

推薦論文

擬似的な前後方向の運動視差付き映像が社会的テレプレゼンスに及ぼす影響の評価

川路 崇博^{1,a)} 坂本 竜基²

受付日 2013年10月22日, 採録日 2014年2月14日

概要: 遠隔地どうしをネットワークで結んだ環境における音声と映像を介するコミュニケーションにおいて、閲覧者の視点の動きに即した運動視差を映像に適用すると社会的テレプレゼンスが増強されるという研究結果がある。なかでも視点の前後方向の動きに同調して、対話相手のカメラを物理的に前後移動させるシステムは著名である。しかしそのシステムが必要とするクリアランスの問題から一般的なデスクトップPC環境には適用できない。一方、カメラの代わりに深度センサを利用し、閲覧者の視点位置検出および、対話相手の映像を擬似的な運動視差付き映像に変換するシステムがある。しかし、このシステムで得られる映像は、深度センサの性質に起因する見え方に不自然な部分をともなう。そこで、この完全ではない運動視差映像が社会的テレプレゼンスに与える影響を明らかにするため評価実験を行った。結果、擬似的な運動視差付き映像も、カメラを前後方向に動かすことにより得られる運動視差と同じく社会的テレプレゼンスを増強することが明らかとなった。

キーワード: ビデオチャット, 社会的テレプレゼンス, 擬似的な運動視差

Effectiveness for the Social Telepresence on Videos with Virtual Motion Parallax Moving Back and Forth

TAKAHIRO KAWAJI^{1,a)} RYUUKI SAKAMOTO²

Received: October 22, 2013, Accepted: February 14, 2014

Abstract: In communication through sounds and images in the environment that bound a distant place fellow together in a network system, a research shows that social telepresence is reinforced when we apply the motion parallax accorded with the movement of the viewpoint of the reader to image. Most especially, there is a well-known system that moves the partner's camera back and forth physically, by tracking the front and back direction movements of the viewpoint. However, it's not applicable to general desktop PC environment due to issue of the requirements for the system. On the other hand, there is a system that using a depth sensor in substitution for the camera to detects the viewpoint position of the reader, and converts the image of the partner into a virtual motion parallax video. But then, the image which obtained by this system involves partial unnatural visual performance due to the nature of the depth sensor. Consequently, we had evaluation experiment to clarify about the influence of the defective motion parallax image gives to social telepresence. The experiment results were revealed; the virtual motion parallax video also reinforces social telepresence, just as same as the motion parallax as if by moving a camera in the direction of back and forth.

Keywords: video chat, social telepresence, virtual motion parallax

¹ 大月市立大月短期大学経済科
Department of Economics, Ohtsuki, Yamanashi 401-0012,
Japan

² ヤフー株式会社 Yahoo!JAPAN 研究所
Yahoo! JAPAN Corporation, Minato, Tokyo 107-6211,
Japan

a) kawaji@ohtsuki.ac.jp

1. はじめに

遠隔地どうしをネットワークで結んだ環境における音声

本論文の内容は2013年5月のグループウェアとネットワークサービス研究会にて報告され、同研究会主査により情報処理学会論文誌ジャーナルへの掲載が推薦された論文である。

と映像を介するコミュニケーションは相手の存在感を獲得できないとされている [1]. この問題を解決するため, 同室感や対面感を演出する手法が研究されており [2], [3], [4], その一要因として運動視差に着目した研究がなされてきた [5], [6], [7], [8], [9]. 運動視差とは, 観測者と観測対象の位置や角度が変わった場合に観測者が知覚する「見え」の変化である.

たとえば, 加藤らはビデオチャットの相手側のカメラを物理的に前後に移動させると社会的テレプレゼンスが増強されると報告している [7]. しかしここで提案しているシステムでは, モニタの前部にカメラを可動させるレールの設置スペースが必要であり, 一般的な環境におけるチャットシステムにそのまま応用することは難しい.

そこで, 平岡らは, 卓上に設置可能な深度センサ付きカメラをもちいて, 余分なスペースがない一般的なデスクトップ PC 環境でも加藤らのシステムで得られるような前後方向の運動視差が付いたかのような映像を生成する手法を提案している [10]. しかし, ここでの運動視差付き映像とはあくまで疑似的な運動視差であり, 物理的に前後方向にカメラを動かした映像と完全には一致しないが, 平岡らは評価実験を行っていないため, この差異が社会的テレプレゼンスに及ぼす影響は検証されていない.

そこで本論文では, 平岡らのシステムを用いて評価実験を行い, 深度センサ付きカメラによる疑似的な運動視差映像が社会的テレプレゼンスに及ぼす影響を報告するとともに, このような省スペースで設置可能なシステムが通常の映像よりも高い社会的テレプレゼンスを伝達することを示す.

以下, 2章で先行研究を示し, 3章で実験概要を述べる. 4章で実験条件を述べたうえでその結果を示す. 5章では追加で行った実験の実験条件と結果を示し, 擬似的な運動視差の効果について述べる. 最後に6章で全体を総括する.

2. 先行研究

遠隔地コミュニケーションにおける社会的テレプレゼンスは映像の有無に左右されることは古くから確認されている [11], [12], [13]. その後, この社会的テレプレゼンスの増強は映像の内容や種類によっても, その度合いや性質が変化することが明らかにされてきており, たとえばアイコンタクトや立体映像は通常の映像に比べて社会的テレプレゼンスが増強されることが報告されている [14], [15]. その中でも運動視差は, 着目すべき条件の1つとして注目されており, 近年さかんに研究されてきている [7], [8], [9].

