

視覚障がい者のVR体験質向上に向けた アバター特徴提示における足音表現の検討

丸尾 悠暉[†] 横山 皓祐^{††} Siriaraya Panote^{†††} She Wan-Jou^{†††} Cheng Asaka Lan[†]
中島 伸介[†]

[†] 京都産業大学情報理工学部情報理工学科 〒603-8555 京都府京都市北区上賀茂本山

^{††} 京都産業大学先端情報学研究科先端情報学専攻 〒603-8555 京都府京都市北区上賀茂本山

^{†††} 京都工芸繊維大学 情報工学・人間科学系 〒606-8585 京都府京都市左京区松ヶ崎橋上町

E-mail: [†]{g2354503,i2486237,a007532,nakajima}@cc.kyoto-su.ac.jp, ^{††}{spanote,wjs2004}@kit.ac.jp

あらまし 近年、仮想環境内でのさまざまなアクティビティや交流が容易になった。しかし、視覚障がい者はアバターの服装や体格といった視覚的な非言語情報を取得できず、他者の特徴把握が困難である。我々は先行研究として、他者の存在や属性を音声によって提示することで、会話開始のきっかけを支援するシステムを開発したが、相手の外見に由来する雰囲気や印象といった非言語的なニュアンスを伝達するには至らなかった。そこで本研究では、こうした情報を補完する手法として足音に着目し、聴覚情報のみでアバター特徴を直感的に把握可能か検討した。実験では、アバターの属性を強調するためにデフォルメを加えた複数の足音を作成し、聴取による印象評価を行った。本稿では、得られた回答データの分析結果に基づき、足音の音響的特徴と想起されるアバター属性との関連性を検証するとともに、視覚障がい者が直感的に他者を認識できる足音表現の要件について議論する。

キーワード VR, ユーザー支援, 立体音響, ソーシャルインタラクション

1 はじめに

視覚障がい者は、令和4年度の厚生労働省の調査によると、日本に27.3万人いるとされている[1]。多くの最新技術が視覚障がい者のアクセシビリティの向上を試みているが[2][3][4][5]、主に視覚中心の体験に焦点を当てたVR開発における肯定的な例が少ないのが現状である[6][7][8][9]。

一方で、VRを含むメタバース分野は今後ますますの発展が期待されている。令和6年の総務省の調査によると、世界のメタバース市場は2022年の461億ドルから2030年には5,078億ドルまで拡大すると予測されている[10][11]。その中でも、ソーシャルVRは注目すべきジャンルの一つであり、インターネットを介して仮想空間内で様々なユーザーと会話や社会的な交流を行うことが可能である。ソーシャルVRは単なる通話手段を超え、アバターを介した社会的な交流の場として定着しつつある。晴眼者は、対話相手のアバターの外見——例えば、重厚な鎧や軽やかな妖精の姿といった視覚的特徴——から、相手の属性やキャラクター性を瞬時に把握し、これを非言語的な手がかりとして円滑なコミュニケーションを行っている。しかし、視覚障がい者にとって、この「アバターの見た目」に関する情報を得ることは困難であり、相手が持つ「雰囲気」や「印象」までは直感的に理解できない。そのため、VRChatのような没入型空間であっても、視覚障がい者にとってはSkypeやZoom等の音声会議ツールと体験が変わらず、他者との交流における心理的な距離感や没入感の欠如につながる可能性がある。

我々は先行研究において、他者の視線情報から名前、性別情報を音声通知することで、視覚障がい者の会話開始のきっかけを支援するシステムを開発した[12]。このシステムにより、「誰がいるか」という個人の識別は可能となった。しかし一方で、相手の外見（体格や服装など）から醸し出される「雰囲気」や「キャラクター性」といった「相手がどのような印象の人物か」という非言語的なニュアンスまでは伝達できていなかった。そこで本研究では、ソーシャルインタラクションの「質」を高めるために、言語説明に頼ることなく、他者のアバター特徴を聴覚のみで直感的に提示する手法を検討する。そのための手がかりとして、アバターの移動に伴い常時発生し、かつ履物や体格に由来する「重さ」や「硬さ」といった物理的特性を本質的に内包している「足音」に着目した。この足音の特性をVR空間に応用し、アバターの視覚的特徴（性別、体格、服装など）を強調・デフォルメした足音として提示することで、視覚障がい者が相手の人物像を具体的に想起できるか検証する。本稿では、作成した足音を用いた聴取実験を行い、足音と想起されるアバターイメージの関連性について評価した結果を報告する。

2 関連研究

本章では、視覚障がい者のためのVRアクセシビリティ支援、足音による人物特性の認知、および音響設計による印象形成に関する先行研究を整理し、本研究の位置づけを明らかにする。

