

360 度映観察時の注視対象について

—カット割の有無と観察姿勢の効果の検討—

名手久貴

映像学科

What does the observer look at in viewing a 360-degree video ? —Effect of presence or absence of multiple cuts and observation posture —

NATE Hisaki

Department of Imaging Art

(Received October 29, 2021 ; Accepted January 12, 2022)

キーワード：360 度映像、カット割、観察姿勢、視線、ヘッド・マウント・ディスプレイ

要旨

In this study, I investigated whether the effect of multiple cuts and observation postures (standing and sitting) on the subjects' gaze when observing 360-degree videos using VR headsets. In the experiment, subjects wearing a VR headset (Vive PRO EYE : HTC) watched a 360-degree video (1 minute 48 seconds) created with 3DCG software six times. Gaze analysis software (Tobii Pro) was used to display a video and to collect and analyze gaze data. Six experimental conditions were created by combining the presented image (multiple cuts, close-up, bird's-eye view) and factors of observation posture (Standing posture and sitting posture). As a result of the experiment, the tendency to look at the speaker in both male and female vocalizations was indicated. In particular, there was a tendency to see the speaker in close-up condition that required turning of the head in order to see the speaker. This result suggests that an object to be watched may be placed in front of the viewer as one of the methods for producing a video work in which the viewer is not confused even with a 360 degree image with multiple cuts. The effect of posture was not strong because it was observed only in female vocalization, not in male vocalization.

1. はじめに

これまで360度映像を表示するには、ドーム型スクリーンや多数のプロジェクターなど大型の装置が必要であったが、昨今、HTC社のVIVE¹⁾やOculus社のRift²⁾など360度すべての情報を表示する360度映像表示装置が、コンシューマーレベルで購入可能である。これらのシステムの多くは、VRゴーグルとその位置・向き情報を取得し、それに応じた映像を送出するPCを組み合わせたシステムを用いている。現在、映像、ゲーム、携帯アプリ、医療、その他のシミュレーション・システムへの応用など様々な分野の人々が、VRゴーグルを用いた360度映像の利用可能性を探っている。

映像製作の分野においても様々な試みが行われてきた。360度映像の映像作品にはドキュメンタリー³⁾やイベントの記録⁴⁾が多かった。しかし、最近は360度映像を用いた旅番組⁵⁾やドラマ⁶⁾など一般的な映像を製作する試みも見られた。これは、モニターやスクリーンで鑑

賞されてきた従来の映像と同様の映像を360度映像でも鑑賞されるよう、製作が試みられたことを意味していた。

2. VR ゴーグルを用いた 360 度映像の特性

現実の空間では、人の頭部は3軸 (x 軸、y 軸、z 軸) の位置情報に加えて3方向 (yaw、pitch、roll) の合計6自由度 (6DoF) で動くことができる。VR ゴーグル自体は6DoFを実現しているが、それを実現するためには表示する映像もそれに対応する必要がある。ゲームのようにリアルタイムに映像を生成するタイプの3DCGでは、6DoFに対応した映像を表示することが可能である。しかし、事前に360度映像を準備する必要がある実写や事前に用意した3DCGの場合には、頭部の方向に応じた映像を表示することは可能であるが、頭部位置の移動への対応は完全ではない。これは、撮影時にカメラの位置は決定されており、表示時に撮影時のカメラを移動させることができないからである。このように360度映像の作品は、完全には現実空間と一致していない。しかしながら、

頭部の方向に応じた映像を表示することができるため、モニターやスクリーンに表示させる従来の映像を観察する場合よりは現実空間を観察する状況に近い。以上のように360度映像は、頭部の向きに対応した映像を表示できることに加え、視野に占める映像の領域が広いことから従来の映像よりは現実空間の見え方に近く、高い臨場感を体験することができる。

