

## 3次元音場再現システムでの コンテンツ選択における指針の提案

小林 まおり<sup>\*1,2</sup> 福井 誠人<sup>\*2,3</sup> 上野 佳奈子<sup>\*2,3</sup>

### Classification of Reproduced Sound Contents for 3-D Reproduction Sound Systems

Maori Kobayashi<sup>\*1,2</sup>, Masato Fukui<sup>\*2,3</sup> and Kanako Ueno<sup>\*2,3</sup>

**Abstract ---** It has been reported that the “sense of presence” is affected by the reproduced sound contents in auditory virtual environments. The selection of the reproduced sound contents is important for the evaluation of the auditory virtual systems appropriately. This paper describes our trials to classify the sound contents depending on experienced realities in a 3-D reproduction sound system. After participants listened to the sound contents, they were asked to evaluate the contents by twelve evaluation points. Principle component analysis and cluster analysis revealed that sound reproduced contents were classified into three categories: music, voices, and moving sound objects. These categories are linked to different elements of sense of presence. From these results, we proposed that three kinds of the reproduced contents are needed for the evaluation of auditory virtual systems.

**Keywords:** auditory 3-D reproduction system, auditory virtual environment, sound recorded contents, sense of presence, subjective evaluation

#### 1 はじめに

音響技術の発展により、高精細かつ立体化された新たな音情報提示システムの試みがなされている[1-7]。これらの音情報システムでは、聴取者があたかもその場に存在するかのような感覚までもを提供することが期待されている。今後、システムの性能を評価するうえで、「あたかもその場にいるような感覚」、いわゆる「臨場感」が重要な評価要素になることは間違いない。そのための評価指標・計測技術の確立や認知メカニズムの解明が求められている[8]。

従来の研究では臨場感の創出要因について数多く検討されており[9-14]、要因の1つに再生コンテンツの影響が報告されている[11, 15]。たとえば小澤らはバイノーラルシステムを用い、オーディオシステムの再現精度は同じであっても、対象となる音源によって臨場感の主観的な評価値が異なることを示した[15,16]。彼らの研究では、コンテンツによる臨場感の違いは最も大きいもので、ダミーヘッドの外耳特性を補正したバイノーラルシステムとモノラルシステムにおける臨場感の違いにも匹敵するほどであったという。これらはシステムの性能として臨場感の評価を行う場合に、使用するコンテンツの選

択に注意が必要であることを示唆する。

また臨場感は複数の感覚要素の複合体と捉えられている[17]。そのため、たとえ「臨場感」に対する評価値が同等だったとしても、その評価に影響した感覚要素はコンテンツによって異なる可能性がある。コンテンツによって臨場感の感覚要素が異なるならば、システム性能として臨場感を適切に評価するために数種類のコンテンツを用意する必要があるだろう。しかし、どのようなコンテンツがどのような臨場感をもたらすのか、コンテンツを検証した研究は少なく、コンテンツを選択するうえで指標となる手がかりも得られていない。実際にコンテンツを制作、あるいは選択する立場から考えると、音楽や音声といった録音対象や音源数、音源配置など客観的に観察可能な特性による感覚要素の予測は有益である。

そこで本研究では、コンテンツ選択の指針構築を目的とし、その第1段階として、経験される臨場感の感覚要素によるコンテンツの分類を試みた。実験では主成分分析を用いて臨場感を規定する成分を抽出し、その主成分得点を用いてクラスター分析を行い、コンテンツを分類した。各クラスターに分類されたコンテンツの共通点を探ることで、コンテンツの特徴と臨場感の感覚要素との対応関係を探索的に検討した。もし、コンテンツと感覚要素に何らかの対応関係が見られるならば、最終的には各クラスターにおけるコンテンツ特徴を1つの指針として、コンテンツ選択に適用できるのではないかと考えた。

加えてコンテンツ選択を行ううえで、高精細かつ立体

\*1 明治大学 研究知財戦略機構

\*2 CREST, 科学技術振興機構

\*3 明治大学 理工学部

\*1 Organization for the Strategic Coordination of Research and Intellectual Properties, Meiji University

\*2 CREST, Japan Science & Technology Agency

\*3 School of Science and Technology, Meiji University

化された音情報提示システムを使用して、どのような音を聞きたいか、といった視点も必要である。現在、普及しているオーディオシステムは映画や音楽などのエンターテインメント作品の再生が主要な目的であり、「娯楽性」もオーディオシステムに対する社会的な要請となっている。そのため「聞いていて楽しい」や「聴く価値がある」といった評価によるコンテンツの分類および選択についても検討した。これらの検討を踏まえ、本研究では3次元音場再現システムにおいてコンテンツを選択するうえでの指針を提案する。

なお、臨場感は多次元的な概念であり、受け手によって意味が異なるため、事前に臨場感という言葉を選定した上で評価者に呈示することが求められる[8]。そこで本論文では、臨場感を「あたかもその場にいるような感覚」と定義し、実験を行った。また、コンテンツによって臨場感が異なることを示した上述の研究[16]においても、「臨場感」という語を定義せずに使用しており、「あたかもその場にいる感じ」という評価がコンテンツによって異なるのか、実際のところは不明である。そこで本研究では「その場にいる感じ」という評価がコンテンツによって左右されるのかについても併せて確認した。

## 2 実験

### 2.1 装置

本研究では、境界音場制御の原理[6]に基づく音場再現システム(以下、BoSC: Boundary Surface Control System[18])を用い、音刺激を作成し提示した。この音場再現システムでは、(1) 原音場内の境界  $S$  上で信号を測定する。(2) 次に再生音場の伝達関数を測定し、

