

高品質 3D 立体視映像を用いた QoE 評価に関する検討

猪野 春記 由井 若菜 藤井 哲郎

3D 立体視映像は臨場感あふれる映像を遠隔地に提供でき、通信の分野ではテレビ会議或いは遠隔操作などにおいて有効である。3D 映像の伝送方式は、BS 3D 放送に代表されるように、サイドバイサイド方式と呼ばれる 960×1080 画素の左目及び右目の映像 2 枚を 1920×1080 画素の HD 映像 1 フレームとして伝送する方式が一般的である。ところがこの方式では、片眼の映像の解像度が水平方向半分に落ちてしまう。本稿では、片眼の映像の解像度を 1920×1080 画素のフル HD 映像の品質で伝送する方式の実現性を検討し、その特性を評価した結果を報告する。特に、IP ネットワークを用いて伝送された 3D 映像のビットレート及びパケットロスと映像品質の関係を主観評価実験により明らかにする。さらに、ネットワークを介したロボット操作を 3D 立体視映像にて行い、遠隔操作の作業性に関する評価結果も報告する。これらは、実用性のある 3D 立体視映像伝送を目指した基礎的な QoE 評価実験である。

キーワード：3D 立体視映像，3D 映像伝送，QoE 評価，伝送遅延，遠隔操作

1 まえがき

3D 立体視映像は通信の分野でも注目を集めており、遠隔地のオフィス間でコミュニケーションを行う際に高い臨場感でやりとりすることができると考えられている [1,2]。また、遠隔地の映像を立体的に映し出すことは、ロボット等の遠隔操作では必須条件である。災害現場或いは原子力発電所内での作業の様に人がアクセスできない場所において、遠隔操作ロボットを活用する機会も増えており、遠隔操作をサポートする 3D 立体視映像伝送技術の重要性はますます重要となっている。これに対応するためには、3D 立体視映像の伝送品質に関する正確な評価が大切である。

3D 立体視映像技術の最大の魅力は従来の 2D 技術には無かった“奥行き表現”が可能となることであり、表示される人物や物体との距離感、位置関係などを自然な形で伝えることができると期待されている。しかし、遠隔地との間で映像通信のやり取りをする際には映像のビットレートによりこの奥行き感が大きく変わってくる。また、最近幅広く利用されている IP ネットワークを用いた通信では、パケットロス或いは伝送遅延が必ず発生する。このような環境において、実用的な 3D 立体視映像

通信を実現するためには、伝送レート、伝送遅延、パケットロス等による劣化と利用者が感じる QoE (Quality of Experience) 品質との関係を明確にしておく必要がある [3]。本稿では、これらの基礎的な評価を行う。

遠く離れた場所で活躍するロボットを遠隔操作する為には、3D 立体視映像が非常に有効である [4,5,6]。この遠隔操作に必要な 3D 映像品質の評価も行う。伝送距離に応じた遅延も作業性に大きくかかわってくる。本稿では、3D 立体視映像を用いた遠隔地からの操作を行う場合の作業性についても定量的な評価を試みた結果を併せて報告する。これらの検討を通し、3D 立体視映像の伝送にかかわる諸問題点を QoE 評価の観点から明確にすることを目的とする。

2 高品質 3D 映像伝送システムの構築

従来、3D 立体視映像の実用的な伝送方式は、衛星放送と同様なサイドバイサイド方式に限られてきた。この方式では、片眼の解像度が 960×1080 画素に限られ臨場感にかけてしまう。本稿では、片眼の解像度が 1920×1080 画素のフル HD 映像での伝送の実現を目指す。片眼あたり 1920×1080 画素のフル HD 映像での伝送を実現する方式として、左右の映像を分けて並列に伝送する並列伝送方式、更に左右の映像を SDI 上で一つに束ねて伝送する多重伝送方式を実際に構築し、その特性を評価する。