しかし、従来の映像手法を用いて360度映像を製作すると混乱が生じる可能性がある。360度映像は、顔を横に向けると従来の映像では見切れていた部分が見られるが、従来の映像では見ることができない。この違いは、カットの切り替わり時に大きな見え方の違いが生じる可能性がある。つまり、観察者が正面ではなく横を向いている時にカットが切り替わると撮影時にカメラの正面の被写体が観察者の正面に表示されていない事態が生じる。HMDによる360度映像観察中の行動特性を検証した先行研究では、没入型視環境において見回すなどの観察行動に至るまでには準備期間が必要であるとした⁷⁾。映像中の環境を把握するのにある程度の時間が必要であるとすると、カットが切り替わる映像においては、製作者の意図した映像を鑑賞者が見ていないことや製作者の意図が理解されない可能性がある。そこで、カットが切り替わる360度映像を観察した場合、観察者はどのような観察行動をとるのであろうか。本研究ではカットが切り替わる映像観察時の被験者の注視対象について検討した。

360度映像観察時の注視対象に影響を及ぼす要因として観察姿勢も考慮する必要がある。観察姿勢により身体の稼働範囲に変化が生じるため、注視対象が変化する可能性があるからである。360度映像観察時の観察姿勢について検証した先行研究では、固定された椅子と回転可能な椅子に座って観察した場合を比較した⁸⁾。結果、座面の回転が垂直方向の視覚情報の受容に影響していた。座面の回転だけでも得られる視覚情報に影響することが示唆されるため、本研究ではさらに姿勢の変化を大きくした場合（起立姿勢と着座姿勢を比較）に、注視対象に影響を及ぼすか検討した。

3. 本研究の目的と実験条件

360度映像観察時、映像の特性及び観察姿勢により注視対象に影響を及ぼすかについて観察者の視線を計測することで検討した。映像特性の要因としてカット割の有無の影響を検討した。カット割無映像については、被写体の近くにカメラを配置した近接映像と被写体を俯瞰する位置にカメラを配置した俯瞰映像を設け、カット割有映像と合わせて3つの条件を設けて比較した。観察姿勢の条件は、起立と着座を設けた。

4. 方法

(1)



(2)



(3)



図1 実験で使用された映像

(1)、(2)、(3)は、それぞれ被験者が観察したカット割有映像、近接映像、俯瞰映像の例である。(2)は、被験者が正面を向いた時に表示される画像である。近接映像では、被験者が頭部を左右に傾けないと人物を見ることができなかった。

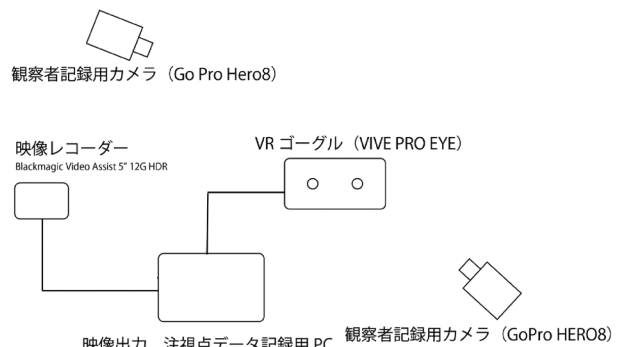


図2 実験システムの概要

(1)



(2)



図3 実験の様子

(1) は起立、(2) は着座で 360 度映像を観察した様子を示す。

HMD (VIVE PRO EYE: HTC 社) を装着した被験者が、3DCGソフト (iClone7: Reallusion 社) で製作した1分49秒の360度映像を観察した (図1、図2、図3)。作成した360度映像の解像度は3840pixel × 1920pixelであり、HMDの表示解像度は片目あたり1440pixel × 1600pixelであった。製作した映像は3種類あり、従来の映像のようにカット割がある映像 (カット割有映像)、人物に近い位置の定点カメラから撮影された映像 (近接映像)、俯瞰位置の定点カメラで撮影された映像 (俯瞰

映像) であった。映像の内容はソファに座った男女の会話シーンであり、カメラの位置が異なる以外、3種類の映像は全て同じであった。男女の会話は、男性、女性、男性と交互に話す構造であり、男女とも6回ずつのセリフがあった。表1の奇数番号の時間帯は、男性がセリフを話しており、偶数番号の時間帯では女性がセリフを話していた (付録参照)。男性がセリフを話した時間は合計78.22秒、女性がセリフを話した時間は合計30.11秒であった。