視覚障がい者がVR空間内で他者と円滑にコミュニケーションを行うためには発話内容に加えて、視線やうなずきといった

非言語的な手がかりの理解が重要であることが指摘されている。Jung らは、視覚障がい者がソーシャル VR 内で会話に参加する際に生じる困難に着目し、対話相手の視線方向やうなずきといった非言語行動を音響および触覚信号として提示する手法を提案した [13]。実験結果から、非言語情報の代替提示が会話のタイミング把握や参加感の向上に寄与することを示している。

これらの研究は、視覚障がい者がソーシャル VR において直面する課題を明確にし、非言語情報を別感覚で提示することの有効性を示している。一方で、提示される情報は主に「会話の成立」に関わる機能的な手がかりに限定されており、対話相手が持つ雰囲気や人物像といった質的な情報の提示については十分に検討されていない。

人間は視覚情報が存在しない場合であっても、音情報のみから他者の身体的・社会的属性を推定する能力を持つことが報告されている。Clausen らは、足音の周波数成分を操作した音刺激を用い、聴取者が歩行者の性別や体重をどのように知覚するかを検証した [14]。その結果、低周波成分を強調した足音は「男性」や「体重が重い」といった身体イメージを想起させることが示され、足音が人物特性の認知に直接影響を与えることが明らかとなった。

また、Sun らは、スニーカー・ブーツ・ハイヒールといった履物の違いおよび歩行リズムが、聴取者の第一印象に与える影響を調査した [15]。実験の結果、履物ごとの音響特性が、「支配性」や「親しみやすさ」などの性格的評価の形成に寄与することが示されている。

これらの研究は、足音が単なる環境音ではなく、聴取者に対して人物像や印象を想起させる社会的な情報源となり得ることを示している。

足音を含む音響情報を意図的に設計することで、人物の属性を表現しようとする試みも行われている。Cunningham らは、映画制作で用いられるフォーリー技術を応用し、足音を加工することで歩行者の「自信」や「体格」といった抽象的な人物特性がどの程度伝達可能であるかを検証した [16]。その結果、視覚情報が限定された状況であっても、音響設計によって特定の人物特性を一貫して想起させることが可能であることを示した。

一方、ロボットのような非人間的対象においても、動作に伴って発生する随伴音が人間の印象形成に与える影響が報告されている。Allen らは、ロボットの駆動音の質感が人間側の信頼感や親しみやすさに影響を与えることを示し、音響設計が社会的受容性を左右する要因となり得ることを明らかにした [17]。

以上のように VR における視覚障がい者支援、足音の音響特性と人物属性の関連、および音響設計による印象形成については、それぞれ個別に研究が進められてきた。しかし、既存研究の多くは日常的な足音からの属性推定や会話成立を支援する機能的な非言語情報の提示に留まっており、アバターが持つ極端な視覚的特徴を意図的にデフォルメした足音として提示し、視覚障がい者に直感的な人物像を想起させることを目的とした研究は限られている。

そこで本研究では、スニーカー、ブーツ、裸足、ロボットと

いった明確に質感の異なる足音を用い、それらの特徴を強調・デフォルメして提示する手法を提案する。本研究は、視覚障がい者が視覚情報に依存することなく、他者のアバターイメージを即時かつ情緒的に把握できる可能性を検証する点に独自性を有する。

3 足音刺激の生成と提示方法

本章では、VR 空間におけるアバターの足音を通じて、キャラクターの属性や実在感を非言語的に伝達するための足音デザイン手法、およびその実装について述べる。

3.1 音源素材の選定

多様な足音の印象を検証するために、本研究では実在する履物（革靴、スニーカー、ヒール、ブーツ、サンダル、裸足）および架空のキャラクター（ロボット、マスコット）を含む計 8 種類の条件を設定した。刺激の作成にあたり、実在履物およびロボットの計 7 条件については、web 上のフォーリーサウンドライブラリ「ZAPSPLAT」[18] より各条件の特性に合致した音源素材を選定した。このうち実在履物条件に関しては、雪や土などの軟質な地面では足音固有の音響的特徴が不明瞭になることを避けるため、コンクリート地面を歩行した際の音源に統一した。これは、標準的な硬質地面においてアバターの特徴が認識可能かを明らかにするための基礎的な検証とするためである。ロボット条件に関しては、同ライブラリ内の金属的な重量感を想起させる素材を採用した。一方、マスコット条件に関しては、Web 上の効果音ライブラリ「DOVA-SYNDROME」[19] より、歩行音を選定し採用した。また、単一音源の反復による聴覚的な単調さを回避するため、各条件につきニュアンスの異なる 2 種類の足音バリエーションを用意した。

