

ロボット視覚と聴覚の連携による部屋の容積の推定

日大生産工(院) ○福田 怜央
日大生産工 霜山 竜一

1. はじめに

人は両耳による聴覚情報を用いて音源までの距離を推定できる. Bronkhorst らは心理物理学実験を通じ, 直接音と反射音のエネルギー比が人の距離知覚に顕著な影響を与えることを示したり, 櫻井らは位相差スペクトルのばらつきを標準偏差で表し, 距離と周波数特性の違いの関係について報告した²⁾. しかし, 標準偏差の周波数特性は距離だけでなく, 部屋の容積に依存する. 本研究では, ヒューマノイド型ロボットの頭部を用いてカメラで物体までの距離を推定し, マイクロホンで検出された, 距離 1[m]における音圧の位相差スペクトルから未知の部屋の容積をリアルタイムで推定する方法について検討した.

2. 聴覚情報による距離推定

反響のある室内では, 音源から放射された音は, 直接耳に到達する直接音と周囲の壁などに反射されて到達する反射音に分けられる. 音源までの距離が遠いほど直接音に対する反射音の割合が大きくなるため, 音圧の位相差スペクトルはばらつく. このばらつきを式(1)の標準偏差値で表す. 中心周波数 f_i の近傍の周波数帯域 Δf に含まれるデータ数を $2m+1$ とすると標準偏差は,

$$\rho_i = \sqrt{\frac{1}{2m+1} \sum_{j=i-m}^{i+m} (\varphi_j - \overline{\varphi_i})^2} \quad (1)$$

で表される. ここで φ_j は周波数 f_j における位相差, φ_j は Δf 内の位相差の平均値である. 中心周波数 f_i を変えて位相差値の標準偏差 ρ_i を求めると, 標準偏差の周波数特性が得られる. この周波数特性を周波数帯域 ΔF で平均した値でばらつきを評価し, 音源までの距離との関係を実験的に求める.

3. 視覚情報による距離推定

音源となる物体の検出に, 比較的頑健な幾



図1 ヒューマノイド型ロボットの頭部



図2 ロボット視覚による音源位置の検出

何学的パターンマッチングを用いた. 距離が変わると画像上の物体の大きさが変わるため物体の中心が正確に求められない. そこで, 検出対象となる物体の画像を, 異なる距離で3種類用意してパターンマッチングに用いた. 検出された画像から物体の方向を計算し, 三角法を用いて距離を求めた. 物体までの距離は実測値を用いて校正してある.

4. 実験条件

オートフォーカス型 Web カメラ (Microsoft, Q2f-00008) とヘッドホン型マイクロホン (Brüel&Kjær, Type4101) を装着したヒューマノイド型ロボット (ヴイストン, M3-neony) の頭部を電動台に固定した. ロボット頭部の外観を図1に示す. 24bit 分解能の A/D 変換ボード (NI, PCI-4474) を装備したワークステーショ

ン(DELL,T3500)で2系統の音圧波形を同時に収録した。ロボット頭部の正面に、高さを合わせてスピーカ(YAMAHA, NS-10MMT)を設置し、連続的に広帯域雑音を発生させた。サンプリング周波数は24[kHz], 各chの1フレームは4800個のデータを含む。 m は250個, Δf は1.25[kHz], ΔF は4~11[kHz]とした。複数の部屋で測定した。測定環境である部屋Aは体育館, 部屋B, C, Dは講義室である。部屋の容積はそれぞれ, $119 \times 10^2 [\text{m}^3]$, $367[\text{m}^3]$, $347[\text{m}^3]$, $113[\text{m}^3]$ である。

