

VR コンテンツにおけるマルチモーダル情報の質と鑑賞者の 視行動パターン・VR コンテンツ記憶の間の関係分析

Relationship Between the Level of Sensory Stimuli in VR Content and the Degree of Immersion Based on Eye Movement Analysis

中川 太一[†]
Taichi Nakagawa

T.L.A.N. Chandrasiri[†]
T.L.A.N. Chandrasiri

小川 信之[‡]
Nobuyuki Ogawa

矢島邦昭[§]
Kuniaki Yajima

北島 宗雄[†]
Muneo Kitajima

中平 勝子[†]
Katsuko T. Nakahira

1 はじめに

人は、感覚器官で周囲の環境から絶えず知覚情報を受け取り、それらを処理することで情報を獲得し、記憶を形成する。morenoらはこのプロセスを、感覚記憶に入力された知覚情報が注意によって選択され、ワーキングメモリで整理・統合された後、長期記憶に保存されるという一連の流れであると説明した[1]。北島らは、全方位映像の鑑賞時に音声ガイドを付与することで、視覚情報と聴覚情報が適切に統合され、体験が記憶に残りやすくなることを示した。[2] この先行研究は、映像コンテンツを鑑賞する際の情報獲得行動を対象としている。この状況は、いわば一方的に提示される知覚情報を処理し、それらを記憶するといった形になっていると言える。しかし、観光や買い物など、人間が現実の物理世界で情報を獲得する際は、興味を持った対象に人間が自ら能動的に働きかけなければ情報を獲得することはできない。そこで本研究では、人間が物理空間において能動的な情報獲得行動をとる際、知覚情報のクオリティが情報記憶に与える影響を調べる。しかし、物理空間で外界の知覚情報のクオリティを変化させることは容易ではないため、提示刺激を自由に制御でき、かつ被験者に能動的な行動を取らせることができる実験環境として、VRを使用する。なお本稿では、視覚情報のクオリティの高さをVRコンテンツの情報量の多さと考え、VR空間に設置する3Dオブジェクトの量で定義する。聴覚情報のクオリティはコンテンツ内音声ファイルの音量、音質で定義する。

2 仮説

次に、知覚情報のクオリティが被験者の情報獲得行動に与えると考えられる仮説を述べる。VR空間の視覚クオリティが高い場合、VR空間の情報量は増えるが、配置されているオブジェクト数が多いため被験者が注意を向ける対象が多くなるため、コンテンツ設計者がターゲットとして想定していたオブジェクトに向けられる注意が減少し、結果として覚えるべき内容が記憶に残らないということが考えられる。一方で、ターゲットとなるオブジェクトの周辺に存在するオブジェクトの存在が記憶

のフックとなり、コンテンツに関する情報が想起されやすくなることも考えられる。また、視覚クオリティが低い場合は、VR空間の情報量が過度に低くなることで現実らしさや臨場感が感じられず、VRコンテンツの内容が被験者の印象に残らないということが考えられる。一方で、周りに余分なオブジェクトが存在しないことで被験者が自然と注意を向けるべき箇所に集中でき、結果としてコンテンツの情報をより記憶できるといったことも考えられる。このように、視覚情報のクオリティと記憶には一定の関係が存在することが予想される。

聴覚情報に関しては、クオリティが高いほど情報記憶に適していると考えられる。VRコンテンツの音声が過剰に大きな音量で提示された場合や音質が悪い場合、純粋に聴覚情報そのものの処理が困難になる。そして、適切でない聴覚情報が付与されると認知負荷が大きくなってしまい、認知リソースが聴覚情報に多く割かれるため、視覚情報の処理が阻害されることで、視覚情報も記憶に残りづらくなるという状況も考えられる。

以上の仮説を検証するため、複数の見た目のVR空間と、複数の音質の音声コンテンツを含むVRコンテンツを作成し、実験を行った。

3 実験

実験方法は[3]に基づき、使用するVRコンテンツに変更点を加えた。

HMDはMeta Quest 2を使用し、視線計測機はTobii Pro Glasses 3を使用し、実験環境のVR空間はUnityを用いて作成した。VR空間の概要を図1に示す。視覚情報についてはVR空間に配置する3Dオブジェクトの数によってクオリティを定義し、高い(H)、普通(M)、低い(L)の3種類のVR空間を作成した。なお、視覚クオリティと神社のモデルの種類は組み合わせに偏りが無いよう選択し、カウンターバランスをとった。聴覚情報については解説音声の音量、音質を変数とした。表1に各クオリティの視覚、聴覚情報の詳細を示す。

本実験におけるVRコンテンツは、VR空間で神社の建築様式を観察するコンテンツを使用した。被験者はVR空間内を自由に動き回ることができ、神社付近に配置されているスピーカーのアイコンをトリガーすることで、神社に関する解説音声再生される。被験者が3つのスピーカーを探し出し、全ての

