

Title	ヒトの方向知覚能力に着目したモノラル音源方向推定法に関する研究
Author(s)	安藤, 将
Citation	
Issue Date	2014-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/12051
Rights	
Description	Supervisor: 鶴木祐史, 情報科学研究科, 修士

修 士 論 文

ヒトの方向知覚能力に着目した
モノラル音源方向推定法に関する研究

北陸先端科学技術大学院大学
情報科学研究科情報科学専攻

安藤 将

2014年3月

修士論文

ヒトの方向知覚能力に着目した モノラル音源方向推定法に関する研究

指導教員 鷓木 祐史 准教授

審査委員主査 鷓木 祐史 准教授
審査委員 赤木 正人 教授
審査委員 党 建武 教授

北陸先端科学技術大学院大学
情報科学研究科情報科学専攻

1210004 安藤 将

提出年月: 2014年2月

概要

我々の暮らしの中で、音源方向を知覚することは重要な役割を果たしている。例えば、後方から接近してくる自動車の走行音を察知することで、我々は回避行動をとることができる。これは、我々がどの方向から音が到来してくるかを瞬時に判断しているからである。このようなヒトの音源方向知覚能力は、音源信号が両耳の鼓膜に到達するまでの時間差や音圧差が重要な手がかりであると言われている。また、音源信号が頭部を回折したり、耳介を反射することによって周波数成分が変化することで生じる、スペクトルキューも重要な手がかりの一つとされている。さらに、先天的あるいは後天的に単耳の聴力を失ったヒトでも、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。単耳では耳介の反射によるスペクトルキューの情報が利用できるが、これだけでは情報が不十分であり、他の手がかりが利用されていると考えられる。Thompsonらは、聴取実験によって、両耳において両耳間変調レベル差および両耳間変調時間差もまた方向知覚に重要な手がかりであると報告している。この報告はモノラル変調スペクトル(MMS)が単耳における方向知覚の手がかりであると示唆している。これらのような、ヒトの音源方向知覚能力を解明することは、聴覚のメカニズムを知る上でも非常に重要な課題である。さらに、工学的に応用することができれば、単一マイクロフォンによる音源方向推定の実現が可能である。

工学的な音源方向推定では、室の反射特性等の音響的特徴を大規模に事前学習するような仕組みがない限り、単一マイクロフォンで正確な音源方向を推定できない。これに対し、Klipperらは、MMSパターンを利用した手法を提案した。この手法では、スペクトラルキューではなく、MMSパターンを機械学習することで音源方向推定を実現している。しかし、この手法では統計的手法を用いてMMSパターンを分類しているため、どのようなMMSが方向知覚の役割を果たしているかその仕組みを説明できない。

本研究では、ヒトの単耳方向知覚の手がかりを解明することを目的とする。また、解明された手がかりを利用したモノラル音源方向推定法の検討も同時に行う。まず、ヒトの頭部インパルス応答(HRIR)の変調伝達関数(MTF)に着目し、MMSの値が音源信号の到来角度によってどのように変化するか調査した。その結果、音源信号をAM信号とした場合において、水平角では観測した耳と同側ではMMSが耳位置を頂点とする上向きの弧状となり、対側ではMMSが反対の耳位置を頂点とする下向きの弧状となる規則性がみられ、頭部回折や肩部反射の影響もみられた。また、仰角では頭頂部付近から後頭部にかけて、MMSが減衰する傾向がみられた。これらの結果は、ヒトの方向知覚に十分に利用可能であるため、ヒトはMMSを音源方向知覚に利用している可能性が高いことが示唆された。

以上の調査結果に基づき、回帰近似を利用した音源方向推定法を検討した。理想的な条件を設けた上で、シミュレーションによって検証を行った結果、マイク(単耳)と同側の水平角についてはモノラル音源方向推定法の有効性を示された。

目次

第1章	序論	1
1.1	はじめに	1
1.2	本研究の目的	3
1.3	本論文の構成	3
第2章	ヒトの音源方向知覚の手がかり	4
2.1	頭部インパルス応答 (HRIR)	4
2.2	バイノーラル	7
2.3	モノラル	8
2.4	本研究の着眼点	10
第3章	モノラル音源方向知覚の手がかり	11
3.1	変調伝達関数 (MTF) の概念に基づくモデルの定式化	11
3.2	モノラル変調スペクトルの角度変化特性の調査	13
3.2.1	調査方法	13
3.2.2	調査結果: 水平面	15
3.2.3	調査結果: 水平面 (低域成分に限定したもの)	24
3.2.4	調査結果: 水平面 (肩部の影響)	28
3.2.5	調査結果: 正中面	31
3.2.6	調査結果: 横断面	34
3.3	考察	37
3.4	まとめ	38
第4章	変調スペクトルに着目したモノラル音源方向推定法	39
4.1	水平角に着目した音源方向推定法	39
4.2	仰角に着目した音源方向推定法	39
4.3	水平角に着目した音源方向推定の実現例	40
4.4	提案法	40
4.5	シミュレーションによる提案法の検証	44
4.5.1	検証方法および条件	44
4.5.2	検証結果	44
4.6	前後弁別法	48

4.7	前後弁別の検証	48
4.7.1	検証方法および条件	48
4.7.2	検証結果	49
4.8	考察	53
4.9	まとめ	53
第5章	結論	54
5.1	本研究で明らかになったことの要約	54
5.2	残された課題	54

目次

2.1	水平方向 θ と仰角方向 ϕ , 音源からの距離 r で表される頭部インパルス応答 (HRIR) の模式図	5
2.2	水平面の HRTF の一例	6
2.3	ITD	7
2.4	ILD	8
2.5	正中面の HRTF の一例	9
3.1	音源信号と鼓膜での観測信号の伝達関数モデル化: (a) 信号波形に対する HRTF と (b) パワーエンベロープに対する HRMTF	12
3.2	分析のイメージ図: (a) 水平面, (b) 正中面, (c) 横断面	14
3.3	MMS の変化: 水平面の場合, AM 信号	16
3.4	MMS の変化: 水平面の場合, AM 雑音	17
3.5	変調周波数ごとの MMS の変化: 搬送波 AM 信号, 変調周波数 2 Hz	18
3.6	変調周波数ごとの MMS の変化: 搬送波 AM 雑音, 変調周波数 2 Hz	19
3.7	変調周波数ごとの MMS の変化: 搬送波 AM 信号, 変調周波数 20 Hz	20
3.8	変調周波数ごとの MMS の変化: 搬送波 AM 雑音, 変調周波数 20 Hz	21
3.9	変調周波数ごとの MMS の変化: 搬送波 AM 信号, 変調周波数 200 Hz	22
3.10	変調周波数ごとの MMS の変化: 搬送波 AM 雑音, 変調周波数 200 Hz	23
3.11	MMS の変化: 水平面の場合 (低域成分, 回折の影響), AM 信号	25
3.12	MMS の変化: 水平面の場合 (低域成分, 回折の影響), AM 雑音	26
3.13	頭部回折の影響の模式図	27
3.14	MMS の変化: 水平面の場合 (肩部有無の差, 反射の影響)	29
3.15	肩の反射の影響の模式図	30
3.16	MMS の変化: 正中面の場合, AM 信号	32
3.17	MMS の変化: 正中面の場合, AM 雑音	33
3.18	MMS の変化: 横断面の場合, AM 信号	35
3.19	MMS の変化: 横断面の場合, AM 雑音	36
4.1	提案法: (a) 事前学習 (b) 方向推定	40
4.2	回帰曲線 AM 信号	42
4.3	回帰曲線 AM 雑音	43
4.4	音源方向推定結果: 2 Hz	45

4.5	音源方向推定結果: 20 Hz	46
4.6	音源方向推定結果: 200 Hz	47
4.7	前後弁別後の音源方向推定結果: 2 Hz	50
4.8	前後弁別後の音源方向推定結果: 20 Hz	51
4.9	前後弁別後の音源方向推定結果: 200 Hz	52

第1章 序論

1.1 はじめに

ヒトは優れた音源方向知覚能力を有している．例えば，後方から接近してきた自動車の走行音を察知することで，我々は回避行動を取ることができる．さらに，その自動車が自身の横を通りすぎていく様子を視認することなく知覚可能である．また，多数の話者がいる環境であっても，各話者がどこから話しているかを把握することができる．これらは音源方向を瞬時に判断する能力および音源を追従する能力をヒトが備えていることを示している．これらのようなヒトの優れた音源方向知覚は，頭部伝達関数（Head Related Transfer Function: HRTF）に含まれる両耳間時間差（interaural time difference: ITD），両耳間レベル差（interaural level difference: ILD），スペクトラルキューを主な手がかりとしていることが，これまでの研究によって明らかにされている [1, 2]．それぞれの詳細については第二章で述べるが，我々は，バイノーラルで受聴した音声信号の左右での時間差やレベル差，あるいは音声信号のスペクトルに含まれるノッチやピークを知覚することで，音声信号の到来方向を知覚している．

一方，ヒトは先天的，あるいは後天的に片方の聴力を失った場合でも，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 [3, 4]．これらの知見から，ヒトは音を単耳のみで聴取した場合でも何らかの手がかりを元に音源方向知覚が可能であることがわかる．音源方向知覚の主な手がかりの中で，単耳での方向知覚に利用できる手がかりはスペクトラルキューだけである [2]．しかし，スペクトラルキューはHRTFスペクトルのピークやノッチであるため，音源信号のスペクトルによっては，前者のピークやノッチが発生する周波数が変化したり，消失してしまう．したがって，単耳での音源方向知覚を行うための情報が欠落してしまうため，スペクトラルキューだけでは単耳の方向知覚の手がかりとして不十分であり，他の手がかりが利用されていると考えられる．

Thompson[5]らは，両耳間変調レベル差および両耳間変調時間差もまた方向知覚に重要な手がかりであると報告している．また，古川は，ITDおよびILDを個別に処理する情報処理機構では，少なくとも変調周波数 50 Hz 程度までの方向に関する情報を保持していると報告している [6, 7, 8]．これらの報告はモノラル変調スペクトル (monaural modulation spectrum: MMS) が単耳における方向知覚の手がかりである可能性を示唆している．しかし，MMS が手がかりであるかどうかは，現在までに解明されていない．

以上のように，ヒトの優れた音源方向知覚能力は未だに解明されていないことが多い．そのため，これらを解明することは，ヒトの聴覚のメカニズムを知る上で非常に重要な課

題である．さらに，ヒトの優れた聴覚のメカニズムを工学的に応用することができれば，高精度な音源方向推定法の実現が可能である．

音源方向推定法は，音声分野の中でも基幹的な技術であり，雑音除去や音声認識などの音源方向が必要な応用技術の前処理段階で利用される．そのため，音源方向推定法の推定精度は直接応用技術の性能に影響を与える．そこで，高精度な音源方向推定法は1900年代後半から数多く研究されてきた．音源方向推定法は，マイクロフォンアレーを用いた技術が広く知られている [9, 10, 11]．例えば，マイクロフォンアレーで収録された観測信号間の位相差を利用した遅延和アレー法 [12] 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 [13] がある．さらに，近年では周波数両耳聴モデルに基づく音源方向推定 [14, 15] など，ヒトの方向知覚能力を応用した音源方向推定法の研究が注目されている．しかし，これらのようなマイクロフォンアレーを用いた技術では，マイクロフォンを一定の間隔で複数配置しなければならないという機械的制約が必ず発生するため，専用のシステムが必要かつ大規模になりやすい．そのため，すでに普及している機器を音源方向推定に利用したり，小型の端末を持ち運んで容易に音源方向推定を行ったりするためには，単一マイクロフォンでの音源方向推定技術が必要である．

近年では，広く普及しているスマートフォンなどの携帯端末に，少なくとも一つマイクロフォンが搭載されている．したがって，単一マイクロフォンによる音源方向推定（以下，モノラル音源方向推定とよぶ）が実現できれば，特殊な装置を用いることなく，聴覚が不自由な人へのサポートが可能になる（例えば，自動車の音源方向を推定し視覚的に危険を促すなど）．さらに，単一マイクロフォンを球状の監視カメラ等と組み合わせることで，異音方向に即座にカメラを向けるなどの防犯性能の向上も期待できる．

現在までに研究されているモノラル音源方向推定法として，音源とマイクロフォン間の音響モデルを学習し，統計的手法によって音源方向を推定する手法 [16, 17] がある．これらの手法では，音源から発せられた音声信号が，壁や障害物を反射することで変化する音響的特徴を利用し，音源や音源位置，室空間を様々に変化させた場合の観測信号を学習し，統計的に分類することで方向推定を実現している．また，単耳の HRTF を用いて音源方向を推定する手法が Fuchs ら [18] によって提案されている．さらに，Klipper ら [19] は，鼓膜で観測された信号の MMS パターンを機械学習し，統計的分類することで音源方向推定を実現している．しかしながら，これらの手法ではヒトの鼓膜で観測される信号を利用してモノラル音源方向推定を実現しているが，いずれも方向を推定する段階で統計的な手法を用いているため，どのような手がかりがヒトの方向知覚に利用されているかまでは明らかにされていない．

1.2 本研究の目的

本研究の目的は、二つある。一つ目は、ヒトのモノラル音源方向知覚の手がかりを明らかにすることである。二つ目は、明らかにした手がかりを利用したモノラル音源方向推定法を提案することである。バイノーラルでは、音源信号の振幅包絡の時間差や音圧差を左右方向の知覚に利用していることが現在までに明らかとなっているが、モノラルでは、音源信号の振幅包絡情報を音源方向の知覚に利用されているかどうか明らかでない。そこで、MMSに着目し、頭部インパルス応答 (Head Related Impulse Response: HRIR) の変調伝達関数 (Modulation Transfer Function: MTF) に着目し、鼓膜で観測される信号の変調スペクトルの角度変化を調査することで、一つ目の目的の達成を図る。続いて、調査した結果を利用したモノラル音源方向推定法を検討し、シミュレーションによって提案法の有効性を検証することで、二つ目の目的の達成を図る。

1.3 本論文の構成

本論文は5章で構成される。

第1章 本研究の背景，問題点，目的を述べる。

第2章 これまでに研究されてきたヒトの音源方向知覚能力をバイノーラル，モノラルについてそれぞれ解説する。その上で，本研究で着目するMMSについて述べる。

第3章 ヒトのモノラル音源方向知覚の手がかりが，MMSであるかどうかを明らかにするための調査について述べる。まず，調査の前段階としてヒトのモノラルモデルの定式化について解説する。次に，MMSの角度変化の調査方法および結果を示す。最後に，全ての調査結果について包括的な考察をする。

第4章 第3章の調査結果を利用した音源方向推定法を検討し，シミュレーションを用いて有効性を検証する。また，提案法の改善策である前後弁別法について検討し，シミュレーションを用いて有効性を検証する。

第5章 本研究により明らかになったことを述べる。加えて残された課題について議論する。

第2章 ヒトの音源方向知覚の手がかり

2.1 頭部インパルス応答 (HRIR)

音源から鼓膜までのインパルス応答を HRIR という。ある音源から発せられた音源信号は、肩や顔、耳介の影響を受けて鼓膜に到達する。図 2.1 に HRIR の模式図を示す。HRIR は、仰角 ϕ 度、水平角 θ 度、音源位置から頭部中心位置までの距離 r_m をパラメータとする変数である [1, 20]。図 2.1 では、左耳の HRIR の模式図が示されているが、右耳の HRIR も同様に求められる。

HRIR を周波数領域で表すと HRTF となる。HRTF はヒトの音源方向知覚の手がかりとなる物理特性であり、ITD, ILD, スペクトラルキュー (ピークやノッチ) が含まれる。また、ヒトの頭部や外耳の形状、肩などによる影響で音源位置によって変化する伝達関数である。ヒトの頭部や外耳の形状は個人差があるため、HRTF も個人差がある。図 2.2 に仰角 0 度、水平面の左耳の HRTF の一例を示す (a) から (c) はそれぞれ、0 度、315 度、270 度 (左耳位置) の HRTF を示す。ここで、それぞれの図の 3 から 5 kHz にピークが、6 から 8 kHz に鋭いノッチ (谷) が観察される。これらは、耳介特有の形状によって音源信号が反射することで生じる。ヒトはこれらのスペクトラルキューの変化を知覚していることが、現在までに知られている [2]。

