

マルチメディアコンテンツ（楽器演奏シーン）における音源定位

高橋一裕・太原泰介・阿山みよし

目加田慶人・長谷川光司・春日正男

宇都宮大学 工学部 情報工学科

〒321-8505 宇都宮市石井町 2753

1. はじめに

視覚と聴覚のように、性質の異なる感性が、互いにどのような作用を及ぼし合うかという問題は、心理学の分野で、古くから研究対象として取り組まれている¹⁾。今日では、視覚と聴覚の相互作用の問題が、実際に映像と音楽を一体のものとして楽しむマルチメディアコンテンツにおける感性や臨場感の問題として関心を持たれるようになってきた²⁾。

しかし、音声と視覚刺激の相互関係に関する詳細な実験は、LEDなどによる光点と、ブザーや白色雑音などの無意味音によるものが多く^{3,4)}、そのような無意味な刺激による所見が映像や音楽などを視聴する実際の状況にどの程度適用できるのかはわかっていない。

そこで、本実験では、音楽演奏シーンという有意義な刺激を用いて、映像と音源の間のような関係があるかを検討した。

実験では、バイオリンとトライアングルの演奏シーンを用い、各々の音の呈示位置と演奏者の映像の整合、不整合条件及び2種類の音質の各条件における水平方向と奥行き方向の知覚的音源位置を測定し、音源定位に対する映像と音質の影響を調べた。

2. 実験

2.1 実験原理

映像刺激は、バイオリンとトライアングルの演奏シーンにおいて、各々の奏者を左右、

前後にそれぞれ配置したものを用いた。音声刺激は、映像に合わせて各々別に演奏されたバイオリンとトライアングルの音を用いた。被験者には、映像と音声の左右又は前後の位置が一致した刺激（整合刺激）と、一致していない刺激（不整合刺激）を呈示し、バイオリンとトライアングルの音がどこから聞こえてくるかを答えてもらう実験を行った。また、音源定位に対する音質の影響を検討するために、音刺激としては2種類の音質レベルを設定した。

2.2 装置

装置図を図1に示す。スピーカは、図のように左、右、前、後に置いた。スピーカを左、右、前、後の順にSL, SR, SM, SDとした。被験者から30度の角度の位置にSLとSRを置いた。また、スピーカの位置が見えない

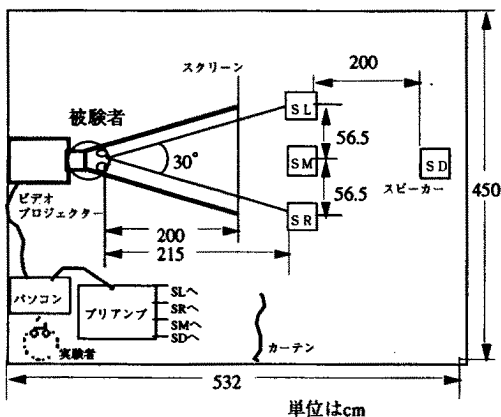


図1 装置図

ようにカーテンで隠した。映像はコンピュータのVGA出力をビデオプロジェクタによりスクリーンに写し出した。また、音刺激はコンピュータのサウンド出力をプリアンプに接続し、4チャンネルセクタによりどのスピーカーから出すかを制御した。

2.3 刺激

バイオリンとトライアングルの演奏シーンのムービーを映像刺激とし、演奏者の位置により各々を条件A～Dとする(表1)。図2に条件Aの例を示す。A～Dの映像刺激を各々5種類ずつ用意した。曲目はバイオリンの初心者向き教本から、似た印象の曲を選び、1小節に1回又は2回トライアングルを打ってもらった。刺激の呈示は3sとした。これは、トライアングル音が2回呈示される長さである。映像刺激のサイズは、縦25.9度横33.4度とした。

音源はバイオリンとトライアングルの2種類を使用した。音刺激は左右(図1のSLとSR)または前後(図1のSMとSD)のいずれかから呈示した。音刺激の条件を表2に示

表1 映像刺激の条件

	A	B	C	D
バイオリン	左	右	前	後
トライアングル	右	左	後	前

左、右、前、後：映像のなかでの演奏者の位置。



図2 映像刺激の例(条件A)

す。まず、バイオリンとトライアングルの合奏シーンを撮影し、各々の演奏者別々にその合奏シーンに合うように演奏してもらった音をコンデンサマイクを使用して録音した。音質は、4 KHzのローパスフィルタをかけた音とかけてない音の2種類用意した。以後、フィルタをかけた音をフィルタあり、かけていない音をフィルタ無しと呼ぶことにする。