加藤らは, 被験者の前方向の移動に同期して動く, 実験者側に設置された可動式カメラを用いて, 前後方向の運動視差が社会的テレプレゼンスに及ぼす影響の測定を試みた [7]. この実験では, 被験者は 2.5m ほど前方向に歩行すると, 実験者の前のレール上に置かれたカメラが約 75cm, 実験者側に同期して動く環境を用意している. この結果, 被験者は, カメラが実際に動いた運動視差付きの映像を見

ると, それをズーム映像と混同してしまうものの, 実際にズームさせた映像に比べて運動視差付きの映像を提示した場合のほうが社会的テレプレゼンスは増強されるとしている. これに対して田中らは, 実験者側の動きにあわせて, 被験者側のディスプレイを物理的に前後 6cm に移動させる実験を行い, 被験者側の表示デバイス自体の動きも社会的テレプレゼンスを増強させることを示した [8].

一方, このような物理的にカメラや表示デバイスを動かすのではなく, 画像処理によって, 運動視差を模した映像を生成するシステムも提案されている. たとえば, 石井らは, MoPaCo (Motion Parallax Communication video system) [9] というシステムにおいて, ユーザの横方向の視点移動に応じて, まるで画面を窓枠として, 向こう側にいる相手を見たかのように, 相手側の背景を合成することで運動視差を模した映像を提示している. この実験では, 通常のテレビ電話と MoPaCo を用いた条件で比較が行われ, 映像処理による画像の破綻や画質の低下は特に問題にならず, 対面感や実在感が増したとされている.

MoPaCo では左右方向の運動視差の提示を主としているが, これに対して平岡らは, 深度センサ付きカメラを用いて, 加藤らの実験のような前後方向の運動視差を模した映像をチャットシステムに応用する提案をしている. このシステムでは, 深度センサ付きカメラとして Microsoft 社の Kinect*¹ をもちいている. ここでの深度センサは, 以下の2種類の役割を担うとされている.

役割 a ユーザの顔がモニタにどれだけ近づいたかを計測する.

役割 b 相手側を映したカメラの映像のうち前景領域をセグメンテーションする.

ここでチャットをしている2者をユーザ A, ユーザ B とすると, システムは, ユーザ A の顔の位置をユーザ A 側に設置された深度センサで検知 (役割 a) し, その近さに応じてユーザ B 側のカメラの映像を疑似的な運動視差付き映像 (以下, 疑似運動視差映像と呼ぶ) に変換してユーザ A のモニタに提示する. この変換は, ユーザ B を映した映像を「ユーザ B (前景)」と「それ以外 (背景)」に領域分割 (役割 b) し, それぞれ下記の透視投影の比から計算される拡大率 α , β で, 拡大したうえで合成することで実現される.

$$\alpha = \frac{l_{userA} + l_{userB}}{l_{userA} + l_{userB} - d_{userA}} \quad (1)$$

$$\beta = \frac{l_{userA} + l_{backgrounduserB}}{l_{userA} + l_{backgrounduserB} - d_{userA}} \quad (2)$$

ただし, l_{userA} , d_{userA} , l_{userB} , $l_{backgrounduserB}$ は, 図 1 に示したユーザ A, ユーザ A の顔, ユーザ B, ユーザ B の背景, それぞれの Kinect からの距離である. ここで, ユーザ A とユーザ B は入れ替わっても問題なく, 双方向で同じ動作が可能であるため文献ではチャットシステムで応用

*1 <http://www.xbox.com/ja-JP/kinect>

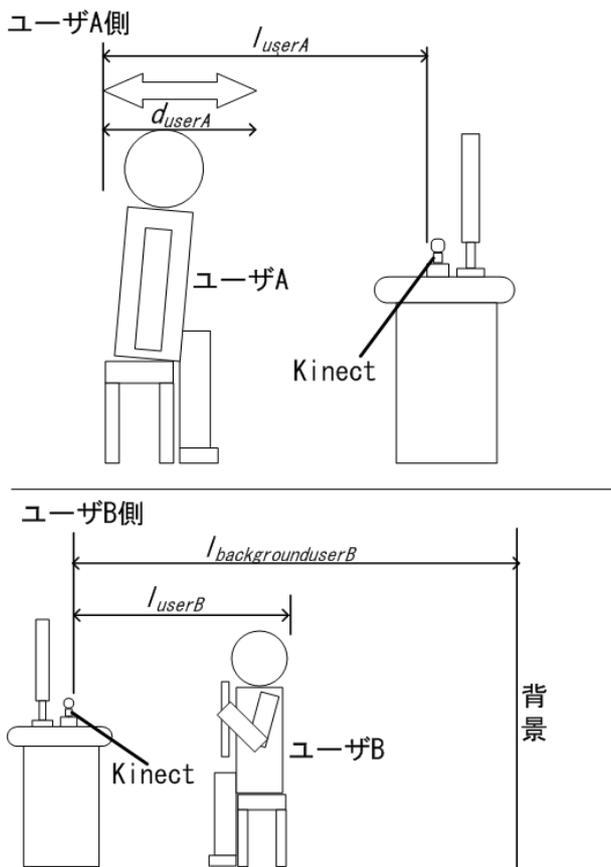


図 1 ユーザ、ユーザ B、背景および Kinect の位置関係
 Fig. 1 Distances on Eq. (1) and Eq. (2).



