

頭部伝達関数と音像定位を巡る諸問題*

○ 平原達也（富山県立大学／東北大学電気通信研究所）、大谷 真（東北大学電気通信研究所）

1 はじめに

音源信号に頭部伝達関数（HRTF: Head-Related Transfer Function）が畳み込まれたバイノーラル信号によって立体音像空間を再現する場合に、HRTF やバイノーラル信号再生系の音響的な厳密さが重要視されている。しかし、受聴者の頭部運動に追従する動的バイノーラル信号を用いると、HRTF やバイノーラル信号再生系に要求される音響的な厳密性を緩和できる。

本稿では、複数箇所を実施した HRTF の計測^[1]を通じて、HRTF の音響計測に係る問題点と HRTF の比較尺度について論じる。また、様々なバイノーラル信号を用いた一連の音像定位実験の結果^[2]に基づいて、静的バイノーラル信号と動的バイノーラル信号が再現する立体音像空間の差異についても論じる。これらを通じて、HRTF の計測とバイノーラル信号による音像定位を巡る諸問題について論考する。

2 頭部伝達関数計測の miss と myth

2.1 頭部伝達関数

HRTF は、頭部がある場合の位置 $S(r_o, \theta, \varphi)$ にある音源から外耳道入口 E に置いたマイクロホンまでの音響伝達関数 $H_E(S, \omega)$ を、頭部が無い場合の同じ音源から頭部中心位置 O に置いたマイクロホンまでの音響伝達関数 $H_O(S, \omega)$ で除した伝達関数 $H_E(S, \omega)/H_O(S, \omega)$ で与えられる。

H_E を H_O で除するのは音響系と電気系の諸

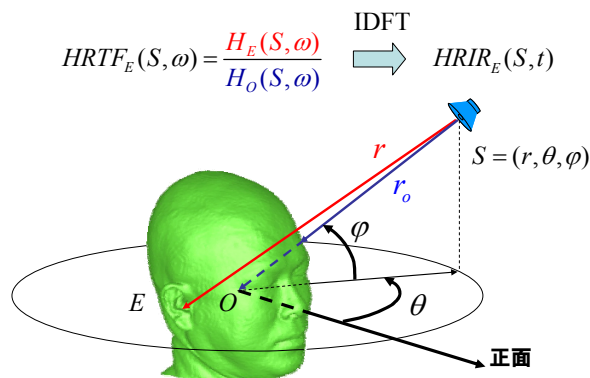


Fig.1 頭部伝達関数 HRTF の定義

特性を相殺するためである。また、 $HRTF_E(S, \omega)$ を離散フーリエ逆変換して得られるインパルス応答は HRIR (Head-Related Impulse Response) と呼ばれる。

r は S から E までの距離、 r_o は S から O までの距離だが、両耳間距離は 0.15 m 程度であり、 $r_o = r$ として差し支えない。マイクロホンは無指向性と仮定して $H_O(S, \omega)$ は真正面の 1 点で代用するが多い。

2.2 頭部伝達関数の計測系

HRTF を計測するシステムの基本部分は Fig. 2 に示すようなもので、音源となるラウドスピーカから頭部中心位置および外耳道入口においたマイクロホンまでの音響伝達関数すなわち音響インパルス応答を正確に測ることが重要である。音波は媒質の圧力変化という物理的な現象であり、様々な物理量に影響を受ける。マイクロホンは十分に感度が高く、それらの影響を検出できるために、音響計測は測定系の物理的な状態が正直に反映される。

そのため、音響計測に際しては、可能な限り反射物体を除去すること、電気音響機器と電気音響変換器のダイナミックレンジを考慮した適切な電圧信号のレベル配分、ハム等の電磁雑音の低減と計測空間の暗騒音を下げることが最低限必要である。これに加えて、HRTF の計測には正確な音源と受音点の相対位置の設定が必要である。

