

令和 6（2024）年度

理学部 卒業論文

卒業論文タイトル

VR におけるストリームバウンス効果：視聴覚同期性の影響

分野名 物質科学

学籍番号 212093

氏名 野沢 優

指導教員 Micheletto Ruggero

令和 7 年 1 月 31 日 提出

VR におけるストリームバウンス効果：視聴覚同期性の影響

Stream bounce effect in VR: the influence of audiovisual synchronization

分野名	物質科学
学籍番号	212093
氏名	野沢 優
指導教員	Micheletto Ruggero

Cross-modal phenomena involve illusions caused by the interaction of multiple sensory inputs, helping to understand complex human information processing. This study focuses on the stream-bounce effect, where two planes appear to bounce off each other when a sound is presented at the moment they pass. Using virtual reality, the study quantifies the impact of sound timing and three-dimensional positioning on this illusion. Results show that presenting sound before visual collision increases bounce response probability, with no significant differences based on sound position. These findings elucidate human perception characteristics and mechanisms in the stream-bounce effect.

＜研究背景と目的＞

複数の感覚情報から起きる錯覚であるクロスモーダル現象の研究はヒトの複雑な情報処理を理解するのに役立つ。クロスモーダル現象の一種であるストリームバウンス効果とは、二つの平面（円または正方形）が互いにすれ違う瞬間に音を提示することで、実際にはすれ違っているのに跳ね返ったように錯覚する現象である。音が視覚情報と同期することで錯覚が強くなることが知られている[1]。

本研究では、ストリームバウンス効果をバーチャルリアリティ上で再現し、音の提示タイミングや音源位置の三次元化を通じて錯覚への影響を定量的に測定し、ストリームバウンス効果におけるヒトの知覚の特性とメカニズムを理解することを目的とする。

二つの実験で条件だけを変更し、同じ方法で Stream/Bounce 評価をした。

＜共通の実験方法＞

1. 固定ターゲット表示 (2000 ms)
2. すれ違う立方体を観察した
3. 2つの立方体がすれ違って見えた(X ボタン)のか、跳ね返って見えた(A ボタン)のか、コントローラーのボタンを押して答えた

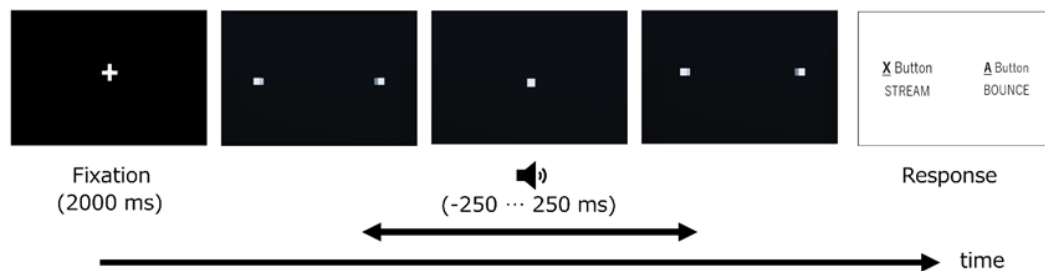


図 1, 実験のイメージ図

＜実験条件＞

実験1: SOA(ms) が ± 250 , ± 225 , ± 200 , ± 175 , ± 150 , ± 125 , ± 100 , ± 75 , ± 50 , ± 25

および 0 ms の 21 条件から毎試行ランダムに決め、音を提示するタイミングが変化させた。SOA がマイナスである場合、物体が衝突するよりも前に音が提示された。

実験2: 各試行において、音を提示する位置は、頭の位置を中心として、前、後、左、右、上、下、中心からランダムに決まった。SOA は 0 ms で固定した。

＜結果と考察＞

音の提示タイミングが視覚情報である物体の衝突に先行すると Bounce 応答確率が高くなった(図 2-左)。ヒトは音の提示から認識までのラグがある[2]ためであると考えられる。また、音の提示位置による有意差は見られなかった(図 2-右)が、これは腹話術効果[3]によって音の提示位置が視覚刺激の位置にシフトしたためであると考えられる。

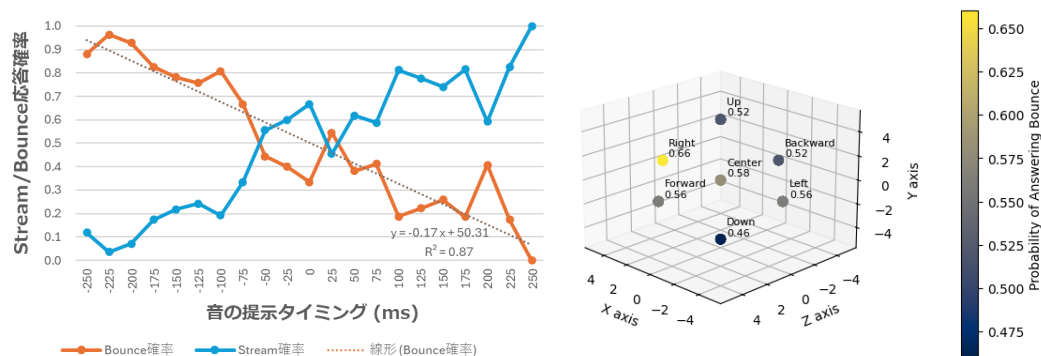


図 2, 音提示タイミング(左)または音提示位置(右)による Bounce 応答確率

【参考文献】

- [1] Sekuler R, Sekuler AB, Lau R. Nature, 385(6614), 308, 1997.
- [2] Hayashi, R., Murakami, I. Sci Rep 9, 3829 (2019).
- [3] Bertelson, P., Vroomen, J., De Gelder, B. et al. Perception & Psychophysics 62, 321–332 (2000)

目次

1	序論.....	1
1.1	背景.....	1
1.2	目的.....	1
2	理論的背景.....	2
2.1	クロスモーダル現象.....	2
2.2	ストリームバウンス効果.....	4
3	先行研究.....	5
3.1	視聴覚刺激の時間同期処理.....	5
3.2	音の空間位置特性.....	7
4	立体物の錯覚への影響.....	9
4.1	Unity について.....	9
4.2	実験環境.....	9
4.3	実験方法.....	9
4.4	実験結果.....	10
4.5	考察.....	11
5	VR におけるストリームバウンス効果.....	13
5.1	VR について.....	13
5.2	実験環境.....	14
5.3	実験方法.....	15
5.4	実験結果.....	17
5.5	考察.....	18
6	刺激音提示位置による錯覚への影響.....	23
6.1	実験環境.....	23
6.2	実験方法.....	23
6.3	実験結果.....	25
6.4	考察.....	27
7	結論と今後の展望.....	31
8	参考文献.....	33
9	謝辞.....	35

1 序論

1.1 背景

ヒトは五感(視覚, 聴覚, 味覚, 嗅覚, 触覚)から得られる情報を知覚することで普段の生活を送ることができる。知覚研究は脳の五感処理メカニズムを解明するためにこれまでも数多く行われてきたが、単独の感覚(ユニモーダル)についての研究が多く、複数の感覚によって認識をする、マルチモーダルな知覚研究はあまりされていない。ヒトは一日の大半を複数の感覚器を同時に活用しながら生活していることはあきらかであり、テレビを見るときは視覚と聴覚、ご飯を食べるときは視覚、嗅覚、味覚、触覚を使い行動している。マルチモーダルな情報処理がユニモーダルな感覚の単純な和ではない[1]ことはヒトの多感覚統合が重要であることを示しており、複数の感覚情報が脳内でどのように処理され統合されるのかを定量的に調べる研究はヒトの複雑な情報処理システムを理解するうえで重要である。

また、ヒトが錯覚することについて研究することはヒトの知覚特性や脳の処理を調べる近道であると考えられる。なぜならヒトは外部から得た感覚情報によりスパイクを起こし、脳に情報が伝達されるからである[2]。つまり普段とは違う脳の活動やメカニズムを錯覚では研究することができるということである。特に錯視や錯聴などのユニモーダルな錯覚ではなく、複数の感覚情報から起きる錯覚であるクロスモーダル現象はヒトの複雑な情報処理の理解に役立つまたは知覚を増幅させることが期待される。

1.2 目的

本研究では、クロスモーダル現象の一種であるストリームバウンス効果を取り扱い、VR技術を用いることで、先行研究例の少ない仮想現実空間上でのストリームバウンス効果の再現をする。そして、視聴覚同期性という観点から、錯覚の実験内で音を提示するタイミング、音源位置の三次元化という二つのアプローチによって、錯覚への影響を定量的に測定し、ストリームバウンス効果におけるヒトの知覚の特性とメカニズムを解明することを目的とする。

2 理論的背景

ここでは、本研究の重点であるクロスモーダル現象およびストリームバウンス効果の基礎的内容について紹介する。

2.1 クロスモーダル現象

クロスモーダル現象とは、ある感覚情報が他の感覚情報(メカニズム)に干渉して感覚情報自体が変化する現象を指す[1]。視覚と触覚の感覚情報によって起こるクロスモーダル現象の一例として、見かけの大きさによる重さの錯覚[3]というものがある。Ellisが発見したこの現象は紐にぶら下げた同じ重さで違う大きさの物体が二つあるとき、物体を紐ごしに持ち上げると小さい物体のほうが重く感じるという錯覚である(図1)。この研究のルーツとなるシャルパンティエ効果[4]は1891年に Charpentier によって報告された。シャルパンティエ効果は錯触の一つであり、二つの同じ重さで違う大きさの物体を、目を閉じた状態で持ち上げたとき小さい物体のほうが重く感じるという錯覚である(図2)。

両者の結果は一見同じように見えるが錯覚のメカニズムは異なる。シャルパンティエ効果では物体をじかに触り、大きさと重さを感じ判断するが、Ellisの実験では物体を間接的に持ち上げているため、重さは感じるができるが、大きさによる判断は目で見ることになる。つまり、錯覚のメカニズムは、実際には同じ重さの物体(触覚情報)が目で見た物体の大きさ(視覚情報)によって変化したということである。小さい物体が重く感じる要因としては、物体の大きさから期待される重さとの齟齬により、重さが強調されていると考察されている。また、物体の密度を重さの判断に用いていることが錯覚に関係していると考えられている[5]。

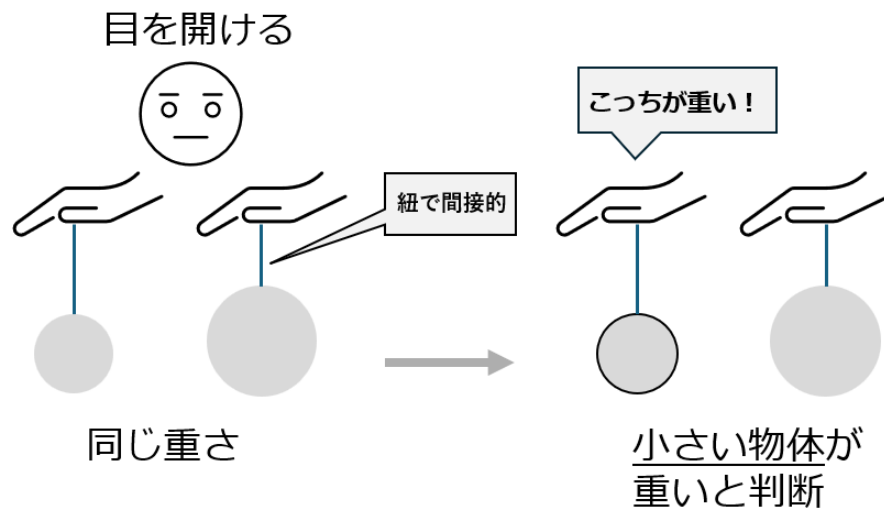


図1, Ellisの実験モデル図 [3]

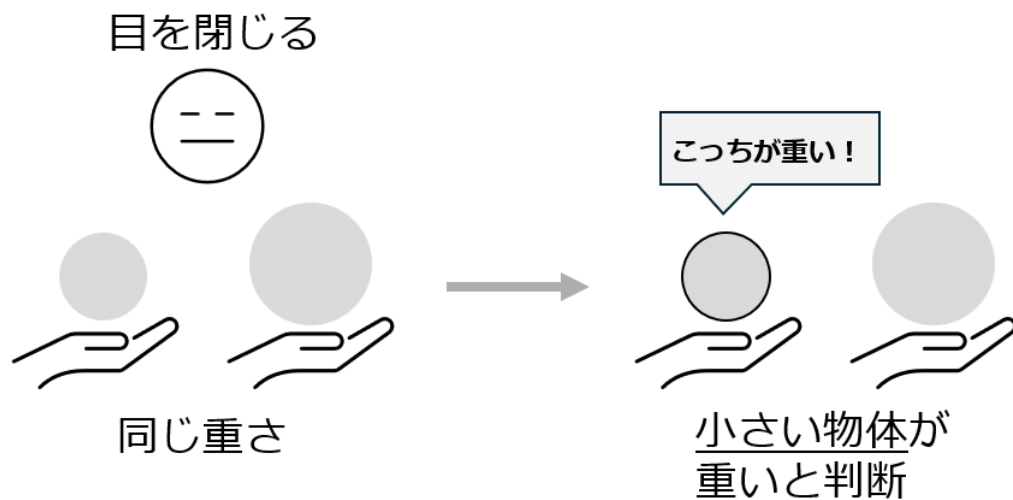


図 2, シャルパンティエ効果のモデル図 [4]