図 2 擬似的な運動視差生成で不自然な部分をともなって変換された映像。破線で囲んだ部分が特に不自然な部分

Fig. 2 The video is changed unnaturally in the virtual parallax generator. The portion surrounded with a circle of the dashed line is especially unnatural.

可能とされている。なお、具体的な実装方法など詳細は文献 [10] を参照されたい。

一方、このシステムを実際に稼働させてみると、深度センサの精度や特性から、完全な領域分割が行われず、やや不自然な映像が提示されることがある (図 2)。また、そもそも前景と背景の倍率を変えただけでは、被写体の中での

凹凸、たとえば顎と首の距離により発生するはずの比較的軽微な運動視差を再現することができない。

加藤らは、相手側のカメラを物理的に動かす、つまり運動視差を付けた映像は、動かないカメラで写す通常の映像に比べてテレプレゼンスが増強されることを示した。しかし、平岡らのシステムは、上述した運動視差としての不完全さがあるため、これが影響して通常の映像に比べて社会的テレプレゼンスを増強するかどうかは不明である。この影響を調べることが本論文の主目的である。

3. 実験概要

3.1 実験目的

まず社会的テレプレゼンスの伝達において通常の映像に対する疑似運動視差映像の優位性を確認したい。検証方法は、先行研究である加藤らの実験手段 [7] にならう。加藤らは通常の映像と運動視差付き映像に加え、ズームの映像も比較しているため、本実験でも頭部位置に応じてデジタルズームする機能を用意して 3 条件を比較することにした*2。

3.2 実験環境

一般的なデスクトップ PC どうしで平岡らのチャットシステムを利用することを想定して図 3 のような実験環境を設定した。実験は、被験者 1 名に対して著者らが指名した説明者 1 名のペアで、システム越しにコミュニケーションをとるタスクを行った後、アンケートに回答してもらう形式をとった。図 3 の上段は被験者側、下段は説明者側の環境を表している。被験者と説明者の間は壁で仕切り、互いにネットワーク接続された PC とディスプレイ、深度センサ付きカメラとして Kinect を設置した。

被験者側では、WUXGA の 24 インチディスプレイを使用した。被験者に提示する映像の解像度は VGA であり、これを全画面表示した。なおフレームレートは約 30 fps であった。これらの条件で実際に説明者を表示してみると、通常状態において上半身が約 230 mm、拡大状態において約 260 mm で表示された。被験者と説明者は Skype を用いて、音声のやりとりを行えるようにしたが、環境音がコミュニケーションの妨げとならないように被験者にはノイズキャンセリングイヤフォンを与えた。

3.3 手順

加藤らの実験では、説明者が小型ロボットを顔の横に掲げ、そのロボットの説明を被験者に対して行うコミュニケーションを設定している。本実験ではこれにならい、説明者が日本の城の写真を胸の前に持った状態 (図 6 右側) で、その城に関する説明を被験者に対して行うコミュニ

*2 ただし、加藤らの実験はデジタルズームではなく、光学ズームである。Kinect には光学ズームの機能がないためデジタルズームで代用した。

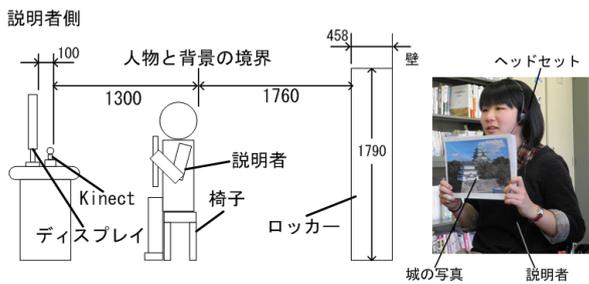
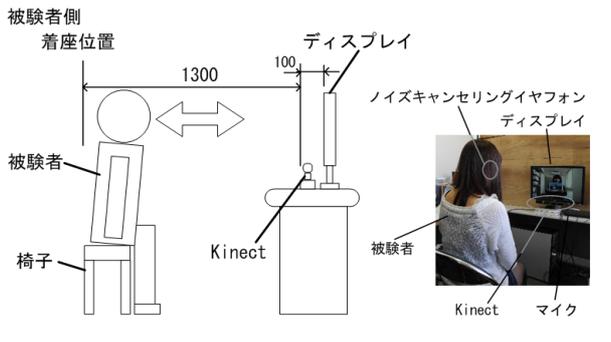


図 3 実験環境 (単位: mm)

Fig. 3 Setup for the experiment (the unit of length is the millimeter).

説明者: こんにちは
 被験者: こんにちは
 説明者: これから日本のお城について説明します。
 被験者: はい
 説明者: (1つ目の城の写真を見せる)
 説明者: この写真のお城は何城か分かりますか?
 被験者: (身を乗り出して映像を見る)
 被験者: (分かるまたは分からない旨回答)
 説明者: (城の名前を教える, または別名を問う)
 説明者: このお城には特徴的な部分があります。
 説明者: それがなにか分かったら教えてください。
 被験者: (身を乗り出して映像を見て, 特徴を探す)
 被験者: (思いついた特徴を言う, または分からないと答える)
 説明者: (お城の特徴を教える)
 説明者: このお城には他にも (城の特徴や逸話など) があります。
 被験者: (感想を言う)
 説明者: これで1つ目のお城の説明を終わります
 説明者: アンケートにご回答ください
 被験者: (アンケートを記入する)
 (アンケートの記入が終わったら, 映像の条件を切替える)
 (説明者: 2つ目の城の写真を見せる)
 説明者: これから2つ目の城の説明をします。
 説明者: この写真のお城は何城か分かりますか?
 [中略]
 (3条件で城の説明を行なう)
 説明者: 以上で終わります。ありがとうございました。

図 4 説明者と被験者の典型的な会話プロトコル

Fig. 4 The typical conversation protocol.

