

研究論文

頭部伝達関数におけるノッチ幅と中心周波数の偏移が 音色知覚と音像定位に及ぼす影響

山高 正烈*, 坂本 修一**, 鈴木 陽一**

(2021年9月30日受理)

Effect of notch width and center frequency deviation in human head-related transfer function on timbre perception and sound localization

Zhenglie CUI*, Shuichi SAKAMOTO**, Yôiti SUZUKI**

(Received September 30, 2021)

Abstract

The frequency spectrum of the head-related transfer function (HRTF) plays an important role in sound localization, particularly in the median plane. The timbre of sounds is also closely related to their frequency spectra. The perception of the position of sound sources is called sound localization. Interestingly, humans can separately perceive both the timbre and position of sound sources even though their frequency spectra are mixed as they reach the ears. This study, we focus on the spectral notches in the HRTF (N1), the central frequency of which may systematically depend on elevation angle. Our study also investigated the influence of the notches on sound localization and timbre perception when the bandwidth of N1 and center of frequency changed. The results show that the increase in bandwidth of N1 is perceptible as a change in timbre. Yet the influence of the center frequency deviation of N1 on the sound localization is small. It suggests sound localization are affected only when the center frequency deviation of N1 is relatively large.

キーワード: 音空間知覚, 頭部伝達関数, 音像定位, 音色, 知覚手がかり

Keywords : Spatial hearing, HRTF (head-related transfer function), Sound localization, Timbre, Perceptual cue

1. はじめに

我々人間は音を聞く際に、音色と音の到来方向を同時に知覚できる。音色は、音源の周波数スペクトルと密接に関係しており[1-3]、音を聞いた主体が音から受ける印象の総称と定義される[4]。一方、音の到来方向を含む3次元空間位置の知覚を音像定位[5]と呼んでおり、人間が音の空間特性を知覚する能力と定義される。耳に到達する音信号には、音源からの直接音だけでなく反射音や回折音等が含まれており、耳介周りの外耳道や耳甲介腔による共振も生じている[6]。このため、両耳で観測される音の周波数スペクトルは音源の相対方向に応じて変化する。この周波数スペクトルの変化を伝達関数として表現したものを頭部伝達関数(HRTF: head-related transfer

function[7])と呼ぶが、頭部伝達関数は音像定位の重要な手がかりとなっている。特に、両耳に入力される音の時間差や強度差がほとんどない正中面や、その変化があまりない矢状面上(両耳からの距離差が一点の円上, cone of confusion [8])に位置する音源の音像定位については、頭部伝達関数のスペクトル形状が非常に重要な手がかりとされている(以降、スペクトラルキューと称する)。

このように、人間の耳に入力される音信号には、音色の手がかりと音像定位の手がかりが混在した状態であり、両方とも耳入力信号の周波数特性が手がかりであると考えられている。しかし、人間は音色と音像の双方を明確に区別して知覚することが可能であり、音の方向が変わっただけで別の音色に聞こえることを日常生活で経験することはあまりない。即ち、人間は耳入

* 愛知工科大学工学部情報メディア学科, 〒443-0047 愛知県蒲郡市西迫町馬乗 50-2

Department of Media Information, Aichi University of Technology, 50-2 Manori, Nishihasama-cho, Gamagori 443-0047, Japan

** 東北大学電気通信研究所, 〒980-8577 仙台市青葉区片平 2-1-1

力信号から音色と音像定位を何らかの仕組みで切り分けており、それぞれの情報を別々に利用していると考えられる。この仕組みの解明は、頭部伝達関数に含まれる音像定位の手がかり、スペクトラルキューを調べる上で極めて有用であると考えられる。