2.4 手順

As, Bs, Cs, Dsのいずれかの音刺激をスクリーンの後ろに並べた別々のスピーカーから呈示する。

映像有りのときは映像を見ながら、バイオリンとトライアングル各々について、左右前後のどこから聞こえたかを答えてもらう。1セッションでは、各々の刺激条件に対して5試行ずつ80試行行い、これを4セッション行った。よって、各々の刺激条件に対して合計20試行ずつ行ったことになる。図3に音声と映像の条件を示す。映像なしのときには映像を呈示せずに各々の音声条件に対して、映像有りのときと同様に答えてもらう。1セッ

表2 音刺激の条件

	As	Bs	Cs	Ds
バイオリン	左	右	前	後
トライアングル	右	左	後	前

左：SL, 右：SR, 前：SM, 後：SDのスピーカーから提示。

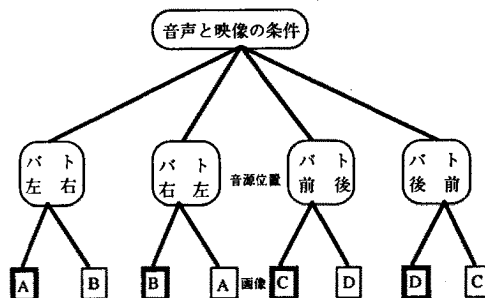


図3 音声と映像の組み合わせ：太い枠の条件のとき整合、細い枠の条件のとき不整合。

ションでは、各々の刺激条件に対して10試行ずつ80試行を行い、これを2セッション行った。よって、各々の刺激条件に対して合計20試行を行った。

なお、被験者の耳を慣れさせるために、セッションを始める前にバイオリンとトライアングルの音を各々のスピーカより呈示し、左、右、前、後のスピーカからの呈示がどのようなものかと確認してもらった。

2.5 被験者

被験者は20代の男性10名と、女性1名の計11名である。いずれも正常な視力と聴力を有している。

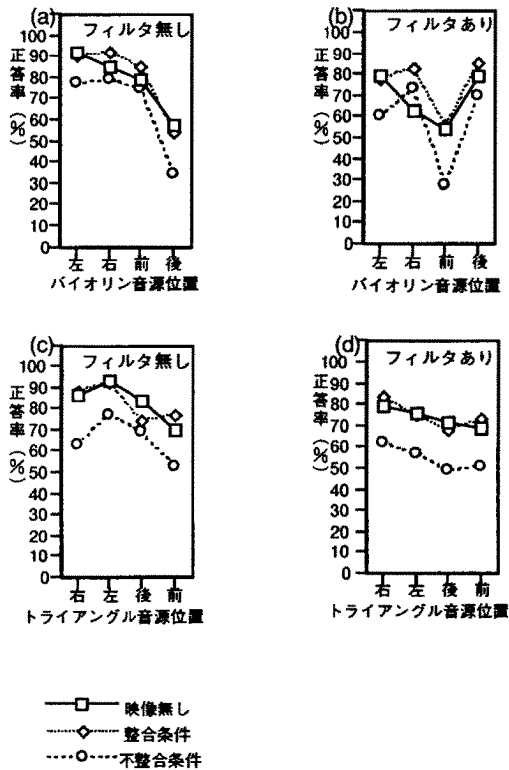


図4 音源定位実験の結果：横軸はバイオリン又はトライアングルの音源位置で、□は映像無し条件、◇は映像が音源位置と整合している条件（整合）、○は映像が音源位置と逆転している条件（不整合）を示している。(a)と(b)はバイオリンの、(c)と(d)はトライアングルの正答率を示している。(a)と(c)はフィルタなし、(b)と(d)はフィルタあり条件の結果である。

3. 結果及びまとめ

図4に被験者11名の平均の結果を示す。映像無しの条件と音源位置が映像に整合している条件（整合条件：図4の□と◇）の結果はほとんど有意差がなかった。ただし、条件によっては、整合条件の方が映像なしより良い成績となる場合もあった。また、図4に見られるように、音源位置が映像と整合していないとき（不整合条件：図4の○）、全体として正答率は低くなっている。不整合条件において、音源定位がどのように映像に影響されているかをみるために、ひっぱられ率を計算してみた。ここで、ひっぱられ率とは、不整合条件において音源位置を映像と同じと答えた値を全誤答数で割った値と定義する。図5に引っぱられ率の結果を示す。このひっぱられ率をみると、左右方向に比べ、前後方向での条件においては、ひっぱられ率が高い。以上のことから、特に前後方向の条件におい

