

透過型 HMD によるミュージックエフェクトの知覚実験

山田千尋¹ 伊藤彰教² 三上浩司²

概要: 透過型 HMD で再生された効果音と、実環境で再生される効果音は明確に分離されて聞こえるのかを確認するため、知覚実験を行った。実験刺激はブロックを積む課題の AR コンテンツを 1 つ準備した。2 分時点と 3 分時点でそれぞれ別のミュージックエフェクトを再生し、可聴性と再生デバイスについてアンケートを行った。19 歳から 24 歳の 15 名が被験者として参加した。McNemar の検定を実施した結果、可聴性については、三角波を用いた長短のミュージックエフェクトで、環境音と HMD 再生音の回答に $p=0.041$ という結果が得られた。

キーワード: 効果音, ミュージックエフェクト, MR, AR

Experiment in Perception of Music Effect with Optical See-Through Head Mounted Display

CHIIHIRO YAMADA^{†1} AKINORI ITO^{†2}
KOJI MIKAMI^{†2}

Abstract: Perceptual experiments were conducted to confirm whether the sound effects played by the transparent HMD and those played in the real environment are clearly separated and audible. Experimental stimuli included one AR content for the task of stacking blocks; different music effects were played at 2 and 3 minutes, respectively, and participants were asked about audibility and playback device. Fifteen subjects, aged 19 to 24 years, participated in the study. McNemar's test was conducted, and the result for audibility was $p=0.041$ for the responses of environmental sounds and HMD playback sounds with long and short music effects using triangular waves.

Keywords: sound effect, music effects, mixed reality, augmented reality

1. はじめに

日常生活において、リビングでくつろいでいる時などに TV から鳴った音なのか、自身のスマートフォンから鳴った音なのかわからなくなるというシチュエーションが存在する。本研究者は、透過型 HMD でも同様の事例が起きる可能性があるのではないかと考えた。透過型 HMD の装着時、意識していない場所からミュージックエフェクトが鳴った場合、実際に再生されたデバイスではなく、他のデバイスから再生されたような錯覚を起こすのではないかとこの事案である。以上を本研究の背景とし、知覚実験を通して本テーマを明らかにする。

透過型 HMD の研究は、視覚呈示やインタラクションに関しては盛んであるが、音の知覚に関する研究は少ない。関連研究としてヒアスルーのヘッドホンと眼鏡型のヘッドセットを使用し、ヘッドホンと眼鏡の外界の音を Acoustic Transparency としている研究が挙げられる[1]が、音のみが対象でありバーチャルで投影されるような視覚的な要因は考慮されていない。また、類似する概念としてゲームサウンドの研究である K. Jørgensen は物語空間と現実空間を

越境させるような音という意味の, transdiegetic sound を提唱している[2]。本研究はゲームとは直接的な関連がないため, transdiegetic sound を応用して透過型 HMD のコンテンツから再生される音について独自の概念を生み出す必要がある。以上のとおり透過型 HMD 向けコンテンツの音を対象とした研究手法が確立されておらず, 研究手法の探索から始めることとする。

2. 透過型 HMD によるミュージックエフェクトの知覚実験

2.1 実験目的

本実験では、透過型 HMD で再生されるミュージックエフェクトと、環境下のスピーカーで再生されるミュージックエフェクトを分離して聞くのか、または連携して聞くのか、を観察することを目的として、印象評価実験を実施した。

2.2 実験方法

被験者は 19 歳から 24 歳の男女 15 名である。

透過型 HMD である HoloLens2 を使用し、ブロック積みアプリを 4 試行実施した。最初のチュートリアルでブロックを掴む動作と離す動作を確認させ、その後 Level 1 か

1 東京工科大学大学院
Tokyo University of Technology Graduate school
2 東京工科大学
Tokyo University of Technology

ら Level14 を順番に体験させた。各 Level で「以下の画像と同じようにブロックを積んでください」というテキストの下に、ブロックが積まれた画像を表示させ、被験者に画像と同じようにブロックを積みさせた。積ませる形状は Level ごとに異なる。Level1 の様子を図 1 に示す。

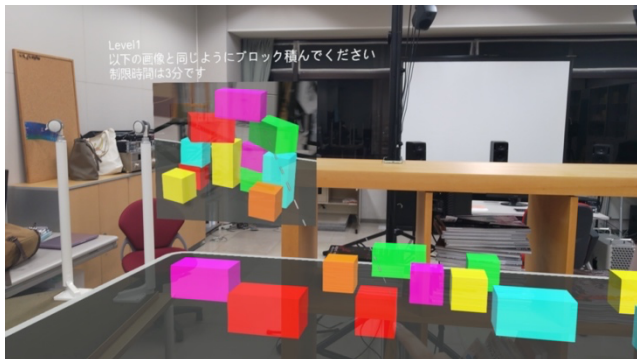


図 1 Level 1 の様子

被験者には xR の操作についての実験であるということのみを知らせ、音に関しての言及は一切しなかった。これは音に対する過度な注意を避け、バイアスがかからないようにする意図がある。

