

「迫真性」を規定する時空間情報

寺本 渉^{*1*2} 吉田 和博^{*1*3} 日高 聡太^{*4} 浅井 暢子^{*2} 行場 次朗^{*2}
坂本 修一^{*1} 岩谷 幸雄^{*1} 鈴木 陽一^{*1}

Spatio-temporal characteristics responsible for high “Vraisemblance”

Wataru Teramoto^{*1*2} Kazuhiro Yoshida^{*1*3} Souta Hidaka^{*4} Nobuko Asai^{*2} Jiro Gyoba^{*2}
Shuichi Sakamoto^{*1} Yukio Iwaya^{*1} and Yōiti Suzuki^{*1}

Abstract – For virtual reality systems, the enhancement of a sense of presence (a subjective experience of being in one place even when one is physically situated in another) has been the most important issue. Both theoretically and empirically, the sense of presence has been found to relate dominantly to background components contained in a scene. In contrast, the reality or virtuality which can be assumed to link essentially to foreground components in a scene has not been investigated in detail. The present study defined the latter type of sense as *vraisemblance* (verisimilitude), and made an exploratory investigation into spatio-temporal characteristics responsible for the higher *vraisemblance* by using a scene containing *Shishi-odoshi* (a traditional Japanese fountain made of bamboos) in a Japanese garden as audio-visual stimuli. In Experiment 1, the effects of the field size of view and the sound pressure level of the background were investigated. Higher *vraisemblance* was observed with the middle field size of view with the original sound pressure level of the background, whereas higher sense of presence was observed with the larger field size of view with the larger background sound. In Experiment 2, the effect of temporal asynchrony between the foreground audio-visual stimuli produced by *Shishi-odoshi* was investigated. The results show that the range of temporal-window for the audio-visual stimuli necessary for high *vraisemblance* was different from those for high presence. These findings suggest that the sense of *vraisemblance* can be distinguishable from the sense of presence, and deeply involved to the foreground-based aesthetic impression in a scene.

Keywords : *vraisemblance*, foreground, background, virtuality

1. はじめに

情報通信技術の進歩に伴い、伝達可能な音声や映像データの量は日々増加し、今や自然でリアルなバーチャルリアリティシステムの実現に対する期待が高まっている。従来、そのシステムが実現する音声や映像の評価として、「臨場感」の有無・強度がよく取り上げられ、創出要因や評価法に関する研究も盛んに行われてきた(例えば, [1] [2] [3] [4] [5])。しかし、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、すなわち、背景的な「場」の本物らしさに関する感性印象であり、必ずしも対象や事象など前景的要素を中心とした

本物らしさに関する感性ではないと考えられる。例えば、バーチャルリアリティシステムにおいて、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。知覚心理学の分野では、前景 (図) と背景 (地) の処理には別々の処理システムが関与している証拠が示されている [6] [7]。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、意識的に詳細な分析を受ける対象となると考えられてきた [7]。これらの点を踏まえると、背景的要素が中心的役割を果たすと考えられる臨場感のほかに、前景的要素と関連した場面の感性印象を抽出し、正しく評価する試みが必要である。本研究では、後者の感性を「迫真性 (*vraisemblance* または *verisimilitude*)」と定義し、その評価可能性について、時空間パラメータを操作し、臨場感と比較しながら、探索的に検討を行った。実験素材には、ししおど鹿威し

*1: 東北大学電気通信研究所

*2: 東北大学大学院文学研究科

*3: 東北大学大学院情報科学研究科

*4: 立教大学現代心理学部

*1: Research Institute of Electrical Communication, Tohoku University

*2: Graduate School of Arts and Letters, Tohoku University

*3: Graduate School of Information Sciences, Tohoku University

*4: College of Contemporary Psychology, Rikkyo University

を用いた。鹿威しは、現実世界の対象の中で特徴的な視聴覚刺激を有し、時間情報を操作しやすく、また、単純な仕掛けでありながら日本の美や心地よさなど豊富な感性情報を含むので実験素材としてふさわしいと考えた。