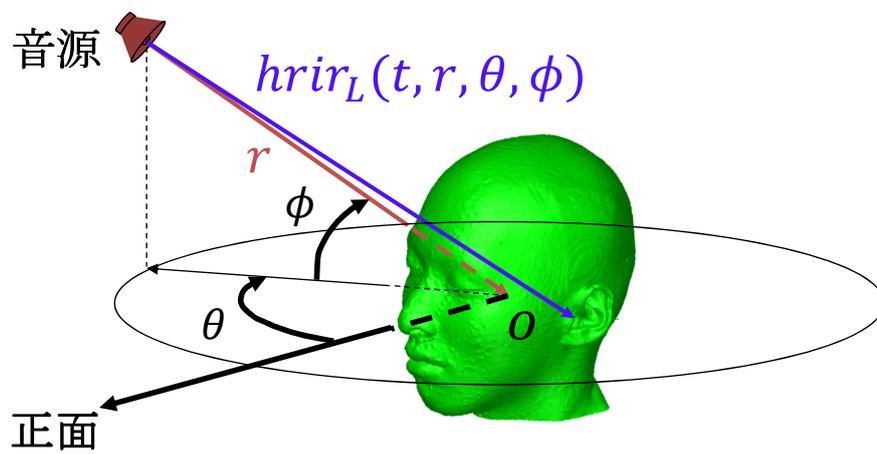


図 2.1: 水平方向 θ と仰角方向 ϕ , 音源からの距離 r で表される頭部インパルス応答 (HRIR) の模式図

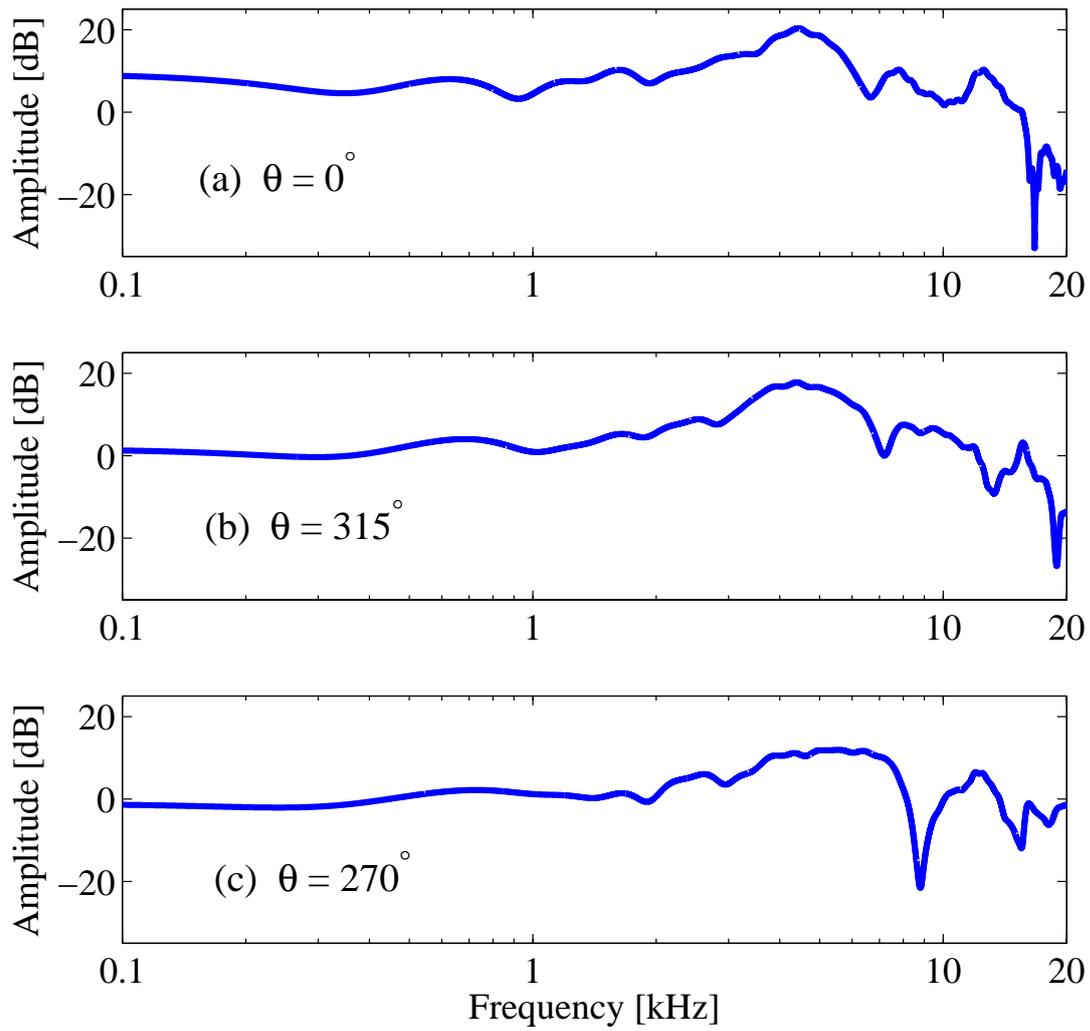


図 2.2: 水平面の HRTF の一例

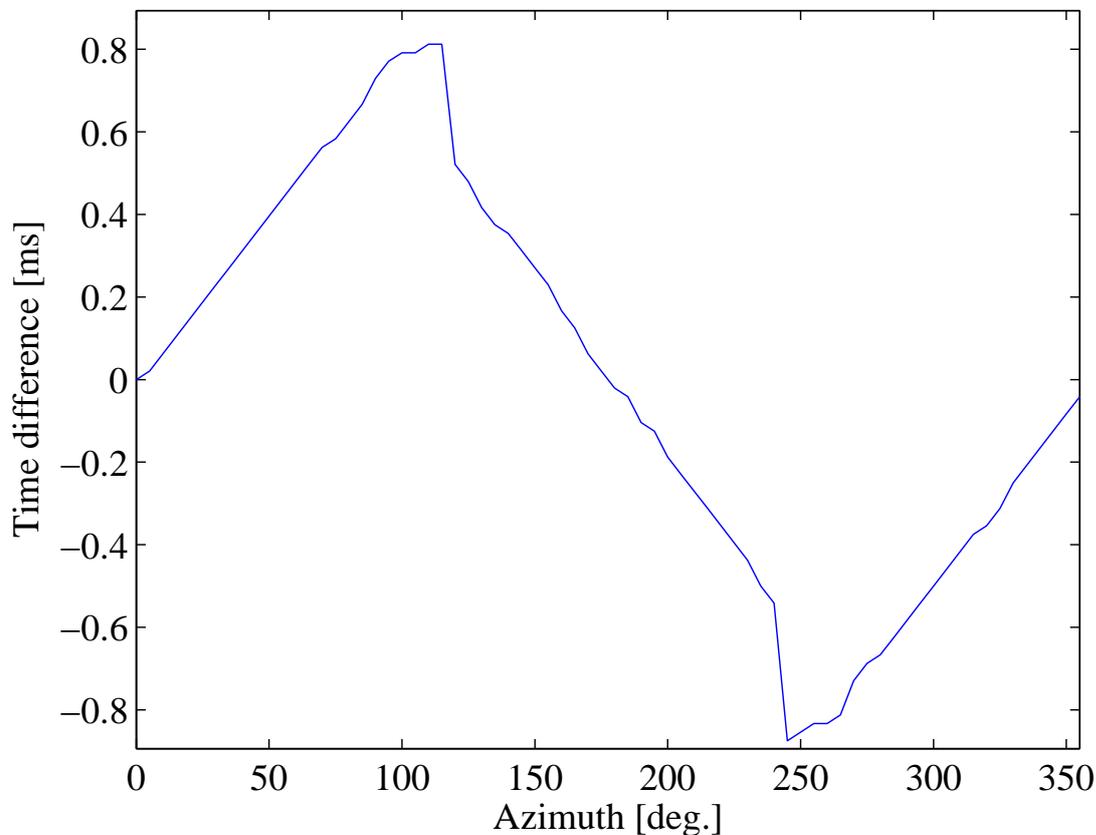


図 2.3: ITD

2.2 バイノーラル

ヒトの両耳は頭部の両側についているため、音源信号が側方から入射した場合、両耳への到達時間および音圧に差が生じる。左右の方向知覚の手がかりは、このような頭部伝達関数により生じる ITD および ILD であるといわれている [1, 2],

両耳の HRIR から求めた時間差を図 2.3 に示す。正面および背面では 0 ms、側方で約 1 ms である。ただし、両耳への入力信号そのものの時間差が左右方向の知覚の手がかりになっているのは、約 1600 Hz 以下に限られる。それ以上の周波数帯域では、後述するが両耳入力信号の包絡線の時間差が手がかりとなっている [5, 21]。

両耳の HRIR から求めたレベル差を図 2.4 に示す。ILD は ITD と同様に、正面および背面では差が最小となり、側方では差が最大となる。また、ILD は可聴周波数全域にわたって左右方向の知覚の手がかりとなっている。

一方で、Thompson らは、ヒトの方向知覚能力には両耳観測信号の振幅包絡成分の時間差やレベル差もまた重要な手がかりであると報告している [5]。これに関連して、古川は、ヒトは ITD および ILD が変動した場合は変調周波数が 10 Hz 程度であっても、その変調

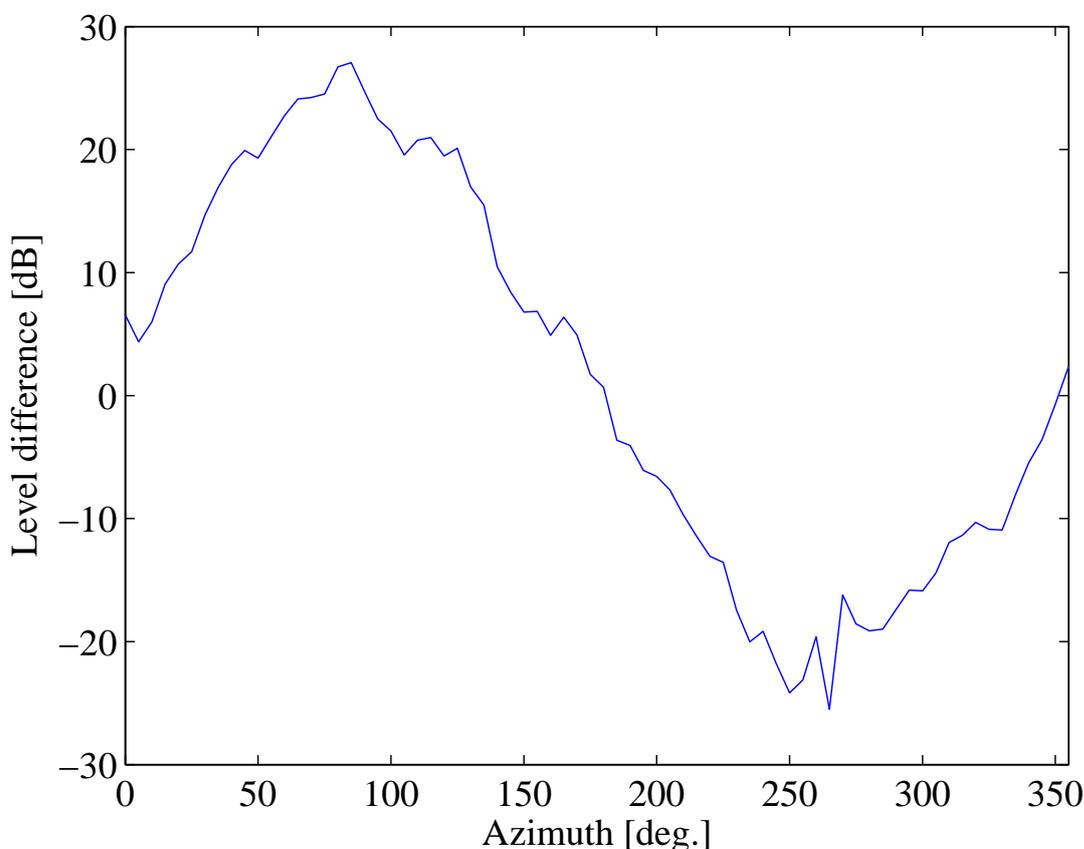


図 2.4: ILD

に追従できない（緩慢性）としているが，ITD および ILD を個別に処理する情報処理機構（脳幹の上オリーブ内側核および外側核に相当する [22]）では，少なくとも変調周波数 50 Hz 程度までの方向に関する情報を保持していると報告している [6, 7, 8]．つまり，緩慢性は ITD および ILD の個別の情報を統合する機構に起因する可能性があるとし唆している．これらの報告は，ヒトが言語知覚のために変調周波数 20 Hz 程度までの変調成分を利用しているという報告と類似している [23]．

2.3 モノラル

バイノーラルでは ITD や ILD などの情報を主に利用しているが，モノラルでは利用することができない．そのため，バイノーラルの音源方向知覚の手がかりの中でも，スペクトラルキューが主に利用されている．図 2.5 に，正中面の HRTF の一例を示す．図中には前後および上下方向の知覚の主な手がかりである第一ピーク（P1）が 3 から 4 kHz に，第一ノッチ（N1）が 6 から 8 kHz に，第二ノッチ（N2）が 11 kHz 付近にそれぞれ示され

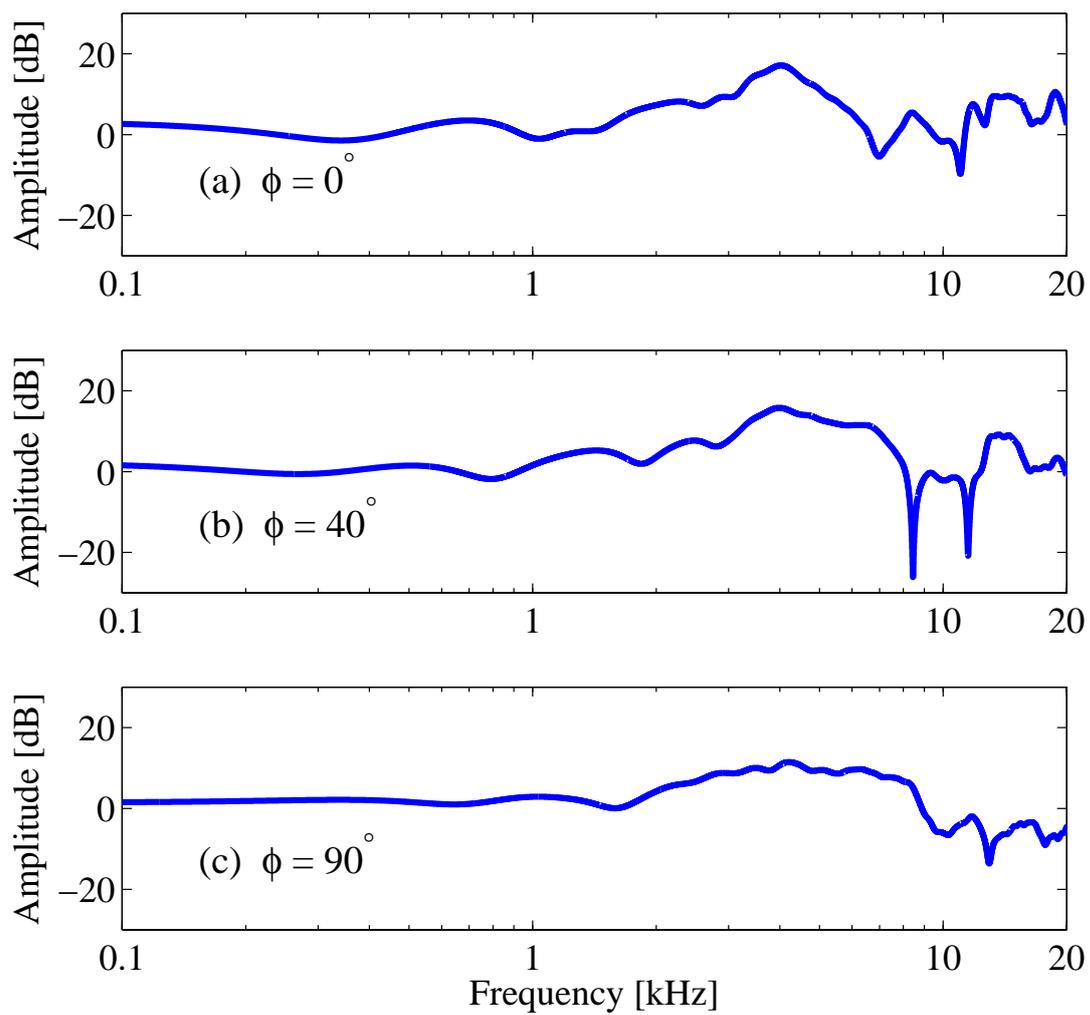


図 2.5: 正中面の HRTF の一例

ている．上下方向で N1 および N2 の周波数が推移していることがわかる．

また，左右方向においても，音源が正面から単耳と同じ側に変化するとピークやノッチの生じる周波数がやや上昇することがわかっている（図 2.2 参照）しかしながら，HRTF は伝達関数であるため，実環境音などの音源信号が単耳に到来した場合，スペクトラルキューのピークやノッチが不明確になる他，音源信号に由来するピークやノッチが生じてしまう．つまり，HRTF のスペクトラルキューに含まれていた音源方向の情報をヒトが知覚することが困難となってしまう．そのため，ピークやノッチだけでは左右方向のモノラル音源方向知覚の手がかりとして不十分であると考えられる．

2.4 本研究の着眼点

前節までに，ヒトのバイノーラルおよびモノラルの音源方向知覚能力について述べた．バイノーラルでは，音源信号の振幅包絡の時間差や音圧差が左右方向の知覚に利用されていることが現在までに明らかとなっているが，モノラルでは，音源信号の振幅包絡情報が音源方向の知覚に利用されているかどうか明らかでない．古川の報告では，ヒトが変調された ITD および ILD の情報を個別に処理する過程では保持しているが，高次の機構ではそれらの情報が打ち消されるとしている．

本研究では，モノラルでは個別に処理する過程で保持されている情報，すなわち MMS に着目し，モノラル音源方向知覚の手がかりを調査する．具体的には，MMS の角度変化を調査し，モノラル音源方向知覚の手がかりが含まれているかどうか確認する．さらに，どの変調周波数の MMS にその手がかりが含まれているのかも確認する．また，単耳の観測信号には音源から鼓膜に直接到来した音（直接音）や，頭部や耳介を回折して鼓膜に到来した音（回折音），肩などを反射して鼓膜に到来した音（反射音）が含まれると考えられる．そこで，これらが MMS にどのような影響を及ぼしているのかを同時に調査する．

第3章 モノラル音源方向知覚の手がかり

3.1 変調伝達関数 (MTF) の概念に基づくモデルの定式化

モノラル音源方向知覚の手がかりを調査する前段階として、音源信号と単耳の観測信号の関係を、時間領域および変調周波数領域でモデルを定式化する。

図 3.1(a) が示すように、音源信号に HRIR が畳み込まれた観測信号を鼓膜で受聴する。観測信号 $y(t, \theta, \phi)$ は次式で表される。

$$y(t, \theta, \phi) = h(t, \theta, \phi) * x(t) \quad (3.1)$$

ここで、 $h(t, \theta, \phi)$ は HRIR、 $x(t)$ は音源信号である。また、 $*$ は畳み込み演算子である。式 (3.1) を周波数領域で表現すると次式となる。

$$Y(f, \theta, \phi) = H(f, \theta, \phi)X(f) \quad (3.2)$$

ここで、 $Y(f, \theta, \phi)$ は観測信号のスペクトル、 $H(f, \theta, \phi)$ は HRTF、 $X(f)$ は音源信号のスペクトルである。また、 f は信号の周波数である。

時間領域におけるこれらの関係は、MTF の概念に基づき [24], [25]、パワーエンベロップ領域では図 3.1(b) のように示される。図 3.1(b) の各パラメータは以下のように導出される。

$$x(t) = e_x(t)c_x(t) \quad (3.3)$$

$$h(t, \theta, \phi) = e_h(t, \theta, \phi)c_h(t) \quad (3.4)$$

$$y(t, \theta, \phi) = e_y(t, \theta, \phi)c_y(t) = h(t, \theta, \phi) * x(t) \quad (3.5)$$

ただし、 $e_x(t)$ 、 $e_h(t, \theta, \phi)$ 、 $e_y(t, \theta, \phi)$ はそれぞれ $x(t)$ 、 $h(t, \theta, \phi)$ 、 $y(t, \theta, \phi)$ の振幅包絡線 (エンベロップ)、 $c_x(t)$ 、 $c_h(t)$ 、 $c_y(t)$ は白色ガウス雑音の特性を有するランダム変数である。ランダム変数には次式の特性がある。

$$\langle c_l(t)c_l(\tau) \rangle = \delta(t - \tau), \quad l \in x, h, y \quad (3.6)$$

ここで、 $\langle \cdot \rangle$ は集合平均を表す。この特性を活用し、 $x(t)$ 、 $h(t, \theta, \phi)$ および $y(t, \theta, \phi)$ の 2 乗集合平均を求めると、観測信号 $y(t, \theta, \phi)$ のパワーエンベロップ $e_y^2(t, \theta, \phi)$ は次式のよ

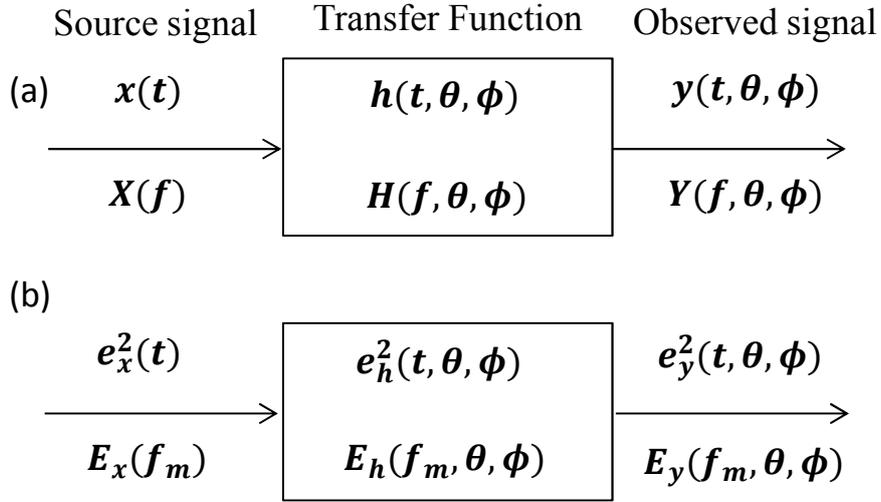


図 3.1: 音源信号と鼓膜での観測信号の伝達関数モデル化: (a) 信号波形に対する HRTF と (b) パワーエンベロープに対する HRMTF

うに導出される .

$$\langle x^2(t) \rangle = e_x^2(t) \quad (3.7)$$

$$\langle h^2(t, \theta, \phi) \rangle = e_h^2(t, \theta, \phi) \quad (3.8)$$

$$\langle y^2(t, \theta, \phi) \rangle = \langle (h(t, \theta, \phi) * x(t))^2 \rangle \quad (3.9)$$

$$= \langle h^2(t, \theta, \phi) * x^2(t) \rangle \quad (3.10)$$

$$e_y^2(t, \theta, \phi) = e_h^2(t, \theta, \phi) * e_x^2(t) \quad (3.11)$$

また , 式 (7) を変調周波数領域で表現すると , 次の関係式となる .

$$E_y(f_m, \theta, \phi) = E_h(f_m, \theta, \phi) E_x(f_m) \quad (3.12)$$

ここで , $E_y(f_m, \theta, \phi)$ は $y(t, \theta, \phi)$ の変調スペクトル , $E_h(f_m, \theta, \phi)$ は head-related MTF(HRMTF) , $E_x(f_m)$ は $x(t)$ の変調スペクトルである . HRMTF は次式で定義される .

$$E_h(f_m, \theta, \phi) = \int_0^\infty h^2(t) \exp(-j2\pi f_m t) dt \quad (3.13)$$

本研究では , $y(t, \theta, \phi)$ から $E_y(f_m, \theta, \phi)$ を導出するために , 次式を利用してパワーエンベロープ $e_y^2(t, \theta, \phi)$ を抽出した .

$$e_y^2(t, \theta, \phi) = \text{LPF} \left[|y(t, \theta, \phi) + j \text{Hilbert}[y(t, \theta, \phi)]|^2 \right] \quad (3.14)$$

ここで , $\text{LPF}[\cdot]$ は低域通過フィルタであり , $\text{Hilbert}[\cdot]$ は Hilbert 変換である . この方法は , 信号の瞬時振幅の計算に基づいたものであり , パワーエンベロープに含まれる高周波成分を取り除くための後処理として低域通過フィルタを利用している . AM 変調の概念では ,

搬送波周波数は変調周波数よりも十分大きくなければならない．ここで，ヒトの可聴域を考慮すると，音源信号（搬送波信号）の周波数は最大で 20 kHz である．そのため，LPF のカットオフ周波数を 200 Hz とした．最後に， $e_y^2(t, \theta, \phi)$ を FFT することで， $E_y(f_m, \theta, \phi)$ を得る．

3.2 モノラル変調スペクトルの角度変化特性の調査

3.2.1 調査方法

観測信号の変調スペクトルがどのような角度変化特性となるかを明確にするために，MMS の角度変化を調査した．音源信号と HRIR の畳み込み積によって算出した単耳の観測信号から式 (3.14) によってパワーエンベロープを得て，それを FFT することで $E_y(f_m, \theta, \phi)$ を求めた．

音源信号は，変調度 1.0，変調周波数 1 Hz から 200 Hz，搬送波を 10 kHz の正弦波とした AM 信号，および搬送波を白色雑音とした AM 雑音を用いた．HRIR は，東北大学電気通信研究所の HRTF データベースを用いた．このデータベースには 114 名 (228 耳) 分の HRIR が収録されている．収録点は水平角 5 度刻み，仰角 10 度刻みの真下を除く計 1225 点である．サンプリング周波数は 48 kHz である．なお，データベース内の HRIR のうち，ハムなどの雑音や隣合う角度間の低域成分の変化が著しく大きいものを除外した，56 名分の HRIR を分析に用いた．除外したデータの一部を付録 A に記載する．HRIR には測定室内のスピーカなどの反射波が含まれるため，HRIR の最大値の 10 サンプル前から，立下り部分を 63 サンプルの余弦関数とした 140 サンプル（約 3 ms）の矩形窓で切り出す処理を施した [26]．

