

時間軸上における映像と音楽のずれと一体感に関する検討

綱本 光洋[†] 松井 淑恵^{††} 風井 浩志[‡] 片寄 晴弘[†][†] 関西学院大学理工学部 ^{††} 関西学院大学理工学部ヒューマンメディア研, JST, CrestMuse Project [‡] 関西学院大学理工学部ヒューマンメディア研

1 はじめに

一般に映画やテレビのような映像には必ず音楽が伴っており、音楽と映像の結び付きは様々な効果をもたらす。先行研究では、円や三角など抽象的な映像素材を用いて、音楽の印象が映像の印象に作用することを示している [1]。また、テレビの音と映像がどの程度ずれていると違和感を覚えるかに関して、様々な視聴覚実験が行われている。[2]。しかし、既存の研究では、対応関係にある視覚刺激と聴覚刺激の一体感を扱ったものであり、映像の動きと音楽のリズムのように現実では対応関係のない視聴覚刺激の一体感について扱ったものではない。

本研究では、映像の動きと音楽のリズムのずれを検知できているかについて、映像と音楽の一体感が芸術性に必要とされるバレエを題材に実験を行った。

2 実験

被験者: バレエ未経験者 16 名が実験に参加した。

2.1 刺激

バレエの DVD および記録映像「ドンキホーテ (L. ミンクス作曲 松山バレエ団公演)」より 3 シーン、「ローザンヌ '89 in Tokyo」より 1 シーン、「白鳥の湖 (P. I. チャイコフスキー作曲 ボリショイ劇場バレエ団)」から 1 シーンの合計 5 シーンをオリジナルムービーとして抜粋し、0 フレーム、4 フレーム、24 フレーム分音声と映像をずらしたのから 8 秒間を切り出して刺激を作成した (図 1)。ずらし方は、音声映像が先行する刺激 (-24 フレーム、-4 フレーム)、音声と映像のタイミングがオリジナルのままの刺激 (0 フレーム)、遅延した刺激 (+4 フレーム、+24 フレーム) とした。フレームレートは 30fps、画面のサイズは 480 × 720 ピクセルとした。

-4 フレーム、+4 フレームは、準備段階でバレエ経験者にアンケートした際の「オリジナルと比べた場合に違いはわかるが、オリジナルを知らなければとくにずれているとは思わない」というコメントにより決定した。

これら 5 種類のすべての順列 20 組 × オリジナルのムービー 5 本で、計 100 組の刺激組み合わせを用意した。また、特定の映像や音声の開始・終了が手かがりになることを避けるため、刺激ごとに切り出す箇所を 3 パターン用意し、刺激組み合わせごとにランダムに 1 種類を用いた。切り出す箇所、刺激組み合わせの提示順序は、被験者ごとにランダム化したものを使用した。音声には、刺激の開始終了 1 秒間に線形のフェードイン・アウトをかけた。また、素材ごとの音声のレベルの差は実験者が耳で聞いて揃えた。

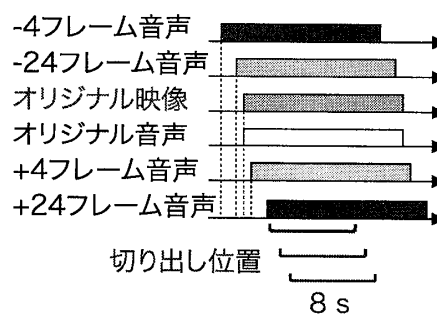


図 1: 本研究に用いた刺激の作成方法。オリジナルの映像部分と 5 種類の音声部分を組み合わせ刺激を作成した。

2.2 手続き

2 つの刺激からなる刺激組み合わせを連続して提示し、どちらの刺激が「映像と音声の一体感があるか」と質問した。判断の基準はとくに指定しなかった。刺激提示のプログラムおよび回答用の GUI は、Max/MSP4.6 を用いて作成した。音声の刺激提示はヘッドフォン Sennheiser HD650 を通して行った。音圧レベルは個々の被験者が快適に実験を行えるレベルに調整した。

3 結果

図 2 はオリジナル刺激とずれのある刺激の比較において、ずれのある刺激を「一致している」と回答した割合を示す。縦軸は映像と音声一致と感じた割合、横軸は本実験に用いた刺激の音声映像に対

Effect of time lag on the synchronization of dance and music in Ballet

[†] Mitsuhiro TSUNAMOTO(ctt60702@ksc.kwansei.ac.jp)

^{††} Toshie MATSUI(tmatsui@kwansei.ac.jp)

