

低域のITD情報を操作したバイノーラル信号による水平面音像定位*

©森川大輔, 平原達也 (富山県立大 工学部)

1 はじめに

両耳間時間差(ITD)と両耳間音圧差(ILD)は音源から両耳までの音の到達時間の差、到達した音圧レベルの差として、通常、全帯域から何 ms、何 dB と算出する。しかし、実際にはILDはもちろん、ITDも周波数によって異なる^[1,2]。また、聴覚系でも、上オリーブ各複合体では周波数チャンネルごとにITDとILDを検出し、周波数チャンネル間の両耳間情報を統合しているのは下丘以上^[3]である。

本報告では、HRTFを低域(1 kHz以下)と高域(1 kHz以上)に分割し、低域のITDを系統的に操作した変形HRTFを用いて合成したバイノーラル音で音像定位実験を行い、低域と高域でITDが異なることが、音像の知覚にどのような影響を及ぼすかを検討した結果について述べる。

2 方法

2.1 刺激音

被験者の水平面の実測HRTFを $HRTF_{org}$ とする。まず、 $|HRTF_{org}|$ の低域を左右とも0 dBとする。そして、 $HRTF_{org}$ の低域成分から算出されるITD(以下 ITD_{low} という)を、 $HRTF_{org}$ の高域から算出されるITD(以下 ITD_{high} という)の a 倍($a=0, 0.5, 1, 1.2, 1.5, 2.0$)に設定した6種の $HRTF_{mod}$ を作成する。刺激音は、これらの $HRTF_{mod}$ を白色雑音に畳みこんだ合成バイノーラル音である。

具体的な $HRTF_{mod}$ 作成手順をFig. 1に示す。まず、 $HRTF_{org}$ の時間領域表現である左右のインパルス応答 $h_{rirL}(t)$ と $h_{rirR}(t)$ をHPF($f_c=1$ kHz, FIR)に通して $h_{rirL}(t)_{high}$ 、 $h_{rirR}(t)_{high}$ を得、それらの時間差から ITD_{high} を算出する。次に、単位インパルス信号 $\delta(t)$ を、 ITD_{high} の a 倍の時間だけシフトした $irL(t)$ と $irR(t)$ を算出する。そして、 $irL(t)$ 、 $irR(t)$ をLPF($f_c=1$ kHz, FIR)に通して $irL(t)_{low}$ と $irR(t)_{low}$ を得る。最後に、 $h_{rirL}(t)_{high}$ と $irL(t)_{low}$ 、 $h_{rirR}(t)_{high}$ と $irR(t)_{low}$ を加算し、 $HRTF_{mod}$ に対応したインパルス応答

を得る。以下、低域のITDを高域のITDの a 倍にした $HRTF_{mod}$ を $HRTF_{moda}$ と表す。

2.2 実験システム

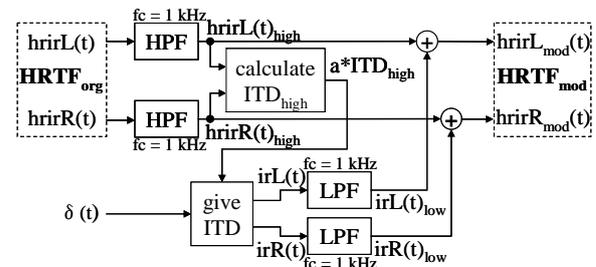
刺激音は、D/A変換器(UA-101, Roland)から出力し、ヘッドホンアンプ(AT-HA20, audio-technica)を通してヘッドホン(HDA200, Sennheiser)から呈示した。サンプリング周波数は44.1 kHz、量子化精度は24 bitとした。

刺激音の持続時間は3 s、ISIは7 s、正面の $HRTF_{org}$ を白色雑音に畳みこんだバイノーラル信号の音圧レベルを70 dBとした。

2.3 実験方法

頭部静止条件では、 $HRTF_{org}$ と6種類の $HRTF_{mod}$ について、正面から45°間隔の水平面8方向を用いて合成した計56種類のバイノーラル音を用意した。1セッションではそれらを1回ずつランダムに呈示し、被験者は4セッションの実験を行った。被験者には、音像位置を、8方位と3段階の距離(頭部表面、頭部近傍、遠方)あるいは頭内、に強制選択させた。また、音像が2つ聴こえた場合には高域音と低域音をわけて回答させた。なお、刺激音の呈示中は、被験者に頭部を静止するよう教示した。

頭部運動条件では、各 $HRTF_{mod}$ を頭部の動きに応じて合成した動的バイノーラル音^[4]を呈示する実験を行った。被験者は前面、右側面、背面、左側面の順に呈示される刺激音を、頭部を自由に回転させながら何度でも聴取し、音像がどのように聴こえたかを $HRTF_{mod}$ 毎に回答した。

Fig. 1 $HRTF_{mod}$ の作成手順

* Horizontal sound localization of binaural signals whose low-frequency ITD was manipulated, by MORIKAWA Daisuke and HIRAHARA Tatsuya (Toyama Prefectural University).

