

EV/HEV の報知音を想定した漸増刺激音の定位の主観評価 A Subjective Evaluation of Sound Source Localization using granularly increased stimulus for sign signal of EVs and HEVs.

竹内 大貴† 大山 貴紀† 金子 格†
Daiki Takeuchi Takayuki Oyama Itaru Kaneko

1. はじめに

静音性に優れている EV/HEV は、歩行者が車両の接近に気づけないことによる事故が懸念されており、安全性の面からは走行音が必要と考えられている 1)。適切な走行音としては、エンジン音に近い物以外に様々な可能性があるが、音質や周囲の環境がいかんして、どの程度走行音の認知に影響するか、まだほとんどわかっていない。音響信号の主観、客観評価手法にはさまざまな方法があるが、このような目的に適した評価方法も確立していない。

本研究では、音源定位が時間、音質、音量などに影響を受けることを考慮し、漸増的に音量があがる信号に対し、瞬時に回答させることにより、自動車走行の安全性の評価に適した定位特性の評価を試みた。

2. 目的

Hebrank 等の研究 2)をきっかけに、HRTF とスペクトラルキューの分析が行われてきた。垂直方向の定位に、頭部伝達関数の、7kHz と 9kHz のピークとノッチが寄与していることは広く知られている。

一方山口等の研究によると、反射のある環境ではスペクトラルキュー以外の要素も定位に寄与していると報告している 3)。反射のある環境で 4kHz 以内や 1kHz 以内の周波数帯域を持つ刺激音でも、動的な上下定位の認知が可能であった 4)。同じ条件下において周波数帯域幅を狭めていった場合、音源定位が失われる閾値を調査した。また上下定位だけでなく左右定位についても同様の調査をした。

同一の環境で実験を行う為、今回の実験では漸増刺激と瞬時回答を用いたサイン音の定位の主観評価 4)の方法を用いた。

3. 実験

3.1 刺激音

これまで音源定位についての研究は、一定方向から定期的に放射される音源定位に関する研究がされていたが、ハイブリッド自動車等のサイン音を検討する場合、音の認知性ととも、音源定位の動的特性が重要と考えた、すなわち音が聞こえはじめてから、どれだけ瞬時に被験者が反応するかを評価する方法を採用している。

そこで、車両が接近する状況に似せて、急速に音量を増す試験音を用い、音源定位の判定速度を比較する、

という方法で音源定位に関する測定を行うこととした。

図 1 に刺激音の生成に用いた Scilab のプログラムコード表 1 にプログラムの変数一覧を示す。このプログラムは 0.5 秒間隔で音量が 5dB ずつ上昇するホワイトノイズと、1kHz、2.5kHz、4kHz、5.5kHz、7kHz の 5 種類のローパス

フィルターを生成している。これらから周波数帯域幅の異なるホワイトノイズを生成し、それぞれ左右単体の刺激音を書き出し、計 10 種類生成している。この刺激音は回答時間が 500ms 増加した場合、

同じ定位を得るための音量が 5dB 余分に必要である。ただし先の研究 4)とは違い、アタック時間は 0.00001 秒、リリース時間は 0.1 秒に固定した。

表1 プログラムで使われた変数の説明

fs	サンプリング周波数
n	1 秒のサンプリング数
t	サンプリング周期
w	帯域幅
a	アタック時間
r	リリース時間

```
fs=48000;
d=fs/2;
n=0:(d-1);
t=n/fs;
a=0.00001;
r=0.1;
w=[1000,2500,4000,5500,7000];|
sig_f=440;

for px=1:5;
    x=(1-2*rand([t,1:(2^14+1)]))*0.1;
    flt=ffiltz("lp",2^14,w(px)/fs,w(px)/fs);
    x2=convol(flt,x);
    x3=x2(n+2^13);
    v=min(exp((t-0.1)/a),exp((0.1-t)/r));
    y=x3.*v;

    trl=0;
    for i=0:10
        db=10^(i/20*5)/100;
        trl(i*d+d)=0;
        trl(i*d+n+1)=trl(i*d+n+1)+y'*db;
    end
    savewave(sprintf("p1-%02d-l.wav",px),[trl,trl*0],fs);
    savewave(sprintf("p1-%02d-r.wav",px),[trl*0,trl],fs);
end
```