過去の研究では、2 kHz以上の周波数帯域における頭部伝達関数の概形を再現すれば高精度の音像定位が可能であり、そのスペクトラルキューとして、仰角によって系統的に変化する複数帯域のレベル差が示唆された[9]。また、仰角によって系統的に変化する7 kHz付近に現れるノッチ（以下N1）がスペクトラルキューであるという報告もなされている[8]。しかし、聴覚系における周波数分析機構を司る聴覚フィルタの特性を考慮すると、N1 のように鋭いスペクトル上のノッチは、元の形状よりもはるかに緩やかで浅い興奮パターンを示すと考えられる[10]。したがって、音像定位の手がかりはN1 そのものの形状やその変化ではなく、興奮パターンの形状にあると考えるべきである。実際、頭部伝達関数の2 kHz以上の帯域を平滑化しても音像定位が可能であるという報告や[9]、8 kHzを中心とする3/4 オクターブ帯域において、それ以外の帯域とは仰角の変化に伴う帯域のレベルの変化傾向が異なる様相を呈することを示した報告もなされている[11]。

このように、音像定位の工学的制御にN1が重要であることは明らかであるが、それに起因する聴覚興奮パターン上のスペクトラルキューが何なのかについては統一的な理解が得られていない。我々の先行研究[12-14]では、広帯域雑音中に付加したN1ノッチフィルタの帯域幅や中心周波数を系統的に変化させ、それに対応する音色変化の検知限と聴取者自身の頭部伝達関数を比較することで、音色と音像定位のスペクトラルキューについて検討した。その結果、N1の中心周波数の変化は、音色の変化としては知覚されるが、仰角方向の定位方向に充分寄与するとは言えないことが示唆された。

しかし、先行研究[12-14]においては、音色弁別が充分可能なノッチ幅を用いてN1中心周波数の偏移による音色及び音像定位への影響を調べた。そのため、音色と音像定位の弁別閾の区別ができておらず、N1のノッチ幅や中心周波数の偏移が、音色に影響するのかそれとも音像定位に影響し得るのかを示すことができなかつた。先行研究の結果、N1中心周波数が7 kHzの場合の音色弁別閾は1.37 ERBであった。そこで本研究では、これらの先行研究の知見を踏まえ、音色弁別が不可能と考えられるN1帯域幅=0.5 ERBの場合と、音色弁別が充分可能なN1帯域幅=3.0 ERBを実験条件として設定し、N1中心周波数を7～

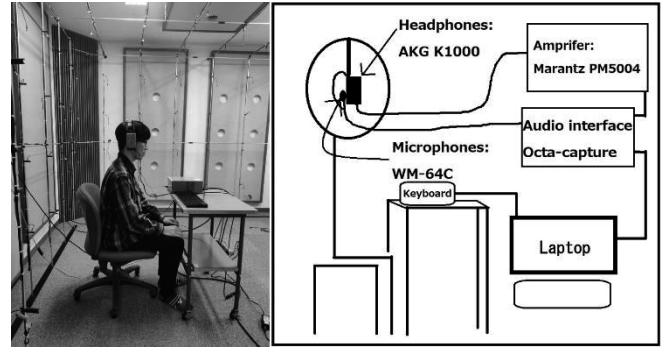


Fig.1 Experimental setup

11 kHzに系統的に変化させた場合の、N1帯域幅と中心周波数の偏移が音色及び音像定位に及ぼす影響を調べた。

2. 実験1(音色知覚実験)

2.1 目的

実験1ではピンクノイズにN1を模擬した反共振形ノッチを付加し、その帯域幅と中心周波数の偏移が音色弁別閾にどのような影響を及ぼすかを検討した。

2.2 聴取者と実験環境

正常な聴力を有する、20代の成人男性4名が聴取実験に参加した。実験環境の概略をFig.1に示す。実験は東北大学電気通信研究所内の防音室で行った。音の提示には耳介開放型ヘッドフォン(SENNEISER, AKG K1000)を用いた。実験では、先行研究[12-14]に習って、聴取者の耳に測定用小型マイクロフォン(Panasonic, WM-64C)を耳珠の真横に設置した。これは、ヘッドフォンから外耳道入口にかけての伝番特性を打ち消すための逆フィルタ作成に必要なインパルス応答の計測[15]、及び、実験中における提示刺激の観測を行うためであった。