[†] 長岡技術科学大学

[‡] 岐阜工業高等専門学校

[§] 仙台高等専門学校

表 1 知覚情報のクオリティ

	低い (L)	普通 (M)	高い (H)
視覚情報	平面に神社だけが置かれている	平面に神社、木、鳥居が置かれている	草、木、遠景、鳥居、手水舎が置かれている
聴覚情報	電話越しのような音質	通常の音声	音量が過剰に大きい

解説音声を再生すると、次の空間へ移動するためのボタンが出現するようになっていく。なお、VR 酔い対策のため VR 空間内での移動はテレポート移動とした。また、被験者は HMD を装着する際、視線計測デバイスも同時に装着し、VR コンテンツ体験時の視線情報を計測した。

実験開始前に、チュートリアル空間を用意し、実験の概要を説明し、VR 空間内での操作を練習する時間を設けた。そして、コンテンツ体験後、被験者に対し記憶想起課題を行い、VR コンテンツの内容で覚えていることを筆記と口頭にて説明させた。

実験開始前の説明では、被験者には、実験後にコンテンツの内容に関するアンケートに答えてもらうとだけ説明した。理由は、実験後に想起テストを行うことを事前に伝えることで被験者が意識的にコンテンツの内容を記憶しようと行動することを避けるためである。

4 結果

3 名分のデータを分析した。全ての被験者の記録に対して、VR コンテンツ体験中に注視していたオブジェクトを時系列順に示すタイムラインを作成した。一例として、図 2 に 1 名の被験者のタイムラインを示す。以下に、実験中の HMD 映像と想起課題の回答からなる行動データと、視線データについてそれぞれ結果をまとめる。

4.1 行動データ

3 名分のデータに共通して見られた特徴的な行動を表 2 に示す。まず、神社の概形を図で想起していたのは、視覚情報のクオリティが高い時に 2 名、中程度の時に 3 名、低い時に 2 名であった。視覚情報のクオリティが高い、もしくは中程度のとき、3 名中 2 名が手水舎や鳥居などのオブジェクトに近づいて観察しており、その 2 名は想起課題にてそれらのオブジェクトについて言及していた。そして、3 名中 2 名が地面や遠景などの周りの環境の情報量の違いに言及していた。また、3 名全員が住吉造の室内の構造について想起しており、内 1 名は聴覚情報のクオリティが低い条件だったが、解説音声の内容と 3D モデルの構造の両方を記憶していた。3 名が神明造の高床式の構造を想起しており、内 1 名は聴覚クオリティが低い条件で視聴していた。また、2 名が大社造の高床式の構造を想起しており、内 1 名は聴覚クオリティが低い条件で視聴していた。

4.2 視線データ

本稿では、各視覚クオリティ条件における 1 秒間あたりの平均サッカード回数、3 秒間隔のサッカード発生回数の時系列分

表 2 被験者に共通して見られた行動データ

内容	人数
神社の概形を想起していた (視覚クオリティ: High)	2/3
神社の概形を想起していた (視覚クオリティ: Middle)	3/3
神社の概形を想起していた (視覚クオリティ: Low)	2/3
視覚クオリティが高いときと普通るとき、周りの環境に実体のあるオブジェクトに興味を示し、観察する	2/3
オブジェクトに実際に近づいて観察した上で、記憶が想起されている	2/3
視覚クオリティが低いとき、周りの環境に興味を示すことなく、スピーカーを探すことに集中していた	3/3
想起課題で、周りの環境の視覚的な情報量が多かったことに言及している	2/3
住吉造の部屋の構造を想起していた	3/3
神明造の構造 (奥行きより横幅が足掻く、長方形の形をしている) を想起していた	2/3
神明造の高床式の構造を想起していた	3/3
大社造の高床式の構造を想起していた	2/3

表 3 視覚クオリティ別の平均サッカード回数 (回/秒)

被験者	視覚情報クオリティ		
	H	M	L
05	5.7	6.8	7.2
20	8.4	7.5	7.2
35	8.9	8.6	9.1

布、各視覚クオリティ条件内での主要イベントにおける視線停留時間とサッカードのそれぞれの合計時間割合を視線データとして使用する。表 3 に、視覚クオリティ別の、毎秒あたりの平均サッカード回数を示す。被験者 05 と 35 は視覚クオリティが低い時にサッカード回数が最大となっており、被験者 20 は視覚クオリティが高い時にサッカード回数が最大となっている。

次に、視覚クオリティ別サッカード回数の時系列分布を図 3 に示す。縦軸は 3 秒間隔の平均サッカード発生回数の増減を表している。本稿では、一例として 1 名分のデータを示す。

行動データと照合した結果、以下の場面でサッカード回数が増加していた。

- 神社の正面から側面を通して背面まで移動し、スピーカーを発見するまで
- 神社外部から神社の中に入ってスピーカーを発見するまで
- 解説音声再生中、最後まで聞かずに途中で次の行動へ移行するとき

また、スピーカーを発見し、解説音声の再生が始まるとサッカードの回数が減少していた。

最後に、各視覚クオリティ毎に主要な行動イベントを切り出し、イベントの長さに対する各イベント中の合計視線停留時間の割合を横軸、合計サッカード時間の割合を縦軸としたグラフを図 4 に示す。切り出したイベント数は、視覚クオリティが高い、中程度、低い場合でそれぞれ 19 個、15 個、13 個である。

