

基礎論文

視聴覚コンテンツの臨場感と迫真性の規定因

本多 明生^{*1} 神田 敬幸^{*2} 柴田 寛^{*3} 浅井 暢子^{*4} 寺本 渉^{*5}坂本 修一^{*6} 岩谷 幸雄^{*7} 行場 次朗^{*8} 鈴木陽一^{*6}

Determinants of Senses of Presence and Verisimilitude in Audio-Visual Contents

Akio Honda^{*1} Takayuki Kanda^{*2} Hiroshi Shibata^{*3} Nobuko Asai^{*4} Wataru Teramoto^{*5}Shuichi Sakamoto^{*6} Yukio Iwaya^{*3} Jiro Gyoba^{*7} and Yôiti Suzuki^{*6}

Abstract — Senses of presence and verisimilitude are affected differently by temporal asynchrony between audio-visual components of audio-visual content. To investigate whether this result is valid more generally, we conducted an experiment using a clip of western orchestral music. Results revealed that the sense of verisimilitude is more sensitive to audiovisual synchronicity than to the display size, whereas the sense of presence is more sensitive to the spatial size than the temporal property. These findings corresponded well with those of the previous study, which indicated that the sense of verisimilitude is distinguishable from the sense of presence. Furthermore, we discussed important considerations related to measurement for Kansei information such as the sense of presence and verisimilitude.

Keywords: sense of presence, sense of verisimilitude, audio-visual contents, orchestral music

1 はじめに

映画や音楽、スポーツなどのエンターテインメントを楽しみたい、自然でリアルなコミュニケーションを体験したいという思いは、世界中のマルチメディアユーザが抱いている願いである。昨今のバーチャルリアリティ(以下VR)技術の発展は著しいが、私たちがVRシステムからどのように本物らしさを知覚しているのか、その規定因は何かを明らかにすることは、VRシステムの技術革新や総合的な性能評価を行ううえで有益な知見を供給する。そこで、我々は、視聴覚コンテンツの臨場感(sense of presence)と迫真性(sense of verisimilitude)に着目し、

両感性の特徴の分析や規定因の解明に取り組んできた(寺本ら[1, 2], 鈴木ら[3])。

1.1 臨場感と迫真性

VRシステムの登場以来、高い臨場感を生み出す要因は何かという問いに対しては数多くの研究が行われてきた。例えば、臨場感を生起させる要因は、内部要因(ユーザ側要因)と外部要因(メディアやコンテンツ側要因)に大きく分けられる[1, 4]。内部要因には、ユーザの知覚・認知能力、運動能力、パーソナリティ、年齢、性別、観察時のムードなどが含まれ[4-6]、外部要因には、感覚情報の提示範囲[7, 8]、ユーザによる環境の変更可能性[9]、コンテンツの内容[10]などが含まれる。特に、高い臨場感を効率的に生み出すVRシステムの設計においては、感覚情報の提示手法(範囲や強度)の解明が重要な課題のひとつに位置づけられてきた。例えば、江本ら[7]は、視聴者に水平方向の提示視野角を変化させた静止画像を提示して、同画像の臨場感、力量感、快適感等を主観評価するように求めた結果、臨場感や力量感の評定値が視野角の拡大とともにほぼ単調増加したことを報告している。また、小澤[8]の実験では、映像の有無に関わらず、本来の音圧レベルよりも高い再生音圧レベルを提示したときに、視聴者の臨場感評定値が最大化したことが示されている。しかしながら、臨場感とは「あたかもその場にいるような感じ」という辞書的定義が示すように、人間(=評価者)を取り巻く空間、す

*1 東北福祉大学総合福祉学部

*2 リオン株式会社

*3 東北文化学園大学医療福祉学部

*4 京都文教大学総合社会学部

*5 室蘭工業大学情報電子工学系学科

*6 東北大学電気通信研究所

*7 東北学院大学工学部

*8 東北大学大学院文学研究科

*1 Faculty of General Welfare, Tohoku Fukushi University

*2 Rion Co., Ltd.

*3 Faculty of Medical Science and Welfare, Tohoku Bunka Gakuin University

*4 Faculty of Social Relations, Kyoto Bunkyo University

*5 Department of Information and Electrical Engineering, Muroran Institute of Technology

*6 Research Institute of Electrical Communication, Tohoku University

*7 Faculty of Engineering, Tohoku Gakuin University

*8 Graduate School of Arts and Letters, Tohoku University

なわち、背景的な「場」の本物らしさに関係する感性印象であり、ユーザが VR システムで体験するものの一部しか抽出していないと考えられる。例えば、VR システムにおいて、オペラの情景を実現する場面を考えた場合、劇場全体を再現すると共に、場面の主役である歌手の挙動もまた感性的にリアルに再現される必要があり、後者によって場面の印象が強く左右されることが予測される[2]。したがって、臨場感とは異なる新たな指標を用いることで、ユーザが VR システムで体験する本物らしさを多角的に検討していくことが求められる。

本物らしさに関係する概念のひとつに迫真性があげられる。迫真性は verisimilitude 以外に fidelity として述べられることもあるが[11]、主としてコンテンツの本物らしさを表す概念として理解されており[11, 12]、いくつかの展望論文[11-13]や探索的研究[14]においては VR との関連性について言及されている。しかしながら、従来の研究においては、迫真性は臨場感とは異なる本物らしさに関係する感性なのかを実証的な知見をもとに論考した研究は行われてこなかった。寺本ら[2]は、背景的な「場」の本物らしさに関係する感性印象である臨場感に対して、対象や事象など前景情報を中心とした本物らしさに関係する感性を「迫真性」と定義し、それぞれが独立した感性印象として測定可能であることを明らかにしている。具体的には、視聴覚刺激として鹿威しを用い、その時空間パラメータを操作して視聴者が感じる臨場感と迫真性の特性の違いを調べた結果、(1)臨場感は現実を超えるような刺激強度でも高まること、(2)迫真性は刺激の強度を高めても必ずしも増加しないこと、(3)臨場感と迫真性の間では視聴覚情報の時間ずれの許容範囲が異なることを示した。知覚心理学の分野では、背景(地)と前景(図)の処理には異なる処理システムが関与している証拠が得られており[15, 16]、背景要素は、無意識的・自動的に処理される一方で、そこから分離された前景要素は、意識的で詳細な分析を受けると考えられている[16]。これらの知見は、背景的要素と強く関連する臨場感のほかに、前景的要素と関連した場面の感性印象である迫真性を取り上げることで VR システムの性能をより総合的に評価することが可能であることを示唆している。迫真性は、刺激強度を高めても必ずしも増加しないことから[2]、還元された情報によって創出される本物らしさの感性[17]として捉えることも可能である。近年、刺激強度が小さいほど種々の感覚間の相互作用が大きくなる逆理的効果 (inverse effectiveness) とよばれる現象が報告されている[17, 18]。現実性を適切に捨象することによって刺激量を減らすことができれば、多感覚間の相互作用が強まることで創発効果が生まれ、またユーザによる補完処理が生起することで、より迫真性のあるコンテンツとして感受される可能性がある[17]。このように、迫真性は、今後求められるであろう VR システ