図1 刺激音のホワイトノイズ生成と5種類のローパスフィルターを含むScilabのプログラムコード

3.2 機材配置

表 2 に機材の一覧表、図 2 に機材配置側面図、図 3 に実験環境と室内の配置図を示す。実験は東京工芸大学 9 号館 3 階にある小ゼミ室で行った。被験者は同じ研究室に所属する 8 名を用意した、全員男性で年齢は 20 もしくは 21 歳、聴覚に障害を持つなどはなく、普通に生活が送れる健康状態である。椅子は机から 100cm 離れた位置にあり、被験者はそこに着席した。

機材は全て机の上に配置され、上下定位の実験ではステレオスピーカの右を机の上に設置し、左はスピーカスタンドを使い右スピーカの 50cm 上方に設置した。左右定

位では左右のスピーカを 3cm 離して、被験者正面に配置した。この時上下定位、左右定位どちらの実験でも、被験者にはどちらか片方のスピーカに偏りすぎない位置へ着席する事と、姿勢をとる事を心掛けさせた。

表2 実験で使用した機材の一覧表（スピーカの高さを調節するために使った机やスタンド、外付けキーボードは含んでいない）

機材	型番など
ステレオスピーカ	BOSE Computer MusicMonitor
PC	Dell Vostro 3700 (AVML-PC29) Core i7 Q740 1.73GHz メモリ 6GB
OS	Windows7 home 64bit
刺激音生成	scilab-5.3.0
アプリケーション開発	Microsoft Visual Studio 2008

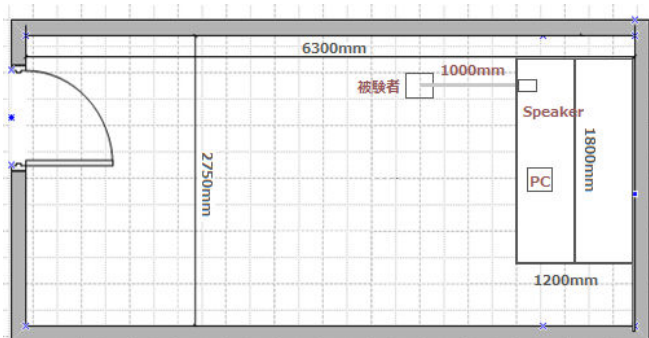


図2 東京工芸大学9号館3階小ゼミ室内実験2機材配置図（上下定位、左右定位共に同一の配置）

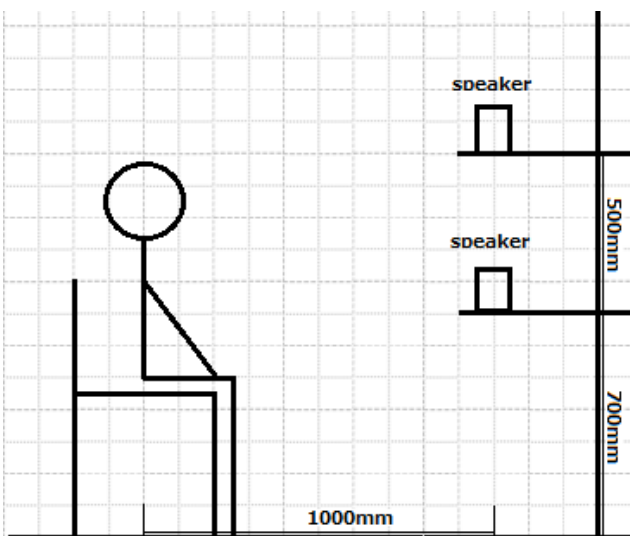


図3 東京工芸大学9号館3階小ゼミ室内実験機材配置側面図（上下定位の場合の配置、左右低位では高さ700mmの位置に左右30mm離して設置）

3.3 被験者回答とアプリケーション

図 4, 5, 6 に Microsoft Visual Studio 2008 で作成した刺激音再生のためのアプリケーションを示す。このアプリケーションは、起動時にあらかじめ設定された数の刺激音を読み込み、同時に発生させた乱数を使い、刺激音が再生される順番をランダムに再配置する。OK ボタンを押すと先頭に配置された刺激音が、2 秒の無音の後に再生される。刺激音は 4 秒間再生され、その間に被験者が回答すればその回答内容と反応時間がテキストファイルに出力される。回答しても再生が自動的に停止することではなく、刺激音は 4 秒間再生され、その後待機状態になる。待機状態で OK ボタンを押すと、2 秒の無音の後に 4 秒間次の刺激音が再生され回答を受け付ける。以後最後の刺激音まで、これを繰り返す。