刺激音を再生するデバイスと再生するタイミングを表 1 に示す。本研究は日常生活における音が対象であるため、環境下のスピーカーにはスマートフォンを選定した。体験から 2 分経過した時点で 2 分の経過を通知する音として Level 1 と Level 2 では音 A, Level 3 と Level 4 では音 C を再生した。このとき、視覚的な表示は行わずに刺激音のみの呈示を行った。体験から 3 分経ったタイミングで、「体験は終了です」の視覚表示と同時に刺激音を再生した。この時の刺激音は、Level 1 と Level 2 は音 B, Level 3 と Level 4 は音 D である。また、刺激音は Level 1 と Level 3 の 2 分経過時にのみスマートフォンで再生し、それ以外の刺激音は HoloLens2 から呈示した。

表 1 各 Level の
刺激音を再生するデバイスと刺激音の種類

	2分経過時		3分経過時	
	音再生デバイス	刺激音	音再生デバイス	刺激音
Level1	スマートフォン	音 A	HoloLens2	音 B
Level2	HoloLens2	音 A	HoloLens2	音 B
Level3	スマートフォン	音 C	HoloLens2	音 D
Level4	HoloLens2	音 C	HoloLens2	音 D

Level 4 終了後に Google form でアンケートを実施した。質問項目は「体験時に刺激音は聞くことができたか」の問いに対して「聞こえた」と「聞こえなかった」を片方選択する内容と、「刺激音が聞こえた場合、どこから聞こえたか」の問いに対して「HoloLens2 から聞こえた」と「部屋のどこから聞こえた」の片方を選択させる内容である。これら

の質問は 4 つの Level ごとに回答させた。また、音 A～音 D に関しての説明は文章で行わず、各音源を Google form 上で再生させて聞かせた。

2.3 実験環境

スマートフォンは iPhone12 mini を使用し、2 分経過時に Apple Watch の遠隔操作によって刺激音を再生した。

スマートフォンは高さ 131cm の棚の上に設置した。被験者からスマートフォンの距離は約 180cm である。被験者と各デバイスの位置関係を図 2 に示す。被験者はデモの体験時、一つの場所にとどまらず自由に移動が可能である。両デバイスの音量は、HoloLens2 は 40%, iPhone12mini は 87.5%とした。

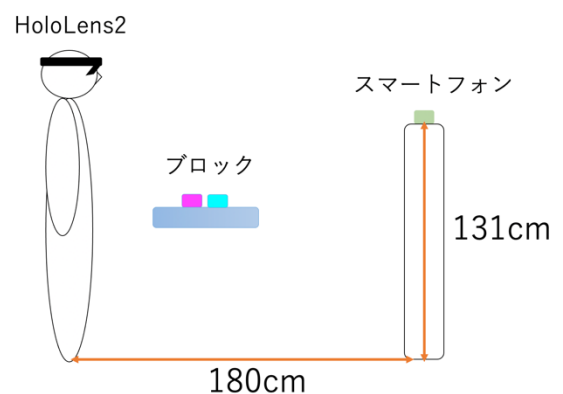


図 2 被験者と各デバイスの位置関係

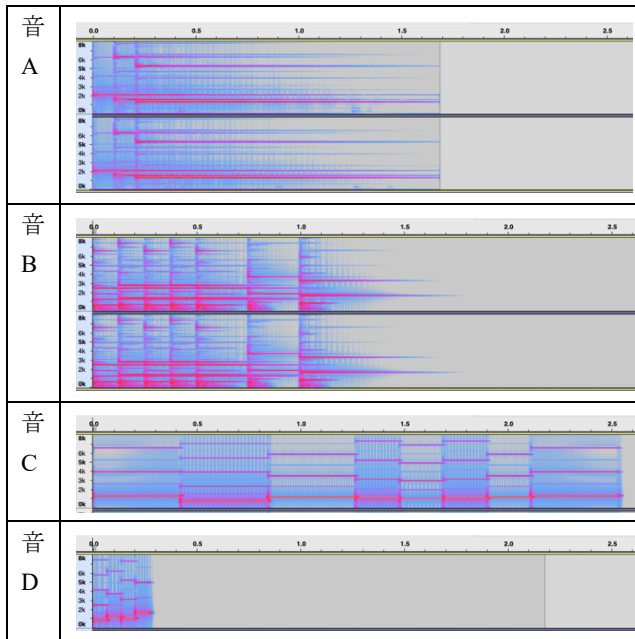
2.4 実験刺激

実験で使用した刺激音の楽譜を表 2, スペクトログラムを表 3 に示す。刺激音の制作にあたって、世の中にある家電などの電子機器から再生されるミュージックエフェクトの傾向を事前調査した。音 A と音 B は高いベルの音色を利用して、家電やスマートフォンから再生されるようなミュージックエフェクトとした。音 C は三角波を利用してメロディフレーズがはっきりと感じられるミュージックエフェクトとした。音 D は三角波を利用して、短いフレーズを高速で再生することによって効果音として聞こえる音とした。