ムのより総合的な評価において重要な役割を果たすことが期待されている。しかしながら、迫真性に関する実証的な知見は依然として乏しいことから、この感性の特徴を理解するための更なる研究が必須とされている。

1.2 本研究の目的

寺本ら[2]の研究は、迫真性が臨場感とは異なる感性である点を明らかにし、同感性が VR システムの総合的な性能評価で果たす役割を示したことで VR 研究に画期的な知見を供給したといえる。しかしながら、この実験結果の一般性には疑問も残されている。寺本ら[2]の実験では、臨場感は迫真性よりも視聴覚コンテンツの強度(視野や音圧レベル)の影響を受けやすいことが示されたが、鹿威し以外の視聴覚コンテンツで同様の現象が観察されるのかは不明確である。事実、視聴者の臨場感に感覚情報の強度が及ぼす効果は視聴覚コンテンツの内容によって異なるという結果も報告されている[19, 20]。そのため、臨場感と迫真性の特性の違いが視聴覚コンテンツの内容やジャンルに依存せずに観察されるのかを明らかにする必要がある。そして、臨場感と迫真性の特徴を明らかにするうえでは、視聴者が許容できる視聴覚コンテンツの時間ずれの範囲についても検討することが極めて重要である。従来の研究は、視聴覚統合の時間限界が -200ms から $+250\text{ms}$ の範囲であることを指摘しているが[21, 22]、発話音声を用いた場合は -320ms から $+480\text{ms}$ という報告[23]があるほか、打楽器の一種であるマリンバを使用した場合では範囲が 0ms から $+700\text{ms}$ へと後方に拡大することが報告されている[24]。これらの研究結果は、視聴覚統合の時間窓の範囲が、視聴覚コンテンツの内容によって変動することを示唆している。上述した視聴覚統合の時間窓に関する問題は、高い臨場感と迫真性を創出するうえで考慮しなければならない問題のひとつに位置づけられる。しかしながら、現状では、身近な視聴覚コンテンツを使用した研究は乏しいことから、多くのユーザにとってより身近な視聴覚コンテンツを刺激に用いることで、視聴覚統合の時間窓に関する新たな知見を得ることが必須である。

以上の議論に基づき、本研究では、新たな視聴覚コンテンツを用いることで、臨場感と迫真性という感性の特性の違いについての更なる知見を得ることを試みた。寺本ら[2]の実験刺激は、情景に関する視聴覚コンテンツであり、日本文化的で、人工物を映像の対象としていた。そこで、本研究では、西洋文化を象徴し、人間を映像の対象とする、オーケストラの演奏映像を実験刺激として使用した。同刺激は、寺本ら[2]の刺激と対照的であり、現象の一般性を調べるうえでふさわしく、多くのユーザにとってより身近な実験刺激であると考えた。更に、本研究では、視聴者が感じる臨場感と迫真性の変化をより詳細に測定することで、臨場感と迫真性の規定因を

明らかにし、さらに視聴覚統合の時間窓の範囲に関する新たな知見を得ることを目的とした。

2 実験 1

視聴覚コンテンツの提示視野範囲と音圧レベルが視聴者の臨場感ならびに迫真性に及ぼす効果を調べた。寺本ら[2]の実験では、提示視野範囲が大きい条件(28.9deg)、鹿脅し全体が見える最少の視野範囲である小条件(16.6deg)、その中間にあたる中条件(22.8deg)を設けて提示視野範囲の効果を調べている。そして音圧レベルに関しては、集音時の音圧レベル条件、集音時の音圧レベルを基準に 10dB 音圧レベルを上げた条件、そして映像に音を加えない条件という三条件でその効果を検証している。本研究では、視聴覚コンテンツの音圧レベルの操作に伴う、臨場感と迫真性の連続的な変化を詳細に記録するために、寺本ら[2]よりも操作する音圧レベルの刻み幅を小さく設定して実験を行った。

2.1. 方法

2.1.1 実験参加者

視聴覚に問題のない大学生 20 名(男性 9 名、女性 11 名)で、心理学を専攻する学生が多かった。

2.1.2 実験刺激

R. ワーグナー作曲、楽劇ニュルンベルクのマイスタージンガーより第 1 幕への前奏曲の一部を撮影した 9.3 秒間のオーケストラ映像を実験刺激とした。映像の解像度は 640×480 であり、その内容は(1)トロンボーン奏者の映像(徐々にトロンボーン奏者に映像がズーム: 映像開始後 0~4 秒)、(2)シンバル奏者の演奏の様子(映像開始後 4~7.8 秒)、(3)バイオリン奏者の演奏場面(映像開始後 7.8~9.3 秒)から構成されていた(図 1)。

