

聴取者の頭部動作による複数話者の方向推定

田中章悟[†] 苮木禎史[‡] 宇佐川毅[‡]^{†‡}熊本大学大学院自然科学研究科

1 はじめに

人間の聴覚が持つ能力の一部として、人は雑音環境下においても目的とする音源のみを選択的に聴取することが可能である。これは左右耳で観測された信号に生じる時間差やレベル差から音源の方向を捉え、その方向からの音響信号に着目し、他方向からの信号を抑圧するような処理を行っているためと考えられている。このような両耳聴の機能の一部をモデル化したもの1つとして周波数領域両耳聴モデル (Frequency Domain Binaural Model : FDBM) が提案されている [1, 2]。FDBM は音源方向推定とそれに基づく音源信号強調を行うことができるが、複数音源があるとき誤った方向を推定することがあるという問題がある。

この問題を解決する方法の一つとして、人間の身体動作を用いることが考えられる。コミュニケーションにおいて、会話中の頭部動作と対話者の話を聴くための空間領域は重大な関わりがある。また昨今の多軸モーションセンサの小型化・高精度化により、情報機器のみならず日常生活で利用する装身具にセンサを装着することが可能になりつつある。このようなセンサを用いて頭部動作の推定を行うことで、信号強調のための空間領域を推定することができ、FDBM と関係することで、音源方向推定の精度をより良くすることが出来る可能性がある。

本稿では、加速度・地磁気センサによる頭部の方位検出、頭部回転角度の推定を行った。また、ある会話状況における頭部動作を計測し、得られたデータから対話者の方向と頭部動作の関係の検討を行った。

2 加速度、地磁気による頭部回転角度検出

本節では、センサによる頭部動作検出方法について述べる。2.1 項では用いたセンサと加速度・地磁気の検出について述べる。2.2 項ではセンサから得られた加

Estimation of speakers' direction by listener's head movement

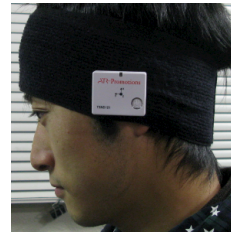
Shogo TANAKA[†], Yoshifumi CHISAKI[‡] and Tsuyoshi USAGAWA[‡]^{†‡}Graduate School of Science and Technology, Kumamoto University, 860-0862, kumamoto, Japan (shogo@hicc., chisaki@, tuie@) cs.kumamoto-u.ac.jp

図 1: A subject wearing a magnetic field sensor TSND121 manufactured by ATR-Promotions.

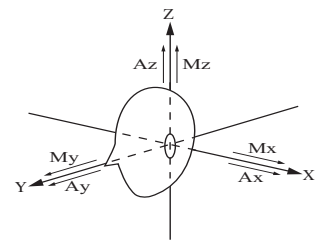


図 2: Relations of acceleration, geomagnetism and the coordinate system when a subject attached a sensor to the left ear.

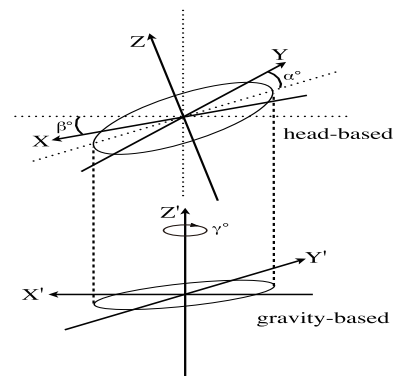


図 3: Mapping onto a gravity-based horizontal plane.

速度、地磁気データによる方位の算出方法について説明する。

2.1 小型無線多機能センサ

本稿では、ATR-Promotions 製の小型無線多機能センサ TSND121 を図 1 に示すように左耳の位置に装着し、3 軸の加速度と地磁気を検出した。図 2 に示す座標系で、加速度 A_x, A_y, A_z 、地磁気 M_x, M_y, M_z が計測される。

2.2 加速度、地磁気による方位算出

身体動作における回転角は、身体に装着したセンサにより計測された重力加速度と地磁気から方位を算出することで得られる [3]。本稿ではこの手法を用いて頭部の回転角を検出した。

頭部が傾いているとき、図 3 のように XYZ 軸が $X'Y'Z'$ 軸に対してそれぞれ $\alpha^\circ, \beta^\circ$ 回転し、頭部動作により、水平方向に γ° 回転したと考える。

センサが傾斜すると、求める方位に誤差が生じてしま

う。そこで、加速度により α, β を求め、下記の式 (1) を用いることでセンサから得られる地磁気の値 M_x, M_y, M_z から傾斜誤差を計算し、重力場を基準とした座標系における水平面に写像した地磁気の値 $M_{x'}, M_{y'}, M_{z'}$ を算出する。

$$\begin{bmatrix} M_{x'} \\ M_{y'} \\ M_{z'} \end{bmatrix} = \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ \sin\alpha\sin\beta & \cos\alpha & -\sin\alpha\cos\beta \\ -\cos\alpha\sin\beta & \sin\alpha & \cos\alpha\cos\beta \end{bmatrix} \begin{bmatrix} M_x \\ M_y \\ M_z \end{bmatrix} \quad (1)$$