カット割有映像では、話している人物がカメラの正面に来るようにカメラの位置が切り替えられた。近接映像、俯瞰映像では、カメラの位置が固定され、1カットの映像であった。近接映像では、カメラの位置は男性と女性の間、方向はカメラの正面にバーカウンター、右手に男性、左手に女性が映る方向であった (図1)。俯瞰映像では、カメラは部屋の上部から男性と女性が映る位置に配置された。また、観察条件として、起立と着座の二つの条件を設けた。3種の映像 (カット割有映像、近接映像、俯瞰映像) と観察姿勢の条件 (起立、着座) を組み合わせた6条件を設けた。

被験者は合計6回、映像を観察した。被験者が映像を観察している際、その様子はカメラで撮影し (GoPro Hero8)、被験者が観察しているリアルタイム映像はBlackmagic Video Assist 15" 12G HDRで記録された。360度映像の表示、視線データの収集、分析には視線分析ソフト (Tobii Proラボ VR360エディション) を使用した。映像観察中の被験者の注視点を1秒間に120回測定した (図2、図3)。被験者は11人 (男性4人、女性7人、平均年齢24才) であり、全ての被験者の両眼視機能は正常であり、視力は0.5以上であった。被験者の内、5人は着座で3種の映像を観察したのちに起立で3種の映像を観察した。残りの6人は反対に起立で3種の映像

表1 時間帯番号ごとの開始時間と終了時間及び、当該時間帯にセリフを話した人

時間帯番号	開始時刻 (分:秒)	終了時刻 (分:秒)	セリフを話した人
1	00:00.00	00:06.83	男性
2	00:06.83	00:13.17	女性
3	00:13.17	00:20.36	男性
4	00:20.36	00:22.59	女性
5	00:22.59	00:41.63	男性
6	00:41.63	00:44.46	女性
7	00:44.46	01:08.36	男性
8	01:08.36	01:11.33	女性
9	01:11.33	01:25.33	男性
10	01:25.33	01:28.36	女性
11	01:28.36	01:35.63	男性
12	01:35.63	01:48.33	女性

を観察したのちに着座で3種の映像を観察した。3種の映像の順序はランダム順であった。

5. 結果

実験の結果、男性発声時と女性発声時で、注視対象の傾向が異なるため、それぞれ独立に分析を行なった。表1に従い男性発声時(表1の奇数番号)と女性発声時(表1の偶数番号)に分割し、被験者の注視対象を男性、女性、周囲(男性、女性以外の部分)と3つのカテゴリーに分類した。その後、男性発声時と女性発声時において、注視対象ごとに注視合計時間を算出した(表2)。この時、表1の各時間帯で男性、女性、周囲、いずれの対象も注視していなかったデータは削除した。各条件ごとの注視時間を示した表2より、男性発声時では男性を注視する時間が最も長く、女性発声時では女性を注視する時間が最も長いことが示された。これは、セリフを発する対象を被験者が注視する傾向を示していた。

次に男性発声時において、注視対象(男性、女性、周囲)、提示映像(カット割有映像、近接映像、俯瞰映像)、

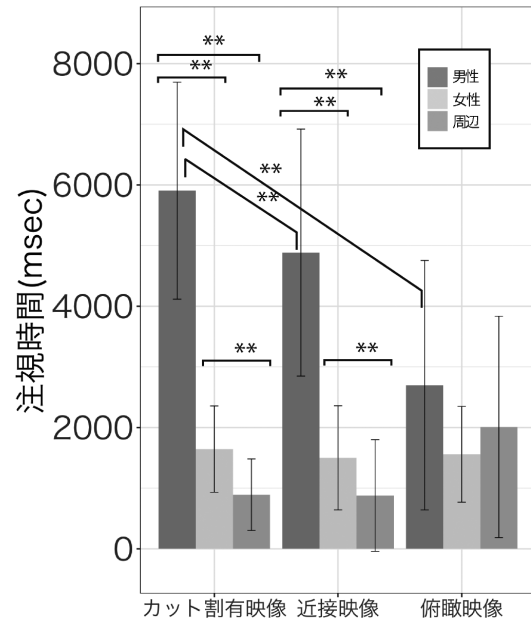


図4 男性発声時における注視時間
横軸は映像、縦軸は注視時間を示す。(**:p<0.01)