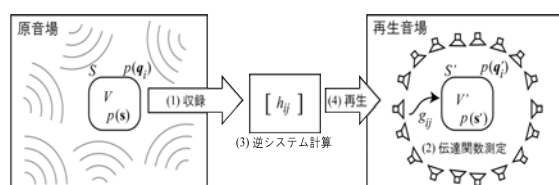


図1 境界音場制御の原理

Fig. 1 The Boundary Surface Control system

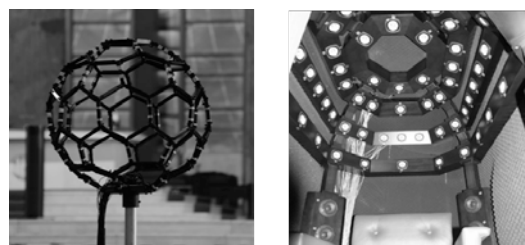


図2 システム構成, 左: BoSC 収録システム, 右: BoSC 再生システム

Fig.2 Configuration of the system, Left: the BoSC recording system, Right: the BoSC reproducing system

(3) その伝達関数の逆システムを(1)で測定した信号に畳み込む。そして(4) 再生音場の境界  $S'$  で信号の音圧を原音場  $S$  の音圧と同等になるように操作することで、原音場における領域  $V$  内の音場を再生音場の  $V'$  において再現するというものである(図1)。

実験で用いた BoSC システムは、原音場を収録するための収録システムと、収録された音場を再現するための再生システムから構成された。収録システムは、無指向性マイクロホン(DPA 4060bm)を C80Fullerene 分子構造の各頂点に配置した 80ch. マイクロホン・アレー

(直径 46 cm, 素子間隔 8 cm)とした(図2)。再現システムは 1.5 畳ほどの防音室(YAMAHA Woody Box, 遮音性能 D-30, 内部に吸音材 30 mm を装備)に、スピーカ・アレイを設置したものであった(残響時間 0.07 s, システム駆動時の暗騒音レベル 25.5 dB(A))。スピーカ・アレイは、4 層の木製エンクロージャから構成され、各層には上から 6, 16, 24, 16ch のフルレンジスピーカ(FOSTEX FE83)を配置した(スピーカ配置の詳細は文献[19])。本システムでは聴取者の頭部周辺の音場の収録・再現を想定しており、音場再現ルーム内のリフトを用いて、聴取者の頭部を再現領域まで上昇させ、音刺激を呈示した。

### 2.2 刺激

BoSC 収録システムで収録した音に、再生音場の伝達関数の逆システムを畳み込んだ 20 種類の音を刺激として用いた(表1)。コンテンツは音臨場感や環境音の先行研究[16, 20]で用いられた音源と類似している状況と考えられる 11 コンテンツを選択した(表1, \*1)。また、これまでの BoSC システムのデモンストレーションでの評価に基づき、3次元音場再現システムの特性を生かすと考えられる 9 コンテンツを加えた。各刺激の時間長は 1min とし、音圧レベルは原音場と近い値になるように再生システム内の頭部中心位置でそれぞれの刺激で設定した。

### 2.3 手続き

**被験者** 聴覚健全な成人 18 名(年齢 20-23 歳, うち女性 9 名)が実験に参加した。

**評価項目** 聴覚臨場感を因子分析した研究に小澤らによる一連の研究[15,16,21]があるが、彼らの研究では臨場感を定義せずに評価語を選出する実験を行っているため、選出された項目が「その場にいる感じ」ではなく、「心揺さぶられる」という印象を評価する項目であった可能性がある。加えて個人補正を行っていないバイノーラルシステムのみが用いられている。この場合、聴取者の頭部や耳の形状がコンテンツ収録用のダミーヘッドと異なると音場が忠実に再現されないため、臨場感の評価要素が本研究で使用

表 1 実験に用いた音刺激

Table 1 Auditory stimuli used in the experiment

|                 | 主な音                                        | 再生音圧<br>レベル<br>(dB L <sub>Aeq</sub> ) | 録音場所          | マイクの位置 (音<br>源との位置関係)   | マイクと<br>音源の距<br>離 (m) | 録音対<br>象の音<br>源数 | 音源の<br>移動 | 人音声<br>の有無 | 楽器<br>音・歌<br>声の有<br>無 |
|-----------------|--------------------------------------------|---------------------------------------|---------------|-------------------------|-----------------------|------------------|-----------|------------|-----------------------|
| A               | オーケストラの演奏                                  | 78.5                                  | コンサート<br>ホール  | ステージ上<br>(指揮者の前)        | 3                     | 30 以上            | 無         | 無          | 有                     |
| B               | オーケストラの演奏 <sup>*2</sup>                    | 79.5                                  | コンサート<br>ホール  | 客席                      | 13                    | 30 以上            | 無         | 無          | 有                     |
| C               | ギターのメロディにのっ<br>て拍手をしながら人が<br>移動する音         | 73.2                                  | 多目的<br>ホール    | 移動軌跡の中心                 | 1 (最接<br>近時)          | 7                | 有         | 無          | 有                     |
| D               | 能 <sup>*2</sup>                            | 75.9                                  | 能舞台           | 舞台下                     | 5                     | 14               | 有         | 有          | 有                     |
| E <sup>*1</sup> | バイオリン ソロ演奏                                 | 78.0                                  | コンサート<br>ホール  | 客席                      | 13                    | 1                | 無         | 無          | 有                     |
| F               | バンド演奏 (歌入り,<br>邦楽)                         | 69.0                                  | 多目的<br>ホール    | 客席                      | 7                     | 4                | 無         | 有          | 有                     |
| G               | バンド演奏 (歌入り,<br>洋楽)                         | 74.7                                  | 多目的<br>ホール    | 客席                      | 7                     | 7                | 無         | 有          | 有                     |
| H <sup>*1</sup> | パイプオルガン演奏 <sup>*2</sup>                    | 84.6                                  | 教会            | 客席                      | 22                    | 1                | 無         | 無          | 有                     |
| I               | 合唱 (カエルの歌の輪<br>唱)                          | 71.7                                  | 多目的<br>ホール    | 輪状に位置した歌<br>い手の中心       | 2                     | 8                | 無         | 有          | 有                     |
| J <sup>*1</sup> | 会話音声                                       | 64.8                                  | オフィス          | 音源脇, 正面                 | 0.6                   | 3                | 無         | 有          | 無                     |
| K <sup>*1</sup> | 自然音 (カラスの鳴き<br>声, 飛行機の音など)                 | 65.1                                  | 屋外 (庭)        | 地面から1~2 m <sup>*3</sup> | *4                    |                  | 有         | 無          | 無                     |
| L               | 会話音声<br>(歩きながらの会話)                         | 53.9                                  | 屋外            | 話者の背後                   | 不明                    | 5                | 有         | 有          | 無                     |
| M <sup>*1</sup> | 会話音声 (ざわめき)                                | 70.0                                  | オフィス          | 音源脇, 正面                 | 0.8~                  | 7                | 無         | 有          | 無                     |
| N <sup>*1</sup> | 蟬の鳴き声                                      | 68.2                                  | 屋外 (大学<br>構内) | 地面から1~2 m               | *4                    |                  | 無         | 無          | 無                     |
| O <sup>*1</sup> | 鈴虫の鳴き声                                     | 62.5                                  | 屋外 (森<br>林)   | 地面から1~2 m <sup>*3</sup> | *4                    |                  | 無         | 無          | 無                     |
| P               | わらべ歌<br>(かごめかごめ)                           | 62.5                                  | 屋外 (庭)        | 輪状に位置した歌<br>い手の中心       | 2                     | 4                | 有         | 有          | 有                     |
| Q <sup>*1</sup> | 商店街の喧騒 (自転車の<br>ベル, アーケードのチャ<br>イム, 車の音など) | 61.9                                  | 街頭            | 地面から約170 cm             | *4                    |                  | 有         | 無          | 無                     |
| R <sup>*1</sup> | 人の走行音                                      | 53.0                                  | 屋外<br>(砂利道)   | 移動軌跡の中心                 | 不明                    | 1                | 有         | 無          | 無                     |
| S <sup>*1</sup> | 人の歩行音                                      | 60.7                                  | オフィス          | 移動軌跡の中心                 | 1 (最接<br>近時)          | 1                | 有         | 無          | 無                     |
| T <sup>*1</sup> | 人の動作音 (PCの操作<br>音, 雑誌のページをめく<br>る音など)      | 53.5                                  | オフィス          | 音源脇, 正面                 | 0.6                   | 6                | 有         | 無          | 無                     |

