

正中面定位における頭部伝達関数の平滑化に関する2,3の考察*

○飯田一博(松下通信・AV&C研) 森本政之(神戸大・工)

1. はじめに

頭部音響伝達関数(以降HRTFsと記述)を再現すれば、実際の音源方向とは無関係に正中面内の任意の方向に音像をシミュレートできることが報告されている[1]。しかし、HRTFsは(特に仰角の知覚にとって重要な高域において)個人差が大きく、従来の知見では本人のHRTFsを使わない限り、精度よく正中面内の音像方向をシミュレートすることはできない。本研究では、複数の受聴者に有効な(個人のHRTFsを使わない)正中面内の音像シミュレーションの第一ステップとして、HRTFsに聴覚系の信号処理メカニズムを考慮した平滑化を導入し、まず個人内の正中面定位の精度を検証した。

2. HRTFsの平滑化方法

2.1 シミュレーションの原理

正中面内の仰角 θ の音源から受聴者の外耳道入口までのHRTFsを $H(\theta, \omega)$ 、音源信号を $S(\omega)$ とすると、耳入力信号 $P(\omega)$ は式(1)で表される。

$$P(\omega) = S(\omega) \cdot H(\theta, \omega) \quad (1)$$

したがって、正中面内の仰角 θ_s のスピーカから、 $H(\theta, \omega)/H(\theta_s, \omega)$ なる処理を施して信号を提示すれば、任意の仰角に対する耳入力信号を合成できる。さらに、正中面内の音像方向はスペクトラルキューで決定されるとの立場をとれば、HRTFsはその絶対値だけを取り扱えばよい。

2.2 WMA(Weighting Moving Average)法による平滑化処理

ヒトの聴覚系の周波数分解能は可聴帯域で一定ではなく、低い周波数においては細かく、高い周波数では粗いことが知られている。正中面定位においても、周波数が高くなるにしたがってHRTFsの微細構造は必要ではなくなり、マクロなふるまいで十分であると推論できる[2]。そこで、HRTFsの絶対値に対して、移動平均により平滑化を行う。ただし、一般的な移動平均ではなく、高域になるにしたがって連続的に平滑化の度合いが大きくなるような処理(WMA; Weighting Moving Average)を導入した。以下にその方法を示す。

一般に、移動平均は式(2)で表される。

$$|H_s(\theta, \omega_i)| = \frac{1}{M} \sum_{j=-L}^L |H(\theta, \omega_{i+j})|, \quad M=2L+1 \quad (2)$$

ここで、 M は移動平均に用いるサンプル数である。移動平均により形成される低域(この場合は周波数特性の変動の低域成分)通過フィルタの遮断周波数 f_c は式(3)で表される(Δs はサンプリング間隔)。

$$f_c = \alpha / (M \cdot \Delta s), \quad \alpha = 0.443 \quad (3)$$

周波数が高くなるにしたがってHRTFsの平滑化の度合いを大きくするには、それにしたがって M を大きくすればよい。ただし、 f_c は単純にHRTFsの周波数と反比例の関係とするのではなく、式(4)のように臨界帯域幅 Δf_G と反比例の関係であるとする。

$$f_c = \beta / \Delta f_G, \quad \beta = \text{const.} \quad (4)$$

臨界帯域幅 Δf_G は、500Hz付近までは100Hzで一定で、それ以上では中心周波数の約20%増で増加していくことが知られている[3]。

$$\Delta f_G = \begin{cases} 100 + 0.2(f - 500) = 0.2f, & 500 < f \\ 100, & f < 500 \end{cases} \quad (5)$$

以上より、式(2)における M は、

$$M = \alpha / (f_c \cdot \Delta s) = (\alpha \cdot \Delta f_G) / (\beta \cdot \Delta s) \quad (6)$$

となり、 $\gamma = \alpha / (\beta \cdot \Delta s)$ とおくと以下のようなになる。

$$M = \begin{cases} 0.2\gamma f, & 500 < f \\ 100\gamma, & f < 500 \end{cases} \quad (7)$$

3. 音響心理実験

平滑化が定位精度におよぼす影響を検証した。

3.1 実験方法

シミュレートした仰角 θ は0-180°(30°間隔)の7方向である。被験者のHRTFsは、無響室でプローブマイクロホンを用いてインパルス応答を測定し、それをフーリエ変換して求めた。サンプリング周波数は44.1kHz、応答長は256サンプルである。

音源信号は100-10kHzの white noiseで、 $\theta = 90^\circ$ に設置したスピーカから式(8)に示す $G(\theta, \omega)$ なる補正処理を加えて提示した。実験で用いた刺激を Table 1 に示す。

$$G(\theta, \omega) = |H(\theta, \omega)| / |H(90, \omega)| \quad (8)$$

*Some consideration on simplification of head-related transfer functions in median plane localization. By K.Iida(AV&C Research Lab. Matsushita Comm.) and M.Morimoto(Fac. of Eng., Kobe Univ.)