他にも、視覚と味覚から起きるクロスモーダル現象の例を挙げると、白ワインの色を赤く着色して飲んでもらうとワインに詳しい人でも赤ワインだと勘違いをするという研究報告がある(図 3)[6]。これは、ワインの味(味覚)がワインの見た目(視覚情報)によって、変化したということである。

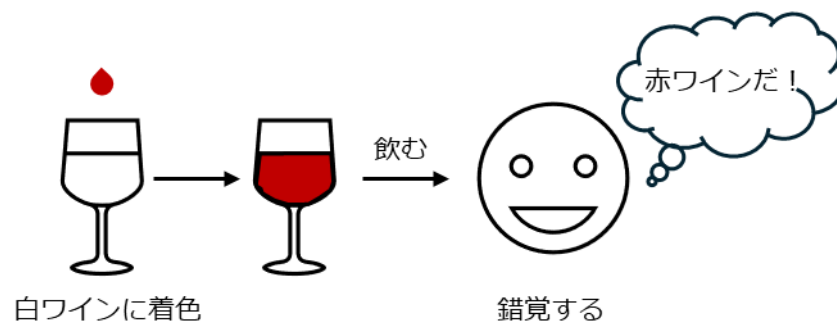


図 3, Morrot による実験モデル図 [6]

このようにクロスモーダル現象とは、ある感覚情報における知覚が同時に受け取る他の感覚に対する刺激によって変化した結果生じた知覚であり、ヒトの知覚を意図的に変化させることのできる唯一の方法ともいえる。近年では、クロスモーダル現象やマルチモーダルに関する研究が多様化しており、今後 VR や AR 技術を取り入れていくとで、クロスモーダル現象について定量的な測定による定式化や新たな知見が得られ、ヒューマンインタフェース開発などに役立つことが期待されている[1]。

2.2 ストリームバウンス効果

ストリームバウンス効果とは、クロスモーダル現象の一種であり、視覚と聴覚が関与する錯覚である。この効果は、二つの平面（円または正方形）が同一直線上で互いにすれ違う瞬間に音を提示することで、実際にはすれ違っているにもかかわらず、跳ね返ったように錯覚する現象である。

ストリームバウンス効果の元祖となる現象を Sekuler は 1997 年に報告している[7]。実験では、コンピューターによって、互いに近づき、一致し、そして離れて移動する二つの同じ物体が表示された。物体が一致したとき、または一致の前後 150 ms で短いクリック音を鳴らすことにより、音を鳴らさない時に比べて、物体が跳ね返る知覚(Bouncing)が促進された。さらに物体の一致と一致の前に音を提示する場合 Bouncing 確率は同程度に高いが、一致の後に音を提示すると Bouncing 確率が低くなることが明らかになった(図 4)。

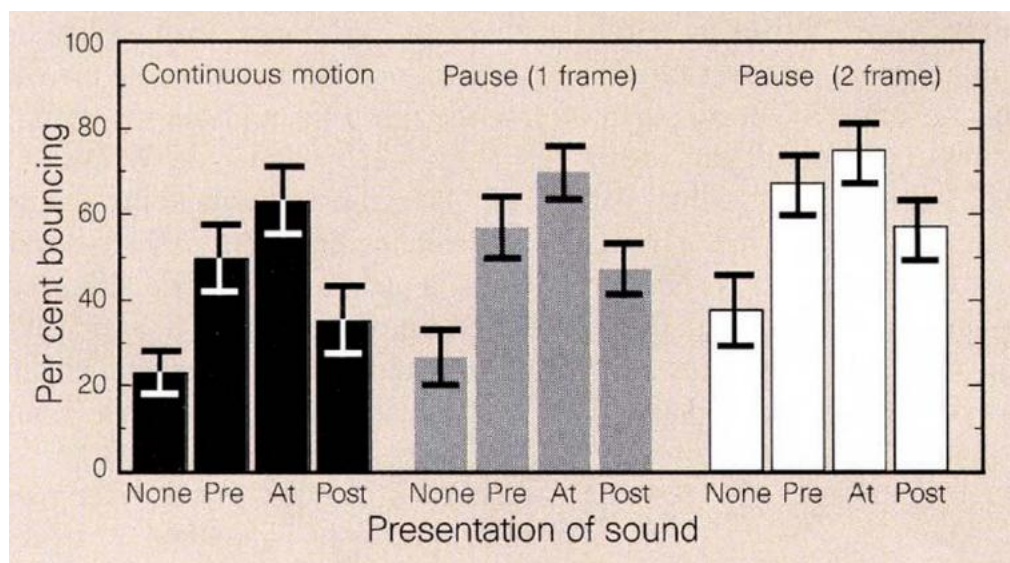


図 4, Bouncing と報告した確率。物体の動きが連続的なのか、一致の瞬間に何フレームか停止するか条件に分け、音無し、一致、一致の前後の音についての評価を示す。[7]

また Sekuler は、視覚運動に対して、音の影響が単に注意効果を表している可能性があると考え、新たな実験を行った。音を鳴らし続け、一致の瞬間だけ音を消すパターンを取り入れることで、ストリームバウンス効果に対する影響を調べたが、音のオフセットは音を提示しない場合と同程度の結果をもたらした。つまり、ストリームバウンス効果において、音が視覚運動に与える影響は、一致の瞬間に注意を引き付けているというわけではなく、物体間の衝突を知らせる音響のイベントが必要であると示唆している[7]。

3 先行研究

ここでは、研究テーマに関連する先行研究について紹介する。内容は主に二つあり、ストリームバウンス効果に対する刺激提示時間間隔(Stimulus onset asynchrony : SOA)の影響および視聴覚同時性判断における音の空間位置特性に関する研究についてまとめる。

3.1 視聴覚刺激の時間同期処理

ストリームバウンス効果について、Sekuler は音の提示が視覚情報と同期または前後することにより Bounce 応答確率が変化することを明らかにしました[7]が、詳しい SOA による反応の違いは未明であった。Takei らはヒトの知覚情報処理時間がどのように決定されているか、またその処理メカニズムを調べるため、ストリームバウンス効果を対象に音のような修飾刺激の前後の入力がどのように影響を与えるか調べた[8]。実験では音修飾刺激条件として、無音条件と音とディスクの接触の SOA を変化させた条件を設けた。SOA は-140, -70, 0, 70, 140 ms の 5 条件設けられ、知覚の処理時間を検討するために Stream か Bounce を判断するまでの反応時間を取得した。

実験の結果、音が先行する条件において、Bounce 応答確率が最も高くなった(図 5)。また、Bounce 知覚が生じた場合にのみ、反応時間が変化した(図 6)。

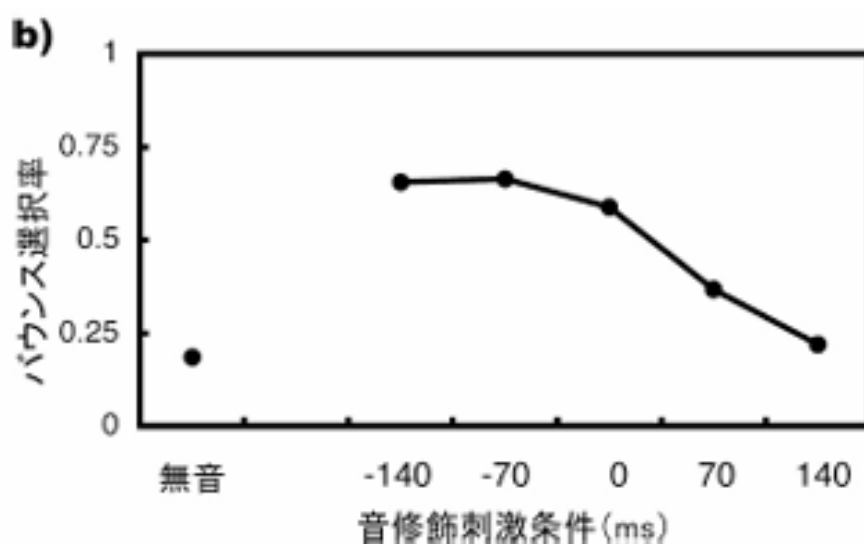


図 5, 音修飾刺激条件に対するバウンス選択率 [8]

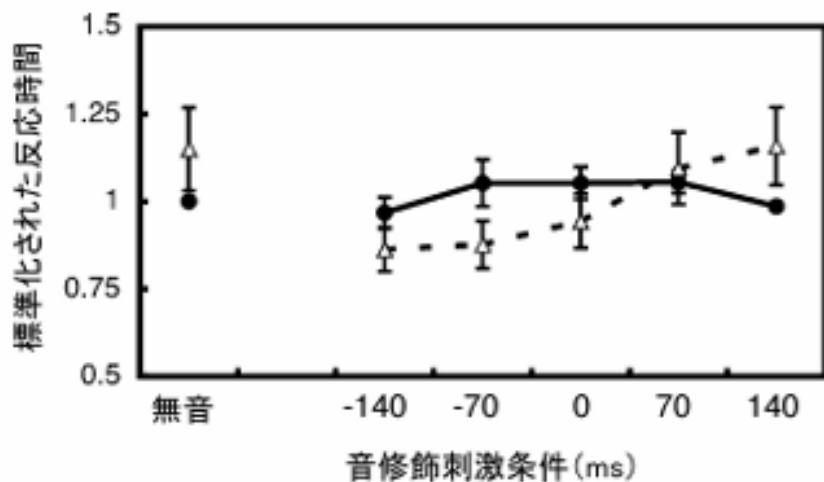


図 6, 音修飾刺激条件に対する反応時間 [8]

結果から、あるイベントの知覚処理時間は一定ではなく、イベントに関係する入力によって処理時間が変化することを示唆している。

Takeshima はストリームバウンス効果を用いて、視野ごとの主観的同時性(PSS)、時間結合窓(TBW)、また迅速な再キャリブレーションの調査を行った[9]が、PSS と TBW は中心視野、周辺視野のどちらにおいても差がなく 0 ms にほぼ等しかった。そして迅速な再キャリブレーションは中心視野条件でのみ起きることを明らかにした。また Bounce 応答確率について、 ± 510 、 ± 260 、 ± 130 、 ± 50 、および 0 ミリ秒の 9 つの SOA 条件で求められた(図 7)。

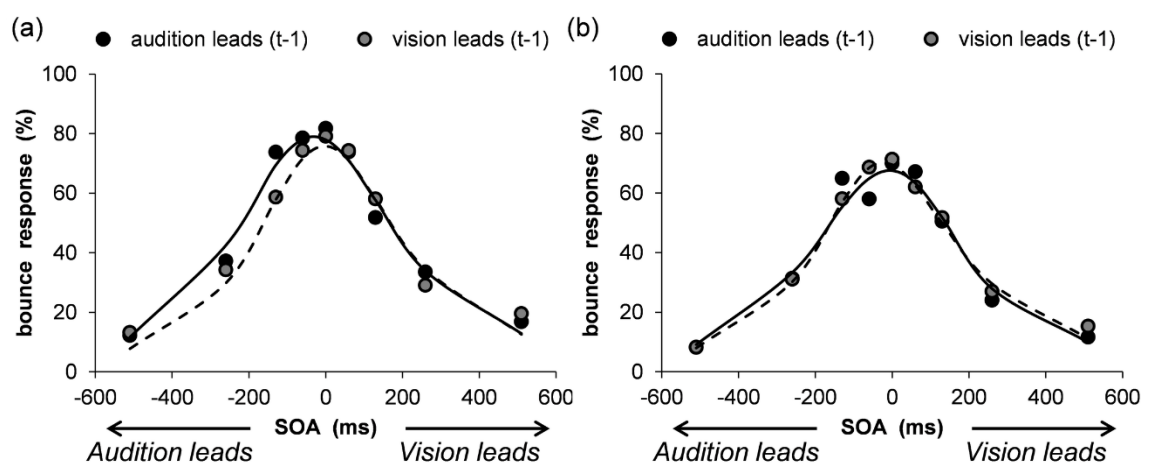


図 7, SOA 条件ごとの Bounce 応答確率 (a)中心視野 (b)周辺視野 [9]

3.2 音の空間位置特性

視聴覚同時性判断において Zampini らは、刺激の位置による多感覚統合への影響を調べた[10]。実験 1 では、視覚刺激と聴覚刺激が同じ位置から提示される場合と異なる位置から提示される場合で、同時性の知覚がどのように変化するかを調査した。実験は完全に暗い防音室で行われ、赤色の LED が参加者の前方 62cm に配置され、視覚刺激として使用された。聴覚刺激は 9 ミリ秒のホワイトノイズバーストで、視覚刺激は 9 ミリ秒の LED 点灯だった。

参加者は、視覚刺激と聴覚刺激が同期して提示されたか、同期せずに提示されたかを判断するタスクを行った。刺激は同じ位置から提示される場合と異なる位置から提示される場合があった。各試行の開始時に固定光が点灯し、参加者はこの中央の赤色 LED に視線を固定するよう指示された。最初の刺激（または SOA が 0 ミリ秒の試行では視覚と聴覚両方の刺激）が左または右から提示され、SOA が 0 ミリ秒を超える試行では、指定された SOA 後に 2 番目の刺激が提示された。参加者は、刺激が同時に提示されたか連続して提示されたかを判断した(図 8)。SOA は -200, -70, -30, -20, 0, 20, 30, 70, 200 (ms) の 9 条件が設けられた。

