

多感覚情報の同期ずれが体験中の高次感性知覚に与える影響

柳生 寛幸^{*1} 崔 正烈^{*1} 坂本 修一^{*1}
 大谷 智子^{*2} 鈴木 陽一^{*1} 行場 次朗^{*3}

Effects of multimodal contents on temporal shifts in reality perception

Hiroyuki YAGYU^{*1} Zhenglie CUI^{*1} Shuichi SAKAMOTO^{*1}

Tomoko OHTANI^{*2} Yôiti SUZUKI^{*1} and Jiro GYOBA^{*3}

Abstract – Understanding the mechanism of virtual reality perception is crucial for the development of high-definition communication systems in multimedia contents. Recent results show that the perception of reality can be enhanced by not only audio-visual contents but adding full-body vibratory information. However, the effects of stimulus onset asynchrony (SOA) on reality perception have not been clearly understood. In this study, we investigated the effects of audio-visual SOA and visual-vibratory SOA on the intensity of reality perception. Results indicated that reality perception was more sensitive to audio-visual SOA than to visual-vibratory SOA.

Keywords : presence, verisimilitude, multimodal integration, stimulus onset asynchrony, reality perception

1 はじめに

今日の情報通信技術の発展により、高品質な音声や映像を利用者に届けることが容易となりつつある。このような技術の高度化にともない、より自然でリアリティのある空間を実現する技術を基礎とした様々な技術応用が期待されている (例えば [1])。その実現には「あたかもその場にいるような感じ」や「本物らしい感じ」といった空間に対する高次感性知覚を利用者に生起させることが欠かせない。しかし、そのような高次感性知覚はいかにして創出され、どのような特性を有するのかについては不明な点が多く、その解明が期待されている。

これまでに、高次感性知覚の感性指標として、臨場感 (sense of presence) と迫真性 (sense of verisimilitude) に関する研究が行われてきた。臨場感には「あたかもその場にいるような感じ」という辞書的意味合いがあるが、実際には多義的な印象を持つことがわかっている ([2] [3] [4] [5] など)。一例としては、心を揺さぶる事象に対しても用いられる傾向があることが示されている [4]。臨場感の特性については、映像の視野サイズ ([6] [7] [8] など) や、提示音の総合的な音圧レベ

ル [7] [9] などの量的増加により印象強度が増加する特性を持つと報告されている。このことは、臨場感が主に背景的な「場」に関する感性指標であることを示している。

一方、空間の評価には背景的要素に関する臨場感についてだけでなく、前景的要素に関する高次感性知覚についても理解することが重要であると考えられる。例えば、コンサートホールで音楽鑑賞をする場合、我々は周辺の観客やホールといった背景的な「地」の要素に囲まれながら、ステージ上の演奏者という前景的な「図」の要素に注意を向けている。この空間をバーチャルリアリティシステムで実現する場合、前景的要素である演奏者に関する情報もリアルに再現されることが求められる。また、前景 (図) と背景 (地) とでは別々の知覚情報処理系が働いている [10] という知見から、前景的要素に関する高次感性知覚の存在が予想される。寺本ら [7] は、そのような前景的要素の本物らしさを評価する高次感性指標として迫真性を定義した。実験の結果、迫真性は主に前景情報を中心としながらも、背景情報との調和も関与する感性指標であることが示された。例えば視野サイズや音圧などについては、前景的情報と背景情報の調和のとれた適度な大きさの感覚情報で印象強度が飽和し、ピークを形成する特性を持つと報告されている [7] [8]。なお、臨場感と迫真性が異なる特性を示すことは、様々なコンテンツにおいても確認されている ([8][11][12] など)。これは、研究で用いられた視聴覚コンテンツ以外においても臨場感と迫真性の定義が一般性を失わないこと

^{*1}東北大学 電気通信研究所/大学院情報科学研究科

^{*2}東北大学 電気通信研究所/大学院情報科学研究科 (現所属: 東京藝術大学芸術情報センター)

^{*3}東北大学 大学院文学研究科

^{*1}Research Institute of Electrical Communication, Tohoku University

^{*2}Art Media Center Tokyo University of the Arts

^{*3}Graduate School of Arts and Letters, Tohoku University

を示していると考えられる。

以上のように臨場感と迫真性に関する従来研究から、臨場感と迫真性の生起要因やその特性の異なることが、部分的ではあるものの明らかにされてきた。さらに、臨場感、迫真性をマルチモーダルな情報処理過程と捉えて、近年では視覚や聴覚だけでなく他の感覚情報も含むコンテンツを用いた研究が行われるようになってきた。それにより、視聴覚情報だけではなく全身振動情報も臨場感および迫真性の創出要因となることが報告されている [13]。全身振動情報を含む多感覚情報コンテンツを用いた結果、臨場感では全身振動情報の強度が大きいほど印象強度が強くなり、迫真性では適度な大きさで印象強度が飽和する特性を持つことが報告されている。また高橋ら [11] は、視聴覚情報を同期した状態で振動情報の同期ずれ (SOA: stimulus onset asynchrony) を変化させた結果、臨場感と迫真性ともに同期ずれ条件によって印象強度が変化すると報告している。

ここまで述べてきた従来研究はいずれも実験刺激を体験後の時点で、観測者に臨場感または迫真性について回答させた結果による知見である。感性知覚は体験中に逐次的に生起されると考えられることから、体験中における臨場感および迫真性の印象強度に関する知見は、高次感性知覚の生起要因およびその特性を明らかにするのに重要である。しかし現在のところ、体験中における臨場感および迫真性について検討した研究は少ない。体験中に関する検討としては、聴覚臨場感を心理学的測定法で測定した研究 ([14] [15]) や、CG映像空間についての臨場感を生理学的指標によって測定した研究 ([16] [17]) があるが、いずれも単独の感覚情報による臨場感についての知見であり、多感覚情報により生じた臨場感および迫真性に関しては明らかになっていない。

本研究では、従来研究で検討が見られない、視聴覚に全身振動を加えた多感覚コンテンツ体験中の臨場感および迫真性の印象強度について検討する。具体的には、多感覚実験刺激を体験中に生起する臨場感と迫真性の印象強度の時間推移特性に着目し、提示情報との関連を特に感覚間の同期ずれが与える影響を中心に検討する。あわせて、多感覚コンテンツ体験後の臨場感および迫真性の総合的印象強度について、同様の検討を行う。

