



HMD を用いたメタバースにおける 聴覚的な擬似気配の提示の試み

木下栞¹⁾, 舟橋健司²⁾

1) 名古屋工業大学 (〒 466-8555 名古屋市昭和区御器所町, s.kinoshita.385@stn.nitech.ac.jp)

2) 名古屋工業大学 (〒 466-8555 名古屋市昭和区御器所町, kenji@nitech.ac.jp)

概要: 本研究では, メタバースにおいて HMD のヘッドフォンを用い, ユーザーの後方などにいる他者の気配を聴覚的に提示する手法の提案と評価を行った. 片側からの音の周波数を短時間だけ変化させることで他者の存在を無意識に感じさせることを試みた. 体験者による評価では, 適切な変化において, その変化自体を明確に感じることなく気配のようななにかを感じ取れる傾向が確認され, 提案手法による気配提示の有効性が示唆された.

キーワード: 気配, HMD, メタバース

1. はじめに

VR 技術は, テーマパークや家庭用ゲームのみならず, 医療, 建築, 教育など多様な分野で応用が広がっている. それに伴い, 視覚や聴覚を中心とした没入型体験を可能にするメタバースも注目を集めており, HMD (ヘッドマウントディスプレイ) はこれらの体験において中心的な役割を果たしている. メタバースの体験では, 主に HMD が利用されるが, 場合によっては PC モニタやスピーカが使われることもある. HMD の種類や PC モニタのサイズによって異なるものの, ユーザの視野はおおよそ 110 度程度に限定される. この狭い視野のため, 後方はもちろん, 横から他の参加者が近づいてきても視覚的に気付くことはできない. メタバース環境であればシステムが周囲の情報を把握しているため, 接近情報を音声で通知することも可能である. しかし, このような手段は情報の量が多くなりがちで, ユーザが目の前のタスクやコミュニケーションに集中できなくなる恐れがある. そこで「気配」のような不確実な情報を提示することができれば, 目の前のタスクを大きく妨げることなく, 周辺情報を適度に把握できると考えられる. 気配とは, 明確に見えたり聞こえたりしないものの, 周囲の状況から漠然と感じ取る様子を指す. 気配の主な要因としては, 物や人の影による視覚情報, 足音などの小さな音による聴覚情報, 風や空気の動き, 準静電界 [1] などによる触覚情報などによると考えられている. ただし, これらの情報を明確に知覚している場合は, 気配とは表現されない. 様々な気配の人工的な再現の試みが行われている. 文献 [2] では水蒸気を利用している. 文献 [3] ではブラウン管テレビが発する準静電界を利用して体毛を刺激している. 文献 [4] では他者などの接近音の 3D 音場を再現できる没入型聴覚ディスプレイ「音響樽」を提案し, 気配の提示を試みている. これらのような方法はバーチャルリアリティやメタバースへの導入も可能であり有効だろう. しかし, これらは特別な機器

や環境を必要とするため, 家庭用のメタバース環境での導入は困難である. 他者などによる環境音の遮蔽を想定して, 可聴音の低周波領域を強調し, また高周波領域を抑制することによる気配表現の試みも報告されている [5]. しかし, 体験者が意識下において可聴音の変化に気づいていないことを前提としておらず, 明確な「合図」と受け取っていたかもしれない. 当研究室ではこれまでに, 特別な機器や環境を必要とせず, 視覚提示機能のみを利用して, 気配のような不確実な情報提示手法を新たに提案している [6]. 気配自体が科学的に定義できているわけではないが, ここではいわゆる気配そのものではなく, 視覚提示により体験者が「気配のようだ」と感じる何かを提示できる可能性について確かめている. 例えば視野外に人が現れた場合に, 薄い人型のシルエットを短い時間だけ重畳表示する. 体験者の意識下においては見えたとは知覚されないような薄さ, 短さにした上で, 無意識に何かを感じ取れることを目指している. ところで, 薄いシルエットを表示する場合には, 背景の色や模様などに大きく影響を受けるだろう. 背景に合わせてシルエットの薄さや提示時間を自動的に調節することも考えられるが, 背景によっては, 意識下においては知覚されずに無意識に何かを感じ取れるような, 適切なパラメータが存在しないかもしれない. そこで本研究では, 前述の手法と相互に補完し合えるような, 視覚に頼らない, 標準的な聴覚提示機能のみを利用する気配の情報提示手法を新たに提案し, 有効性を確認する. 例えば, 背後などの視野外に人が現れた場合に, 短い時間だけ音を提示したり提示中の音を変化させたりする. 体験者の意識下においては聞こえた, あるいは変化したとは知覚されないような小ささ, 短さにする必要がある. 視覚を対象とした場合と同様に, 無意識に何かを感じ取れる必要もある. 体験者による評価を行い, この擬似気配提示手法の有効性や, 擬似気配を感じられるような音の変化や短さについて検討する. 具体的には,

