

# 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

代表研究者 森川大輔 北陸先端科学技術大学院大学 助教  
共同研究者 鵜木祐史 北陸先端科学技術大学院大学 准教授

## 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、統計的な学習が必要になるために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ものある音源方向推定法の十減可能性について調査を行った。その結果、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、周波数ごとの位相の違いを音源方向推定に利用できる可能性があること、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。さらに、変調スペクトルの情報のうち、水平方向の角度変化に着目し、実際に水平方向の音源方向推定をシミュレーション上で行った。その結果、水平面に置いて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、音源から聴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

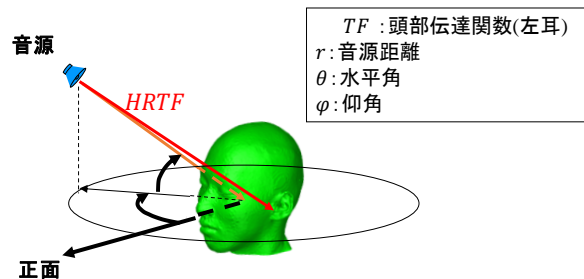


図1: 頭部伝達関数

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]なども提案されている。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。

そこで本研究では、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。

## 2 調査内容

本研究では、ヒトが音の到来方向の知覚に利用している HRTF に含まれる情報のうち、以下の二つの情報に着目してモノラル音源方向推定法の実現可能性を検討した。

1. ヒトの音源方向推定能力に重要な音響的特徴のうち、片耳受聴で利用可能な音響的特徴である HRTF スペクトルのピークやノッチ情報であるスペクトラルキュー。
2. モノラル変調スペクトルの角度毎の関係性とその頭部運動による時間的な変化に着目した、音源方向推定の実現可能な範囲とその性質の調査

### 2-1 スペクトラルキューの角度差とその頭部運動による時間的な変化に着目した、音源方向推定の実現可能な範囲とその性質の調査

調査研究の結果、スペクトラルキューが生じる周波数は到来方向によって複雑に変化するが、その変化は仰角方向での変化が大きく(図 2)、水平方向の変化は小さい(図 3)ことがわかった。そのため、仰角方向の到来方向推定には有用であるが、水平方向の推定には課題が残った。スペクトル全体のレベルが利用可能かについても検討を行ったが、刺激音自体の音圧の影響を受けやすく、ロバストではないことがわかった。

一方、蘆原らの研究[15]では両耳受聴の場合に周波数ごとの位相差から周波数ごとに到来方向を推定し、周波数ごとに合成する伝達関数を変えることで立体的な刺激音を作成できることを明かにし、そのシステムを作成していた。この結果からヒトは周波数ごとの位相も音源方向推定に利用していることが予測される。そこで、周波数領域ではなく時間領域に着目し、周波数ごとの位相情報が両耳受聴だけでなく単耳受聴でも利用できないか検討を行っている。位相は波長が短くなると何波長ずれた上で位相差が生じているかがわかりにくくなってしまいうため、単独で音源方向推定に用いることは困難であるが、他の情報を用いておおまかな方向が分かっているならば、おおまかな方向の近傍範囲の中から到来位置を高い精度で推定可能であることが予測される。これによって、特に水平方向の到来方向推定が可能になることが期待されたため調査を続けていく予定である。

