

遠隔手話通訳のための実写立体映像の伝送方法に関する検討

筑波技術大学産業技術学部産業情報学科

若月大輔, 塩野日剛亮, 加藤伸子, 河野純大, 村上裕史, 皆川洋喜, 西岡知之, 内藤一郎

要旨 :

聴覚障害者の社会参加を支援するための遠隔コミュニケーションに関する研究開発を進めている。本研究では遠隔手話通訳の読みやすさを向上させるために、立体映像を用いた遠隔手話通訳システムの試作した。映像に奥行映像を分割して付加することによって、通常の映像伝送装置を用いた立体映像の伝送方法を提案し、その評価を行った。また、試作システムを実際の遠隔講義へ導入した事例、およびイベントでの一般公開の実施について紹介する。

キーワード : 実写立体映像, 手話通訳, 遠隔情報保障, 聴覚障害者支援

1 はじめに

社会参加する聴覚障害者を支援するために、図1に示すような遠隔情報保障システムを構築し、遠隔から手話通訳や字幕通訳を提供する遠隔コミュニケーション技術に関する研究や実験の支援を進めてきた [1, 2, 3, 4, 5]。本研究では、より臨場感の高い遠隔情報保障システムの実現を目指して、実写立体映像を用いた遠隔手話通訳について検討を行っている。

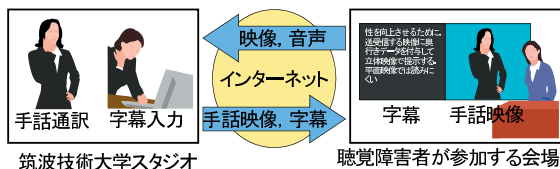


図1 遠隔情報保障システム

手話は手指の形や動きに加えて、頭の動きや表情などの空間的な表現を用いる言語であり、形や動きの奥行が内容を正確に伝える上で重要となる場合が多い。従来の遠隔手話通訳では、テレビ会議システム等の映像伝送装置を介して情報保障が必要なユーザ（被情報保障者）に提示される。しかし、映像になった手話は平面的になるため読みにくかったり、誤解を生じたりする原因となる。一方、平面的な映像でも手話を読みやすくするように奥行をできるだけ使わないように手話表現を工夫する通訳者もみられる。しかし、スムーズに表現するには熟練が必要であり、マニュアル等も整備されていないため習得が容易ではないことが指摘されている [6]。

本研究では、遠隔手話通訳をはじめとした遠隔コミュニケーションの可読性や品質を改善するために、手話の

立体映像を被情報保障者へ提示するシステムの研究開発を進めている。これまで、立体映像をリアルタイムに撮影するカメラシステムを試作し検討を進めてきた [7]。本報告では、映像に奥行情報を付加して伝送し、立体映像として提示する方法を提案する。提案法を実装したシステムを試作してコンピュータグラフィックスによる立体映像の奥行提示実験を実施し、立体映像の伝送と提示の特徴と有効性について基礎的検討を行った。

2 関連技術・研究

3 立体映像の撮影と伝送

立体映像を撮影する主な方法として、2台以上のカメラを水平に配置したステレオカメラやカメラアレイを用いて視差のついた映像を撮影する方法がある [8, 9]。撮影した視差映像を観察者の左右の目に対応させて提示することで、立体映像を観察することができる。この方法で撮影した映像を伝送する場合、左右の映像を同期させる必要があるため、図2のように左右の画像のサイズを縮小して1フレームに収める方法が実用化されている [10]。しかし、再生時に元のサイズへ伸張して表示するため実質的な解像度低下が発生し手話や表情が読みにくくなる可能性がある。図2の場合では横方向の解像度が1/2になる。

撮影対象の奥行を取得する方法としては、視差映像を用いて三角測量や影から画像処理によって得る方法が代表的であり、製品化されているものもある [11]。ステレオカメラによる方式はカメラシステムが高精度にキャリブレーションされていれば広範囲の距離計測が可能であるが、特徴の少ない映像の場合は正しく奥行が得られな