調査は仰角 0-180 度の水平面，水平角 0-180 度の正中面，水平角 90-270 度の横断面について行った．図 3.2 に各面の模式図を示す．また，音源信号の周波数成分による影響を明らかにするため，水平面の調査では搬送波周波数を 1.5 kHz とした AM 信号および搬送波の白色雑音を 1.5 kHz のローパスフィルタに通した低域通過雑音についても調査を行った．さらに，肩部がない HRIR と肩部がある HRIR の MMS の角度変化特性を調査した．肩部がある HRIR を用いた場合と，肩部がない頭部だけの HRIR を用いた場合の MMS を算出し，前者から後者を減算する．これにより，MMS の差分のみが抽出される．ここでは，頭部だけの影響が含まれる HRIR と，頭部と肩の影響が含まれる HRIR を利用した．HRIR は，東北大学電気通信研究所の無響室にて，Brüel&Kjær の HATS を利用して直接法で測定された富山県立大学平原研究室の HRTF データベースを用いた．他の結果で利用した HRIR と異なり，水平角 10 度刻みで収録されている．サンプリング周波数は 48 kHz である．音源信号は AM 雑音を用いた．

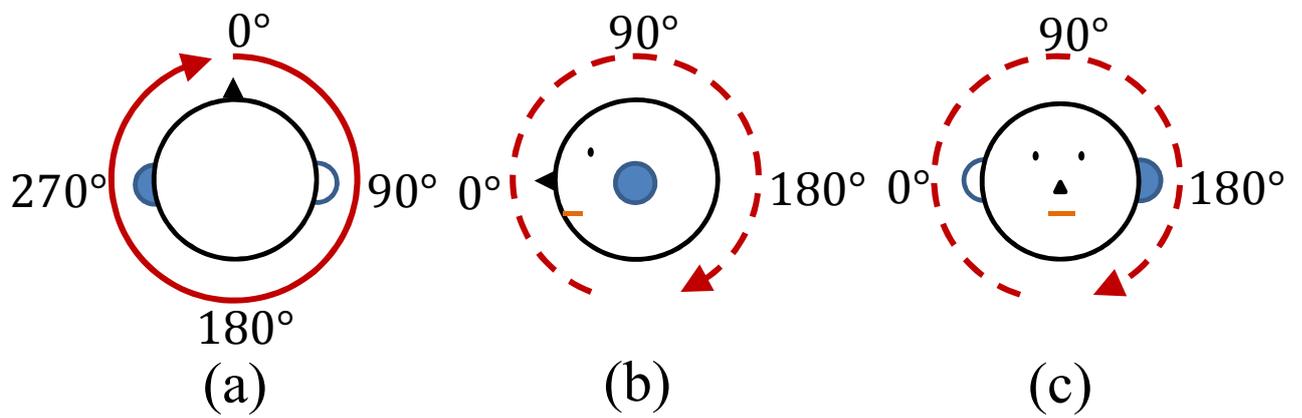


図 3.2: 分析のイメージ図: (a) 水平面, (b) 正中面, (c) 横断面

3.2.2 調査結果: 水平面

図 3.3 から図 3.4 に、代表的な 2 名分の左耳の分析結果を示す。音源信号の変調周波数が 2, 20, 200 Hz の結果のみを示す。横軸は水平角を、縦軸は変調スペクトルの値を示す。図 3.3 および図 3.4 から、AM 信号と AM 雑音共に、水平面では MMS は角度に伴って正弦波のように変化した。ここで、左耳がある半面（180 度から 360 度）に着目すると、左耳がある位置（270 度）付近を頂点とした上向きの弧形となることが分かった。一方、左耳と反対側の半面（0 度から 180 度）に着目すると、右耳がある位置（90 度）付近を頂点とした下向きの弧形となることが分かった。この結果は、ITD および ILD の形状と類似しており、頭部の遮蔽による影響であると考えられる。また、AM 信号よりも AM 雑音では、弧の形状がなめらかであることが分かった。また、変調周波数が変化しても、ダイナミックレンジはあまり変化せず、およそ 20 dB 程度であることが分かった。右耳の HRIR や、他者の HRIR を用いた場合でもこの傾向は類似していたが、弧の頂点の位置や弧の鋭さが異なった。これらの変化は HRIR の個人性に依存するものだと考えられる。

続いて、図 3.5 から図 3.10 に音源信号の変調周波数が変化したときの、変調周波数ごとの MMS の角度変化特性の代表的な 1 名分の結果を示す。これらの図は、図 3.3 から図 3.4 についてそれぞれ、変調周波数 1 Hz から 200 Hz の結果を一系列に並べたものと等しい。ここで、各図の横軸は変調周波数、縦軸は水平角を示す。また、図中の色は MMS の値を示しており、赤が値が高く、青が値が低い。図 3.5 および図 3.6 は変調周波数 2 Hz で変調した AM 信号を観測した信号の MMS の角度変化特性である。そのため、変調周波数 2 Hz 付近の MMS の値が左耳がある半面で高くなっている。

次に図 3.7 および図 3.8 を見る。これらの図は変調周波数 20 Hz で変調した AM 信号を観測した信号の MMS の角度変化特性である。そのため、変調周波数 20 Hz 付近の MMS の値が左耳がある半面で高くなっている。さらに、変調周波数 0 Hz も同様に MMS の値が高いことがわかる。また、他の変調周波数帯では、倍音構造のようになっている。

次に図 3.9 および図 3.10 を見る。これらの図は変調周波数 200 Hz で変調した AM 信号を観測した信号の MMS の角度変化特性である。そのため、変調周波数 200 Hz 付近の MMS の値が左耳がある半面で高くなっている。さらに、変調周波数 0 Hz も同様に MMS の値が高いことがわかる。また、図 3.7 および図 3.8 と同様に他の変調周波数帯では、倍音構造のようになっている。また、右耳の場合や、残りの 55 名分も同様の傾向であった。