表2 男性発声時、女性発声時の注視時間 (msec)

男性発声時		起立					
	カット割有映像		近接映像		俯瞰映像		
	平均	SD	平均	SD	平均	SD	
男性	5994.23	1949.74	4458.08	2043.83	2423.61	2051.57	
女性	1598.33	754.05	1426.69	664.20	1504.42	800.62	
周囲	946.77	680.99	1075.26	1114.18	2007.62	1945.87	
着座		起立					
	カット割有映像		近接映像		俯瞰映像		
	平均	SD	平均	SD	平均	SD	
男性	5814.91	1704.69	5309.43	2032.90	2972.79	2124.20	
女性	1687.56	698.88	1573.37	1046.50	1610.27	814.30	
周囲	838.32	509.13	680.71	673.32	2011.18	1789.76	
女性発声時		起立					
	カット割有映像		近接映像		俯瞰映像		
	平均	SD	平均	SD	平均	SD	
男性	933.03	384.43	749.40	345.90	488.24	365.73	
女性	2156.98	741.21	1698.56	610.40	1264.22	590.66	
周囲	283.85	226.53	380.60	297.76	725.18	799.81	
着座		起立					
	カット割有映像		近接映像		俯瞰映像		
	平均	SD	平均	SD	平均	SD	
男性	1158.62	555.01	597.83	270.29	445.44	314.04	
女性	2081.14	592.99	2003.56	671.57	1834.18	711.32	
周囲	187.38	154.04	261.50	263.19	664.56	682.77	

観察姿勢（起立、着座）を要因とする三元配置被験者内計画の分散分析を行った。結果、注視対象と提示映像の間に有意な1次の交互作用が見られた ($F(4,40)=11.04, p<.01$)。下位検定の結果、カット割有映像、近接映像において注視対象間、注視対象が男性において提示映像間に有意な差が見られた ($F(2,20)=62.02, p<.01$)、($F(2,20)=27.96, p<.01$)、($F(2,20)=14.92, p<.01$)。

Shaffer法による多重比較の結果、カット割有映像、近接映像提示時、ともにセリフを話していた男性を注視する時間が女性、周囲よりも有意に長く、女性は周囲よりも長かった(全て $p<.01$) (図4)。また、注視対象が男性の時、カット割有映像、近接映像は俯瞰映像よりも注視時間が長かった ($p<.01, p<.01$)。

カット割有映像で男性の注視時間が長かった原因として、男性発声時には男性がいつも被験者の正面に表示されるため、被験者が正面を向いていれば他の対象に注意を意識的に向けない限り男性を正面に見据えることができたためと考えられた。近接映像においても男性の注視時間が長かったことは、セリフを発する人を注視する傾向があることを示唆していた。近接映像では、被験者が真っ直ぐ正面を向くとパーカウンターを見ることになり、男性も女性も視界に入らなかった (図1(2))。近接映像では、カット割有映像のように被験者が正面を向くと男性を見られないので、男性を見るためには頭部を左に向ける必要があった。これは、被験者が意識的にセリフを話す男性を注視したことを示しており、360度映像観察時、観察者は発話者を見る傾向があることを示していた。カット割有映像、近接映像では女性を注視する時間も周囲よりも長かった。これは、セリフを話さなくても人への注視時間はそれ以外の対象よりも長くなることを示唆していた。一方、俯瞰映像では、注視対象間の注視時間に違いが見られなかった。他の二つの映像と異なり、注視時間に差が見られなかった原因として、同時に男性と女性が見ることができ、男性と女性の画像中の大きさが他の二つの映像よりも小さく、容易に男性、女性、周囲と注視対象を移動させることができるため、注視時間に差が見られなかったことが考えられた。また、男性と女性を見るためには頭部を下に向ける必要があるため、発話者を観察しづらかった可能性も考えられた。