2 方法

2.1 実験刺激

実験刺激として、走行する電車の映像、音、および振動情報を記録した収録素材を用いた。この収録素材は、走行する電車が通過する映像、音、振動を線

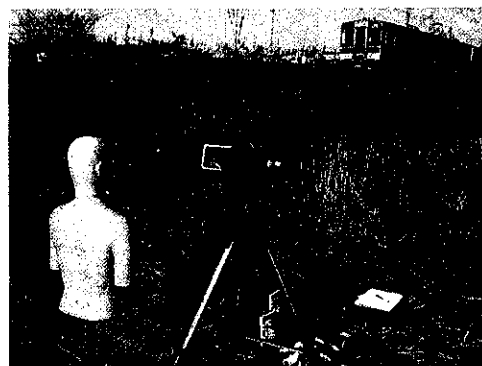


図1 収録風景

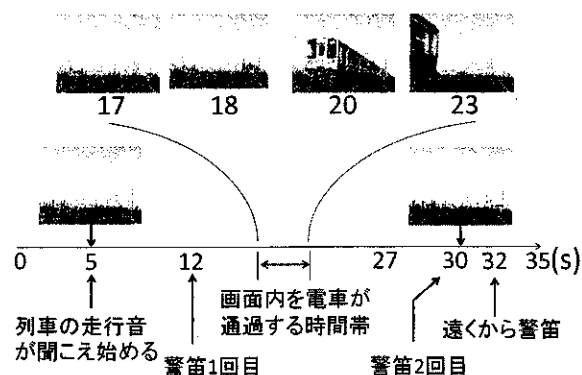


図2 実験刺激の流れ

路沿いの位置で記録したものである。実際の収録風景を図1に示す。映像の録画はDVカメラレコーダ (Panasonic AG-DVX100A) で行い、音はカメラの横に設置したダミーヘッド (高研 SAMRAI) 両耳の耳孔にコンデンサマイク (B & K 4101) を取り付けてバイノーラル録音し、振動はカメラ設置場所付近の地面に設置した加速度ピックアップ (RION VM-80) を用いて収録した。

高橋ら [11] は収録素材のうちの連続した 15 s を取り出して実験刺激に用いていたのに対し、本実験では臨場感および迫真性の継時変化を測定するために 35 s の実験刺激を用いた。この 35 s 間は、高橋らの実験刺激の前に 12 s、後ろに 8 s 加えたものである。図2は実験刺激 35 s 間の情景の変化の概要と主なイベントをまとめたものである。音圧および振動は列車がカメラの前を通過する時間帯 (およそ 17~23 s) で大きくなる。本実験では、映像を基準として音と振動情報それぞれの SOA (stimulus onset asynchrony) を操作して提示した。

音と振動情報は、映像と同じ場面の素材をもとに作成した。これらを SOA が大きい状態でも空白が出ないように 35 s よりも長い時間を切り出した上で、そこから映像を基準として設定した SOA に基づいて 35 s 分

を抽出して、映像と組み合わせて提示した。さらに、映像、音、全身振動の提示終了後の継時変化についても測定を行うため、提示終了後に映像のみ黒色の画面を3 s間追加した。よって1試行あたりの実験刺激は合計38 sとなった。

2.2 実験条件

実験条件は、 $2 \times 9 \times 9$ 水準の3要因混合計画とした。第1の要因は感性印象（臨場感・迫真性：被験者間要因）、第2の要因は映像-音間 SOA 条件（ $-800 \sim +800$ ms まで：被験者内要因）、第3の要因は映像-振動間 SOA 条件（ $-800 \sim +800$ ms まで：被験者内要因）である。SOA 条件は、映像と音、振動の測定時のタイミングを基準（0 ms）とし、映像に対して、音、振動刺激それぞれを 100, 200, 400, 800 ms ずつ遅延（+100, +200, +400, +800 ms）、および 100, 200, 400, 800 ms ずつ先行（-100, -200, -400, -800 ms）させて、同時（0 ms）を含めて合計9種類とした。

2.3 観測者

正常な視覚、聴覚を有する大学生16名（男性10名、女性6名）の観測者を、各感性印象のグループ間で男女比が同じとなるように割り当てた。臨場感、迫真性は共に8名（男性5名、女性3名）となった。臨場感を評価したグループの年齢は22.0歳（標準偏差1.8歳）。迫真性を評価したグループの年齢は22.1歳（標準偏差3.1歳）であった。なお、観測者全員の利き手は右であった。

2.4 実験環境

実験環境の概要を図3に示す。観測者の立ち位置からスクリーンまでの距離を3.0 mとした。提示画面角は観測者の立ち位置から26 deg.（水平方向）、18 deg.（垂直方向）であった。映像刺激は観測者の視野中央に映像刺激の中心が提示されるように設定した。収録素材の視覚刺激（解像度： 960×720 、フレームレート：30 fps）、聴覚刺激（サンプリング周波数：48 kHz、量子化ビット数：16 bit、チャンネル：2 ch、コーデック：PCM）および全身振動となる振動刺激（サンプリング周波数：48 kHz、量子化ビット数16 bit、チャンネル：4 ch、コーデック：PCM）を、それぞれDLPプロジェクタ（SANYO PDG-DHT100JL）、密閉型ヘッドフォン（SENNHEISER HDA 200）、モーションプラットフォーム（D-BOX MASTERING MOTION:D-BOX）から提示した。聴覚刺激と振動刺激は、測定時と等しい音圧レベルおよび加速度で提示した。聴覚は最大音圧レベル106.7 dBとなった。振動は、振動加速度の最大値がおよそ 20 cm/s^2 となった。

観測者は暗室の中に、ヘッドホンを装着した状態で、設置されたモーションプラットフォームの上に自然な

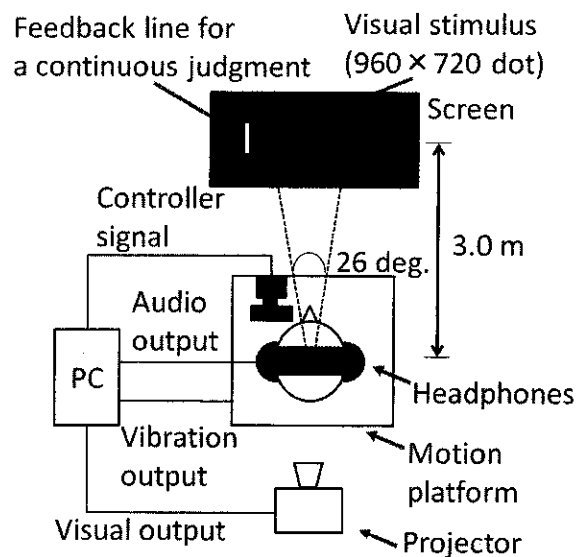


図3 実験環境

体勢で立ち、スクリーンの中央方向を向いて実験刺激を体験した。スクリーン上に投写される視覚刺激の横に表示される縦線の線分長操作用に、スロットレバー型コントローラ（SAITEK Throttle Quadrant）を観測者の利き手側に設置した。このコントローラは上下方向のみに滑らかに可動させることができ、任意の角度にレバー位置を保持することもできる。コントローラの上下に対応して線分長はスクリーン上部方向に伸縮し、2 dotの解像度で2~400 dot（観測者位置からの画角換算でおよそ0.04~15.19 deg.）の範囲で線分長を変化させることができる（線分の横幅は常に2 dot）。線分の表示更新周期は30 ms、線分長の記録周期は100 msとした。