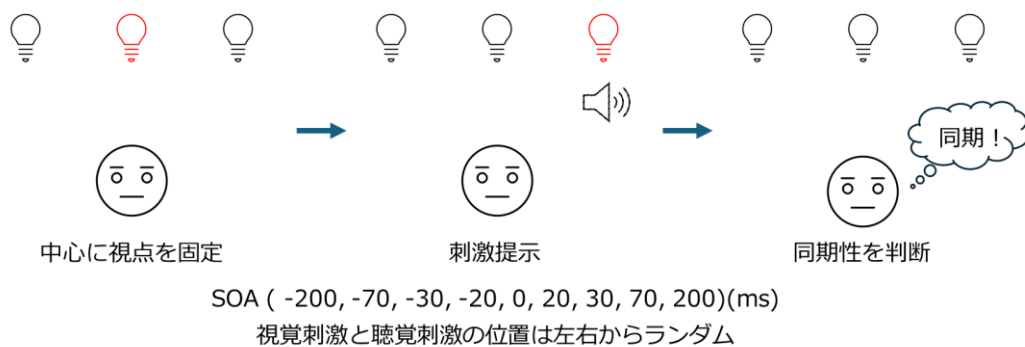


図 8, Zampini の実験モデル図 [10]

結果として、参加者は視覚刺激と聴覚刺激が同じ位置から提示された場合、異なる位置から提示された場合よりも同時性を感じやすいことが示された(図 9)。この研究は、視覚と聴覚の統合が空間的な要因によってどのように影響を受けるかを明らかにするものであり、マルチモーダル知覚の理解に重要な貢献をしている。

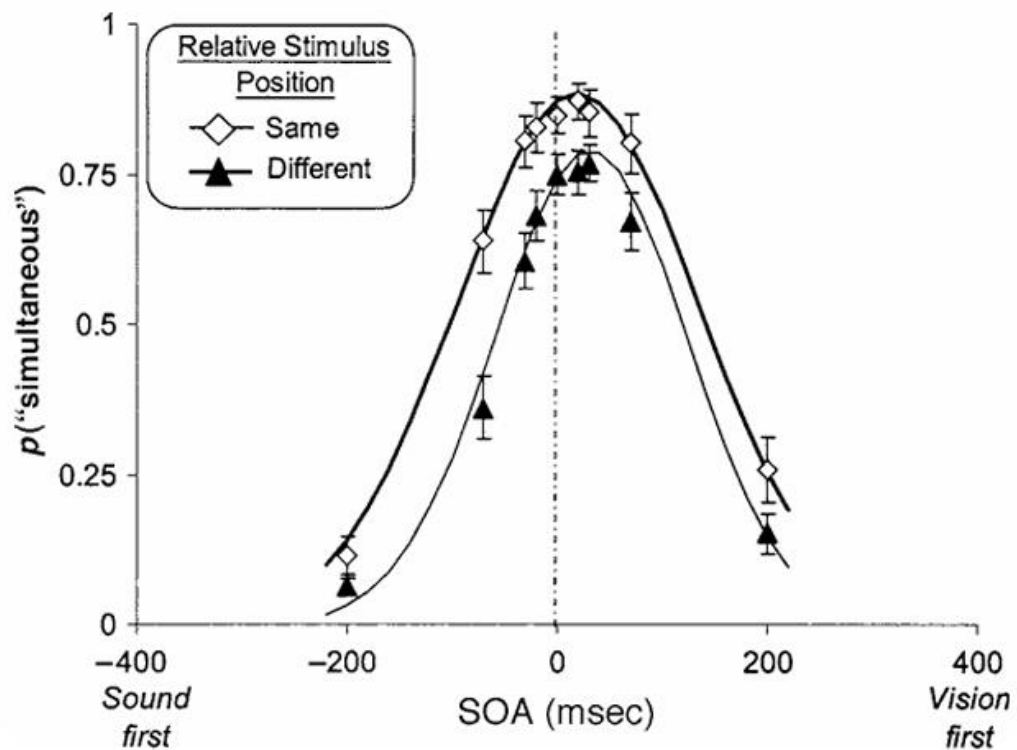


図 9, 刺激提示位置および SOA による同時性判断 [10]

4 立体物の錯覚への影響

二次元平面上に立体物を用いることで、奥行き知覚がストリームバウンス効果にどのような影響がでるのかを調査した。

4.1 Unity について

本研究の実験はすべて Unity によって制作した。Unity とは、ユニティ・テクノロジー社が開発・提供するゲーム開発プラットフォームである。2D, 3D, VR, AR, MR に対応しており、ゲーム制作だけでなく、映像、医療、建築、自動車、製造業、教育などの多種多様な分野で活用されている[11]。

4.2 実験環境

被験者は成人男性 3 名で、全員晴眼者（眼鏡、コンタクトレンズ等の矯正も含む）である。実験開始前には、口頭で PC の操作方法や実験の構造について被験者に指示し、各被験者が適切に実験を操作できるようにした。実験は、横浜市立大学理学研究棟の一室で行った。また、被験者が静かに実験に集中できるように、必要以上に話しかけることはしなかった。本研究で用いたノート PC のディスプレイサイズは、15.6 インチである。実験では、音を提示するため、ヘッドホン(Steel Series Arctis7)を使用した。

4.3 実験方法と条件

実験の方法、条件とそのモデル図を以下に示す(図 10)。試行回数は一人 100 回行った。

実験方法

1. 固定ターゲット表示 (2000 ms)
2. すれ違う立方体を観察した
3. 2 つの立方体がすれ違って見えた(F キー)のか、跳ね返って見えた(J キー)のか答えた

実験条件

- 本研究の SOA は、先行研究と比較して幅広く、そして細かく ± 250 , ± 225 , ± 200 , ± 175 , ± 150 , ± 125 , ± 100 , ± 75 , ± 50 , ± 25 および 0 ms の 21 条件が毎試行ランダムに決められ、音を提示するタイミングが変化させた。
- SOA がマイナスである場合、物体が衝突するよりも前に音が提示された。

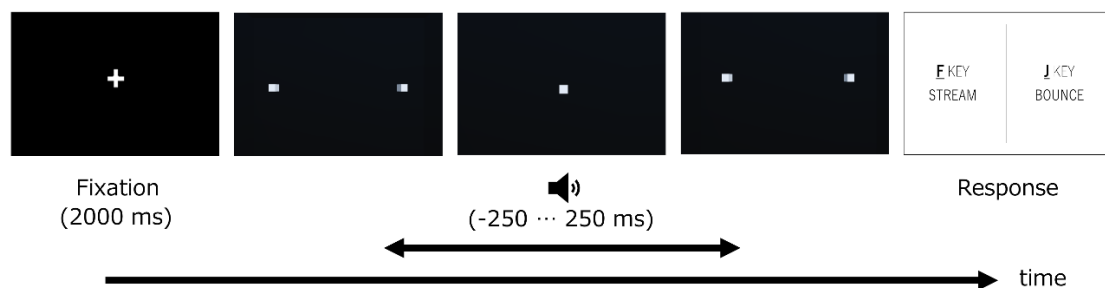


図 10, 実験のモデル図

4.4 実験結果

被験者が各試行で Stream または Bounce と答えたときの SOA の条件ごとにデータを取
得した。データは Unity のプログラムによって実験終了後に csv ファイルとして保存され
た。得られた各被験者のデータをまとめ、音を提示するタイミングごとの Bounce 応答と
Stream 応答の箱ひげ図を示した(図 11)。

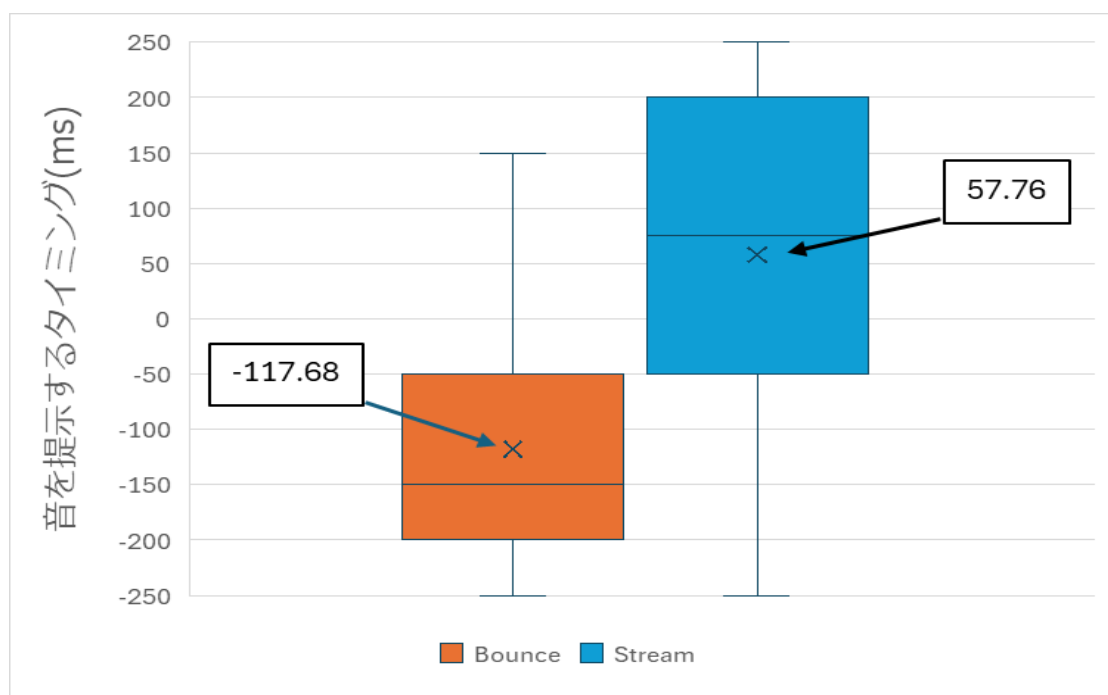


図 11, 音提示タイミングによる Stream/Bounce 応答の箱ひげ図

また SOA(音提示タイミング)の条件ごとに Stream と Bounce の応答する確率を求めた(図 12)。

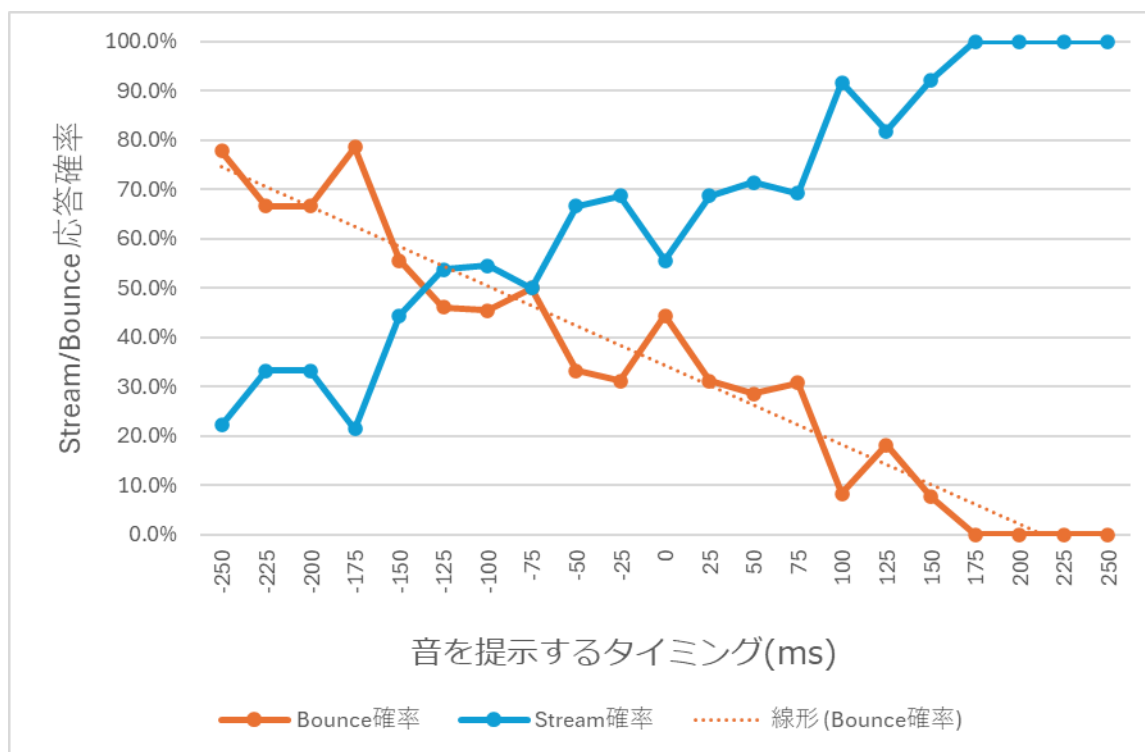


図 12, 音提示タイミングによる Stream/Bounce 応答確率

4.5 考察

はじめに、ストリームバウンス効果は動く対象が平面でなくても起きることが確認できた。図 11 から、刺激音を提示するタイミングが立方体の衝突よりも早い時に Bounce の応答が多く、衝突より後に音が提示された場合 Stream の回答が多くなった。また、図 12 より Bounce 応答確率が最も高かったのは SOA が -250 ms, -175 ms の時であり、175 ms 以降の SOA 条件では Bounce 応答確率は 0 %であった。