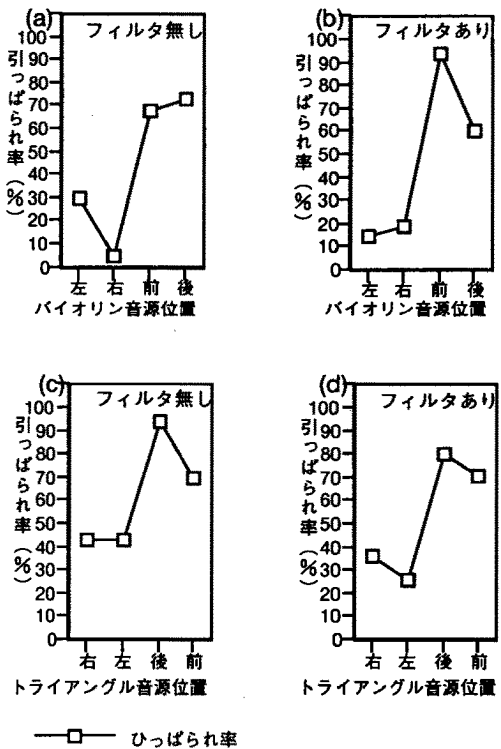


図5 ひっぱられ率：ひっぱられ率とは、不整合条件において音源位置を映像と同じに答えた値を全誤答数で割った値を指す。

て、音源位置を決定する際に、映像があることによってその映像に音源位置がひっぱられているということを示している。左右方向の条件では、そのような傾向は顕著には現れなかった。

また、バイオリンの定位では、フィルタ無しでは前スピーカ (SM) から音が呈示される (図 4 (a) の前) のよりも、後ろスピーカ (SD) から音が呈示される方 (図 4 (a) の後) が定位が悪く、反対に、フィルタありでは SD よりも SM から呈示される方が定位が悪い (図 4 (b) の前と後)。この原因は、次のように考えられる。フィルタありと無しの条件を同一セッション内で行っているため、被験者にとっては、フィルタありのバイオリンの音はフィルタ無しの音と比較し、小さく聞こえていると考えられる。フィルタ無しの音が SD から呈示されても、フィルタありの音と比較し大きく聞こえるため、SM から呈示されているように知覚される。逆に、フィルタありの音が SM から呈示されても音は小さく聞こえるので SD から呈示されているように知覚されていると考えられる。したがって、フィルタ無しでは SD から、フィルタ有りでは SM から聞こえるという回答が少なくなり、各々の音源位置が後と前の場合の正答率が低くなった。このため、バイオリンの定位では、フィルタあり無しで、前と後の条件間で逆転する結果となったと考えられる。しかし、トライアングルの定位では、フィルタによる影響は現われてい

ない (図 4)。これは、バイオリンと違いトライアングルはフィルタをかけても音の大きさがあまり変わらない様に知覚されたからだと考えられる。

フィルタありとフィルタ無しでは、全体的にフィルタありの方が音源定位能力は悪い。したがって、正確な音源定位がなされる為には、4kHz 以上をカットされた帯域での音質では不十分で、高音質な音情報が必要である。また、前後方向では映像に引っ張られやすいことから、音響再生系において容易にまたは安価には実現できない音の奥行感を感じさせるのに、奥行感のある映像が効果がある可能性がある。

バイオリンとトライアングルを用いた音源演奏シーンという人間にとって有意味な刺激を用いた場合でも、音源を定位する際に映像に影響されることが分かった。

文 献

- 1) 丸山欣哉：感覚間相互作用。亭坂良二 (編)：講座心理学，3 感覚。東大出版，267-297, 1969.
- 2) 二階堂誠也：私説：A と V の心理学，その 3 耳と目が同時に働くとき…。JAS Journal, 12月号, 22-26, 1984.
- 3) M. Radeau and P. Bertelson: Auditory-visual interaction and the timing of inputs, Tomas (1941) revisited. *Psychological Research*, 49, 17-22, 1987.
- 4) 北島律之，山下由乙男：音源定位への光刺激の影響における同期性の効果の空間異方性。Vision, 9, 1-12, 1997.