2.5 実験手続き

実験に際し、観測者には臨場感を「その場にいる感じ」、迫真性を「本物らしい感じ」と定義して評価するように求めた。ここで臨場感または迫真性の定義を教示するとき、指定された感性指標について観測者間で共通のイメージを持つように、各感性指標に関する例文をまとめたインストラクションを同時に提示した。インストラクションの一例は次の通りである。臨場感の場合は「野球場のバッテリーボックス近くの席に自分が座り、バッテリー、ピッチャー、キャッチャー、観客の歓声など全ての挙動が体感できる感じ」であり、迫真性の場合は「落語家が、扇子でそばを食べる真似をしている。本当にそばをおいしそうに食べているように見える」という例文であった。また、観測者には割り当てられたグループ側の感性印象についてのみ教示を行った。例えば、臨場感を回答するグループの実験参加者に教示を行う場合、臨場感についてのみ教示を

行い、迫真性については一切触れないこととした。

実験の流れは以下のとおりである、初めに観測者の正面方向となるスクリーン位置に注視点が表示され、その後、38 s 間の実験刺激が提示された。実験刺激の提示終了時には観測者に終了の合図音が提示された。この間、試行中の臨場感または迫真性の印象強度について、手元のコントローラを操作してリアルタイムで回答するように教示した。また、1 試行が終了ごとに試行全体の臨場感または迫真性の印象強度について回答を求めた。体験中および体験後の回答を行う際は、実験刺激の中の映像だけや音だけといった単独の情報で判断するのではなく、実験刺激全体について注意を払って臨場感または迫真性の回答を行うよう教示した。

観測者は練習として 2 試行を体験した後、各 SOA の組み合わせ 81 試行を 1 セッションとし、全 3 セッションで構成される本実験に参加した。セッション内の試行順はランダムとした。実験を行うに際し、セッション内での 40 試行目終了後および各セッションの境目で必ず休憩を取ることとした。休憩を含めて全セッションを行うのにかかる時間はおよそ 4 時間であった。

2.6 回答方法

体験中の回答方法として、線分長との連続的クロスモダリティマッチングを用いた [18]。これは画面に表示される線分の長さの印象と、刺激の感覚量をリアルタイムで一致させて回答する手法である。画面に表示される線分長を操作して、その線分長の印象と試行中に体感した臨場感または迫真性の印象強度をリアルタイムで一致させて回答するよう観測者に教示した。この手法は、線分の最大長と最小長の定義を行う必要があるため、線分の最大長（線分長 400 dot）を「日常生活で経験する最大の臨場感または迫真性に対応」、最小長（線分長 2 dot）を「臨場感または迫真性が全くない状態に対応」として考えるよう求めた。また、試行開始時は必ず線分長を最小の状態から開始し、回答はコントローラのレバー角度ではなく、あくまでスクリーン上の線分長を印象強度とマッチングさせるように指示した。

体験後の回答方法として、観測者には 1 試行終了ごとに、試行全体の臨場感または迫真性の印象強度について評定尺度法による回答を求めた（0: 全くない～6: 非常にある、の 7 段階）。

3 実験結果

3.1 印象強度の経時変化

印象強度の経時変化について、体験中の臨場感と迫真性の観測者平均を図 4 に示す。図の縦軸は印象強度として回答した線分長で、横軸は時間経過を示す。ここでは、全ての感覚情報が同期した条件（SOA: 0 ms）

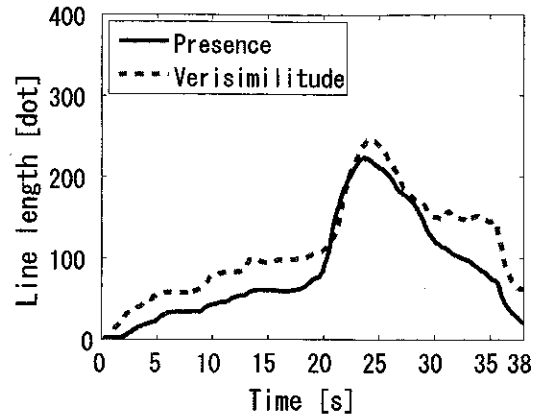
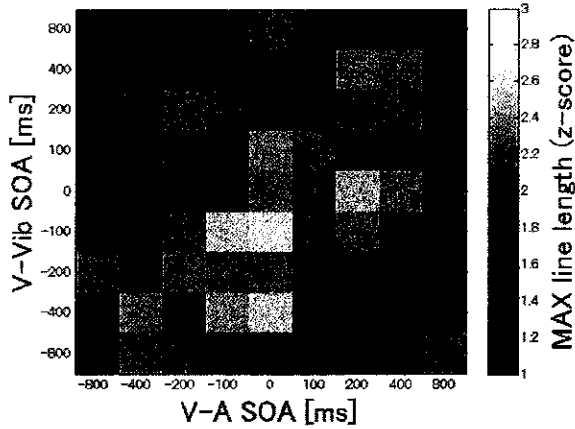


図 4 体験中の印象強度の平均値（映像-音間 SOA 条件および映像-振動間 SOA 条件 0 ms 時）

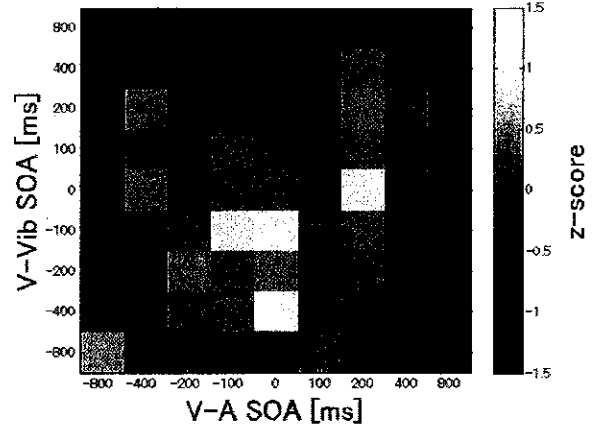
のみを示すこととする。臨場感、迫真性ともに提示開始から 20 s までなだらかに上昇し、20 s から 30 s の間で印象強度が山型となり、印象強度が最大となる傾向が見られる。ただし、印象強度が最大となった以降は、臨場感と迫真性では 30 s から 35 s までの経時変化が異なり、臨場感は比較的速度やかに減少する傾向を示す一方で、迫真性は印象強度を保持するように減少が少ない傾向が見て取れる。映像、音、振動の提示が終了となる 35 s から 38 s では急激に印象強度は減少するものの、臨場感と迫真性がわずかに存在する状態で終了する。これらの傾向は、SOA 条件が 0 s 以外のいずれの条件でも観測された。