以上の結果から、音源の変調周波数が上昇しても、AM 信号の種類に依らず、低域の変調周波数帯に図 3.3 から図 3.17 と同様の傾向があることが分かった。

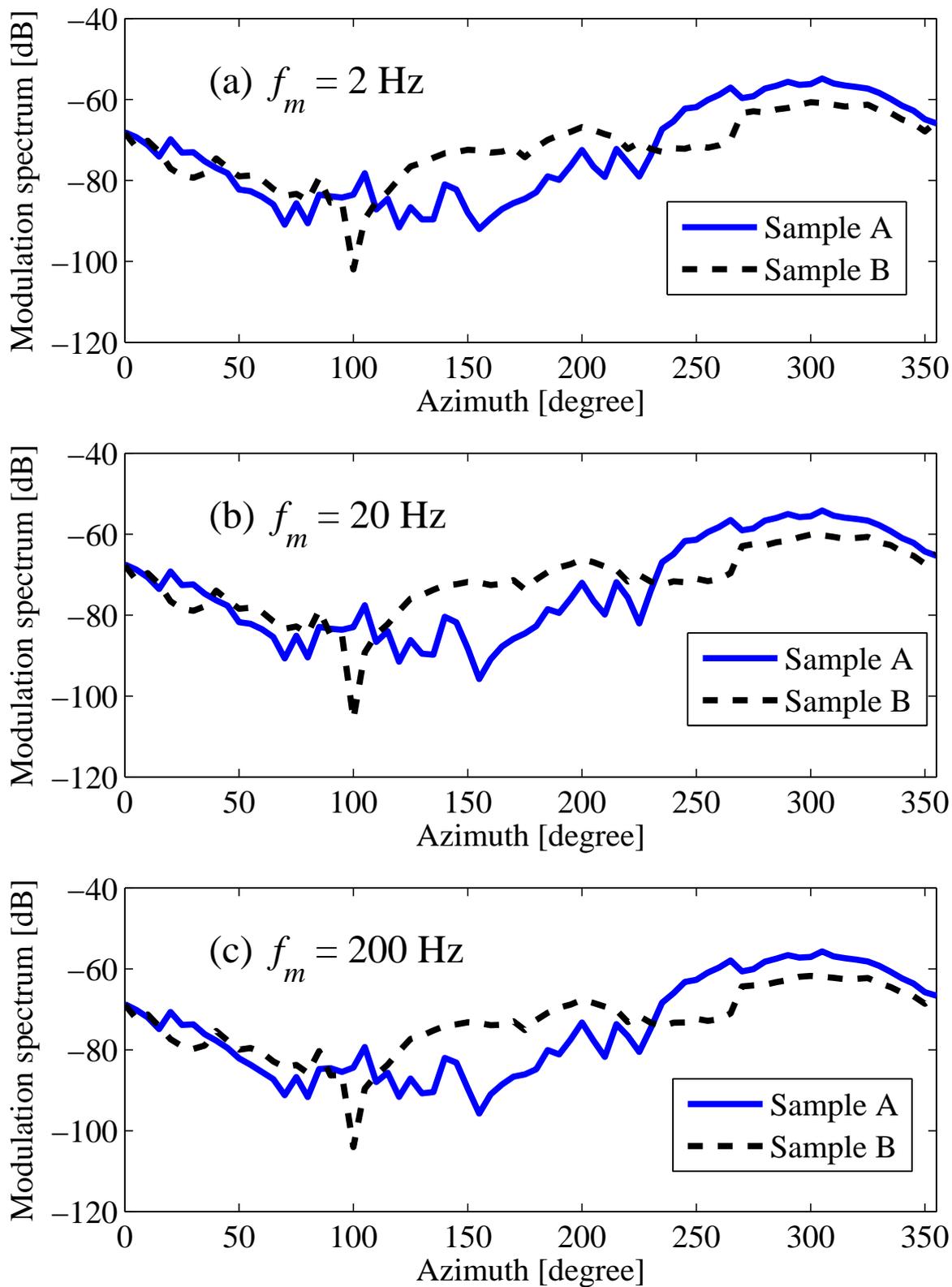


図 3.3: MMS の変化: 水平面の場合, AM 信号

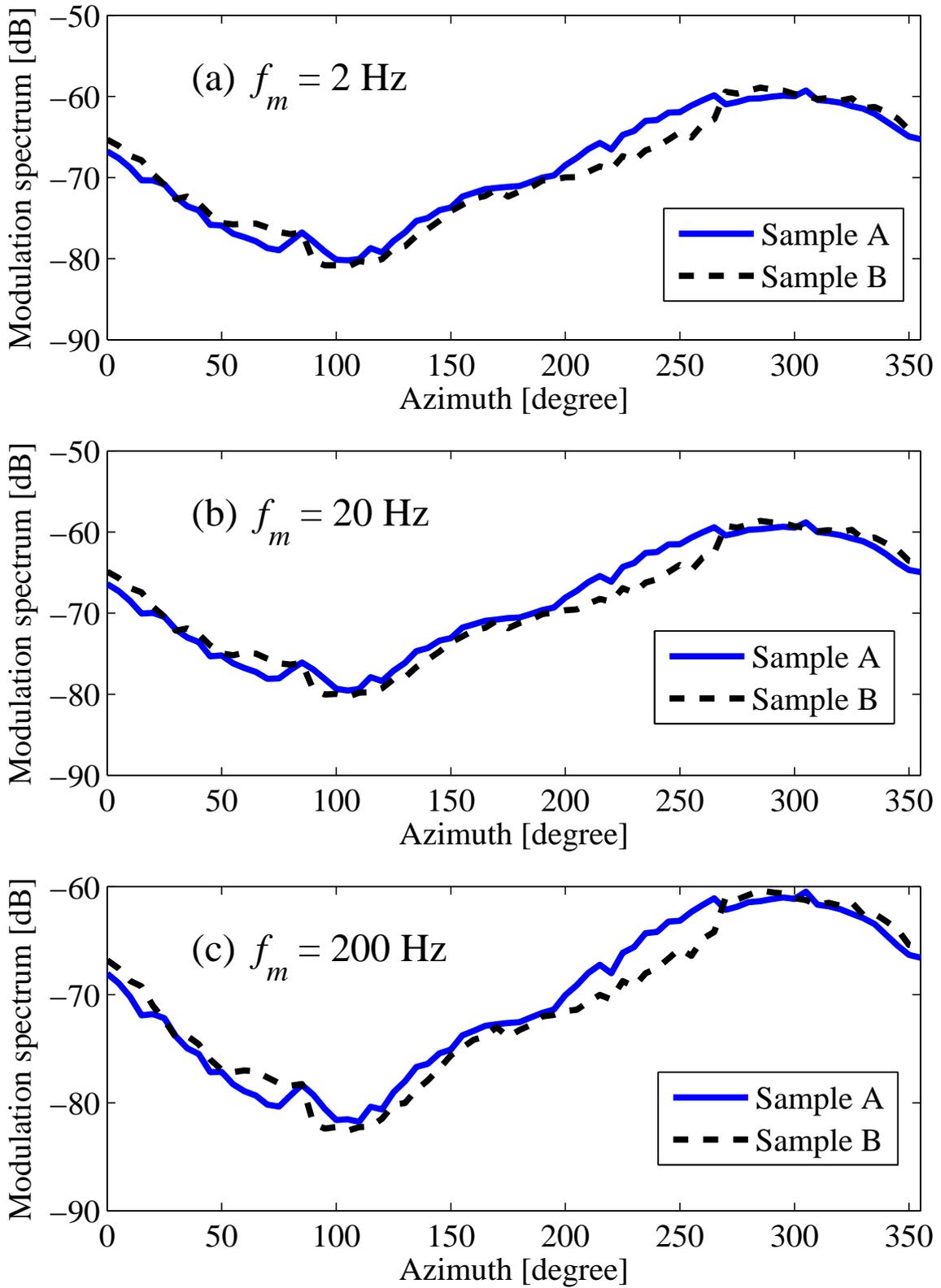


図 3.4: MMS の変化: 水平面の場合, AM 雑音

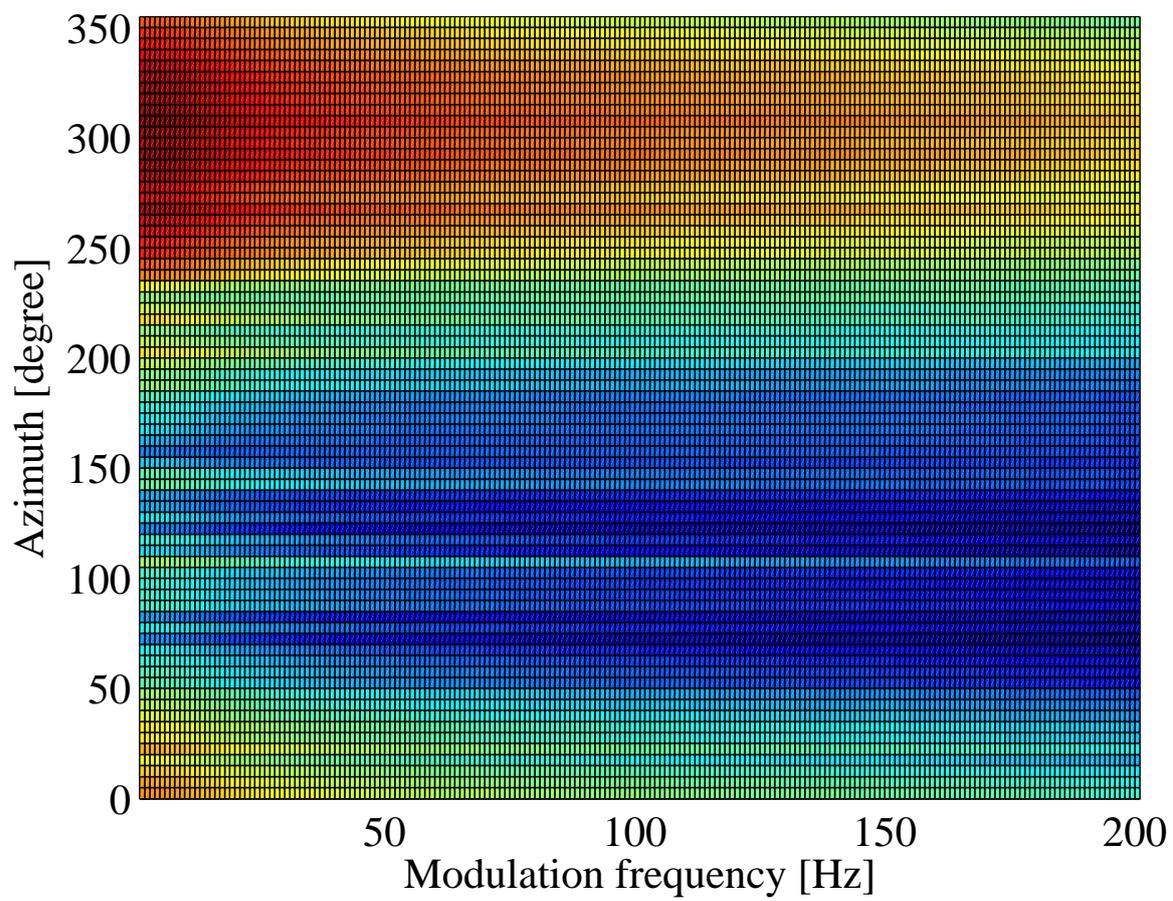


図 3.5: 変調周波数ごとの MMS の変化: 搬送波 AM 信号, 変調周波数 2 Hz

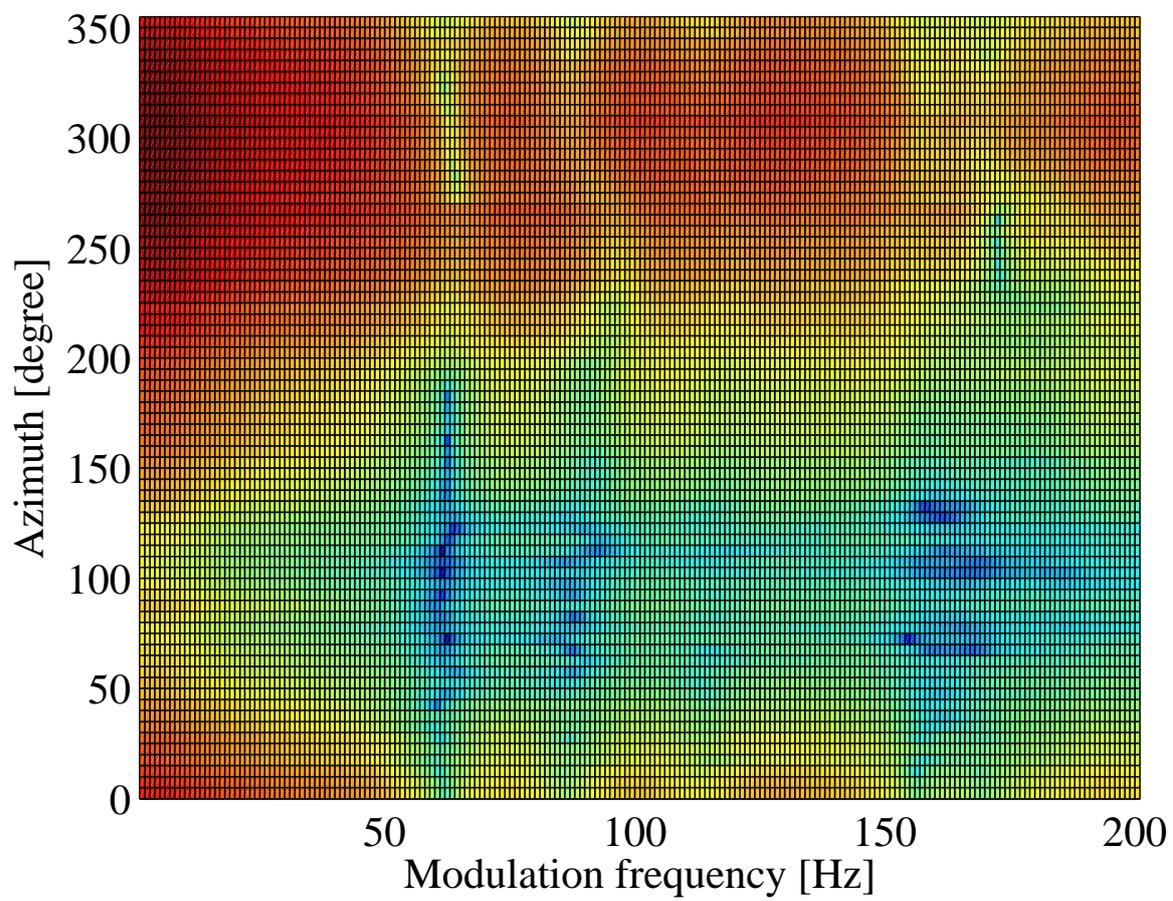


図 3.6: 変調周波数ごとの MMS の変化: 搬送波 AM 雑音, 変調周波数 2 Hz

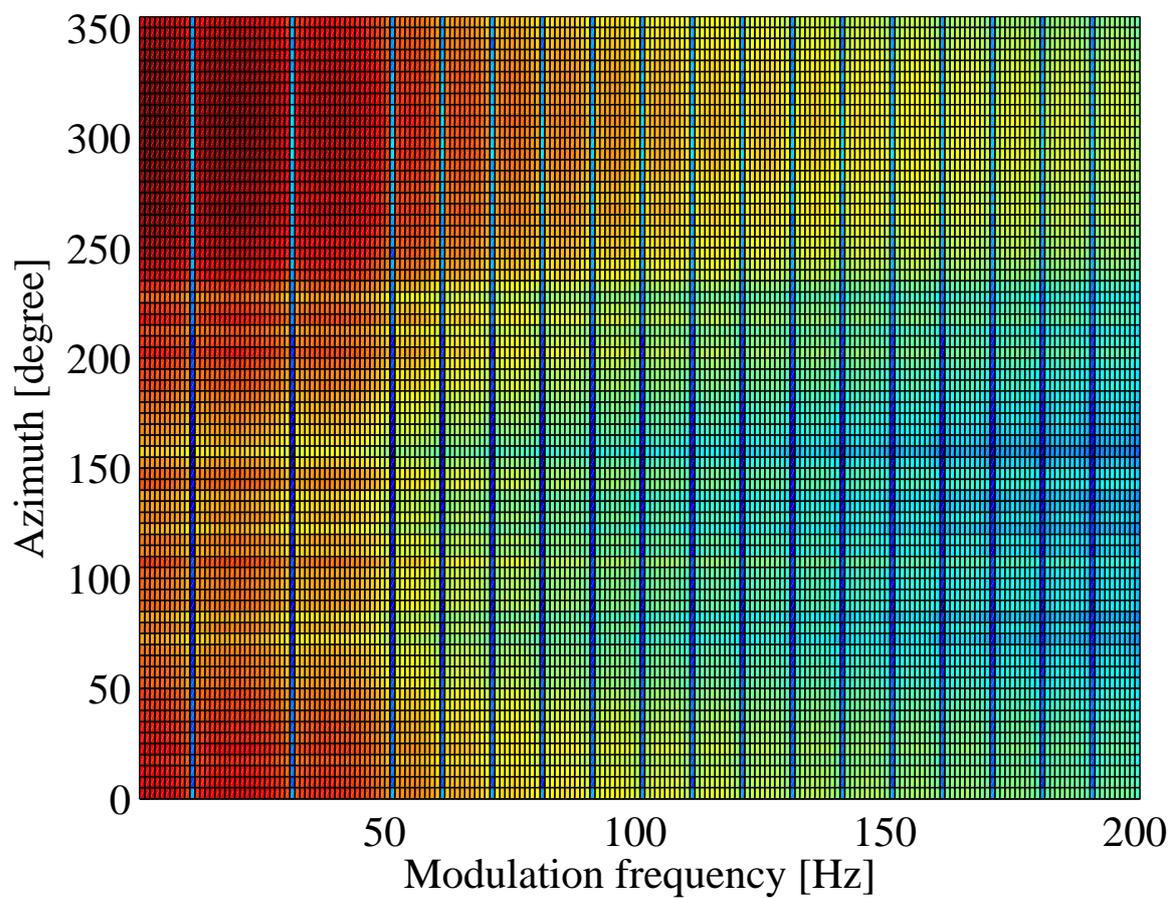


図 3.7: 変調周波数ごとの MMS の変化: 搬送波 AM 信号, 変調周波数 20 Hz

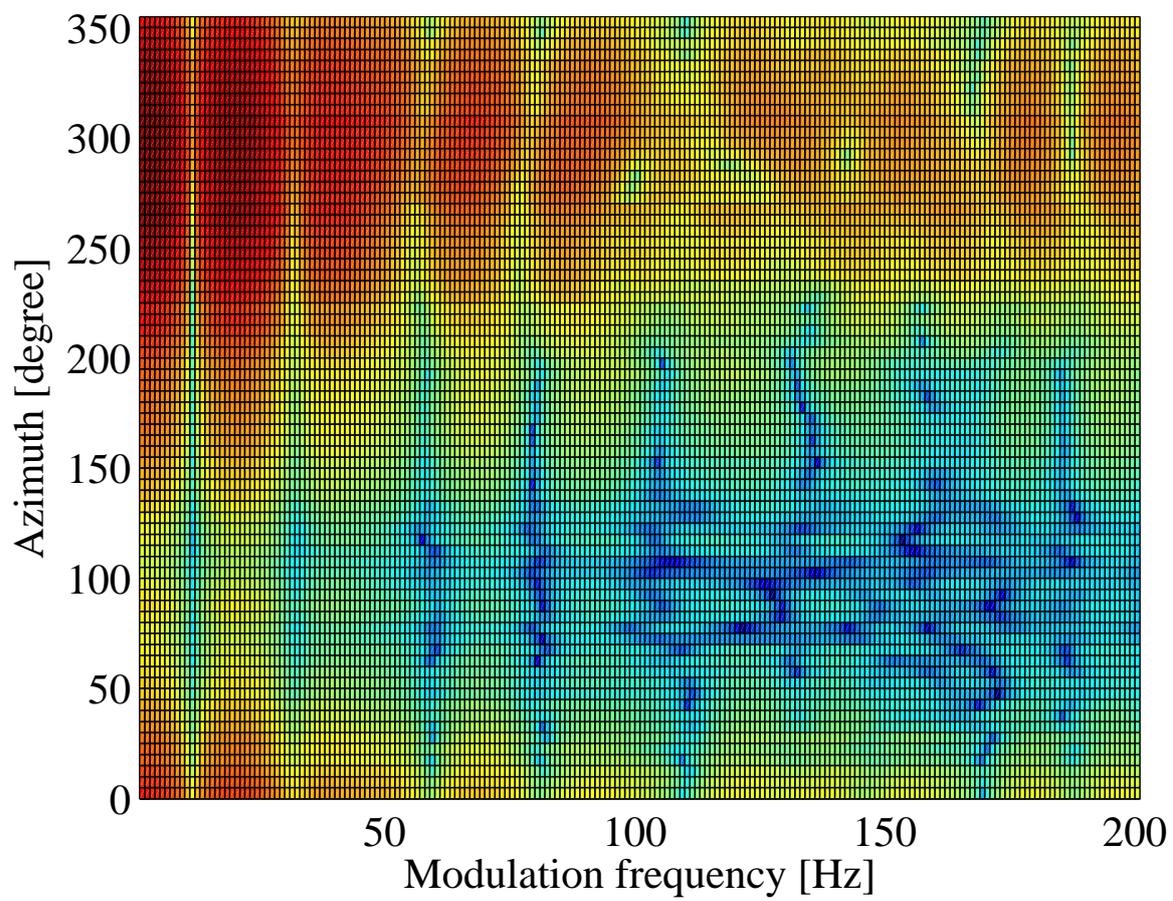


図 3.8: 変調周波数ごとの MMS の変化: 搬送波 AM 雑音, 変調周波数 20 Hz

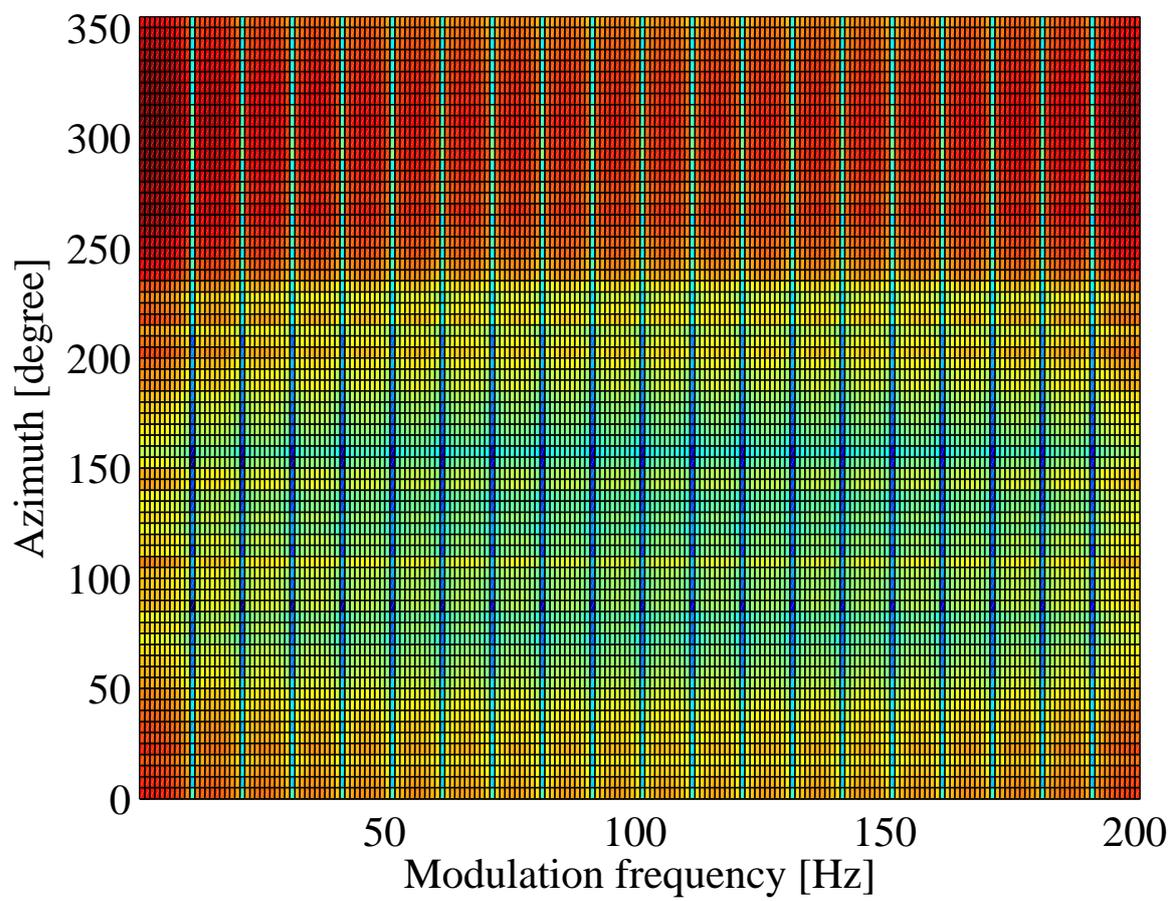


図 3.9: 変調周波数ごとの MMS の変化: 搬送波 AM 信号, 変調周波数 200 Hz

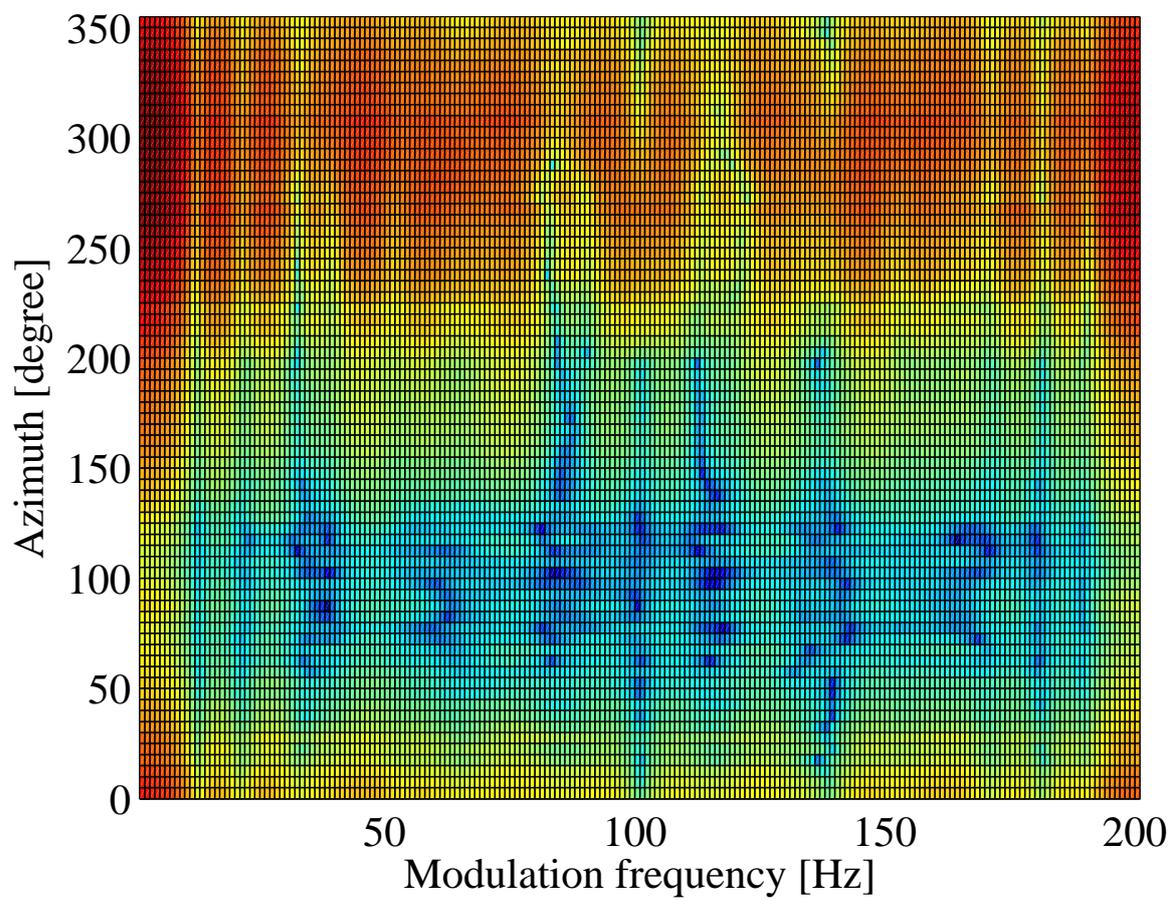


図 3.10: 変調周波数ごとの MMS の変化: 搬送波 AM 雑音, 変調周波数 200 Hz

3.2.3 調査結果: 水平面 (低域成分に限定したもの)

図 3.11 から図 3.12 の (a) から (c) より, 搬送波が変化した場合においても前節の結果と同様に, AM 信号と AM 雑音共に水平角によって正弦波のような形状になった。さらに, 右耳がある位置近傍 (90 度前後) では, 低い搬送波周波数の場合において, MMS が盛り上がっている部分が見られた。この特徴は, 搬送波周波数が上昇するにつれて確認されなくなった。右耳の HRIR や, 他者の HRIR を用いた場合でもこの傾向は類似していた。これは, 図 3.13 のように, 1500 Hz 以下の波長が頭部より長いことによって生じる回折の影響であると考えられる。

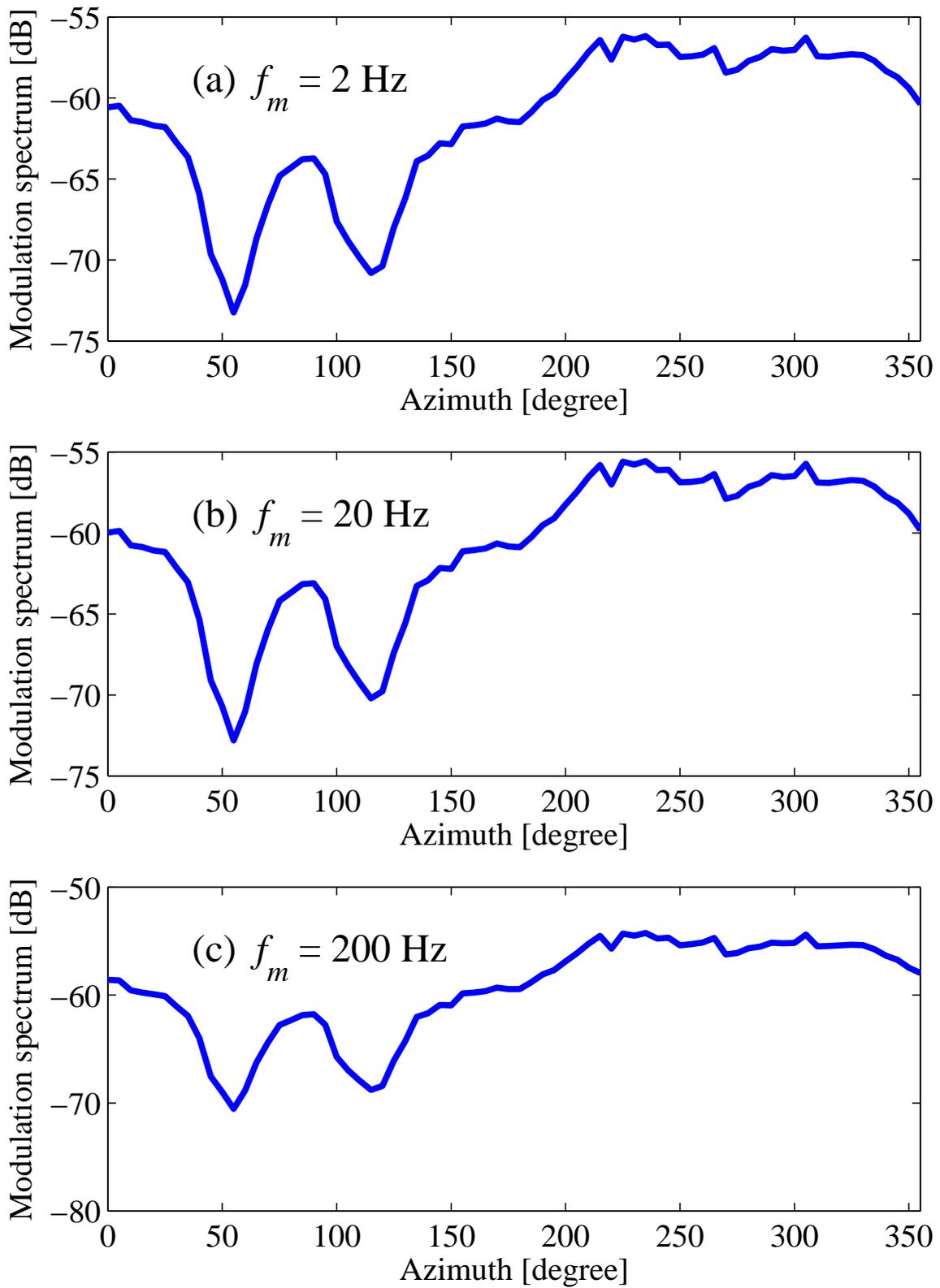


図 3.11: MMS の変化: 水平面の場合 (低域成分, 回折の影響), AM 信号

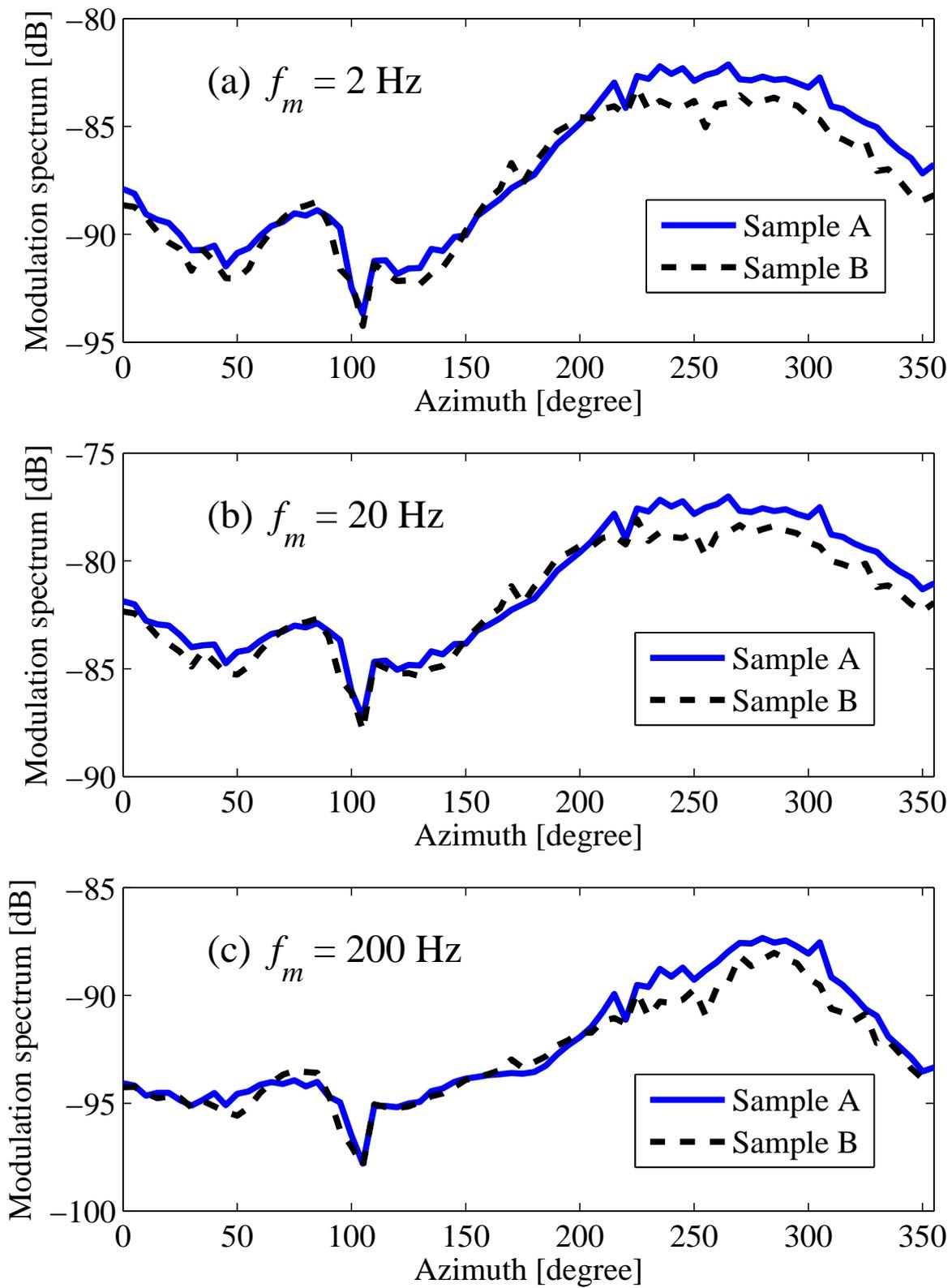


図 3.12: MMS の変化: 水平面の場合 (低域成分, 回折の影響), AM 雑音

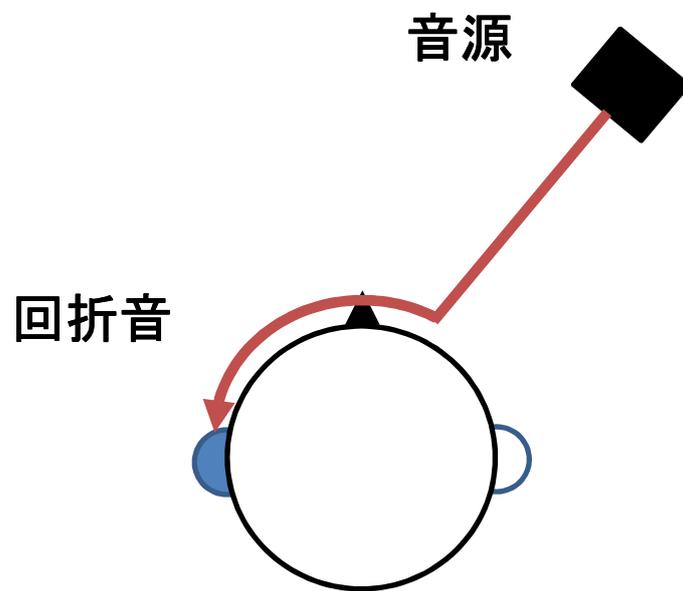


図 3.13: 頭部回折の影響の模式図

3.2.4 調査結果: 水平面 (肩部の影響)