被験者に直感で回答される事を避けるため、どちらかわからない場合の無回答を認め、間違えた場合は実験を始めからやり直しとして、確実な回答を得た。また無回答が続く回答が 3 割以上得られなかった場合にも、実験のやり直しを行った。

また被験者は 100cm 離れた位置から回答するため USB 外付けタイプのキーボードを用意した。上から再生されていると判断したら上キー、下から再生されていると判断したら下キーを押すことで瞬時に回答出来る。被験者の回答はテキストファイルに出力されるが、OK ボタンの上のラベルにも、上ならば U、下ならば D と表示し、ステータスバーに表示された現在再生している刺激音と照らし合わせ、即座に被験者の回答の正誤を判断する事を可能にした。

ステータスバーに表示されている情報は、xyz の右の数値が全体の刺激音の数を表しており、p から左は再生中の刺激音の情報である。x は用意されたファイルの総数、y は刺激音の音色パターン数(今回はホワイトノイズのみなので 1 になっている)、z はエフェクトパターン数(今回は 1 になっている)。一方 p の後には、再生中の音色パターン番号とエフェクトパターン番号、その横にはスピーカの LR どちらから再生されているかを表示している。実験中にはこの情報を被験者には知らせていない。

またデバッグ用として、ランダムに再配列された刺激音の再生順を確認できる一覧や、今回の実行で使った乱数もウィンドウを拡張すると表示することができる。音色パターンやエフェクトパターンの追加は、ファイルさえ用意すれば Microsoft Visual Studio 2008 から非常に簡単に行える。

上下定位については 2 回実験を行っている、この差については後述する。

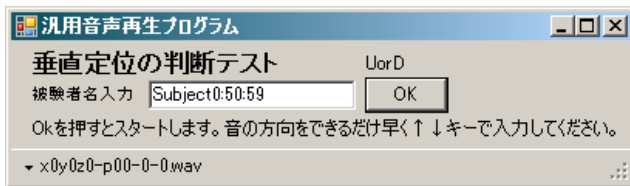


図4 Microsoft Visual Studio 2008で作成した刺激音を再生するために汎用音声再生プログラム

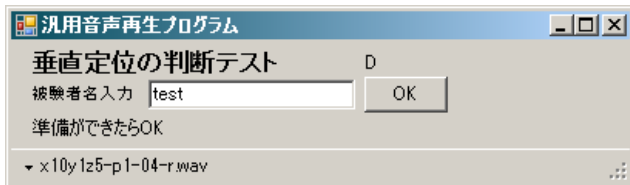


図5 待機中の汎用音声再生プログラム(被験者の回答はOKボタンの上に表示され、ステータスバーに表示されている再生された試験音と比較して正誤判定が可能になっている)

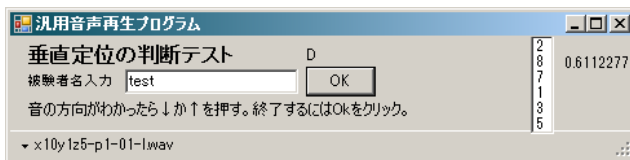


図6 動作中の汎用音声再生プログラム(デバッグ用のトラックリストと乱数を表示している)

4. 結果

図7には被験者が各刺激音の左右定位の感知に必要な平均時間の95%信頼区間のグラフを示す。16回の回答のうち無回答や誤答は除外している。左右定位は全ての刺激音で10以上の標本が取れた。周波数帯域幅が狭まるにつれて反応時間は大きくなっていったが、信頼区間の幅にはそれほど差がなく、周波数帯域幅を狭めていっても、反応に時間を要するが同じように定位が得られたことが見て取れる。

図8には被験者が各刺激音の上下定位の感知に必要な平均時間の95%信頼区間のグラフを示す。同様に16回の回答のうち無回答や誤答は省いている。1回目の実験は正答率が5割前後になり、回答が2択の実験としては信頼できる結果ではなかった。そのため2回目の実験を行い、以上の問題への対策として、被験者が回答を誤った場合その時点で実験をやめ、最初からやり直すことで正答率100%のデータを得た。