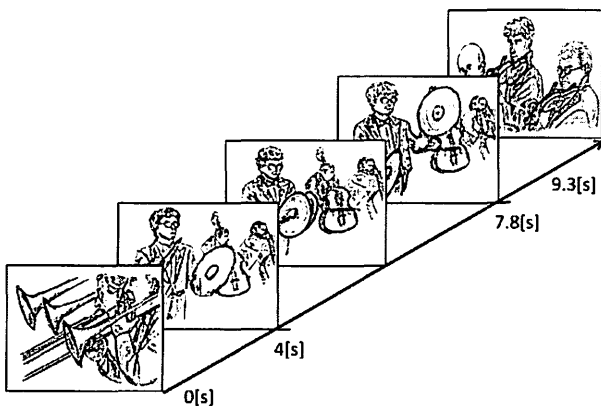


図 1 実験刺激の概要図

Fig. 1 Schematic diagram of stimulus sequence

寺本ら[2]は、鹿威しの竹筒の末端部が敷石をたたき、カコーンという特徴的な打叩音を発するシーンを実験刺

激として使用している。本研究では、(1)音楽コンテンツ内のシンバル奏者演奏が上記の鹿威しと同様に特徴的な視聴覚情報を含んでいること、(2)シンバルは単発音を発する楽器であるにも関わらず、音楽の印象にアクセントを加える役割を担っており、感性情報が豊富に含まれていることが期待できたことから、上述のシンバル演奏を中心的な内容とするオーケストラ映像を実験刺激として選定した。本実験では、視聴覚コンテンツ内のシンバル演奏の開始から映像の終了までの音刺激(映像開始 6.3~9.3 秒)に Matlab(The Math Works, Inc.)を用いて操作を加えた。さらに、音圧レベルの操作に伴う、演奏のつながりの不自然さを軽減させるために、音圧レベルの減衰の直前と直後に 10ms のコサイン窓をかけた。

2.1.3 実験条件

視野サイズ(3 水準: 大, 中, 小)×音圧レベル(11 水準: -15dB~+15dB まで 3dB 間隔)の二要因被験者内デザインを用いた。視野サイズは、映像の左右端に黒帯の枠を付け加えることで操作した。視野大条件(28.9deg)は画面に収まる最大のサイズであり、シンバル奏者とその周囲が見える条件であった。また、視野小条件(16.6deg)は、シンバル奏者が見えるのみで背景がほとんど含まれない視野範囲であり、視野中条件の視野範囲は上記 2 条件の中間サイズであった(22.8deg)。音圧レベルは、オーケストラ映像に収録されていた音(実験時には等価音圧レベルで 81.5dB で提示)を基準に 15dB 音圧レベルを下げたものから 3dB 間隔で 15dB まで音圧レベルを上げたものの計 11 種類であった。

2.1.4 実験手続き

実験参加者は、暗室内でヘッドホン(STAX SR-202)と CRT モニタ(SONY, CPD-G520, 60Hz)によって刺激を視聴した。実験参加者は、あご台に頭部を固定された状態で実験刺激を視聴し、各刺激の提示直後に、その臨場感と迫真性を、0(全くない)から 6(非常にある)までの視覚的アナログ尺度で回答した。その際、臨場感と迫真性の多義性を避けるために、臨場感については「その場にいる感じがする」、迫真性については「本物らしい感じがする」という注釈を提示した。この操作は寺本ら[2]に準じたものである。ただし、同研究では鹿威しのもつ風情や趣を醸し出す鑑賞用装置としての特性から、迫真性には「本物らしい趣がある」という注釈を使用していたのに対し、本研究では、特定刺激に依存しにくい注釈を作成することが望ましいと考えたことから、その注釈を「本物らしい感じがする」に変えた。実験は、全 2 セッションから構成されていた(全 66 試行)。1 セッションは視野サイズ(3 水準)×音圧レベル(11 水準)の 33 試行で、刺激の提示順序はランダム化した。第 1 セッションで各実験刺激に対して臨場感(もしくは迫真性)から先に評価した場合、第 2 セッションでは両感性の評定順序を入れ替えることで感性評価における順序効果の相殺を

図った。実験時間は 25 分程度であった。

2.2 結果と考察

寺本ら[2]と同様に臨場感と迫真性の直接的な比較を行うために、実験データをz得点に標準化したうえで(臨場感の素点平均評定値 2.97, 標準偏差 1.37; 迫真性の素点平均評定値 2.94, 標準偏差 1.33), 感性(2) × 視野サイズ(3) × 音圧レベル(11)の三要因分散分析を行った。その結果, 視野サイズの主効果($F_{2, 38} = 15.54, p < .001$), 音圧レベルの主効果($F_{10, 190} = 60.45, p < .001$), 感性 × 視野サイズの交互作用($F_{2, 38} = 8.62, p < .001$; 図 2), 感性 × 音圧レベルの交互作用($F_{10, 190} = 5.11, p < .001$; 図 3)が有意であった。図の縦軸は、値が大きいほど臨場感(Presence)および迫真性(Verisimilitude)が高いことを示す。

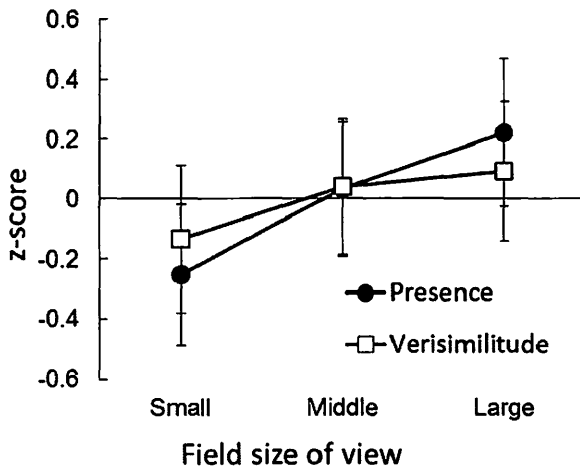


図 2 感性と視野サイズの関係

Fig. 2 Relationship between Kansei and field size of view

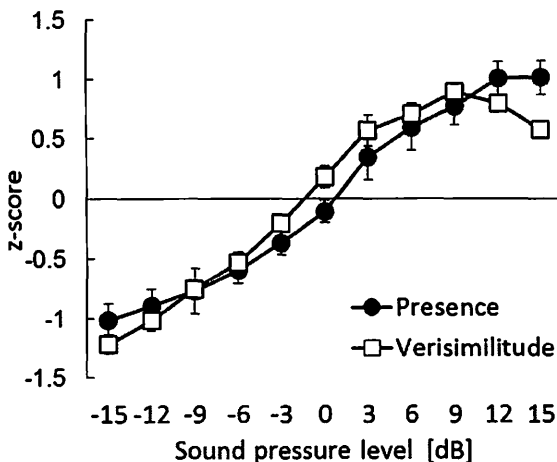


図 3 感性と音圧レベルの関係

Fig. 3 Relationship between Kansei and sound pressure level

感性 × 視野サイズの交互作用について単純主効果の検定ならびに多重比較(Ryan 法)を行った結果, 臨場感は視野サイズが大きくなるにつれて高まる一方, 迫真性は視野大および中条件のときに最も高まることが示された。また, 感性 × 音圧レベルの交互作用に対する下位検定(Ryan 法)からは, 臨場感は 12dB 付近まで単調増加する一方で, 迫真性は 9dB 付近から低下することが示された。実験結果は, 寺本ら[2]の知見と一貫するものであった。