図 3.14 に左耳の場合の結果を示す。搬送波周波数 1000 Hz, 変調周波数が 2, 20, 200 Hz の結果を示す。図より, 30 度および 150 度付近に 2 dB 程度の差がみられた。これらの位置は, 左耳とは反対側であるが, 音源と肩までの経路に頭部がない。そのため, 図 3.15 のように左肩を反射して耳介に到達した成分であると考えられる。一方で, 左耳と同側では肩の反射とみられる影響がほとんど見られない。これは, 直接音や回折音が反射音よりも支配的であるため, 反射音の影響が抑えられているためであると考えられる。

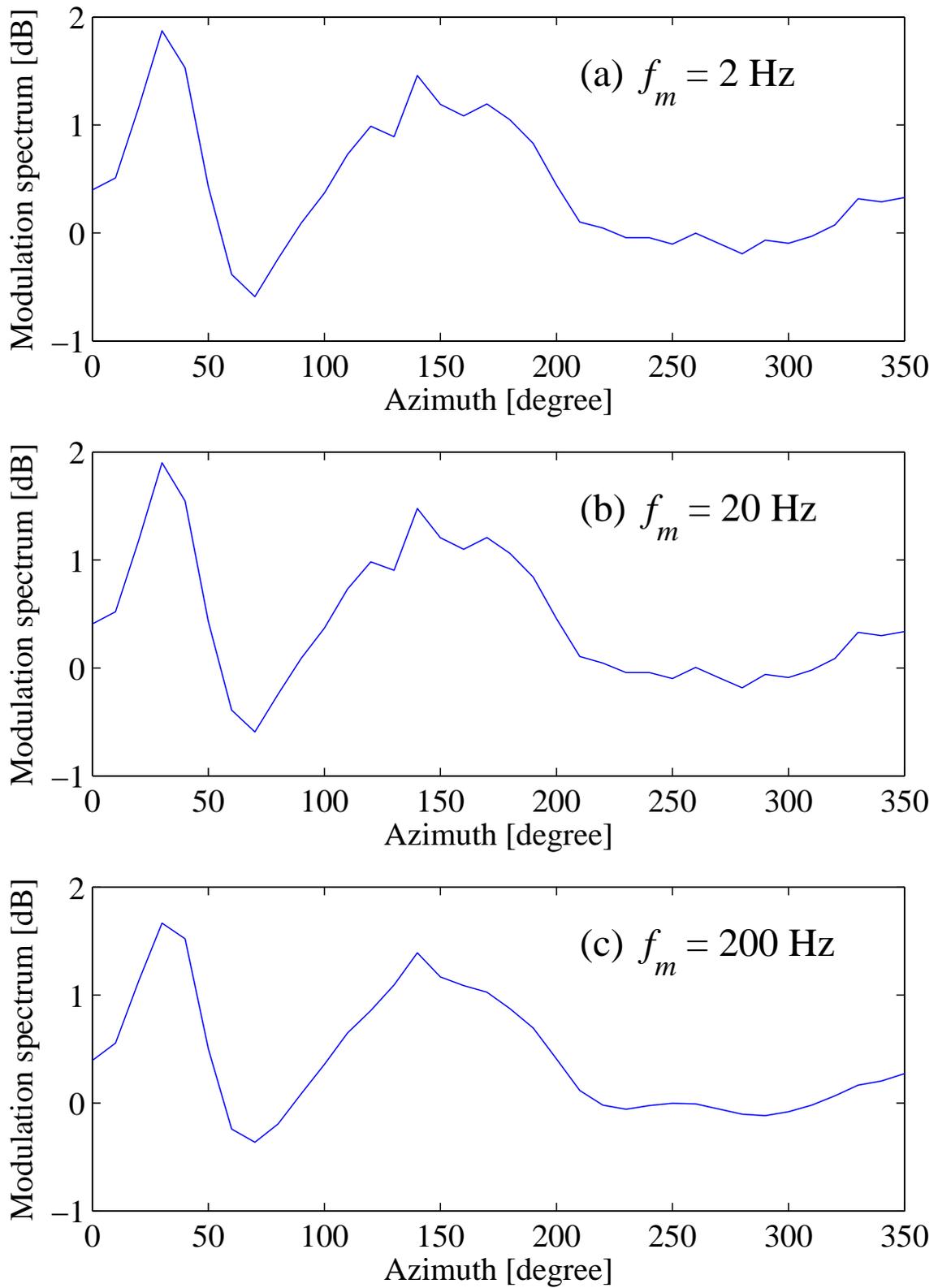


図 3.14: MMS の変化: 水平面の場合 (肩部有無の差, 反射の影響)

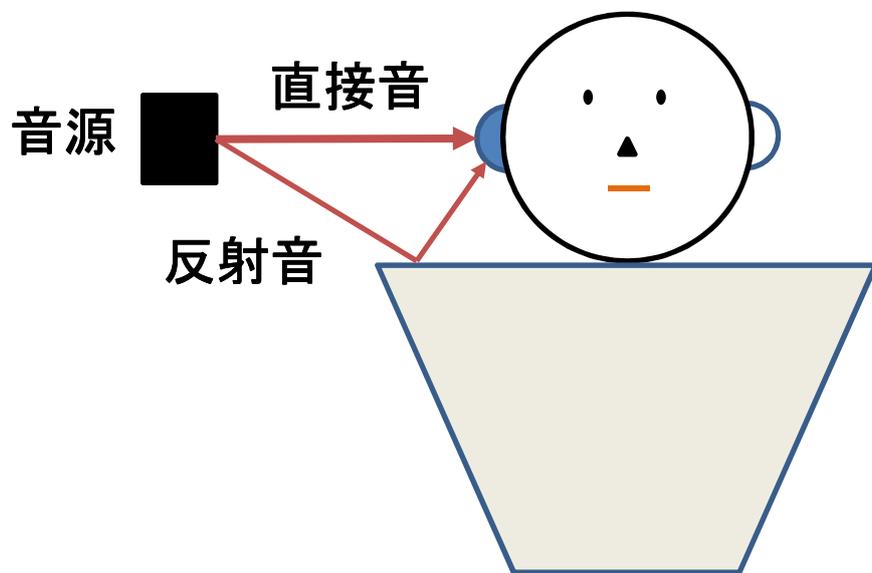


図 3.15: 肩の反射の影響の模式図

3.2.5 調査結果: 正中面

図 3.16 から図 3.17 から，正中面の正面側から頭頂部では，サンプルによって MMS の角度変化の形状が異なること，頭頂部から後頭部にかけて MMS は減衰する傾向であることがわかった．右耳の HRIR や，他者の HRIR を用いた場合でもこの傾向は類似していたが，正面部から頭頂部でサンプルによる違いが大きくなったのは，HRIR の個人性に依存するものだと考えられる．

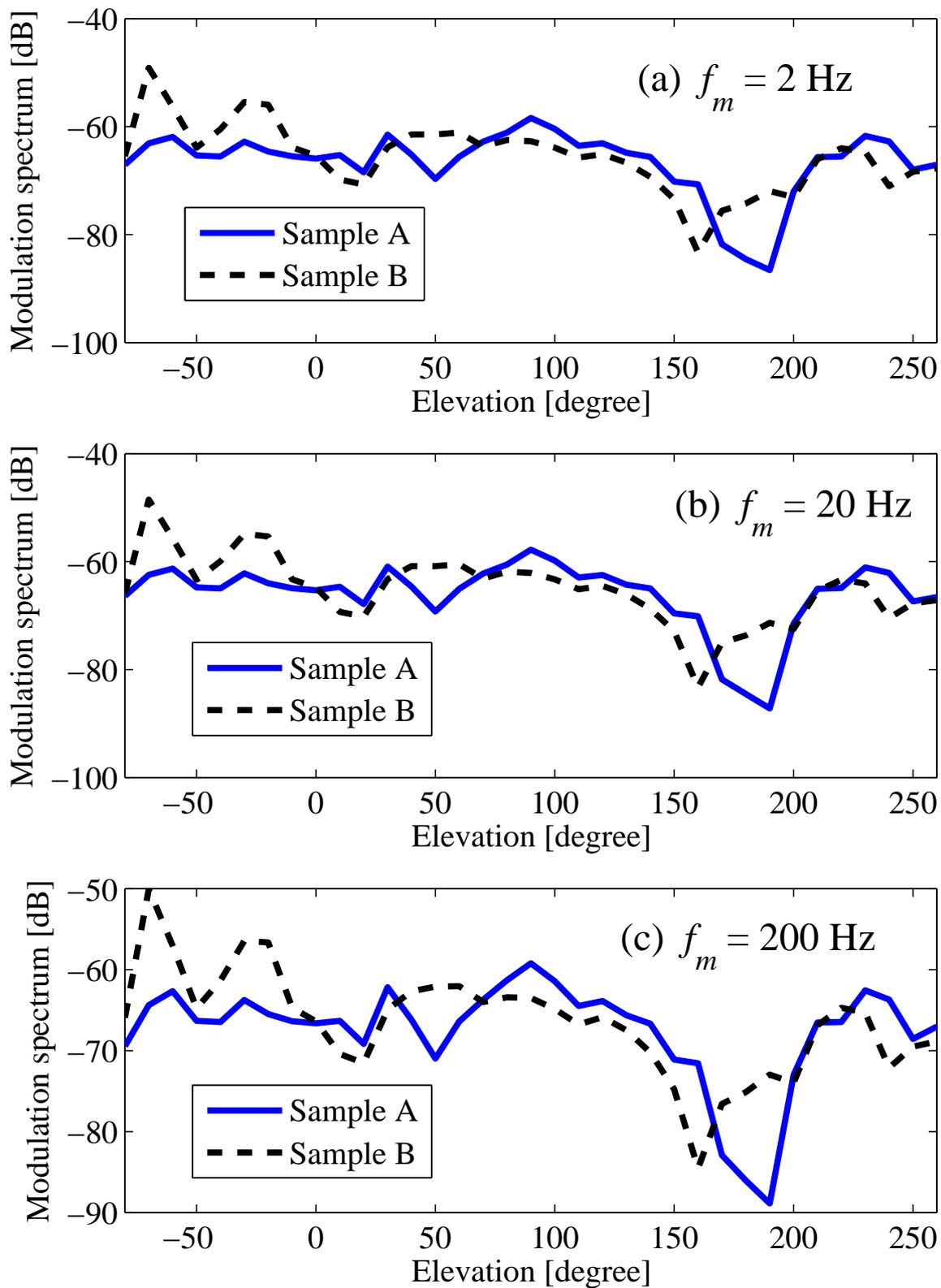


図 3.16: MMS の変化: 正中面の場合, AM 信号

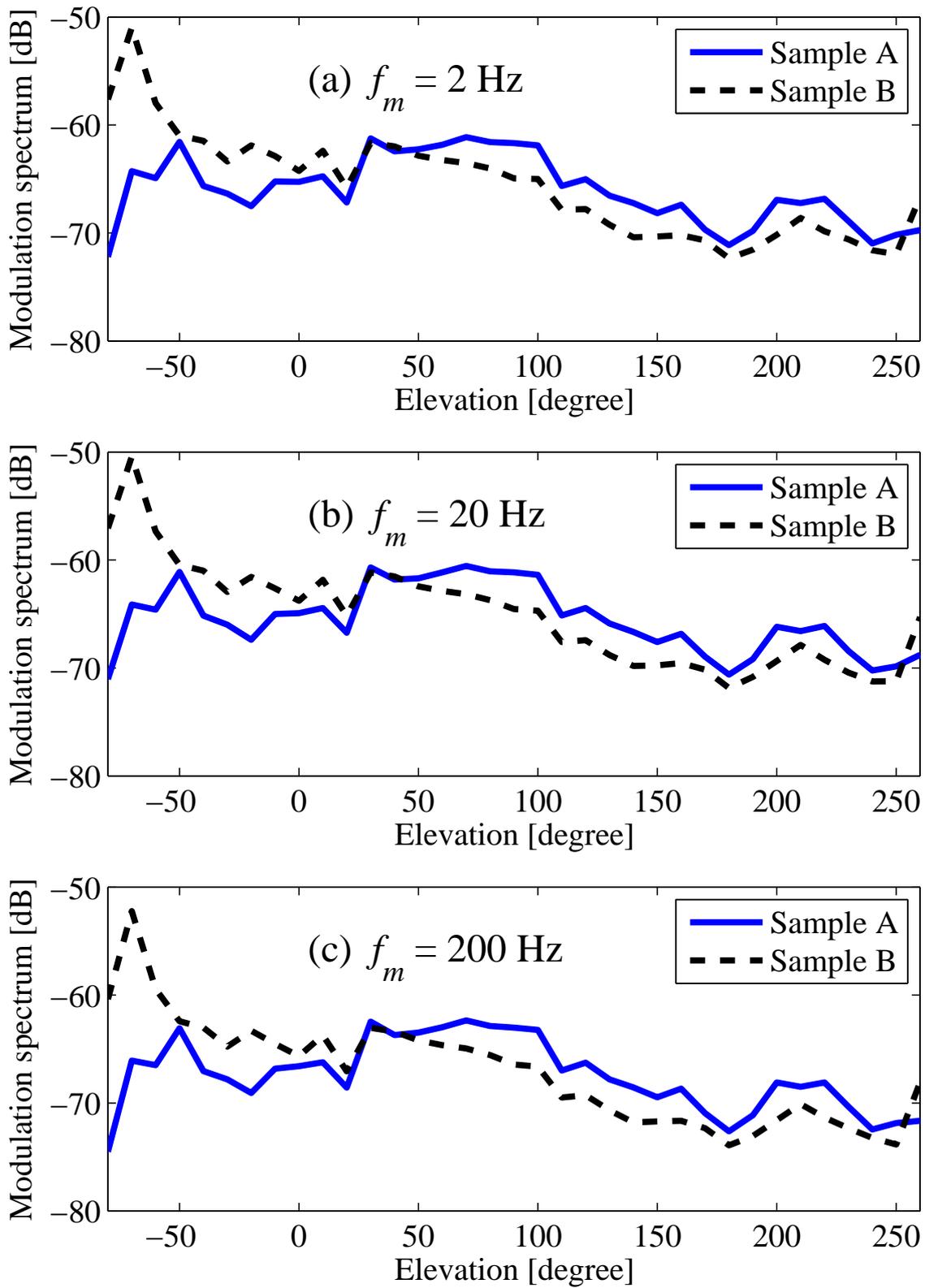


図 3.17: MMS の変化: 正中面の場合, AM 雑音

3.2.6 調査結果: 横断面

図 3.18 から図 3.19 から，横断面では右耳側から頭頂部にかけて MMS が上昇すること，頭頂部から左耳側では MMS が一定になることがわかった．右耳の HRIR や，他者の HRIR を用いた場合でもこの傾向は類似していた．右耳側から頭頂部にかけて MMS が上昇したのは，頭部の遮蔽によるものであると考えられる．また，頭頂部から左耳側で MMS が一定になったのは，遮蔽物がなく音源信号がそのまま鼓膜位置に届くためであると考えられる．

