

ヒトの聴覚能力に基づく汎用モノラル音源方向推定法の実現可能性(継続)

代表研究者 森川大輔 富山県立大学 講師
共同研究者 鵜木祐史 北陸先端科学技術大学院大学 教授

1 研究調査の要旨

これまでの研究から、ヒトは片耳だけである程度正確に音の到来方向を知覚できる。一方、音源方向推定技術では多チャンネルのマイクを使うことが多く、機械的な制約から一般への普及が難しい。マイク1本での音源方向推定法も提案されているが、統計的な学習が使用環境ごとに必要になるため環境の変化に弱く、持ち運んでの利用には適していない。そこで本研究では、ヒトの片耳での音源方向推定の方略を模擬することによるモノラル音源方向推定法の実現可能性について調査を行った。昨年度までの助成によって、変調スペクトル上の情報が音源方向推定に利用できること、シミュレーション上で水平面の音源方向推定が可能であることを示した。今年度の研究では、変調スペクトル上の情報をより詳しく調査することで、変調伝達関数の周波数変化が方向推定に利用できることを明らかにした。また、昨年度発見した情報をより詳しく調査し、これを周波数変化と組み合わせることで、水平面以外を含んだ三次元の全方位の音源方向推定が可能であることを示した。

2 研究背景

我々は普段の生活の中で、音源方向を知覚し役立てている。例えば、他者から声をかけられればその方向を振り向くことができ、後方から迫る自動車や自転車の走行音や警告音を察知することで危機回避に繋がっている。このようなヒトの音源方向推定能力には、音源から耳に届いた音に含まれる、左右での音圧差である両耳間時間差(ILD: Interaural Level Difference)、左右での時間差である両耳間時間差(ITD: Interaural Level Difference)、およびスペクトラルキューとよばれるスペクトルの特定の周波数のピークやノッチが重要であると言われている。これらの音響特徴は、音源から両耳に至るまでの音響伝達関数である頭部伝達関数(HRTF: Head Related Transfer Function)に含まれている[1](図1)。

この中で、片耳のHRTFに含まれている特徴はスペクトラルキューだけであり、ILD、ITDは単耳での音源方向推定に利用することはできない。そのため、健聴者が耳栓で片耳を塞ぎ音源方向推定を行うと、音源は耳栓をしていない耳側に偏って知覚される[2, 3]。一方で、先天的あるいは後天的に片耳の聴力を失ったヒトでは、単耳だけである程度の音源方向知覚が可能であることが知られている[4, 5]。また、Strelnikovらは、訓練によって単耳での音源方向知覚能力が向上することを示している[6]。スペクトラルキューは特定周波数のピークやノッチであるため、実際に到来する音によってスペクトラルキューが打ち消されてしまうことや、到来音のピークやノッチをスペクトラルキューとして誤検出してしまう可能性が高い。さらに、ILDとITDは主に水平方向の知覚に、スペクトラルキューは主に仰角方向の知覚に寄与していることが知られ、スペクトラルキューだけで水平方向の知覚が可能であるかは不明である。したがって、先天的あるいは後天的に片耳の聴力を失ったヒトは、スペクトラルキューを水平方向の知覚にも利用できるように、もしくは他の手がかりを方向の知覚に利用できるように、健聴者とは異なる情報、異なる戦略を用いて方向を知覚していると考えられる。ILD、ITD、スペクトラルキューとは異なる情報として、変調ITD、変調ILDが方向知覚

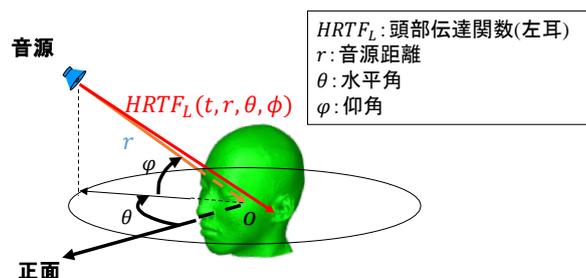


図1: 頭部伝達関数

の手がかりである、という報告が Thompson ら [7] や Furukawa [8] によってなされている。変調 ITD、変調 ILD は両耳間情報のため単耳受聴時には利用できないが、これらの報告から変調領域の情報で単耳でも利用可能な、モノラル変調スペクトルが単耳での方向知覚の手がかりである可能が示唆される。