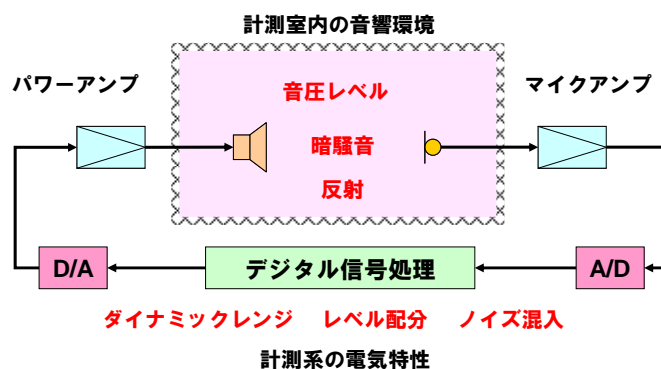


Fig.2 HRTF 計測系と音響計測の外乱要因

* Myth and miss of head-related transfer functions and sound image localization, by Hirahara, Tatsuya (Toyama Prefectural Univ., RIEC, Tohoku Univ.), Otani, Makoto (RIEC, Tohoku Univ.)

2.3 反射波成分

Fig. 3は、3箇所でのHRTF計測サイトにおける、真正面から出力したTSP (Time Stretched Pulse) 信号を頭部中心位置に設置したマイクロホンで受信し加算平均したTSP 応答信号のスペクトログラムの一例である。(a)では反射はほとんど認められないが、(b)では2つの強い反射波が、(c)では弱いながら1つの反射波が認められる。(a)(b)は無響室であるが(c)は通常の部屋の床・壁・天井にポリエステル不織布吸音材(ササクラ、厚さ50mm、密度22kg/m³)を貼り付けただけの部屋である。(b)の計測サイトではフレームに多数のスピーカを設置した球状スピーカアレイを用いており、このフレームやスピーカコーンからの6.3~16.7msの遅延を伴う反射波が生じている。このように、無響室を使用しているからといって安心してはいけない。

HRIRとして用いられているインパルス応答の長さは512点程度なので、遅延時間の長い反射波成分の影響は受けないが、48kHzでサンプリングしている場合10.7ms以下の短い遅延時間を持つ反射波成分はインパルス応答長に含まれてしまう。このような反射波成分があるといわゆるカラーレーションが生じ、HRTFの振幅スペクトルは波打つ。その結果、水平面HRTFの方位周波数パターンを描くと縦縞が現れ縞々になる。それを防止するために、しばしばHRIRに窓をかけて反射波成分を取り除くことが行われるが、HRIRに含まれるべき応答の一部を無視することになる。

また、頭部中心位置にマイクを置いたときの周囲の状況と、実頭やダミーヘッドを置いたときの状況とが大きく異なると、音響系の相殺が機能せず、面妖なHRTFが得られてしまう。特

に、反射物の有無や計測室の温度には最大限の注意を払うべきである。頭部中心位置での耳栓マイクのインパルス応答を計測する際に、被験者を座らせる椅子の上に約80cmのマイクロホンスタンドを立て、その椅子を吸音材で覆わずに測定してしまっただけである。そのときは、椅子の座面からの数msの遅延を持つ強い反射波がインパルス応答に混入し、結果は惨憺たるものとなり、再計測を余儀なくされた。

2.4 高調波歪

Fig.3に示されるように、いずれのスペクトログラムにもTSP信号の主成分の他に、より急峻に周波数変化する高調波歪成分が現れている。(a)では3次高調波歪成分が主成分の-55dB、(b)では3次高調波歪成分が主成分の-60dB、(c)では2次、3次、4次高調波歪成分も見え、最も強い3次歪成分は主成分の-50dB程度である。これらの歪は主として小型のラウドスピーカで発生しているものであり、信号のS/Nを向上させようとして入力音圧を上げる前に、精査が必要である。

2.5 雑音

Fig.3に示すように、(a)と(b)には低域に断続的な強い雑音成分が、(b)と(c)には定常的な弱い雑音成分が現れている。低域の雑音成分は無響室のような密閉空間に生じる建築物や空調器の振動によるもの、あるいは信号ラインに混入したハム雑音である。前者の本質的な低減は困難であるが、後者は努力次第で除去できる。一方、高域の雑音成分は様々な成因によるもので、努力次第で除去できる。(b)では收音系の改善によって、(c)では部屋の蛍光灯を消灯することによって高域雑音は除去できた。これらの雑音対策は、ケースバイケースで対処せざるを得ない。