ケーションを設定した。このときコミュニケーションが一方向的にならないように図 4 のような質疑応答も含む粗筋をあらかじめ用意しておいた。

粗筋には、被験者が身を乗り出す動作が発生するよう、城の写真をよく見ようとする動作を2回以上誘発する問い

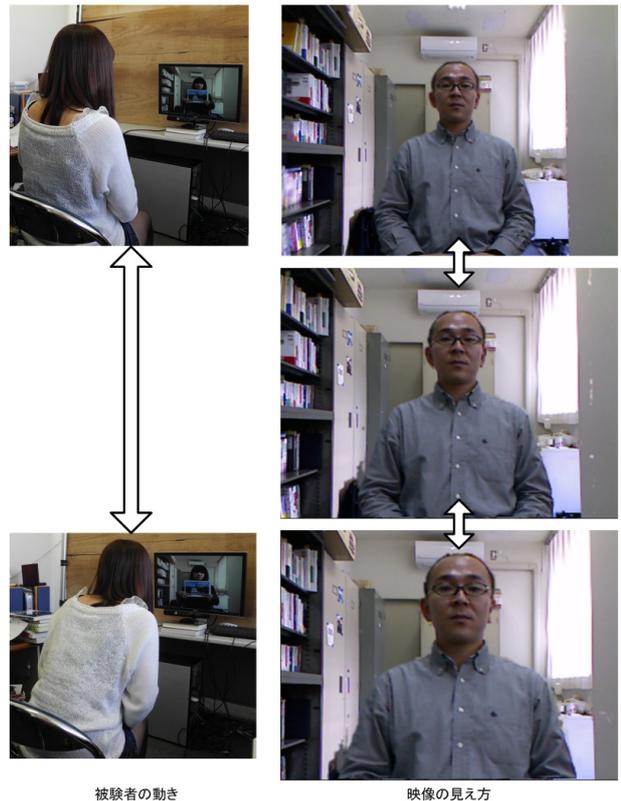


図 5 被験者の動きに対応した疑似運動視差映像の見え方

Fig. 5 The virtual motion parallax video corresponding to the movement of the subject.



図 6 実験の様子

Fig. 6 The conditions of an experiment.

かけを入れた。意図したとおり、実験では、このような問いかけがあった場合、被験者は 30 cm ほど身を乗り出すことが観察された。

図 5 に被験者側に表示する疑似運動視差映像の例を示す。上段は被験者が普通に椅子に座っている状態のときに表示される映像で、下段は身を乗り出したときに表示される映像である。

4. 実験 1

上述したとおり、本実験では、通常の映像、デジタルズーム映像、疑似運動視差映像という3種類の映像を比較する。これらの映像を用いた実験条件を以下のようにそれぞれ条件 N, 条件 Z, 条件 P と呼ぶ。



図 7 各条件の映像の見え方。上段が初期状態，下段が身を乗り出した状態（左：条件 N，中：条件 Z，右：条件 P）

Fig. 7 The image visibility in the three conditions. The upper row is in an initial state and the lower berth is in the state toward which the subject leaned out (left: condition N, mid: condition Z, right: condition P).

条件 N (Normal)：一般的なビデオチャットシステムで用いられる映像を用いたコミュニケーションである。被験者が身を乗り出しても表示される映像には変化がない。

条件 Z (Zoom)：深度センサ付きカメラから生成された、被験者の前後移動に同期してデジタルズームをする映像（以下、デジタルズーム映像とする）を用いたコミュニケーションである。デジタルズームでは画像の一部を切りだして拡大するため人物と背景の拡大率に違いがない。

条件 P (Parallax)：深度センサ付きカメラで変換した平岡らの提案する疑似運動視差映像によるコミュニケーションである。被験者の頭部とディスプレイ間の距離に応じて変化する疑似的な前後方向の運動視差付きの映像を表示する。

各条件の映像の見え方の例を図 7 に示す。上段が初期状態，下段が身を乗り出した状態において表示されるフレームである。この図では、たとえば条件 Z と条件 P の間では下段における背景の位置関係*3 に多少の差異しか現れていないが、実際の映像において上段と下段の間における連続的な変化がつくと両者はかなり異なって見える。実験は、1名の被験者に対して3条件を続けて行った。説明者が3条件を選択する順番は被験者ごとにランダムで選び、カウンターバランスをとった。説明するお城は6種類用意して、それぞれの内容は大きく異ならないように留意したうえで、条件にかかわらず毎回ランダムに選び、説明の内容が結果に影響しにくい設定とした。

*3 たとえば、上部にあるエアコンの室内機の位置。

アンケートは9段階のリッカート尺度を用い、以下の各質問項目に対して1はまったくあてはまらない、3はあてはまらない、5はどちらともいえない、7はあてはまる、9は非常にあてはまる、に対応させた。

- Q1：映像は十分にきれいだと感じた。
- Q2：音声は十分にきれいだと感じた。
- Q3：城の説明は分かりやすかった。
- Q4：実際に机越しに直面して会話している感じがした。
- Q5：説明者を机越しに眺めている感じがした。
- Q6：説明者から机越しに眺められている感じがした。
- Q7：机越しに直面している距離感を感じた。