WMA法では、2.2で述べたように周波数が高くなるほど平滑化の度合いが進むが、この処理を開始する下限周波数をパラメータとして4種類の刺激を作成した。また、比較のため通常の移動平均(Constant Moving Average, 以降CMAと記述)による刺激も作成した。ただし、100-10kHzの帯域内で移動平均に用いた総サンプル数 ΣM がWMA法のそれとほぼ同じになるように M の値を選んだ。したがって、WMA1(CMA1)がWMA(CMA)法のなかで最も平滑化の度合いが大きい。さらに、平滑化処理を行わない刺激(NS)と実音場(各 θ に設置したスピーカから提示)による刺激(RSF)を作成した。

無響室内でこれら70個の刺激(10方法 \times 7方向)をランダムに提示した。刺激の提示時間は1秒、刺激と刺激の間隔は5秒である。被験者は椅子に座り、頭を固定されている。マッピング法により各刺激に対して20個の回答を得た。提示音圧レベルは50dB(A)(SLOW, Peak)である。被験者は男子学生1名である。

Table 1 Kinds of stimulus used in the experiment.

Method	Simplification		
		Frequency range(Hz)	ΣM
Real Sound Field, RSF	—	—	—
No Simplification, NS	—	—	—
Weighting Moving Average,	WMA1	500-10k	1152
	WMA2	1000-10k	563
	WMA3	2000-10k	271
	WMA4	4000-10k	130
Constant Moving Average,	CMA1	—	1160
	CMA2	100-10k	580
	CMA3	—	290
	CMA4	—	116

3.2 結果と考察

式(9)により、各刺激について知覚方向とシミュレーション方向の誤差 e を求めた。

$$e_{j,k} = S_{j,k} - R_{j,k} \quad (9)$$

ここで、 S は知覚方向、 R はシミュレーション方向、 j は音源方向の番号、 k は試行番号である。誤差 e は正規分布をするとみなせるので、その平均値と標準偏差を求めた(Table 2)。*および**は各シミュレーション方法が実音場(RSF)と比較して差があるとみなした場合の有意水準(それぞれ5%, 1%)である。ただし、標準偏差(SD)に付与したアスタリスクは分散に差があるか否かの検定結果である。これより、平均誤差でRSFと差があるとみなせないのはWMA3とCMA2であるが、分散では差があるとみなせないシミュレーション方法はない(いずれの方

法も分散は増大している)。平均誤差および標準偏差の双方においてRSFに最も近いのはWMA3である(平均値では有意差がなく、標準偏差の差が約10°)。RSFとWMA3の知覚方向の分布をFig. 1に示す。WMA3のほうがややばらつきが大きい。分布の中心はほとんどかわらず、WMA法によるシミュレーションの可能性を示している。

Table 2 Error of elevation. (* $p < 0.05$, ** $p < 0.01$)

Method	Mean error(deg.)	SD(deg.)
RSF	-6.0	22.8
NS	-18.0 **	42.2 **
WMA1	20.8 **	54.7 **
WMA2	5.6 **	43.2 **
WMA3	-5.7	32.7 **
WMA4	-15.6 *	40.8 **
CMA1	15.5 **	46.7 **
CMA2	-2.0	36.8 **
CMA3	-18.2 **	42.5 **
CMA4	-13.4 *	37.5 **

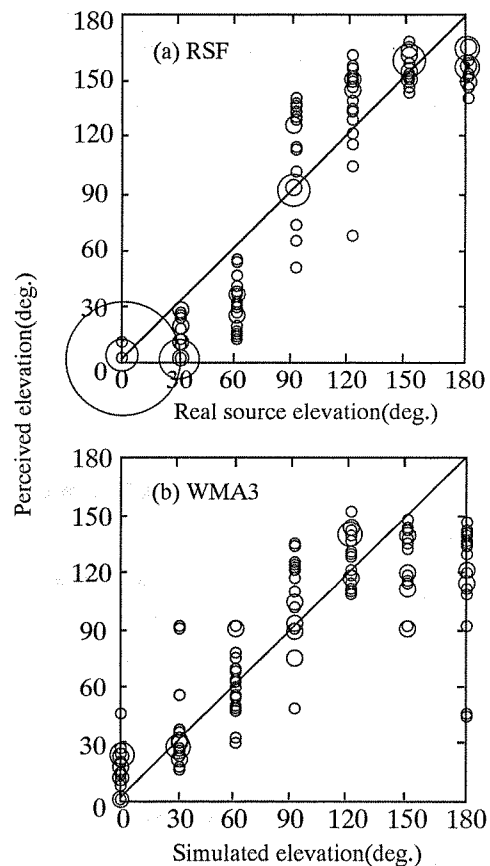


Fig. 1 Response to simulated and real sound sources. Area of circle is proportional to the number of responses

文献

- [1] M. Morimoto and Y. Ando, J.A.S.J(E)1, 167-174 (1980)
- [2] 例えば, F. Asano *et al.*, J.A.S.A. 88, 159-168 (1990).
- [3] E. Zwicker and H. Fastl, *Psychoacoustics Facts and Models*, Springer-Verlag (1990).