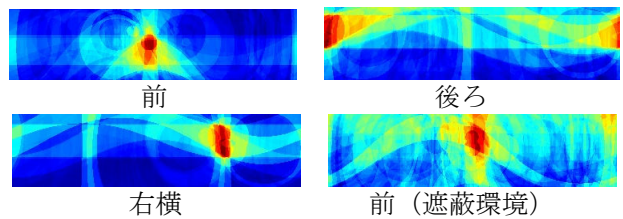
実験は，様々な方向から人物がロボットに向かって接近し，そのときのILD値と音源尤度マップを取得するという方法で行った．一回の接近に対してILDおよび音源尤度マップを5回(5歩分)取得し，それを3セット行い平均をとる．これを前，後ろ，右横，前(遮蔽環境)のそれぞれから人物が近付く4つのパターンについて測定し，全部で12セット測定した．

この実験で取得したILDと音源尤度マップは図4のような値となった．音源尤度マップは色が濃い赤に近いほど，その方向に音源がある可能性が高いことを表している．また，音源尤度マップはロボットの周囲360度の尤度を表しており，中央はロボットの正面，両端は真後ろを表す．



(時間の経過に対応)

(a) ILDの変化



(b) 音源尤度マップ

図4. 動作実験における取得データ

各音源尤度マップを見ると，それぞれ人物がいる方向に高い尤度が観測できていることが分かる．遮蔽環境の場合は遮蔽物の端から来る回折音などを観測することになり，尤度はその位置に出るため，図4では正面から少しずれている．

一方，図4よりILDはどの場合においても時間的に上昇しているのので，音は接近していると判断できる．この結果から，視野内に人物がいなくても，物陰から人物が接近していることを認識することができる．動作実験では，これらの情報からロボットが衝突の危険を察知して回避する様子を観察できた．

#### 6. おわりに

本論文では，マイクロホンアレーで受信した音のILDの変化から視覚では捉えられない物陰の状況を予測し，視野外にいる人物との衝突回避を行う手法を提案した．また，音源尤度マップや視覚情報と組み合わせることで，状況に応じた適切な動作を選択させ，動作の無駄を省くことができた．今後の課題としては，足音以外の音源にも対応できるようにし，歩いている人物に限らず，より一般的に人物や物体を対象とした衝突回避を実現することが挙げられる．

#### 参考文献

- [1] 奥迫伸一，坂根茂幸：“レーザーレンジファインダを用いた移動ロボットによる人の追跡，”日本ロボット学会誌，Vol. 24, No. 5, pp. 605-613, 2006.
- [2] M. Kobilarov, G. Sukhatme, J. Hyams and P. Batavia: "People tracking and following with mobile robot using an omnidirectional camera and a laser," Proceedings of the 2006 IEEE International Conference on Robotics and Automation, pp. 557-562, 2006.
- [3] 野戸広之，安部正人，城戸健一：“帯域信号の包絡線を利用した音源位置推定，”信学技報，EA92-86, Dec. 1992.