音源方向推定技術は、先天的な聴覚障害者、高齢者のサポートのための自動車走行音等の到来方向の視覚的呈示や、防犯装置・警告装置への音の到来方向情報の利用等多くのニーズがある。さらに、雑音除去や音声認識といった技術にも前処理段階で音源方向情報が利用される。そのため、音源方向推定技術は古くから数多くの研究がなされてきた。特にマイクロホンアレーを用いた技術が広く知られ、例えばマイクロホン観測信号のマイク間の位相差を利用した遅延和アレー法 [9] や、観測信号と雑音信号の固有空間および固有値を利用した高分解能到来方向推定が可能な MUSIC 法 [10] がある。また、ヒトの方向知覚能力に着目した 2 チャネルの音源方向推定法 [11] など提案されている。しかし、これらの技術ではマイクロホンを複数用いる、一定の間隔で配置するといった制約があり、システムが専用の大規模なものになりやすい。そのため、音源方向推定技術の利用は一部の業務用製品に留まり、個人のニーズに答えることはできていない。音源方向技術を一般に普及させるためには、小型の端末による単一マイクロホンでの音源方向推定技術が必要である。

現在までに研究されている単一マイクロホンでの音源方向推定法として、音源とマイクロホン間の音響モデル、つまり、音源から発せられた音が床、壁、天井あるいは障害物を反射することで音響特徴が変化すること、を利用しこれを統計的に学習することで音源方向推定を行う方法がある [12]。しかし、これらの手法は、音自体や、音源の位置、部屋の環境を様々に変化させた場合の信号を学習する必要があり、環境が変われば再度学習を行う必要があるため、ユーザーが移動しながら用いるといった条件では利用することはできない。Fuchs らの研究 [13] では HRTF、Klipper らの研究 [14] ではモノラル変調スペクトルといったヒト由来の音響特徴を用いて検討を行っているが、統計的に学習していることは他の単一マイクロホンでの音源方向推定法と変わらないため、他の手法と同様に環境の変化には対応できない。

本研究では、単一マイクロホンで様々な環境に対応可能なモノラル音源方向推定法を確立することを目的として、ヒトの単耳受聴時の音源方向推定の方略を模擬したモノラル音源方向推定法の実現可能性について調査を行った。昨年度までの調査では、ヒトが音源方向推定に主に利用している特徴量であるスペクトラルキューだけでは音源方向推定は困難であること、変調スペクトル上の情報が音源方向推定に利用できることがわかった。そして、発見した変調スペクトル上の情報を用いてシミュレーションを行い、水平面の音源方向推定が可能であることを明らかにした。今年度の調査では、変調スペクトル上の情報をより詳しく調査するとともに、音源の変化に対する頑健性についても調べた。また、変調領域以外に利用可能な情報がないか、システムの実装時の課題についても調査した。

2 調査内容

2-1 HRTF の変調伝達関数と受聴信号の変調スペクトル

昨年度の調査において、水平面、正中面、環状面それぞれについて、音源方向推定に変調スペクトルの角度変化が利用可能であることが明らかになったが、それぞれが独立の場合しか考慮していなかった。また、その変化が音源方向推定にロバストに利用可能であるかも課題として残された。そこで、変調スペクトルの角度変化がロバストに利用可能であることを確認するため、周波数による変調伝達関数について調査した。その結果、周波数によって大きさは異なるものの変化の形状は類似していた (図 2, 3)。したがって、これらの変化はロバストに利用可能であると考えられる。次に、三次元全方位に対応しすべての面を同時に扱うために、全方位の変調スペクトルの最大値を網羅的に調査した (図 4)。さらに、この調査の過程で各角度の変調スペクトルの周波数による変化も音源方向推定に利用可能であることがわかった。そこで、この周波数変化についても全方位を網羅的に調査した (図 5)。その結果、それぞれの特徴だけでは角度の特定ができないものの、2 つの特徴を組み合わせることで、三次元全方位から 4 方位のどこかであることを限定できるようになること、運動による時間変化を組み合わせることで、4 方位から 1 方位を選ぶことができ、三次元全方位に対する音源方向推定が可能になることがわかった。実際にこの二つの特徴と頭部運動による特徴の変化を用いた音源方向推定のシミュレーションを行うと、平均誤差 21.9° 、最大誤差 35.1° で音源方向推定を行うことができた。ヒトの両耳受聴での音源方向推定が 30° 間隔のスピーカアレーを用いた場合に 100% 正しく回答することができず、正答率は 90% 程度となることから、十分な精度であると考えられる。