Sekuler の結果(図 4)と比較すると、SOA の条件が -150 ms のとき Bounce 応答確率は二つの実験で 50 %を示している。しかし、0 ms, 150 ms のとき本研究の Bounce 応答確率は低く、Sekuler の確率分布が山型であるのに対し、本研究では 0 ms 条件で Bounce 応答確率が最高にならず、直線に近似することが可能である。また、Takei らの結果(図 5)と比較すると、全体の Bounce 応答確率は低くなっているが、SOA の条件範囲内において音の提示が早いほど Bounce 応答確率が高くなる傾向は一致した。Takeshima の結果(図 7)は Bounce 応答確率分布が山型を描いており、本研究の結果とは違う傾向が見られた。特に -260 ms の SOA 条件では、本研究の -250 ms の SOA 条件における Bounce 応答確率が

80 %であるのに対し、Takeshima の結果では、40 %以下である。本研究の結果の中で-175 ms と-150 ms のバウンス応答確率を比較すると約 25 %の違いがあるが、10 ms の違いによって、ここまでの差があることは-250 ms がヒトの視覚と聴覚を同期させ、錯覚させるための許容範囲であり、その先は崖のように Bounce 応答確率が下がるということが考えられる。そのため、SOA の条件範囲についてはより細かく調べる必要がある。

ただ、この実験は立体物を使用した場合においてもストリームバウンス効果は確認できるのかを調べるための予備実験として行ったものであり、被験者の数が少ないため SOA 条件ごとのデータの数が多分ではなかったないかと考えられる。また、Takeshima の研究から、視野条件によって Stream/Bounce の応答確率が変わることから、ディスプレイの大きさや周辺環境の複雑さ、キーボードによる操作が実験の結果に影響した可能性がある。

二次元平面上に立体物を用いることでストリームバウンス効果にどのような影響が出るのかを調査したが、二つの立方体がぶつかり重なった瞬間(図 13)、二次元平面上において、奥行き感覚のない平面の正方形に見え、予備実験としての目的であったストリームバウンス効果における立体視の影響が測定できていたかは定かではない。

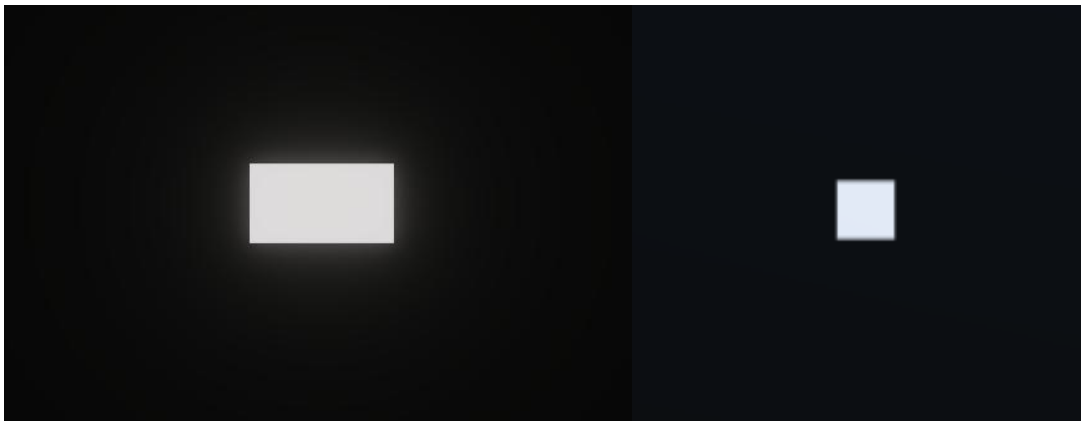


図 13, 立方体の衝突の瞬間または直後（左）と衝突し重なった瞬間（右）

5 仮想現実空間におけるストリームバウンス効果

4 章にて行った実験をバーチャルリアリティ空間上に移し、実際に奥行きのある空間内にてストリームバウンス効果にどのような影響が出るのかを調べる

5.1 VR について

バーチャルリアリティ (Virtual Reality:VR) とは、「見かけは現実ではないが、効果としてあるいは実質的に現実である」[12]という意味である。バーチャルという単語の使い方の実例として、近年普及率が高まっているキャッシュレス決済や PASMO や Suica などの交通系 IC カードはバーチャルマネーであり、その意味は「見かけ上お金ではないが、効果としてお金である」ということである。仮想現実という翻訳は正しくないとされている。なぜなら、バーチャルを仮想と翻訳する場合、修飾された名詞はこの世に存在しないことになってしまうからである。ただ日本語には本来の意味に相当する単語が存在しないため、仮想という使われ方をしている。

バーチャルリアリティの歴史は長く、用語が使われ始めたのは 1989 年のことである。世界に広く認知されだしたのは、2016 年ソニー・インタラクティブエンターテインメント (SIE) が、家庭用ゲーム機である PS4 向け VR ヘッドマウントディスプレイ (HMD)PlayStation VR の価格と発売日を発表したことによる影響が大きい。この年には他にも、Oculus(現 Meta)社から、Oculus Rift が販売され、またバーチャルリアリティを体験するイベントや施設が次々と展開された。

主にゲーム業界において、ユーザーの関心が集まる VR だが、医学教育では HMD を用いた手術のシミュレーション、製造業の分野では、製品設計や製造工程をバーチャルに行う技術「デジタルエンジニアリング」として活用されています。また、研究分野としてのバーチャルリアリティも成長しており、日本ではバーチャルリアリティ学会、北米では IEEE や SIGGRAPH などでも国際会議が開かれている[13]。

実質的な現実を表現することのできる VR は知覚研究に役立つと考えられており、特にユニモーダルな知覚ではなく、マルチモーダルな知覚に対して、人間の複雑な知覚特性を探ることに適している。

近年では、メタバースという「多人数が同時にオンラインで社会的な活動が可能な 3D バーチャル空間」[14]の市場が急激に拡大しており、日本においてもメタバース市場（メタバースプラットフォーム、プラットフォーム以外（コンテンツ、インフラ）、メタバースサービスで利用される XR (VR、AR、MR) 機器の合計）は、2027 年度には 2 兆円まで拡大すると予測されている[15](図 14)。

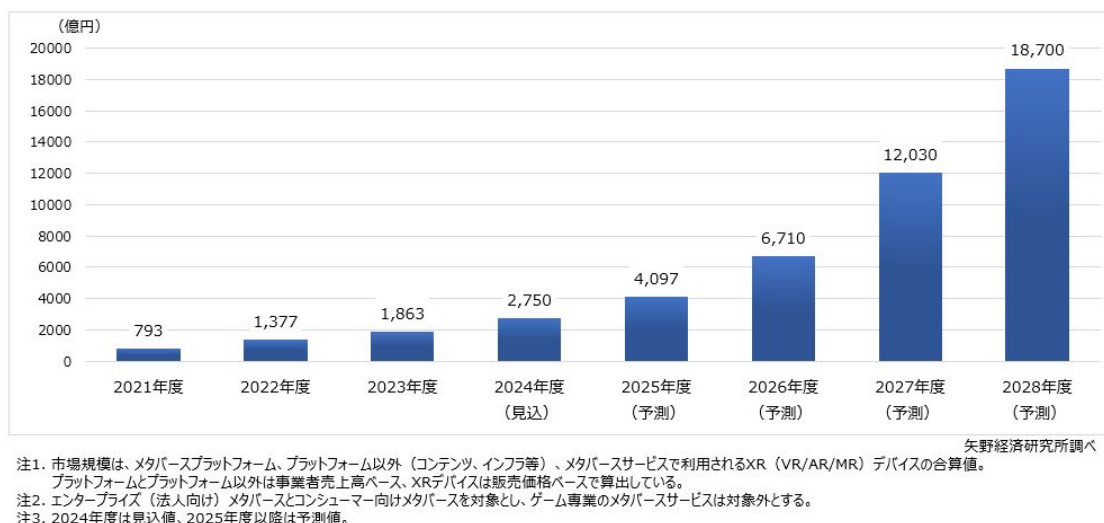


図 14, 日本のメタバース市場規模（売上高）の推移と予測 [15]

5.2 実験環境

被験者は成人男性 5 名、女性 1 名で、全員晴眼者（眼鏡、コンタクトレンズ等の矯正も含む）である。実験開始前には、口頭で VR の機器にういて操作方法や実験の構造について被験者に指示し、各被験者が適切に実験を操作できるようにした。必要に応じて実験のデモンストレーションを行った。実験は、横浜市立大学理学研究棟の一室で行った。また、被験者が静かに実験に集中できるように、必要以上に話しかけることはしなかった。本研究で用いた VR 機器は Meta Quest 3 である[16](図 15)。実験では、音を提示するため、ヘッドホン(Steel Series Arctis7)を使用した。



図 15, Meta Quest 3 [16]

5.3 実験方法

実験の方法、条件とそのモデル図を以下に示す(図 16)。また、実験内の環境を理解しやすくするため空間の写真に掲載する(図 17, 18)。試行回数は一人 105 回行った。

実験方法

1. 固定ターゲット表示 (2000 ms)
2. すれ違う立方体を観察した
3. 2つの立方体がすれ違って見えた(X ボタン)のか、跳ね返って見えた(A ボタン)のかコントローラーのボタンを押して答えた

実験条件

- 本研究の SOA は、先行研究と比較して幅広く、そして細かく ± 250 , ± 225 , ± 200 , ± 175 , ± 150 , ± 125 , ± 100 , ± 75 , ± 50 , ± 25 および 0 ms の 21 条件が毎試行ランダムに決められ、音を提示するタイミングが変化させた。
- SOA がマイナスである場合、物体が衝突するよりも前に音が提示された。
- 被験者は実験中空間内を自由に見渡すことが可能でしたが、実験中は固定のターゲットに目を向けるように指示した。
- 被験者が操作のミスをした場合、その数だけ多く試行回数を増やした。