映像-音間 SOA、もしくは映像-振動間の SOA が、体験中における臨場感および迫真性の最大印象強度にどのような影響を及ぼすのかを調べるために、観測者ごとに 35 s 間の全ての印象強度を基に z 得点を算出し、その中で SOA 条件の組み合わせごとに最大となる印象強度の z 得点を感性印象のグループの全観測者について平均した。先行研究では、臨場感はその「場」にいる感じを表し背景的要素と強く結びついた感性指標であるのに対し、迫真性は「対象」の本物らしさを表し前景的要素と強く結びついた感性指標であることが報告されており、物理パラメータの変化に対するそれぞれの感性指標の特性も異なることが示されている ([4], [19] など)。本研究では、臨場感と迫真性の評価尺度を「全くない状態」から「日常生活で経験する最大」の範囲であると定義したものの、得られた数値は両者では異なる意味を示すものであり、両者の特性を比較する際に数値自体を直接比較することはできないと考えられる。このような理由から、z 得点を用いて SOA に対する印象変化傾向を調べた。

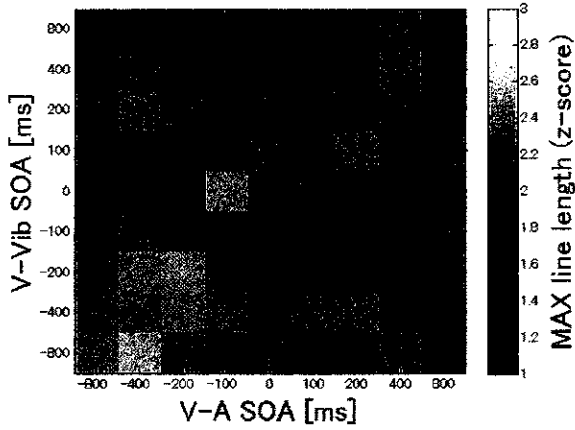
体験中における臨場感および迫真性の印象強度の最



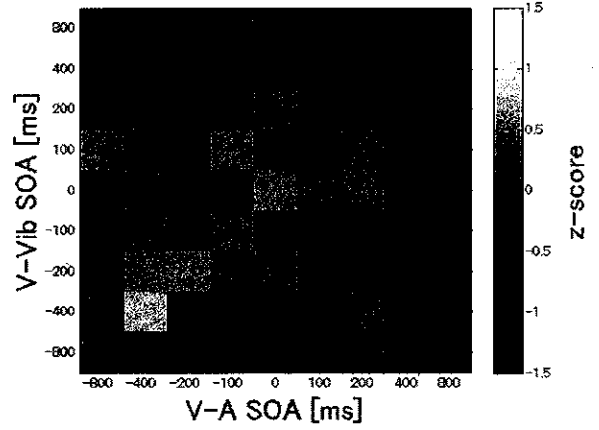
(a) 臨場感



(a) 臨場感



(b) 迫真性



(b) 迫真性

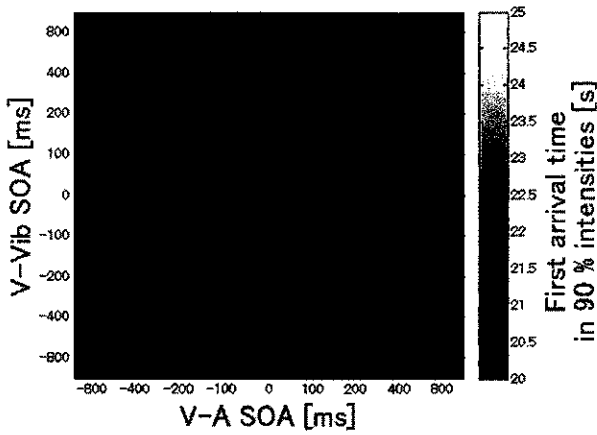
図5 体験中の最大印象強度に SOA 条件が与える影響 (z 得点)

図6 コンテンツ全体を体験した後の印象強度 (z 得点)

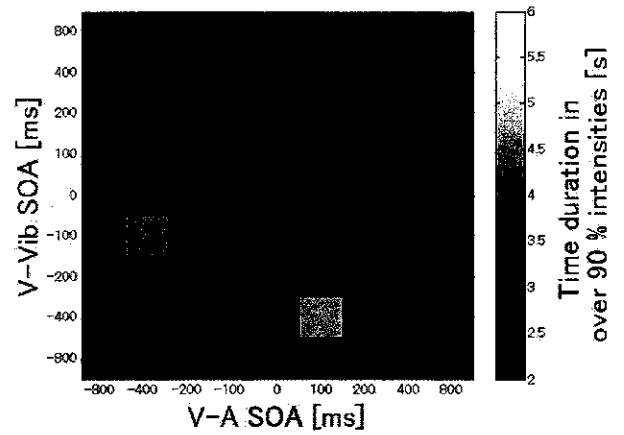
大値に SOA 条件が与える影響を図5に示す。最大値について、感性印象 (2 条件) ×映像-音間 SOA 条件 (9 条件) ×映像-振動間 SOA 条件 (9 条件) の3要因分散分析を行った。その結果、2 次の交互作用および1 次の交互作用は認められなかった (2 次の交互作用: $F_{64,896} = 0.900, n.s.$; 1 次の交互作用: 感性印象 ×映像-音間 SOA 条件 $F_{8,112} = 0.699, n.s.$, 感性印象 ×映像-振動間 SOA 条件 $F_{8,112} = 1.099, n.s.$, 映像-音間 SOA 条件 ×映像-振動間 SOA 条件 $F_{64,896} = 1.080, n.s.$)。一方で、映像-音間 SOA 条件の主効果 ($F_{8,112} = 2.170, p < .05$)、映像-振動間 SOA 条件の主効果 ($F_{8,112} = 3.157, p < .001$) が有意であった。そこで、映像-音間 SOA 条件の多重比較 (Ryan 法, $p < .05$) を行った結果、0 ms 条件では +800 ms よりも有意に印象強度が大きくなること が認められた。また、映像-振動間 SOA 条件の多重比較を行った結果、-400 および -200 ms 条件では +800 ms 条件よりも有意に印象強度が大となった。以上から、臨場感と迫真性ともに、音情報の同期ずれおよび振動情報の同期