上下定位2回目の実験では、1kHz以内の周波数帯域を持つ刺激音は1回しか回答が得られなかった。また他の周波数帯域を持つ刺激音では複数回答は得られているが、全ての刺激音で1回目の実験より反応時間が大きくなる結果になった。2.5kHz以内と5.5kHz以内の周波数帯域を持つ刺激音、4kHz以内と7kHz以内の周波数帯域を持つ刺激音で似たような平均とばらつきの結果が得られた。また広い周波数帯域幅を持つ、5.5kHz以内や7kHz以内の刺激音では大幅に回答時間が大きくなっている。

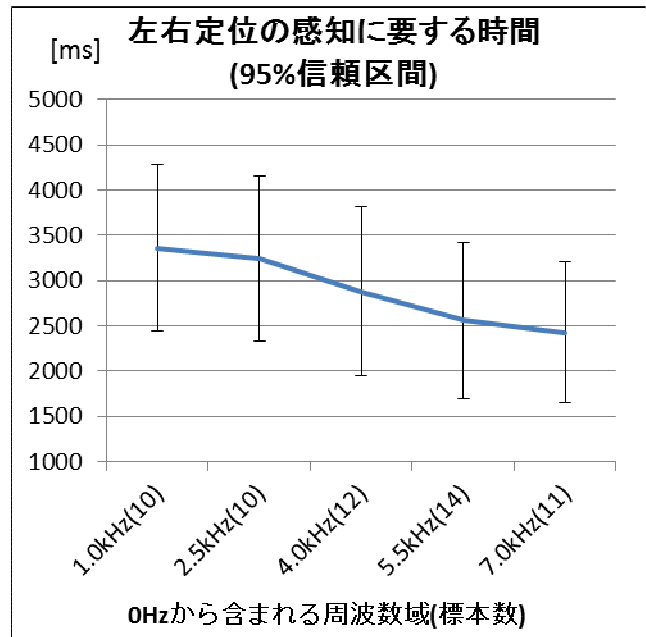


図7 被験者が各刺激音の左右定位の感知に必要な平均時間の95%信頼区間のグラフ

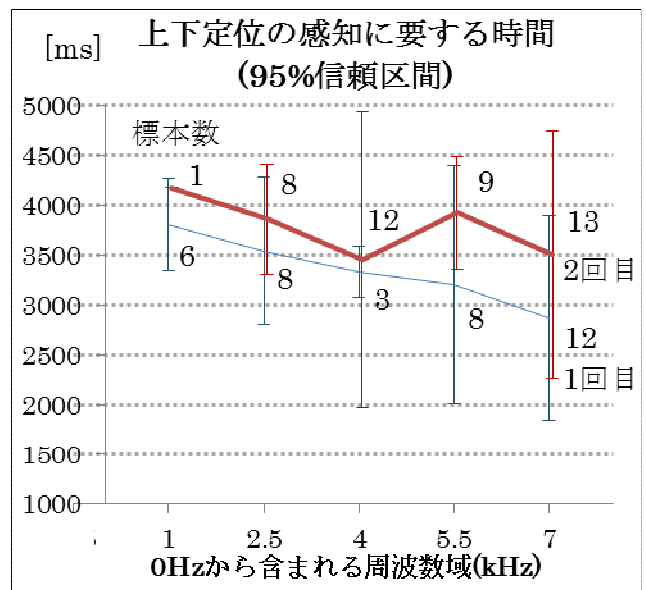


図8 被験者が各刺激音の上下定位の感知に必要な平均時間の95%信頼区間のグラフ

5. 考察

左右定位の実験では、誤答があった場合のやり直しを導入していなかったが、どの刺激音でも高い正答率が得られ、周波数帯域幅が狭まっても定位が得られることが分かった。周波数帯域幅が狭くなっていくにつれ、反応時間が大きくなったのは、人の聴覚が周波数によって感度が違うからと考えられる。

上下定位1回目の実験では、敏速な聞き分けが困難であったのか、実験で慌てて回答し回答後にその答えでは無かったと発言する被験者もいた。結果は正答率が5割前後になり、回答が2択の実験としては、信頼できない結果になった。しかし被験者が早く答えようとして、勘違いや

直感で答えるなどをしてしまうほど、上下定位は左右定理に比べて、即座に定位を判断する事が容易でないことが分かった。

上下定位 2 回目の実験は、正答率が 5 割では定位の閾値や実際の反応時間は測定できないため、その対策として、「少しでも不安があれば無回答を勧めておき、回答を誤った場合はその場でやり直しをしろ」と、速さではなく確実に回答するよう被験者に指示をした。