実験では、1条件の説明に約3分、アンケートの記入に約1分ほど要していた。

4.1 実験結果

まず、全被験者 (N = 19) のアンケート結果を集計し、3条件に対して一元配置分散分析を行った。先行研究 [7], [8] では、Q1 から Q3 を「説明の品質」、Q4 から Q7 を「社会的テレプレゼンス」に関するアンケート項目と分類している。

一元配置分散分析の結果、「説明の品質」の項目、Q1 ($F(2,54) = 1.80, p = 1.74$), Q2 ($F(2,54) = 2.96, p = .06$), Q3 ($F(2,54) = 1.73, p = .19$) で、有意差は認められなかった。「社会的テレプレゼンス」の項目、Q4 ($F(2,54) = 41.74, p < .01$), Q5 ($F(2,54) = 59.31, p < .01$), Q6 ($F(2,54) = 39.57, p < .01$), Q7 ($F(2,54) = 38.94, p < .01$) では、有意差が認められた。

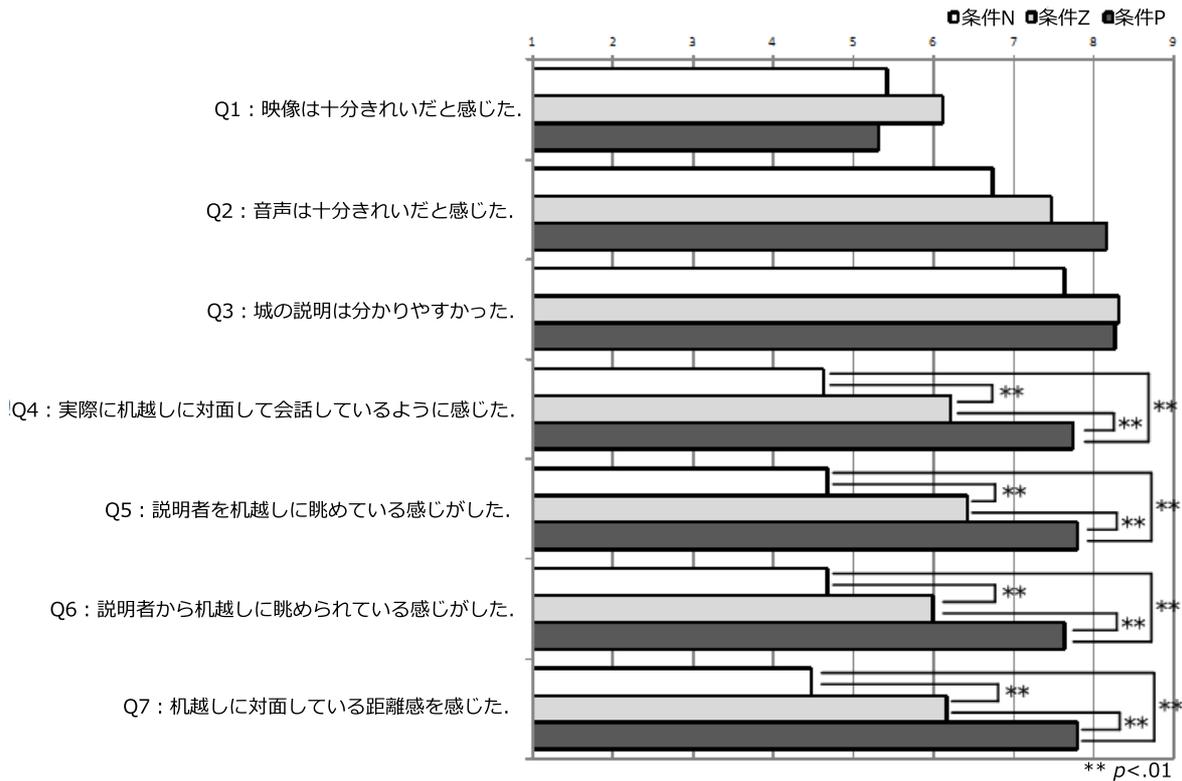


図 8 アンケート項目および実験の結果

Fig. 8 The items and the result of questionnaire for experiment.

有意差が認められた社会的テレプレゼンスの全項目で、各条件について Tukey の多重比較を行った結果を図 8 に示す。まず条件 N と条件 Z を比較したところ、有意差が認められた ($p < .01$)。これは、通常の映像よりもデジタルズーム映像が、テレプレゼンスを増強することを示唆している。

次に条件 N と条件 P を比較したところ、これも有意差が認められた ($p < .01$)。これは、通常の映像よりも疑似運動視差映像がテレプレゼンスを増強することを示唆している。

さらに条件 Z と条件 P の比較でも有意差が認められた ($p < .01$)。つまり、これらの結果は通常の映像はもとより、デジタルズーム映像よりも疑似運動視差映像が社会的テレプレゼンスをより増強することを示している。

一方、アンケート項目「映像は十分にきれいだと感じた。」では、3 条件に有意差が認めれない。したがって疑似運動視差映像は、図 2 のような深度センサの性質に起因する画像の乱れがあるものの、それは認知的にも問題がなく、また、社会的テレプレゼンスの伝達要因としても致命的な影響はないと考えられる。

5. 実験 2

前章の実験において、被験者が画面に近づいた場合、条件 Z と条件 P の映像での説明者は、条件 N の映像よりも大きく表示される。よって、条件 N に対して、条件 Z と条