2.1 並列伝送方式

3D カメラで撮影した映像を BlackMagic 社製

INO Haruki

東京都市大学環境情報学部情報メディア学科 2015 年度卒業生
YUI Wakana

東京都市大学環境情報学部情報メディア学科 2015 年度卒業生
FUJII Tetsuro

東京都市大学環境情報学部情報メディア学科教授

DeckLink 4K Extreme を搭載した PC により左右の眼に対応する 2 本の映像ストリーム (1080/60i) に分離する。この 2 本の映像は夫々がフル HD であり、この 2 本の映像を夫々 Nimbus 社製の映像伝送装置で別々に符号化伝送する。このシステム構成を図 1 に示す。エンコーダーが WiMi6400T、デコーダーが WiMi6400R である。並列に伝送された 2 本の映像は、Matrox 社の MC-100 によりフレームシーケンシャル 3D 映像信号に変換され、HDMI ケーブルにより 3D 液晶 TV に送られ、スクリーン上に映し出される。

2. 2 多重化伝送方式

MC-100 には、2 本の 1080/60i 映像を SDI として多重化し、一本のストリームに変換する機能が有る。Nimbus 社の WiMi6400 はこの多重化された映像を符号化し伝送できる。この機能を活用し、1 ストリームに多重化された 3D 映像の伝送を行う。このシステムの構成を図 2 に示す。

並列伝送方式と同様に、撮影した 3D 映像を DeckLink 4K Extreme を搭載した PC により左右の眼に対応する 2 本の映像ストリームに分離する。この 2 本のフル HD 映像ストリームを MC-100 にて多重化された SDI 信号に変換する。この多重化された SDI 信号を WiMi6400R&T を用いることにより映像符号化伝送する。復号化された多重化伝送された映像は、MC-100 により 2 本のフル HD 映像ストリームに分解され、再度 MC-100 を用いてフレームシーケンシャル 3D 映像信号に変換し、3D 液晶 TV 上に映し出す。

2. 3 高品質 3D 映像伝送システムの構築

3D 立体視映像の撮影には、左右の映像を夫々フル HD で撮影可能な Panasonic 社製の HDC-Z10000 を用いる。2 個のレンズ間距離は 41mm である。このカメ

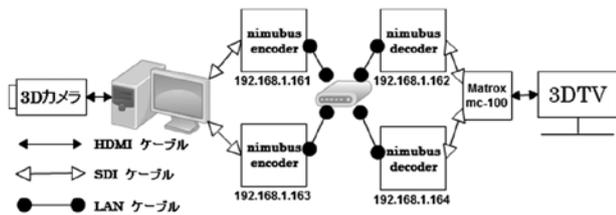


図 1 3D 映像並列伝送方式のシステム構成

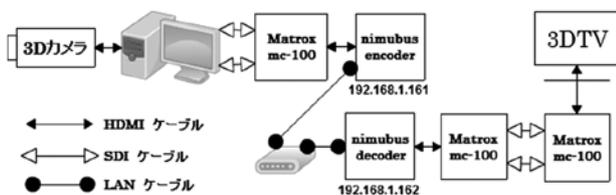


図 2 3D 映像多重化伝送方式のシステム構成

ラ HDC-Z10000 はそれぞれフル HD 映像の左右の眼の映像 (1080/60i) をフレームシーケンシャル 3D 映像信号として HDMI より出力できる。残念ながら、この HDMI 信号を直接伝送できる映像伝送装置は存在しない。そのために、前述の並列伝送或いは多重伝送方式のシステムを構築する必要がある。

3D カメラを本学情報基盤センターの映像メディアルーム内のスタジオに設置し、同センター内の映像編集室に伝送し、高品質に 3D 映像を表示するシステムを構築した。この 2 か所は 10GbE のイーサネットにて接続されている。この 2 箇所間にネットワークシミュレータである Netem 機能を搭載したパソコンを挿入し、様々な IP ネットワーク環境をエミュレートできる環境を実現する。伝送遅延を 1msec から 1 秒までの間で設定可能であり、パケットロスも 0% から 100% の間で任意に発生させることができる。