3.2 再生制御システムの実装

各条件間の歩行リズムや接地間隔を完全に統一するため、刺激の作成は実時間での操作ではなく、ゲームエンジン Unity を用いたレンダリングおよび録音によって行った。具体的には、Unity 上で一定速度の歩行アニメーションを再生し、周辺を一周する動作イベントに合わせて前述の 2 種類の足音バリエーションを交互に発音させる機構を構築した。この際、VR 空間における実在感や方向感を担保するため、Unity 標準の Audio Source コンポーネントに加え、空間音響プラグイン「Meta XR Audio SDK」[20] を導入した。具体的には、Spatial Blend を完全な 3D に設定するとともに、HRTF（頭部伝達関数）に基づくバイノーラル処理を有効化することで、アバターの足元から発せられる音が、聴取位置に対して上下前後を含む立体的な定位感を持つように設定した。この制御により、単調さを回避して歩行音の自然さを担保しつつ、生成された音声をデジタル録音してベースとなる音声データを作成した。また、作成にあたっては Unity 上の Audio Source コンポーネントにおいてゲイン調整を行い、聴感上の音量感が条件間で同等になるよう設定した。この生成プロセスにより、歩行速度のゆらぎや操作による個人差といった音色以外の交絡要因を排除し、純粋な音響特性

による印象評価が可能となった。

3.3 音声刺激の加工

録音により生成された音声刺激に対し、各条件が想定する物理的特性（重さ・硬さ）を聴覚的により明確化するため、波形編集ソフトウェア「Audacity」[21]を用いた信号処理を施した。具体的には、表1に示すパラメータに基づき、「ピッチ変更」および「イコライジング」による調整を行った。イコライジングにおいては、重量感に寄与する低周波数帯域（60Hz - 100Hz）と、接地感や鋭さに寄与する高周波数帯域（4kHz - 8kHz）を操作対象とした。「スニーカー」条件においては、標準的な体格を想定し、ピッチを上昇（+8%）させることにより軽快さを演出する一方で、高音域をカット（-3dB）し低音域を補強（+3dB）することでゴム素材の柔軟性を表現した。「ブーツ」条件では、「力強さ」や「身体的な大きさ」を表現するため、低周波数帯域を強調（+6dB）し、地面に重く沈み込むような重量感を補強した。対照的に「ヒール」条件では、「女性らしさ」といった属性を明確化するため、ピッチの上昇（+9%）と高周波数帯域の強調（+9dB）を行い、鋭利で硬質な接地感を際立たせた。「裸足」条件では、高周波数帯域を著しく強調（+9dB）する調整を行った。これは、通常の歩行音よりも皮膚と床面の摩擦音を際立たせることで、靴を履いていない状態特有の「肉感的な生々しさ」を強調し、他の条件との明確な差別化を図るためである。また「サンダル」条件については、低音域のカット（-3dB）と高音域の強調（+6dB）を組み合わせることで、靴底の薄さと軽さを強調し、「日常的でリラックスしたラフな印象」の提示を試みた。「ロボット」条件では「金属的な非生物感」,「マスコット」条件では「コミカルな親しみやすさ」,「革靴」条件では「フォーマルな標準性」を提示することを意図し、これらの条件については元音源が十分にその特性を有していたため、加工を行わず原音を採用した。

表 1: ピッチ変更と EQ 設定

条件	ピッチ変更	EQ Lows (60-100Hz)	EQ Highs (4-8kHz)
スニーカー	+8%	+3 dB	-3 dB
ブーツ	±0%	+6 dB	±0 dB
ヒール	+9%	±0 dB	+9 dB
マスコット	±0%	±0 dB	±0 dB
ロボット	±0%	±0 dB	±0 dB
裸足	+5%	±0 dB	+9 dB
サンダル	+5%	-3 dB	+6 dB
革靴	±0%	±0 dB	±0 dB

以上の工程を経て完成した8条件の音声ファイルを、本検証における評価用刺激とした。

4 評価実験

前章で提案・実装した足音刺激がアバターの物理的特徴やキャラクター性を正確に伝達し、アバターの種類を識別するための有効な手がかりとなり得るかを検証するため評価実験を実

施した。実験には、晴眼者である17名（20代男性11名、20代女性6名）の被験者が参加した。

4.1 実験手順

本実験は、Google フォームを用いた Web アンケート形式にて実施した。参加者は各自の PC またはスマートフォン等の端末を用いて実験に参加した。必ずヘッドホンまたはイヤホンを着装して聴取し、アンケートに回答するよう参加者に指示を行った。