件 P が高得点を得ていた理由は、ズームや運動視差などではなく、単に説明者が大きく表示されていたという要因のみが影響していた可能性を否定できない。そこで、説明者が大きく表示される映像を用いる新たな条件を追加し、これを条件 NB として本実験を行った。この映像は、デジタルズーム映像で、最大限までズームした状態に倍率を固定することで生成した。

また、デジタルズーム映像と疑似運動視差映像においても、拡大率が社会的テレプレゼンスに影響する可能性があるため、条件 Z と条件 P のそれぞれの拡大率を上げた設定にした条件 ZB、条件 PB も用意した。まとめると以下のようなようになる。

条件 NB (Normal(Big)): 被験者が前後しても拡大率を変化しない映像によるコミュニケーションである。ただし、条件 N よりも説明者は大きく映るような設定となっている。

条件 ZB (Zoom(Big)): 条件 Z よりも拡大率を大きく設定したデジタルズーム映像をもちいたコミュニケーションである。被験者が近付いた場合に表示される説明者の大きさは、条件 NB と同程度になるように設定した。

条件 PB (Parallax(Big)): 条件 P よりも拡大率の大きい疑似運動視差映像である。条件 ZB 同様、拡大したときの表示される説明者の大きさは、条件 NB と同程度になるように設定した。

実際には、条件 PB の調整が最も困難であるため、式 (1)



図 9 実験 2 での各条件の映像の見え方. 上段が初期状態, 下段が身を乗り出した状態 (左: 条件 NB, 中: 条件 ZB, 右: 条件 PB)

Fig. 9 The image visibility in the three conditions. The upper row is in an initial state and the lower berth is in the state toward which the subject leaned out (left: condition NB, mid: condition ZB, right: condition PB).

における α , 式 (2) における β を 2α , 2β と置き換えて条件 PB を出力し, そのときの説明者の最高の拡大率となる大きさにおおよそ合致するように, 条件 NB および条件 ZB の最高拡大率を設定した. なお, これらの係数 (2α , 2β) は, 本実験環境において, 説明者が画面に収まる最大の拡大率であることを目視によって事前に確認して決定した. 条件 NB で説明者の上半身は約 290 mm, 条件 ZB および条件 PB では, 通常状態において上半身は約 230 mm, 拡大状態において約 290 mm で表示された.

本実験環境は, 前章の実験と同一 (図 3), 実験方法やアンケートも同一としたが, 被験者が重複しているため城の写真は異なるものを用意した. 各条件における説明の見え方を図 9 に示す. 上段が初期状態, 下段が身を乗り出した状態において提示される映像のスクリーンショットである.

実験 2 のアンケート結果と前章での実験結果から, 条件 N, Z, B, NB, ZB, PB の 6 条件に対して一元配置分散分析を行った. この結果, 「説明の品質」の項目, Q1 ($F(5,108) = 2.30, p < .01$), 「社会的テレプレゼンス」の項目, Q4 ($F(5,108) = 29.84, p < .01$), Q5 ($F(5,108) = 32.24, p < .01$), Q6 ($F(5,108) = 23.34, p < .01$), Q7 ($F(5,108) = 29.34, p < .01$) で, 有意差が認められた. 「説明の品質」の項目, Q2 ($F(5,108) = 2.96, p = .07$), Q3 ($F(5,108) = 1.33, p = .26$) では, 有意差は認められなかった. 有意差が認められたアンケート項目に対し, Tukey の多重比較を適用した. この結果を図 10 に示す.

まず, 同じ種類の映像, つまり, 通常どうし, デジタルズームどうし, 疑似運動視差映像どうしで倍率が異なる条件間の検討を行った. 結果, 社会的テレプレゼンスの項目 Q4, Q5 で, 条件 N と条件 NB に有意差が認められた ($p < .05$) が, 条件 Z と条件 ZB, 条件 P と条件 PB では認められなかった. また Q6, Q7 では, 条件 N と条件 NB, 条件 Z と条件 ZB, 条件 P と条件 PB のすべてで有意差は認められなかった.

これは, デジタルズームや疑似運動視差映像などを適用しない通常の映像では, 相手の見た目の大きさが社会的テレプレゼンスに影響を与えるが, デジタルズームや疑似運動視差映像では, 見た目の大きさは影響を与えないことを意味している. この理由を考察すると, 有意差が認められた Q4, Q5 は, 「説明者と机越しに対面して会話している感じがした」「説明者を机越しに眺めている感じがした」であることから, 説明者の目元や口元など細部の視認性も得点に影響をあたえらると思える. また, 通常の映像では, 説明者の見た目が大きくなると, これらの視認性が向上するため高得点となったのではないかと考えられる. 一方, 「条件 NB」と「条件 ZB, PB」との差は被験者の動きに同期して映像が変化するか否かであり, 大きくはなるものの, たえず大きさが変化してしまうこと, さらに, 「条件 NB」ではつねに説明者が大きく表示されているが, 「条件 ZB, PB」では, 被験者が近付いたときのみ大きく表示されることから細部の視認性が低下したと考えると説明がつく.