2. 実験 1

実験 1 では、臨場感に強く寄与するとされる提示視野範囲および背景音圧 (例えば, [2] [8]) が迫真性に与える影響に注目し、検討を行った。

2.1 実験刺激

宮城県白石市郊外にある鎌先温泉の旅館の庭園に設置されている鹿威しの映像と音を収録し、編集を行った (Fig. 1; 距離: 1.6 m, 焦点距離: 65 mm)。撮影にはデジタルビデオカメラ (Panasonic, AG-DVX100A)、收音マイク (Bruel & Kjar, Type4165) 及び計測用アンプ装置 (Bruel & Kjar, Type2209) を用いた。音信号はデジタルビデオカメラに外部入力として取り込み、映像とともにデジタルビデオテープに記録した。映像の編集にはノンリニア編集ソフトウェア (Avid Xpress Pro 5.6) を用い、映像全体の長さが 8.14 s、竹筒の末端部が敷石を叩く部分 (打叩部) が動画の開始部分から数えて約 5.97 s 後になるよう設定した。音源の編集には Matlab (The MathWorks, Inc.) を用い、前景音 (鹿威しの流水音発生から、竹筒の末端部が敷石を叩く音までの 2.4 s 間) と背景音 (前景音以外の環境音であり、虫や鳥の鳴き声が含まれている 8.14 s 間の音源) を分けて抽出・編集したものを組み合わせて使用した。また、前景音の立ち上がりとしち下がりには 10 ms のコサイン窓をかけた。

2.2 実験条件

視野サイズ (大, 中, 小) × 背景音 (大, 中, なし) の 2 要因実験参加者内デザインを用いた。視野範囲は、視野制限のための窓枠を付け加えることによって操作した。視野大条件 (直径 28.9 deg) は画面に収まる最大の窓枠サイズであり、視野小条件 (同 16.6 deg) は鹿威し全体が見える最小の視野範囲、視野中条件の視野範囲はそのちょうど中間であった (同 22.8 deg)。背景音の音圧レベルは、集音時の音圧 (背景音中条件) を基準に、10 dB 音圧を上げたもの (背景音大条件) と背景音が全くないもの (背景音なし条件) の 3 種類であった。それら背景音と前景音 (集音時のもの) を重畳した。

1: 鹿威しは水力によって音を発生する装置である。複数の節をもつ竹筒の中心に支点を設け、竹筒の一方を開放し、水を引き入れる。竹筒に蓄えられた水が一定量を超えると、その重みによってシーソーのように竹筒が傾き、水がこぼれて元の位置にもどる。その際、竹筒の末端部が敷石を叩き、カコンという特徴的な音が発生する (Fig.1 参照)。



図 1 実験で用いた鹿威し画像 (視野大条件)

Fig. 1 An image of Shishi-odoshi in the Large field-of-view condition.

2.3 実験手続き

正常な視覚、聴覚を持つ大学生 20 名が実験に参加した。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, 60 Hz) によって刺激を視聴した。実験ではシェッフェの一対比較法に基づき、2つの条件の映像を連続提示し、「後の映像が前の画像に比べ、臨場感 (あるいは迫真性) があるか」を -3 (後の方がない) ~ +3 (後の方がある) の 7 段階尺度で判断させた。なお、臨場感や迫真性として受け取られる感性の多義性を回避するため、それぞれ「その場にいる感じがする」「本物らしい趣がある」という注釈を提示した。9 条件のすべてのペアおよびその提示順序を逆にしたペアをランダムな順序で提示した (臨場感ブロック、迫真性ブロックそれぞれ 72 試行)。半数の実験参加者に臨場感を先に評定させ、残り半数に迫真性を先に評定させた。

