

バイノーラル音を処理する聴覚の仕組み

平原達也 (富山県立大学・工学部)

1. はじめに

バイノーラル (binaural) とは両耳で音を聴くことである。私たちの多くは日常の音をバイノーラル受聴しているし、一つのラウドスピーカから音を出すモノフォニック方式や二つのラウドスピーカから音を出すステレオフォニック方式の再生音を耳で直接聴く場合もバイノーラル受聴である。また、両耳イヤホンを用いて再生される音を聴く場合もバイノーラル受聴である。

1 チャンネルのモノラル信号や左右のマイクロホンで収録した 2 チャンネルのステレオ信号をバイノーラル再生しても、音像は頭の中にしか聴こえない。一方、ダミーヘッドの耳に設置した二つのマイクロホンで収録したり、モノラル信号に左右耳の頭部伝達関数 (HRTF: head-related transfer function) を畳みこんで合成した 2 チャンネルのバイノーラル信号をバイノーラル再生すると頭外に音像が聴こえ、その 3 次元的な位置が判る。

本稿では、バイノーラル音に含まれる音響的特徴と、それらがどのように皮質下の聴覚神経系で処理されて、立体的な音の聴こえに脳が必要とする情報を「計算」しているかについて概説する。

2. バイノーラル音の音響的特徴

バイノーラル音の音響的特徴は、頭部の存在によって左右チャンネル間に生じる両耳間時間差 (ITD: interaural time difference) と、両耳間音圧差 (ILD: interaural level difference) と、スペクトルの特徴 (SF: spectral feature) である。

ITD は、通常、左右チャンネル信号波形の相互相関から算出される信号全体の時間差である。ITD を支配するのは、頭部を回りこむ低周波成分である。ヒトの両耳間距離は約 150 mm で、それに π を乗じた両耳間の測地線距離 236 mm を音速で除した約 $\pm 700 \mu\text{s}$ が ITD の最大値である。

ILD は、通常、バイノーラル収録した左右チャンネル信号波形のパワーの差、あるいは HRTF の振幅スペクトルの差から算出される信号全体の音圧差である。ILD を支配するのは、頭部により遮蔽される高周波成分である。ヒトでは ILD の最大値は約 $\pm 16 \text{ dB}$ である。

SF は HRTF の振幅スペクトルに現れるスペクトルピークとディップパターンである。SF を支配するのは、肩や耳介の凹凸による反射に参与する 1 kHz 以上の周波数成分である。

ITD も ILD も SF も音源位置の関数となる。

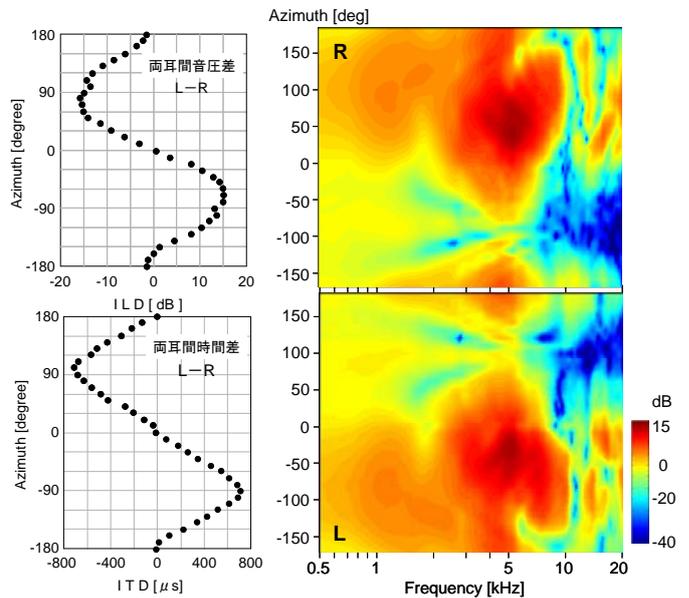


図 1 実頭で測定した水平面頭部伝達関数 (HRTF) のコンター図 (右) と、それより算出した両耳間音圧差 ILD (左上) と両耳間時間差 ITD (左下)