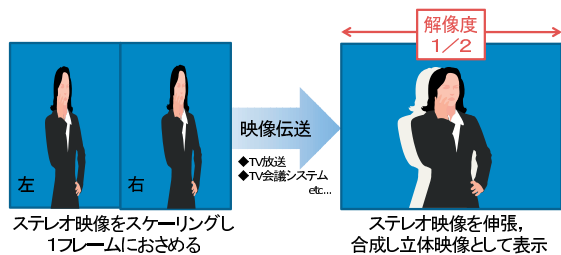


図2 視差映像を伝送する場合の立体視の問題点

い場合がある。一方、対象に赤外線光を照射して反射光の位相のずれによって対象物までの距離を得る方法が近年実用化されてきた[12]。赤外線光を照射する必要があるため計測範囲が制限されるが、処理が単純なことから高速かつ精度よく奥行を取得することが可能である。今回の撮影対象は手話であり限られた空間のみの撮影になるため、後者の方法を採用し立体映像撮影カメラシステムを試作した。

3.1 立体映像の表示

立体映像を観察者に提示する代表的な方法として、液晶シャッターや偏光フィルタを用いる方法がある。これらの方法はユーザが専用のメガネを着用して立体映像を観察する。一般的に通訳を含めた手話によるコミュニケーションでは手指の形や動きだけでなく表情也非常に重要である。メガネ等で被情報保障者の表情が隠れた場合、通訳者と被情報保障者の間で十分なコミュニケーションがとれない可能性がある。そこで、本研究ではIP方式(Integral Photography)の裸眼立体視ディスプレイを用いて立体映像の提示を行った。

4 試作した立体映像撮影・伝送・表示システム

4.1 立体映像撮影カメラシステム

筆者らはこれまでステレオカメラ[11]を用いて奥行映像を取得し、カラー映像をテクスチャマッピングすることによって立体映像を提示する方法を提案し検討を行ってきた[13]。奥行映像とは、各画素にカメラからの撮影対象までの距離(奥行値)を記録した画像である。立体映像をリアルタイムに撮影し、表示することが可能となったが、手話通訳時の無地の背景など特徴が少ない部分で奥行きを正しく計測できない場合が多く立体映像が乱れるなどの問題も明らかになった。

この問題を解決するために試作した立体映像撮影カメラシステムを図3に示す。奥行映像を撮影するカメラとしてSR-3000(MESA社)[12]を採用した。同カメラは赤外線を対象に照射し反射光の位相のずれによって奥行値を計測するため、特徴が少ない面でも安定に奥行映像を得ることができる。カラーカメラGrasshopper

(PointGreyResearch社)と奥行カメラSR-3000を光軸方向を平行に配置してカメラシステムを構成した。両カメラの内部、外部パラメータについては手動で調整を行った。Grasshopperの最大解像度1600×1200であり、一般的な解像度を自由に選択できる。また、SR-3000の解像度は176×144であり、計測可能奥行が7.5m、奥行方向の分解能は1mの距離で6mm、2mの距離で13mm程度である。



図3 立体映像撮影カメラ(左：奥行、右：カラー)

4.2 立体映像の表示

奥行き映像から対象物のサーフェイスモデルを生成し、そのサーフェイスモデルにカラー映像をマッピングして3次元コンピュータグラフィックス(CG)で立体映像の表示を行う。本カメラシステムでは図4(a)に示すようにカメラの位置が水平方向にならんでいるため、このオフセットを考慮した立体映像を生成する必要がある。

奥行映像の撮影位置 O_d と各画素の奥行値 p_{di} ($i = 0 \sim n$)から逆投影によって各画素に対応した3次元頂点 v_{di} を求める。互いに隣接する頂点を結びポリゴンメッシュを生成し、撮影対象のサーフェイスモデルを構築する。次にサーフェイスモデルの3次元頂点 v_{di} をカラーカメラの撮影位置 O_c に対して投影し、カラー画像上の座標値 p_{di} を求める。 p_{di} をテクスチャ座標 t_{di} に変換し、対応する v_{di} に t_{di} を対応づけてカラー画像をテクスチャマッピングして描画する。これにより、図4(b)に示すように奥行映像から生成されるサーフェイスモデルに対して、カラー映像を正確にマッピングすることで3次元CGを立体映像としてリアルタイムに生成する。