提示刺激には、3章で作成した8種類の足音条件を用いた。足音の音声ファイルはフォーム内に埋め込まれ、参加者は任意のタイミングで再生が可能であった。これらの音声は、3.2節で述べた通り、聴取者の周囲をアバターが一周歩行するように空間音響処理が施されている。

事前聴取の終了後、本番となる「評価フェーズ」へ移行した。参加者は、8種類の足音刺激を一つずつ順次再生し、提示された足音ごとに、アバターのイメージおよび印象評価アンケートに回答を行った。

4.2 評価指標

参加者の抱いた印象を多角的に分析するため、以下の2種類の方法で回答を求めた。

4.2.1 SD法による印象評価

足音から受ける印象の傾向を定量化するために、SD法（Semantic Differential Method）を用いた。評価項目として、足音の物理的特性およびキャラクター性に関連する形容詞対を選定した。物理的な特性に関する項目として「重さ（軽い–重い）」「大きさ（小さい–大きい）」「硬さ（柔らかい–硬い）」を設定した。また、キャラクター性や性格に関する項目として「女性らしさ（男性的–女性的）」「若さ（年若い–若々しい）」「人間らしさ（非人間的–人間的）」「親しみやすさ（敵対的–友好的）」を、動作や気配の性質として「力強さ（か弱い–力強い）」「騒がしさ（静かな–うるさい）」「荒々しさ（おだやか–荒々しい）」「乱暴さ（慎重–乱暴）」を採用した。

参加者は、提示された足音がそれぞれの形容詞対のどちらに近いかを、1から7までの7段階のリッカート尺度（例：1=非常に軽い、7=非常に重い）で回答した。

4.2.2 自由記述回答

SD法では捉えきれない具体的なイメージや連想語を収集するために、自由記述欄を設けた。参加者は、その足音から具体的にどのようなキャラクターや情景を想起したかについて、自由に記述した。

4.3 分析方法

収集した回答データに対し、アバターの印象構造を明らかにするため、以下の定量的および定性的分析を行った。

4.3.1 印象評価の分析

SD法により得られた数値データについては、まず条件ごとの平均値を算出し、レーダーチャートを用いて各足音の特性（重さ、硬さ、大きさ等）の傾向を視覚化した。さらに、評価に用いた複数の形容詞対の背後にある潜在的な因子を抽出し、足音の

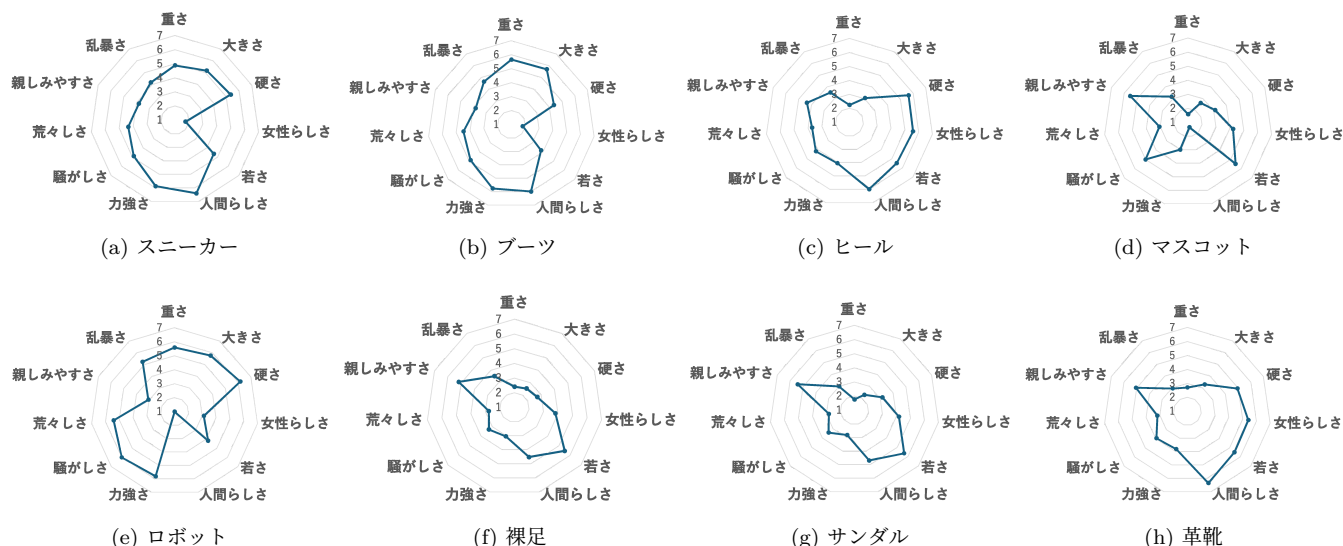


図 1: 各条件における印象評価のレーダーチャート (n=17)