3. バイノーラル音の聴こえ

適切な ITD と ILD を持つバイノーラル信号を両耳イヤホンで再生すると立体的な音が再生できるといわれている。しかし、話はそう単純ではない。

白色雑音信号 WN に実頭で実測 ITD を付与し ILD を 0 とした信号 WN_{ITD} と、WN に実測 ILD を付与し ITD を 0 とした信号 WN_{ILD} の音像は頭内で左右の耳の間を往復するだけである。WN に実測 ITD と ILD を付与した信号 WN_{ITLD} の音像も頭の中で動くだけである。WN に実測 HRTF を畳みこんだ信号 WN_{HRTF} は頭外に定位する。

ただし、 WN_{HRTF} においても、正面附近の音像は頭外定位しても距離感が出ず、頭部近傍に定位することが多い。その理由は定かではないが、正面は視覚情報を利用できるので聴覚情報だけで方位と距離を正確に計算する必要がないからとも言われている。腹話術効果として知られるように、音像の定位は音響情報だけで決まるわけではない。視覚が使える場合、音像と映像の方向が 10° 以下の場合、音像は映像によって大きく引き寄せられ、 20° 以上になると映像による音像の方向支配の程度は著しく減少する [1]。

なお、HRTF を畳みこむということは、単に SF を与えるということではなく、各周波数における ITD と ILD をも同時に付与するということである。

3. バイノーラル音の神経処理

左右の耳に届いた音は、まず、蝸牛の基底膜振動系が構成する機械的な帯域通過フィルタ (BPF) で周波数分析される。この BPF を聴覚フィルタと呼ぶ。基底膜の各場所の振動は、基底膜上に載る 3,500 個の内毛細胞の電位変化に変換される。その電位変化は、内毛細胞に接続する約 3 万個の一次聴神経 (AN) の発火パターンとして神経符号化される。各 AN の発火頻度は、入力音のワースペクトル情報を、各 AN の発火間隔は入力音圧波形の周期情報を表現する。ただし、AN は 3 ~ 4 kHz 以下の低周波数成分波形にしか位相同期発火しない。AN は蝸牛神経核 (CN) の複数の神経細胞に投射している。

蝸牛神経核前腹側核 (AVCN) の神経細胞 (Busy Cell) は、AN の発火間隔のジッターを減らし、より正確に位相同期した発火パターンを両側の上オリブ複合体の神経細胞に中継する。蝸牛神経核後腹側核 (PVCN) と背側核 (DCN) の神経細胞 (Stellate cell, Giant cell, Fusiform cell) は、AN の発火パターンの周波数・時間的特徴を強調あるいは抽出して対側の下丘中心核 (ICC) へ送る (図 2)。特に、ネコなど耳介を随意に動かせる動物の DCN にある Fusiform Cell はスペクトルの谷を検出して、耳介の動きに伴う HRTF 上の高域のスペクトルの谷の位置変動を補正し音源定位に役立っている [2]。耳介が動かないヒトの DCN は退化している。

上オリブ複合体主核は外側核 (LSO) と内側核 (MSO) で、両側の蝸牛神経核から投射を受け、ILD と ITD を聴覚フィルタの帯域ごとに計算し、その結果を主として下丘中心核 (ICC) へ送る。

LSO には高い最適周波数を持つ神経細胞が多く、主として高周波数成分の発火頻度情報から ILD を「計算」している。その原理は、同側 AVCN の小型 Spherical Busy Cell からの信号と、抑制性の台形体内側核 (MNTB) を経由して受ける対側 AVCN の Globular Busy Cell からの信号を LSO で加算するというものである。つまり、左右の LSO では L-R および R-L の ILD 情報が発火頻度で表現される (図 3)。