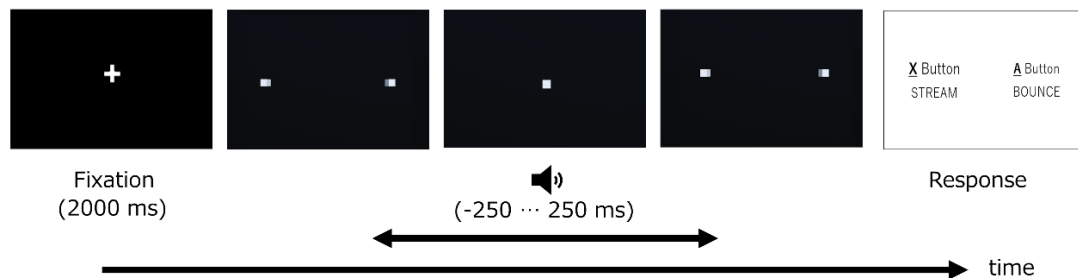


図 16, 実験のモデル図

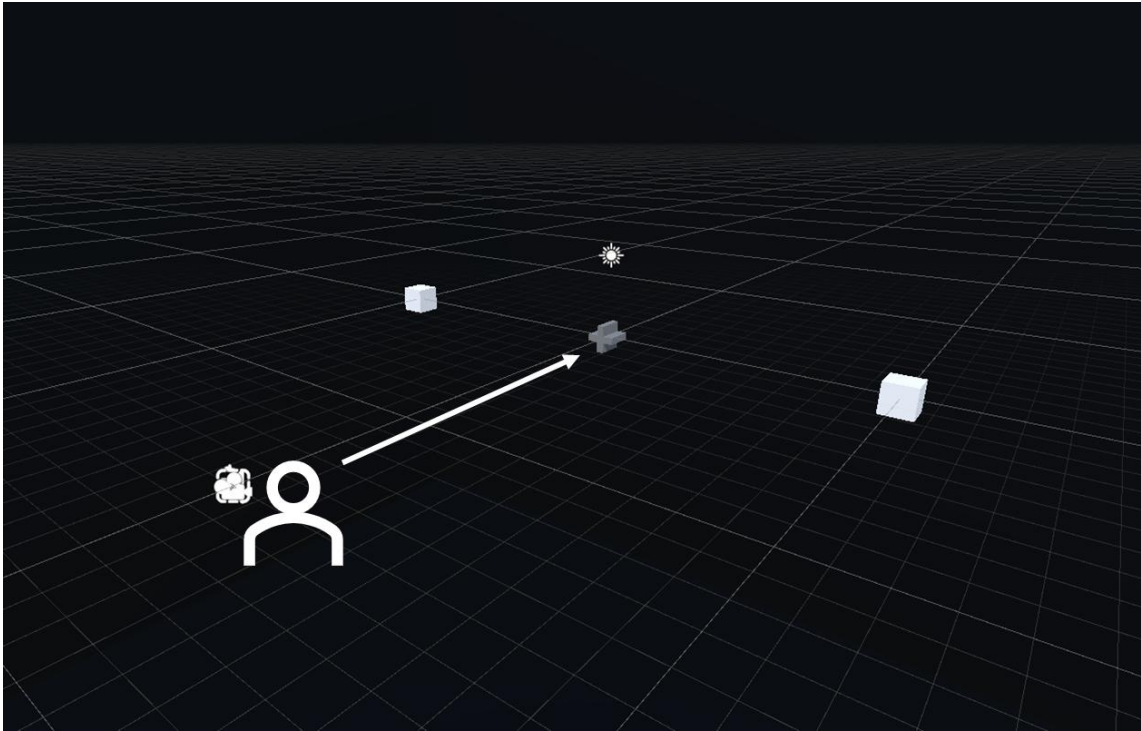


図 17, 実験環境 実際にはグリッド線は描画されなかった。

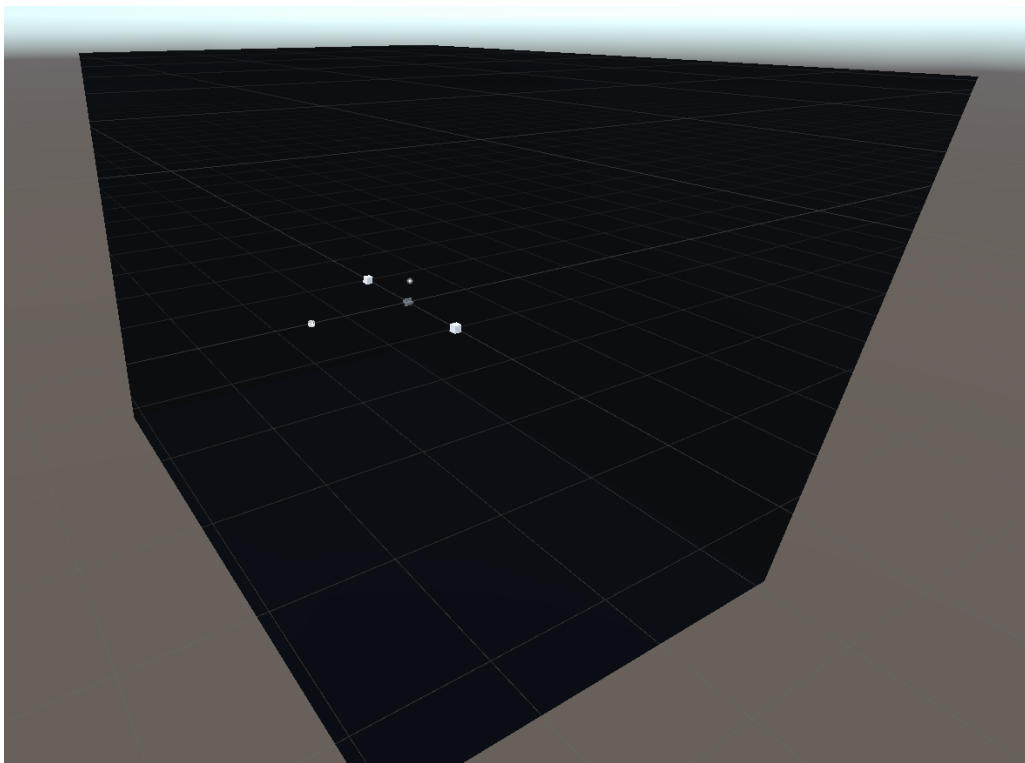


図 18, 実験環境は黒い部屋に囲まれており、被験者から部屋の外を見ることはできなかった。

5.4 実験結果

被験者が各試行で Stream または Bounce と答えたときの SOA の条件ごとにデータを取
得した。データは Unity のプログラムによって実験終了後に csv ファイルとして保存され
た。得られた各被験者のデータをまとめ、音を提示するタイミングごとの Bounce 応答と
Stream 応答の箱ひげ図を示した(図 19)。

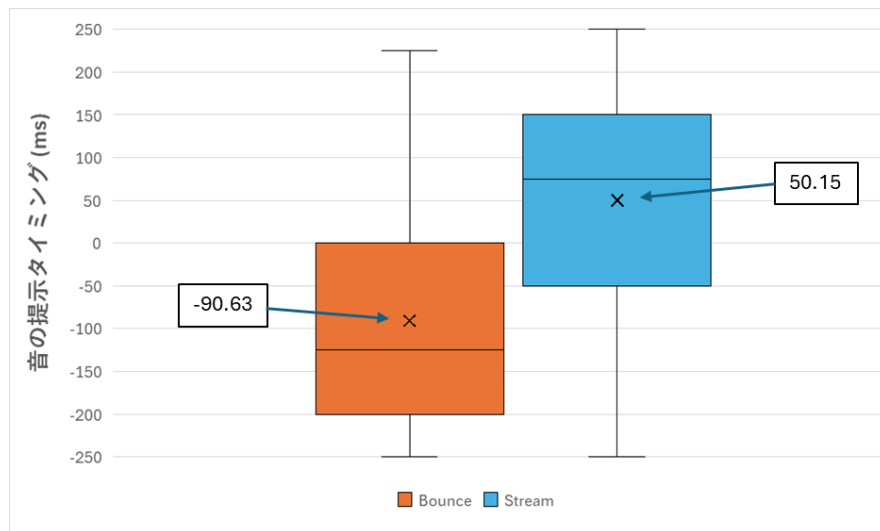


図 19, 音提示タイミングによる Stream/Bounce 応答の箱ひげ図

また SOA(音提示タイミング)の条件ごとに Stream と Bounce の応答する確率を求めた
(図 20)。

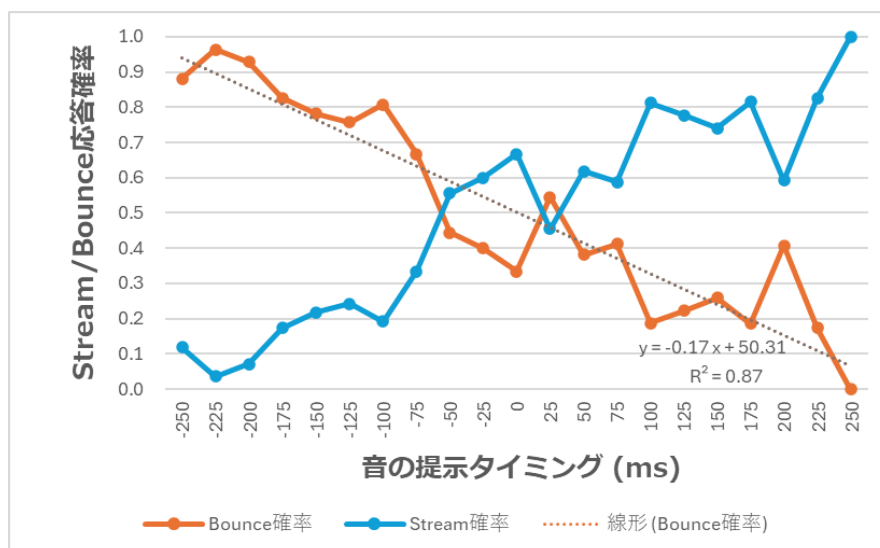


図 20, 音提示タイミングによる Stream/Bounce 応答確率

5.5 考察

はじめに、ストリームバウンス効果はバーチャルリアリティ上でも起きることが確認できた。図 19 から、刺激音を提示するタイミングが立方体の衝突よりも早い時に Bounce の応答が多く、衝突より後に音が提示された場合 Stream の回答が多くなった。また、Bounce 応答確率が最も高かったのは図 20 から -225 ms の SOA 条件であった。Bounce 応答確率の近似直線の式について、 y を Bounce 応答確率 P_B 、 x を SOA t_{SOA} とすると、

$$P_B = -0.17t_{SOA} + 50.31 \quad (1)$$

であり、SOA が 0 ms であるとき Bounce の応答確率が約 50 %であることを表している。

4 章の実験結果(図 11, 12)と比較すると、Stream/Bounce 応答の傾向に変わりはないことが分かる。厳密にグラフを見比べると、バーチャルリアリティでの実験では SOA の条件が -100 ms 以上早いとき Bounce 応答確率が 70 %を超えており、二次元ディスプレイ上での実験に比べて Bounce 応答確率が高まっている。また、二つの実験から SOA の条件が -50 ms よりも遅い場合 Stream 応答の確率が Bounce 応答確率を上回っており、これはどの先行研究と比較しても違っている。つまり、本来視覚と聴覚が同期する衝突の瞬間である 0 ms のタイミングに音を提示した場合でも Stream と回答する場合が多かったということである。

しかし、Sekuler はストリームバウンス効果の Bounce 応答確率を最も高めるのは音刺激と視覚情報が同期するときであると考えている。また、先行研究者はなぜ、ストリームバウンス効果は音が視覚情報に先行するときに Bounce 応答確率が高くなるのかについては考察をしていない。そのため、本研究では視覚と聴覚の同期メカニズムや 2 次元と 3 次元空間での知覚特性の違いを踏まえ、なぜ本研究の結果では完全な同期のタイミングである 0 ms の条件で Bounce 応答確率が下がるのか、そしてなぜ、音が視覚刺激(衝突)に先行する場合 Bounce 応答確率が高くなるのかを考察していく。

● 3 次元と 2 次元間の判断時間差

Kawamura は二次元画像と三次元画像のカテゴリー判断において三次元画像のほうが判断に時間がかかることを明らかにした[17](図 21)。これは、三次元画像が奥行きの知覚を理解しやすい一方で奥行き処理による負荷が増大し画像の認識判断までにかかる時間が増えた可能性があることを示唆している。また試行回数によって反応時間の差がなくなっていくことが分かる。3 次元画像は両眼視差情報を持ち、奥行き方向を知覚する手がかりが多い。このため、対象の立体的な特徴が理解されやすく、視覚教材として優れた認知的機能をもっていると考えられる[18]。

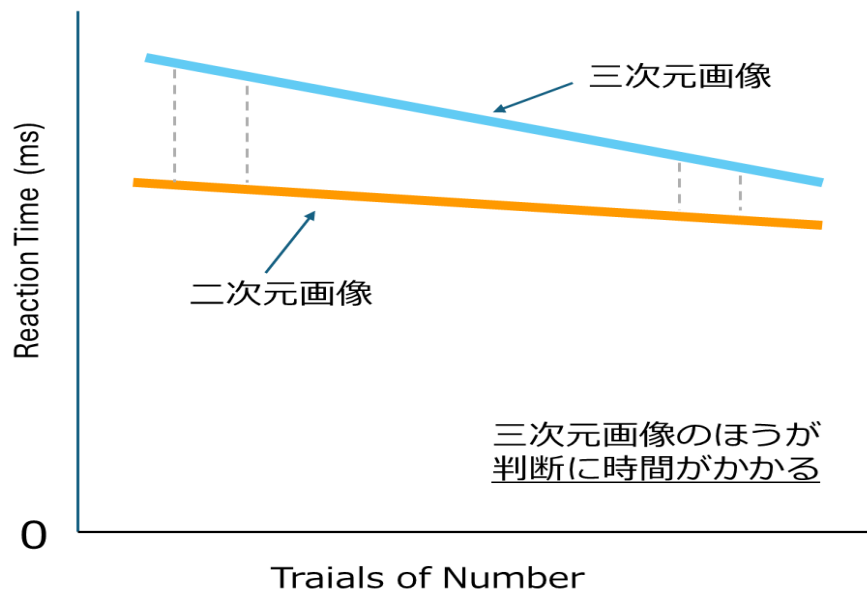


図 21, 試行回数による二次元画像と三次元画像の判断時間差 [17]を参考に作成

つまり、本研究において立体物を用いたまたはバーチャルリアリティでの実験を行ったことにより、視覚の三次元情報処理に時間がかかり、完全な同期のタイミングである 0 ms の条件で音を提示した場合でも Bounce 応答確率が下がるのではないかと考えられる。しかし、立方体同士が衝突する瞬間を認識するまでにラグが生じる場合、視覚と聴覚の同期するタイミングが遅れ Stream/Bounce 応答確率のグラフ(図 12, 20)は、SOA がプラス側の条件において Bounce 応答確率が高くなると考えられる。

図 21 は SOA の条件がプラスだった場合(衝突のあとに音が提示される条件であるとき)、視覚刺激である衝突を認識するタイミングに同期することが可能であることを示す。また、図 22 には、SOA の条件がマイナスだった場合(衝突の前に音が提示される条件であるとき)、聴覚刺激には視覚刺激と同期することができる時間の範囲があり、衝突を認識するころには聴覚刺激による Bounce 応答の誘発効果が切れているのではないかと考えられる。本研究の結果(図 20)では、SOA 条件がプラスの場合、25 ms であるときのみ、Bounce 応答が Stream 応答を上回っている。これは、視覚刺激の認識にかかるまでの時間のラグが 25 msであることを示している可能性がある。

しかし、この考察は本研究の結果に矛盾する部分を持ち合わせている。なぜなら、実際には Bounce 応答確率は SOA の条件がマイナスであるときに高く、プラスであるときに低い。これは単に二次元の条件と三次元の条件による視覚の認識時間の差によるものではないことが考えられる。