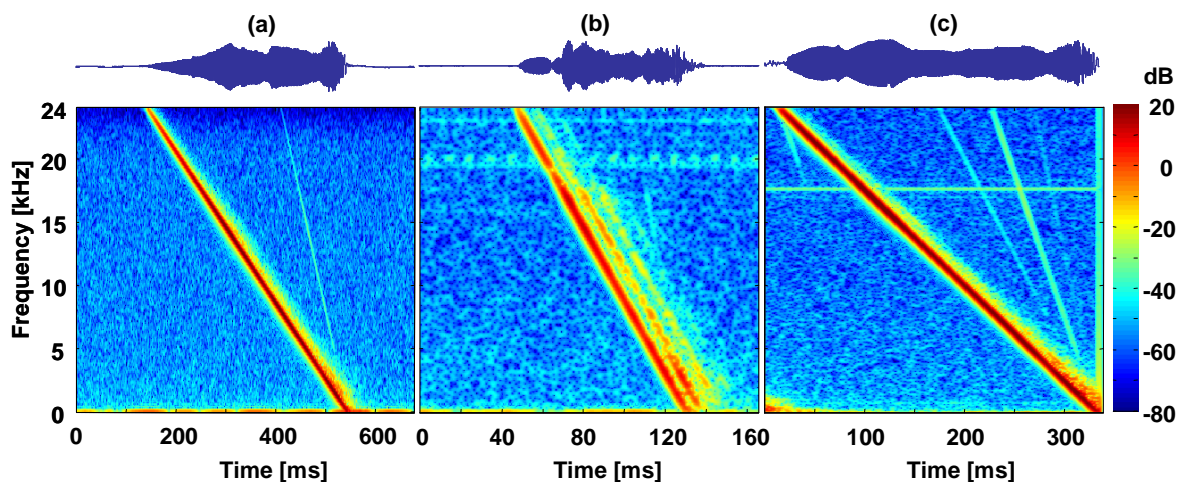


Fig.3 三箇所でのHRTF計測サイトにおけるTSP 応答信号波形とスペクトログラム

2.6 被測定物の位置合わせ

HRTF の計測時にマイクロホンや頭部の位置を正確に定めることは、データの再現性を確保するために重要である。著者らはレーザーポインタを用いて実頭やダミーヘッドや頭部中心マイクの位置合わせを行っていたが、最近では、レーザー墨出し器を用いて位置合わせを短時間で確実にしている (Fig. 4)。

しかし、一旦設置すると動かないダミーヘッドやマイクロホンと異なり、実頭は HRTF の計測中に頭部位置を固定することは容易ではない。頭部を全く固定しない場合には、95 分間に渡る HRTF 計測時に頭部は大きく動く。また、後頭部を保持するサポート具を用いた場合でも頭部は動く。Fig. 5 は、後頭部を保持した状態で 60 分間に渡る HRTF 計測の最初 ($\theta = 0^\circ$) と最後 ($\theta = 360^\circ$) に 2 回測定した真正面の HRTF を比較したものである。ダミーヘッドの HRTF はほとんど一致するのに対して、実頭ではスペクトルノッチの周波数が異なり、頭部が動いていたことがわかる。

頭部の位置が変化すると HRTF は変化するので、固定具の影響を評価した上で頭部を固定するか、常に頭部位置を補正するか、頭部が動かないうちに HRTF の測定を終えるか、あるいは、誤差があるのを承知の上で計測データを利用するか、である。

2.7 HRTF の比較尺度

二組の HRTF の差は平均振幅スペクトル距離 SD (Spectral Distance) を尺度として評価する場合が多い。音声情報処理の分野では、多数の実データによってこの SD の有効性が実証されている。しかし、実測した HRTF の SD は 4~8 dB と狭い範囲の値しかとらないし、そこには計測系の誤差や位置ずれに伴う誤差が 3 dB 以上含まれるために、頭部形状の違いを HRTF の SD で評価することは難しい。

また、下丘より下位の神経核では聴覚フィルタ帯域内での情報処理が行われているために、聴覚フィルタの帯域幅で積分した聴覚系内スペクトルを用いると音の聴こえ方をうまく説明できることが多い。音声の聴覚系内スペクトルは物理的なスペクトルとは様相が大きく変化するが、HRTF スペクトルはもともと低ケレンシー成分しかないので、Fig.6 に示すように聴覚フィルタ処理を行っても大きな変化はない。対数周波数軸や ERB-rate 上で等間隔にと