3 実験 2

実験 1 では, 視聴覚コンテンツの提示視野範囲と音圧レベルが視聴者の臨場感と迫真性に及ぼす効果を調べた。その結果, 臨場感は視野サイズや音圧レベルが高まると概ね単調増加するが, 迫真性は強度を上げすぎるとその効果が減衰することが示された。この結果は, 迫真性が感覚様相間の調和を基盤とする感性であることを示唆している。

そこで, 実験 2 では, 視聴覚コンテンツの音と映像の時間的な調和性, すなわち同期性が視聴者の臨場感と迫真性に及ぼす効果を検討した。寺本ら[2]の実験 2 では, 実験刺激の SOA (Stimulus Onset Asynchrony) を操作した結果(-300/-150/0/+150/+300/+450ms の 6 水準), 最適な視聴覚情報の時間窓が臨場感と迫真性では異なることが報告されている。具体的には, 臨場感は一150ms から+150ms の範囲で時間ずれが許容され, 迫真性は收音時と同じ音圧レベルの条件では 0ms から+150ms, 音圧レベルを高めた条件では-150ms から+300ms で時間ずれが許容されたことを報告している。この知見が, 彼らの用いた実験刺激に特有のものでないのであれば, 本実験でも, 臨場感と迫真性の視聴覚情報の時間窓に特徴的な差異が認められると予測した。さらに, 従来の研究においては, 視聴覚統合の時間窓の範囲を一200ms から+250ms とする考えが支配的であった[21, 22]。しかしながら, 発話音声では-320ms から+480ms という報告[23]があるほか, 打楽器のひとつであるマリンバを視聴覚刺激に用いた研究からは, 視聴覚統合の時間窓の範囲が 0ms から+700ms へと後方に拡大したことが報告されている[24]。そのため, 実験 2 では, 実験 1 と同様にシンバル演奏を中心的な内容とするオーケストラ映像を用い, 寺本ら[2]よりも SOA の測定範囲をより長く設定することで, 視聴覚統合の時間窓の範囲に関する新たな知見を得ることを試みた。

3.1 方法

3.1.1 実験参加者

視聴覚に問題のない大学生 15 名(男性 7 名, 女性 8 名)で, 心理学を専攻する学生が多かった。

3.1.2 実験条件

実験 1 の視聴覚刺激をもとに実験刺激を作成した。実験刺激の視野サイズは実験 1 の視野大条件 (28.9deg) と小条件 (16.6deg) を用いた (2 水準)。音圧レベルは、オーケストラ映像に収録されている音 (実験 1 と同じく実験時には等価音圧レベルで 81.5dB で提示) を基準に音圧レベルを 2dB 上げたものと 10dB 上げたものを使用した (2 水準)。なお、実験 1 で臨場感が最大化した音圧レベルよりも低めの 10dB に設定したのは、実験刺激を高い音圧レベルで繰り返し視聴することで、参加者の聴覚への負荷が生じる可能性に配慮したためである。SOA は、シンバル奏者がシンバルをたたく際の視覚的打叩時刻と聴覚的打叩時刻のずれと定義して -300ms から +1050ms までを 150ms 間隔で設定した (10 水準)。SOA がマイナスの場合は、シンバル音 (聴覚的打叩) が映像 (視覚的打叩) よりも早く、プラスの場合は映像よりも音が遅れて再生されることを意味している。上述した各条件を組み合わせ、40 種類の視聴覚刺激を作成して実験に用いた。

3.1.3 実験手続き

実験参加者は、実験 1 と同一の環境と手法で、各映像の臨場感と迫真性を評価した。実験は、全 2 セッションから構成されていた (全 80 試行)。1 セッションは視野サイズ (2 水準) × 音圧レベル (2 水準) × SOA (10 水準) の 40 試行で、実験刺激の提示順序はランダムとした。第 1 セッションで各実験刺激に対して臨場感 (もしくは迫真性) から先に評価した場合、第 2 セッションではその評価順序を第 1 セッションとは入れ替えた。実験時間は 30 分程度であった。

3.2 結果と考察

実験 1 と同様に実験データを z 得点に変換し (臨場感の素点平均評定値 2.59, 標準偏差 1.28; 迫真性の素点平均評定値 2.34, 標準偏差 1.41), 感性 (2) × 視野サイズ (2) × 音圧レベル (2) × SOA (10) の四要因分散分析を行った。その結果、視野サイズの主効果 ($F_{1, 14} = 23.25, p < .001$), 音圧レベルの主効果 ($F_{1, 14} = 14.16, p < .01$), SOA の主効果 ($F_{9, 126} = 33.56, p < .001$), 感性 × 視野サイズの交互作用 ($F_{1, 14} = 6.52, p < .05$), 感性 × 音圧レベルの交互作用 ($F_{1, 14} = 4.73, p < .05$), 感性 × SOA の交互作用 ($F_{9, 126} = 10.84, p < .001$), 視野サイズ × 音圧レベルの交互作用 ($F_{1, 14} = 11.43, p < .01$), 音圧レベル × SOA の交互作用 ($F_{9, 126} = 2.86, p < .01$) が有意であった。

本研究は、臨場感と迫真性の差異とその規定因に関する知見を得ることを主たる目的としていたため、以降の分析では感性要因とその他の要因の交互作用に焦点をあてた。図の縦軸は、値が大きいほど臨場感および迫真性が高いことを示す。感性 × 視野サイズの交互

作用に対して単純主効果の検定を行った結果、臨場感では視野サイズが大きくなると増加したが ($p < .001$), 迫真性では視野サイズの効果は示されなかった (図 4)。この結果は、実験 1 の結果と一貫する。