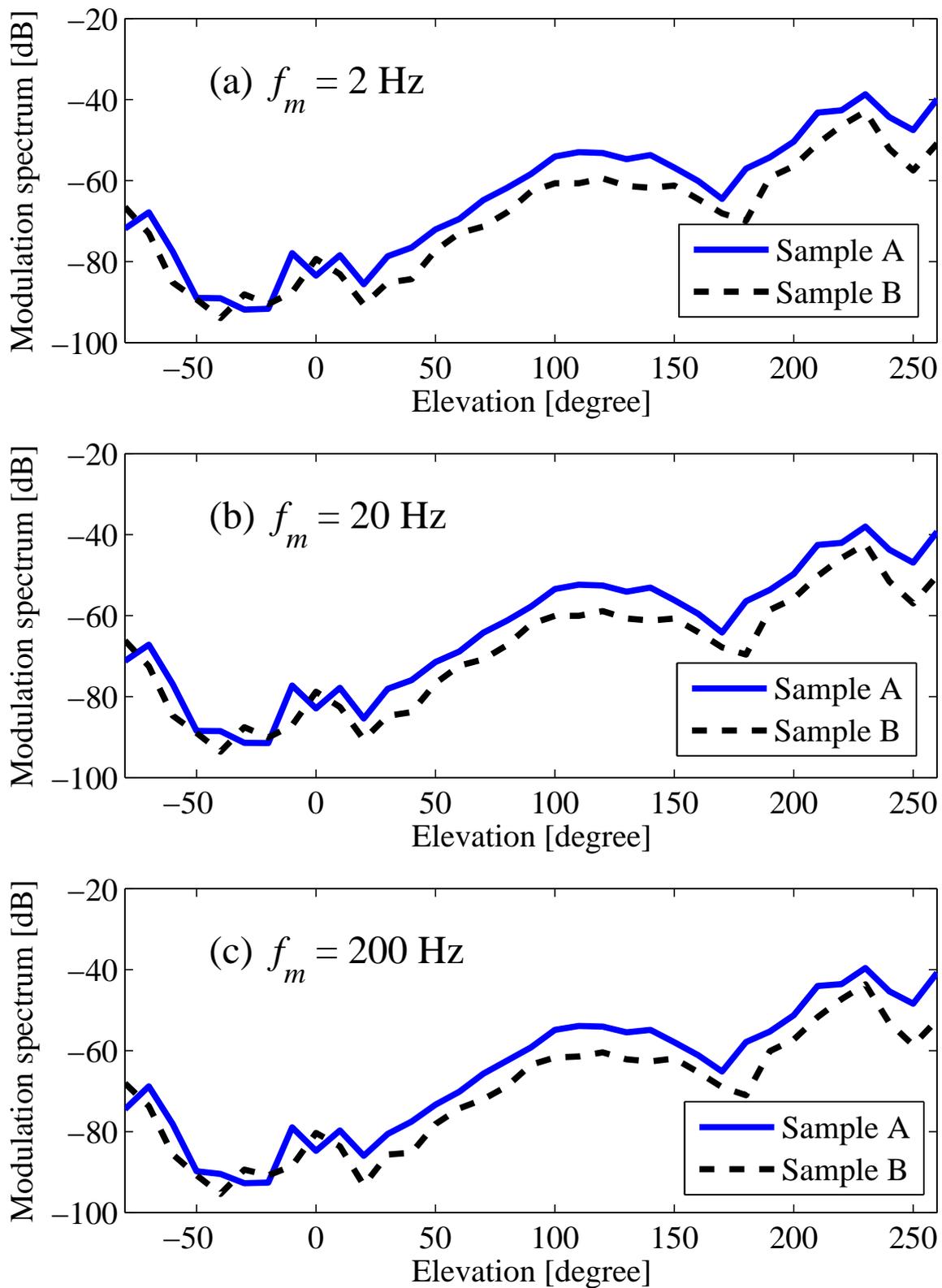


図 3.18: MMS の変化: 横断面の場合, AM 信号

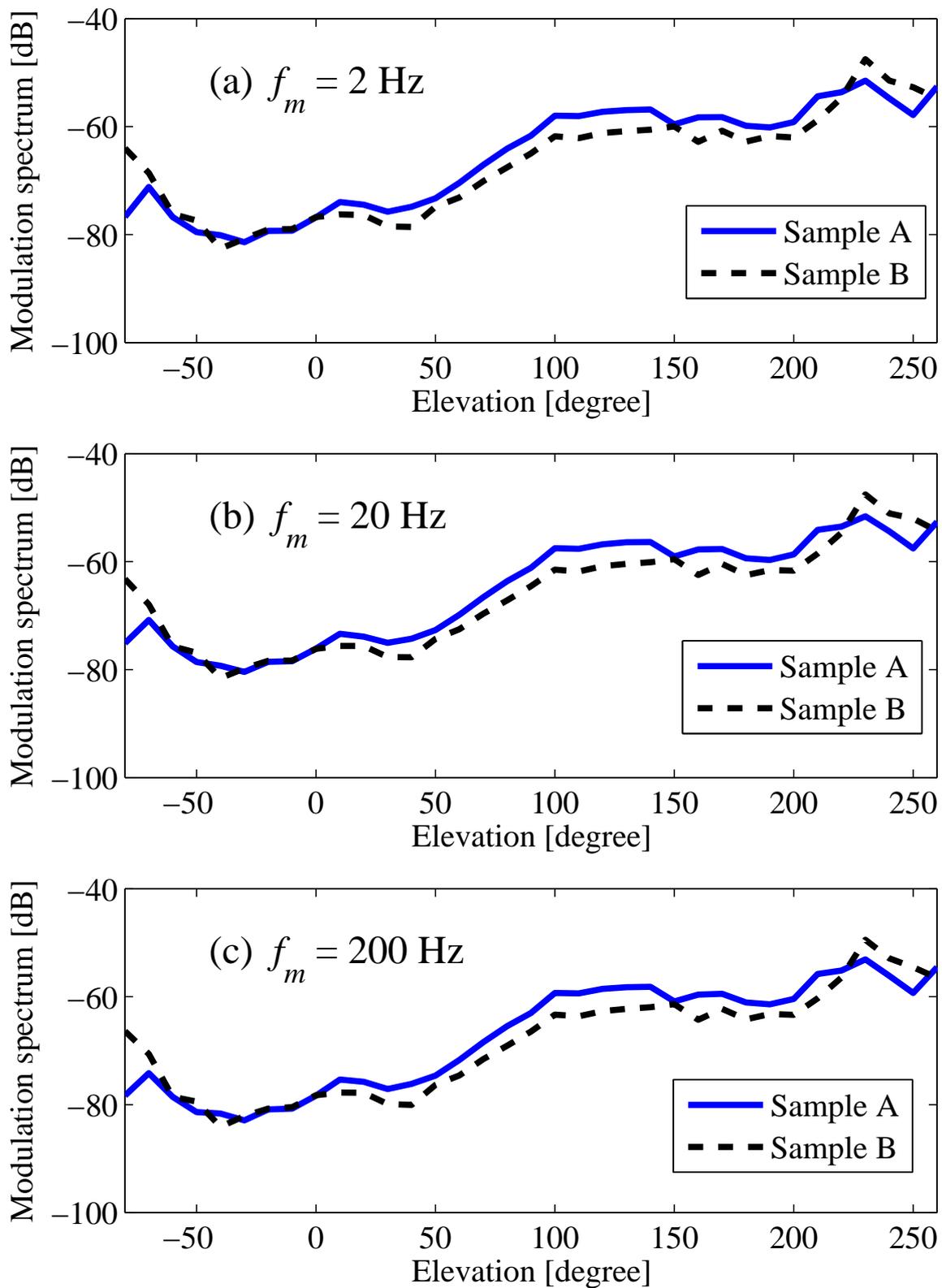


図 3.19: MMS の変化: 横断面の場合, AM 雑音

3.3 考察

水平面の調査結果より，水平面では MMS は角度に伴って正弦波のように変化した．この結果は，ITD および ILD の形状と類似しており，頭部の遮蔽による影響であると考えられる．また，変調周波数が変化しても，ダイナミックレンジはあまり変化しないことが分かった．これらのことから，単耳がある位置と同じ側から到来した音源信号に対しては，ヒトが音源方向知覚の手がかりを得ていると理解できる．また，音源信号の変調周波数が変化しても，低域には上記と同じ傾向が得られた．そのため変調周波数帯域に依存せずに，ヒトが音源方向知覚の手がかりを得ている可能性も示唆された．また，AM 信号よりも AM 雑音では弧の形状がなめらかであることから，搬送波に低域成分を多く含んでいる白色雑音は，より頭部や耳介を回折していると考えられる．

水平面の回折特性の調査結果より，低域通過雑音を用いた場合，図 3.4(a) と異なり，右耳がある位置（90 度）付近の MMS が盛り上がっていることが分かった．これは，1500 Hz 以下の波長が頭部より長いことによって生じる回折の影響であると考えられる．

水平面の反射特性の調査結果より，単耳と反対側から音源信号が到来した場合において肩の反射の影響が見られたことから，音源信号の搬送波周波数によっては回折音と同様に反射音の影響が多いと予想される．ただし，直接音や回折音のほうが反射音よりも支配的であるため，反射音の影響は低いと考えられる．一方で，搬送波周波数が高い信号だと，鼓膜ではほとんど直接音しか観測しないということが考えられる．

正中面の調査結果より，正面側から頭頂部では，サンプルによって MMS の角度変化の形状が異なること，頭頂部から後頭部にかけて MMS は減衰する傾向であることがわかった．右耳の HRIR や，他者の HRIR を用いた場合でもこの傾向は類似していたが，正面部から頭頂部でサンプルによる違いが大きくなったのは，HRIR の個人性に依存するものだと考えられる．

横断面の調査結果より，右耳側から頭頂部にかけて MMS が上昇すること，頭頂部から左耳側では MMS が一定になることがわかった．右耳の HRIR や，他者の HRIR を用いた場合でもこの傾向は類似していた．右耳側から頭頂部にかけて MMS が上昇したのは，頭部の遮蔽によるものであると考えられる．また，頭頂部から左耳側で MMS が一定になったのは，遮蔽物がなく音源信号がそのまま鼓膜位置に届くためであると考えられる．

以上の議論から，MMS の特徴のうち最も容易に音源方向推定に利用可能な特徴は，水平面の角度変化に関する MMS の変化であると考えられる．また，この角度変化は，ITD および ILD の角度変化と類似しており，ヒトが音源方向を知覚する手がかりとしても十分利用できると思われる．さらに，音源信号の変調周波数帯に依らず，MMS を知覚している可能性も示唆された．

3.4 まとめ

本章では，モノラル音源方向知覚の手がかりを明らかにするために水平角および仰角の角度変化特性の調査を行った．その結果，水平角では，角度変化に伴うMMSの変化は耳位置を頂点とした上向きの弧形（下向きの孤形）となる傾向がみられた．また，頭部回折の影響や肩部および頭部反射の影響もみられた．仰角では，頭頂部付近から後頭部にかけて，MMSが減衰する傾向が見られた．また，頭頂部付近において，音源に対して対側から同側にかけてMMSが増加する傾向がみられた．これらのことから，角度変化に伴うMMSの変化は，モノラル音源方向推定の手がかりに十分利用可能であることが示唆された．

第4章 変調スペクトルに着目したモノラル音源方向推定法

4.1 水平角に着目した音源方向推定法

前節の分析結果を利用した音源方向推定モデルを検討する。図 3.11 および図 3.12 のように到来した音源信号に低域成分が多く含まれている場合は、単耳の反側に頭部回折の影響がみられた。頭部回折によって音源が到来していない側では、MMS の変化がスムーズでない。そこで、全帯域成分を含む観測信号の MMS と、観測信号から LPF などで低域成分のみ（本稿では、1.5 kHz 以下）を抽出した信号の MMS を比較し、その差を利用することで、左右の弁別が可能であると考えられる。

左右を弁別した上で、次の方法を適用すれば、水平角半面のモノラル音源方向推定が可能となる。図 3.3 および図 3.4 から、水平角が変化した場合、AM 雑音を用いた観測信号の MMS には耳位置を頂点とした弧形となる傾向がみられた。そのため、水平角の変化特性を回帰近似を利用して学習することで、音源方向推定が可能であると考えられる。ただし、回帰近似を利用することから推定角度が一意に定まらない問題が発生してしまう。この問題は、頭部運動を想定し、MMS の時間変化を利用すれば改善可能であると考えられる。

以上の水平角に着目した音源現方向推定モデルを図 4.1 に示す。(a) では、MMS を回帰近似し、このとき得られた回帰係数を学習し、データベース化する。(b) では、観測信号の MMS の変調周波数に対応する回帰係数を呼び出し、回帰多項式の逆関数に MMS および回帰多項式を代入することで角度を推定する。

4.2 仰角に着目した音源方向推定法

図 3.16 および図 3.17 から、頭頂部付近から後頭部にかけて、MMS がなめらかに減衰する傾向が見られた。そのため、この傾向を利用すれば後頭部付近の音源方向推定が可能であると考えられる。

図 3.18 および図 3.19 から、頭頂部付近において、右耳側から左耳側にかけて MMS が一次関数的に増加する傾向がみられた。そのため、この傾向を利用すれば頭頂部付近の音源方向推定が可能であると考えられる。

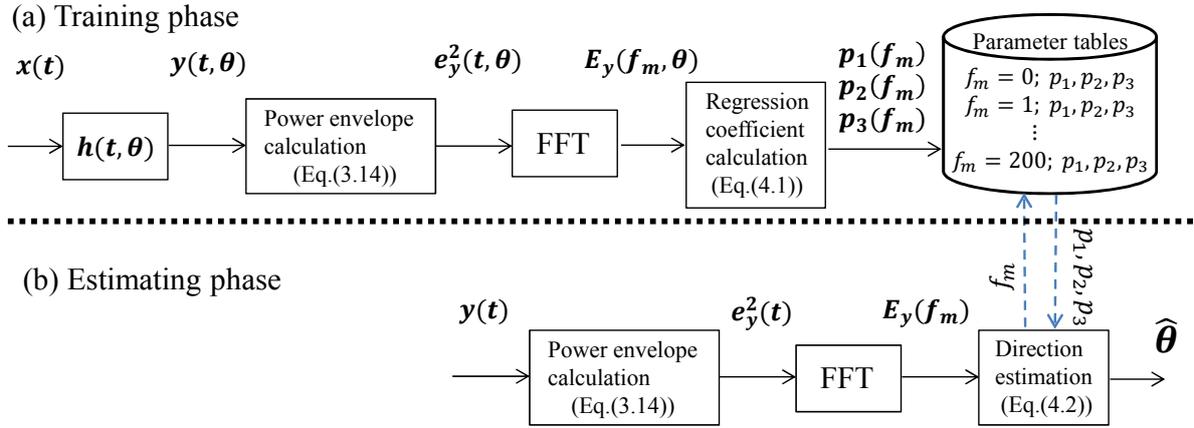


図 4.1: 提案法: (a) 事前学習 (b) 方向推定

仰角においても，図 4.1 と同様に，図 3.18 および図 3.19 を回帰近似することで，音源方向推定が可能であると考えられる．

4.3 水平角に着目した音源方向推定の実現例

4.4 提案法

検討した音源方向推定モデルのうち，水平角に着目した音源方向推定モデルを，計算機シミュレーションによって実現可能性を確認する．シミュレーションにさきがけて，次の理想条件を設ける．1. MMS の角度変化特性を得るために，マイクロフォンに頭部模型や耳型のアタッチメントを装着すると仮定する．2. 音源信号を AM 雑音とし，その変調周波数も既知とする．3. 学習に利用した HRIR と評価シミュレーションに利用する HRIR を同じものとする．4. 単耳と同側（右耳では水平角 0 度から 180 度，左耳では水平角 180 度から 355 度まで）を対象範囲とする．

図 4.1 に MMS の規則性を利用した音源方向推定法（提案法）の処理フローを示す．提案法は (a) 事前学習と (b) 方向推定の 2 段階で構成される．まず事前学習の段階を解説する．MMS は弧状の規則性ではあるが，実測値であることから値にバラつきが生じている．そのために，まず MMS を式 (4.1) の回帰多項式を用いて（図 4.2 と図 4.3 の○印）を音源方向について，最小二乗誤差で近似する．

$$\hat{E}_y(f_m, \theta) = p_1(f_m)\theta^2 + p_2(f_m)\theta + p_3(f_m) \quad (4.1)$$

ただし， $p_1(f_m)$ ， $p_2(f_m)$ および $p_3(f_m)$ は回帰係数， $\hat{E}_y(\theta)$ は変調スペクトルの近似値である．図 4.2 と図 4.3 に回帰近似によって得られた変調スペクトルの値の回帰曲線を実線で示す．ここで， f_m および単耳と同側の HRIR が既知であれば，回帰係数 $p_1(f_m)$ ， $p_2(f_m)$

および $p_3(f_m)$ が算出される．算出された回帰係数を f_m ごとにデータベース化する．これで，事前学習の段階は完了である．

次に，方向推定の段階を解説する．未知の角度 θ から到来した音源信号がマイクで観測された場合，提案法には観測信号 $y(t)$ が入力される．ここで，観測信号の f_m は既知であると仮定する．次に，式 (3.14) より $y(t)$ からパワーエンベロープが抽出され，FFT により観測信号の MMS $E_y(f_m)$ が導出される．ここで， $E_y(f_m)$ の f_m に対応する回帰係数 $p_1(f_m)$ ， $p_2(f_m)$ および $p_3(f_m)$ を事前学したデータベースから参照し， $E_y(f_m)$ および回帰係数を式 (4.2) に代入することで音源方向が推定される．

$$\hat{\theta}(f_m, E_y) = \frac{-p_2(f_m) \pm \sqrt{p_2^2(f_m) - 4p_1(f_m)(p_3(f_m) - E_y)}}{2p_1(f_m)} \quad (4.2)$$

ここで，式 (4.2) は式 (4.1) の逆関数であり， $\hat{\theta}$ は推定角度である．以上で，方向推定が完了する．

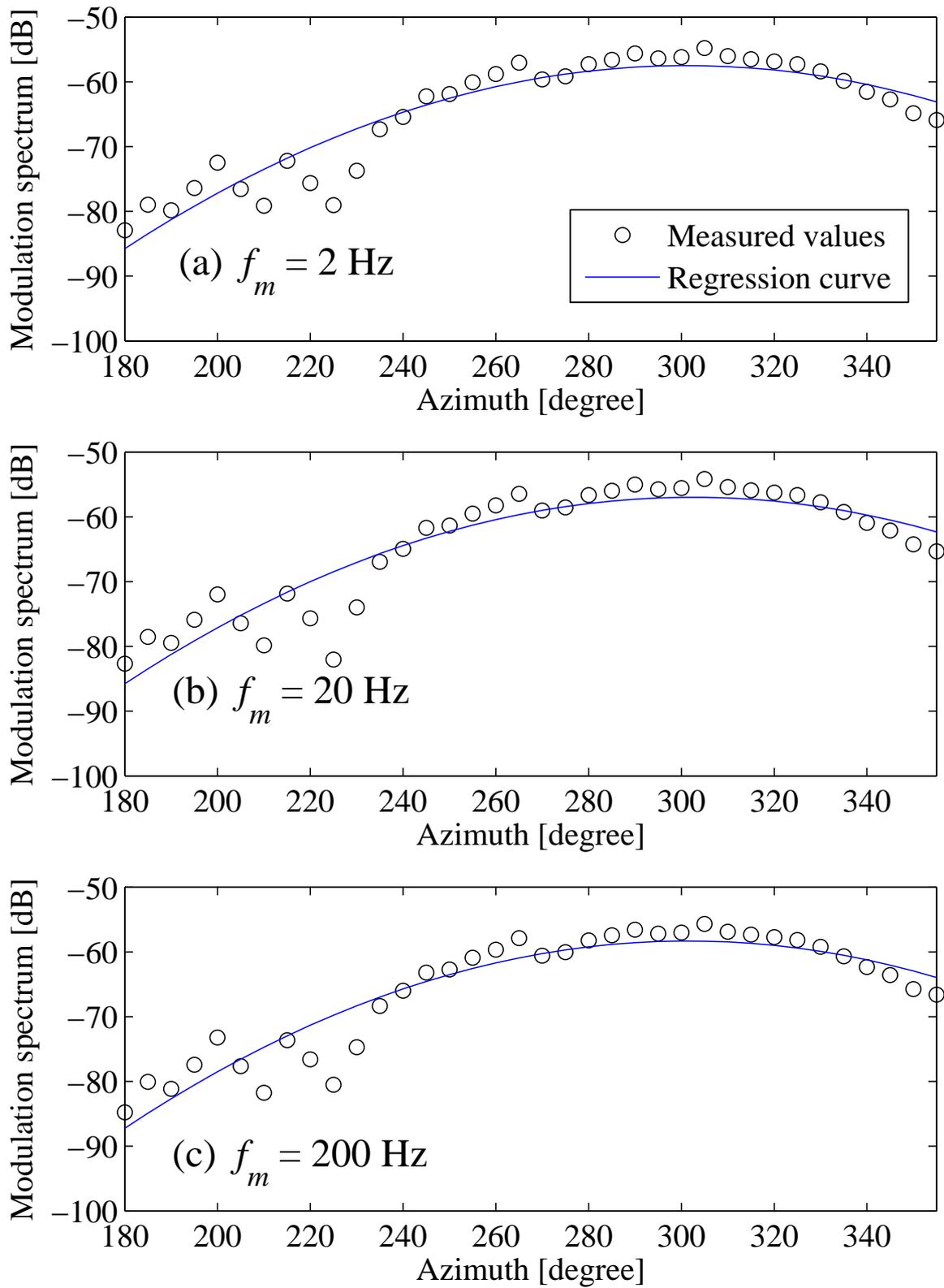


图 4.2: 回歸曲線 AM 信号

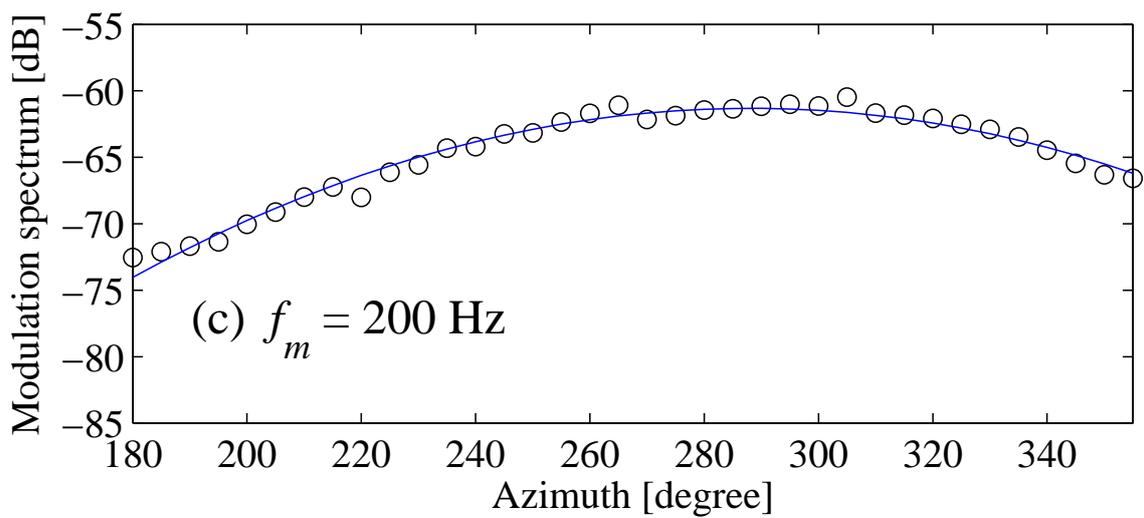
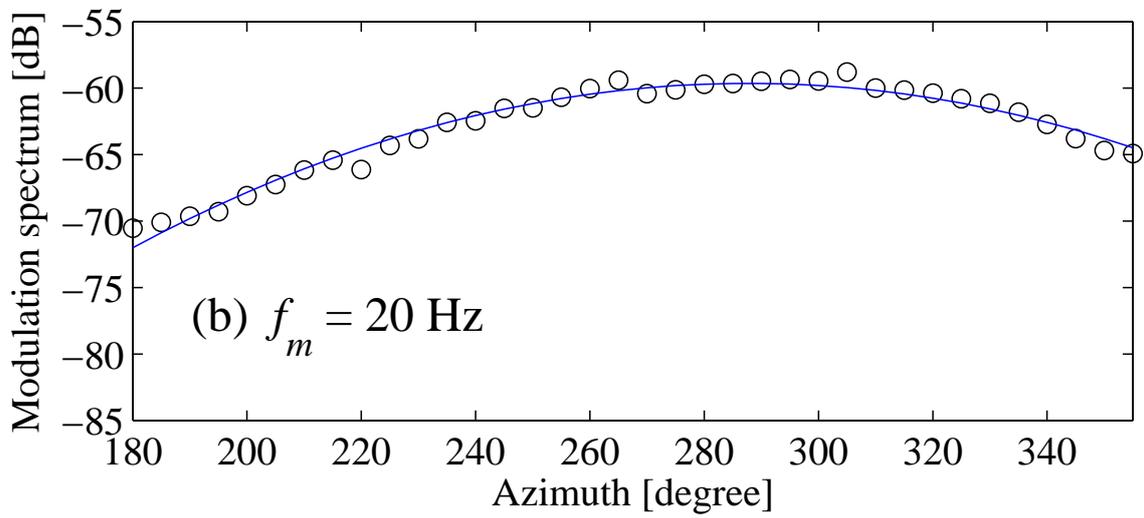
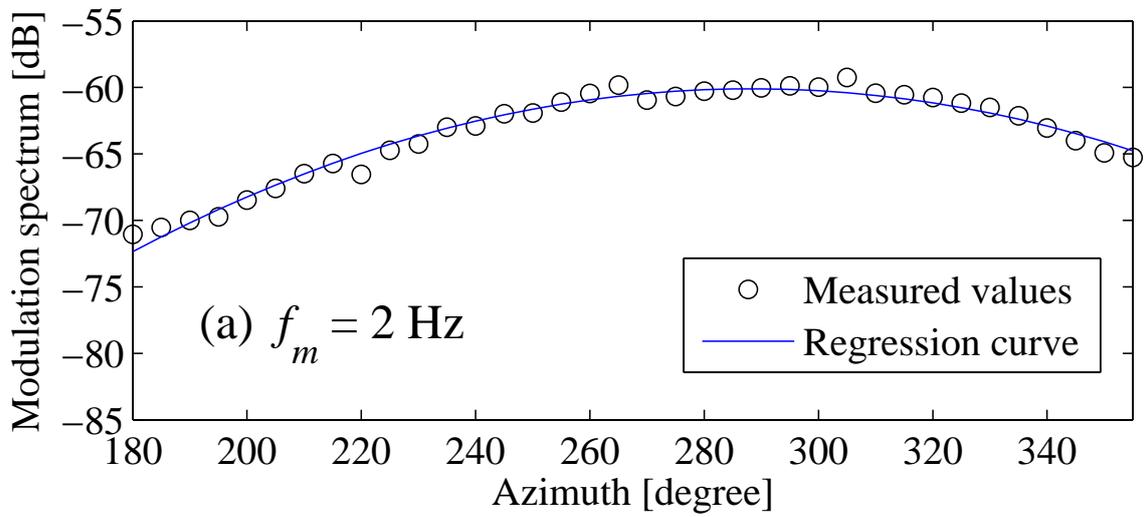