次に拡大率が変化しても, 前章の実験結果, つまり, 通常の映像よりも, 疑似運動視差映像のほうが社会的テレプレ

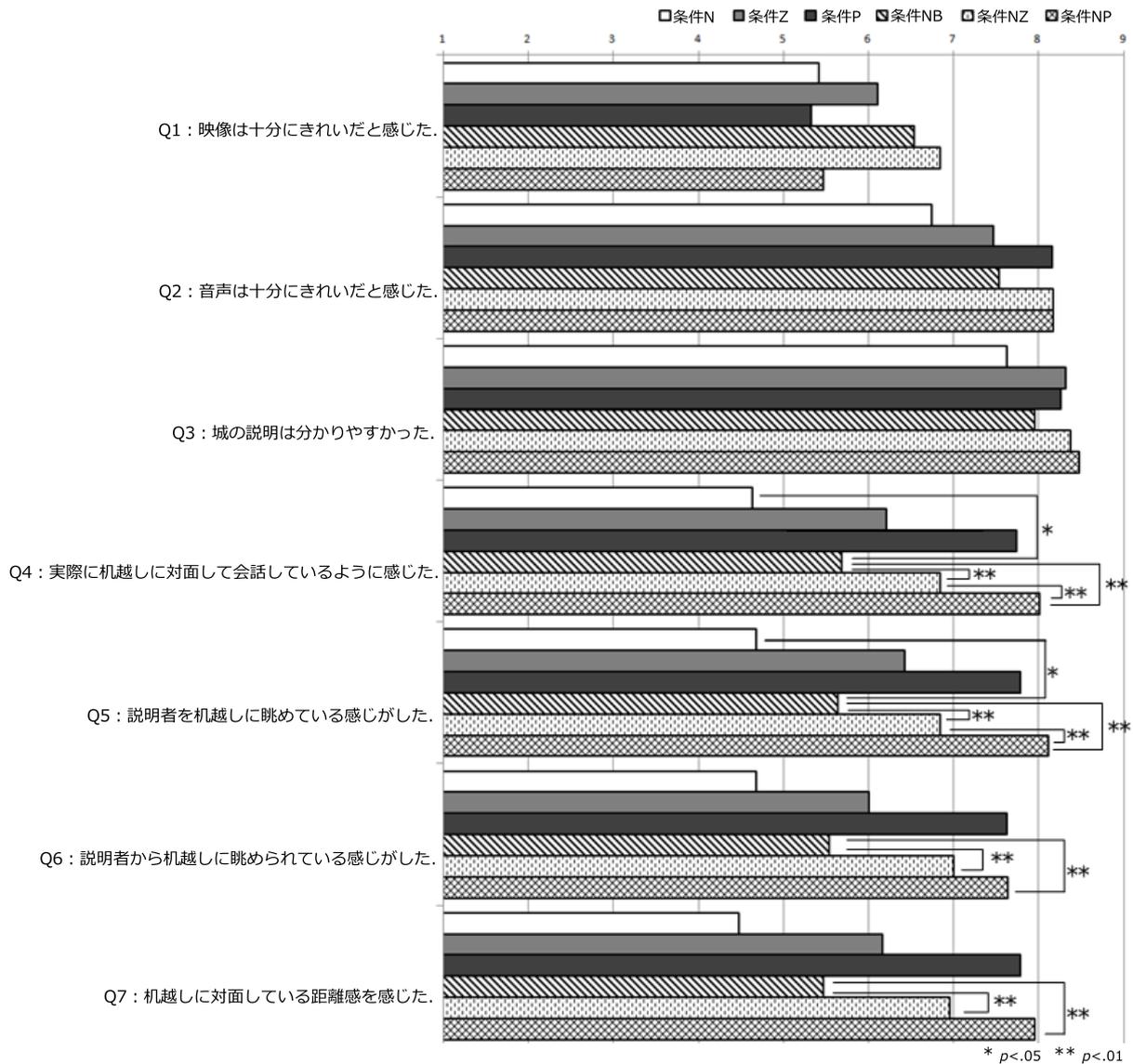


図 10 アンケート項目および実験の結果
 Fig. 10 The items and the result of questionnaire for experiment.

レゼンスが増強される傾向が維持されているかどうか検証した。結果、社会的テレプレゼンスの項目 Q4, Q5, Q6, Q7 において「条件 NB, 条件 ZB 間」, 「条件 NB, 条件 PB 間」のそれぞれで有意差が認められた ($p < .01$)。つまり、拡大率の大きいデジタルズーム映像と拡大率の大きい疑似運動視差映像は、拡大率を高めた通常の映像よりも社会的テレプレゼンスを増強させる。

さらに Q4, Q5 で、条件 ZB と条件 PB に有意差が認められた ($p < .01$)。これは拡大率の大きい疑似運動視差映像が、拡大率の大きいデジタルズーム映像よりも社会的テレプレゼンスを増強する傾向にあることを示唆する。よって表示される説明者の大きさが異なっても、それが及ぼす社会的テレプレゼンスへの影響は前章の実験の結果とほぼ同傾向であるといえる。

6. まとめ

本論文では、深度センサ付きカメラを用いて、ユーザの前後方向の運動視差を模した映像を疑似的に生成し、ビデオチャットの映像とするシステムにおいて、その視差が社会的テレプレゼンスに及ぼす影響を実験により明らかにした。この結果、疑似的運動視差映像も、物理的に前後方向にカメラを動かすことにより得られる運動視差と同じく、社会的テレプレゼンスを増強させることが判明した。

一方、本実験では説明者と実験者に役割を分けて実験環境を設定したが、実際のビデオチャットシステムとして運用した場合は、このような役割は流動的である。また、協調作業やコールセンタで顧客対応をする場合など、運用には様々なシチュエーションが考えられる。このようなより実際の運用に即した場合でも同じ効果を得られるのかどうかは不明である。また、本論文では見た目の大きさを要因