注視対象が男性の時、カット割有映像は、近接映像、俯瞰映像よりも注視時間が長かった (図4)。カット割有映像では体を動かさなくてもセリフを発する人物を注視できるが近接映像ではセリフを話す人物を見るためには体を動かす必要があった。体の移動が必要ないこととセリフを話す人物を注視する傾向の両方の要因を活かせるため、カット割有映像は近接映像や俯瞰映像よりも注

視時間が長かったと考えられた。俯瞰映像もセリフを話す人物を見ることが可能であったが、セリフを話さない人物も同時に見ることができたために注視時間が分散し、カット割有映像よりも注視時間が短くなった可能性があった。

次に女性発声時についても注視対象（男性、女性、周囲）、提示映像（カット割有映像、近接映像、俯瞰映像）、観察姿勢（起立、着座）を要因とする三元配置被験者内計画の分散分析を行った。結果、注視対象と提示映像の間と姿勢と提示映像の間に有意な1次の交互作用が見られた ($F(4,40)=6.42, p<.01$)、($F(2,20)=3.59, p<.05$)。注視対象と提示映像間の下位検定の結果、提示映像がカット割有映像、近接映像、俯瞰映像において有意な差が見られた ($F(2,20)=50.06, p<.01$)、($F(2,20)=44.16, p<.01$)、($F(2,20)=12.62, p<.01$)。また、注視対象が男性と女性においても提示映像間に有意な差が見られた ($F(2,20)=11.00, p<.01$)、($F(2,20)=5.10, p<.01$)。

提示映像がカット割有映像、近接映像、俯瞰映像のそれぞれにおいてShaffer法による多重比較の結果、カット割有映像、近接映像提示時、ともにセリフを話していた女性を注視する時間が男性、周囲よりも有意に長く、男性は周囲よりも長かった(全て $p<.01$)。俯瞰映像では、女性を注視する時間が男性、周囲よりも有意に長かった ($p<.01, p<.05$) (図5)。これらの結果は、男性発声時と同様の傾向であり、セリフを話す人物を注視する傾向があることやセリフを話さなくても人物であればセリフを話す人物ほどではないが注視する傾向があることを支持していた。

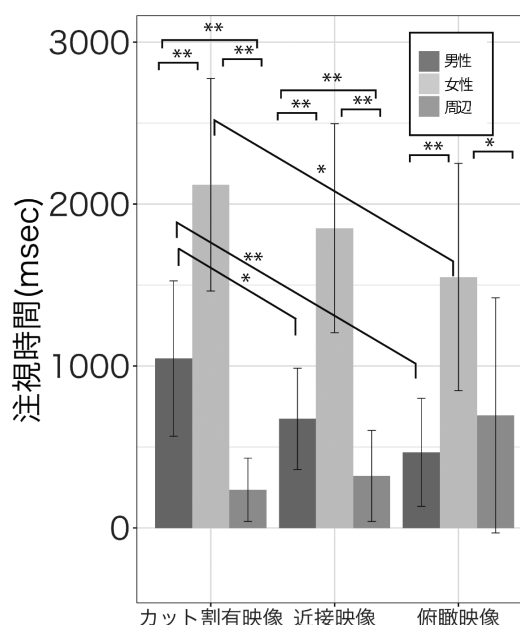


図5 女性発声時における注視時間
横軸は映像、縦軸は注視時間を示す。(**:p<.01, *:p<.05)

注視対象が男性と女性においてそれぞれShaffer法による多重比較を行った結果、注視対象が男性ではカット割有映像は近接映像、俯瞰映像よりも注視時間が長く ($p<.05$, $p<.01$)、カット割有映像は俯瞰映像よりも注視時間が長かった ($p<.05$)。これらの結果は、男性発声時と同じく、カット割有映像では体の移動が必要ないこととセリフを話す人物を注視する傾向の両方の要因を活かせるために注視時間が1番長くなったと考えられた。