\*1: 先行研究[15, 26]で用いられたものと類似している状況と考えられるもの

\*2: プロの演奏者・表現者によるコンテンツであることを示す。(無印の場合はアマチュア)

\*3: 収録に使用したマイクがBoSCマイクではないことを示す。詳細は文献[19]を参照。

\*4: 対象音源が複数存在するため、記述不可能なコンテンツであることを示す。

する 3次元音場再現システムと異なることも考えられる。

そこで本実験と同じシステムを使用して予備実験を行い、その結果をもとに評価語を抽することとした。予備実験では評価グリッド法を用いて、被験者に「その場にい

る感じ」の創出要因と考えられる項目を申告させた(付録A-1参照, [22])。このうち先行研究[8, 16]で用いられた形容詞対と対応すると考えられるもののなかから 6 項目を設定した(表 2, \*1)。その他, 出現頻度が高かった 3 項目も加え, 9 項目を設定した(表 2)。総合評価項

表 2 実験に用いた評価項目

Table 2 Evaluation items used in the experiment

| No.      | 質問項目                                    |
|----------|-----------------------------------------|
| Q1       | 聞き慣れた音である (音の情報に関わる因子) *1<br>(日常性因子) *2 |
| Q2       | 人の存在を感じる                                |
| Q3       | 音に距離感を感じる (定位性因子) *1                    |
| Q4       | 音に方向感を感じる (定位性因子) *1                    |
| Q5       | 音に包まれている                                |
| Q6       | 情景がわかる (音の情報に関わる因子) *1                  |
| Q7       | いろいろな音を感じる                              |
| Q8       | リアクションしたくなる                             |
| Q9       | 音に動きを感じる (定位性因子) *1<br>(活動性因子) *2       |
| <総合評価項目> |                                         |
| Q10      | 自分がその場にいるように感じる                         |
| Q11      | 楽しめる (評価性因子) *1,2                       |
| Q12      | 緊張感がある音である (心的負荷因子) *1                  |

\*1: 先行研究[16]で抽出された因子

\*2: 先行研究[21]で抽出された因子

目として、先行研究[16,22]で用いられた評価語のうち 2 項目(「Q11: 楽しめる」「Q12: 緊張感がある音である」)および「Q10: 自分がその場にいるように感じる」を加えた。これらの評価項目群は、3 次元音場再現システムにおける「コンテンツ自体の魅力」を説明するために必要十分な項目と考えた。

**手続き** 12 評価項目についてリッカート法(5 段階)を用いて主観評価実験を行った。被験者に各音源に対して、それぞれの評価項目について、「1: まったく感じない」から「5: とても感じる」までの 5 段階で回答させた。刺激はランダムな順番で呈示し、その際の質問項目の順序も試行毎に変更した上で回答させた。実験に要した時間は休憩も含めて 60 分程度であった。

### 3 結果と考察

各評価項目についてコンテンツごとに全被験者の回答を累積したところ、Q12:「緊張感がある音である」という項目を除いて全ての項目で正規分布に従った。そのため Q1~Q9 に対する回答の分析には平均値を使用し、本来順序尺度であるが、間隔尺度と見なし統計処理を行った。一方、Q10~Q12 の評価項目を総合的に扱う場合には中央値を使用し、分析を行った。

表 3 にコンテンツごとに Q10「自分がその場にいるように感じる」評価値の平均値を算出した結果を示す。全てのコンテンツにおいて平均評価値が 3.2 以上(全コンテンツ平均: 4.01; 標準誤差: 0.05)であった。このことから、今回の実験で用いたコンテンツはおおむね臨場感を感じられるものであったといえる。加えてコンテンツによって「その場にいる感じ」の評価値が変わるのか統計的検定を行ったところ、コンテンツ間に有意差が認められた( $F(19, 323) = 35.72, p < 0.01$ )。これらの結果から、「臨場