4.3 立体映像の伝送

図2で示したように視差映像を1フレームにおさめて映像の伝送や記録を行う場合、伸縮した分だけ解像度が低下してしまう。そこで、カラー映像と奥行映像による立体映像の送受信を提案する。ここでは奥行映像の解像度はカラー映像の解像度にくらべて低くても十分な立体

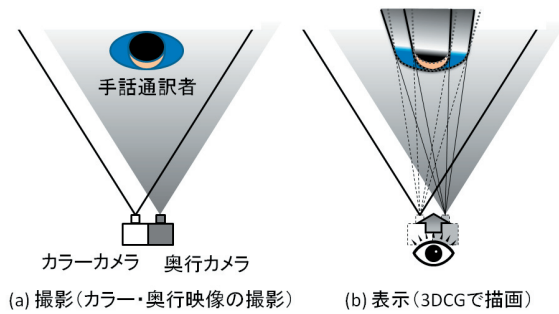


図 4 立体映像の撮影と表示

感が得られる立体映像を提示できると仮定する。

図 5 に立体映像の伝送の流れを示す。今回はカラー映像の解像度を 800×600 に設定した。また、奥行映像の最大解像度 176×144 である。奥行映像の高さはカラー映像の約 $1/4$ であるため、奥行映像を分割して埋め込み、映像を (a)～(c) の流れで伝送する。

(a) 奥行映像の分割とカラー映像との合成

(b) 合成した映像の伝送

(c) 奥行映像の分離と復元

(a) 奥行映像を図 5(a) のように 4 等分割して、カラー映像の横に伸張して合成する。(b) 映像伝送に用いるテレビ会議システムが解像度へ合成した映像をスケーリングして遠隔地へ伝送する。(c) 受信した合成映像の奥行部分を分離し奥行映像の復元を行う。分離したカラー画像と復元した奥行映像から前節で述べた方法で立体映像を表示する。

前節で述べた立体映像撮影カメラシステムと立体映像表示システムを実装し、インターネットを介さず両システムを直結してテストを行った。その結果、立体映像の撮影、伝送用映像の合成、復元処理、および表示までの一連の流れを遅延なく 30fps で実行可能であった。

5 実験・結果、公開

5.1 奥行映像の解像度と立体感についての調査

第??節で述べたように、提案した立体映像の伝送方法ではカラー映像の解像度よりも奥行映像の解像度のほうが低くなる。このため、奥行映像の解像度低下が立体映像に与える影響に関する実験を行った。

被験者は聴覚障害を持つ大学生の男女 7 名（男性：5 名、女性 2 名）とし、実験用立体映像は 42 インチの裸眼立体視ディスプレイ（PHILIPS 3D Display）に表示する。被験者はそれを 2 メートル離れた位置から観察する。図 6 で示すように CG で描画した回転するティーポット