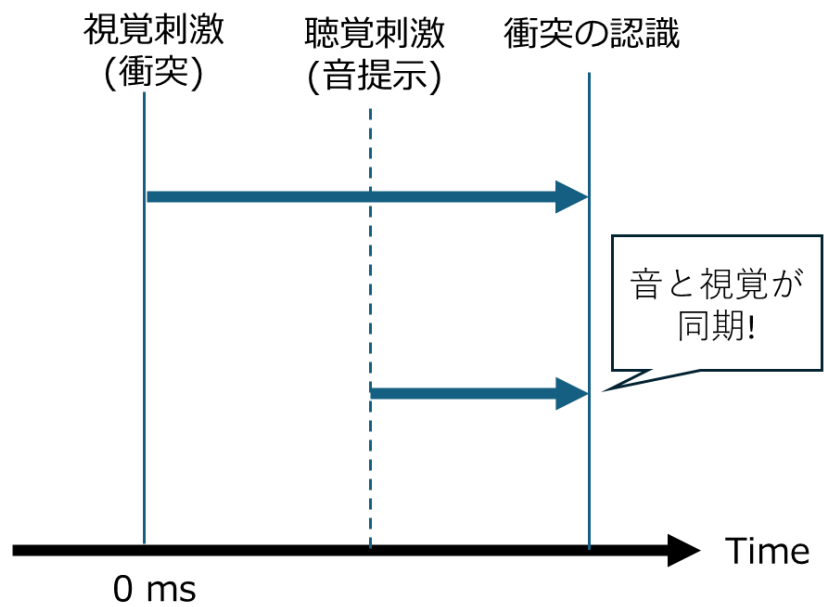


図 21, SOA がプラスであるときの視覚と聴覚の結合

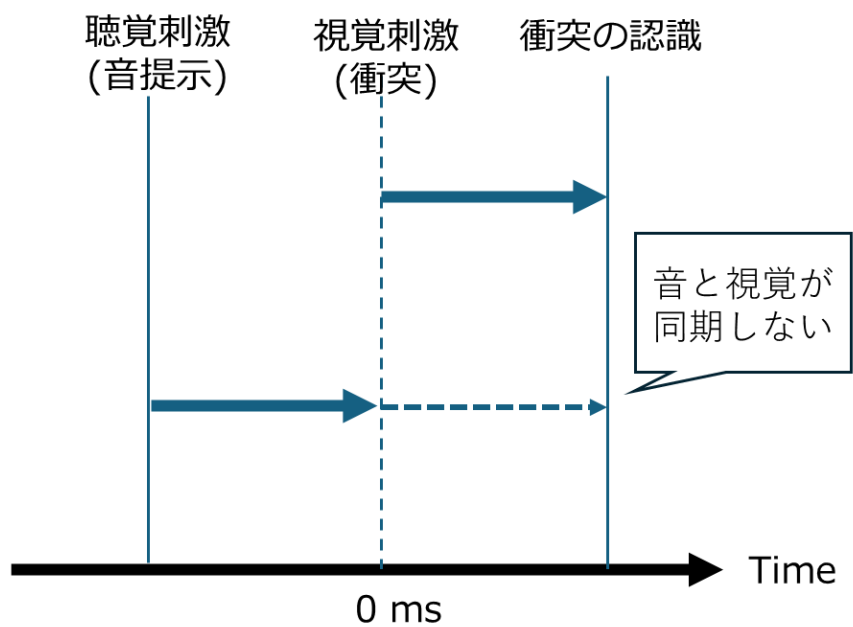


図 22, SOA がマイナスであるときの視覚と聴覚の結合

● 視覚と聴覚の時間的結合

三次元情報に対する視覚の認識時間の変化について、考察したが、ここでは聴覚によるストリームバウンス効果への影響を考察する。

Hayashi らは視覚と視覚のように同一感覚から得られた複数の情報を統合して時間を判断する場合と、聴覚と視覚のように異なる感覚から得られた複数の情報を統合して時間を判断する場合では、同時と感ずる仕組みが異なることを初めて実験により明らかにした[19]。実験では、画面に映る 10 ms ごとに切り替わる映像を見せ、ある一転でクリック音を鳴らし、被験者は音と視覚情報が同期したタイミングの映像を報告した。結果、クリック音は映像に対し、いずれもクリック音よりも後の時刻が同時と判断された。異なる感覚間では、情報を直接統合することができないため、クリック音のような短い聴覚刺激が提示されてから認識したタイミング(音が提示されてから少し後)に画面に表示されていた映像を同時と判断すると考えられている(図 23)。

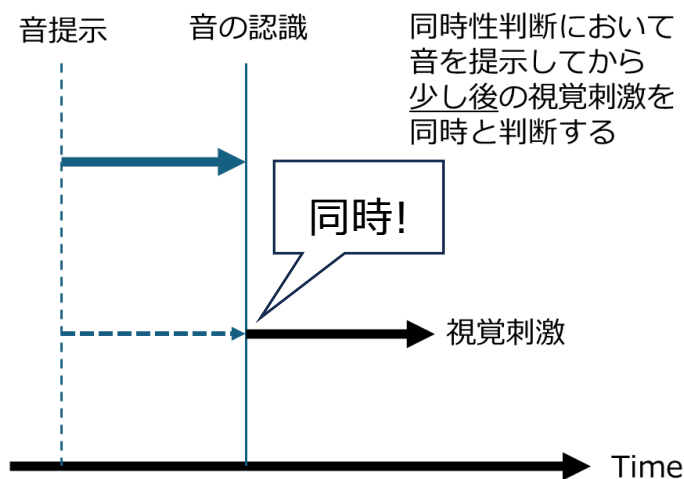


図 23, 視覚と聴覚の同時性判断 実験中映像は流れており、ある一点で音刺激を提示する [19]を参考に作成

Zampini らの結果(図 9)と比較すると、Hayashi らの結果は間違っているように感じる。なぜなら、音が視覚刺激の後に提示され、その認識がさらに遅れるならば、同時だと判断することは難しいということが SOA の条件が 0 ms から離れるほど同時であると判断する確率が下がっていることから明らかである。しかし、Hayashi らの実験は連続的な視覚刺激に対する単発の音が同時性判断に与える影響を測るものであり、Zampini らの実験は単発的な視覚刺激と単発的な聴覚刺激の同期、刺激の前後が影響するのかを調べるものであるため、両者の視覚刺激の違いから、比較の対象にするべきではないと考えられる。

本研究のストリームバウンス効果は連続的な視覚刺激であり、聴覚刺激も単発的な音を提示していることから、Hayashi らの結果を参考にストリームバウンス効果の SOA 条件による Bounce 応答確率の変化がなぜ起きるのか考察する。

図 23 に示した通り、ヒトは連続的な映像と単発的な音に対する同時性判断において、音の提示よりも後の視覚刺激である映像を同時と判断することが明らかである。そこから、本研究のストリームバウンス効果においても音の提示よりも後の視覚刺激である衝突を同時と判断し、視覚と聴覚が同期することで Bounce 応答確率が高まっているのではないかと考えられる。つまり、音の提示が立方体の衝突よりも早い時に Bounce 応答確率が高くなり、音の提示が立方体の衝突と同時の時ヒトは音の認識が遅れ視聴覚の同期が起きず Bounce 応答確率が低くなるということである。本研究の結果(図 20)では、-250 ms から -75 ms の間で Bounce 応答確率高くなっており、そこから先は Stream の応答確率が 25 ms の条件を除いて高くなっている。これは聴覚刺激の認識にかかるまでの時間のラグが 75 ms 以上であることを示している可能性がある。

二つの考察から、ストリームバウンス効果において視聴覚同期のタイミングは二か所存在し、視覚刺激が先行する場合衝突の認識がされるタイミングに音が同期することで Bounce 応答確率が高まり、聴覚刺激が先行する場合聴覚刺激の認識がされるタイミングに視覚刺激が提示されることで Bounce 応答確率が高くなっていると考えられる(図 24)。また、聴覚刺激が先行する場合の視覚刺激認識までの遅れについてだが、Takei らの研究結果(図 6)から音が先行する場合の Bounce 知覚において Bounce 判断時間が短くなることが明らかであることから、音による後のイベント(衝突)の予測のようなことができている可能性があると考えられる。

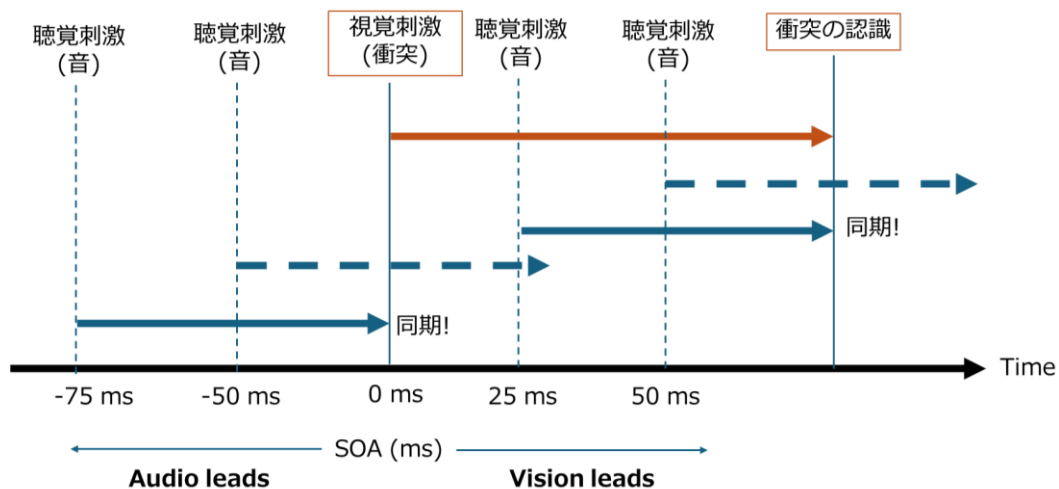


図 24, ストリームバウンス効果における視聴覚の時間的結合 結果から $SOA \leq -75$ ms, $SOA = 25$ ms であるときに同期がしやすいと考えられる。($-250 \leq SOA \leq 250$)

6 刺激音提示位置による錯覚への影響

ストリームバウンス効果において、刺激音の立体的な提示位置による錯覚への影響は研究されていない。そのため、バーチャルリアリティ上に立体音響を再現し音の提示位置によって錯覚にどのように影響するのか調べた。

6.1 実験環境

被験者は成人男性 4 名、女性 1 名で、全員晴眼者（眼鏡、コンタクトレンズ等の矯正も含む）である。実験開始前には、口頭で VR の機器にういて操作方法や実験の構造について被験者に指示し、各被験者が適切に実験を操作できるようにした。必要に応じて実験のデモンストレーションを行った。実験は、横浜市立大学理学研究棟の一室で行った。また、被験者が静かに実験に集中できるように、必要以上に話しかけることはしなかった。本研究で用いた VR 機器は Meta Quest 3 である。実験では、音を提示するため、ヘッドホン (Steel Series Arctis7) を使用した。

6.2 実験方法

実験の方法、条件とそのモデル図を以下に示す(図 25)。また、実験内の環境を理解しやすくするため立体音響のモデル図を掲載する(図 26)。試行回数は一人につき 70 回行った。

実験方法

1. 固定ターゲット表示 (2000 ms)
2. すれ違う立方体を観察した
3. 2 つの立方体がすれ違って見えた(X ボタン)のか、跳ね返って見えた(A ボタン)のかコントローラーのボタンを押して答えた

実験条件

- 本研究の SOA は、0 ms に固定し、音を提示するタイミングを変化させなかった。
- 各試行において、音を提示する位置は、頭の位置を中心として、前、後、左、右、上、下、中心からランダムに決まった。
- 被験者は実験中空間内を自由に見渡すことが可能でしたが、実験中は固定のターゲットに目を向けるように指示した。
- 被験者が操作のミスをした場合、その数だけ多く試行回数を増やした。

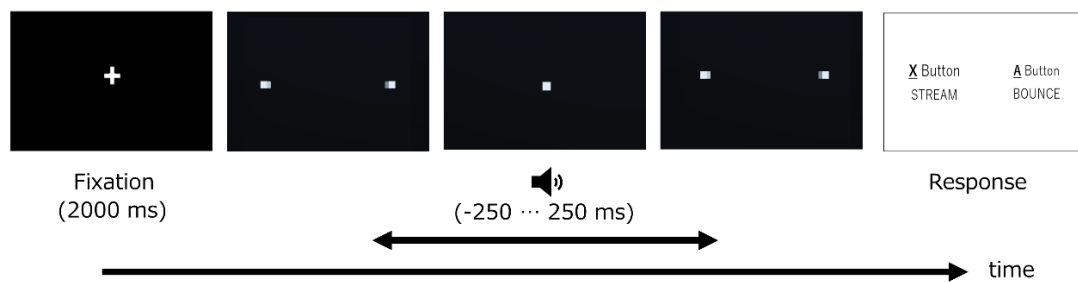


図 25, 実験モデル図

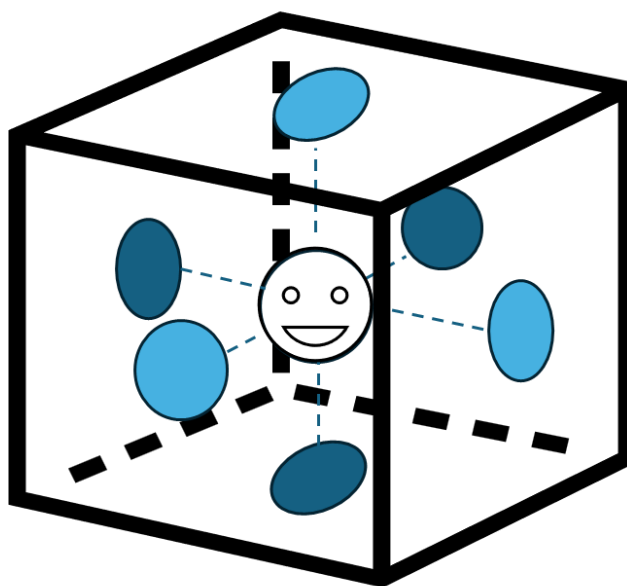


図 26, 立体音響モデル図 頭の位置を中心として、前, 後, 左, 右, 上, 下, 中心の七か所

6.3 実験結果

被験者が各試行で Stream または Bounce と答えたときの音提示位置の条件ごとにデータを取得した。データは Unity のプログラムによって実験終了後に csv ファイルとして保存された。得られた各被験者のデータをまとめ、音を提示する位置ごとの Bounce 応答確率を表す三次元プロット図を作成した(図 27)。