表 3 各コンテンツにおける Q10 (自分がその場にいるように感じる)の平均評価値と標準誤差

Table 3 Mean score and SE of each sound contents for Q10

|   | 平均   | 標準誤差 |   | 平均   | 標準誤差 |
|---|------|------|---|------|------|
| A | 4.22 | 0.22 | K | 4.17 | 0.27 |
| B | 3.89 | 0.28 | L | 3.78 | 0.24 |
| C | 4.06 | 0.21 | M | 3.50 | 0.27 |
| D | 4.33 | 0.23 | N | 4.56 | 0.18 |
| E | 3.94 | 0.25 | O | 4.11 | 0.21 |
| F | 4.00 | 0.18 | P | 3.67 | 0.24 |
| G | 4.33 | 0.18 | Q | 3.56 | 0.30 |
| H | 4.44 | 0.25 | R | 4.28 | 0.18 |
| I | 3.78 | 0.29 | S | 4.56 | 0.17 |
| J | 3.89 | 0.24 | T | 3.28 | 0.32 |

感」を「その場にいる感じ」と定義した場合においても、コンテンツによって評価値が変わることが確認された。

### 3.1 主成分分析とクラスター分析によるコンテンツの分類

#### 3.1.1 主成分分析による臨場感に関する要素の抽出

まず主成分分析を行い、臨場感を構成する成分を抽出し、各コンテンツの主成分得点を算出した。Q10~12 を除いた Q1 から Q9 の評価項目に対する各コンテンツの平均評価値をサンプルとして、主成分分析を行った結果、固有値が 1 以上の主成分は 3 つであり、寄与率は第 3 主成分までで累積 54.0%であった(表 4)。第 1 主成分については「距離感」、「方向感」、「動き」「情景がわかる」「いろいろな音」といった項目で負荷量の値が高いことから定位性に関係する主成分と解釈した。第 2 主成分については、「音に包まれている」で値が突出して高かったことから、包囲感に関係する主成分と解釈した。第 3 主成分では「聞き慣れた音」「リアクションしたくなる」で値が高いことから経験や日常性に関係する主成分と解釈した。研究の結果と、「臨場感」を定義せず聴覚臨場感の生起要因を因子分析により抽出した先行研究[16, 22]との比較を行う。先行研究では既知の音であるかといった「音の情報に関わる因子」[16, 22]と「音像の定位に関する因子」[16]が抽出されている。これらの因子は本研究で抽出された第 3 主成分(経験や日常性)と第 1 主成分(定位性)に対応するものと考えられる。一方、先行研究では、「評価制因子」や「心的負荷因子」が抽出されているが、本研究では抽出されなかった。これは、本研究における主成分分析(「その場にいる感じ」に関わる要素の分析)では、「評価因子」や「心的負荷因子」に関連する評価語を除外したために、先行研究で抽出されたこれらの因子と同等と考えられる成分が抽出されなかったものと考えられる。さらに、先行研究[16, 22]で得られてない包囲感が、本実験では第 2 主成分として抽出された。これは実験で用いた再生システム(パイノ

表 4 主成分分析表

Table 4 Table of principle component analysis

|             | 主成分負荷量 |       |       |
|-------------|--------|-------|-------|
|             | 第1主成分  | 第2主成分 | 第3主成分 |
| 1. 定位性成分    |        |       |       |
| 音に方向感を感じる   | 0.75   | 0.10  | -0.19 |
| 音に距離感を感じる   | 0.73   | 0.03  | -0.39 |
| 音に動きを感じる    | 0.66   | 0.19  | -0.40 |
| 情景がわかる      | 0.57   | -0.17 | 0.43  |
| いろいろな音を感じる  | 0.47   | 0.28  | 0.27  |
| 2. 包囲感成分    |        |       |       |
| 音に包まれている    | 0.04   | 0.83  | 0.15  |
| 3. 日常性成分    |        |       |       |
| 聞き慣れた音である   | 0.41   | -0.24 | 0.54  |
| リアクションしたくなる | 0.24   | 0.24  | 0.43  |
| 人の存在を感じる    | 0.49   | -0.51 | -0.02 |
| 寄与率 (%)     | 28.22  | 13.56 | 12.20 |
| 固有値         | 2.50   | 1.20  | 1.10  |

ーラルシステムと 3 次元音場再現システム)の違いによるものと考えられる。「音に包まれている感じがする」については、前後エネルギー比、第 1 波面の法則を超えるエネルギー比や音源の到来方向[23]が関与すると考えられている。このような物理特性を再現するには個人特性の補正を行っていないバイノーラルシステムでは不十分であり、音源コンテンツに情報として含まれていたとしても再現されず、因子として抽出されなかったのではないかと考えられる。3 次元空間音響システムに関する他の研究では、臨場感(リアリティ)を知覚するために定位感、空間の広がりや包まれ感が知覚できるような聴空間の再現性が重要であることが指摘されている[12, 24]。これらのことを考慮すると、聴覚臨場感のうち「その場にいる感じ」について基本的な成分が抽出されたものと考えられる。

### 3.1.2. クラスタ分析によるコンテンツの分類

次に臨場感の感覚要素によってコンテンツの分類を行うため、第 3 主成分までの主成分得点を用いてクラスタ分析(Ward 法)を行った。その結果、得られた dendrogram を図 3 に示す。平方ユークリッド距離 10 をカッティングポイントと設定したところ、コンテンツは 3 つのクラスターに分類された。各クラスターの特徴を示すために、x 軸を第1主成分、y 軸を第 2 主成分、z 軸を第 3 主成分とし、各コンテンツの分布を図 4 に示した。第 1 クラスタは第 2 主成分の成分得点が高い群であり、第 2 クラスタは第 1 主成分得点が高いコンテンツが多かった。第 3 クラスタは第 3 主成分得点が高く、かつ第 2 主成分が低いコンテンツが多いことがわかった。クラスターごとに各主成分の平均得点および標準誤差を算出したものを図 5 に示す。分散分析を行った結果、全ての成分でクラスターの主効果が認められた。Scheffe の多重比較を行った結果も図 5 に記す。これらの結果から、第 1 クラスタでは包囲感が、第 2 クラスタでは定位性が、