2.3 実験刺激

実験は、2つの音を連続して提示し比較する方法で実施した。以降、先に提示する音を参照音、後に提示する音を比較音と呼ぶ。音刺激は、参照音と比較音ともに、周波数が1オクターブ増加するごとにパワーが半減する特徴を持つピンクノイズにノッチを加えて生成した。ノッチフィルタ[16]の帯域幅と深さは参照音と比較音で共通しており、帯域幅は0.5 ERBと3.0 ERBのいずれかで、深さは20 dBの値を用いた。

中心周波数(f_0)については、参照音と比較音で異

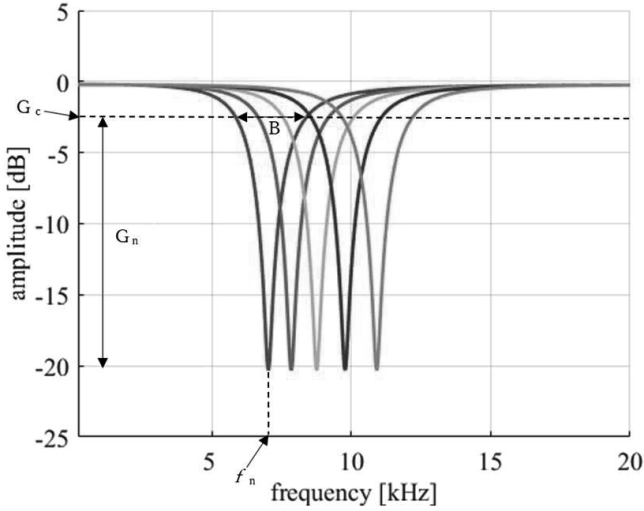


Fig. 2 Shape of N1 notch filter

なる値を用いており，参照音では実験を通じて $f_0 = 7$ kHzで固定した．比較音には5種類の値を用いた．1つは参照音と同一の中心周波数 f_0 で， f_0 から1 ERBずつらし，それぞれ $(f_0 + 1)$ ERB, $(f_0 + 2)$ ERB, $(f_0 + 3)$ ERB, $(f_0 + 4)$ ERBの値を用いた．以下，N1ノッチフィルタを例に，フィルタの設計について説明する．設計に当たっては， g , r , ω_0 をパラメータとした式(1)で表される2次のIIRフィルタを用いて反共振ノッチフィルタを作成した．ここで f_n , G_c , G_n , B は，それぞれ中心周波数[Hz]，帯域幅を定めるカットオフのレベル[dB]，ノッチの深さ[dB]， G_c における帯域幅[Hz]を表す．なお， f_s はサンプリング周波数である．深さ G_n は-20 dB，帯域幅はノッチ両側の平坦な帯域のレベルから-3 dB低下する周波数間の値であり，+1.5 ERBに固定した．作成した刺激の形状をFig.2に示す．

$$H(z) = \frac{1-2rcos\omega_0 z^{-1}+r^2 z^{-2}}{1-2\frac{g}{r}cos\omega_0 z^{-1}+\frac{2g-1}{r^2} z^{-2}} \quad (1)$$

$$g = \frac{1}{1+\tan\left(\frac{B_0}{2}\right)\frac{\sqrt{1-A_c^2}}{A_c}}$$

$$r = 1 - A_n \frac{B_0}{\pi}, \quad \omega_0 = 2\pi \frac{f_n}{f_s}$$

$$A_c = 10^{\frac{-G_c}{20}}, \quad A_n = \frac{1}{A_c} 10^{\frac{-G_n}{20}}, \quad B_0 = 2\pi \frac{B}{f_s}$$

2.4 実験手続き

実験で使用したピンクノイズは同一試行内では同じものを使用し，試行ごとに生成しなおした．これは1試行内でノッチが生じた場合に音色の違いを知覚させるためであった．実験では参照音と比較音をラン