2.4 結果と考察

得られたデータに尺度構成を施した結果を Fig. 2 に示す。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。横軸は提示刺激 (視野サイズ/背景音) である。感性 (2) × 視野 (3) × 背景音 (3) の 3 要因分散分析を行った結果、視野サイズの主効果 ($F_{2,38} = 61.03, p < .001$), 背景音の主効果 ($F_{2,38} = 28.86, p < .001$), 感性 × 視野サイズの交互作用 ($F_{2,38} = 7.05, p < .005$), 感性 × 背景音の交互作用 ($F_{2,38} = 18.22, p < .001$), 視野サイズ × 背景音の交互作用 ($F_{7,76} = 7.55, p < .001$) が見られた。感性 × 視野サイズの交互作用の下位検定および多重比較 (Tukey's HSD, $p < .05$) 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わかった。感性 × 背景音の交互作

2: 鹿威しは、日本庭園などに多く見られ、その実用的な機能よりも、風情や趣を醸し出す観賞用装置としての役割もつ。そのため実験では、鹿威しのある場面の迫真性を表す表現として「本物らしい趣がある」と定義した。

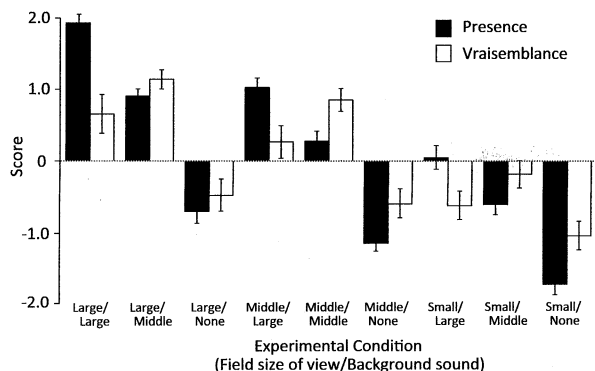


図2 臨場感と迫真性に対する提示画角と背景音圧の効果

Fig. 2 The effects of the field size of view and the sound pressure level of the background on sense of presence and vraisemblance.

用に関しても同様に分析を行った結果、臨場感は背景音が大きくなるにつれて高まり、迫真性は、背景音なし、大、中の順で高まることわかった。臨場感に関して、視野サイズや背景音が大きいほど高まることを示す以上の結果は、迫力や力量性と高い相関をもつという従来の研究に [2] [8] 一致するものであり、背景情報の重要性を示唆する。一方、迫真性に関しては、視野サイズも背景音も中程度がよいことを示しており、前景情報と調和した背景情報の必要性を示唆する。

3. 実験 2

実験 1 では、視野サイズと背景音という事象の背景的要素、すなわち空間的情報を操作することによって臨場感と迫真性の違いを示した。こうした空間的情報以外に、臨場感に強く影響を与えるものとして、各感覚モダリティからの情報の同期性が挙げられる [5] [9] [10]。いくら多くの感覚モダリティに刺激を与えても、その情報が時間的一致性に欠ける場合には、臨場感が高まず、単一感覚モダリティへの刺激よりも低下する可能性もある [10]。そこで、実験 2 では音と映像の同期性を操作し、臨場感、迫真性に対する影響を調べた。

3.1 方法

実験 1 で用いた背景条件のうち、最も臨場感と迫真性の違いが明確に現れた視野サイズ大/背景音大条件と視野サイズ中/背景音中条件を用いた。SOA は、鹿威しが敷石を叩く際の視覚的打叩時刻と聴覚的打叩時刻 (前景音) のずれと定義し、6 つの条件 (-300/-150/0/150/300/450 ms) を設定した。これら 2 (背景条件) × 6 (SOA) に関して、実験参加者内実験デザインを用いた。12 条件のすべてのペアおよびその提示順序を逆にしたペアをランダムな順序で提示した (臨場感ブロック、迫真性ブロックそれぞれ 132 試行)。