ずれが、体験中の印象強度の最大値に影響を及ぼすことが示唆された。

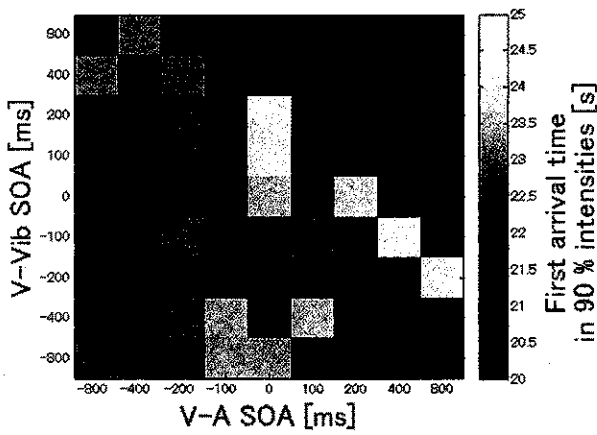
コンテンツ全体を体験した後の印象強度について、臨場感と迫真性の比較検討を行うため評定尺度法により得られた評価値を用いて観測者ごとに z 得点を算出し、感性印象のグループごとに観測者の平均値を算出した。結果を図6に示す。感性印象 (2 条件) ×映像-音間 SOA 条件 (9 条件) ×映像-振動間 SOA 条件 (9 条件) の3要因分散分析を行った。その結果、2 次の交互作用が有意であった ($F_{64,896} = 1.454, p < .05$)。感性印象に注目して2 次の交互作用を分析した結果、臨場感では映像-音間 SOA 条件と映像-振動間 SOA 条件の単純交互作用が有意であった ($F_{64,896} = 1.713, p < .001$)。さらに、臨場感における映像-音間 SOA 条件の単純主効果が有意となる映像-振動間 SOA 条件は、-400 ms のとき ($F_{8,1008} = 1.984, p < .05$)、-100 ms のとき ($F_{8,1008} = 3.786, p < .001$)、および 200 ms のとき ($F_{8,1008} = 2.604, p < .01$) であった。さらに、臨場感における映像-振動間 SOA 条件



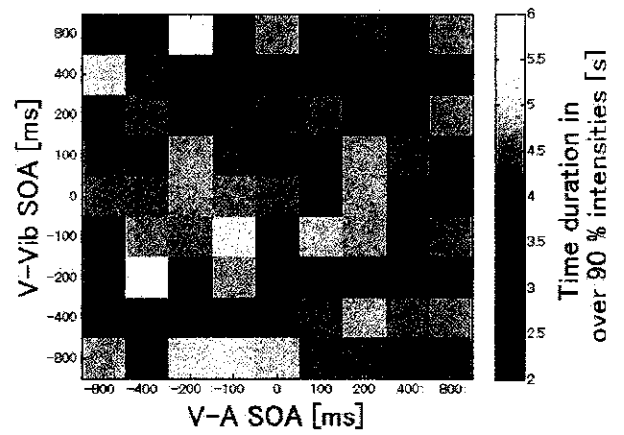
(a) 臨場感



(a) 臨場感



(b) 迫真性



(b) 迫真性

図7 「高印象強度区間」に最初に到達するまでの時間

図8 「高印象強度区間」の時間幅

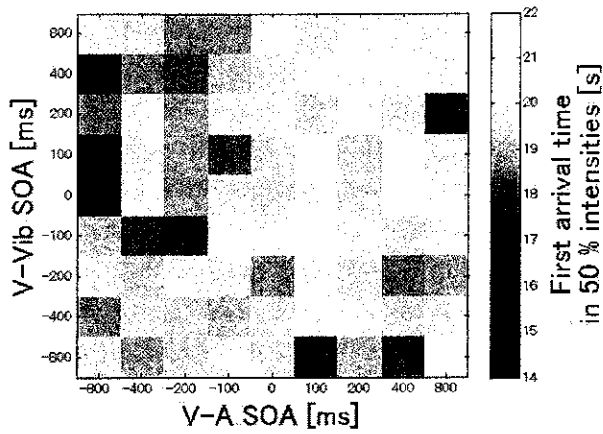
の単純単純主効果が有意となる映像-音間 SOA 条件は、 -800 ms のとき ($F_{8,1008} = 2.267, p < .05$), 0 ms のとき ($F_{8,1008} = 3.162, p < .005$), および 200 ms のとき ($F_{8,504} = 3.989, p < .001$) であった。一方、迫真性では、映像-音間 SOA 条件と映像-振動間 SOA 条件の単純交互作用は有意ではなかった ($F_{64,896} = 1.037, n.s.$)。以上まとめると、コンテンツ全体を体験した後の印象強度は、臨場感の場合、音情報の同期ずれと振動情報の同期ずれの影響を受け、一方、迫真性の場合、音情報の同期ずれと振動情報の同期ずれによる影響は認められなかった。

体験中と体験後の印象強度の対応関係について検討するため、体験中の最大印象強度の z 得点と体験後の印象強度の z 得点との相関を調べた。その結果、臨場感ではやや強い正の相関が認められた ($r = 0.52, p < .01$)。一方、迫真性では弱い正の相関が認められた ($r = 0.22, p < .01$)。

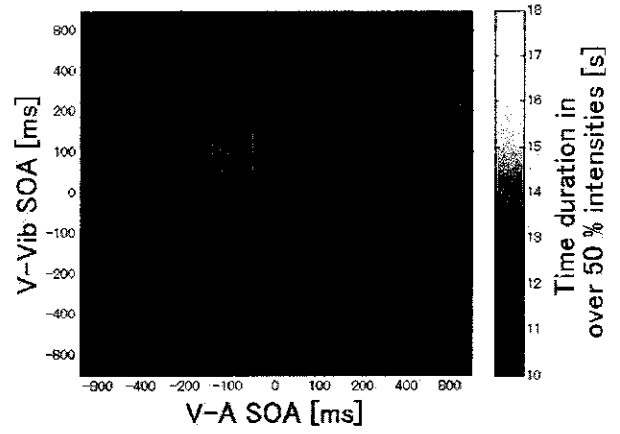
3.2 高い臨場感と迫真性が得られた時間幅

本研究では、印象強度の継時変化と、SOA が継時変化に与える影響を明らかにするために、各印象強度の最大値のみではなく、刺激全体に渡って比較的高い印象強度を維持する時間帯、および、SOA とこの時間帯の関係を分析対象とした。そこで、全区間における最大印象強度の 90 パーセント以上を「高印象強度区間」と定義した。また、図 4 の実験結果から、電車が画面内を通過した後においても印象強度が緩やかに落ち、最大印象強度の半分程度を維持する時間帯 (30~35 s) が見られることから、この区間も対象と含まれる最大印象強度の 50 パーセント以上を「中程度印象強度区間」と定義した。この定義に基づき、最初に 90 パーセントおよび 50 パーセントの印象強度区間に到達するまでの時間と、全刺激に渡ってこれら印象強度が維持する時間幅を算出した。