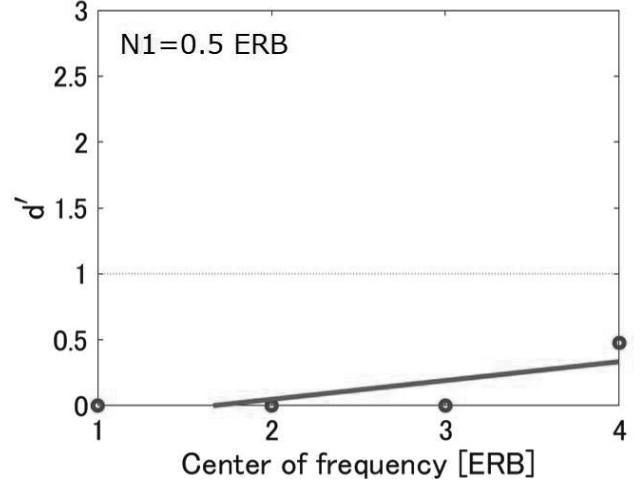


Fig. 4 Timbre discrimination performance (d') for N1=3.0 ERB

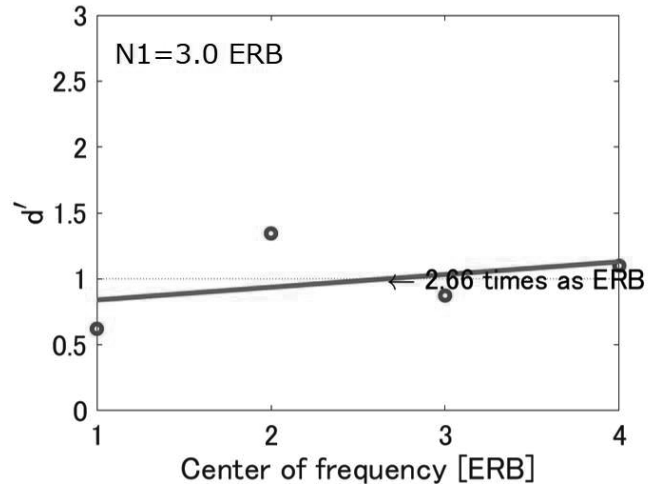


Fig. 3 Timbre discrimination performance (d') for N1=0.5 ERB

ダムな順に 0.5 秒間隔で連続して提示した．これを1試行とし，1セッションに40回行い計4セッション(40×4=160試行)実施した．

4セッション中に参照音と異なる4種類の比較音を各20試行ずつ提示し，残りの80試行は各セッションで20試行ずつ参照音と同一の比較音を用いた．被験者には，最後に提示した音が前の音と比べて音色が同じか異なっていたかを判断させた．

2.5 実験結果と考察

N1帯域幅が0.5 ERB及び3.0 ERB場合の実験結果を，それぞれFig.3とFig.4に示す．図は，聴取者4名のN1中心周波数の偏移に対する音色知覚の平均弁別閾を表しており，弁別閾は $d' = 1$ となる周波数偏

移とした。

d' [17]の算出に当たっては、4セッションの実験結果に対し、N1中心周波数が高くなるにつれてそれぞれの音色弁別の正答率を求めた。グラフの横軸はN1中心周波数の偏移量を $ERB_{Nnumber}$ スケール[14]で示した値であり、単位はCamである。グラフの縦軸は d' である。グラフ中の●点は各中心周波数の偏移パターンにおける実験結果を表しており、実線は最小二乗回帰直線である。

実験結果より、N1帯域幅=0.5 ERBの場合は、中心周波数を偏移しても d' が1を上回らないことが見て取れる(Fig.3参照)。即ち、先行研究[14]と同様に、N1=0.5 ERBは音色の変化を弁別できないほど小さい帯域幅であることが示された。一方、N1帯域幅=3.0 ERBの場合は、 d' が1を上回る結果となった(Fig.4参照)。これから、帯域幅の増加自体が音色知覚に影響を及ぼした可能性が示唆される。中心周波数の影響については、中心周波数の偏移が2.66 kHzのときに d' が1を上回っており、グラフ全体を見ると中心周波数の偏移による音色弁別への影響が限定的であることが分かる。以上の結果より、音色知覚には、中心周波数の偏移よりもノッチ帯域幅の変化による影響が大きいと考えられる。