3 実験結果

Fig. 2 は頭部静止条件で高域音と低域音の音像を分離定位した割合を描いたものである。どちらの被験者も、高域音と低域音の音像を分離定位する割合は $HRTF_{mod}0$ では高く、 $HRTF_{mod}1.5$ では低かった。

Fig. 3 は頭部静止条件で音像を知覚した距離の割合を描いたものである。高域音と低域音の音像を分離して定位した場合は低域音の距離を示してある。どちらの被験者も、最も遠方に定位する割合が高かったのは $HRTF_{mod}1.5$ で、低域の ITD が短くなるにつれて、遠方に定位する割合は減少した。

頭部運動条件では、どちらの被験者も、 $HRTF_{mod}0 \sim HRTF_{mod}1.0$ に対しては高域音と低域音の音像を分離定位し、高域音の音像は頭外に、低域音の音像は頭内に定位した。一方、 $HRTF_{mod}1.2 \sim HRTF_{mod}2.0$ に対しては1つの音像を頭外に定位した。なお、条件毎の音像の距離は被験者で異なった

なお、頭部静止条件における低域音の音像はあまりはっきりしないが、頭部運動条件における低域音の音像ははっきりわかる、という内観報告を sub.1 から得た。また、 $HRTF_{mod}0$ に対しては頭外に定位する高域音の反対側に低域音が定位するという報告を、sub. 1 では頭部静止条件、sub.2 では頭部運動条件で得た。

4 考察

頭部静止条件、頭部運動条件どちらの結果からも、低域の ITD を高域の ITD の 1.5 倍に長くした $HRTF_{mod} 1.5$ 前後がもっとも自然な頭外定位を与える $HRTF$ であった。これは、実際の ITD が低域になるほど長いこと^[1, 2]に関係すると考えられる。

$HRTF_{mod}0$ は低域の ILD が 0 dB、ITD も 0 s であるにもかかわらず、低域音は頭内中心に定位せず、高域音と反対側に定位した。これは、2つの音を同時に呈示すると、ターゲットの音がもう一方の音から離れる方向にシフトするという藤井らの結果^[5]と一致する。

5 まとめ

$HRTF$ の低域 (1 kHz 以下) の ITD を系統的に操作した変形 $HRTF$ で合成したバイノーラル音を用いて、頭部静止・運動条件で、音像定位実験を行った。その結果、低域の ITD を高域 (1 kHz 以上) の ITD の 1.5 倍程度の

長さにすることにより、合成バイノーラル音の音像は自然に頭外定位することがわかった。

謝辞

本研究の一部は科研費(22300061)の助成を受けた。

参考文献

- [1] 平原 他, “両耳間時間差の周波数依存性の成因—音響計測編—,” 音講論(春), 2012.
- [2] 大谷, 平原, “両耳間時間差の周波数依存性の成因—数値解析編—,” 音講論(春), 2012.
- [3] J. Pickles, *An Introduction to the Physiology of Hearing (Second Edition)*, Academic Press, 1988.
- [4] 大谷, 平原, “Windows 上で動作する動的聴覚ディスプレイ,” 音講論(春), 711-712, 2007.
- [5] 藤井, 寛, “2音同時呈示条件における音源定位の偏向の規定因,” 信学技報, TL2001-21, 39-44, 2001.

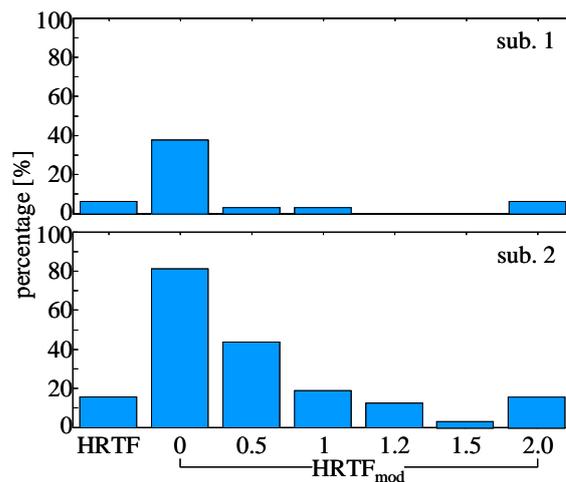


Fig. 2 音像を分離して定位した割合

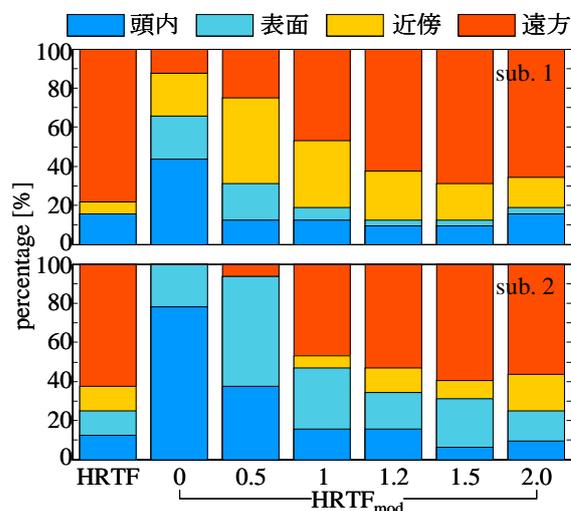


Fig. 3 音像距離の判定結果の割合