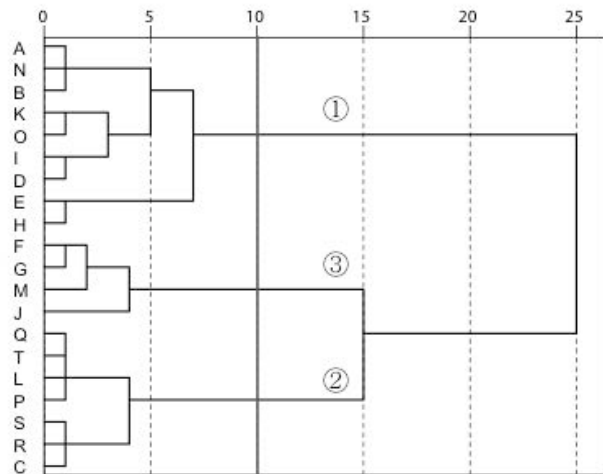


図 3 主成分得点によって分類されたデンドログラム

Fig.3 The dendrogram by the PCA score

第 3 クラスタでは日常性の主成分得点が高く、各クラスターで臨場感を左右する要素が異なることが示唆された。

各クラスターに分類されたコンテンツについて見てみると、同一クラスター内のコンテンツに含まれる対象音源の種類に共通点があることがわかった。まず第 1 クラスタではセミや鈴虫などの環境音(K, N, O)や音楽系コンテンツ(A, B, D, E, H, I)が多かった。第 2 クラスタは人の走行音(S)や歩行音(R)など対象となる音が移動するコンテンツ(C, L, P, Q, T)が多く、第 3 クラスタはバンド演奏(F, G)や会話音(J, M)など音声が含まれるコンテンツであった。

### 3.1.3. コンテンツ分類結果に関する考察

主成分得点によって分類されたクラスターごとに特徴的な感覚要素があり、かつ各クラスターのコンテンツに共通性が認められたことから、対象音源の種類と臨場感を左右する感覚要素には、ある程度の対応関係があると考えられる。このことは対象音源の種類によって対応する臨場感の感覚要素が異なる可能性を示す。そのため、システムの臨場感を評価する際には異なる感覚要素を生起する複数のコンテンツを用意したほうがよいと思われる。

また今回の研究では、音楽系・環境音系コンテンツ(第 1 クラスタ)と包囲感が、移動音源(第 2 クラスタ)と定位性が、音声コンテンツ(第 3 クラスタ)と経験や日常性という対応関係がみられた。このような対応関係が見られた理由として以下のことが考えられる。まず、音楽・環境音のコンテンツでは、音楽系コンテンツでは音源数が多い(A, B, D, K, N, O)、残響時間が比較的に長い場所(A, B, E, H)、音源がマイクを取り囲む配置(A, I, K, N, M)が多く、そのため包囲感の評価が高まったと考えられる。次に移動音源のコンテンツでは、音源の移動によって定位精度が向上する[25]ため、定位性に関わ

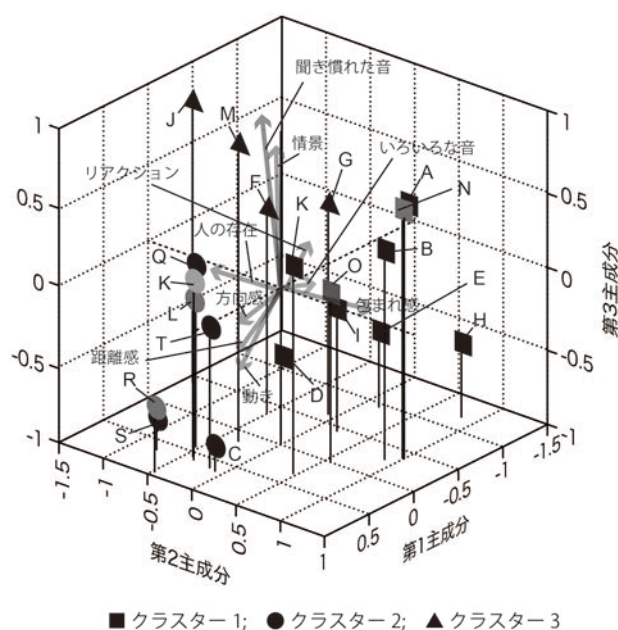


図4 各コンテンツの主成分得点による3次元プロット図

Fig.4 PCA score of each sound contents

る評価が高まったためと考えられる。また音声コンテンツでは、会話(L, M)、バンド演奏(F・G)など日常で接する機会が多いため、日常性の評価が高かったと考えられる。以上のことから、評価用コンテンツ選択の指針として、音楽・環境音、音声、そして移動音源の3種類からそれぞれ選択する、といったことが有効と考えられる。

一方で、一見クラスターの分類結果と含まれる音源が整合していないコンテンツもみられた。たとえばI(合唱)は音声が含まれているものの第3クラスター(音声)ではなく第1クラスター(音楽・環境音)であった。Iは対象音源の種類からみれば、どちらのクラスターにも分類可能である。しかし、Iではマイク(聴取者)を取り囲むように音源(歌手)が配置されている。そのため「音に包まれている」の評価が高くなり、結果として包囲感を主要な感覚要素とする第1クラスターに分類されたと推測できる。同様に、L(会話音声)も音声であるものの、第3クラスターではなく、第2クラスター(移動音源)に分類されている。Lは屋外での会話音声であり、かつ、話者が歩きながら話しているコンテンツであった。第2クラスターに分類されたのは、移動によって定位性の評価が高まったためと推測できる。複数のクラスターにまたがる音源を含むコンテンツに関しては、対象音源と対応関係にある主成分のうち得点が最も高い主成分によって分類されていた。このように対象音源と感覚要素が一对一に対応するとは限らない。そのため評価用コンテンツとして複数の対象が含まれるコンテンツを選択する場合には、音源配置や録音場所などの情報から主要な感覚要素を推測する必要がある。