片側からの音の周波数を短時間だけ変化させることで他者の存在を無意識に感じさせることを試みた。適切な変化において、その変化自体を明確に感じることなく気配のようななにかを感じ取れる傾向が確認され、提案手法による気配提示の有効性が示唆された。

2. 擬似気配の提示方法

人体の感覚は、外部や内部からの刺激を受け取り、それを脳に伝えることで成立する知覚である。これらの感覚を担う感覚器官や神経系は、それぞれ特定の刺激に対して応答するが、あらゆる強度の刺激に反応するわけではなく、また刺激の違いをすべて正しく感じ取れるわけではない。まず、刺激が感覚を生じさせるために越えなければならない最低限の強度を「刺激閾」という [7]。例えば聴覚の場合、周波数により一定の音の強さ以上でなければ音として感知されない。この感知可能な最小の音の強さ（音圧レベル）が刺激閾である。この仕組みにより、感覚器官は不要な刺激に反応せずすむ。また感覚の弁別には、ある刺激とそれに近い別の刺激を区別するために必要な最小の差が存在する。この最小差を弁別閾という [8]。例えば聴覚の場合、音の強さが異なる場合に、弁別閾を超えた差であれば、2つの音が異なると知覚される。その差が弁別閾を超えない場合、音は同一であると知覚される。なお、感覚器官では感知しているものの、意識レベルでは認識されないこともあり、この境界を意識閾という。さらに、感覚には時間的な解像度も存在する。短い時間間隔で生じた2つの刺激が、実際にはわずかに時間差があったとしても、一定の範囲内であれば同時に生じたものとして知覚される。この時間幅を同時性の窓と呼ぶ [9]。刺激閾付近、例えば聴覚においては完全に「聞こえる」状態と全く「聞こえない」状態の間に、曖昧な中間状態が存在しているかもしれない。弁別閾付近では、2つの刺激の違いをはっきりと知覚できる場合と、ほとんど区別がつかない場合の中間的な感覚が存在するかもしれない。同時性の窓付近でも、同様である。このような意識閾付近を利用して閾下知覚を促すことで、あたかも何か近づいている、あるいは存在しているかのような擬似的な気配の提示を実現できるのではないかと考えた。ところでスピーカでは音圧レベル、インピーダンス、周波数特性は重要な仕様であり、これらの物理的特性は再生音の聞こえ方に影響を与える。特に、再生可能な周波数帯域は音量によって大幅に狭まることもあるため、微細な変化を適切に再現し、知覚させるためには、これらの特性を考慮して周波数と音量を設計する必要がある。本研究では、安価なスピーカでは再生できないかもしれない周波数の音や微弱な音量は想定せずに、HMDに内蔵されたヘッドフォンにBGMをステレオ（左右2チャンネル）で提示していることを前提に、左右どちらかのチャンネルの音をごく短時間変化させることで気配のようななにかを感じさせることを考えた。具体的には、以下の3つの手法を検討した。

1. 音量をわずかに増加，または減少させる。

- 1'. 文献 [5] を参考に、低周波領域の音量をわずかに増加し，また高周波領域の音量を減少させる。
2. 周波数をわずかに上昇，または下降させる。
3. 発声時刻をわずかに遅らせる。