図 4.3: 回帰曲線 AM 雑音

4.5 シミュレーションによる提案法の検証

4.5.1 検証方法および条件

前節で提案した音源方向推定法の有効性を確認するため，シミュレーションによる検証を行った．まず，事前学習段階として，単耳の観測信号を式(3.5)に基づき，音源信号と既知である HRIR を用いて算出する．算出した観測信号から式(3.14)によってパワーエンベロープを得て，それを FFT することで MMS を求める．ここで，HRIR の角度を 0 度から 355 度まで変化させる．そして，式(4.1)によって回帰近似し，各回帰係数をデータベースとして保持する．次に，方向推定段階として，未知の角度 θ から到来した音源信号を単耳で受聴したと仮定し，その観測信号から式(3.14)によってパワーエンベロープを算出する．それを FFT し MMS を求める．ここで，到来した音源信号の変調周波数は 2, 20, 200 Hz であると仮定し，音源信号と同じ変調周波数の回帰係数と求めた MMS を式(4.2)に代入することで，角度を推定する．

3.2 節で利用したものと同じ AM 信号および AM 雑音を音源信号と仮定し，提案法を適用する．また，ここで利用する HRIR は 3.2 節で利用したものと同様であり，56 名分をシミュレーションに用いた．シミュレーションの範囲は，単耳と同側である．

4.5.2 検証結果

代表的な一人分の左耳の検証結果を図 4.4 から図 4.6 に示す．ここで，横軸は音源信号方向を，縦軸は推定角度を表す．AM 信号と AM 雑音共に，式(4.2)の \pm 項の影響によって推定角度が一意に定まっていない．+ の場合は，誤差があるものの耳位置付近より後方の推定ができています．一方 + の場合とは逆に，- の場合は，耳位置付近よりも前方の推定ができています．ただし，+ の場合は，耳位置付近より前方で，- の場合は，耳位置付近より後方で前後の推定誤りが生じています．また，変調周波数の変化によって，推定誤差はあまり変わらないことがわかる．また，AM 雑音のほうが AM 信号よりも誤推定が少ないことが分かる．以上の結果は，右耳の場合および他の 55 名分も同様の傾向を示した．

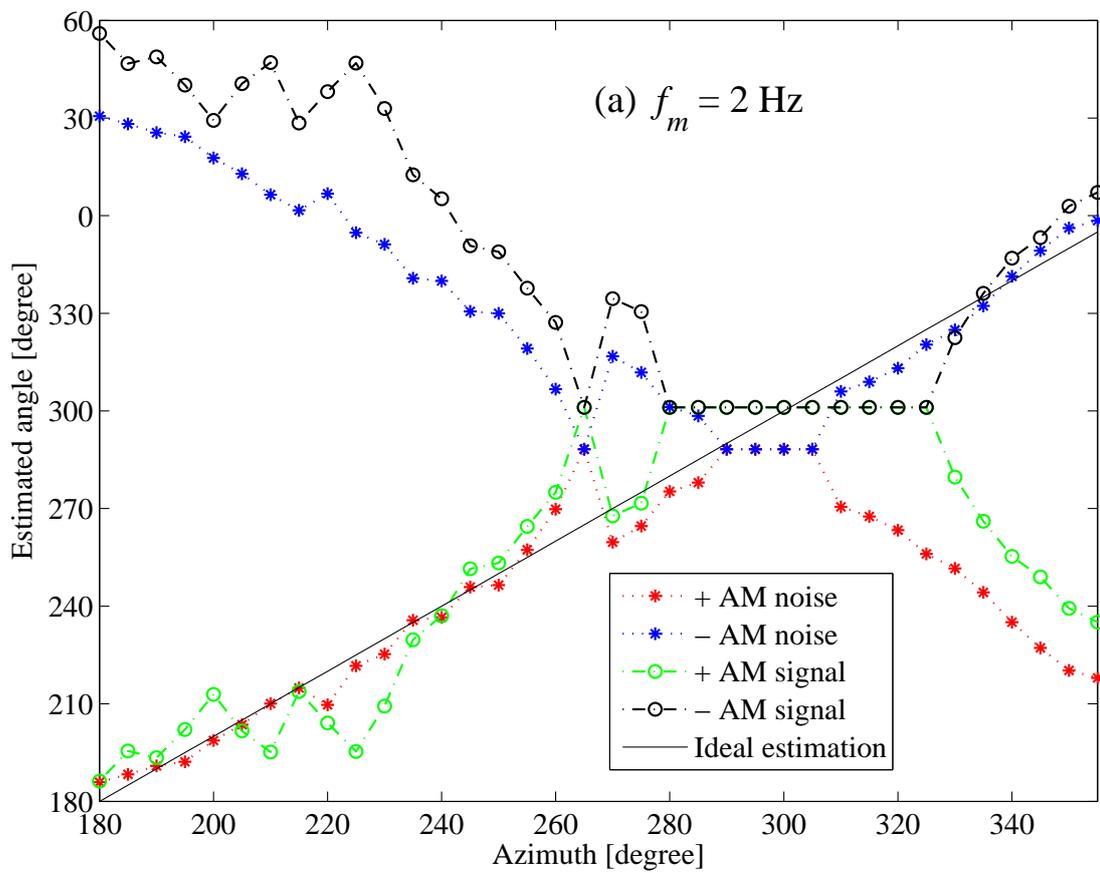


図 4.4: 音源方向推定結果: 2 Hz

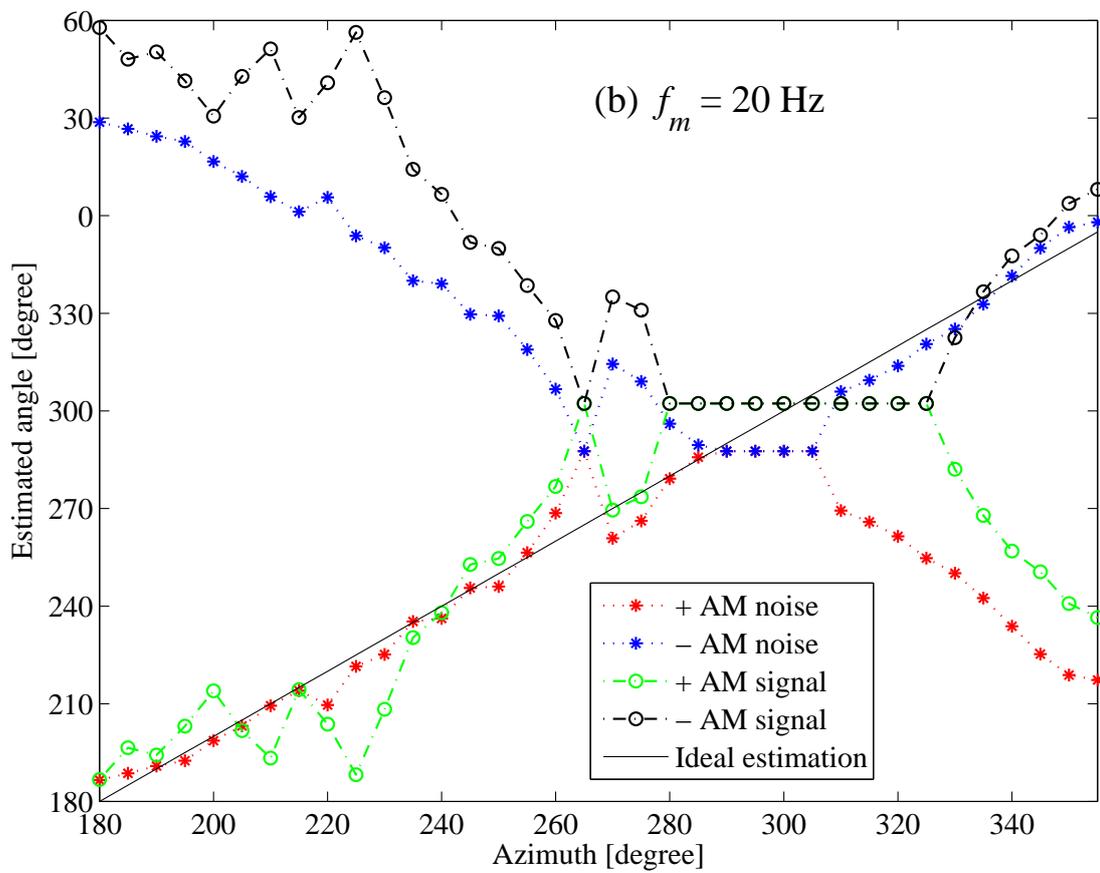


图 4.5: 音源方向推定結果: 20 Hz

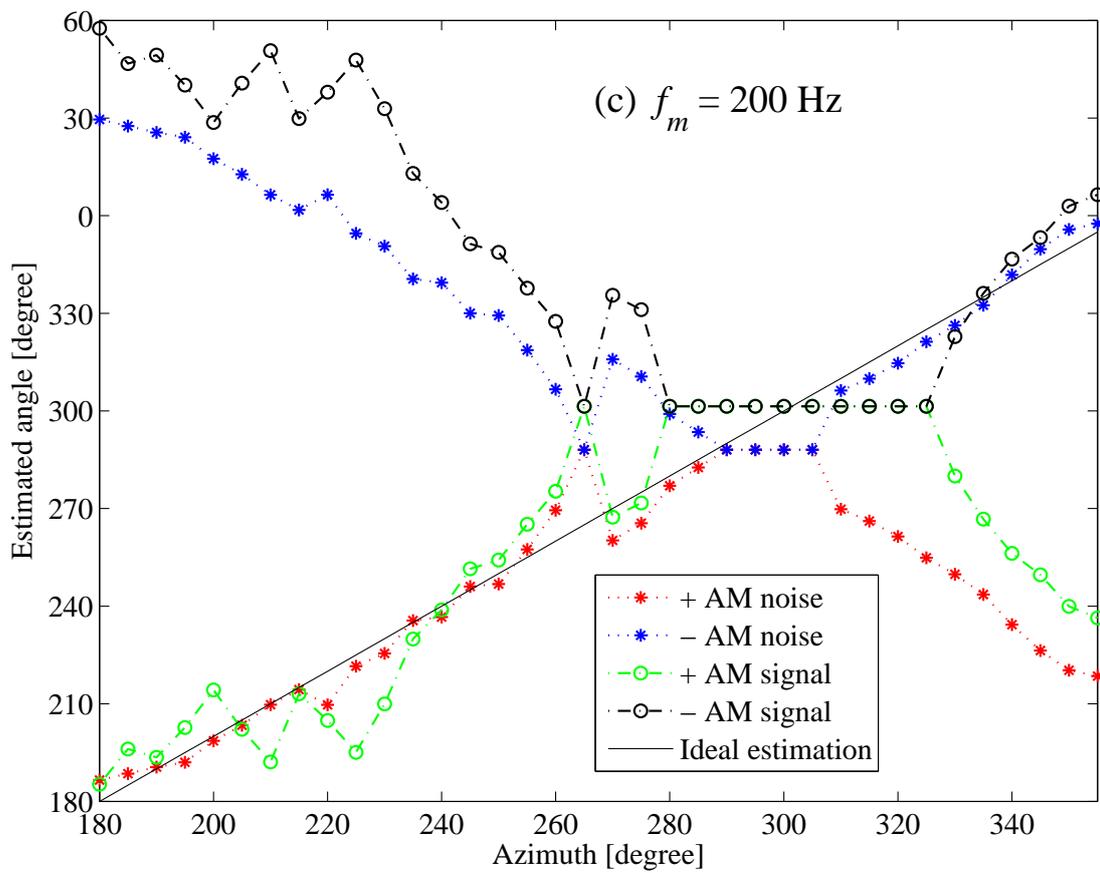


図 4.6: 音源方向推定結果: 200 Hz

4.6 前後弁別法

前節の式 (4.2) にはプラスマイナス項が含まれていたため，前後の推定誤りが生じた．そこで，前後の推定誤りに対して，以下の改善策を検討する．

ヒトが頭部運動をすることで音源方向推定精度が向上すると報告されている [27]．これは，移動前の観測信号と移動後の観測信号の間に擬似的な時間差を生じさせることで，幾何学的な手がかりを増やすことと同義である．そこで，ヒトの頭部運動をマイクの移動で代用し，上述と同様に時間差によって生じた変化を利用する．前節で提案した音源方向推定法に，MMS の時間変化が生じるような条件を設ける．未知の角度 θ から音源信号が到来する際に，マイクを正面方向から背面方向へ半時計回りに水平角 $\Delta\theta$ 度移動したと仮定する．この場合，見かけ上，音源が時計方向へ $\Delta\theta$ 度移動したことになる，移動前の MMS の値と移動後の MMS の値には増減が生じる．ここで，MMS が角度変化によって弧状に変化する規則性を考慮すると，移動前の MMS の値が移動後に増加した場合は必ずマイク位置より後方（180 度から 270 度）である．逆に，移動前の MMS の値が移動後に減少した場合はマイク位置より前方（270 度から 355 度）である．

提案法では二つの角度 $\hat{\theta}_1$ および $\hat{\theta}_2$ が推定されているが，今回の条件を設けることで必ず一つの角度が選択される．以上のように，頭部運動を想定した，マイクの移動による MMS の増減を利用することで前後の判定を行う．

MMS の増減を判定するための式を導出する．式 (4.1) を微分すると次式が得られる．

$$\hat{E}_y'(f_m, \theta) = 2p_1(f_m)\theta + p_2(f_m) \quad (4.3)$$

式 (4.3) に，移動前の推定角度 $\hat{\theta}_1$ ， $\hat{\theta}_2$ および移動後の推定角度 $\hat{\theta}_1 + \Delta\theta$ ， $\hat{\theta}_2 + \Delta\theta$ を代入し MMS を求め，後者から前者をそれぞれ減算した結果より，増減を判断する．ここで，増加した場合は，推定角度が 180 度から 270 度が正しい推定角度であり，減少した場合は 270 度から 355 度が正しい推定角度である．以上のアルゴリズムによって，前後弁別される．

4.7 前後弁別の検証

4.7.1 検証方法および条件

前後弁別法の有効性を確認するため，シミュレーションによる検証を行った．検証方法は 4.5.1 節の検証方法に加えて，マイクを正面方向から背面方向へ半時計回りに水平角 30 度移動したと仮定し，推定角度 $\hat{\theta}_{1,2} + 30$ 度の MMS を式 (4.3) から算出する．次に，移動前 $\hat{\theta}_{1,2}$ 度の MMS をそれぞれ減算することで増減を算出し，前後弁別する．最後に，前後に対応する推定角度を選択する．

3.2 節で利用したのと同じ AM 信号および AM 雑音を音源信号とした．また，ここで利用する HRIR も 3.2 節で利用したのと同様であり，56 名分をシミュレーションに用いた．シミュレーションの範囲は，単耳と同側である．

4.7.2 検証結果

マイクを正面方向から背面方向へ半時計回りに水平角 30 度移動したと仮定した場合の前後弁別検証結果を，図 4.7 から図 4.9 に示す．ここで，横軸は音源信号方向を，縦軸は推定角度を表す．全ての結果において，200 度付近を除いて前後の誤推定を改善していることがわかった．図 4.7 から図 4.9 と図 4.4 から図 4.6 を比較すると，前後弁別法を適用した場合には，理想推定角度に沿った推定角度を選択していることがよく分かる．このことから，前後弁別法の有効性が示された．