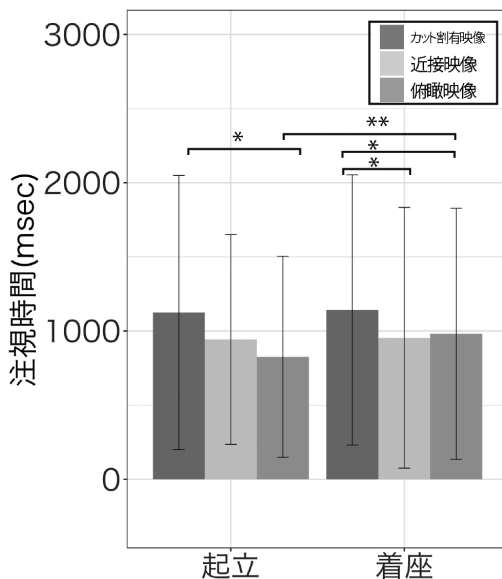


図6 女性発声時における注視時間
横軸は姿勢、縦軸は注視時間を示す。(**: $p<.01$, *: $p<.05$)

姿勢と提示映像間の下位検定の結果、俯瞰映像では着座の注視時間は起立よりも長かった ($F(1,10)=10.51$, $p<.01$) (図6)。俯瞰映像観察時、人物を注視するためには頭部を下に向ける必要があった。起立では頭部を下に向け続ける姿勢では疲労するために注視時間が短くなったと考えられた。また、起立、着座において有意な差が見られた ($F(2,20)=7.05$, $p<.01$)、($F(2,20)=4.90$, $p<.05$)。Shaffer法による多重比較の結果、起立ではカット割有映像の注視時間は俯瞰映像よりも長かった ($p<.05$)。着座では、カット割有映像は近接映像、俯瞰映像よりも注視時間が長かった ($p<.05$, $p<.05$)。起立、着座ともにカット割有映像の注視時間が他の映像よりも長い傾向が見られた。カット割有映像は、正面に発話者が存在するため姿勢を変えずに発話者を注視することができるために他の映像と比べて注視時間が長くなる傾向が見られたと考えられた。しかしながら、起立ではカット割有映像と近接映像に有意な差が見られなかったことから、起立姿勢ではカット割有映像の発話者が正面に位置する効果が減少することを示唆していた。これ

は、弱いながらも起立時、頭部の移動の制約が減少するため、様々な対象に視線が移動することにより注視時間が減少したことを示唆していた。このように女性発声時には姿勢の効果が見られたが、男性発声時には姿勢の効果が見られなかった。男性発声時に比べ女性発声時が短かったことが原因とも考えられ、姿勢の効果について様々な要因を考慮してさらに検討する必要がある。

6. カット割有無と姿勢の効果についての考察

男性発声時、女性発声時ともにカット割有映像が他の映像よりも発話者の注視時間が長かった。実験で使用したカット割有映像は、発話者が被験者の正面に位置したため安定的にセリフを話す人物を注視できたことが考えられた。本実験の結果は、カット割のある360度映像でも鑑賞者が戸惑うことのない映像作品を製作するための手法の一つとして、今回使用したカット割有映像のように注視されやすい対象を鑑賞者の正面に配置させることがあることを示唆していた。

今回の実験では、起立と着座の二つの姿勢で360度映像を観察した。姿勢の影響は男性発声時では見られず、女性発声時のみで見られた。女性発声時、起立、着座ともにカット割有映像の注視時間が長かったことは、発話者が正面に位置するために安定的にセリフを話す人物を注視できることが原因であると考えられた。また、俯瞰映像のみ起立と着座の間に有意な差が見られた。俯瞰映像において人物を見るためには頭部を下に向ける必要があり、この姿勢の持続性による負担の差により起立と着座に差が見られた可能性が考えられた。