フル HD 映像の伝送には Nimbus 社製のエンコーダー WiMi6400T 及びデコーダー WiMi6400R を用いる。この装置は映像符号化方式として H.264 を採用し、1080/60P のフル HD 映像を伝送できる。ビットレートは 0.1 ~ 30Mbps まで自由に設定可能である。伝送プロトコルとして UDP を用いており、低遅延にて双方向の通信が実現できる。この CODEC を標準モードで用いると片方向で 180msec の遅延が発生する。3D 映像の表示には、SONY 社製の KD-55X8500A を用いる。このテレビは、フレームシーケンシャル 3D 映像信号を HDMI より受け取り、3D 立体視表示することができる。3D 立体視映像を観る際は、アクティブシャッター方式の液晶メガネを装着して映像を評価することになる。

3 高品質 3D 立体視映像の QoE 評価

3. 1 SD 法による 3D 立体視映像の評価

本節では、SD 法で 3D と 2D 映像の印象を確認し、3DTV 会議に有効なシーンを選定する。そして、SD 法で選定した映像を使用し、ビットレートを変化させた時とパケットロスが発生した時の影響をそれぞれ調べて 3DTV 会議の有効性の検証を行う。まず、評価画像の撮影を行う。3D カメラを TV の上に設置することを配慮し、被写体との間隔を TV の 3H とした。撮影対象として TV 会議によるグループディスカッションを想定し、1 人、机横 4 人、机縦 5 人、机横 5 人、2 列 5 人と変化させ、それぞれ 10 分間撮影した。撮影した評価用映像から 5 分を切り出し、2D (左目映像のみ) と 3D 映像を交互に被験者に見せる。そして、形容詞とその反対語の組み合わせ (例えば、なめらか、粗い) を 27 組用意し、映像を見ながら回答用紙の該当する箇所を一つ選んで記入させた [7]。

被験者 6 人にて評価を行った結果、机横 4 人が最も

評価が高く、逆に1人は最も評価が低かった。最も評価の高い机横4人の結果を図3に示す。図3-1が3Dの結果を示し、図3-2が2Dの解析である。色の濃い箇所が良い印象を持ったところである。机横4人の映像の印象を詳細に比較すると、3D映像は派手さや迫力点で評価が高くなり、印象に残り親しみやすさを得られるというプラスの結果が出ている。しかし、見易さの点では2Dの方が上回った。これは、3D映像を観る際の3Dメガネの着用や長時間観ていることによる疲労が原因と考えられる。以上の結果より、5段階主観評価実験を行うにあたり、特徴的なシーンとして、最も印象の良い机横4人と悪い1人、さらにその中間である2列5人の3種類の3D立体視映像を選定した。

3. 2 高品質 3D 映像伝送における QoE 評価

SD法の選定より印象の最も良かった机横4人と最も悪かった1人の映像、さらに中間的な5人2列の3種類を用いて主観評価実験を行う。映像伝送の観点より、ビットレートとパケットロスを変化させ、5段階の主観評価実験を行う。まず、図1&2に示された3D映像伝送システムを用いて評価素材の録画を行う。並列伝送方式では、一本のストリームについて1 Mbps, 2.5 Mbps, 5 Mbps, 7.5 Mbps, 10 Mbps, 12.5 Mbps,

15 Mbps と変化させ、多重伝送方式では多重化されたストリームに対して2Mbps, 5Mbps, 10Mbps, 15Mbps, 20Mbps, 25Mbpsと変化させる。多重伝送は並列伝送の2倍のビットレートに設定されている。パケットロスはロス率を0%, 0.01%, 0.03%, 0.05%, 0.1%と設定した。