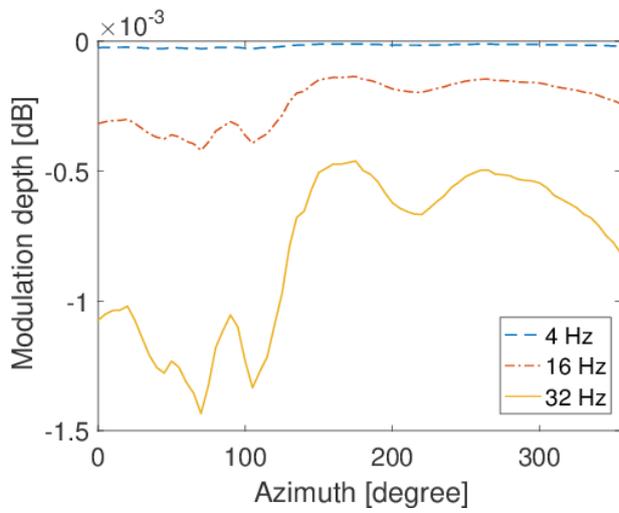


図 2: 水平面変調伝達関数の角度変化の例

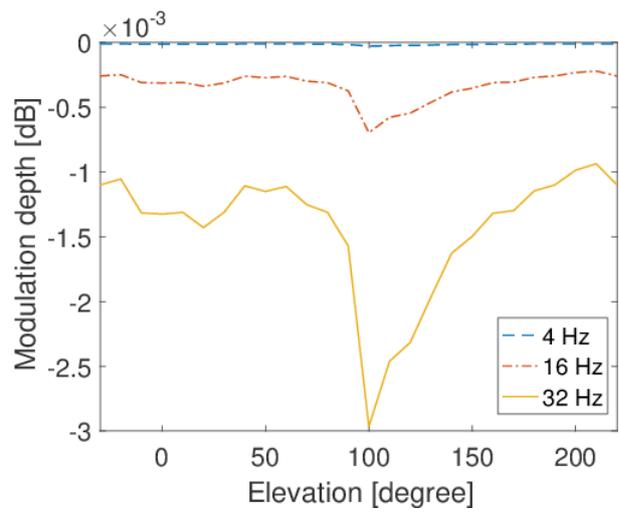


図 3: 正中面変調伝達関数の角度変化の例

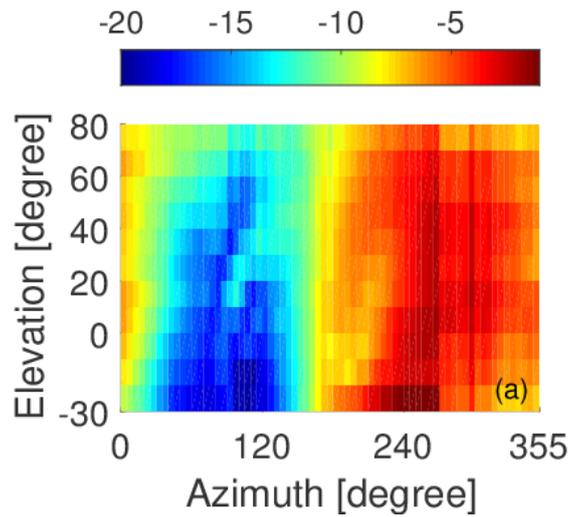


図 3: 全方位の変調スペクトルの最大値の例

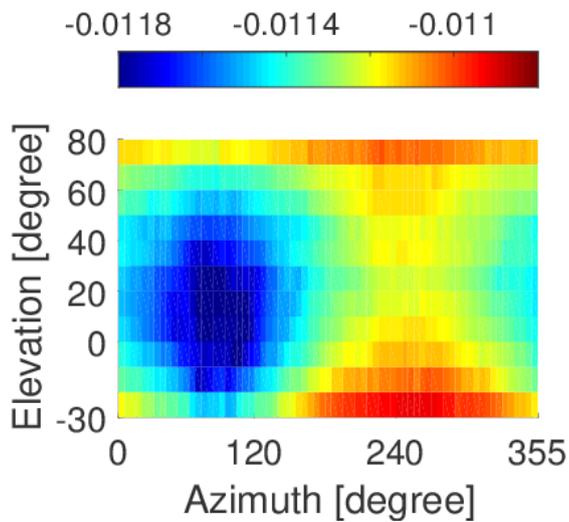


図 4: 全方位の変調スペクトルの周波数による変化の例

2-2 変調領域以外の情報

位相情報を音源方向推定に活用するにあたり、ヒトがどの程度単耳で位相情報を利用できるかを確認するために、モノラルで位相情報を系統的に制御した実験を行うシステムを構築した。構築したシステムを用いて聴取実験を行ったが、健全な聴力をもつ受聴者は単耳での位相情報の利用は難しいようであった。そこで、単耳受聴を普段の生活の中で学習している、単耳しか聴こえていない受聴者で実験を行ったところ、単耳受聴者は音像と音源を別々に知覚可能であるという事象を発見した。この現象は、今回の音源方向推定への利用はできないが、ヒトの聴知覚能力を探るうえで重要な知見であるために、現在別のプロジェクトによってこの現象に関する詳細な調査を進めている[15, 16]。