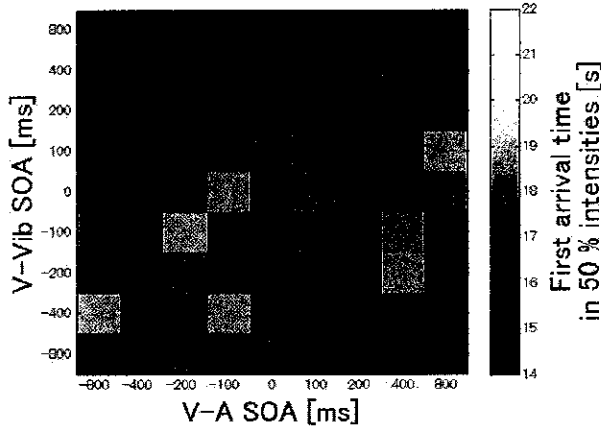
ところで、図 2 に示したように、本研究で用いた実験刺激には、警笛、列車の通過などの複数のイベントが含まれている。印象強度の測定の際に、刺激全体をひとつの流れとして捉えて評価した観測者が大多数で



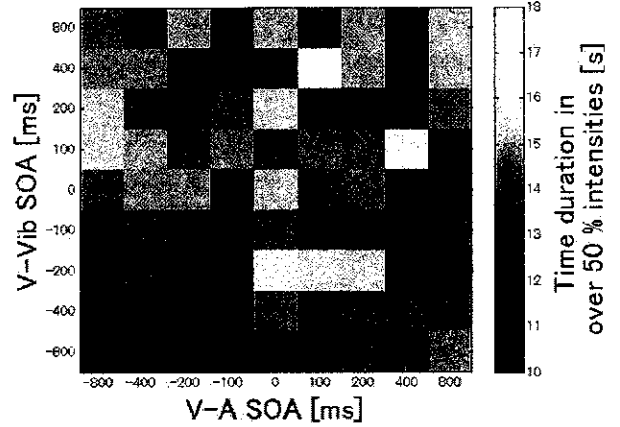
(a) 臨場感



(a) 臨場感



(b) 迫真性



(b) 迫真性

図9 「中程度印象強度区間」に最初に到達するまでの時間

図10 「中程度印象強度区間」の時間幅

あった一方で、実験刺激に含まれる個々のイベントに着目して、それぞれのイベントで臨場感、迫真性を評価した観測者も見受けられた。後者の場合、個々のイベントで印象強度が変化するため、複数の山型形状を示す結果となっていた。そこで、各印象強度区間の時間幅を算出する際には、全時間長(0~35 s)における印象強度に関し、90パーセント以上、もしくは50パーセント以上となった時間の和を分析対象となる時間幅とした。

まず「高印象強度区間」の算出結果について述べる。高印象強度区間に最初に到達するまでの観測者平均を臨場感と迫真性それぞれについて図7に示す。最初に到達するまでの時間とは、0 s から時間経過方向に印象強度を確認し、最大値の90パーセント値を初めに確認するまでの時間幅である。最初に到達するまでの時間は臨場感において平均22.21 s、標準偏差1.08 s、迫真性において平均22.05 s、標準偏差3.11 sであった。感性印象(2条件)×映像-音間SOA条件(9条件)×映像-振動間SOA条件(9条件)

の3要因分散分析を行ったところ、2次の交互作用および1次の交互作用は認められなかった(2次の交互作用: $F_{64,896} = 1.027, n.s.$; 1次の交互作用: 感性印象×映像-音間SOA条件 $F_{8,112} = 0.814, n.s.$, 感性印象×映像-振動間SOA条件 $F_{8,112} = 0.199, n.s.$, 映像-音間SOA条件×映像-振動間SOA条件 $F_{64,896} = 0.974, n.s.$)。一方で、映像-音間SOA条件の主効果が有意であった($F_{8,112} = 3.286, p < .005$)。映像-音間SOA条件において多重比較(Ryan法, $p < .05$)を行った結果、-800 ms条件よりも0, +100, +200, +800 msで有意に遅くなることが認められた。また、映像-音間SOAとその条件における到達時間の関係について相関を求めたが、臨場感と迫真性ともに有意な相関は認められなかった(臨場感: $r = 0.03, n.s.$; 迫真性: $r = 0.02, n.s.$)。

「高印象強度区間」の時間幅を、臨場感と迫真性ごとに図8に示す。時間幅は臨場感において平均3.26 s、標準偏差2.16 s、迫真性において平均4.27 s、標準偏差3.58 sであった。3要因分散分析を行ったところ、交互作用および主効果はいずれも有意にはならなかった。

また、映像-音間 SOA とその条件における時間幅の関係について相関を求めたが、臨場感と迫真性ともに有意な相関は認められなかった（臨場感： $r = 0.10, n.s.$ ；迫真性： $r = 0.04, n.s.$ ）。

次に、「中程度印象強度区間」の結果を述べる。中程度印象強度区間に最初に到達するまでの観測者平均を臨場感と迫真性それぞれについて図9に示す。最初に到達するまでの時間は臨場感において平均 19.46 s、標準偏差 2.919 s、迫真性において平均 17.19 s、標準偏差 6.01 s であった。感性印象（2 条件）×映像-音間 SOA 条件（9 条件）×映像-振動間 SOA 条件（9 条件）の 3 要因分散分析を行ったところ、交互作用および主効果はいずれも有意にはならなかった。また、映像-音間 SOA とその条件における到達時間の関係について相関を求めたが、臨場感と迫真性ともに有意な相関は認められなかった（臨場感： $r = 0.07, n.s.$ ；迫真性： $r = 0.01, n.s.$ ）。

「中程度印象強度区間」の時間幅を、臨場感と迫真性ごとに図 10 に示す。時間幅は臨場感において平均 11.62 s、標準偏差 4.54 s、迫真性において平均 14.08 s、標準偏差 7.39 s であった。3 要因分散分析を行ったところ、2 次の交互作用および 1 次の交互作用は認められなかった（2 次の交互作用： $F_{64,896} = 0.837, n.s.$ ；1 次の交互作用：感性印象×映像-音間 SOA 条件 $F_{8,112} = 1.031, n.s.$ 、感性印象×映像-振動間 SOA 条件 $F_{8,112} = 0.464, n.s.$ 、映像-音間 SOA 条件×映像-振動間 SOA 条件 $F_{64,896} = 0.734, n.s.$ ）。一方で、映像-振動間 SOA 条件の主効果が有意であった（ $F_{8,112} = 2.049, p < .05$ ）。多重比較（Ryan 法、 $p < .05$ ）を行った結果、全ての条件で有意差は認められなかった。また、映像-音間 SOA とその条件における時間幅の関係について相関を求めたが、臨場感と迫真性ともに有意な相関は認められなかった（臨場感： $r = 0.01, n.s.$ ；迫真性： $r = 0.02, n.s.$ ）。