撮影済みの3D映像を3Dカメラで再生し、映像伝送装置で伝送する。この時、ビットレート及びパケットロスを様々に変化させ、受信された映像を録画し、この録画した映像を編集して主観評価実験を行う。受信側の3D映像の録画にはBlackMagic社製Media Expressを用いて3D映像を2チャンネルのフルHD映像として録画する。これを編集し、主観評価実験を行う。これにより、全ての評価者が同じ3D映像を観ながら主観評価実験を行うことが可能となる。

被験者は15名(男11, 女4)、3D立体映像の見易さに関して、1(非常に悪い)、2(悪い)、3(普通)、4(良い)、5(非常に良い)の5段階評価で実験を行った。また、パケットロスに関しては、多重伝送方式のみの評価を行った。並列伝送方式については、片目のみをランダムにパケットロスさせると生理的な不快感が生じ、実験そのものが危険であった為、実験を取りやめた。評価方法に関しては、パケットロスが1(非常に邪魔になる)、2(邪魔になる)、3(邪魔になるが気にならない)、4(わかるが気にならない)、5(わからない)の5段階評価で実験を行った。なお、最適視聴距離3Hであるディスプレイから204cm離れて評価実験を行った。

ビットレートを2~25Mbpsまで変化させ、5段階の主観評価実験を行った結果を図4, 5, 6に示す。いずれの図においても、伝送ビットレートがほぼ10Mbpsを超えるとほぼMOS値も3を超える。3D立体視映像をH.264方式を用いてテレビ会議を行う場合、最低でも10Mbpsは必要と判断される。なお、いずれの図においても、片眼が見る映像に関しては、2D映像は3D映像の倍のビットレートとなり、より精細な映像である。即ち、トータルの伝送レートで比較を行っている。

図4より、並列伝送方式の3D映像の方がより2D映像よりはるかにMOS値が高くなっている。これは、明らかに3D映像により臨場感が高くなり、QoEに関して効果的であると判断できる。多重伝送の評価が低いのは、5人が並んだ映像では、H.264による符号化の効率が上がらないためと推定される。図5においては、多重化伝送方式が最も高いMOS値となっている。これは、多重化された映像の方が符号化効率が上がることを示している。

図6より、臨場感に乏しい一人の映像では、いずれの方式ともに差が無くなること解る。特にビットレートが高いところでは、一眼あたりのビットレートが高くなる2Dの方が僅かではあるがMOS値が高くなっている。