2-3 システムの実装時の課題

3Dプリンタを用いて様々な条件で耳介を印刷し、その伝達関数を調査した。その結果、有用な音響的な特徴は見つからなかったものの、一般向けに販売されている3Dプリンタであっても音響的に十分な精度のアタッチメントを作成することが可能であることがわかった(図5, 6)。したがって、音源方向推定システムを現状のヒトの形状を元に行った研究にそのまま基づいて作成する際には、一般向けの3Dプリンタで印刷したアタッチメントを用いれば十分に可能である。また、耳介形状等の頭部形状がないと音源方向推定ができないこととなったとしても、一般家庭でアタッチメントを作成しシステムを普及させることに繋げられることがわかった。

2 まとめ

変調スペクトル上の情報をより詳しく調査することで、変調伝達関数の周波数変化が方向推定に利用できることがわかった。また、昨年度発見した変調スペクトルの最大値の情報をより詳しく、三次元全体を網羅的に調査した。その結果、周波数変化と最大値を組み合わせることで、水平面以外を含んだ三次元の全方位の音源方向推定が可能であることを示した。また、実際の実装に向けて3Dプリンタで耳介を印刷し、一般向けの3Dプリンタであっても十分に音源方向推定システムのアタッチメント作成に利用できることを明らかにした。

