

## ダミーヘッドに設置したマイクロホンアレーにおける 発話方向検出の検討\*

◎春原卓也<sup>1</sup> 菊池慶子<sup>1</sup> 中島弘史<sup>2</sup> 中臺一博<sup>2</sup> 長谷川雄二<sup>2</sup> 金田豊<sup>1</sup>

1 東京電機大・工 2 (株)ホンダ・リサーチ・インスティテュート・ジャパン

### 1 はじめに

ロボットと人の会話を実現するためには人の発話方向を検出する技術(発話方向検出)は重要である。音情報を利用した発話方向検出は話者が複数いるなど画像情報による検出が場合難しい場合でも利用できるという利点がある。従来、この問題に対しては、室壁に大規模なアレーを設置した方法[1]や室内の反射音を利用した方法[2]が報告されていたが、本研究では、ロボット頭部に設置した小規模アレーで反射音を利用しない、という枠組みで問題を考えることとした。今回はその第一歩として、発話方向によるアレー受信信号の音響的特徴の分析を進めた。

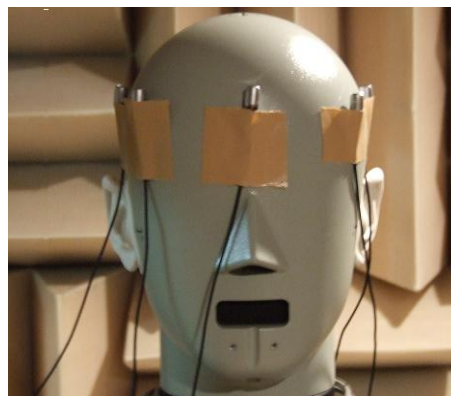


図1 使用したダミーヘッド

### 2 発話方向検出のための收音機器

今回対象とした検出システムは図1に示すように、ダミーヘッド(B&K Type4100)にマイクロホン(SONY ECM-C10) 8個を設置したものである。図2示すように、ダミーヘッド正面部のマイクロホンをM1とし、左回りにM2～M8と定めた。話者は図2に示すように、ダミーヘッドの正面方向で距離 $r = 1\text{m}$ の地点において発話方向 $\theta$ に向いて発話するものとした。

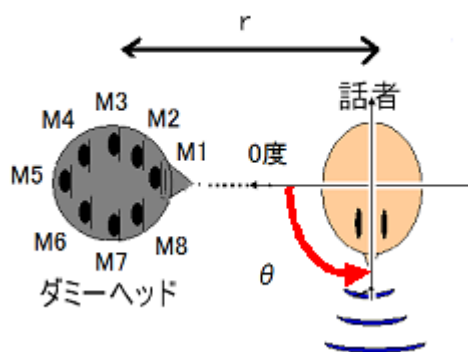


図2 発話方向検出モデル

### 3 分析

#### 3.1 測定条件

発話方向によるマイクロホン出力の特徴を見るために無響室で收音を行った。話者は成人男性3名で、短文「爆音が銀世界の高原に広がる」を発声した。ダミーヘッドの頭部と発声者の頭部は高さがほぼ等しくなるように調整した。発話方向は $\theta = 0$ 度,  $90$ 度,  $180$ 度,  $270$ 度の4方向とし測定時のサンプリング周波数は $16\text{kHz}$ とした。

#### 3.2 予想される結果

- ① 話者がダミーヘッド方向を向いた場合( $\theta = 0$ 度の場合), 高い周波数成分は

ダミーヘッドの陰になるマイクロホンM5に伝わりにくいのでマイクロホンM1とM5の差は大きくなると予想される。一方, $\theta = 180$ 度の場合, 音声は話者の頭を回りこんで伝わるのでマイクロホンM1とM5の差は $\theta = 0$ 度の場合に比べて小さいと予想される。

- ②  $\theta = 0$ 度の場合, ダミーヘッドにおける対称位置のマイクロホンM2とM8, M3とM7, M4とM6からの出力はほぼ等しいと考えられる。これに対して $\theta = 90$ 度方向を向いた場合には音声の放射方向から遠いM2, M3, M4のマイクロホン出力はM6, M7, M8のマイクロホン出力

\* A study of sound source orientation estimation using a microphone array attached to a dummy head, by Takuya HARUBARA<sup>1</sup>, Keiko KIKUCHI<sup>1</sup>, Hirofumi NAKAZIMA<sup>2</sup>, Kazuhiro NAKADAI<sup>2</sup>, Yuji HASEGAWA<sup>2</sup> and Yutaka KANEDA<sup>1</sup> (1 Tokyo Denki University, 2 Honda Research Institute Japan).