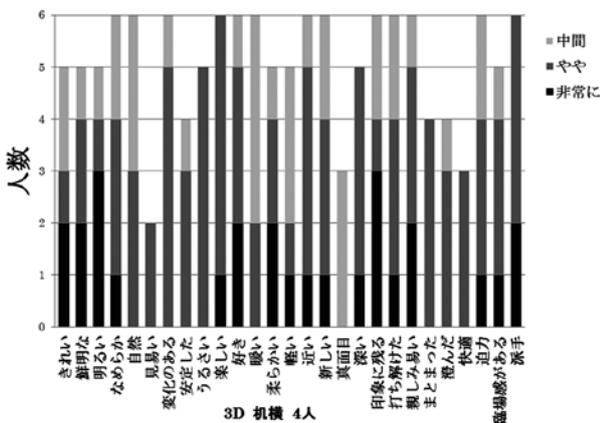


図3-1 「机横4人」シーンの3D映像の印象

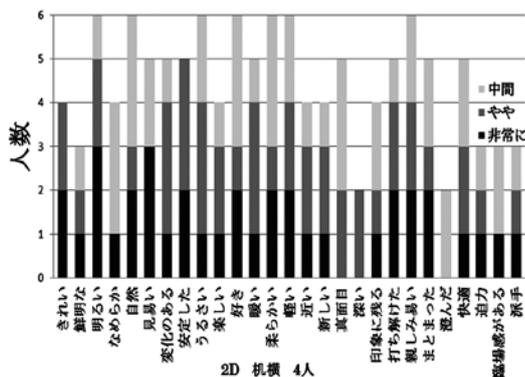


図3-2 「机横4人」シーンの2D映像の印象

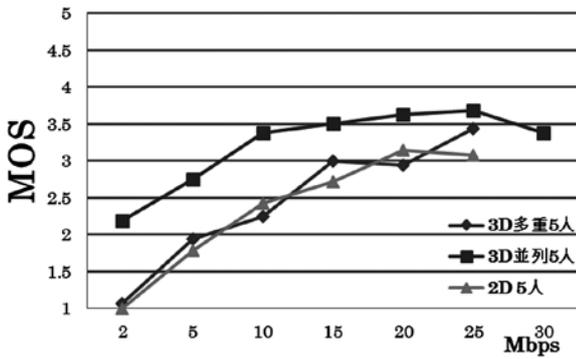


図4 「2列5人」のシーンでの3種類の伝送方式における伝送ビットレート毎の主観評価実験結果

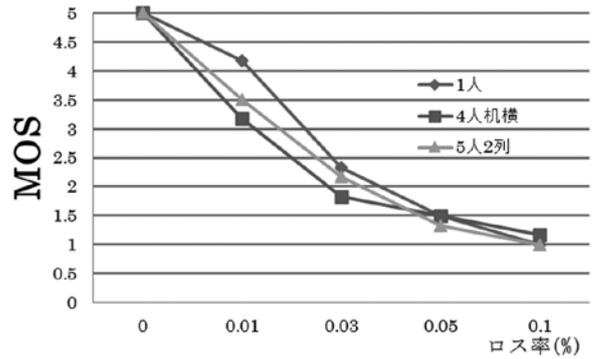


図7 3D 多重伝送方式時のパケットロス

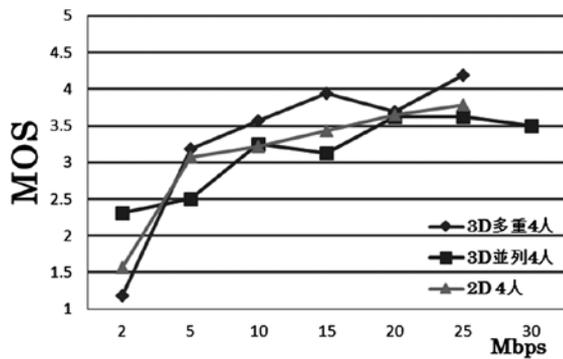


図5 「4人」のシーンでの3種類の伝送方式における伝送ビットレート毎の主観評価実験結果

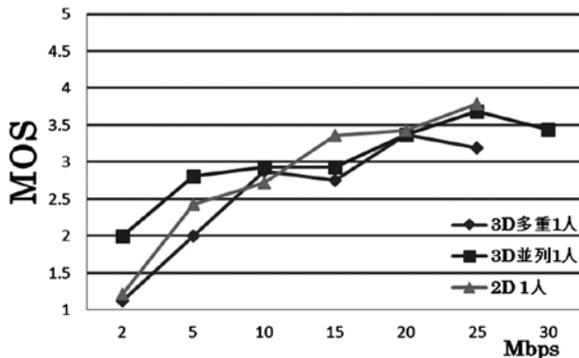


図6 「一人」のシーンでの3種類の伝送方式における伝送ビットレート毎の主観評価実験結果

る。これは、必ずしも 3D であれば、臨場感が高くなり、どの様な場合でも MOS 値が高くなるという訳では無いことを明確に示している。即ち、3D 立体視映像を用いたテレビ会議において、被写体の人数、配置によって見易さが大きく異なることが示されたといえる。

パケットロスを発生させ、3D 映像の主観評価実験を行った結果を図7に示す。3D 映像伝送システムのエンコーダーとデコーダーの間に回線遅延シミュレータを設置し、パケットロスを発生させ、被験者 10 名 (男 6, 女 4) を対象に主観評価実験を行った結果である。3D