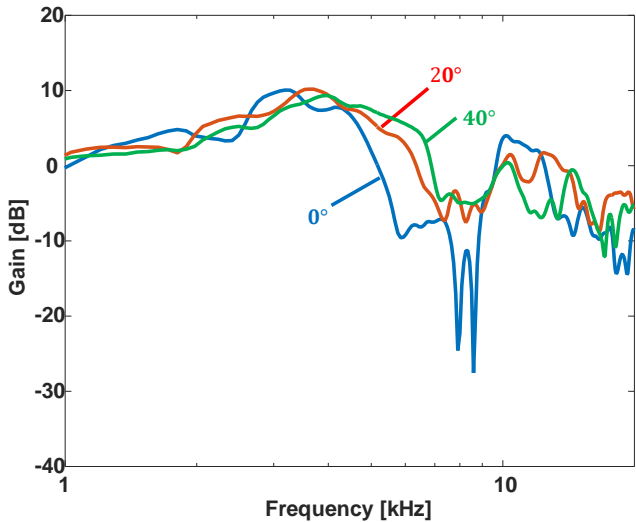


図 2: HRTF の仰角の角度変化の例

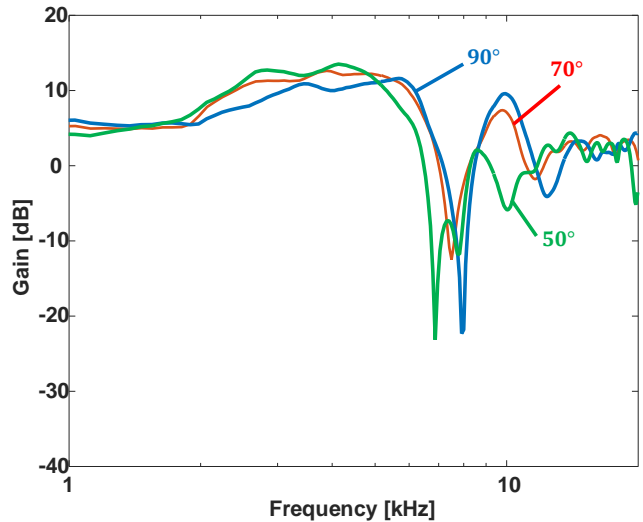


図 3: HRTF の水平角の角度変化の例

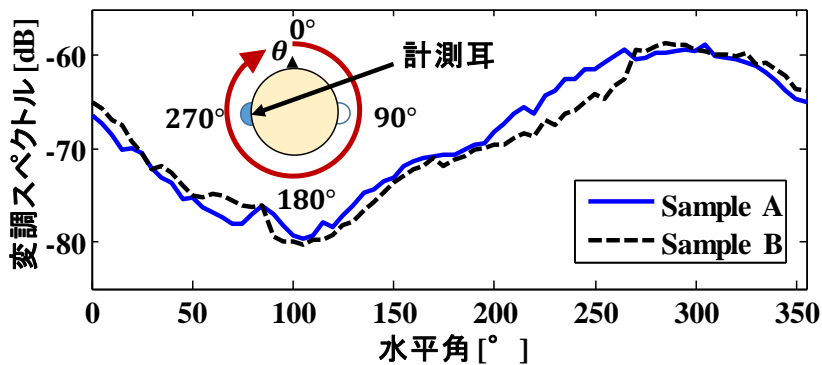


図 4: 変調スペクトルの角度変化の例

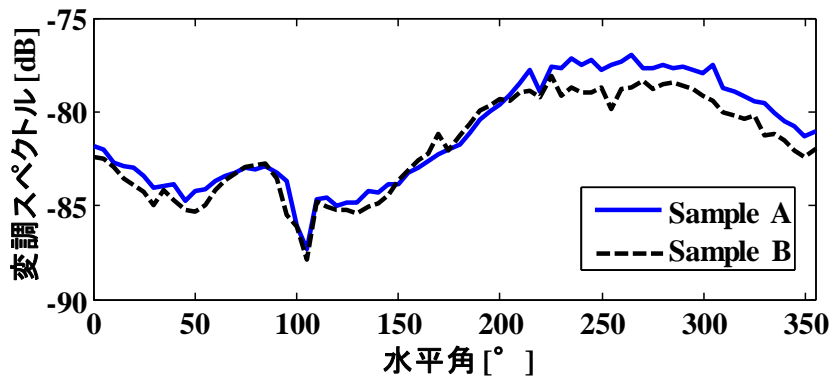


図 5: 変調スペクトルの角度変化の例 (搬送波が低域の場合)

## 2-2 モノラル変調スペクトルの角度毎の関係性とその頭部運動による時間的な変化に着目した、音源方向推定の実現可能な範囲とその性質の調査

変調スペクトルの変化について調査した結果、水平方向では、耳のある側面では頭部中心に対して計測耳の角度を頂点とした上向きの弧を、耳と反対側の側面では頭部中心に対して計測耳でない耳の角度を頂点とした下向きの弧を描くように変調スペクトルが変化することがわかった(図 4)。この傾向は頭部の遮蔽によるものと考えられ、ITD および ILD の傾向と類似していた。また、変調周波数が変化しても変調スペクトルのダイナミックレンジはあまり変化しないことがわかった。ただし、搬送波が低域成分だけの場合、対側側は下向きの弧ではなくなり、計測耳でない耳の角度の変調スペクトルも大きくなった(図 5)。これは波長が