表 2 音 A, 音 B, 音 C, 音 D の楽譜

音 A (ベル, 約 1.7s) ♪ = 142 8	音 B (ベル, 約 1.7s) ♪ = 120
音 C (三角波, 約 2.5s) ♪ = 142	音 D (三角波, 約 0.3s) ♪ = 142

表 3 音 A, 音 B, 音 C, 音 D のスペクトログラム



2.5 実験結果

(1) 提示音を聞くことができたか

集計結果を表 4 に示す. McNemar の検定を実施した結果, Level 3 の音 C と音 D に $p=0.04123$ ($p<0.05$) で有意差が検出された. 音 C は 2.5 秒程度の長めのメロディフレーズであり, 一方の音 D は 0.3 秒程度の効果音的なミュージックエフェクトである. このことから長めのフレーズは「聞こえた」と感じ, 短めのフレーズは「聞こえない」と感じる傾向があることが示唆された. さらに Level 1 と Level 2 とは異なる刺激音が鳴ったことで, 深く印象に残った可能性も考えられる.

表 4 可聴性に対する回答 (単位: 人)

	Yes	No
Level 1 の音 A (スマホ)	12	3
Level 1 の音 B (HMD)	10	5
Level 2 の音 A (HMD)	12	3
Level 2 の音 B (HMD)	10	5
Level 3 の音 C (スマホ)	14	1
Level 3 の音 D (HMD)	8	7
Level 4 の音 C (HMD)	12	3
Level 4 の音 D (HMD)	7	8

Level 3 と Level 4 の音 D に関しては, 「聞こえなかった」と回答する被験者が約半数ほどであった. これは HoloLens2 上に刺激音の再生と同時に, 視覚的なフィードバックを行っていたため, 音への注意が逸れたことが一つの要因として考えられる.

(2) 提示音はどこから聞こえたか

表 5 に集計結果を示す. Level ごと・音ごとの組み合わせで McNemar の検定を実施した. Level 2 の音 A と Level 3 の音

C の組み合わせでは $p=0.07364$ ($p>0.05$) であった. 統計的有意差は認められなかったが, 提示デバイスと提示音が両方切り替わるタイミングは提示デバイスを正しく当てられるという傾向が示唆された.

表 5 再生デバイス判断に対する回答 (単位: 人)

	HoloLens2 から	部屋のどこから	無回答
Level 1 の音 A (スマホ)	8	4	3
Level 1 の音 B (HMD)	5	5	5
Level 2 の音 A (HMD)	10	2	3
Level 2 の音 B (HMD)	5	6	5
Level 3 の音 C (スマホ)	4	10	1
Level 3 の音 D (HMD)	3	5	7
Level 4 の音 C (HMD)	5	7	3
Level 4 の音 D (HMD)	4	3	8

3. 考察

Google form 上で音を聞いた被験者の多くは, 「各音は聞いた覚えはあるが, どのタイミングでどこから鳴っていたかははっきり覚えていない」という状態であることが多かった. 被験者の大半は HoloLens2 のような MR は未体験であり, 今回の実験で初めて体験していた. HoloLens2 での操作で大きな負荷がかかり, 音に対する注意が著しく低下したことが「音をはっきり覚えていない・聞いていない」となった結果の要因として考えられる. また, 音 D のような短いミュージックエフェクトはフレーズ的なミュージックエフェクトより「聞いていない」という印象を抱く傾向があった. 統計的有意差は認められなかったが, 透過型 HMD 上での作業中, ミュージックエフェクトが透過型 HMD から再生されているのか, もしくは環境下のスピーカーから再生されているのか, 必ずしも判断ができるわけではないという傾向が示唆された. HoloLens2 とスマートフォンのどちらから音が鳴っているかの判断材料として, 音源再生時のスマートフォンとの距離と頭の向きによるものが挙げられた.

4. おわりに

今後は, 透過型 HMD 上で作業をしていない負荷の低い状態で音源の再生デバイスを判断できるのかについて検証をするべきである. その後, 透過型 HMD と環境下のミュージックエフェクトが利用者にとってどのような影響を与えるのかを調べるため, 様々なパターンの音刺激を評価する必要がある.

参考文献

- [1]Mark McGill et al., Acoustic Transparency and the Changing Soundscape of Auditory Mixed Reality, CHI '20: CHI Conference on Human Factors in Computing Systems, 2020, p. 1-16.
- [2]Jørgensen, K. A Comprehensive Study of Sound in Computer Games, The Edwin Mellen Press, 2009, p. 97-110.