同様に感性 × 音圧レベルの交互作用に対して単純主効果の検定を行った結果、臨場感では 10dB 条件のほうが 2dB 条件よりも高いことが示されたが ($p < .001$), 迫真性ではそのような音圧レベルの効果は示されなかった (図 5)。実験 1 では、臨場感では 12dB 付近まで単調増加の傾向を示すが、迫真性は 9dB 付近からその効果が減衰する結果が得られている。したがって、臨場感と迫真性で音圧レベルの効果が異なることを示す知見が実験 1 と実験 2 の双方から得られたといえる。

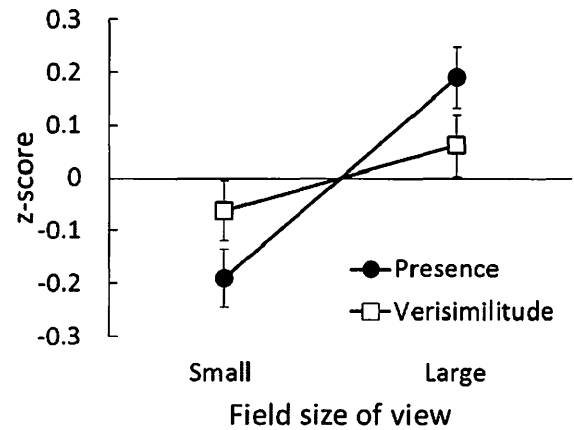


図 4 感性 × 視野サイズの関係

Fig. 4 Relationship between Kansei and field size of view

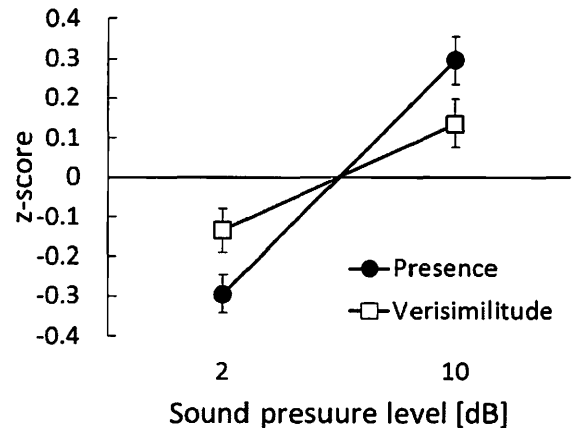


図 5 感性 × 音圧レベルの関係

Fig. 5 Relationship between Kansei and sound pressure level

感性 × SOA の交互作用効果を図 6 に示す。各 SOA における感性の単純主効果を分析した結果、SOA + 150ms ($p < .01$), +300ms ($p < .01$) では、臨場感よりも

迫真性の評定値のほうが高く、 $SOA + 1050ms$ ($p < .05$) では迫真性よりも臨場感の評定値が高かった。次に $SOA 0ms$ 条件を基準にどの程度の時間ずれを許容できるか ($SOA 0ms$ との有意な差がどの程度の時間ずれで認められるか) という視点で下位検定 (Ryan 法) を行ったところ、臨場感は $0ms$ から $+600ms$ の範囲で、迫真性は $0ms$ から $+450ms$ の範囲で時間ずれが許容されていることが明らかとされた。これらの結果は、迫真性のほうが臨場感よりも感覚様相間の調和に対する感度が高いことを意味しており、視聴覚コンテンツの音と映像の時間ずれの影響をより強く受けていることを示唆している。

寺本ら[2]は、臨場感については $-150ms$ から $+150ms$ の範囲で時間ずれが許容され、迫真性は收音時と同じ音圧レベルの条件では $0ms$ から $+150ms$ 、音圧レベルを高めた条件では $-150ms$ から $+300ms$ で時間ずれが許容されたことを報告している。この知見と本研究の結果を考え合わせると、(1) 臨場感と迫真性では視聴覚情報を統合するための時間窓が異なること、(2) 視聴者が許容可能な臨場感と迫真性の時間ずれの範囲は $0ms$ から $+150ms$ を中心としたものであること、(3) 許容可能な時間ずれの範囲の上限と下限は視聴覚コンテンツの内容やジャンルによって変動することが示唆される。

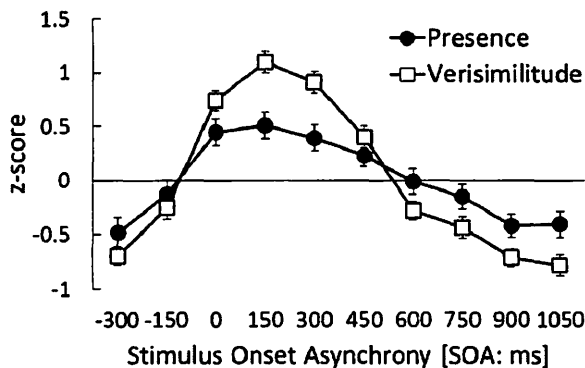


図 6 感性×SOA の関係

Fig. 6 Relationship between Kansei and SOA

4 総合的考察

本研究は、オーケストラの演奏映像を実験刺激として用い、その視野サイズ、音圧レベル、SOA を操作し、視聴者の臨場感と迫真性の変化を視覚的アナログ尺度により測定した。その結果、以下のことが明らかとされた。

1. シンバル演奏を中心的な内容とする視聴覚コンテンツを提示した場合、視野サイズや音圧レベルを大きく設定すると視聴者が感じる臨場感は概ね増加するが、迫真性は刺激強度を上げすぎると、その効果の度合いは減衰する。
2. 迫真性は臨場感よりも感覚様相間の調和に対する

感度が高いことから、視聴覚コンテンツの音と映像の時間ずれの影響を臨場感よりも強く受ける。

3. 視聴者が許容可能な臨場感と迫真性の時間ずれの範囲は $0ms$ から $+150ms$ を中心としたものであるが、その範囲の上限と下限は視聴覚コンテンツの内容やジャンルによって変動する。