頭部よりも長くなることによる回折の影響と考えられる。

仰角方向では、正中面は頭頂方向から背面方向に向けて変調スペクトルが減衰することがわかった。しかし、正面方向から正中面方向では変調スペクトルの変化傾向が検討に用いる HRTF によって異なった。これは HRTF の個人性によるものと考えられる。また、環状面では、耳のある側面に対して耳と反対側の側面で変調スペクトルが小さいことがわかった。

両耳受聴時に主に音源方向推定に利用される特徴量とモノラル変調スペクトルの角度変化が同じ傾向であることから、水平方向の音源方向推定に最も利用できる可能性が高いと考えられたため、水平方向に関する音源方向推定をシミュレーション上で行った。その結果、水平面に置いて音源方向推定が可能であることを示した(図 6)。今後は、これらの傾向がどの程度ロバストに利用可能か調査を続ける予定である。

### 2-3 まとめ

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、周波数ごとの位相の違いを音源方向推定に利用できる可能性があること、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。周波数ごとの位相の違いは、単独で音源方向推定に用いることは困難であるが、他の情報を用いておおまかな方向・距離が分かっているならば、おおまかな方向・距離の近傍範囲の中から到来位置を高い精度で推定可能であることが予測される。一方、変調スペクトル上の情報は、水平面において十分に大まかな方向推定が可能であることが示唆された。したがって、これらを組み合わせることで、音源方向推定法が確立できると考えられる。また変調スペクトルでの音源方向推定が難しくなる仰角については、スペクトラルキューを組み合わせることで、音源方向推定法が確立できると考えられる。

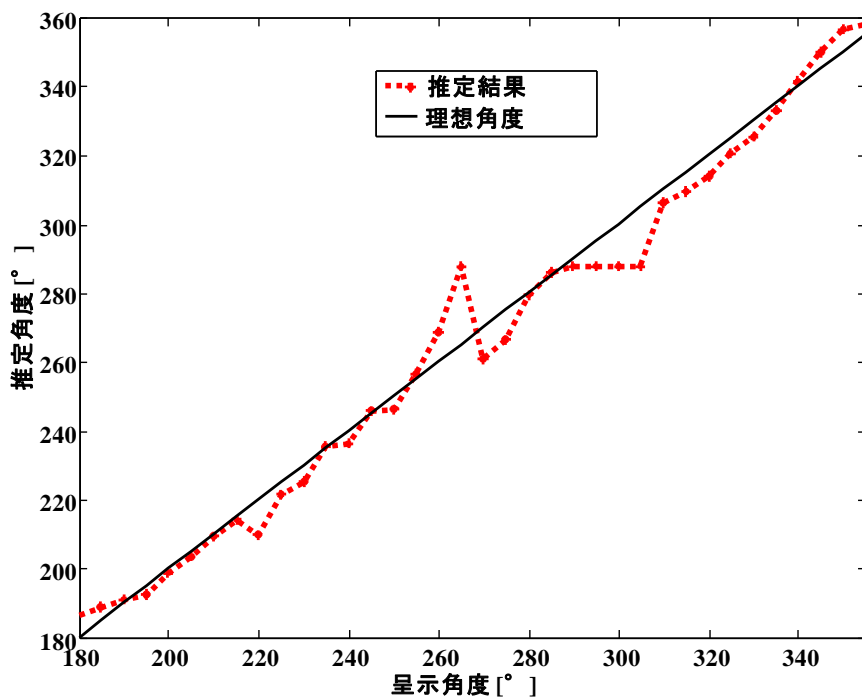


図 5: 音源方向推定のシミュレーション結果

## 【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, 苮木禎史, 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] 岡崎裕人, 蘆原郁, 齋藤毅, 三好正人, “ミニチュアヘッドシミュレータで収録した音空間の再現性について,” 聴覚研究会資料, vol. 43, no. 1, pp.49-54, 2013.

## 〈発表資料〉

題 名	掲載誌・学会名等	発表年月
Feasibility of estimating direction of arrival based on monaural modulation spectrum	The Eleventh International Conference on Intelligent Information Hiding and Multimedia Signal Processing	2015年9月23~25日