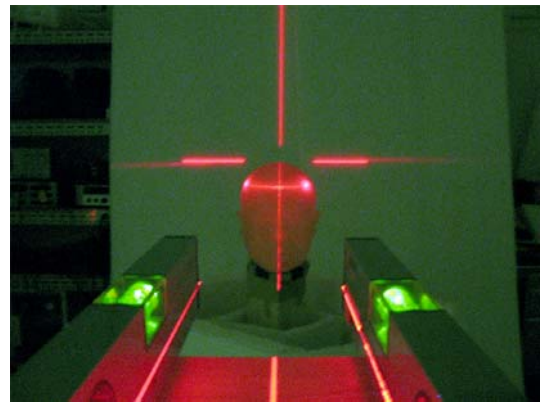


Fig.4 レーザー墨出し器を利用した位置合わせ

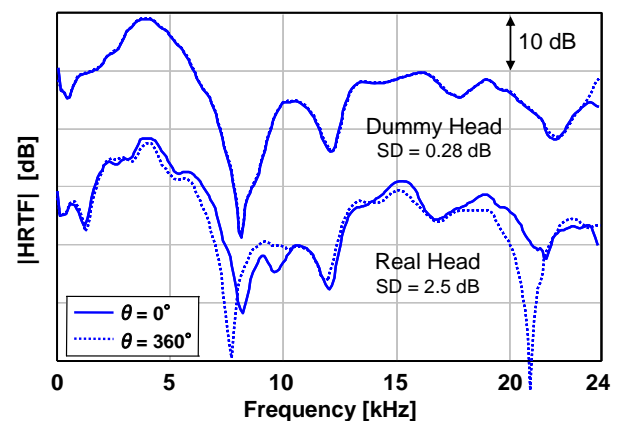


Fig.5 ダミーヘッドと実頭の真正面 HRTF

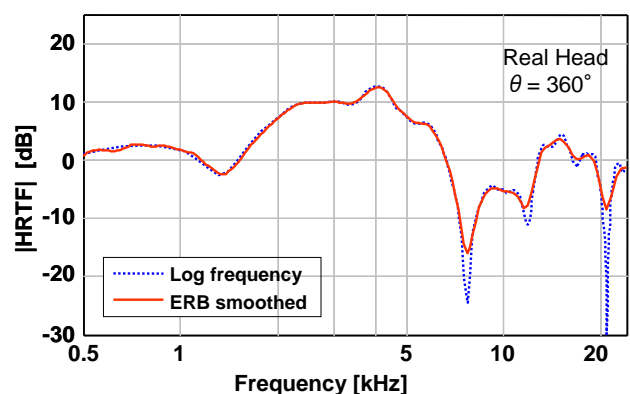


Fig.6 対数 HRTF と ERB 平滑した HRTF

ったスペクトルに基づいて SD_{\log} や SD_{erb} も算出できるが、変化に乏しい低域周波数が強調される SD_{\log} や、スペクトルが平滑化される SD_{erb} は、いずれも 3~5 dB と通常線形周波数軸より算出される SD より狭い範囲の値しか採らない。このように、 SD という物差しを当てると HRTF は似たり寄ったりなのである。

さらに、聴覚は音声信号の位相差には鈍感だが、バイノーラル信号の位相差には敏感で、音像の聴こえ方の違いを HRTF の SD だけで評価することは難しい。ITD を含めた物差しや、特徴抽出という困難はあるが、飯田らが提案するような HRTF スペクトルの構造と知覚手がかりを踏まえた物差し^[3]の検討も必要である。

3 バイノーラル再生の *myth* と *miss*

市販のダミーヘッドで収録したバイノーラル信号や、HRTF を畳み込んで合成したバイノーラル信号をイヤホンで聴くと、確かに立体的な音像が頭の外から聴こえるが、その立体音像空間はしばしば歪む。その理由は、用いた HRTF が受聴者自身の HRTF ではないことや音響再生系が音響的に不完全であることが原因と言われている。そして、HRTF をいかに受聴者に適合させるか、電気的なバイノーラル信号をいかに忠実に再生するか、といった観点から、多くの研究がなされ、HRTF の個人適応化法や選択法やヘッドホン特性の補正法やバイノーラル受聴に適したヘッドホンなどが提案されている。これらは、バイノーラル信号を物理的に正しく再生するための音響的な視点である。