3. 実験 2(音像定位実験)

3.1 実験目的

実験 2 では、広帯域ピンクノイズに付加したN1の帯域幅と中心周波数の偏移が、仰角方向の音像定位にどのような影響を及ぼすかを検討する。

3.2 実験内容

被験者と実験環境は実験 1 と同様であった。実験においては、音像定位の変化を仰角方向に限定し、音色が異なって知覚した場合であってもそれを無視し、参照音に対して比較音の仰角方向の音像が上方に偏移したか、それとも下方に偏移したかを回答するように求めた(2肢強制選択法)。ノッチの帯域幅は、実験 1 と同様に参照音、比較音ともに 0.5 ERBと 3.0 ERBの 2 条件を用意した。

3.3 実験結果と考察

聴取者 4 人の、比較音が参照音と比べて上方に偏移したと回答した割合を算出し、Fig.5 とFig.6 に示す。グラフの横軸は各中心周波数の偏移量を表しており、縦軸は参照音と比べて比較音が上方に偏移したと回答した割合を表す。なお、ここでは頭部伝達関数の特

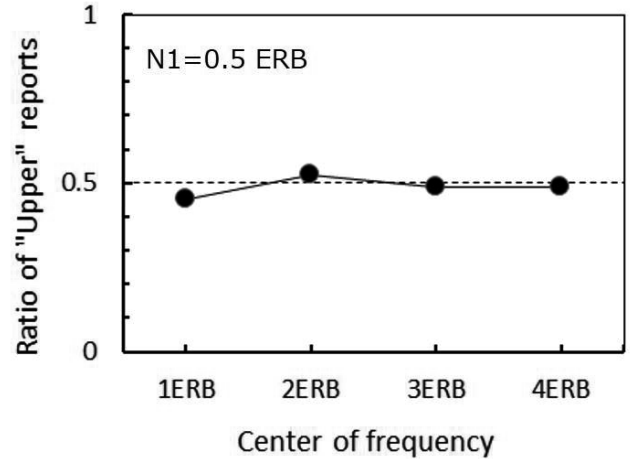


Fig. 5 Sound localization discrimination for N1=0.5 ERB

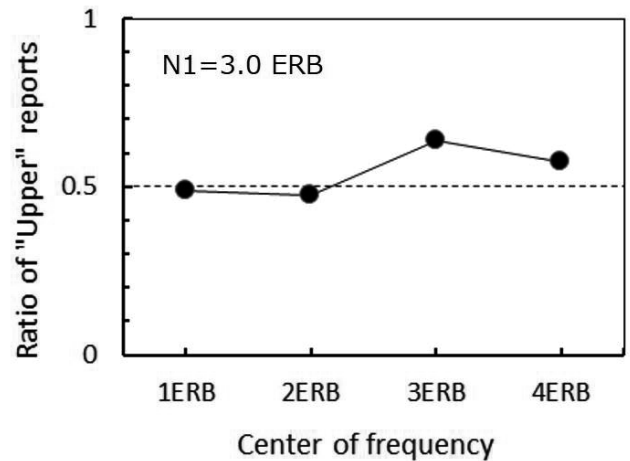


Fig. 6 Sound localization discrimination for N1=3.0 ERB

徴に習って、音像が上方に定位するという仮定の下で正答率を求めた。

実験結果、N1帯域幅=0.5 ERBの場合、参照音に対して比較音が上方に偏移したと回答した正答率が、どの中心周波数偏移条件においても 50% 付近であることが見て取れる(Fig.5 参照)。即ち、N1帯域幅が 0.5 ERBの場合は、仰角方向の音像定位ができていないことが示唆された。一方、N1帯域幅が 3.0 ERBで中心周波数が 3~4 ERBと比較的大きく偏移した場合には正答率が概ね 0.5 を超えており、十分な弁別閾とは言えないものの音像を上方に偏移したと回答した割合が増加傾向にあることが見て取れる(Fig.6 参照)。これらの結果から、N1の帯域幅を 3.0 ERBと固定したまま中心周波数の偏移を更に大きくすれば、仰角方向の音像定位が可能であると考えられる。