MSO には低い最適周波数を持つ神経細胞が多く、主として低周波数成分の位相同期した発火パターンから ITD を「計算」している。鳥類では、遅延線と一致検出器を利用して ITD を検出する多数の神経細胞があり、どの ITD を検出する神経細胞が発火したかという場所情報で ITD を表現している。一方、哺乳類では、MNTB から MSO に対する精密なタイミングでの抑制を利用して、MSO 神経細胞が ITD に応じた頻度で発火する。つまり、哺乳類の MSO では少数の神経細胞の発火頻度で ITD を表現している [3] (図 4)。

下丘では、各聴覚フィルタの周波数帯域を越え

た情報の統合が行われ、ITD の位相多義性の解消に関連したニューロンもある。上丘には視覚と聴覚から得られる空間地図があると言われている。

聴覚皮質の傍帯状領域 (parabelt area) には周波数選択性が低く、音の方向や音源の動きに対して応答するニューロンがある。そして、聴覚皮質から後部頭頂野へ向かう経路は音源定位に係わる聴覚の "Where" 経路であることが確認されている [4]。

頭部運動との相互作用など、これ以上詳細なバイノーラル音の神経処理の仕組みは藪の中である。

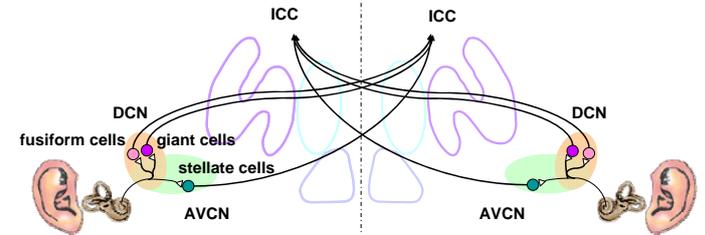


図 2 バイノーラル音のスペクトル的特徴処理経路

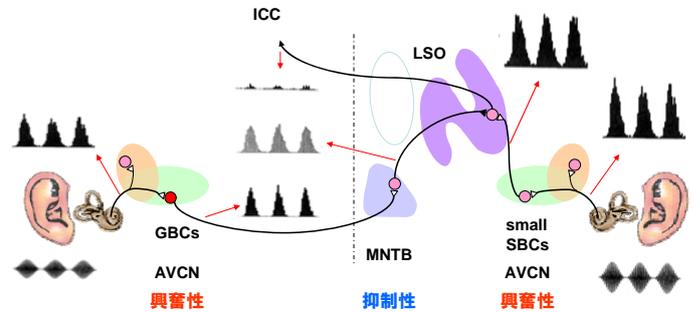


図 3 バイノーラル音の ILD 処理経路

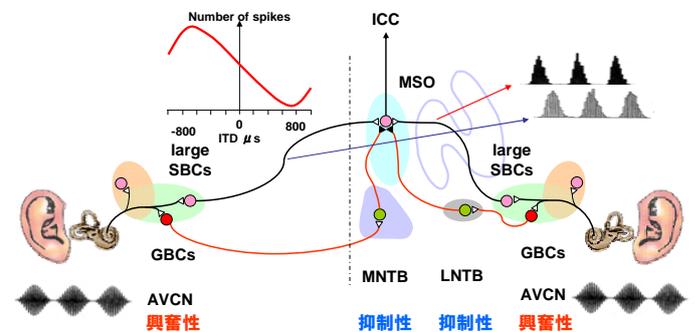


図 4 バイノーラル音の ITD 処理経路

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] 中林克己, "ステレオ音像とテレビ映像の相乗効果," テレビジョン学会誌 37(12), 984-991, 1983.
- [2] D.Oertel and E.D.Young, "What's a cerebellar circuit doing in the auditory system?" TRENDS in Neurosciences 27(2), 104-110, 2004.
- [3] D.McAlpine and B.Grothe, "Sound localization and delay lines – do mammals fit the model?" TRENDS in Neurosciences 26(7), 347-350, 2003.
- [4] J.P.Rauschecker and B.Tian, "Mechanisms and streams for processing of 'what' and 'where' in auditory cortex," Proc. Natl. Acad. Sci. USA, 97, 11800-11806, 2000.