結果 5 名の被験者はやり直すことなく確実に回答したが、3 名の被験者は数回のやり直しを要した。一番やり直しの多かった被験者が誤答した時に質問すると「確実にこちらだと思った。」と回答した。今回取ったデータは正答率 100%ではあるが、一部の被験者はどこかで勘違いして回答するので、全体では 100%にならない事が分かった。その確率については、今回やり直し前のデータは捨ててしまっているのでも求められない、今後は誤答も記録するよう記録方法を改善したい。

正答率が必然的に 100%になる今回の方法では、1kHz 以内の周波数帯域を持つ刺激音を除き、回答率が 5 割から 8 割と程々に認知できる結果になった。1kHz 以内の周波数帯域を持つ刺激音では、1 人の被験者以外は正解出来なかった。

広い周波数域幅を持つ 5.5kHz 以内や 7kHz 以内の周波数帯域を持つ刺激音では、1 回目と異なり他の試験音より更に回答時間が大きくなったが、これは被験者が小さい音量で分かっても即座に判断せず、正解しなければいけないという重圧から定位の方向をよく考え、回答が遅れたのではないかと考える。

また 2.5kHz 以内と 5.5kHz 以内の周波数帯域を持つ刺激音、4kHz 以内と 7kHz 以内の周波数帯域を持つ刺激音で似た平均と信頼区間を持つ結果になった。これについてもどの刺激音が分かりやすいということはない事の表れであるかもしれないが、平均の他にも標本数が 8 と 9、12 と 13 と似たような数値で、似たような信頼区間を持っている。このことから信頼区間が広く標本数が多い刺激音は、よく聞いて考えれば 4 秒間以内に感知できるという刺激音ではないかと考えられる。標本数の少ない刺激音も 4 秒を超えて再生すれば標本数が増える可能性もある。しかし標本数の多い試験音の平均の反応時間が小さい結果になっているので、4 秒以上の再生で標本数が増えてもグラフの折れ線が形を変えることはないように思われる。

ところで、1kHz 以内の周波数帯域を持つ刺激音で正解出来た被験者に質問したところ、「ファーストパーソン・シューティングゲームという、ゲームを普段からよくプレイし、音の方向や距離間には自信がある」と回答した。このジャンルのゲームは一人称視点で 3D のフィールドを歩き回り、相手を倒すという内容が多く、特に銃などを装備して戦闘するゲームの場合は、相手の足音を聞き位置を探るという要素も用意されている。この被験者は日常的にこのような事を行っていたので、定位を感知する能力が上がっていたと考えられる。これと似たような鍛錬を積んだ場合、定位を感知できる確率は上がるが、ほとんどの被験者の場合 1kHz 以内しか周波数帯域を持たない刺激音では定位を感知できない事が分かった。

実験後分かりやすい分りにくい試験音はあったかという質問に、1kHz 以内の周波数帯域を持つ刺激音が分か

りにくいと以外は、特にどの刺激音が分かりやすい事はなかったようなので、それが表れた結果なのだと考えられる。

今回の実験は標本数が最大で 16 しか得られずまだ規模の小さい実験であるので結果の信頼性は高くない。今後同じ条件による実験データを増やし信頼性を増すとともに、実験方法の最適化を図っていく必要がある。

参考文献

- [1]国土交通省、ハイブリッド車等の静音性に関する対策のガイドライン, <http://www.mlit.go.jp/common/000057372.pdf>
- [2]Jack Hebrank et.al, Spectral cues used in the localization of sound sources on the median plane, J. Acoust. Soc. Am. Volume 56, Issue 6, 1829-1834, 1974
- [3]金子 格, 土屋 慶多, 山口 隼人, 本人の HRTF から合成した仮想音源による有響室における音像定位の主観評価試験法: もっと多く, もっと精密に, もっと短時間に, 電子情報通信学会技術研究報告. EA, 応用音響 110(285), 49-54, 2010
- [4]竹内 大貴, 金子 格: 漸増刺激と瞬時回答を用いたサイン音の定位の主観評価, 情報科学技術フォーラム. 巻: FIT2011 頁: 335-336 特殊号: 講演論文集 第 4 分冊, 2011