本研究と寺本ら[2]の結果から、視聴覚コンテンツの視野サイズや音圧レベル、映像と音の同期性は、いずれも臨場感と迫真性の規定因ではあるが、臨場感は視聴覚コンテンツの強度 (視野サイズや音圧レベルの大きさ) に、迫真性は視聴覚コンテンツの時間的な特徴 (映像と音の時間的なずれ) に、他方よりも強く依存することが明らかになった。視聴覚コンテンツの強度が臨場感に及ぼす効果は、視聴覚コンテンツの内容によっても異なることが報告されているが[19, 20]、本実験結果は寺本ら[2]の鹿威し映像を使用した実験結果と類似したものであった。したがって、臨場感と迫真性の違いは、視聴覚コンテンツの内容やジャンルに依存する可能性は低いと考えられる。しかしながら、寺本ら[2]や本研究でも用いた刺激は、風景コンテンツと音楽コンテンツというジャンルの違いはあるものの、視聴覚コンテンツの内容は、ある事物が別の事物に接触するという、打叩事象を中心に構成された実験刺激であることに留意する必要がある。具体的には、本研究ならびに寺本ら[2]で用いた打叩事象を中心とする視聴覚コンテンツは、打叩事象を含まない、もしくは中心的な内容としていない視聴覚コンテンツよりも映像と音の対応付けが容易である。その結果、視聴覚コンテンツの刺激強度の効果がより大きく観察された可能性が存在する。今後の研究では、打叩事象を中心としない視聴覚コンテンツを用いた実験を行うことが求められるだろう。さらに、本研究の実験結果は、視聴者の感覚様相のなかでも視覚と聴覚に関して得られたものであることから、視覚や聴覚以外の感覚様相でも同じく適用可能なかという点に関しては疑問が残る。そのため、我々の研究グループは、感覚様相のなかでも体性感覚に着目し、身体振動情報が多感覚コンテンツの臨場感や迫真性の知覚にどのような影響を及ぼすのかを調べた[25]。その結果、視聴者が映像から知覚する臨場感は振動の振幅に従って単調増加する一方で、迫真性は刺激強度を上げすぎると減衰することが示された[25]。したがって、本研究と寺本ら[2]が明らかにしてきた視聴覚コンテンツの臨場感と迫真性の規定因、特に刺激強度に関する知見は、特定の感覚様相に依存しないことから、マルチモーダルコンテンツ全般に適合する可能性が高いことが示唆される。

本研究と寺本ら[2]においては、臨場感と迫真性の規定因に関して共通する知見が得られた一方で、いくつかの点において差異も認められた。ひとつは、臨場感と

迫真性のピークに関するものである。寺本ら[2]の実験では臨場感と迫真性のピークは音と映像が同期する SOA 0ms 条件において観察されたが、本研究では映像よりも音が少し遅れる SOA +150ms 条件でそのピークが認められた。もう一つは臨場感と迫真性の時間窓に関するものである。時間ずれを許容できる範囲は、寺本ら[2]よりも本研究のほうが広いことが示された。従来の研究からは、視聴覚統合の時間限界は -200ms から +250ms 程度の範囲であること[21, 22]、発話音声を用いた場合は -320ms から +480ms の範囲であること[23]が報告されている。本研究から見いだされた時間窓の範囲は、それを上回るものであるが、近年の研究からは、打楽器のひとつであるマリimbaを使用した場合は視聴覚統合の時間窓が 0ms から +700ms の範囲へと後方へと拡大したことが報告されている[24]。これらの知見と本研究の実験結果をもとに考察すると、本研究で観察された時間窓の変化は、視聴覚刺激として打楽器(打叩音)が使用されたことに起因する可能性が高いことから、打楽器以外のコンテンツを刺激とした実験を行うことで同問題を検証する必要があるだろう。また、撮影地点が固定されている映像を視聴する場合は、自分が撮影対象からどの程度離れているのかを把握しやすいため、映像と音の時間ずれを感知しやすいことが知られている[26]。本研究の刺激は、トロンボーン奏者からシンバル奏者へ、そしてバイオリン奏者へと映像が切り替わるものであった。その結果、視聴者は自分の視聴位置を把握しにくくなり、映像よりも音が少し遅れて再生される SOA +150ms 条件において臨場感と迫真性のピークが認められたのかもしれない。このことから、今後の研究では、コンテンツの前景との距離感によって、臨場感と迫真性に視聴覚情報の時間ずれが与える影響が変化するかといった点についても検討する必要があるだろう。

このように、本研究は、臨場感と迫真性の規定因に関する知見に留まらず、許容可能な視聴覚統合の時間窓が視聴覚コンテンツの内容やジャンルによって異なること、具体的には視聴覚コンテンツとして打楽器を使用した場合は視聴覚統合の時間窓が後方へと拡大することを示す新たな知見を得ることができた。感覚情報の強度が視聴者の臨場感に及ぼす効果は、視聴覚コンテンツの内容によって異なることを示す知見はこれまで報告されているが[19, 20]、どのような内容が視聴覚統合の時間窓に影響を及ぼすのかに関しては不明確な部分が多い。本研究と打楽器を使用した先行研究[23]から見出された現象が、打楽器を使用した場合に限定して認められる特異的な現象に該当するのかについては更なる詳細な検討が求められるであろう。また、本研究の実験参加者は、心理学を専攻する大学生を中心に構成されていたことから、今後の研究においては、ユーザの属性に対しても焦点を当てることが必要だろう。

そして、本研究と寺本ら[2]では、臨場感と迫真性の主観評定値を分析対象としたが、今後の研究では、両感性と生理指標との関連性の検討も進めなければならない。臨場感については、心拍や皮膚電気反応との関係性が議論されているが[27]、迫真性に関しては、現状ではほとんど検討が行われていない。迫真性が前景的要素と関連した場面の感性印象として機能しており、その処理には焦点的注意が向いている対象や事象などの前景的要素の情報が重要[17]であるとの指摘を考えると、迫真性は視聴者の眼球運動や瞳孔の大きさと関連することが予測される。また、臨場感と迫真性に関係する脳内基盤を明らかにすることも重要である。柴田ら[28]は、臨場感評定時には異種感覚情報の統合や視点の変換に関係する下頭頂小葉の活動が認められるが、迫真性評定時には図の抽出や輪郭の処理に関わる中後頭回の活動が見られたと報告している。臨場感と迫真性がどのような生理指標や脳部位と関連するのかを明らかにすること、そして主観評定以外に複数の指標をもとに臨場感と迫真性に対して多角的にアプローチを行うことは、臨場感と迫真性という二つの感性がもつ本質的な意味を理解し、より総合的な VR システムの評価を実現するうえで極めて重要な問題であるといえる。