として考えた実験2を行ったが、たとえばスマートフォンに応用した場合などの表示が極端に小さい場合でも視差が社会的テレプレゼンスを増強するかどうか不明である。どちらも実験により評価可能であると考えられるが、これらについては今後の課題とする。

謝辞 システムの運用に協力いただいた和歌山大学大学院平岡勇作氏と、説明者として主体的に活動していただいた大橋美佐緒氏と細井絵里香氏に感謝する。

参考文献

- [1] Heath, C. and Luff, P.: Disembodied conduct: Communication through video in a multi-media office environment, *Proc. SIGCHI Conference on Human Factors in Computing Systems, CHI '91*, pp.99-103 (1991).
- [2] 森川 治: 超鏡: 魅力あるビデオ対話方式をめざして, 情報処理学会論文誌, Vol.41, No.3, pp.815-822 (2000).
- [3] 平田圭二, 高田敏弘: 超臨場感を達成するための同室感というアプローチ, 電子情報通信学会誌, Vol.93, No.5, pp.410-414 (2010).
- [4] 吉野 孝, 藤田真吾: 重畳表示型ビデオチャットにおける枠の3次元的な移動と存在の効果, 情報処理学会論文誌, Vol.54, No.1, pp.249-255 (2013).
- [5] Towles, H., Chen, W.C., Yang, R., Kum, S.U., Kelshikar, H.F.N., Mulligan, J., Daniilidis, K., Fuchs, H., Hill, C.C., Mulligan, N.K.J., Holden, L., Zeleznik, B., Sadagic, A. and Lanier, J.: 3D Tele-Collaboration Over Internet2, *International Workshop on Immersive Telepresence, Juan Les Pins* (2002).
- [6] 末永 剛, 松本吉央, 小笠原司: 非拘束な運動視差提示3次元ディスプレイの提案と評価, ヒューマンインタフェース学会論文誌, Vol.9, No.2, pp.49-56 (2007).
- [7] 加藤 慶, 村上友樹, 中西英之: 可動式カメラによる社会的テレプレゼンスの強化, 情報処理学会論文誌, Vol.52, No.4, pp.1635-1643 (2011).
- [8] 田中一品, 加藤 慶, 中西英之, 石黒 浩: 人の移動の表現方法: ズームカメラと移動ディスプレイによる社会的テレプレゼンスの向上, 情報処理学会論文誌, Vol.53, No.4, pp.1393-1400 (2012).
- [9] 石井 亮, 小澤史朗, 川村春美, 小島 明, 中野有紀子: 映像コミュニケーションにおける窓越しインタフェースMoPaCoによるテレプレゼンスの増強, 電子情報通信学会論文誌D, 情報・システム, Vol.96, No.1, pp.110-119 (2013).
- [10] 平岡勇作, 宮田慎也, 坂本竜基: 前後移動時の運動視差を模した擬似視差映像の生成, 情報処理学会研究報告, 2012-HCI-146, Vol.2012, No.5, pp.1-5 (2012).
- [11] Buxton, W.: Telepresence: Integrating Shared Task and Person Spaces, *Proc. Graphics Interface 92*, pp.123-129 (1991).
- [12] de Greef, P. and Ijsselstein, W.: Social presence in a home tele-application., *Cyberpsychol Behav.*, Vol.4, No.2, pp.307-315 (2001).
- [13] Isaacs, E.A. and Tang, J.C.: What video can and can't do for collaboration: A case study, *Proc. 1st ACM International Conference on Multimedia, MULTIMEDIA '93*, pp.199-206 (1993).
- [14] Bondareva, Y. and Bouwhuis, D.: Determinants of social presence in videoconferencing, *Proc. Workshop on Environments for Personalized Information Access*, pp.1-9 (2004).
- [15] Prussog, A., Muhlbach, L. and Bocker, M.: Telepresence

in Videocommunications, *Proc. Annual Meeting of Human Factors and Ergonomics Society*, pp.25-38 (1994).

推薦文

本論文は、音声と映像を介したコミュニケーションにおいて、閲覧者の視点の動きに即した運動視差を映像に適用すると、テレプレゼンスが増強されるという研究である。先行研究により、深度センサ付きカメラにより検出した閲覧者の視点位置に基づき、対話相手の映像を疑似的な運動視差付き映像に変換するシステムを用いる方式が提案されているが、映像の見え方に不自然な部分をともなう場合があり、この完全ではない運動視差付きの映像がテレプレゼンスを増強させるかは不明であった。被験者実験により、疑似的な運動視差がテレプレゼンスに与える影響を明らかにした点を評価し、推薦する。

(グループウェアとネットワークサービス

研究会主査 市村 哲)



川路 崇博 (正会員)

1998年長崎県立大学経済学部流通学科卒業。2000年北陸先端科学技術大学院大学知識科学研究科博士前期課程修了。2008年同研究科博士後期課程修了。博士(知識科学)。現在、大月市立大月短期大学経済科准教授。創造性支援システムの研究に従事。日本情報経営学会、日本創造学会各会員。



坂本 竜基 (正会員)

2003年よりATR 知能ロボティクス研究所研究員、2008年より和歌山大学システム工学部講師を経て、2012年よりヤフー株式会社 Yahoo!JAPAN 研究所上席研究員。3次元ビジョンおよびインタフェースに関する研究開発に従事。博士(知識科学)。