それぞれの手法を少人数に対して、様々な提示時間において予備的に評価した。いずれも、左右のうち変化した側に気配を感じることを期待していた。手法1では、音量を増減したときに「明確に違いが分かる」という意見と「違いが分からない」という意見の2つにはっきりと分かれた。人は音量差を手がかりに音源の方向を判断するため、一方の音の音量を増減すると、変化した側に限らず、音が大きい側に音源があると感じた。また、音量の変化により音質が変わるように感じた。手法1'でも同様に2つの意見に分かれた。なお、文献 [10] では400Hz以下の低周波領域を+10dB強調し、10000Hz以上の高周波領域を-10dB抑制していた。そこで手法1'では低周波領域を400と600Hz、高周波領域を8000と10000Hzの、それぞれの組み合わせ計4パターンに対して、さらに±3,5,10dBの複数の加工を試した。我々の予備評価では、10dBの加工では明らかに音が変わったことに気づいた。変化が3または5dBの加工においては提示時間により前述の通り2つの意見に分かれた。また低周波領域の設定が400Hzの場合に耳鳴りのような不快感を感じることもあった。スピーカやヘッドフォンの違いを考慮する必要もあるが、この予備調査の結果を踏まえると、文献 [5] の手法では体験者は音の変化に気づいてしまうかもしれない。手法2では、周波数を昇降したときに「明確に違いが分かる」という意見、「違いが分からない」という意見、に加えて「よく分からない」という意見もあった。また、左右での周波数差により音源の方向について特段の知覚はなかった。手法3では手法1と同様に、発声時刻をわずかに遅らせたときに「明確に違いが分かる」という意見と「違いが分からない」という意見の2つにはっきりと分かれた。また、遅れた側とは反対側から音が聞こえるように感じることもあった。音が片方の耳に遅れて到達すると、脳はその時間差を手がかりに音源の方向を判断するためだと考えられる。定位が不自然になる状態を「不快」と感じることもあった。これらの結果により、手法1と3は「中間状態」があまり期待できなさそうである。さらに手法3は不快感による聴覚疲労が心配である。手法2は右（後方）と左（後方）の気配を区別することは難しそうであるが、違うのか違わないのかよく分からない、という意見が疑似的な気配の実現に最も近そうである。このような聴覚の特性、スピーカの特性、および予備評価の結果を踏まえたうえで、聴覚を通して疑似的な気配を知覚させたい。ステレオ2チャンネルで常に音楽を提示している状態で、一方のチャンネル（右か左）のみ、明確に違いが分からない程度の周波数の変化量と提示時間だけ、音楽を部分的に加工して提示することで、聴覚を通じた擬似気配の提示を試みる。



図 1: 体験の様子



図 2: HMD に投影される風景

3. 評価

提案手法を, Meta Quest 3 を用い, Unity 上で構築した仮想環境において評価した (図 1). 提示すべき情報は聴覚情報だけであるが, 評価時にスピーカを使用すると, 外部の音が影響しやすく, 結果に悪影響を与える可能性がある. ヘッドフォンを使用すれば, 完全な遮音でなくても外部ノイズの軽減が期待できるだろう. また, 体験参加者が過度に音の変化に集中してしまうことは避けたい. そこで, 参加者には, 聴覚だけに過度に集中しないよう, 視覚的サブタスク (スリーセルゲーム) を同時に実施してもらった (図 2). HMD に投影される風景はビデオシースルー機能により実際の体験環境が表示されるが, 壁のみのシンプルなものに限定し, テーブルと 3 つのカップは CG で描画した. HMD の視野角は約 100 度であり, 窓や他の物体が映り込まない位置で体験することを想定した. BGM はポップカントリー調の音楽とし, いずれかのチャンネルの音の周波数を一定時間わずかに変化させた. 周波数差 δ を +0.5Hz から +4.0Hz までの 0.5Hz 間隔で, 変化時間 t を 1 から 5 秒までの 1 秒間隔で設定し, 合計 40 パターンを用意した. 参加者は大学生 (大学院生を含む) 10 人であった. 各参加者に 40 パターン全てを 1 回ずつ, ランダムな順に割り当てた. 全試行数は 400 回だった. 試行中に「気配を感じた」と思ったタイミングでボタンを押すよう指示し, 各試行後に音の変化を認識したかどうか, すなわち音の変化自体に気づいたかど

うか報告してもらった. 気づかなかった場合には, 気配 (のようになにか) を感じたかどうか報告してもらった. なお予備調査において, 方向について特段の知覚はなかったが, 気配を感じた場合には参考としてその方向 (左右どちらか) についても報告してもらった. 各試行の回答を以下の 4 つのカテゴリーに分類した:

- *A*: 音の変化自体に気づいた (audible)
- *P*: 気配を感じる時刻が正しかった (音の変化には気づかず, 気配を感じた) (presence)
- *F*: 気配を感じる時刻が間違っていた (音の変化には気づかず, 気配を感じたと報告したが, 時刻が違った) (false)
- *N*: 気配は感じられなかった (音の変化にも気づかなかった) (not felt)

全ての組み合わせにおける回答数を表 1 に示す. 擬似気配を感じてもらうためには, まず音の変化に気づいてしまっはならない. その上で, 感じた気配のタイミングが正しいことを期待したい. そこで *A* と *P* に注目する. なお本評価では, 気配は常に感じられるわけでも正しく認識されるわけでもないため, *N* および *F* の数は考慮しなかった. 参加者には気配を感じている間中, ボタンを押し続けてもらったが, 特にボタンを押した最初の時刻のみに注目した. すなわち, *P* は周波数が変化している時間の開始時刻以後, かつ少し遅れて気配を感じることも考慮して終了時刻の 3 秒後までにボタンを押したことを意味している. ボタンを離れた時刻については考慮しなかった. 気配提示を主目的とし, その方向は重視しておらず, また予備調査からも周波数が変化した側としなかった側のいずれに気配を感じるのか, そもそも不明であったため左右の方向についての報告も考慮しなかった. なお, サブタスクであるスリーセルゲームの正答率は 93.5% であった. 体験者はサブタスクに集中していたと考えられる. 各 δ および t の組み合わせにおける *P* と *A* の差分を図 3 に示す. 全体として, *P* は $(\delta, t) = (+2.0, 1)$ から $(+0.5, 5)$ の範囲でわずかに *A* を上回り, とくに $(\delta, t) = (+2.0, 3)$ で顕著に上回った. 一方で, +3.0 Hz 以上では音の変化に気づく傾向が強くなることが分かった. 比較的, 周波数差が大きい場合には変化時間を短くすることで提示前後による瞬間的な変化と捉えられ, 反射的な知覚を促されたのだろう. 周波数差が特に小さい場合, 短い変化時間では意識下で知覚されにくい, 変化時間を少しだけ長くすることで無意識下での知覚を促されたのだろう. 今回の評価では, 明らかな有効性を示すことはできなかったが, 周波数差と変化時間 (δ, t) が $(+2.0, 3)$ を含む $(+2.0, 1)$ $(+0.5, 5)$ を中心とする帯状のあたりにおいて, 音の周波数を変化させることが擬似的な気配の提示に有効であると期待できる.

4. むすび

より没入感のあるメタバース体験に貢献するため, 標準的な HMD の聴覚提示機能 (ヘッドフォン) のみを用いて

表 1: 各周波数差 δ , 変化時間 t に対する A, N, P, F の表 ($A/N/P/F$)

$\delta \backslash t[s]$	1	2	3	4	5
+ 0.5	0/10/ 0/ 0	1/ 8/ 0/ 1	1/ 9/ 0/ 0	1/ 9/ 0/ 0	0/ 8/ 1/ 1
+ 1.0	1/ 9/ 0/ 0	0/ 9/ 1/ 0	2/ 6/ 0/ 2	6/ 1/ 2/ 1	4/ 5/ 1/ 0
+ 1.5	0/ 8/ 1/ 1	0/ 9/ 0/ 1	1/ 7/ 1/ 1	1/ 6/ 1/ 2	3/ 6/ 1/ 0
+ 2.0	0/ 9/ 1/ 0	1/ 7/ 1/ 1	1/ 5/ 4/ 0	1/ 8/ 0/ 1	2/ 7/ 1/ 0
+ 2.5	1/ 9/ 0/ 0	8/ 1/ 1/ 0	3/ 7/ 0/ 0	3/ 4/ 1/ 2	4/ 4/ 2/ 0
+ 3.0	4/ 5/ 1/ 0	5/ 3/ 1/ 1	10/ 0/ 0/ 0	3/ 6/ 1/ 0	10/ 0/ 0/ 0
+ 3.5	4/ 3/ 1/ 2	10/ 0/ 0/ 0	7/ 2/ 0/ 1	0/10/ 0/ 0	3/ 4/ 2/ 1
+ 4.0	9/ 1/ 0/ 0	8/ 1/ 0/ 1	10/ 0/ 0/ 0	1/ 8/ 0/ 1	7/ 0/ 0/ 3