5 結論

本研究は、臨場感と迫真性という二つの感性の特性の違いやその規定因についての更なる知見を得ること、そして視聴覚統合の時間窓の範囲に関する新たな知見を得ることを目的とした。オーケストラの演奏映像を実験刺激として用い、その視野サイズ、音圧レベル、SOA を操作し、臨場感と迫真性の変化を視覚的アナログ尺度により測定した結果、打楽器演奏に起因すると考えられるいくつかの現象の相違点が認められたものの、寺本ら[2]と概ね共通する知見を得ることができた。具体的には、視聴覚コンテンツの視野サイズや音圧レベル、映像と音の同期性は、いずれも臨場感と迫真性の規定因ではあるが、臨場感は視聴覚コンテンツの強度(視野サイズや音圧レベルの大きさ)に、迫真性は視聴覚コンテンツの時間的な特徴(映像と音の時間的なずれ)に他方より強く依存することが明らかとされた。したがって、迫真性は臨場感とは異なる感性であり、今後 VR システムの総合的な性能評価を実施するうえで有益な知見を提供する可能性が高いといえる。そして本研究では打楽器演奏を中心的な内容とする音楽コンテンツを実験刺激として用いたが、視聴者が許容可能な視聴覚統合の時間窓の範囲は、寺本ら[2]の風景コンテンツの実験よりも拡大することが示された。これは、ユーザが許容可能な臨場感と迫真性の時間的な特性が視聴覚コンテンツの内容やジャンルによって変動することを示唆している。視聴覚統合の時間特性は、対象までの知覚距離に

よって異なるという報告もなされていることから[26], ユーザがどのような状況で何を体験したのかを詳細に記述することは, VR システムの総合的な性能評価を実施するうえで重要である。したがって, VR システムの総合的な評価においては, 実験に用いたコンテンツの内容や特性にも関しても注意を払うことが望ましい。さらに今後の研究においては, 臨場感と迫真性がどのような生理指標や脳部位と関連するのかを解明すること, そして複数の指標をもとにユーザの臨場感と迫真性に多角的にアプローチすることが求められるといえる。

謝辞

本研究は文部科学省科学研究費補助金特別推進研究(課題番号: 19001004)ならびに基盤研究(A)(課題番号: 24240016)の助成を受けた。また, 本研究成果の一部は, 日本バーチャリアリティ学会第 16 回大会, 12th International Multisensory Research Forum 等において報告された。なお, 本研究は, 東北大学電気通信研究所研究倫理委員会の承認を受けた。

引用文献

- [1] 寺本 渉, 吉田和博, 浅井暢子, 日高聡太, 行場次朗, 鈴木陽一: 臨場感の素朴な理解, 日本バーチャリアリティ学会論文誌, Vol. 15, No. 1, pp. 7-16, 2010.
- [2] 寺本 渉, 吉田和博, 日高聡太, 浅井暢子, 行場次朗, 坂本修一, 岩谷幸雄, 鈴木陽一: 「迫真性」を規定する時空間情報, 日本バーチャリアリティ学会論文誌, Vol. 15, No. 3, pp. 483-486, 2010.
- [3] 鈴木陽一, 寺本 渉, 吉田和博, 浅井暢子, 日高聡太, 坂本修一, 岩谷幸雄, 鈴木陽一: 視聴覚コンテンツの臨場感と迫真性, 電子情報通信学会技術研究報告: EMM, マルチメディア情報ハイディング・エンリッチメント, Vol. 111, No. 74, pp. 29-36, 2011-05-23.
- [4] Slater, M., Steed, A., McCarthy, J., Maringelli, F.: The influence of body movement on subjective presence in virtual environments. *Human Factors*, Vol. 40, No. 3, pp. 469-477, 1998.
- [5] Slater, M., Usoh, M.: Representations systems, perceptual position, and presence in immersive virtual environments. *Presence*, Vol. 2, No. 3, pp. 221-233, 1993.
- [6] Slater, M., Wilbur, S.: A Framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, Vol. 6, No. 6, pp. 603-616, 1997.
- [7] 江本正喜, 正岡顕一郎, 菅原正幸, 野尻祐司: 広視野静止画像による臨場感の提示視覚依存性と評価指標間の関係. 映像情報メディア学会誌, Vol. 60, No. 8, pp. 1288-1295, 2006.
- [8] 小澤賢司: 聴覚臨場感の基礎特性: コンテンツ臨場感とシステム臨場感. 電子情報通信学会技術研究報告, EA2008-115(2008-12), pp. 83-88, 2008.
- [9] Sheridan, T. B.: Musings on telepresence and virtual presence. *Presence: Teleoperators and Virtual Environments*, Vol. 1, No. 3, pp. 120-126, 1992.
- [10] Lessiter, J., Freeman, J., Keogh, E., Davidoff, J.: A cross-media presence questionnaire: The ITC-sense of presence inventory. *Presence: Teleoperators and virtual environments*, Vol. 10, No. 3, pp. 282-297, 2001.
- [11] Lindsay, E. D., Murray, S., Liu, D. K., Lowe, D. B., Bright, C. G.: Establishment reality vs. maintenance reality: How real is real enough?. *European Journal of Engineering Education*, Vol. 34, No. 3, pp. 229-234, 2009.
- [12] Helmick, R.: Virtues of verisimilitude in design and art. *Computers & graphics*, Vol. 19, No. 4, pp. 505-507, 1995.
- [13] Zyda, M.: From visual simulation to virtual reality to games. *Computer*, Vol. 38, No. 9, pp. 25-32, 2005.
- [14] Francis, A., Couture, M.: Credibility of a simulation-based virtual laboratory: An exploratory study of learner judgments of verisimilitude. *Journal of Interactive Learning Research*, Vol. 14, No. 4, pp. 439-464, 2003.
- [15] Weisstein, N., Maguire, W., Brannan, J. R.: M and P pathways and the perception of figure and ground. In J. Brannan (Ed.), *Application of parallel processing in vision*. Elsevier Science Publishers. pp. 137-166, 1992.
- [16] 行場次朗: 図と地の知覚: 視覚の心理, 電子情報通信学会誌, Vol. 74, No. 4, pp. 315-320, 1991.
- [17] 行場次朗, 寺本 渉: 臨場感と迫真性. 映像情報メディア学会誌, Vol. 66, No. 7, pp. 561-563, 2012.
- [18] Senkowski, D., Saint-Amour, D., Höfle, M., Foxe, J. J.: Multisensory interactions in early evoked brain activity follow the principle of inverse effectiveness. *NeuroImage*, Vol. 56, No. 4, pp. 2200-2208, 2011.
- [19] Lombard, M., Ditton, T. B., Grabe, M. E., Reich, R. D.: The role of screen size in viewer responses to television fare. *Communication reports*, Vol. 10, No. 1, pp. 95-106, 1997.
- [20] IJsselstein, W., de Ridder, H., Freeman, J., Avons, S.E., Bouwhuis, D.: Effects of stereoscopic presentation, image motion, and screen size on subjective and objective corroborative measures of presence. *Presence: Teleoperators and Virtual Environments*, Vol. 10, No. 3, pp. 298-311, 2001.
- [21] Shams, L., Kamitani, Y., Shimojo, S.: Visual illusion induced by sound. *Cognitive Brain Research*, Vol. 14, No. 1, pp. 147-152, 2002.
- [22] Dixon, N. F., Spitz, L.: The detection of auditory visual desynchrony. *Perception*, Vol. 9, No. 6, pp. 719-721, 1980.
- [23] Soto-Faraco, S., Alsius, A.: Deconstructing the McGurk-MacDonald illusion. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 35, No. 2, pp. 580-587, 2009.
- [24] Schutz, M., Kubovy, M.: Causality and cross-modal integration. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 35, No. 6, pp. 1791-1810, 2009.
- [25] Sakamoto, S., Hasegawa, G., Honda, A., Iwaya, Y., Suzuki, Y., Gyoba, J.: Effects of full body vibration on the perceived sense of verisimilitude and sense of presence. 13th International Multisensory Research Forum, 19-22 June 2012, Oxford, UK. (*Seeing and Perceiving*, Vol. 25, p. 111, 2012.)
- [26] Sugita, Y., Suzuki, Y.: Audiovisual perception: Implicit