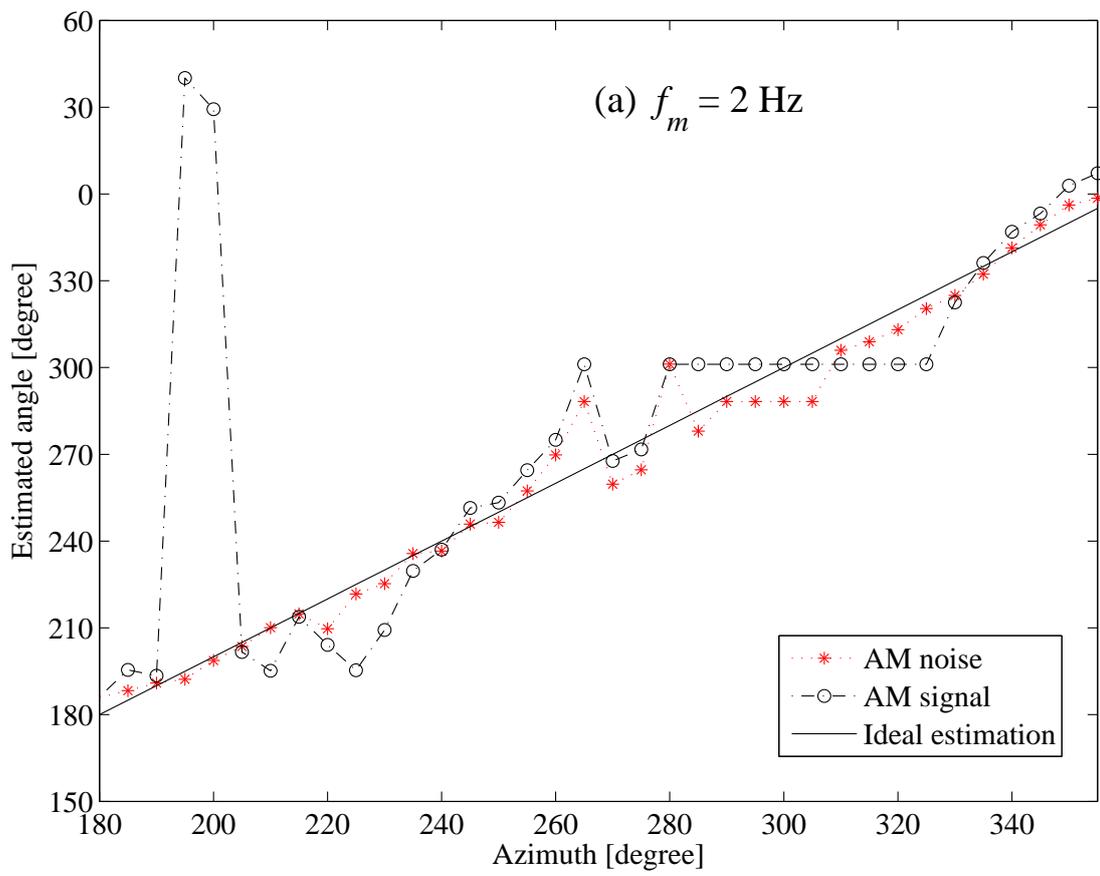


図 4.7: 前後弁別後の音源方向推定結果: 2 Hz

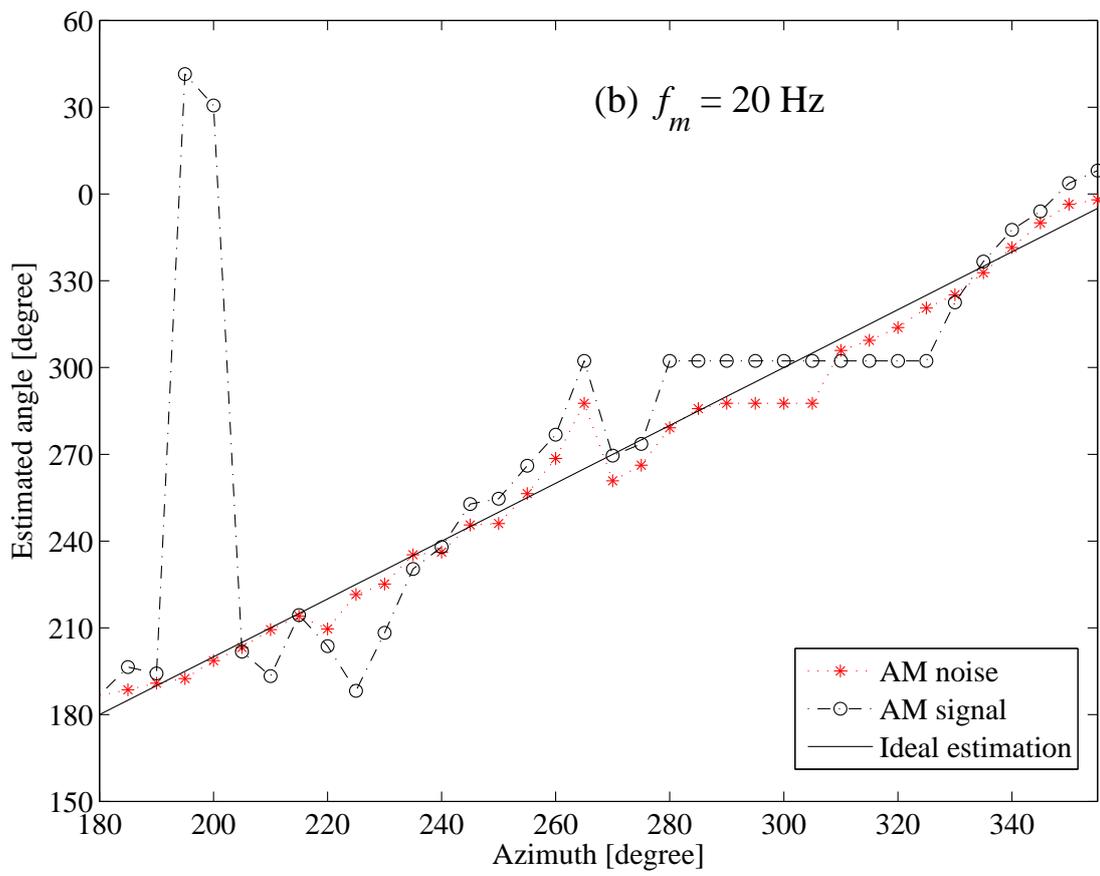


図 4.8: 前後弁別後の音源方向推定結果: 20 Hz

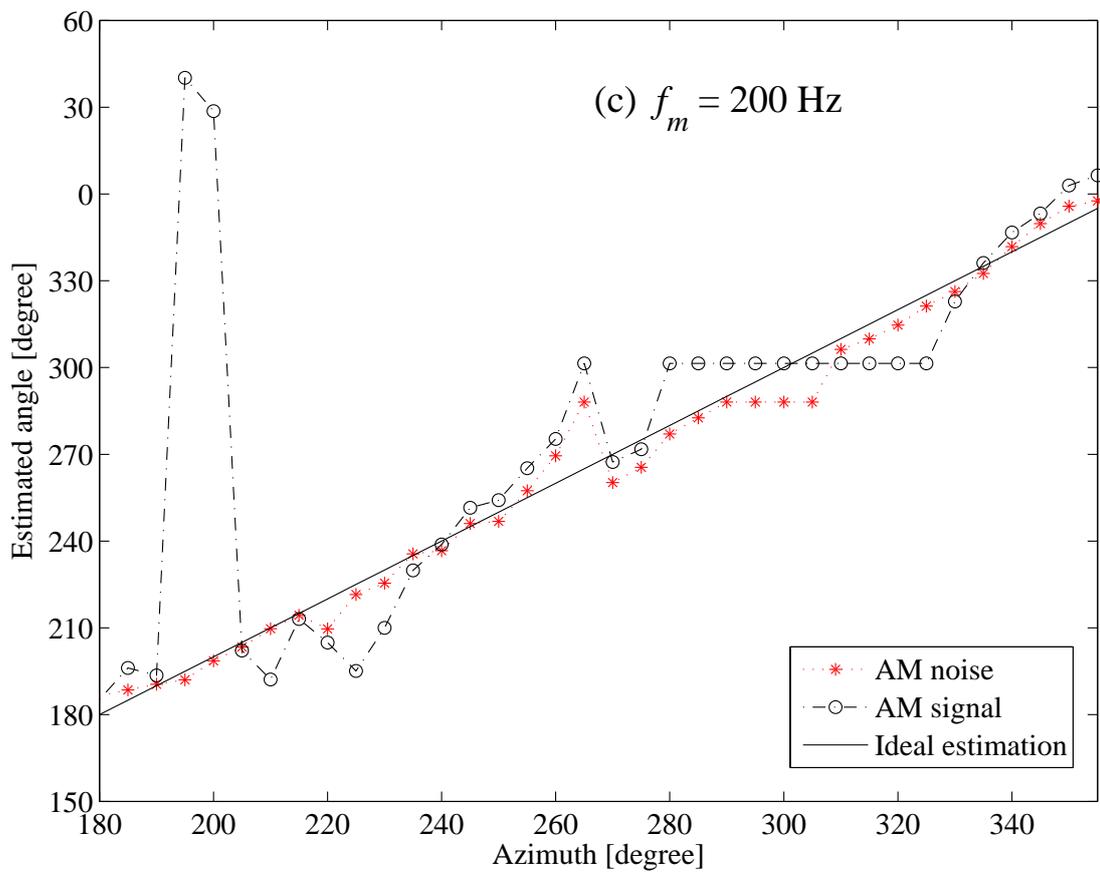


図 4.9: 前後弁別後の音源方向推定結果: 200 Hz

4.8 考察

モノラル音源方向知覚の手がかりを利用した音源方向推定法は、理想条件下では、前後推定誤りが生じているが、概ね音源方向を推定できた。ただし、音源方向推定精度はMMSの値と回帰曲線フィッティング率に依存しているため、これ以上の精度向上は難しいと考えられる。また、実環境音を想定した場合には、音源信号の変調周波数は未知である。しかし、利用可能な変調周波数のMMSを統合することで提案法の適用が可能であると考えられる。

4.7節より、前後弁別法は提案法において有効であることがわかった。200度付近で前後弁別ができていないが、さらに大きなMMSの時間変化をもたらすことで改善が可能であると考えられる。

4.9 まとめ

本章では、前章の調査で明らかになった結果をもとに、MMSの水平角および仰角変化特性を利用したモノラル音源方向推定法を検討した。また、水平角に着目したモノラル音源方向推定法について、工学的応用を想定した実現例を示し、シミュレーションによって検証した。その結果、マイクと同じ側において、モノラル音源方向推定法の有効性を示した。

第5章 結論

5.1 本研究で明らかになったことの要約

本研究では、ヒトのモノラル方向知覚の手がかりを解明することを目的とし、MMSの角度変化を調査した。また、調査した結果から解明された手がかりを利用したモノラル音源方向推定法の検討も同時に行った。まず、ヒトの頭部インパルス応答(HRIR)の変調伝達関数(MTF)に着目し、MMSが音源信号の到来角度によってどのように変化するか調査した。その結果、水平角では、音源信号をAM信号とした場合において、MMSの角度変化には耳位置を頂点とした弧形(あるいは負の弧形)となる規則性が示された。これは音源信号の変調周波数に影響されないことを確認した。また、音源信号の搬送波周波数が変化した場合においては、単耳の反対側に、頭部の回折の影響とみられる変化が見られた。また、少なからず肩などの反射による影響もみられた。仰角では、頭頂部付近から後頭部にかけて、MMSが減衰する傾向が見られた。また、頭頂部付近において、音源に対して対側から同側にかけてMMSが増加する傾向がみられた。これらのことから、直接音や回折音あるいは反射音が含まれたMMSを知覚することは、十分にヒトの方向知覚の手がかりに利用可能であることが示唆された。また、以上の分析結果に基づき、理想的な条件を設けた上で、回帰近似を利用した音源方向推定モデルおよび頭部運動を利用した音源方向推定モデルを検討した。さらに、水平角に着目した音源方向推定の実現例をシミュレーションによって示した結果、水平角半面において音源方向推定が可能であることを示した。

5.2 残された課題

本研究では、分析結果に基づき、理想的な条件を設けた上で、回帰近似を利用した音源方向推定モデルおよび頭部運動を利用した音源方向推定モデルを検討し水平角に着目した音源方向推定の実現例を提案した。しかし、推定範囲が水平角半面であった。よって、水平角・仰角方向を含めたモノラル音源方向推定法の実装ならびに評価を行う必要がある。

また、提案法は理想条件下でシミュレーションを行ったが、音源信号が音声信号などの実環境音を想定した場合を考慮する必要がある。よって、音源信号を実環境音にした場合のシミュレーションを行う必要がある。

謝辞

本研究を進めるにあたり，指導教官である北陸先端科学技術大学院大学情報科学研究科の鷓木祐史准教授には熱意あふれるご指導や叱咤激励を賜りました．ここに深く感謝します．また，研究室会議などで数多くのご助言を頂いた北陸先端科学技術大学院大学情報科学研究科の赤木正人教授，ならびに宮内良太助教に深く感謝します．さらに，公私問わず熱心にご指導，ご討論をしていただいた北陸先端科学技術大学院大学情報科学研究科の森川大輔助教に心より感謝します．

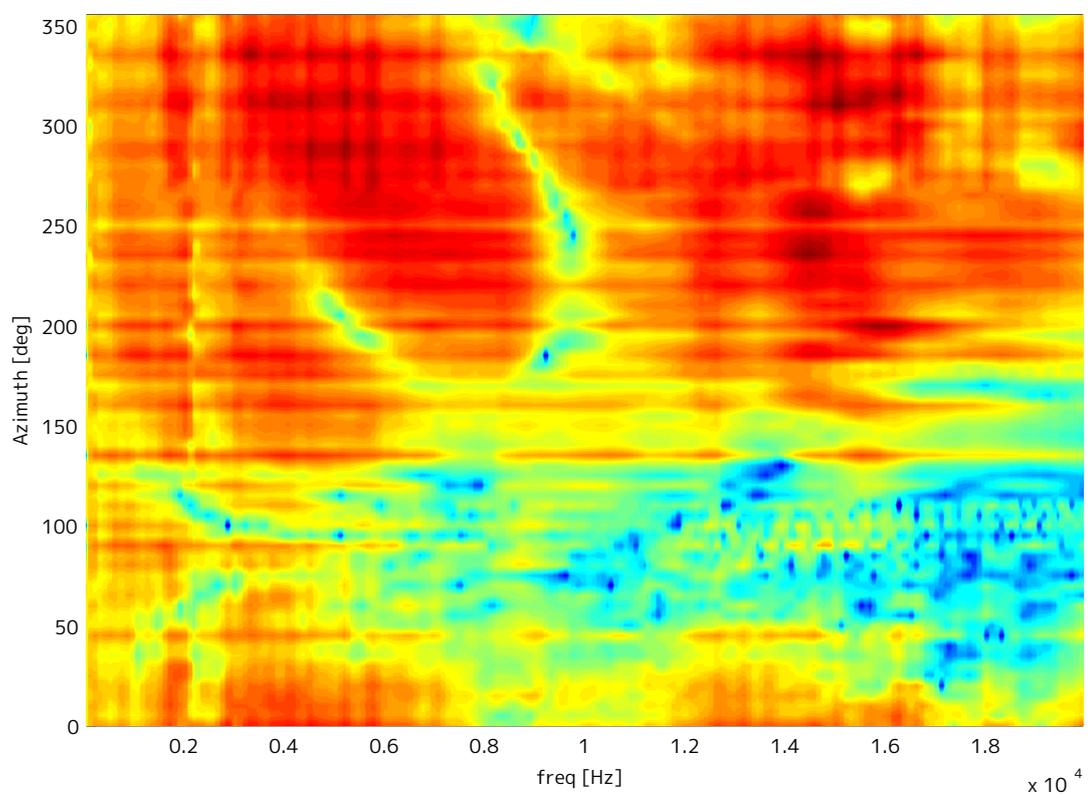
本研究の遂行にあたり，貴重なデータベースを提供していただきました，東北大学電気通信研究所鈴木陽一教授，ならびに富山大学知能デザイン工学科の平原達也教授のご厚意に深く感謝します．

また，著者の研究生生活をおくる上で大変支えとなっていたいただいた，研究室の諸先輩や諸氏に深く御礼申し上げます．最後に，本学での研究生生活を支えていただき，温かく見守っていただいた両親ならびに姉弟に，心から感謝致します．

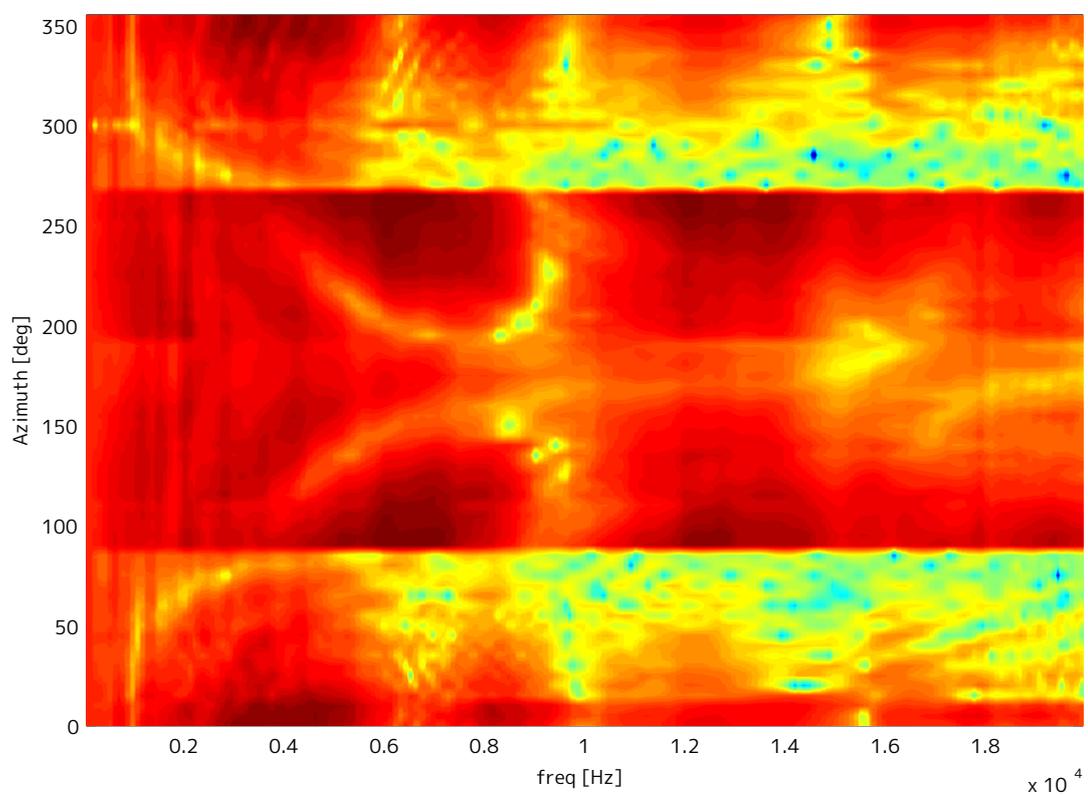
付録 A

除外した HRIR の例を示す．HRTF の Azimuth のスペクトルをカラーマップ表示したもの．一枚目が隣接しあう角度のスペクトルが極端に不連続な例，二枚目が測定ミスと見られる例．

HRTF左耳 振幅スペクトル



HRTF左耳 振幅スペクトル



参考文献

- [1] J. Blauert, Spatial Hearing, The MIT Press, Cambridge, 1974.
- [2] 飯田一博, 森本政之, 空間音響学, コロナ社, 東京, 2010.
- [3] 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.
- [4] 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.
- [5] 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.
- [6] 古川茂人, “動的な両耳間時間差・強度差変調の検出-両耳検出の緩慢性はどこからくるのか-,” 聴覚研究会資料, vol. 39, no. 7, H-2009-93, pp. 527-530, 2009.
- [7] 古川茂人, “両耳間時間差・強度差変調の検出-変調相対位相の効果-,” 聴覚研究会資料, vol. 41, no. 5, H-2011-65, pp. 355-358, 2011.
- [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] 牧野昭二, 荒木章子, 向井良, 澤田宏, “独立成分分析に基づくブラインド音源分離,” 信学技法, vol. 103, no. 129, EA2003-45, pp. 17-24, 2003.
- [10] 辻大亮, 陶山健仁, “逐次更新則による安定な移動音源追尾,” 信学技法, vol. 107, no. 120, EA2007-78, pp. 33-38, 2007.
- [11] 鈴木敬, 金田豊, “サブバンドピークホールド処理を用いた音源方向推定法,” 日本音響学会誌, vol. 65, no. 10, pp. 513-522, 2009.
- [12] 山崎芳男, 金田豊, 音・音場のデジタル処理, コロナ社, 東京, 2002.
- [13] 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.

- [14] 佐保貴哉, 菅木禎史, 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, 2011–3.
- [15] 富田拓郎, 佐保貴哉, 菅木禎史, 宇佐川毅, “前方のみに指向性を生成可能な周波数領域両耳聴モデル-HRTF カタログ間の前後誤判定特性の比較-,” 日本音響学会講演論文集, pp. 835–838, 2012–9.
- [16] 住田雄司, 滝口哲也, 有木康雄, “音響モデルを利用したシングルチャンネルによる音源方向推定,” 信学技法, vol. 106, no. 432, EA2006–90, pp. 7–11, 2006.
- [17] 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.
- [18] A. Fuchs, C. Feldbauer, and M. Stark, “Monaural sound localization,” Proc. Interspeech 2011, pp. 2521–2524, Florence, Italy, 2011.
- [19] 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.
- [20] 平原達也, 大谷真, 戸嶋巖樹, “頭部伝達関数の計測とバイノーラル再生にかかわる諸問題,” Fundamentals Review, vol. 2, no. 4, pp. 68–85, 2009.
- [21] アラン R. パーマー, トレバー M. シャクレトン, “両耳聴の神経機構,” 日本音響学会誌, vol. 58, no. 3, pp. 184–192, 2001.
- [22] 森周司, 香田徹, 聴覚モデル, コロナ社, 東京, 2012.
- [23] Brian C. J. Moore, An Introduction to the Psychology of Hearing Sixth Edition, BRILL, Boston, 2013.
- [24] T. Houtgast and H. J. M. SteeneKen, “The Modulation Transfer Function in Room Acoustics as a Predictor of Speech Intelligibility,” Acustica, vol. 28, pp. 66–73, 1973.
- [25] M. Unoki, “Speech Signal Processing Based on the Concept of Modulation Transfer Function (1) -Basis of Power Envelope Inverse Filtering and Its Applications-,” Journal of Signal Processing, vol. 12, no. 5, pp. 339–348, 2008.
- [26] 今井悠貴, 森川大輔, 平原達也, “相反法による頭部伝達関数の測定,” 信学技法, vol. 112, no. 266, EA2012–72, pp. 43–48, 2012.
- [27] 野島遼子, 森本政之, 佐藤逸人 “方向定位における頭部運動について,” 日本音響学会講演論文集, pp. 515–516, 2009–10.

学会発表リスト

1. 安藤 将, 鷓木 祐史, “モノラル変調スペクトルに着目した音の到来方向推定の可能性,” 平成 25 年度電気関係学会北陸支部連合大会, 講演番号:G-3, 2013.
2. 安藤 将, 森川 大輔, 鷓木 祐史, “変調スペクトルに着目したモノラル音源方向推定法の検討,” 第 28 回信号処理シンポジウム A3-4, pp. 188-193, 山口, 2013.
3. 安藤 将, 森川 大輔, 鷓木 祐史, “変調スペクトルに着目したモノラル音源方向推定法の検討,” 第 4 回音響セミナー in 金沢, 金沢, 2013.
4. Masaru Ando, Daisuke Morikawa, and Masashi Unoki, “Study on DOA estimation method based on human ability of sound localization,” The 6th seminar of A3 foresight program, Ishikawa, 2014.
5. Masaru Ando, Daisuke Morikawa, and Masashi Unoki, “Study on method of estimating direction of arrival using monaural modulation spectrum,” 2014 RISP International Workshop on Nonlinear Circuits and Signal Processing, 2014.
6. 安藤 将, 森川 大輔, 鷓木 祐史, “モノラル変調スペクトルを利用した音源方向推定法の検討,” 信学技法, vol.113 no.503 EA2013-124, pp. 25-30, 2014.