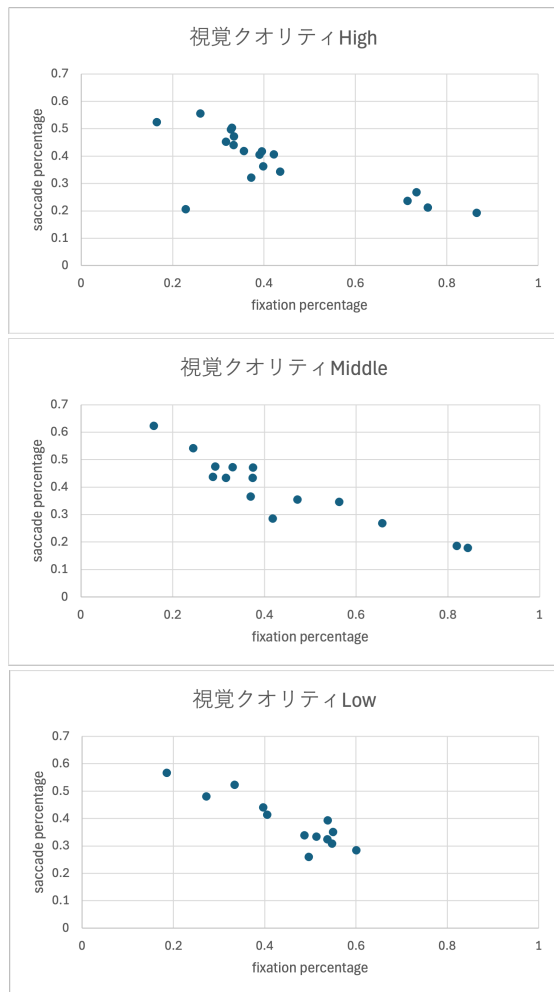


図 4 各視覚クオリティの主要イベントにおける視行動割合

また、図 3 から、サッカード回数の時系列分布の中で値が大きく増減している箇所が見られる。この箇所に関して被験者の映像データと照らし合わせたところ、スピーカーを探しているときにサッカードの回数が増加する傾向が見られた。特に、神社背面の高い位置に配置したスピーカーや、神社内部に入ってスピーカーを探す場面において顕著であった。本実験で被験者に与えた指示は「配置されている 3 つのスピーカーを見つけて再生し、解説音声聞く」というものであり、スピーカーはいわば目標達成のためのターゲットであったことから、このように探索行動が顕著に視線データに表れていると考えられる。また、スピーカーを発見した後狙いを定めるために一箇所を注視することでサッカード回数は減少した。さらに、解説音声を聞きながら目の前の神社を観察する時に、一点を注視するのではなく、扉、屋根、脚など様々な部分に目を向けるので再度サッカード回数が増加するという場面も見られた。以上から、サッカード発生回数の分布と VR 空間内での情報獲得行動の間に対応が見られた。これに加え、注視回数の時系列分布も見ることで、被験者が VR 空間内で情報探索、情報処理のどちらの行動状態にあるかを推定できると考えられる。

6 まとめ

VR コンテンツの知覚情報のクオリティを変化させた結果、被験者の情報獲得行動に差異が見られることが、行動データと視線データからわかった。

視覚情報のクオリティが高い場合、被験者はより周りの環境に注意を向ける傾向が見られ、視覚情報のクオリティが低い場合は、より最適化された探索行動をとる傾向が見られた。

聴覚情報に関しては、仮説で述べたような視覚情報の記憶を阻害するなどの影響は見られず、クオリティが低い場合であっても視覚情報との対応づけがうまくなされることで情報が正確に想起されている例も見られた。

今後は分析するデータの総数を増やしつつ、データ群の分け方を視覚クオリティ毎、神社の種類毎、シーン毎、順番毎、想起課題の記述量などでも分類し、それらの群の間で情報獲得行動に差異が見られるか分析を進める。また、記憶が想起されたと判断する基準を明確に定めつつ、VR 空間の設計とコンテンツの内容を改善する。

謝辞

本研究の一部は科研費 JSPS (22K12284, 代表: 岐阜工業高等専門学校・小川信之, 23K11334, 代表: 長岡技術科学大学・中平勝子) および経営改革促進事業 (代表校・長岡技術科学大学) の助成を受けたものである。

参考文献

- [1] Roxana Moreno and Richard Mayer. Interactive multi-modal learning environments. *Educ Psychol Rev*, Vol. 19, pp. 309–326, 09 2007.
- [2] Muneo Kitajima, Shono Shimizu, and Katsuko T. Nakahira. Creating memorable experiences in virtual reality: Theory of its processes and preliminary eye-tracking study using omnidirectional movies with audio-guide. In *2017 3rd IEEE International Conference on Cybernetics (CYBCONF)*, pp. 1–8, 2017.
- [3] 中川太一, 小川信之, 北島宗雄, 中平勝子, 2024.