写像した地磁気の値を用いて式 (1) を計算することで重力場を基準とした座標系における γ を算出する。この γ が頭部回転角度の情報となる。

$$\gamma = \arctan\left(\frac{M_{y'}}{M_{x'}}\right) \quad (2)$$

3 実験

実験として、会議の状況を想定し、会話をする実験を行った。条件は以下の通りである。

- 椅子に座って 3 分間会話をする
- 会話に集中し、他の行動をしない
- 正面を 0° とし、 $\pm 60^\circ$ 方向にそれぞれ対話者がいる状況で行う

会話中、人は頭部回転運動以外の動作も行う。そのため、人が会話中にどの程度動くのかを測定する予備実験を行った。その結果、話者から見て $\pm 15^\circ$ 以内の範囲で動くことがわかった。そのため、今回の実験では、図 4 のように、 30° 毎の領域に分けて検討を行った。また、領域 X の左右境界をそれぞれ θ_{Xl}, θ_{Xr} 、ある時刻 i におけるサンプルを $S(i)$ とし、 $S(i)$ がどの領域にあるかという判定を、下記の式 (3) で行った。

$$S(i) = \{X | \theta_{Xl} < \gamma < \theta_{Xr}\} \\ X \in \{A, B, \dots, H\} \quad (3)$$

図 5 に、横軸が時間、縦軸が領域として頭部を向けた領域と発話区間を示す。点線の枠が被験者、対話者の発話区間、実線が被験者の向いている方向を示している。また、右側の数値は各領域における $S(i)$ の合計サンプル数を表す。図 5 から、対話者方向の領域である領域 C, G が最もサンプル数が多いことがわかる。また、C の隣である D においても、サンプル数が多く、その他の領域は C, G に比べ少ないことがわかる。以上のことより、被験者は完全に音源の方向を見ているわけではないが、対話者がいない領域に対して、対話者が存在している領域を向いている時間が多いということがわかった。

4 まとめ

本稿では、FDBM による音源方向推定の精度改善を目的とした、頭部動作計測による会話領域の推定の

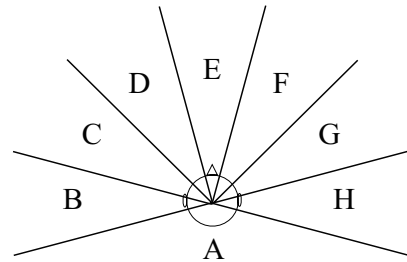


図 4: Partition of spatial zone.

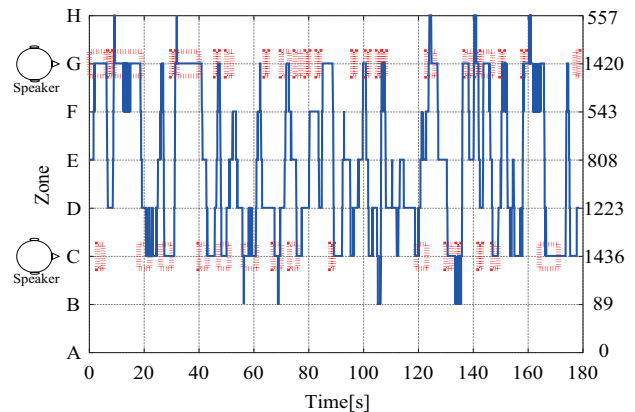


図 5: The head directions and utterance sections when speakers were in zone C and G.

ため、加速度・地磁気センサによる頭部動作の検出、頭部回転角度の推定を行った。また会話実験を行い、センサから得られたデータより、対話者の方向と頭部動作の関係について調べた。その結果、対話者の存在している領域を向いている時間が多いということが示され、頭部動作により対話者の方向が推定できることが示唆された。今後の課題として、センサの入力による会話領域の推定と FDBM による音源方向推定と連係することを目指す。

5 謝辞

本研究の一部は、科学研究費補助金 (基盤 C No.24500147) の支援を受け行ったものである。

参考文献

- [1] Hidetoshi Nakashima, Yoshifumi Chisaki, Tsuyoshi Usagawa and Masano Ebata, "Frequency domain binaural model based on interaural phase and level differences," *Acoustical Science & Technology*, Vol.24, No.4, pp.172-178, 2003.7.
- [2] Yoshifumi Chisaki, Sho Kawano, Kyoko Nagata, Kōtarō Matsuo, Hidetoshi Nakashima and Tsuyoshi Usagawa, "Azimuthal and elevation localization of two sound sources using interaural phase and level differences," *Acoustical Science & Technology*, Vol.29, No.2, pp.139-148, 2008.3.
- [3] 廣瀬圭, 土岐仁, 小田紳介, 永作清, "拡張カルマンフィルタを用いたスピーカ・ターンの運動計測と力学解析に関する研究," *日本機械学会論文集 (C 編)*. 77 巻 774 号 pp. 470-480, 2011.