知覚構造を低次元化して捉えるために、主成分分析 (Principal Component Analysis; PCA) を適用した。これにより、8種類の足音がどのような軸に基づいて聴取者に識別されているかを2次元平面上にマッピングし、各条件の位置関係を分析した。

4.3.2 自由記述の分析

自由記述回答により得られたテキストデータについては、テキストマイニング手法を用いた分析を行った。具体的には、形態素解析により名詞および形容詞を抽出し、出現頻度に応じた文字サイズで単語を描画する「ワードクラウド」を作成した。これにより、各足音から想起される具体的なキャラクター像 (例: 「ロボット」「大男」) や、SD法では捉えきれない微細なニュアンスの違いを可視化し、条件間でのイメージ傾向の差異を比較した。

4.4 結果

本実験により得られた回答データ (SD法による評価値および自由記述) について、多変量解析およびテキストマイニングを用いた分析を行った。

4.4.1 印象評価のレーダーチャート

各条件における印象評価の平均値を算出し、可視化したレーダーチャートを図1に示す。「ロボット」条件および「ブーツ」条件においては、「重さ」「大きさ」「力強さ」のスコアがいずれも顕著に高いスコアを示した。一方で、両者の間では「人間らしさ」のスコアに乖離が見られ、「ブーツ」は高水準であったのに対し、「ロボット」は極めて低いスコアとなった。

「ヒール」条件は、「重さ」や「大きさ」が低いスコアを示す一方で、「硬さ」および「女性らしさ」において高いスコアを示した。「マスコット」「裸足」「サンダル」の3条件は、「重さ」や「力強さ」が低く評価されたが、「親しみやすさ」はいずれも高いスコアとなった。特に「マスコット」条件は、「人間らしさ」が低いスコアである一方、「親しみやすさ」では全条件中で最大のスコアを記録した。

「スニーカー」条件では、「女性らしさ」が低く、「人間らし

さ」が高いスコアを示した。一方で、「若さ」のスコアは他の条件と比較して突出した高さは示さなかった。無加工の「革靴」条件については、概ね女性らしい傾向を示し、「女性らしさ」のスコアは5.41と中立よりもやや高いスコアとなった。

4.4.2 主成分分析

SD法により得られた全回答データに対して主成分分析を行い、足音の印象を構成する主要な因子の抽出を試みた。分析の結果、第1主成分 (PC1) の寄与率は51.1%、第2主成分 (PC2) の寄与率は11.0%となり、これら2つの成分によって全体の分散の約6割 (62.1%) を説明可能であることが示された。算出された主成分負荷量および散布図をそれぞれ表2と図2に示す。

まず、第1主成分 (PC1) においては、「力強さ (0.369)」や「硬さ (0.262)」が正の方向に高い負荷量を示した一方で、「親しみやすさ (-0.329)」や「若さ (-0.248)」が負の方向に負荷量を示した。散布図上での配置を確認すると、PC1の正方向 (右側) には「ロボット」や「ブーツ」といった重量感や威圧感のある条件が配置され、負方向 (左側) には「マスコット」や「裸足」といった軽快で愛嬌のある条件が配置されている。

次に、第2主成分 (PC2) においては、「人間らしさ (0.643)」が正の方向に突出して高い負荷量を示した。散布図の縦軸 (PC2) に着目すると、正の領域 (上方) には「革靴」「スニーカー」「ヒール」など、実在する人間用の履物が一貫して配置されているのに対し、負の領域 (下方) には「ロボット」および「マスコット」という架空のキャラクター条件が明確に分離して配置された。

4.4.3 条件ごとのワードクラウド

自由記述回答から抽出された名詞および形容詞を用い、各条件のイメージ構造を可視化するためにワードクラウドを図3に示す。出現頻度の高い単語群を分析した結果、レーダーチャートで確認された印象傾向を、より具体的なキャラクター像として裏付ける結果が得られた。

「ロボット」条件では「ロボット」「機械」「警備」、「ブーツ」

表 2: 各評価項目の主成分負荷量

評価項目	PC1	PC2
力強さ	0.369	0.056
荒々しさ	0.358	-0.287
重さ	0.354	0.296
大きさ	0.352	0.188
親しみやすさ	-0.329	0.093
乱暴さ	0.329	-0.240
騒がしさ	0.270	-0.374
硬さ	0.262	-0.040
女性らしさ	-0.257	-0.289
若さ	-0.248	-0.296
人間らしさ	-0.024	0.643