と比べて小さいことが予想できる。

### 3.3 特徴の分析と結果

以上の予想に基づいてマイクロホン M1 と M5, および M2 と M8, M3 と M7, M4 と M6 レベル差を分析した。周波数帯域の特徴を得るために、マイクロホン出力の周波数成分を 1kHz 毎の平均パワーのレベル差を算出した。

図 3 にマイクロホン M5-M1 のレベル差の結果を示す。横軸は周波数帯域、縦軸はレベル差を示す。図は 3 人分の発声の 0 度(青), 90 度(黒), 180 度(緑), 270 度(赤)の結果である。1~4kHz までは各方向において差はないが、5~7kHz では 180 度方向のレベル差が -5dB より小さくなっていることがわかる。

図 4 にマイクロホン M3-M7 のレベル差の測定結果を示す。横軸、縦軸は図 3 と同様である。図より 270 度では +2~3dB, 0 度では約 0dB, 90 度では -2~3dB のレベル差があることがわかる。各方向のレベル差は 1~5kHz の範囲で特徴が明確である。

また、マイクロホン M2-M8, M4-M6 のレベル差は M3-M7 と類似の傾向が見られた。 $\theta=180$  度方向についてはレベル差のバラつきが大きいため、この図には描かなかった。

## 4 発話方向検出への応用

今回は上記の結果を利用した第一歩として 4 種類の発話方向 ( $\theta=0, 90, 180, 270$  度) を判別する方法について検討した。

### 4.1 発話方向検出アルゴリズム

本手法では 2 段階の手法となっている。

- ① 180 度方向はマイクロホン M5-M1 の 5~7kHz の平均パワーのレベル差が -5 dB 以下であれば 180 度と判定。
- ② 180 度以外の方向はマイクロホン M3-M7, M2-M8, M4-M6 の 1~5kHz の平均パワーのレベル差平均値 Dif が

$$\text{Dif} < -1\text{dB} \rightarrow 90 \text{ 度}$$

$$-1\text{dB} \leq \text{Dif} \leq 1\text{dB} \rightarrow 0 \text{ 度}$$

$$1\text{dB} < \text{Dif} \rightarrow 270 \text{ 度}$$

と判定。

### 4.2 評価実験

まず、図 3, 図 4 の分析に用いた 3 話者 4 方向のデータに関しては 100%(12/12) の正解率であった(closed test)。その後、同じ

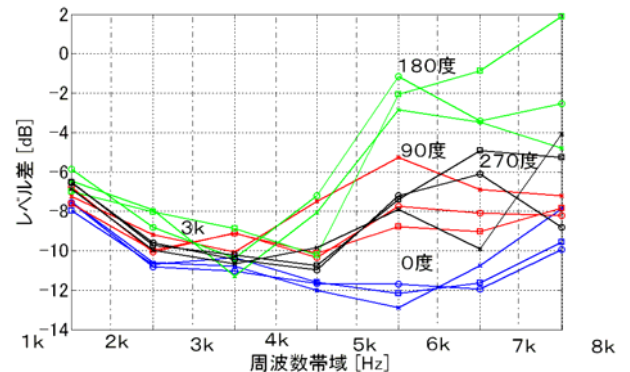


図 3 マイクロホンのレベル差(M5-M1)

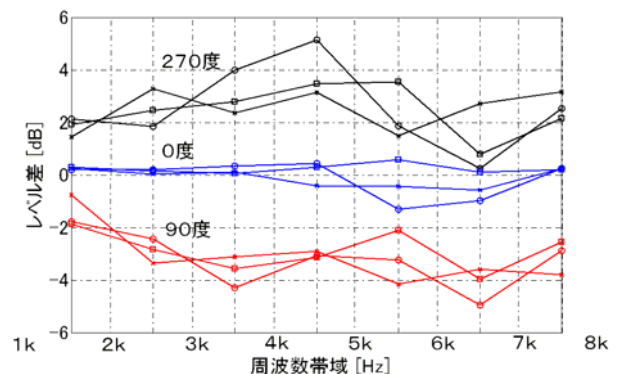


図 4 マイクロホンのレベル差(M3-M7)

3 話者の別発声と別話者男女各 1 名, 2 発声, 計 5 話者 4 方向 2 発声, 計 40 発声に対して判定を行った結果(open test), 95%(38/40) の正解率を得た。

## 5 まとめ

ダミーヘッドに設置したマイクロホンアレーを利用し、発話方向による受信信号の特徴を分析した。分析結果より、180 度方向はマイクロホン M5-M1 のレベル差に、その他の方向はマイクロホン M4-M6, M3-M7, M2-M8 のレベル差に大きな特徴があることが分かった。その結果から発話方向検出の手法を検討し、良好な結果となった。今後は残響や雑音のある実環境での有効性について検討を進めて行く。

### 参考文献

[1] 菊池他, 音講論 (秋) pp.789-790, 2008

[2] 丹羽他, 音講論 (春) pp.775-776, 2009

### 謝辞

本研究のための基礎検討を行った本研究室 OB 井出明氏に謝意を表す。