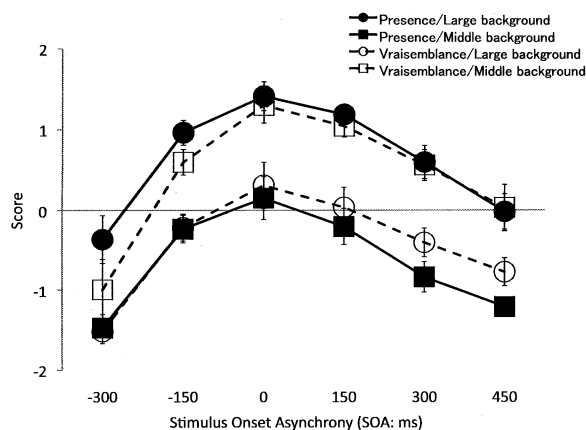


図3 臨場感と迫真性に対する SOA の効果

Fig. 3 Effects of the SOA of audio-visual components contained in Shishi-odoshi on sense of presence and vraisemblance.

3.2 結果と考察

得られたデータに尺度構成を施し (Fig. 3), 感性 (2) × 背景 (2) × SOA (6) の 3 要因分散分析を行った結果、SOA の主効果 ($F_{5,45} = 18.57, p < .001$), 感性 × 背景の交互作用 ($F_{1,9} = 19.66, p < .005$), 感性 × SOA の交互作用 ($F_{5,45} = 3.67, p < .01$), 感性 × 背景 × SOA の 2 次の交互作用 ($F_{5,45} = 2.80, p < .05$) が見られた。2 次の交互作用について下位検定を行ったところ、臨場感に関しては全ての SOA の条件で背景大条件の方が中条件よりも有意に高く、迫真性に関しては -300 ms SOA 条件以外で背景中条件の方が大条件よりも有意に高かった ($p < .05$)。このことは、臨場感にとって背景情報の豊富さが極めて重要な役割を果たしていることを示唆する。

次に SOA の効果を SOA 0 ms 条件を基準にどの程度の時間ずれを許容できるか (0 ms SOA 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、SOA の単純単純主効果について多重比較を行った (Tukey's HSD, $p < .05$)。その結果、臨場感の場合、背景条件によらず、-150 ms ~ 150 ms であった。迫真性に関しては、前景に焦点があたるような背景中条件では 0 ms ~ 150 ms, 背景が強調された背景大条件では -150 ms ~ 300 ms であった。これは臨場感と迫真性の中で最適な視聴覚情報の時間窓が異なることを示唆する。

4. 総合考察

本研究では、「迫真性」の評価特性について探索的に検討した。実験 1 では視覚・聴覚的背景情報を操作した。その結果、臨場感には、背景情報の量的増加が重要であり、現実を超えるような刺激でも高まること

示され、背景情報が前景情報とは独立に作用することがわかった。一方、迫真性に関しては、前景情報と調和した背景情報の重要性が示唆された。実験2では視聴覚刺激の SOA を操作した結果、臨場感と迫真性の間で視聴覚情報の時間ずれの許容範囲が異なることが示された。以上の結果は、迫真性が前景情報を中心とした本物らしさに関連する感性印象であり、提示情報の本物らしさを、臨場感とは別の観点から計測するための指標となりうることを示している。

「場」としての強度を強めることで臨場感が高まることは従来から示されている [2] [8]。例えば、小澤は視聴覚刺激を同時に提示した場合には、実際に収録した音圧以上の音圧で提示することによって臨場感が高まると報告している [8]。また、江本らは、画像提示面積が広いほど高い臨場感が得られることを指摘している [2]。本研究の結果はこれらの知見と一貫する。それに対して、迫真性は「場」としての強度を強めても必ずしも高まらず、むしろ、前景情報や別の感覚モダリティ情報など様々な情報の調和を反映したものであった。この点からも、迫真性が、臨場感とは異なる感性印象であることが示唆される。