映像伝送方式としては、多重伝送のみについて行っている。並列伝送では、パケットロスが発生した映像は生理的に危険な要素があり、今回中止した。パケットロス率は、0%から 0.1%の間で変化させた。0.1%を超えると、映像が激しく乱れ、映像がハングアップした状態になることを確認した。映像の伝送ビットレートは、20Mbpsとした。同図より、パケットロスが 0.03 以上で MOS 値が 3 以下になり、利用に適さないことが解る。最低限でもパケットロス値は 0.02%以下が求められる。また、SD 法の評価の高いシーンほど、パケットロスによる影響を強く受けていることも明らかである。

4 遠隔作業における QoE 評価

4.1 ロボットアーム遠隔操作システムの構築

3D 立体視映像を用いた遠隔操作の作業性評価を行う為に、ネットワークを介してロボットアームを遠隔操作するシステムを構築する。このシステムを用いて、被験者にタスクを課しその作業時間を計測することにより、3D 立体視映像を用いた遠隔操作の作業性評価を行う。遠隔操作するロボットアームにはイーケイジャパン社製の MR-999R を用いる。これは、5つの関節を持つロボットアームであり、Amazon 等で簡単に入手できる。この装置を図1に示された双方向 3D 映像通信システムの方端である情報メディアルームのスタジオ内に設置した。映像編集室のパソコンから操作する為に、同社製の MOVIT-LAB2 IF-100 (制御インターフェイスボード) と付属の制御ソフトを用いる。この制御ソフトの操作画面を Windows パソコンのリモートデスクトップ機能を用いて遠隔地のパソコンの画面に表示し、遠隔地からロボットアームを操作する。即ち、情報メディアルーム内にロボットアームとインタフェースを介して接続されたパソコンを設置し、その画面をリモートデスクトップ機能により映像編集室のノートパソコンに表示し、映像編集室からロボットアームの遠隔操作を行う。

ロボットアームの様子を撮影する 3D 映像カメラは前節と同様に HDC-Z10000 を用いる。映像伝送には、図

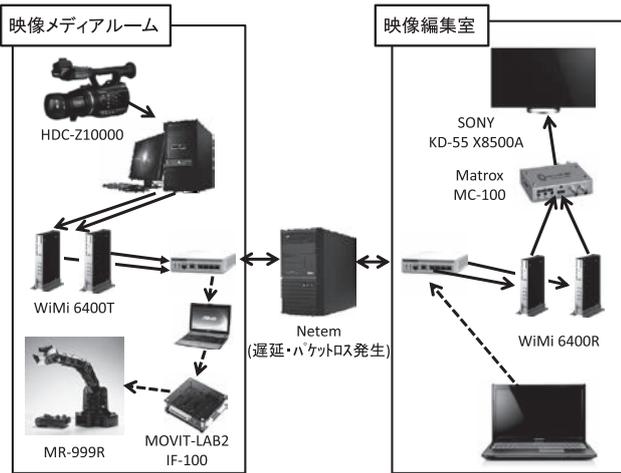


図 11 ロボットアーム遠隔操作システム

1の並列映像伝送システムを用いる。映像は映像編集室に伝送し、55インチの3D液晶ディスプレイに表示する。液晶シャッター付の3Dメガネを装着して3D映像を観る。この映像を見ながら、ロボットアームを遠隔操作する。構築したロボットアームの遠隔操作システムの外観を図12と13に示す。

様々なネットワーク環境下における3D立体視映像を用いた遠隔操作の操作性を評価する為に、ビットレートを変化させその作業時間の計測を行う。さらに、IP伝送における遅延を発生させ計測する。この時、Netemにより0msから250msまでの遅延を発生させ、評価を行った。比較の為に、同じネットワーク環境下における2D映像を用いた操作性の評価実験も行う。なお、映像を伝送する際には、CODECにおいても遅延が発生し、その遅延時間は180msである。