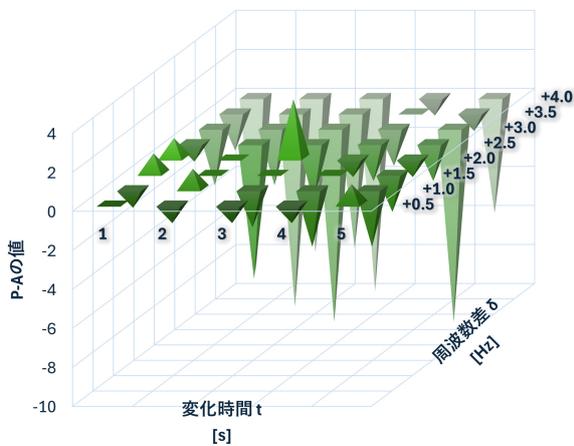


図 3: P と A の差 ($P-A$) のグラフ

擬似的な気配を提示する手法を提案し、その有効性を評価した。ステレオ 2 チャンネルの BGM が流されていることを前提に、一方のチャンネル (右か左) のみ音の高さ (周波数) を短時間だけ変更して提示することで、聴覚を通じた擬似気配の提示を試みた。様々なパターンの、変化させる周波数の差や変化させている時間の組み合わせに対して、効果を検証した。一部の周波数差と変化時間の組み合わせにおいて、音の変化自体に気づかず、しかし気配のようなものを感じて、さらにその時刻が聴覚提示の変更時刻に一致していた試行数が多かった。本手法による気配提示の可能性が示唆された。一方で、ある特定の変化時間において、周波数差の大きさに関わらず多くの試行で変化自体が気づかれており、結果の一部に誤差の可能性も疑われた。今後は、この条件について再調査を行い、結果の信頼性を確認する必要がある。また音量変化や遅延による気配の提示手法も再検討したい。

参考文献

[1] 滝口清昭, 遠山茂樹: “犬は主人を電界で見分ける?—歩行による人体の電界発生とその伝搬—”, 国際生命情報科学会誌, Vol.21(2), pp. 428–441, 2003.
 [2] K. Hokoyama, Y. Kuroda, K. Kiyokawa, H. Takemura: “Mugginess sensation: Exploring its princi-

ple and prototype design”, In Proceedings of IEEE World haptics, pp. 563–568, 2017.
 [3] K. Suzuki, K. Abe, H. Sato: “Proposal of perception method of existence of objects in 3d space using quasi-electrostatic field”, In Proceedings of International Conference on Human-Computer Interaction, 2020.
 [4] 伊勢史郎: “聴覚的リアリティを実現する音のデザイン—ヒトとモノの動きがもたらすリアリティの本質—”, 日本音響学会誌, Vol.74(11), pp. 598–602, 2018.
 [5] S. Zhao, A. Ishii, Y. Kuniyasu, T. Hachisu, M. Sato, S. Fukushima, H. Kajimoto: “Augmentation of acoustic shadow for presenting a sense of existence”, In Proceedings of International Workshop on Modern Science and Technology, pp. 563–568, 2012.
 [6] K. Funahashi, Y. Matsunoo: “A visual approach to pseudo unseen presence in the metaverse using HMD”, In Proceedings of International Conference on Artificial Reality and Telexistence, 2023.
 [7] 岡本康秀, 神崎晶, 井上泰宏, 齋藤秀行, 貫野彩子, 久保田江里, 中市健志, 森本隆司, 原田耕太, 小川 郁: “周波数選択性測定臨床応用”, Audiology Japan, Vol.55, pp. 642–649, 2012.
 [8] 岡本安晴: “弁別閾と感覚尺度”, 基礎心理学研究, Vol.12(1), pp. 60, 1993.
 [9] 一川誠, 政倉裕子: “能動的観察と視聴覚刺激の同時性”, 日心第 70 回大会, 2006.
 [10] S. Zhao: “Augmentation of Acoustic Shadow for Presenting a Sense of Existence”, 電気通信大学大学院修士論文, 2012.