今回用いた近接映像では、姿勢の間に差が見られなかった。着座は起立姿勢に比べ、腰から上の回転が行いづらくなるため、着座では特定対象への注視時間が起立に比べ短くなる可能性もあったが、今回の実験ではそのような傾向は見られなかった。今回用いた近接映像では、頭部の回転のみで人物を見られる可能性があったので、さらに後方に人物を配置した場合、起立と着座による注視時間や注視対象の違いが見られる可能性も考えられた。360度映像の観察姿勢についてはさらに検討する必要がある。

7. まとめ

映像 (カット割有映像、近接映像、俯瞰映像) と観察姿勢 (起立、着座) を要因として、360度映像観察時の被験者の注視時間を計測した。実験の結果、男性発声時、女性発声時ともに話者を見る傾向が示唆された。特に話者を見るためには頭部を回転させる必要がある近接映像においても話者を見る傾向が見られたことから、その傾

向は強いものであることが示された。これは、カット割のある360度映像でも鑑賞者が戸惑うことのない映像作品を製作するための手法の一つとして、注視される対象を鑑賞者の正面に配置させることが有効であることを示唆していた。

本研究の目的の一つである姿勢の効果については、男性発声時には見られず女性発声時のみ見られたことから、強いものであるとは言えなかった。今回用いた映像は1分48秒と短いものであったので身体的な疲労が発生しない程度であった。今後、姿勢要因による鑑賞行動の変化についてさらに検討が必要である。

(謝辞)

本実験で使用した映像のシナリオは、東京工芸大学芸術学部高山隆一教授に作成していただいた。ここに記して感謝いたします。

参考文献

- 1) <https://www.vive.com/jp/>
- 2) https://www.oculus.com/rift/?locale=ja_JP#oui-csl-rift-games=mages-tale
- 3) 『360° Great Hammerhead Shark Encounter』 National Geographic, 2016.
- 4) 『2年ぶりに開催！仙台七夕まつり』NHK, 2021
- 5) 『鶴瓶の家族に乾杯360° VR 小豆島・寒霞溪 エンジェルロードの絶景旅』NHK, 2020
- 6) 『ゴースト刑事 日照荘殺人事件』日本テレビ, 2018
- 7) 塚田、長谷川、伴地、森川、河合「簡易型HMDを用いた360度映像観察中のユーザの身体行動特性の分析」日本バーチャルリアリティ学会誌, vol.21, No.4, page595-603, 2016
- 8) 伴地、吉川、河合「HMDを用いた360度動画視聴時の座面の回転がユーザ体験に及ぼす影響」日本バーチャルリアリティ学会誌, vol.23, No.3, page217-227, 2018

付録 実験で使用した映像中のセリフ (番号は時間帯番号を示す)

1 亮太「大丈夫だった。あんな古い映画」

2 由依「大丈夫、面白かった。でもあんなに古い映画初めて」

3 亮太「そう、デジタルリマスター版で。前から大きなスクリーンで見たかったんだ」

4 由比「すきな」

5 亮太「うん、あの女優さん好きなんだ。むかし、親父がビデオで見てて。それから何回も見てるんだよ。あの人さ、俺のおふくろと同じ年だったんだ。それにおふくろが死んだ年に亡くなったんだ」

6 由比「よく知ってるんだね」

7 亮太「あの人の映画や本読んでさ。とても素敵な人なんだ」

7 亮太「今日、ありがとう。浜田さんみたいな人が一緒に映画付き合ってくれたなんて奇跡みたいだよ。いつも演劇部のかっこいい役者さんとしてさ。俺みたいな映画研究会のオタクみたいな奴がさ」

8 由比「そんなことないよ」

9 亮太「俺、今日あきらめに来たんだ。大好きな女優さんを大きなスクリーンで見て、告白して、ふられて、酒飲んで一晩中泣くんだよ」

10 由比「どうして、そうなの」

11 亮太「決まってるじゃん、俺と浜田さんだよ。浜田さんみたいな綺麗な人」

12 由比「わたし、お母さんのこと、素敵な人っていう人、嫌いになれないよ。木ノ内君、私もお母さんのように撮ってくれる」