#### 3.1.4. クラスターごとの総合的評価値に関する分析

次に、クラスターごとに臨場感の総合的な評価値に差

異があるか確認するため、Q10(自分がその場にいるように感じる)に対する各クラスターの平均評価値と標準誤差を求め、1要因の分散分析を行った結果、クラスターによる評価値の違いは認められなかった( $F(2,19) = 1.26, p = 0.39$ )。この結果は、各クラスターで臨場感の感覚要素は異なっても、「その場にいる感じ」という評価には差異がないことを示している。先行研究では対象となる音源が移動するようなコンテンツで高い臨場感が得られることが示されており[15]、本研究の結果とは一見矛盾する。しかし、先行研究と本研究ではいくつか異なる点がある。たとえば先行研究では「臨場感」を「あたかもその場にいるような感覚」と定義せず、あえてそのまま使用している点である。「臨場感」という言葉には「あたかもその場にいるような感じがする」という辞書的な意味の他に、「迫力がある」などの意味があることが指摘されている[8]。そのため「臨場感」という語をそのまま用いた先行研究の結果はコンテンツの「迫力」に対する評価も反映されたものと考えられる。加えて、先行研究と本研究では用いたコンテンツの種類、再生システムにも大きな違いがある。こうした手法上の違いが結果に影響したと考えられる。

#### 3.2 総合評価項目による分析

「Q10: 自分がその場にいるように感じる」という臨場感の評価に加え、デモンストレーションに適するコンテンツとその特徴を探るために、「Q11: 楽しめる」という評価項目と「Q12: 緊張感がある」という評価項目を設定し、評価値(中央値)を分析した。Q10~Q12については、その評価値をもとに順位付けを行った。1位であれば20、2位であれば19というように順位に応じて重み付けを行

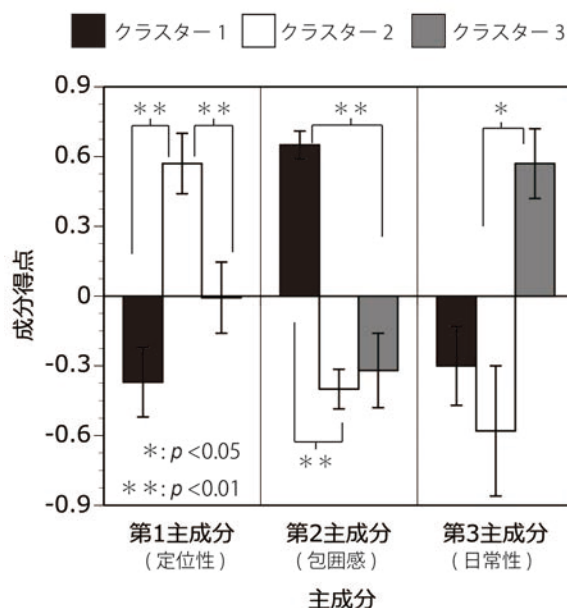


図5 各クラスターの平均主成分得点

Fig. 5 Mean PCA score of each cluster

表 5 Q10 - Q12 の合計評価値の

上位・下位の 3 コンテンツ

Table 5 The top 3 and the worst 3 contents of  
sum of Q10 - Q12 scores

| 順位 | 上位 |       |       |     | 下位 |     |     |       |
|----|----|-------|-------|-----|----|-----|-----|-------|
|    | 総合 | Q10   | Q11   | Q12 | 総合 | Q10 | Q11 | Q12   |
| 1  | A  | M     | A     | H   | T  | F   | S   | J     |
| 2  | B  | P (1) | B (1) | A   | J  | E   | R   | K     |
| 3  | H  | S (1) | D     | D   | Q  | T   | T   | L (2) |

括弧内の数字は同率の順位を示す

い、3 評価項目の合計値を算出した。

表 5 に合計得点が高い、および低いコンテンツをそれぞれ 3 つ示す。合計得点の高いコンテンツはどれも音楽が含まれるエンターテインメント性が高いものであった。一方、得点が低くなったコンテンツは動作音や街頭音といった対象音源が動くものや、会話音声であった。

下位 3 コンテンツはどれも聞き慣れた音であり、その日常性が娯楽性を減じた可能性が考えられる。しかしこれらのコンテンツでは Q1(聞き慣れた音である)の平均評価値は全コンテンツの平均値よりも低く、必ずしも日常的な音と評価されているわけではない。むしろ日常性に関する評価値は音楽などのコンテンツの方が高かった。これらのことを考慮すると、今回の被験者は通常の生活で聞く音というよりはむしろオーディオシステムで聞く経験としての日常性を評価したのではないかと考えられる。従来のシステムでは音声や音楽などのコンテンツは作品として作られており、歴史も長いことから、聴取者はその楽しみ方についても熟知している。一方、単なる走行音や日常会話のみをオーディオ機器で聞く機会は少なく、どのように聞くべきコンテンツなのか、聴取者に理解されていない。そのため、結果として Q1の評価が低くなった可能性がある。

聴取者の「聴き方」がコンテンツ評価に影響する可能性については本研究の結果のみでは明らかにすることはできない。しかし、本研究で利用したコンテンツのうち 6 コンテンツ(A, B, C, D, L, N)を用い、音響技術者や研究者などの専門家 54 名を対象に調査を行ったところ(調査手法などの詳細については付録 A-2 参照)、一般と専門家では異なる傾向があった。専門家に対する調査では「その場にいるように感じるか」について 5 段階評価を、「楽しめる」について順位付けによる評価を行った。(なお、「緊張感がある」という評価項目は用いなかった。)表 6 左列に「その場にいる感じ」に対する評価値を比較したものを示す。また、本実験で各コンテンツに対する「楽しめる」の平均スコアをもとに順位付けを行った結果と専門家対象の調査での順位の比較を表 6 右列に示す。この表から大学生を対象にした本実験では A