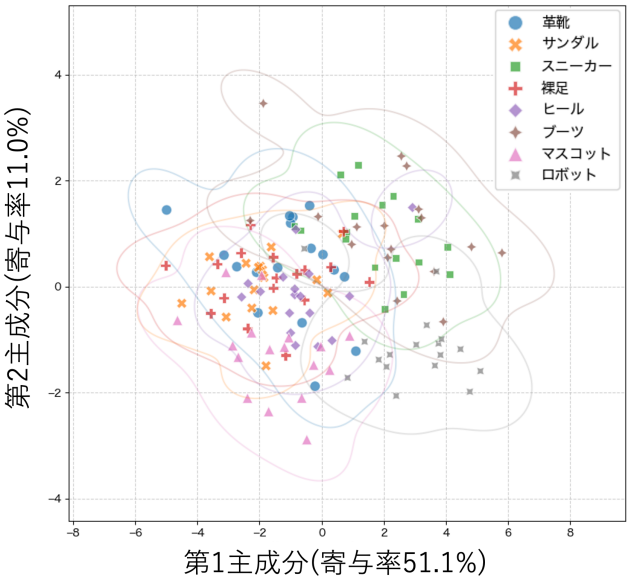


図 2: 主成分分析 (PC1 vs. PC2) の散布図

条件では「男性」「体格」「雪国」といった単語が多く出現した。「スニーカー」条件では「男性」「スーツ」「体格」が支配的であった。「ヒール」条件では「女性」「ヒール」「ハイヒール」といったジェンダーを特定する単語が抽出された。「裸足」および「サンダル」条件では「子供」「動物」「ペンギン」といった語が共通して見られた。「マスコット」条件では「キャラ」「アニメ」「ロボット」といった語が含まれていた。「革靴」条件では「革靴」「スーツ」に加え、「ヒール」「女性」という単語が比較的大きく表示された。

4.5 考 察

4.5.1 設計意図と知覚結果の整合性

実験の結果、各足音音源に対して施した信号処理（ピッチおよびイコライジング）による加工意図は、概ね被験者の知覚結果と整合していた。これは、視覚情報が制限された環境において、音響パラメータの操作を施した足音がアバターの属性やキャラクター性の想起を促す有効な手段であることを示している。

「ロボット」条件および「ブーツ」条件においては、「重さ」や「力強さ」といった物理的強度が強調された結果、いずれも



図 3: 各条件におけるワードクラウド

顕著に高いスコアを示した。特に「ロボット」条件では、意図した機械的な質感が、生物的な人間とは異なるシステム的な存在としての認識を促した。一方、「ブーツ」条件において低周波数帯域（60Hz - 100Hz）を強調（+6dB）したことが、地面への重い沈み込み感として正しく受容された。これにより、同じ重量級の足音であっても、信号処理の差異によって無機質なマシーンと屈強な人間の違いを明確に描き分けられることが示された。また、「サンダル」条件で低域を抑制（-3dB）したことは、重厚感の排除と軽快な印象の形成に寄与した。加えて、「マスコット」条件のように元音源の特性を活かした非人間的な高音は、「親しみやすさ」の提示に成功した。これらの結果は、物理的な音の強さが威圧感に繋がり、親和性とトレードオフの関係にあるという、第 1 主成分における「強度と親しみやすさの負の相関」を裏付けるものである。

「ヒール」条件における高域強調（+6dB）はコンクリート路面での「硬さ」を強調し、典型的な女性像の想起を成功させた。「裸足」条件では、高周波数帯域を大幅に強調（+9dB）し、靴の打撃音とは異なる「皮膚の摩擦音（ペタペタという音）」として知覚され、自由記述回答においても「子ども」や「ペンギン」といった、非装飾的な人間味や生物性を際立たせる結果と

なった。これは、デフォルメを伴う強調処理が、対象の実在感を補強する可能性を示唆している。

無加工の「革靴」条件では、硬質な高域成分が原因で一部の参加者に「ヒール」と誤認される傾向が見られた。「スニーカー」条件では、ピッチの上昇(+8%)によって「若さ」や「軽快さ」の提示を試みたものの、実際の評価では「標準的な男性」に留まり、若さのスコアは向上しなかった。これは、スニーカーの素材感(ゴム底)を表現するために施した高周波数帯域の抑制(-3dB)が、音の明瞭度や鋭さを減退させ、ピッチ上昇による若々しい印象を打ち消してしまったことが要因と考えられる。ここから、ジェンダーのような基本属性の正確な伝達には帯域制御による「素材感のデフォルメ」が有効である一方、年齢層のような詳細な属性提示には、音色操作に加えて歩行テンポ等の時間的要素の制御も必要である可能性があることが明らかとなった。

以上の通り、各条件において施した音響的特徴の強調は、聴取者のイメージ形成を概ね設計意図に沿った方向へと誘導し、信号処理による属性提示の有効性を裏付ける結果となった。この知見は、次節で述べる「物理的な強度」と「人間属性の有無」の2軸に基づく足音設計指針の妥当性を支えるものである。