4. 2 遠隔操作の作業性評価

被験者10人に対し、遠隔操作の作業性に関する評価実験を行う。タスクとして、マス目のついた画用紙の上に高さ10cmの盛槽用の木製棒3本を10cm間隔に設置し、この棒を20cm奥のマークされた位置に移動させる作業を課す。この作業に要する時間を計測し、操作性の評価を行う。実験を繰り返すことで、習熟度が高まり操作時間が短くなる事が予測されるので、事前に一連の作業を2回行い、その後2回テストを行う。各テストの作業時間差が5秒以内だった時に、評価実験を開始する事とした。

ビットレートを1, 3, 5, 10, 20, 30Mbpsと変化させ、2D及び3D立体視映像を用いた遠隔操作の作業性に関する実験結果を図14に示す。同図より、2D映像を用いた実験では、ビットレートの変化による作業性に関して大きな差は見られなかった。これに対し、3D立体視映像を用いた遠隔操作では、ビットレートが上が

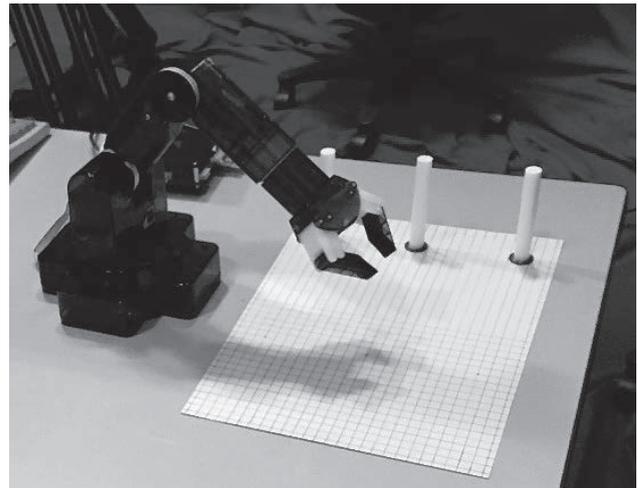


図 12 ロボットアーム(イーケイジャパン社製 MR-999), 但し、棒に関しては盛槽用の棒に変更



図 13 3D 立体視映像を用いた遠隔操作端末

るにつれ、作業時間が確実に短縮された。これは、ビットレートの向上に伴って画質が向上し、3Dとして奥行きを知覚しやすくなり、作業時間が短縮されたと推察される。全体として、2D映像に比べ、3D映像を用いた作業の方が30秒以上作業時間が短く、操作性に優れていることが明かである。

伝送遅延時間を0, 50, 100, 150, 200msと変化させ、2D及び3D立体視映像を用いた遠隔操作の作業性に関する実験結果を図15に示す。伝送レートは、最も高品質な映像伝送が可能となるようにビットレートを30Mbpsに設定した。従って、トータルのビットレートはこのシステムで最高品質の60Mbpsである。同図より、伝送遅延時間が50msでは、作業時間にあまり差がないことがわかる。しかし、遅延が大きくなるにつれ2D映像では大幅に作業時間が長くなってしまいうのに対し、3D立体視映像を用いた場合は作業時間の増加が抑えられている。その差は最終的に1分20秒に達している。明らかに遅延に関して、3D立体視映像を用いた映像の