表 6 専門家と一般聴取者の評価の比較

Table 6 Comparison of the evaluation for 6 contents  
between normal and experts

| コン<br>テ<br>ン<br>ツ | その場にいる (値) |      | 楽しめる (位) |     |
|-------------------|------------|------|----------|-----|
|                   | 一般         | 専門家  | 一般       | 専門家 |
| A                 | 4.55       | 3.6  | 1        | 5   |
| B                 | 4.27       | 3.55 | 2        | 1   |
| C                 | 4.27       | 4.31 | 4        | 2   |
| D                 | 4.02       | 3.97 | 3        | 6   |
| L                 | 3.16       | 4.13 | 5        | 3   |
| N                 | 3.27       | 4.24 | 6        | 4   |

(オーケストラ・ステージ), B(オーケストラ・客席), D(能)といった鑑賞性が高いコンテンツの順位が高いのに比べて、専門家では L(会話音声・屋外)や N(セミ)などの必ずしも鑑賞性が高いとは言えないコンテンツに対する順位が高かった。専門家を対象にした調査ではコンテンツ数が少ないなど本研究と異なる点が多く単純に比較できないが、評価者のオーディオシステムやコンテンツに対する構えが評価に大きく影響する可能性も考えられる。

#### 4 まとめ

「臨場感」を「その場にいるような感じ」と定義した今回の実験から、たとえ総合的な音による臨場感の評価値に差異が認められなかったとしても、コンテンツの種類によって、聴取者が経験している臨場感の感覚要素は異なることが示された。さらにコンテンツの分類を試みところ、音楽、音声、そして移動音源の 3 種類に分類され、それぞれ包囲感、日常性、定位性といったように臨場感の評価を決定付ける感覚要素が異なることが示唆された。

このことから評価を行ううえでは数種類のコンテンツを用意する必要があると考えられる。今回の研究では、音楽、音声、そして移動音源に感覚要素の違いが見られたことから、1つの指針としてこれら 3 種類のコンテンツを選択することが推奨される。ただし、今回の実験で用いたコンテンツには限りがあり、今後、本研究で得られたコンテンツの分類と感覚要素の対応が一般化するものか検証する必要がある。加えて使用した音情報提示システムは 1 種類であったため、他のシステムを用いても同様の傾向が見られるのか、特にシステムの再現性の観点から検討することも必要であろう。

また「その場にいるように感じる」「楽しめる」「緊張感がある」という観点でコンテンツ選択を行うと、音楽系の要素が含まれたコンテンツが高く評価されることがわかった。一方、単純に動いている音や日常的な音に関しては「その場にいるように感じる」という評価は得られるも

の「楽しめる」「緊張感がある」という評価は低くなる傾向が見られた。ただし、本研究で使用したコンテンツのうち 6 コンテンツを用いて行った音響の専門家を対象としたアンケートでは、音楽など鑑賞性が高いコンテンツよりも会話や屋外の音など必ずしも鑑賞性が高いとはいえないコンテンツの評価が高い傾向であった。そのためコンテンツ選択には、対象となる聴取者の「音響システムでの聴取経験」を考慮する必要があると思われる。

これらのことから、今回の実験のように一般の試聴者を対象にした場合には、従来のシステムで試聴されるコンテンツ同様、ある種の作品性が必要といえよう。高評価を受けるためには音楽系の要素を含みつつ、3 次元音場再現システムの良さを引き出すようなコンテンツを制作することが求められるものと考えられる。

### 謝辞

本研究は科学技術振興機構(JST)戦略的創造研究推進事業 (CREST)「音楽を用いた創造・交流活動を支援する聴空間共有システムの開発」(研究代表:伊勢史郎)の助成を受けた。またコンテンツの収録に協力いただいた同プロジェクトの関係者各位に感謝する。

なお、本研究は明治大学理工学部遺伝子組換え実験に関する安全及びヒトを対象とした実験研究に関する倫理委員会の承認(承認番号:理工安倫 11 第 510 号)をうけて実施された。