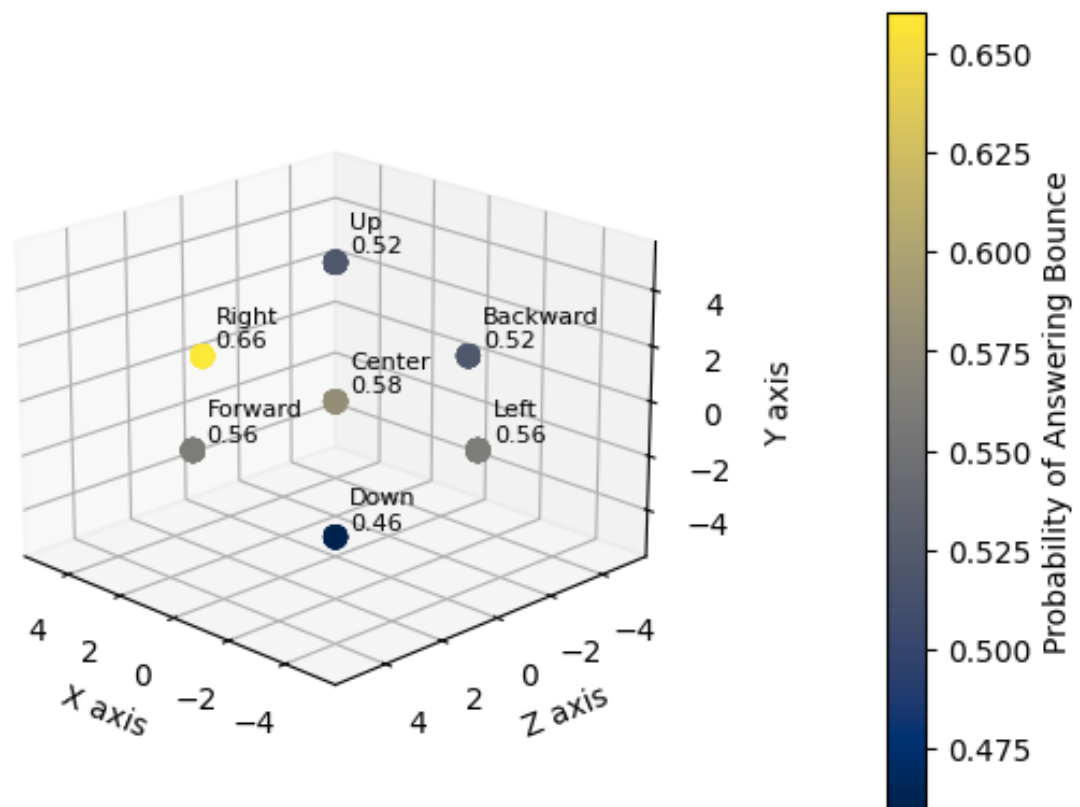


図 27, 音提示位置によるバウンス応答確率の三次元プロット図

また、音提示位置ごとの Stream/Bounce 応答確率を示す箱ひげ図を作成した(図 28)。

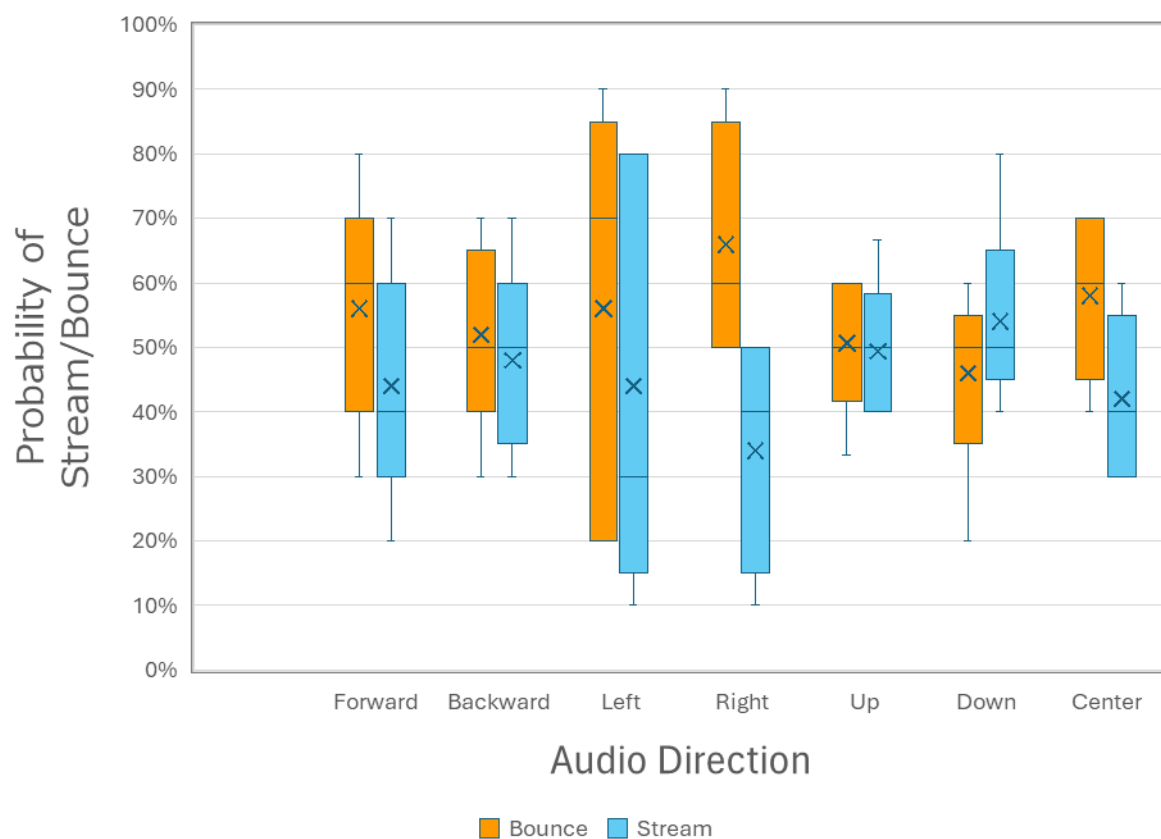


図 28, 音提示位置による Stream/Bounce 確率を示す箱ひげ図

6.4 考察

実験の結果について、音提示位置のそれぞれについて Stream/Bounce 応答に有意差があるかを調べるため、カイ二乗検定を行ったところ、いずれの音の位置による組み合わせでも有意ではなかった(表 1)(図 29)。

表 1, 音の位置による Stream/Bounce 応答の p 値

Pair	p-value
(Right, Left)	0.412163
(Right, Up)	0.222493
(Right, Down)	0.069816
(Right, Forward)	0.412163
(Right, Backward)	0.222493
(Right, Center)	0.536533
(Left, Up)	0.840978
(Left, Down)	0.423618
(Left, Forward)	1.000000
(Left, Backward)	0.840978
(Left, Center)	1.000000
(Up, Down)	0.689098
(Up, Forward)	0.840978
(Up, Backward)	1.000000
(Up, Center)	0.687673
(Down, Forward)	0.423618
(Down, Backward)	0.689098
(Down, Center)	0.316923
(Forward, Backward)	0.840978
(Forward, Center)	1.000000
(Backward, Center)	0.687673

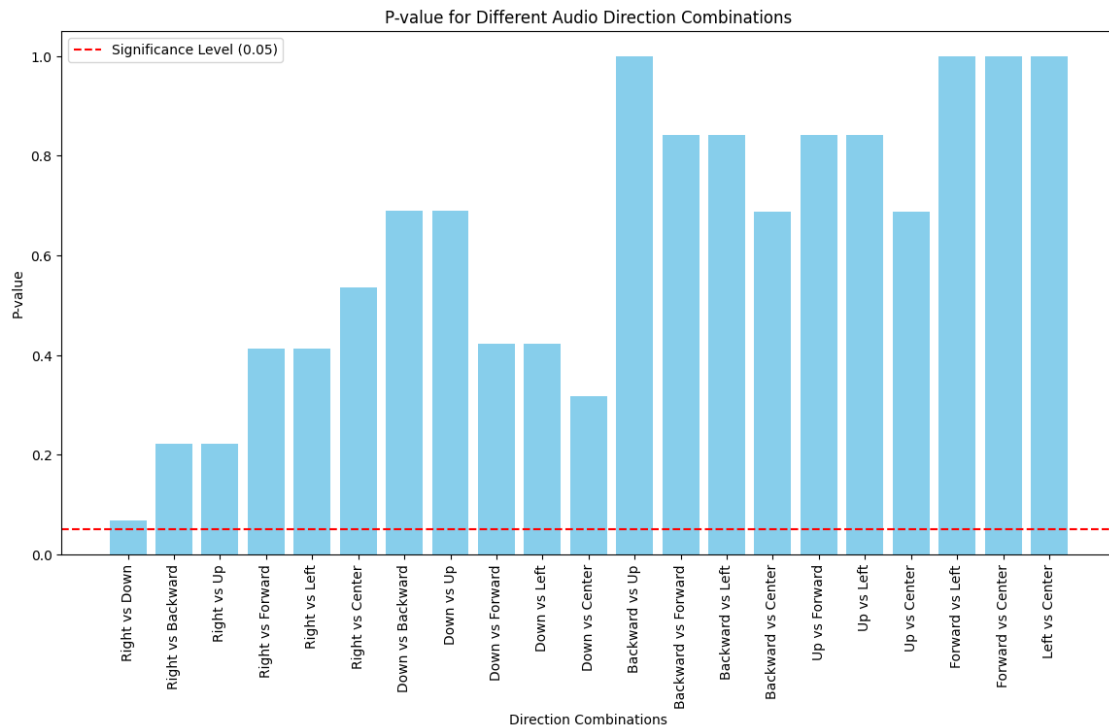


図 29, 音の位置による Stream/Bounce 応答の p 値 有意水準は 0.05

実験の仮説として、音の提示位置を変えることにより、5 章の音源位置であった Center や、実際に立方体の衝突が起きる方向である Forward の条件で Bounce 応答確率が高くなり、他の視覚情報とは位置が同期していない条件では Bounce 応答確率が下がると予想した。これは Zampini らの、視覚刺激と聴覚刺激が同じ位置から提示された場合、異なる位置から提示された場合よりも同時性を感じやすいという結果から考えられた。しかし、本研究の結果からは音の提示位置による有意性はなかった。

そこで、なぜ音の提示位置によるストリームバウンス効果への影響がでないのか、腹話術効果というクロスモーダル現象に関連させ考察する。

● 腹話術効果

腹話術効果とは、腹話術師の口から声が出ているのにも関わらず、手に持つ音声と同期して動かす人形から声が出ているように感じるクロスモーダル現象である[1]。これは腹話術師の声(聴覚情報)が腹話術師の人形の動き(視覚情報)によって、声の位置が人形側にシフトするとされている。

Bertelson らの研究によると、腹話術効果の実験において、視覚刺激の位置が聴覚刺激の位置の知覚を視覚刺激の位置側にずらしてしまうことを示し、また視覚刺激の位置による差はないことを明らかにした[20](図 30)。

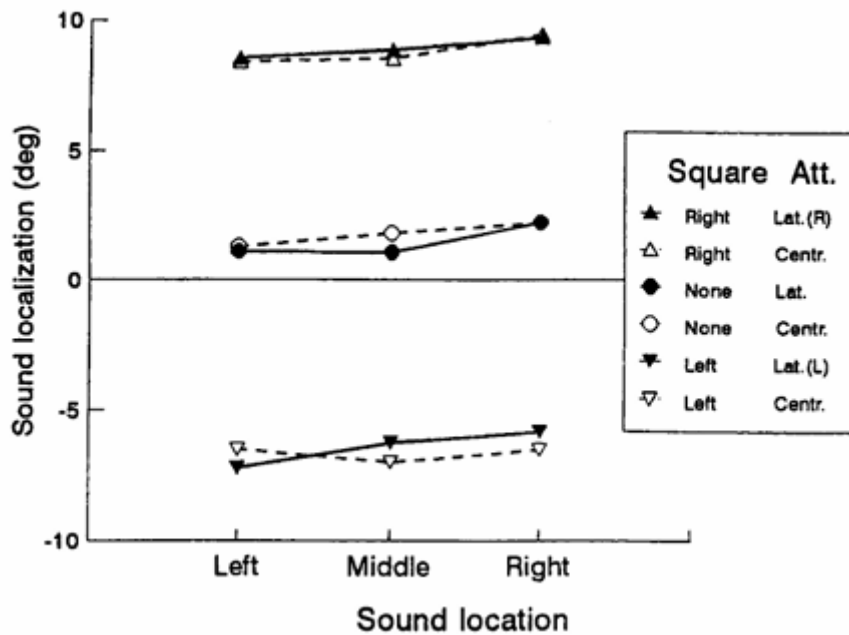


図 30, 視覚刺激と聴覚刺激の位置による音源位置の特定 [20]