以上の結果をまとめる。「高印象強度区間」に関しては、映像-音間 SOA 条件による影響が示された。映像-振動間 SOA 条件による影響は、検討した全ての時間で認められなかった。また、「中程度印象強度区間」では、映像-振動間 SOA 条件による影響が示された。映像-音間 SOA 条件による影響は、検討した全ての時間で認められなかった。

4 考察

4.1 印象強度の継時変化

本実験から、臨場感、迫真性ともに、印象強度が実験刺激にあわせて継時的に変化することが確認できた。また、特に 30 s から 35 s までの継時変化に代表されるように、臨場感と迫真性では継時変化の特性が異なる

ことも確認された。30 s から 35 s までの時間帯で列車通過後も迫真性の印象強度が比較的保持されたのは、観測者が聞こえてくる警笛を前景要素の情報として受け取り、その結果、前景要素に密接に関係する迫真性が生起されたためと考えられる。一方で、臨場感は感覚情報の量的増加に密接に関係するため、通過時よりも音や振動情報が小さくなる列車通過後では印象強度が比較的速やかに減少したと考えられる。また、列車通過後も意識の中に存在する補完された列車の心像が、列車通過後の音や振動情報に触発された結果、臨場感および迫真性に反映された可能性が考えられる。

体験中と体験後の印象強度の対応関係について検討するため、体験中の最大印象強度と体験後の印象強度の相関について調べた。その結果、臨場感と迫真性ともに正の相関が認められた。Ozawa ら [14] は、カテゴリ連続判断法 [20] を用いて臨場感の印象強度を測定した。10 パーセント 時間率臨場感値 S_{10} （評価結果が存在する時間帯のうち 10 パーセントの時間で S_{10} 以上の値をもつような値）を指標として、コンテンツ全体に対する臨場感評価値との関係を分析したところ、正の相関関係（ $r = 0.96$ ）があったことを報告している。本実験の結果では、臨場感だけでなく迫真性についても有意な正の相関が認められた。しかし、Ozawa らの結果のような高い相関係数は得られなかった（臨場感： $r = 0.52, p < .01$ 、迫真性： $r = 0.22, p < .01$ ）。これは Ozawa らは音コンテンツについて評価したのに対し、本実験では映像、音、振動情報を含む多感覚コンテンツについて評価したためであり、含まれる感覚情報が増加したことによる影響が現れたと考えられる。以上から、臨場感のみならず迫真性に関しても、体験中と体験後の結果に何らかの関連があるものと考えられる。しかし、その詳細については不明であるため今後更なる分析を行う必要があると考えている。

4.2 高い臨場感と迫真性が得られた時間幅

高い臨場感および迫真性が得られる時間幅について検討した結果、臨場感において印象強度が最大値の 90 パーセント値に最初に到達するまでの時間は映像-音間 SOA 条件に影響されることが示された。一方で、高印象強度時間帯では、SOA 条件による影響は認められなかった。以上の結果を総合的に考察すると、高い臨場感と迫真性が得られる時間幅は映像に対する音の同期ずれに影響を受けると考えられる。さらにこの変化において、映像-振動間 SOA 条件の影響は認められず、映像-音間 SOA 条件のみが支配的に働く。このことは、体験中で高い臨場感を得られる時間については、振動情報は音情報ほどは影響しないことを示唆している。

5 まとめ

振動情報よりも音情報の方が強く影響する理由について以下のように考察する。寺本ら [4] は臨場感創出に対する感覚モダリティの重要度は上位から聴覚、視覚、平衡感覚と報告している。このことから、臨場感の重要度の高い視覚および聴覚情報に強く依存しており、そのため平衡感覚に関係する振動情報よりも聴覚情報（音情報）と映像情報との間の同期ずれに強く影響されたのではないかと考えられる。

先にも考察した通り、今回の実験では振動情報の時間的なずれが、臨場感、迫真性の継時変化や、視聴後の判断に影響を及ぼすとは言えない結果となった。この理由としては振動情報の変化に対して臨場感と迫真性は鈍感であるからと考えられる。高橋ら [11] は、視聴覚情報が同期された状態で振動情報の SOA 条件を変えた場合、 $-400 \sim +200$ ms の範囲で印象強度が大きくなると報告している。一方で、振動情報を含まない視聴覚情報間での SOA 条件を変えた場合、寺本ら [7] や本多ら [19] は、 $0 \sim +150$ ms の範囲で大きくなると報告している。これらの結果は、振動の同期ずれは音の同期ずれよりも許容範囲が広く、振動情報の変化に対して臨場感と迫真性は鈍感であることを示唆している。また他には、車両の車軸間は前後で等間隔であり、電車に起因する振動は周期的となるため同期ずれの知覚が難しいことから、振動情報による影響が認められなかった可能性も考えられる。

一方で、「中程度印象強度区間」に関しては、時間幅のみ映像-振動間 SOA によって変化する結果が得られ、映像-音間 SOA は影響を及ぼさないことが示唆された。これは、振動情報は音情報と異なり、刺激の全時間において継続的に存在するため、ある程度の印象強度を得るための基盤として作用して、SOA が影響を及ぼしたものと考えられる。このことは、高い印象強度を得るためのイベントの評価自体には振動情報が寄与しないともいえる結果であり、映像、音、振動の各感覚情報が、感性評価においてどのような役割を果たすのかについて、今後研究を進めていく必要があると考える。

なお、本実験では観測者の住所と電車の走行する環境との距離は統制できていなかったため、電車に対する馴染みの個人差による影響は不明であった。しかし、日々の生活の中で電車に関する経験の多い観測者であれば、今回用いたような電車が通過するコンテンツに対して、経験の少ない観測者とは異なる傾向の臨場感および迫真性を生起する可能性がある。このような、観測者の持つ経験や事前知識がトップダウンとして臨場感、迫真性の知覚にどのような影響を及ぼすかは興味深いテーマであり、これについては、今後新たに検討を行う必要があると考えている。

本研究では、視覚、聴覚、全身振動情報で構成される多感覚情報間の同期ずれを操作して、実験刺激を体験中に生起する臨場感および迫真性の印象強度の変化および各感覚情報の寄与について検討した。

電車が接近する視聴覚刺激と合わせて全身振動情報を提示し、刺激提示中の臨場感と迫真性を測定した結果、臨場感、迫真性ともに印象強度が継時変化することを確認した。さらに、高い臨場感および迫真性が得られる時間幅では、映像に対する音の同期ずれに影響を受けるが、振動の影響は認められなかった。以上から、音および振動情報は臨場感、迫真性の継時変化に影響を及ぼすことが示唆された。

謝辞

本研究の一部は、それぞれ、NICT 委託研究「革新的な三次元映像技術による超臨場感コミュニケーション技術の研究開発」および文部科学省科研費基盤研究 A (課題番号: 24240016)、日中韓フォーサイト事業「次世代ネットワークにおける超臨場感音響相互通信の実現」、東北大学電気通信研究所の共同プロジェクト研究、文科省特別経費「人間の判断の実現に向けた新概念脳型 LSI 創出事業」による。