[‡] Koji KAZAI(kazai@kwansei.ac.jp)

[†] Haruhiro KATAYOSE(katayose@ksc.kwansei.ac.jp)

School of Science and Technology, Kwansei Gakuin University (†)

Research for human media, School of Science and Technology, Kwansei Gakuin University, JST, CrestMuse Project (††)

Research for human media, School of Science and Technology, Kwansei Gakuin University (‡)

Gakuen, Sanda, Hyogo 669-1337, Japan

してずれているフレーム数を示す。回答の割合が0.5から有意に異なっているかどうか、カイ二乗検定を実施した結果、-4フレームずれた刺激、すなわち音声は4フレーム先行する刺激に対して「一致していない」と正しく判断されていることがわかった。結果は、ずれの絶対量が大きいほどずれが検知しやすいとは限らないことを示唆している。

回答の割合がフレームごとに異なる原因を探るため、本実験に用いた刺激のBPMを小節単位で計測し、ずれのある刺激において映像と音声の時間軸上の関係がどのようになっているかを調査した。図3および図4は、ドンキホーテから抜粋した1シーン（ドンキホーテ02）の映像と音声のずれの状況と、同じくドンキホーテ02における、オリジナル刺激とずれのある刺激の比較において、被験者がずれのある刺激を「一致している」と回答した割合を表す。アスタリスクはカイ二乗検定において0.5より有意に異なっていることを示す（ $p = 0.05$ ）。

5種類のオリジナルムービー全体に対する回答（図2）と回答の傾向は似ており、また図3からは、被験者は映像より音声が遅れている刺激のほうを「一致している」と感じていることが分かる。この結果はドンキホーテ02のムービーのみならず、他の4つのムービーに対しても同様の傾向が見られた。

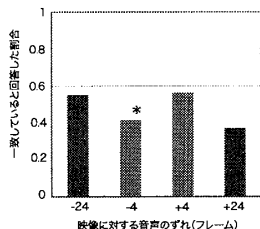


図2: オリジナル刺激とずれのある刺激の比較において、ずれのある刺激を一致していると回答した割合。アスタリスクは回答の割合が0.5より有意に異なっていることを示す（ $p = 0.05$ ）。

4 考察

音声と映像が一致して芸術とみなされるバレエ作品を用いた実験結果より、人は音声映像に先行する刺激より、音声映像に対して遅延した刺激の方が、一体感を感じていることが分かった。

この要因としては、人間は聴覚の方が視覚よりも聴覚の時間分解能のほうが優れているので、音声遅延した刺激に対しては、視覚の時間分解能の窓の中に音声到達するため、両者のずれは検知されない。逆に、音声先行した場合、視覚刺激が到達する前に音声

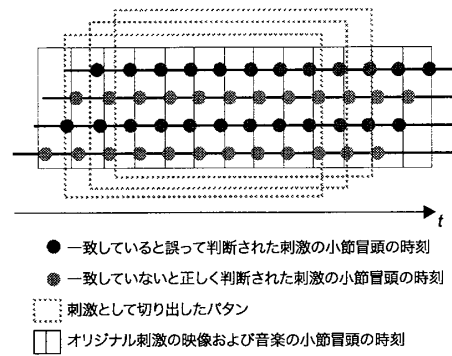


図3: ドンキホーテ02における、オリジナル刺激の映像および音声の小節冒頭の時刻と、ずれのある刺激の小節冒頭の時刻の関係。

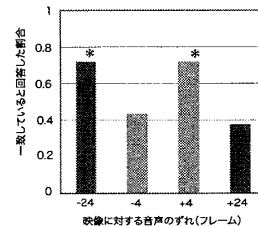


図4: ドンキホーテ02において、オリジナル刺激に対してずれのある刺激を一致していると回答した割合。アスタリスクは回答の割合が0.5より有意に異なっていることを示す（ $p = 0.05$ ）。

到達したことは知覚されるので、両者のずれが検知されるのではないかと考えられる。ただし、音声が遅れると一致していると判断されやすい、という傾向がすべての刺激において観察されているわけではないので、映像と音声の一体感を得るための条件はほかにも存在する可能性はある。

5 参考文献

- [1] Marshall S. K. and Cohen A. J., Effects of musical soundtracks on attitudes toward animated geometric figures, *Music Percept* 6, 95-112, 1988.
- [2] 岩宮眞一郎, 音楽と映像のマルチモーダル・コミュニケーション, 九州大学出版会, 2000.