4.5.2 足音設計における軸

VR空間における足音の印象が「物理的な強度(PC1)」と「人間属性の有無(PC2)」という、互いに独立した2つの次元で構成されていることを示した。これは、アバターの足音を設計する際、「重さ・大きさ」などの物理パラメータと、「人間らしさ」などの属性パラメータは、個別に制御可能であるという重要な設計指針を与えるものである。例えば、本実験において「ロボット」条件と「ブーツ」条件はいずれもPC1で高い値を示したが、PC2では両極端に位置した。この知見は、今後「重厚な鎧を着た騎士(人間)」や「巨大なゴーレム(非人間)」といった未知のアバターを表現する場合において、汎用的な設計プロセスを示唆する。すなわち、まずは低音強調によって「物理的な巨大さ」を規定し、その上で「音の余韻やリズムのゆらぎ」といったテクスチャを操作することで生物か非生物かを書き分けるなどといった、2段階の音響設計アプローチが有効であると考えられる。

4.5.3 音響的な「らしさ」の吟味と聴覚的なデフォルメの必要性

「革靴」条件において生じたジェンダーイメージの混同は、リアルな環境音と、人々が抱くイメージの間にズレが生じる場合があることを示している。本実験において、「ヒール」条件などは高音域を強調する加工を行っているものの、基本的には現実の音に近い特性を持っており、それで十分に女性的な印象が伝達された。これは、現実の音が持つ特徴が、社会的なステレオタイプと合致していたためである。しかし一方で、「革靴」条件のように、現実の音が、意図せぬ別のステレオタイプ(ヒール)を想起させてしまうケースも確認された。これは、「リアルな音を使えば正解」という単純な図式が成立しないことを意味する。したがって、視覚障がい者向けの足音を設計する際には、まず選定した音が一般的なイメージと乖離していないかを吟味

する必要がある。その上で、革靴の例のように紛らわしい成分が含まれる場合には、現実の再現性にこだわらず、誤認される要素を削ぎ落とし、特徴的な成分を誇張する「聴覚的なデフォルメ」を適用するという、柔軟な設計プロセスが求められる。

4.5.4 親しみやすさと安心感のあるデザイン

「マスコット」条件が示した高い親しみやすさは、足音のデザインが、単なる「属性の識別(誰がいるか)」を超えて、「相手に対する心理的障壁の低減(どう感じるか)」に寄与できることを示している。視覚障がい者がソーシャルVRに参加する際、相手の姿が見えないことによる「威圧感」や「不安感」は大きな課題である。本研究の結果は、足音のピッチや音色を調整し、「小ささ」や「柔らかさ」を強調することで、視覚的な愛らしさを聴覚的に翻訳し、他者とのコミュニケーションを促進させられる可能性を示している。これにより、将来的にはユーザーの心理状態に合わせて他者の足音印象を動的に変化させるといった、VRならではのアクセシビリティ支援への応用が期待される。

5 おわりに

本研究では、視覚障がい者がソーシャルVRにおいて他者のアバターが持つ視覚的特徴を直感的に把握することを目的とし、アバターの移動に伴い発生する足音表現によって想起されるアバターイメージについて検討を行った。

足音の質感を強調・デフォルメした表現が聴取者にどのような人物像を想起させるのかを明らかにするため、複数種の足音を用いた聴取実験を実施した。実験の結果、足音の違いによって体格感や非人間的・女性的といった印象等に一定の傾向が確認された。また、足音が与える印象が「物理的な強度」と「人間属性の有無」という独立した2つの軸に分類可能であることが示され、足音表現がアバターイメージの形成に影響を与える要因となり得ることが示唆された。さらに、言語的観点から足音のイメージ構造を可視化することで、確認された印象傾向についての補足を得られた。一方で、一部足音については想起されるイメージにばらつきが見られ、単なる現実音の再現では一貫した人物像を提示することが難しいことも明らかとなった。このことから、現実の再現以上に特徴を意図的に強調する聴覚的なデフォルメが有効である可能性が示唆された。

今後は、実際の視覚障がい者を対象とした評価実験を行うとともに、足音表現によるアバターイメージ提示の可能性について、条件設定や評価観点を拡張した検討が必要である。特に、足音の質感や印象の違いが人物像の想起に与える影響について、より多様な刺激条件を用いた検証を行うことで、本研究で得られた知見の汎用性を高めることが今後の課題である。

謝 辞

本研究の一部は、京都産業大学先端科学技術研究所(人間情報学研究センター)共同研究プロジェクト(M2301)、京都産業大学情報理工学部デジタル制作・研究共創チャレンジプログラムの助成を受けたものである。ここに記して謝意を表す。