4. 総合考察

実験 1 の結果, 音色知覚にはノッチの中心周波数の偏移よりも帯域幅が影響をし得る可能性が示唆された. 一方, 実験 2 からは, N1 の帯域幅を 3.0 ERB に固定し中心周波数を大きく偏移した場合に, 音像を上方に偏移して知覚する可能性が示唆された. 即ち, 音色に充分影響し得るノッチ帯域幅 3.0 ERB を用いて, その中心周波数を 4 ERB 以上と大きく偏移させた場合に, 正中面の音像の変化を知覚できる可能性が考えられる.

これらを確かめるために, 実験 2 の結果を基に最小弁別値(Just Noticeable Difference: JND) を計算しフィッティング曲線を求めた. ここで, JND は「元とは異なる音刺激を提示した際に元の音刺激との変化を知覚できる最小の値」を意味する. 計算結果, JND の値は 5.97 ERB であった. この計算結果から, ノッチの帯域幅 3.0 ERB 条件では, 中心周波数の偏移が約 6.0 ERB 以上であれば音像定位が可能であることが推測される. 即ち, 中心周波数の偏移量が比較的大きい場合のみ, 正中面の音像変化を知覚できると考えられる.

4. まとめ

本研究では, 音色知覚と音像定位知覚のスペクトラルキューの解明を目指して, 頭部伝達関数の周波数スペクトルに典型的に現れるノッチに着目し, N1 を模擬したノッチフィルタを用いてその帯域幅の変化と中心周波数の偏移を様々な値に変化させ, これらの変化が音色知覚と仰角方向の音像定位にどのような影響を及ぼすかを検討した.

実験 1 では, N1 ノッチ幅を 0.5 ERB と 3.0 ERB の 2 条件に設定し, 中心周波数を偏移させながら音色知覚実験を実施した. その結果, 3.0 ERB 条件に限って, 中心周波数を偏移させた場合に音色の違いを弁別できることが分かった. 実験 2 では, 実験 1 と同様のノッチ条件を用いて音像定位実験を実施し, N1 ノッチ幅や中心周波数の偏移が仰角方向の音像定位にどのような影響を及ぼすかを検討した. 実験結果, N1 = 0.5 ERB 条件の場合には, 中心周波数を偏移させても音像定位ができていなかった. しかし, N1 = 3.0 ERB 条件の場合には, 中心周波数を 3~4 ERB と大きく偏移させたときに音像定位が可能であることが示唆された. 実験 2 の結果を基に JND を求めた結果, 中心周波数が約 6.0 ERB 以上であれば仰角方向の音像定位が可能であろうと考えられる.

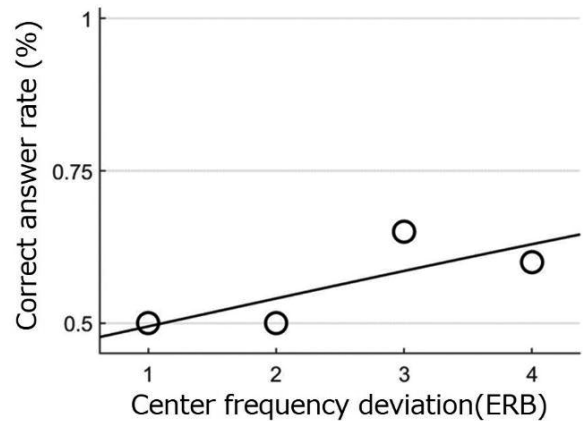


Fig. 6 Regression lines for sound localization (N1=3.0 ERB)

これらの結果から, 中心周波数の偏移による音像定位への影響は小さく, 偏移幅が比較的大きい場合のみ音像定位に影響を及ぼしている可能性が示唆された. しかし, N1 ノッチ幅や中心周波数の偏移による音色知覚への影響については明らかになったものの, 音像定位への寄与は限定的であった. 今後は, N1 の操作のみではなく, 頭部伝達関数に広く現れる P1 や N2 を含めた幅広い特徴量を模擬して同様に検討を行い, 正中面音像定位に及ぼす周波数スペクトラルキューの特定を目指したいと考えている.