図は例えば左から音が提示された場合、視覚刺激が左の場合は左に、右の場合は右に、視覚刺激がない場合は真ん中に音源位置があると判断していることを示している。

このことから、本研究の実験の中でも、立方体の衝突という視覚刺激のイベントに衝突音として提示されている音がどの位置にあった場合においても、中心または前から音がするように感じたのではないかと考えられる。つまり、位置による有意性は見られず全体で同じくらいの Bounce 応答確率になるということである。

また、被験者一人ひとりの音源位置別 Bounce 応答確率のデータを見てみると(図 31)、それぞれ、左耳に偏る人、右耳に偏る人、偏りのない人が見られる。これは、個人ごとの利き耳や特定の音の方向への注意による認知バイアスの影響によるものと考えられる。利き耳の研究について、Ishizu は利き耳にどのような機能があるのか解明することが今後の課題であり、その糸口として、両耳分離聴研究や聴覚の発達研究によって、聴覚の中核および末梢レベルでの非対称性についての検討をすすめることは、利き耳の機能差の解明につながるかもしれないと考えている[21]。

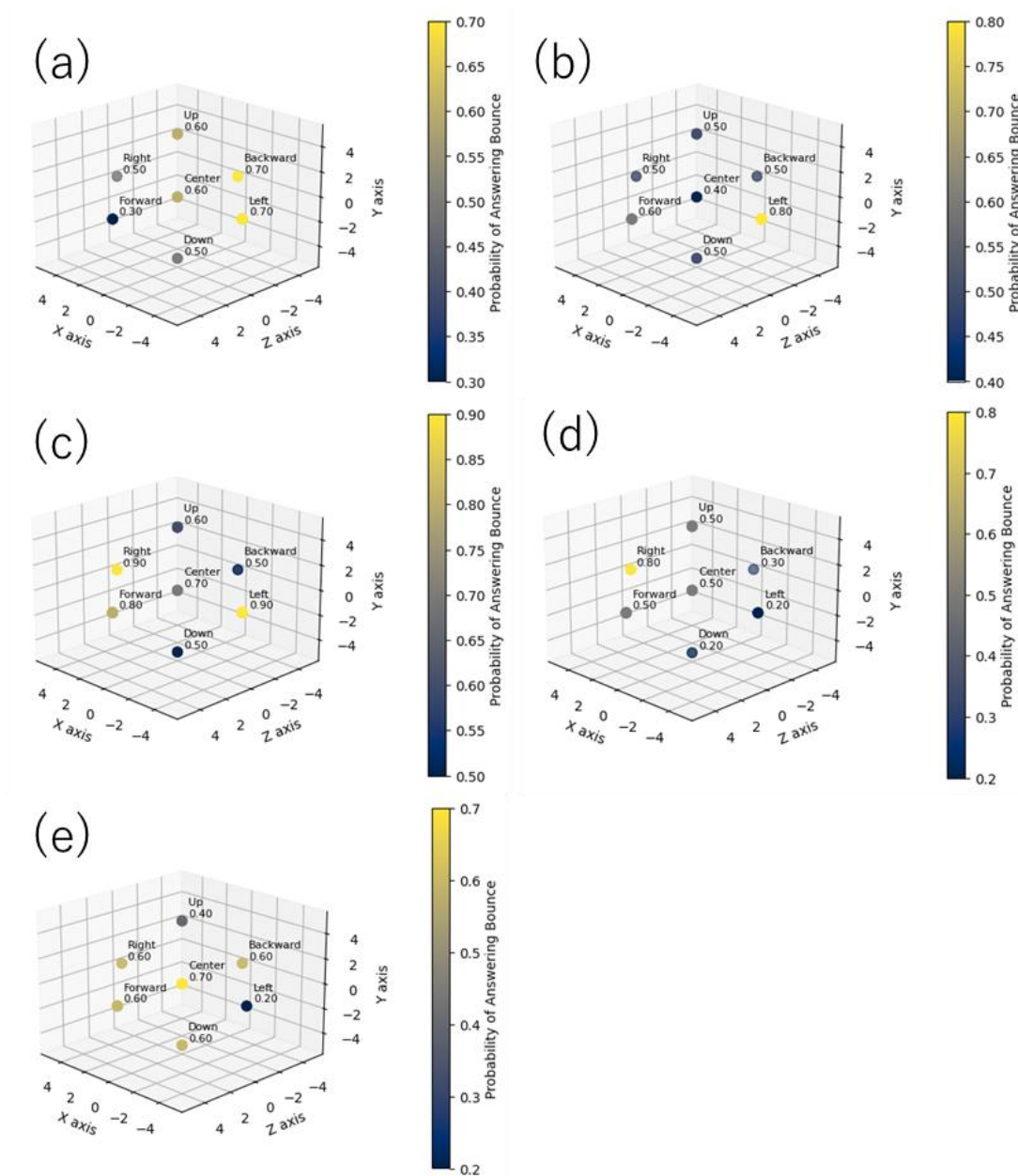


図 31, 被験者ごとの音源位置別 Bounce 応答確率

7 結論と今後の展望

本研究では、まずストリームバウンス効果を二次元ディスプレイ上で立体物を用いた場合の影響を調べた。取得できたデータが少なかったことや物体を立体にすることで単眼性の奥行き知覚による影響を調べたかったが図 13 のように、立方体が中心に近づくとき二次元ディスプレイ上では平面にしか見えないことから、結果が有意であったとはいえない。

その後バーチャルリアリティにて再現し、より現実に近い空間の中で SOA の条件による Stream/Bounce 応答確率がどのように変化するか調べた。刺激音を提示するタイミングが立方体の衝突よりも早い時に Bounce の応答が多く、衝突より後に音が提示された場合 Stream の回答が多くなった。また、Bounce 応答確率が最も高かったのは図 20 から 225 ms の SOA 条件であった。Bounce 応答確率の近似直線の式について、Bounce 応答確率を P_B 、SOA を t_{SOA} とすると、

$$P_B = -0.17t_{SOA} + 50.31 \quad (1)$$

という式が結果から得られた。SOA が 0 ms であるとき Bounce の応答確率が約 50 %であることを表している。

また、聴覚刺激が視覚刺激に先行することで Bounce 応答確率が高くなることや本研究の結果において完全な視聴覚の同期のタイミングであると考えられていた 0 ms の SOA 条件で Bounce 応答確率が先行研究と比較して低くなることについて、二次元情報と三次元情報の認識にかかる判断時間差や連続的な視覚刺激に対して聴覚刺激による同時性判断を行う場合の視覚と聴覚の同期性から理由を考察した。

発展として、刺激音の提示位置による Bounce 応答確率の変化について調査した。提示位置による Stream/Bounce 評価の有意差は見られなかったが、腹話術効果によって、音源位置が視覚刺激の位置へと引っ張られることで全体の Bounce 応答確率が同等になっていると考察した。また、個人のデータを見ると左右に Bounce 応答確率が偏っている被験者や差のない被験者がおり、利き耳の影響による認知バイアスが働いているのではないかと考えた。

以上より、バーチャルリアリティにおけるストリームバウンス効果について、視聴覚同期性という観点から、音のタイミングと提示位置による影響を定量的に調査し、音のタイミングに関しては特定の SOA の条件における Bounce 応答確率のモデル式を求めた。また、それぞれの実験では、今までされてこなかったストリームバウンス効果と視聴覚同期性の関連について人間の知覚特性からなぜ本研究の結果が先行研究と異なるのか考察した。

現実において、物体が衝突する前に音が鳴ることはあり得ない。しかし、ストリームバウンス効果においては音が物体の衝突前になることで跳ね返るように知覚する。この不思議な特性は、ヒトの脳の柔軟性が優れていることを表している。これを生かすためのアイ

デアを考えると車の警告音などは車載カメラなどによる危険判断から危険感知の表示をする少し前に警告音を先行して提示することでドライバーの反応時間の短縮ができる可能性がある。このようにヒトの知覚を刺激提示によって助長することができる可能性が広がっていくことがクロスモーダル現象の研究する意義であると考えている。

今後は、神経科学の分野から Hebbian learning という概念を取り入れたい。その意味は「ニューロン A の発火がニューロン B を発火させると 2 つのニューロンの結合が強まる」[22]ということである。例えば、Takei らのように Stream/Bounce を判断する時間を取ることで、試行を繰り返していくうちに、その判断に使うニューロンの結びつきが強くなり判断時間が短くなっていくのではないかという考えである。ヒトの知覚はニューロンのスパイクによって起きているため、今後ユニモーダルな錯覚やクロスモーダル現象を研究していく際には神経科学や脳科学からのアプローチも取り入れたい。

8 参考文献

- [1] 岡嶋克典, 視覚情報によって誘発されるクロスモーダル効果, 映像メディア学会誌 Vol. 72, No. 1, 2018
- [2] Ruggero Micheletto, 戸坂 亜希, 及川 虎太郎, 英語と日本語で学ぶ知覚情報科学, 共立出版, 2020, 224p
- [3] Ellis, R. R. & Lederman, S. J. (1993). The role of haptic versus visual volume cues in the size-weight illusion. *Perception & Psychophysics*, 53, 315–324.
- [4] Murray, D. J., et al. (1999). Charpentier (1891) on the size—weight illusion. *Perception & Psychophysics*, 61, 1681–1685.
- [5] Wolf C, Bergmann Tiest WM, Drewing K (2018) A mass-density model can account for the size-weight illusion. *PLoS ONE* 13(2)
- [6] G. Morrot., et al. (2001). The Color of Odors. *Brain and Language*, 79, 309-320.
- [7] Sekuler R, Sekuler AB, Lau R.: Sound alters visual motion perception. *Nature*, 385(6614), 308, 1997.
- [8] S, Takei, et al. (2009). Is the Latency for Perceiving Ambiguous Motion Affected by the Timing of Modulatory Stimulus Presentation? *VISION*, 21, 2, 87-99
- [9] Takeshima Y (2021) Visual field differences in temporal synchrony processing for audio-visual stimuli. *PLoS ONE* 16(12)
- [10] Zampini, M., Guest, S., Shore, D.I. et al. Audio-visual simultaneity judgments. *Perception & Psychophysics* 67, 531–544 (2005).
- [11] ユニティ・テクノロジーズ・ジャパン株式会社, 2025, ユニティ・テクノロジーズ・ジャパン株式会社ホームページ, (2025 年 1 月 20 日取得, <https://unity3d.jp/>)
- [12] 日本バーチャルリアリティ学会, バーチャルリアリティ学, 株式会社コロナ社, 2011, 384p
- [13] 廣瀬通孝, 今日からモノ知りシリーズ トコトンやさしい VR の本, 日刊工業新聞社, 2019, 159p
- [14] 雨宮智浩, メタバースの教科書, 株式会社オーム社, 2023, 212p
- [15] 株式会社矢野経済研究所, 2024/12/09, メタバースの国内市場動向調査を実施(2024 年), (2025 年 1 月 20 日取得, https://www.yano.co.jp/press-release/show/press_id/3688)
- [16] Meta, 2025, メタストアホームページ, (2025 年 1 月 20 日取得, <https://www.meta.com/jp/quest/quest-3/>)
- [17] Kawamura, S. Comparison of Impression and Looking Time between 3D and 2D Pictures in the Slide-show. *日本教育工学会論文誌* 36 (Suppl.), 133-136, 2012
- [18] 長峰和仁・中山 実・清水康敬 (1992) メンタルローテーションへの立体テレビ提示の効果. *電子情報通信学会論文誌 A. 基礎・境界*, 75 : 218–225

- [19] Hayashi, R., Murakami, I. Distinct mechanisms of temporal binding in generalized and cross-modal flash-lag effects. *Sci Rep* 9, 3829 (2019).
- [20] Bertelson, P., Vroomen, J., De Gelder, B. et al. The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception & Psychophysics* 62, 321–332 (2000).
- [21] K, Ishizu. What do We Know about Earedness? : A Review of Studies, 日本大学大学院総合社会情報研究科紀要, 8, 325-333, 2007
- [22] N, Takahashi et al. ヘブ則, 脳科学辞典, (2012) (2025 年 1 月 25 日取得, <https://bsd.neuroinf.jp/wiki/%E3%83%98%E3%83%96%E5%89%87>)

9 謝辞

本研究を行うにあたり、ミケレットルジェロ教授には実際にプログラミングや実験をレビューしながら研究の方向性や実験の設定について、熱心にご指導を賜りました。ありがとうございました。

また、ミケレット研究室博士課程の及川虎太郎さんには、日ごろのディスカッションや実験のレビューによって研究のアイデアと一緒に探っていただきました。一月下旬に開催された日本視覚学会の冬季大会への参加を進めてくださり、特に忙しい時期であるのにも関わらず、ポスター制作のコツや発表のポイントなどを教えていただきました。日本視覚学会にてポスター発表をしたことでいろいろな方から本研究に対する新しい着眼点や考察を得ることができました。ありがとうございました。

同期の小泉さんや修士課程の河野さんには、研究に対するアイデアや資料制作に対するアドバイスを頂きました。ありがとうございました。

ミケレット研究室の卒業生である吉田瞬良さんには、研究室に所属してから python の使用方法や知覚情報科学への応用について熱心にレクチャーしていただきました。また、VR を用いて研究するにあたり、機器のセットアップの手伝いや研究テーマの調べ方について教えていただきました。吉田さんの研究への取り組む姿勢を見られたからこそ、本研究へのモチベーションが維持できたのだと思います。ありがとうございました。

最後に、日ごろの生活を支えてくれた家族、そして大学生活を共にした友人に感謝を申し上げます。ありがとうございました。