参考文献

- [1] 榎並和雅, 岸野文郎: 今後の超臨場感にかかわる研究はどこを目指すべきか, 電子情報通信学会誌, Vol. 93, No. 5, pp. 363-367, 2010.
- [2] M.Lombard, T.Ditton: At the heart of it all: The concept of presence, Journal of Computer - Mediated Communication, Vol. 3, No. 2, 1997.
- [3] M.Slater: Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments, Phil. Trans. R. Soc. B, Vol. 364, pp. 3549-3557, 2009.
- [4] 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論文誌, Vol. 15, No. 1, pp. 7-16, 2010.
- [5] 福江一智, 小澤賢司, 木下雄一朗: 視聴覚コンテンツ臨場感の多次元性に関する検討, 日本感性工学会論文誌, Vol.11, No. 2, pp. 183-192, 2012.
- [6] 正岡顕一郎, 江本正喜, 菅原正幸, 岡野文男: 広視野・超高精細平面ディスプレイの画角と臨場感の関係, 映像情報メディア学会技術報告, Vol. 28, No. 31, pp. 17-20, 2004.
- [7] 寺本渉, 吉田和博, 浅井暢子, 日高聡太, 行場次朗, 鈴木陽一: 「迫真性」を規定する時空間情報, 日本バーチャルリアリティ学会論文誌, Vol. 15, No. 5, pp. 483-486, 2010.
- [8] 神田敬幸, 本多明生, 柴田寛, 浅井暢子, 寺本渉, 坂本修一, 岩谷幸雄, 行場次朗, 鈴木陽一: 視聴覚コンテンツの臨場感と迫真性に関する実験的研究~シンバル演奏場面における視野制限と音圧レベルの効果~, 電子情報通信学会技術研究報告, HIP, ヒューマン情報処理, Vol. 111, No.283, pp. 7-10, 2011.

- [9] K.Ozawa, M.Miyasaka : Effects of reproduced sound pressure levels on auditory presence, *Acoustical science and technology*, Vol. 25, No. 3, pp. 207-209, 2004.
- [10] 行場次朗: 図と地の知覚-視覚の心理-, 電子情報通信学会誌, pp. 315-320, 1991.
- [11] 高橋恵美, 坂本修一, 鈴木陽一, 行場次朗: 多感覚コンテンツの高次感性知覚に影響を与える振動情報の強度ならびに時間特性, 日本バーチャルリアリティ学会第 18 回大会論文集, pp.92-95, 2013.
- [12] 本多明生, 神田敬幸, 柴田寛, 坂本修一, 岩谷幸雄, 行場次朗, 鈴木陽一: 視聴者が感じるスポーツコンテンツの臨場感と迫真性は再生速度によって変化する, 日本バーチャルリアリティ学会第 17 回大会論文集, pp.499-502, 2012.
- [13] S. Sakamoto, G. Hasegawa, A. Honda, Y. Iwaya, Y. Suzuki, J. Gyoba : Body vibration effects on perceived reality with multi-modal contents, *ITE Transactions on Media Technology and Applications*, Vol. 2, No. 1, pp. 46-50, 2014.
- [14] K. Ozawa, S. Tsukahara, Y. Kinoshita, M. Morise : Instantaneous evaluation of the sense of presence in audio-visual content, *IEICE Transactions on Information and Systems*, Vol. E98-D, No. 1, pp. 49-57, 2015.
- [15] 大出訓史, 安藤彰男, 谷口高士: 聴覚的臨場感に関する階層的印象の時間的変化, 日本音響学会講演論文集(秋), pp. 887-888, 2013.
- [16] M. Meehan, B. Insko, M. Whitton, F. P. Brooks. Jr : Physiological measures of presence in stressful virtual environments, *ACM Transactions on Graphics*, Vol. 21. No. 3, pp. 645-652, 2002.
- [17] M. Slater, P. Khanna, J. Mortensen, I. Yu, : Visual Realism Enhances Realistic Response in an Immersive Virtual Environment, *Computer Graphics and Applications*, IEEE, Vol. 29, 3, pp. 76-84, 2009.
- [18] T. Kato, S. Namba, S. Kuwano : Continuous judgment of level-fluctuating noise, *Proc. INTER-NOISE 94*, pp. 1081-1084, 1994.
- [19] 本多明生, 神田敬幸, 柴田寛, 浅井暢子, 寺本涉, 坂本修一, 行場次朗, 鈴木陽一: 視聴覚コンテンツの臨場感と迫真性の規定因, 日本バーチャルリアリティ学会論文誌, Vol. 18, No. 1, pp. 93-101, 2013.
- [20] S. Namba, S. Kuwano : The relation between overall noisiness and instantaneous judgment of noise and the effect of background noise level on noisiness, *Journal of the Acoustical Society of Japan (E)*, Vol. 1, pp. 99-106, 1980.

(2015年3月9日受付)

[著者紹介]

柳生 寛幸 (学生会員)



2009年山形大学工学部電気電子工学科卒業。2011年東北大学大学院情報科学研究科博士課程前期課程修了。運輸会社勤務を経て、現在、東北大学大学院情報科学研究科博士課程後期課程在学中。マルチモーダル知覚の研究に従事。

崔 正烈 (正会員)



2004年九州大学大学院システム情報科学府博士課程修了。同年、同大学 COE 研究員・ユーザーサイエンス機構研究員を経て、2009年より東北大学電気通信研究所研究員、現在に至る。人工現実感環境生成法、人の感性・生理・心理特性に関する研究に従事。ヒューマンインターフェース学会、バーチャルリアリティ学会等の会員、博士(工学)。

坂本 修一 (正会員)



1997年東北大学大学院情報科学研究科修士課程修了。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)。単語知覚過程に関する研究に従事するほか、近年ではマルチモーダル感覚情報処理過程に関する研究にも注力。日本音響学会、電子通信情報学会、日本バーチャルリアリティ学会、日本感性工学会等の会員。

大谷 智子 (正会員)



2004年、聖心女子大学大学院文学研究科人間科学専攻博士後期課程満期退学。東京大学大学院情報学環特任助教、東北大学電気通信研究所助教を経て、2015年度から、東京藝術大学芸術情報センター助教。人間の色知覚に関する研究や、インタフェースやデジタルコンテンツの心理学的評価を行うとともに、当該分野の一般社会への理解増進事業に従事。博士(心理学)。

鈴木 陽一 (正会員)



1981年東北大学大学院工学研究科修了。工博。東北大学大型計算機センター助教等を経て、1999年から東北大学電気通信研究所教授、現在に至る。2011年日本バーチャルリアリティ学会フェロー、理事。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、理事。米音響学会フェロー。著書に聴覚モデル(コロナ社)など。

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。