謝辞

本研究の一部は JSPS 科研費 19H02396 の助成を受けた. 本研究の聴取実験にご協力いただいた愛知工科大学の春日部剛士君に感謝する.

参考文献

- [1] J. R. Miller and E. C. Carterette, "Perceptual space for musical structures," J. Acoust. Soc. Amer., vol.58, no.3, pp.711-720, 1975.
- [2] J. Grey, "Multidimensional perceptual scaling of musical timbres," J. Acoust. Soc. Amer., vol.61, no.5, pp.1270-1277, 1977.
- [3] J. Grey and A. Moorer, "Perceptual evaluations of synthesized musical instrument tones," J. Acoust. Soc. Amer., vol.62, no.2, pp.454-462, 1977.
- [4] 日本音響学会, "音響用語辞典," コロナ社, 2003.
- [5] J. Blauert, "Spatial Hearing- The Psychophysics

of Human Sound Localization,” Revised edition (MIT Press, Cambridge, Mass.), 1996.

[6] E. A. G. Shaw, “Transformation of sound pressure level from the free field to the eardrum in the horizontal plane,” *J. Acoust. Soc. Am.*, vol.56, no.6, pp.1848-1861, 1974.

[7] M. Morimoto and Y. Ando, “On the simulation of sound localization,” *J. Acoust. Soc. Jpn. (E)*, vol.1, pp.167-174, 1980.

[8] K. Iida, M. Itoh, A. Itagaki and M. Morimoto, “Median plane localization using a parametric model of the head-related transfer function based on spectral cues,” *Applied Acoustics*, no.68, pp.835-850, 2007.

[9] F. Asano, Y. Suzuki and T. Sone, “Role of spectral cues in median plane localization,” *J. Acoust. Soc. Am.*, no.88, pp.159-168, 1990.

[10] 北島周, 鈴木陽一, 坂本修一, “周波数振幅スペクトル上のノッチによる音色変化の検出限と弁別限,” *日本音響学会講演論文集*, 2-P-9, pp.819-820, 2015.

[11] 曲谷地哲, 岩谷幸雄, 鈴木陽一, “仰角依存性の強い帯域のレベルを模擬した頭部伝達関数による正中面定位,” *聴覚研究会資料*, vol.42, no.1, pp.47-52, 2012.

[12] 齋藤優一, 鈴木陽一, 坂本修一, “ノッチフィルタが広帯域雑音の音色弁別に及ぼす影響,” *日本音響学会聴覚研究会資料*, pp.77-82, 2016.

[13] 日吉啓, 鈴木陽一, 坂本修一, “広帯域ノイズ中のノッチ中心周波数の偏位が音色と音像定位に与える影響,” *電子情報通信学会技術研究報告*, EA2018-44, pp.103-108, 2018.

[14] 山高正烈, 坂本修一, 鈴木陽一, “ノッチフィルタが広帯域雑音の音像定位に及ぼす影響,” *愛知工科大学紀要第18巻*, pp.21-26, 2021.

[15] Y. Suzuki, F. Asano, H. Y. Kim, T. Sone, “An optimum computer-generated pulse signal suitable for the measurement of very long impulse response”, *J. Acoust. Soc. Am.* vol.97, no.2. pp.1119-1123, 1995.

[16] 川浦純一, 鈴木陽一, 曾根敏夫, 相馬次郎, “悪条件の伝達関数の逆特性推定のための新しい方法,” *電子通信学会論文誌*, J69-A, pp.914-917, 1986.

[17] G.A. Gescheider, “Psychophysics: the fundamentals,” Psychology Press, 2013.