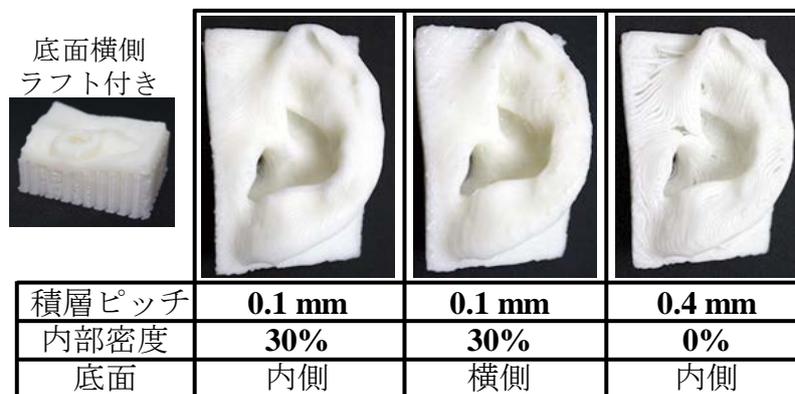


図5:水平面変調伝達関数の角度変化の例

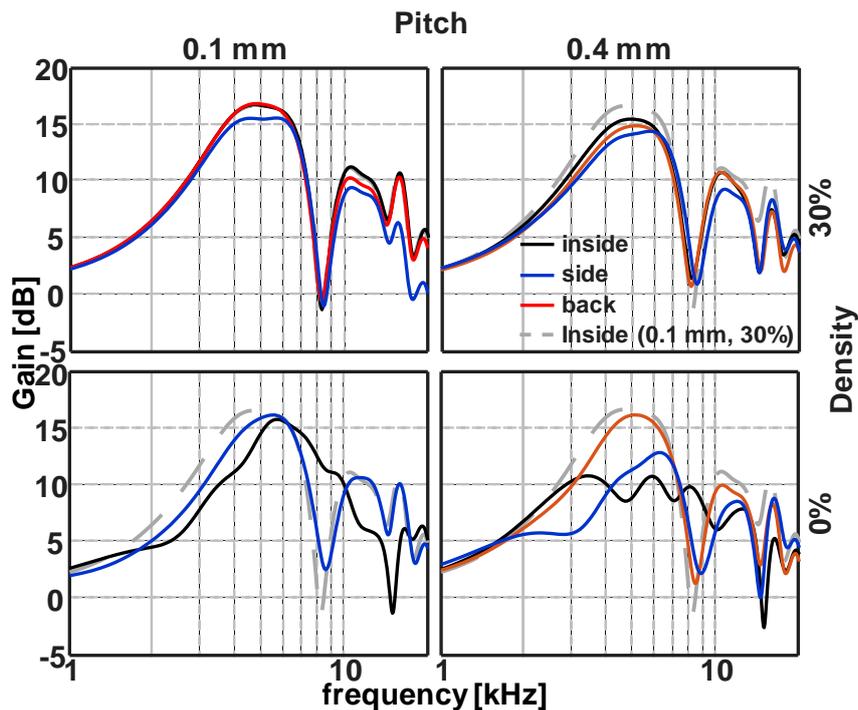


図 6:印刷条件による耳介伝達関数の変化の例

【参考文献】

- [1] イェンス・ブラウエルト, 森本政之, 後藤敏幸, 空間音響, 鹿島出版会, 東京, 1986.
- [2] F. Wightman, D. Kistler, "Monaural sound localization revisited," JASA, vol. 101, no. 2, pp. 1050-1063, 1997.
- [3] 小島大輝, 平原達也, "モノラル水平面音像定位," 信学技報, vol. 115, no. 359, EA2015-12, pp. 31-36, 2015.
- [4] 安藤四一, 森本政之, 依藤庸正, 服部浩, "片耳による定位," 聴覚研究会資料 H-47-2, 1977.
- [5] 佐藤亮祐, 降旗建治, "頭蓋骨損傷が聴覚系に及ぼす影響の一例," 信学技報, vol. 112, no. 266, EA2012-71, pp. 37-42, 2012.
- [6] K. Strelnikov, M. Rosito, and P. Barone, "Effect of Audiovisual Training on Monaural Spatial Hearing in Horizontal Plane," PloS one, vol. 6 no. 3, pp. 1-9, 2011.
- [7] E. R. Thompson and T. Dau, "Binaural processing of modulation interaural level difference," J. Acoust. Soc. Am., vol. 123, No. 2, pp. 1017-1029, 2008.
- [8] S. Furukawa, "DitECTION of simultaneous modulation of interaural time and level differences: Effects of modulation rate and relative phase (L)," J. Acoust. Soc. Am., vol. 132, pp. 1-4, 2012.
- [9] 山崎芳男, 金田豊, 音・音場のデジタル処理, コロナ社, 東京, 2002.
- [10] 大賀寿郎, 山崎芳男, 金田豊, 音響システムとデジタル処理, コロナ社, 東京, 1995.
- [11] 佐保貴哉, 荻木禎史, 宇佐川毅, "狭帯域信号を用いた音源方向の前後弁別手法・周波数両耳聴モデルに基づく補聴システムへの実装の検討," 日本音響学会講演論文集, pp.761-764, 2011.
- [12] 高島遼一, 滝口哲也, 有木康雄, "音響モデル合成を用いた単一マイクによる 2 話者位置推定," 電子情報通信学会論文誌, vol. J96-D, no. 3, pp. 675-685, 2013.
- [13] A. Fuchs, C. Feldbauer, and M. Stark, "Monaural sound localization," Proc. Interspeech 2011, pp. 2521-2524, Florence, Italy, 2011.

- [14] R. Kliper, H. Kayser, D. Weinshall, I. Nelken, and J. Anemuller, “Monaural azimuth localization using spectral dynamics of speech,” Proc. Interspeech 2011, pp. 33–36, Florence, Italy, 2011.
- [15] K. Takahashi and D. Morikawa, “Horizontal localization of sound image and source in monaural congenital deafness,” Proc. 2017 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing, Hawaii, USA, 2017.
- [14] 高橋響子, 森川大輔, “先天性単耳受聴者の水平面における音像定位と音源定位,” 日本音響学会秋季研究発表会, 2017.

〈発 表 資 料〉

題 名	掲載誌・学会名等	発表年月
3D プリンタの印刷条件による耳介伝達関数の変化	日本音響学会秋季研究発表会	2016年9月
Investigation on the head-related modulation transfer function for monaural DOA	The Twelfth International Conference on Intelligent Information Hiding and Multimedia Signal Processing	2016年11月
Investigation on the head-related modulation transfer function for monaural DOA	Advances in Intelligent Information Hiding and Multimedia Signal Processing: Proceeding of the Twelfth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Nov., 21-23, 2016, Kaohsiung, Taiwan, Volume 1	2016年11月
Investigation of head-related modulation transfer function for monaural sound localization	E M M 研 究 会	2017年1月