しかし、これらのほとんどは、静的バイノーラル信号を用いた音像定位実験に基づいて検証されている。古くから知られているように、受聴者の頭部運動に追従させた動的バイノーラル信号を用いると、再生される立体音像空間の様相は大きく変わる。すなわち、自分で頭部を動かすと、他人のダミーヘッドを使っても、縮小ダミーヘッドを用いても、他人の HRTF を畳み込んで、ステレオマイクを用いても、ヘッドホンの実耳応答特性を補正しなくても、FEC ヘッドホンでなくても、頭部運動と音像運動に 100 ms 程度遅延があっても、再生立体音像空間の歪は激減する。

動的バイノーラル信号を用いても、ヘッドホンの片側を逆相接続した場合と、他人の頭部運動に追従したステレオマイクの出力をバイノーラル受聴した場合には、立体音像空間は歪む。Fig.7 は、左ドライバを逆相接続した HDA200 を用いて、自分の HRTF を畳み込んだ合成バイノーラル信号で水平面の音像定位実験を行った結果である。頭部を動かすと音像は頭外に出るものの、HDA200 とは思えないほど頭部の近傍に定位しその位置は曖昧であった。Fig.8 は 300 mm 離れたステレオマイクロホンを用いて自分が頭部を動かした場合と、他人が頭部を動かした場合の水平面の音像定位実験結果である。他人が頭部を動かすと音像は動いて聴こえるが、その動きは自分が頭を動かした場合と全く異なり、音像空間の歪は大きい。仕組みは不明であるが、音像の知覚に運動系が深く関与していることは間違いない。

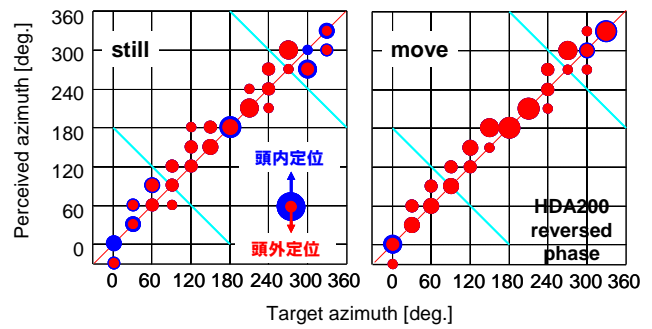


Fig.7 逆相接続した HDA200 による音像定位

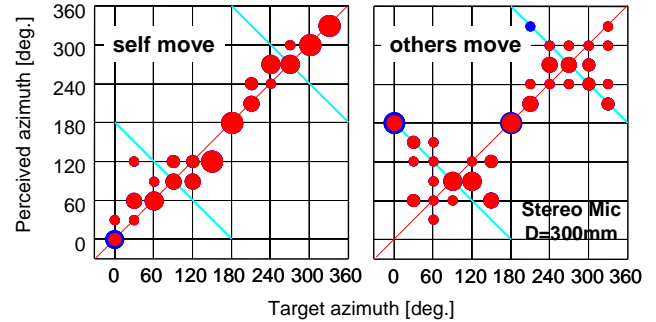


Fig.8 他人が動かしたステレオマイクによる音像定位

どの程度の音像定位精度が必要かはアプリケーション次第であるが、その精度を獲得するためには音響的な厳密さを向上させる方法だけでなく、聴き手の情報処理の仕組みを利用する方法もある。立体音像の妥当な聞こえを生じさせるために必要な手がかりは何か、どのようなメカニズムでその手がかりから音像の位置を計算しているのか、という聴覚的な視点についてはまだまだ未解決問題が山積している。

4 まとめ

HRTF の計測に際しては音響計測の基本を忘れてはならない。また、HRTF の SD は HRTF を形成する頭部形状の差や HRTF によって生成される音像の定位性能の差を必ずしも反映しない。この三者を結びつける適切な物差しの探求が必要である。さらに、音像定位精度を改善するためには音響的視点とともに聴覚的な視点も大切である。

謝辞：本研究の一部は科学研究費補助金 (18300042) および東北大学電気通信研究所共同プロジェクト研究による。

参考文献

- [1] 平原 他, 頭部伝達関数論考, 聴覚研究会資料 38(1), H2008-4, pp.19-24, 2008.
- [2] 平原 他, 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 2(4), 2009.
- [3] 飯田 他, 方向知覚メカニズムに基づいた HRTF の物理評価指標, 音講論, 505-508, 2008-09.