### 参考文献

- [1] ITU-R Rec. BS.775-2: Multi-channel stereophonic sound system with or without accompanying picture; ITU, (2006)
- [2] Hamasaki, K., Hiyama, K., Nishiguchi, T., Okumura, R: Effectiveness of Height Information for Reproducing Presence and Reality in Multichannel Audio System; AES 120<sup>th</sup> Conversation, Paris, France, Convention paper 6679, (2006)
- [3] Berkout, A.J., de Vries D., Vogel, P: Acoustic control by wave field synthesis; J. Acoust. Soc. Am., 93 (5), 2764-2778 (1993)
- [4] Cooper, D. H., Shiga, T: Discrete-matrix multichannel stereo; J. Audio Eng. Soc., 20 (5), 346-360 (1972)
- [5] Gerzon, M.A: Periphony: With-height sound reproduction; J. Audio Eng. Soc., 21 (1), 2-10 (1973)
- [6] 伊勢史郎: キルフホッフヘルムホルツ積分方程式と逆システム理論に基づく音場制御の原理; 日本音響学会誌, 53 (9), 706-713 (1997)
- [7] Takane, T., Suzuki, Y., Sone, T: A new method for global sound field reproduction based on Kirchhoff's integral equation; Acustica, 85, 250-257 (1999)
- [8] 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, 15 (1), 7-16 (2010)
- [9] Slater, M., Steed, A., McCarthy, J., Maringelli, F: The influence of body movement on presence in virtual environment; Hum Factors, 40, 469-477 (1998)
- [10] Sheridan, T.B: Musings on telepresence and virtual presence; Presence (Camb), 1, 120-126 (1992)
- [11] Lessiter, J., Freeman, J., Keogh, E., Davidoff, J: A cross-media presence questionnaire: The ITC Sense of Presence Inventory; Presence (Camb), 10, 282-298 (2001)
- [12] Lombard, M., and Ditton, T: At the heart of it all: The concept of presence; J. Comput. Mediat. Commun., 3 (2), 1-43 (1997)
- [13] Slater, M., and Usoh, M: Representations systems, perceptual position, and presence in immersive virtual environments; Presence (Camb), 2, 221-233 (1994)
- [14] Slater, M., Wilbur, S: A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments; Presence (Camb), 6, 603-616 (1997)
- [15] 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-; 信学技報, 108-333, 83-88 (2008)
- [16] Ozawa, K., Chujo, Y., Suzuki, Y., Sone, T: Psychological factors involved in auditory presence; Acoust. Sci. & Tech, 24 (1), 42-44 (2003)
- [17] 安藤広志, 日本バーチャルリアリティ学会(編): バーチャルリアリティ学; 日本バーチャルリアリティ学会, (2011)
- [18] Enomoto, S., Ikeda, Y., Ise, S. Nakamura, S: Three-dimensional sound field reproduction and recording systems based on boundary surface control principle; Proc. 14<sup>th</sup> ICA, ICAD08-1-8 (2008)
- [19] 池田雄介, 榎本成悟, 伊勢史郎, 中村哲: 多チャンネル 3次元音場再現システムの再生方式の違いによる再現性能評価; 信学技報, EA2008-114, 77-82 (2008)
- [20] 阿部幸治, 小澤賢司, 鈴木陽一, 曾根敏夫: 音色表現語, 感情表現語及び音情報関連語による環境音評価; 日本音響学会誌, 54 (5), 343-350 (1998)
- [21] 福江一智, 小澤賢司, 木下雄一郎: 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, 11(2), 183-192 (2012)
- [22] 福井誠人, 小林まおり, 上野佳奈子: 多チャンネル 3次元再現音場における聴覚印象の評価構造; 日本音響学会講演論文集, 871-872 (2012)
- [23] 飯田一博, 飯田一博・森本政之(編): 空間音響学; コロナ社, (2010)
- [24] Biocca, F., Delaney, B: Immersive virtual reality technology. In Biocca, F. Levy, M. R. (Eds.), Communication of virtual reality, Hillsdale, NJ: Lawrence Erlbaum Associates.
- [25] Wightman F. L., Kistler, D.J: Resolution of front-back ambiguity in spatial hearing by listener and source movement; J. Acoust. Soc. Am., 105 (5), 2841-2853 (1999)
- [26] Ikeda, Y., Ise, S: Sound Cask: Music and voice communication system with three dimensional sound reproduction based on boundary surface control principle; Proc. 21<sup>st</sup> ICA, 19, (2013)
- [27] 山下真依, 中島宏毅, 小林まおり, 池田雄介, 榎本成悟, 上野佳奈子, 伊勢史郎: “音響樽”の音像定位に関する実験的検討-BoSC 再生システム 96 ch と 62 ch の比較-; 日本音響学会講演論文集, 719-720 (2013)

## A 付録

## [著者紹介]

### A.1 評価グリッド法による予備実験

何らかの「好み」「総合評価」の理由とそれに関連する要素をインタビュー形式で問い、個人の評価構造を探索する一連の測定手法を評価グリッド法という。予備実験では、今回提示したコンテンツと同様のコンテンツを6名にそれぞれ提示した後に、評価グリッド法によるインタビューを行った。最初にインタビュー対象者に聴空間の印象が似たコンテンツをグルーピングさせ任意の数のグループを作らせた。そのうち2つのコンテンツグループを比較して、「どちらがよりその場にいるように感じるか」を選択させ、その理由について述べさせ、「その場にいる感じ」に関連する評価語を抽出した(総当たり形式で全てのグループの比較を行った)。

### A.2 音響の専門家を対象としたアンケート調査

音響の専門家を対象にアンケート調査を行った。手続きは下記の通りである。

**調査年月** 2012年10月(AES ジャパンコンファレンス・仙台 2012 開催中に設置、調査を行った)

**調査場所** 宮城県仙台市(せんだいメディアテーク)

**調査対象者** 音響の専門家 54名であった。

**装置** 本実験と同様に境界音場制御の理論をもとにして音場再現システムを用いた。ただし、再現システムは96chのスピーカを内蔵した音響樽[26]を用いた。なお本実験で用いた62ch再生システムと96ch再生システムでは、定位性能について差異はない[27]。

**刺激** 本実験で用いた20コンテンツのうち6コンテンツ(A, B, C, D, L, N)を呈示した。また本実験には用いていないコンテンツ X(歩きながらのギターの演奏)も併せて呈示した。各コンテンツの試聴時間は60s程度であり、全7コンテンツを試聴するのに7min程度を要した。

**手法** 聴取者は音響樽内の椅子に座り、コンテンツを試聴した。各コンテンツを試聴中に、そのコンテンツの「その場にいる感じ」について「1:全く感じない」～「5:とても感じる」の5段階で評価した。またコンテンツ終了後に「楽しめた」コンテンツを上位3つ挙げるよう求めた。

**結果の整理** 「その場にいる感じがするか」に対する5段階評価は被験者全54名の平均値をコンテンツごとに算出した。また、「楽しめた」という評価に関しては、1位として選択されたコンテンツには3点、2位には2点、3位には1点というように順位に応じて重み付けを行い、被験者54名の合計得点を用いて順位付けを行った。

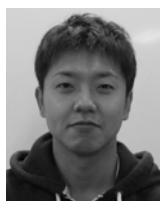
(2013年7月19日受付)

### 小林 まおり (正会員)



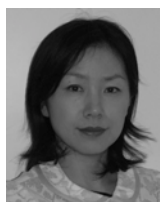
2006年立教大学大学院文学研究科修了、同年東北大学電気通信研究所研究員、2011年明治大学研究知財戦略機構研究員、現在に至る。聴覚心理学およびマルチモーダル知覚情報処理に関する研究に従事。博士(心理学)。

### 福井 誠人 (非会員)



2012年明治大学理工学部建築学科卒業、同年4月より明治大学大学院理工学研究科建築学専攻入学、現在に至る。日本音響学会学生会員。

### 上野 佳奈子 (非会員)



1998年東京大学大学院工学系研究科修了、1999年東京大学生産技術研究所助手、2008年明治大学理工学部専任講師、2010年同大学准教授、現在に至る。建築音響、環境心理に関する研究に従事。博士(工学)。