5. 実験結果

スピーカまでの距離を変えて音圧を測定した場合の距離と標準偏差の平均値の関係を図3に示す。測定した4つの部屋すべてで、標準偏差の平均値は距離に比例して増加した。距離に対する標準偏差の平均値の傾きは部屋によって異なり、容積が最も大きい部屋Aは他の部屋と比べて傾きが最も小さい。音源までの距離が1[m]の場合の、標準偏差の平均値との標準偏差の平均値の傾きの関係について図4に示す。部屋の容積が大きいほど、標準偏差の平均値とその傾きはともに小さくなる傾向がある。また、距離1[m]における標準偏差の平均値の傾きは標準偏差の平均値に対して傾きが約 $0.8[\text{m}^{-1}]$ の比例関係にあることが分かる。図5に距離1[m]における標準偏差の平均値 $\bar{\rho}$ と部屋の容積 V の関係を示す。べき乗で近似すると、式(2)が得られた。

$$V = e^{(17.38 - 3.93 \times \ln \bar{\rho})} \quad (2)$$

近似曲線との誤差は最大で11[%]で、部屋の容積が小さいほど誤差が大きくなる傾向がみられた。式(2)から部屋の容積が推定できる。

6. あとがき

カメラとマイクロホンを搭載したヒューマノイド型ロボットの頭部を用いて、視覚と聴覚を連携させて部屋の容積を推定する方法について検討した。距離1[m]における標準偏差の平均値と部屋の容積の関係を表す近似式を導いた。この近似式を用いて未知の空間の容積が推定できる可能性がある。今後は、様々な環境で部屋の容積を推定する予定である。

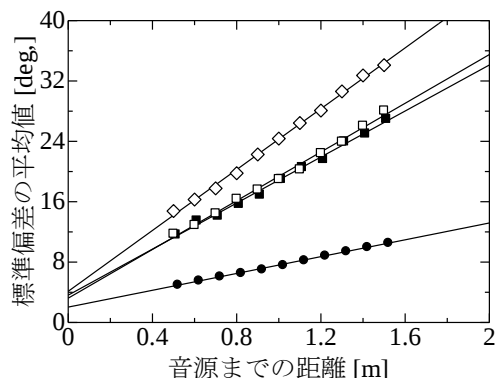


図3 距離と標準偏差の平均値の関係
(●:部屋A ■:部屋B □:部屋C ◇:部屋D)

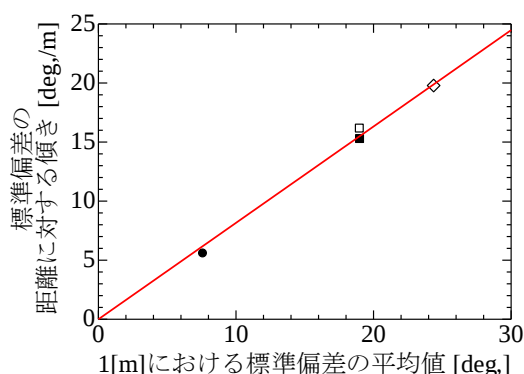


図4 標準偏差の平均値と傾きの関係
(●:部屋A ■:部屋B □:部屋C ◇:部屋D)

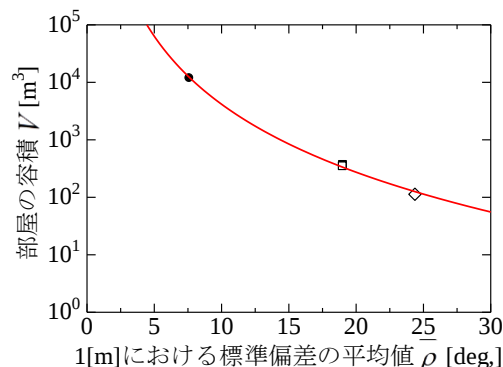


図5 標準偏差の平均値と部屋の容積の関係
(●:部屋A ■:部屋B □:部屋C ◇:部屋D)

「参考文献」

- 1]A.W.Bronkhorst & T.Houtgast
“Auditory distance perception in rooms.”
Nature,397,1999,pp517-520
- 2]櫻井, 霜山 “複数の環境に適応する両耳聴による距離推定法” 電子情報通信学会大会講演論文集,2013,pp137