文 献

- [1] 厚生労働省. 令和 4 年生活のしづらさなどに関する調査 (全国在宅障害児・者等実態調査). 2024 年 12 月 18 日確認.
- [2] Rafael Jeferson Pezzuto Damaceno, Juliana Cristina Braga, and Jesús Pascual Mena-Chalco. Mobile device accessibility for the visually impaired: problems mapping and recommendations. *Universal Access in the Information Society*, 17:421–435, 2018.
- [3] Roberto Vaz, Diamantino Freitas, and António Coelho. Blind and visually impaired visitors’ experiences in museums: increasing accessibility through assistive technologies. *The International Journal of the Inclusive Museum*, 13(2):57, 2020.
- [4] Venkatesh Potluri, Priyan Vaithilingam, Suresh Iyengar, Y Vidya, Manohar Swaminathan, and Gopal Srinivasa. Codetalk: Improving programming environment accessibility for visually impaired developers. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–11, 2018.
- [5] Konstantinos S Papadopoulos and Dimitrios B Goudiras. Accessibility assistance for visually-impaired people in digital texts. *British Journal of Visual Impairment*, 23(2):75–83, 2005.
- [6] Neveen I Ghali, Omar Soluiman, Nashwa El-Bendary, Tamer M Nassef, Sara A Ahmed, Yomna M Elbarawy, and Aboul Ella Hassanien. Virtual reality technology for blind and visual impaired people: reviews and recent advances. In *Advances in Robotics and Virtual Reality*, pages 363–385. Springer, 2012.
- [7] Yuhang Zhao, Cynthia L Bennett, Hrvoje Benko, Edward Cutrell, Christian Holz, Meredith Ringel Morris, and Mike Sinclair. Enabling people with visual impairments to navigate virtual reality with a haptic and auditory cane simulation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2018.
- [8] Catherine Todd, Swati Mallya, Sara Majeed, Jude Rojas, and Katy Naylor. VirtuNav: A virtual reality indoor navigation simulator with haptic and audio feedback for the visually impaired. In *2014 IEEE Symposium on Computational Intelligence in Robotic Rehabilitation and Assistive Technologies (CIR2AT)*, pages 1–8, 2014.
- [9] Michael Zyda. From visual simulation to virtual reality to games. *Computer*, 38(9):25–32, 2005.
- [10] Statista. Metaverse - worldwide. 2024 年 12 月 25 日確認.
- [11] 総務省. 令和 6 年度版情報通信白書. 2024 年 12 月 25 日確認.
- [12] 中島 伸介 横山 皓祐, Panote Siriaraya. 仮想現実空間における視覚障がい者向け交流支援システムの構築. In **第 17 回データ工学と情報マネジメントに関するフォーラム (DEIM2025)**, pages 1G–01, March 2025.
- [13] Crescentia Jung, Jazmin Collins, Ricardo E. Gonzalez Penuela, Jonathan Isaac Segal, and Andrea Stevenson Won. Accessible nonverbal cues to support conversations in vr for blind and low vision people. In *Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility (ASSETS)*, 2024.
- [14] Sünje Clausen, Ana Tajadura-Jiménez, et al. Auditory cues and social perception: How sound shapes impressions of others. *Frontiers in Human Neuroscience*, 2021.
- [15] Qihan Sun, Lu Chen, Minshi Fu, and Linzi Chang. Influence of footstep sounds induced by varying shoe types and gait frequencies on initial impressions. In *Proceedings of the AHFE International Conference on Physical Ergonomics and Human Factors*, 2025.
- [16] Stuart Cunningham and Iain McGregor. Audience perceptions of foley footsteps and 3d realism designed to convey walker characteristics. *Personal and Ubiquitous Computing*, 2024.
- [17] Aimee Allen, Tom Drummond, and Dana Kulić. Robots have been seen and not heard: Effects of consequential sounds on human perception of robots. *IEEE Robotics and Automation Letters*, 2025.
- [18] Zapsplat. Free sound effects & royalty free music. <https://www.zapsplat.com/>. 2026 年 1 月 4 日確認.
- [19] DOVA-SYNDROME. フリー bgm (音楽素材) 無料ダウンロード. <https://dova-s.jp/>. 2026 年 1 月 4 日確認.
- [20] Meta. Meta xr audio sdk. https://developers.meta.com/horizon/documentation/unity/meta-xr-audio-sdk-unity/?locale=ja_JP. 2026 年 1 月 4 日確認.
- [21] Audacity Team. Audacity(r): Free, open, cross-platform audio software. <https://www.audacityteam.org/>. 2025 年 1 月 4 日閲覧.