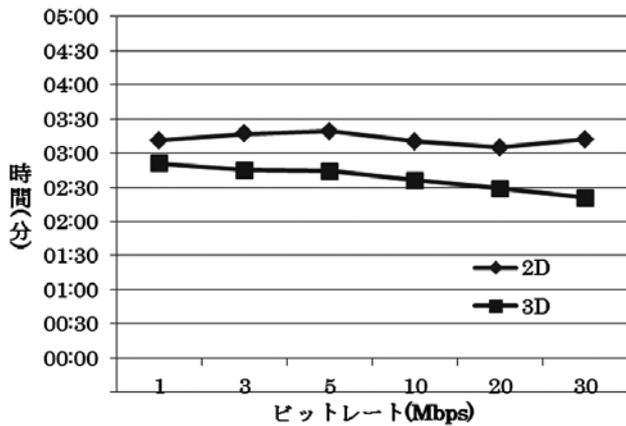


図 14 ビットレートの変化に伴う 3D 立体視映像を用いた遠隔操作の操作性評価

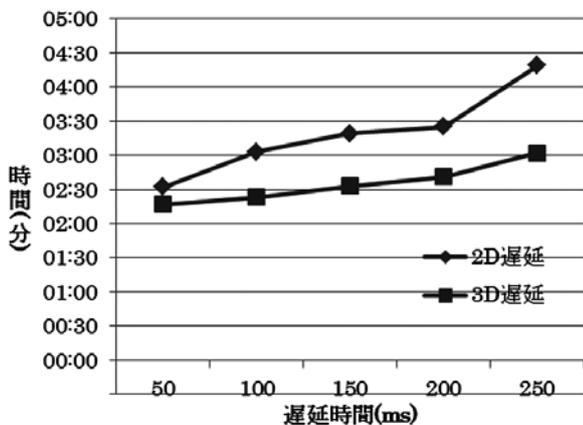


図 15 遅延時間の変化に伴う 3D 立体視映像を用いた遠隔操作の操作性評価

作業優位性が確認された。これは被験者が 3D 立体視映像を基に、空間的な位置の予測が立て易い為と思われる。この奥行き感が重要なポイントである。

ロボット遠隔操作において、映像品質、遅延特性共に 3D 立体視映像を用いた方が、短い操作時間で効率良く遠隔操作を行えることが定量的に明らかとなった。遠隔操作を行う際、3D 立体視映像を用いた作業の奥行き感が大切と思われる。

5 まとめ

3D 立体視映像を用いた遠隔地との通信において、映像の伝送ビットレート及びパケットロスが大きく通信品質に関わることが示された。特に、3D 立体視映像を用いた H.264 方式のテレビ会議においては、実用的な品質を確保するためには 5Mbps 以上の伝送レートが必要なこと、パケットロスは、0.02% 以下であることが望ましいことが示された。それと同時に、被写体映像の構成も大きく係っていることも明かにされた。3D 映像は奥行き感をユーザに提供するシステムであり、奥行き

感の無い映像は 2D 映像伝送にてビットレートを上げる方が適していることも示された。また、ロボットの遠隔操作において、3D 立体視映像は非常に有効であり、操作時間が短くなり、効率よく遠隔操作の作業が行えることが定量的に明らかとなった。

謝辞

本研究は JSPS 科研費 25420388 の助成を受けたものである。

参考文献

- [1] 山崎, 西, 野中「3D 遠隔オフィス間における実用的な 3D 映像通信へ向けた主観評価実験」電子情報通信学会, CQ 研究会, CQ2012-36, 2012 年
- [2] 山崎, 西「パケットロス環境における 3D 映像通信に向けた主観評価実験」電子情報通信学会, CQ 研究会, CQ2013-23, 2013 年
- [3] 池田, 藤井「3D 立体視映像を用いた双方向通信における QoE の一検討」電子情報通信学会, 2015 年総合大会, A-16-5, 2015 年 3 月
- [4] 遊佐, 玉手, 藤井「3D 立体視映像を用いた映像通信の QoE 評価に関する検討」東京都市大学、横浜キャンパス、情報メディアジャーナル、第 16 号、2015 年 4 月
- [5] 古屋, 栗生, 清水「3D 画像と体感型操縦を用いた「次世代無人化施工システム」」大林組技術研究所報, No.76, No.16, 2012 年
- [6] 古屋, 清水「簡易 VR を用いた「次世代無人化施工システム」の開発」建設の施工企画, pp.53-59, 2012 年 11 月
- [7] 宮川洋「テレジョン画像の評価技術」コロナ社, 1986 年