estimation of sound-arrival time. Nature, vol.421, p. 911, 2003.

- [27] Meehan, M., Razzaque, S., Insko, B., Whitton, M., Brooks, F. P.: Review of four studies on the use of physiological reaction as a measure of presence in stressful virtual environments. Applied Psychophysiology and Biofeedback, Vol. 30, No. 3, pp. 239-258, 2005.
- [28] 柴田理瑛, 柴田 寛, 竹島康弘, 河地庸介, 川畑秀明, 行場次朗: fMRI を用いた迫真性及び臨場感評定時の脳活動部位の検討. 第 14 回日本ヒト脳機能マッピング学会, p. 76, 2012.

(2012 年 9 月 10 日受付)

[著者紹介]

本多 明生 (正会員)



2006 年東北大学大学院文学研究科後期課程修了. 博士 (文学). 日本学術振興会特別研究員, いわき明星大学研究助手, 東北大学電気通信研究所研究員等を経て, 現在, 東北福祉大学総合福祉学部助教. 専門は実験心理学, 認知工学, 災害心理学.

神田 敬幸



2012 年東北大学大学院情報科学研究科博士課程前期修了. 修士 (情報科学). 現在, リオン株式会社所属.

柴田 寛



2008 年東北大学大学院文学研究科後期課程修了. 博士 (文学). 日本学術振興会特別研究員, 東北大学電気通信研究所研究員, 東北大学大学院文学研究科研究助手等を経て, 現在, 東北文化学園大学医療福祉学部講師. 専門は認知心理学, 実験心理学.

浅井 暢子 (正会員)



2008 年神戸大学大学院文化学研究科博士課程修了, 博士 (学術). 名古屋大学博士研究員, 東北大学 COE フェローを経て, 現在, 京都文教大学総合社会学部講師. 専門は社会心理学. 特に社会的認知に関する研究に従事.

寺本 渉 (正会員)



2004 年神戸大学大学院文化学研究科博士課程修了. 日本学術振興会特別研究員 (PD), 東北大学電気通信研究所・同大学院文学研究科研究員を経て, 現在, 室蘭工業大学情報電子工学系学科准教授. 知覚心理学, 実験心理学, 特にマルチモーダル感覚情報処理に関する研究に従事. 博士 (学術).

坂本 修一 (正会員)



1997 年東北大学大学院情報科学研究科博士前期課程修了. 同年 (株) 日立製作所入社. 2000 年から東北大学電気通信研究所. 現在, 同所准教授. 博士 (工学). 聴覚・マルチモーダル感覚情報処理過程に関する研究に注力. 日本音響学会, 電子情報通信学会, 日本バーチャルリアリティ学会, 日本感性工学会等の学会会員.

岩谷 幸雄 (正会員)



東北学院大学工学部教授. 博士 (情報科学). 1993 年東北大学博士課程前期課程修了. 秋田大学助手, 講師, 東北大学准教授を経て現職. 音空間のレンダリングと知覚過程の解明に従事.

行場 次朗 (正会員)



東北大学大学院文学研究科心理学講座教授. 1981 年東北大学大学院文学研究科博士課程心理学専攻修了. 信州大学人文学部助教授, 九州大学文学部助教授を経て, 2002 年より現職. 専門は視覚心理学・認知心理学. 共著書, 「知性と感性の心理」(福村出版), 「視覚と聴覚」(岩波書店) など. 日本心理学会, 日本基礎心理学会, 日本認知心理学会, 電子情報通信学会各会員.

鈴木 陽一 (正会員)



1981 年東北大学大学院工学研究科博士課程修了. 現在, 東北大学電気通信研究所教授. 3次元音空間知覚と聴覚ディスプレイ, マルチモーダル知覚などの研究に従事. 日本 VR 学会フェロー, 評議員. 米国音響学会フェロー. 著書に音響学入門 (コロナ社) など. 工博.