また、臨場感は、複数の感覚モダリティ情報の時間的一致性によって強く影響されることも指摘されてきた [5] [9] [10]。実験心理学的研究では、様々な刺激を用いた検討から、視聴覚統合の時間限界は $-200\text{ ms} \sim 250\text{ ms}$ 程度と報告している [11] [12]。今回見いだされた $-150\text{ ms} \sim 150\text{ ms}$ (臨場感)、 $0\text{ ms} \sim 150\text{ ms}$ (迫真性・背景中条件)、 $-150\text{ ms} \sim 300\text{ ms}$ (迫真性・背景大条件) という時間範囲は、先の視聴覚統合の時間限界に含まれるものであり、視聴覚刺激の統合過程が臨場感だけでなく、迫真性とも密接な関係にあることが示唆される。また、迫真性に関しては、背景条件によって視聴覚刺激の時間ずれの許容範囲が異なっていた。背景中条件では、視覚的にも聴覚的にも空間的に調和のとれた刺激が提示されたことで、前景に適切に焦点が向けられ、時間ずれの許容範囲が非常に狭くなったと考えられる。それゆえ、打叩音が視覚的打叩に先行するという物理的にはあり得ない状況が適切に検出され、迫真性の評価に反映されたのだろう。一方、背景刺激が極端に強調された背景大条件では、前景に焦点を向けることが難しかったために、迫真性評定値が全体的に低下し、特にプラス側の時間ずれに対してより緩い減衰を示したと考えられる。ただし、視聴覚統合の時間特性は、対象までの知覚された距離によって異なるという報告もあり [13]、今後、今回見出された時間限界と知覚距離の関係を調べることで、迫真性の性質をより深く議論することが可能になると考えられる。

本研究では、刺激視野面積、背景音圧、SOA の影

響のみを検討した。今後臨場感と迫真性の差違をさらに明らかにするためにも、前景音や反射音、残響音など様々な要素について検討していくことが必要である。また、鹿威し以外の素材を用いて同様の検討を行うことにより、本研究で見いだされた迫真性の性質の一般性を明らかにできると考える。

謝辞

本研究は科学研究費補助金特別推進研究 (19001004) の助成を受けた。また、鎌先温泉・旅館みちのく庵のご協力のもと鹿威しの撮影を行った。

参考文献

- [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; 日本バーチャルリアリティ学会誌, vol.15, 7-16 (2010).
- [2] 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, 1288-1295 (2006).
- [3] 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, 282-298 (2001).
- [4] Slater, M., Steed, A., McCarthy, J., and Maringelli, F.: The influence of body movement on presence in virtual environments; Human Factors, vol.40, 469-477 (1998).
- [5] Sheridan, T. B.: Musings on telepresence and virtual presence; Presence: Teleoperators and Virtual Environments, vol.1, 120-126 (1992).
- [6] Weisstein, N., Maguire, W., and 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. 137-166 (1992).
- [7] 行場次朗: 図と地の知覚-視覚の心理-; 電子情報通信学会誌, 74, 315-320 (1991).
- [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, 108-333, 83-88 (2008).
- [9] Held, R., and Durlach, N. : Telepresence; Presence: Teleoperators and Virtual Environments, vol.7, 109-112 (1992).
- [10] Witmer, B. G., and Singer, M. J.: Measuring presence in virtual environments: A presence questionnaire; Presence: Teleoperators and Virtual Environments; vol.7, 225-240 (1998).
- [11] Shams, L., Kamitani, Y., and Shimojo, S.: Visual illusion induced by sound; Cognitive Brain Research; vol.14, 147-152 (2002).
- [12] Dixon, N. F. and Spitz, L.: The detection of auditory visual desynchrony; Perception; vol.9, 719-721 (1980).
- [13] Sugita, Y. and Suzuki, Y.: Audiovisual perception: Implicit estimation of sound-arrival time; Nature; vol.421, 911 (2003).

(2010年4月26日)