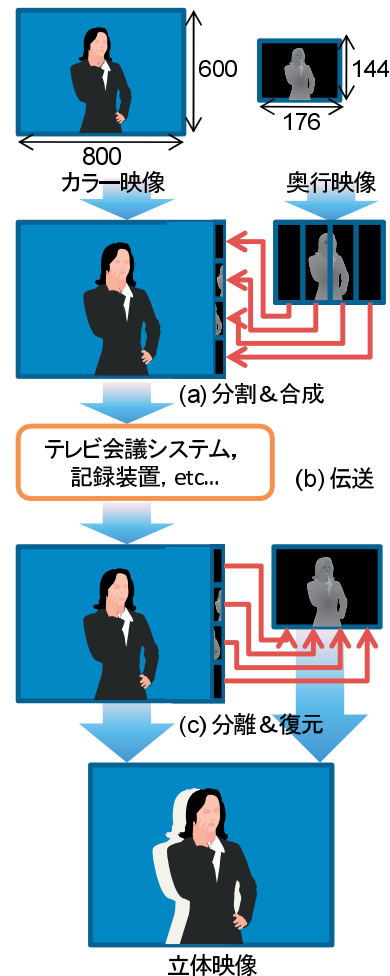


図 5 立体映像伝送の流れ

を実験映像として、立体視ディスプレイに入力する映像の解像度を 800×600 に設定した。映像に対応する奥行映像の解像度を $1/1$ 倍 (800×600)、 $1/2$ 倍 (400×300)、 $1/4$ 倍 (200×150) および $1/8$ 倍 (100×75) の 4 段階に変化させ、それぞれについて一対比較を行った。被験者には奥行映像の解像度を変化させていることを伝えず、各実験映像を 20 秒間見せて評価させた。「自然な立体に見えたのはどちらか」、「きれいな映像に見えたのはどちらか」の 2 項目についてそれぞれ 7 段階で回答させた。

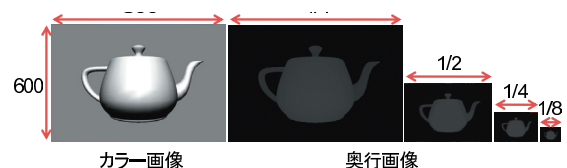


図 6 実験映像と奥行映像の解像度変化

結果を図 7 に示す。グラフの上下に比較した奥行映像の解像度を記し、棒グラフは平均値、エラーバーは標準

偏差を表す。評価値 4 が「どちらともいえない」の評価である。平均値を見ると、解像度が 1/8 の場合（図 7 の (a)～(c)）について評価が低い傾向が見られた。また、「輪郭部分がギザギザに見えた」、「動きに違和感があった」などの記述もあった。これは奥行映像の解像度不足で生成されるポリゴンモデルが粗くなりすぎたことが原因であると考えられる。一方、1/4 以上の場合（図 7 の (d)～(f)）の比較についてはほとんど差が見られない結果となった。したがって、今回試作したシステムの奥行映像の解像度は前節で述べた仮定通りに立体視にほとんど影響がないことが示唆された。

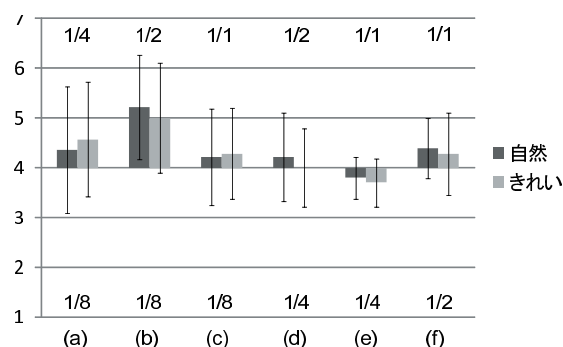


図 7 奥行映像の解像度と立体感に関する評価結果

次に、奥行映像の解像度を徐々に変更させた場合の立体感への影響について実験を行った。カラー映像に対する奥行映像の解像度の倍率が 1.0 → 0.0（解像度：800x600 → 0x0）へ変化する立体映像に対して被験者が違和感を感じる倍率と、逆に 0.0 → 1.0（解像度：0x0 → 800x600）へ変化する立体映像に対して違和感を感じなくなる倍率について調査を行った。奥行き映像の解像度は 90 秒間で変化させ、違和感を感じた／感じなくなった瞬間の奥行映像の倍率を計測した。結果を図 8 に示す。図中の◇は平均値、エラーバーは標準偏差を表す。平均値は解像度高→低で約 0.35 倍（解像度：280x210 程度）、解像度低→高で約 0.5 倍（解像度：400x300 程度）であった。しかし、標準偏差がそれぞれ約 0.3、約 0.24 と大きく非常にばらつきがある。

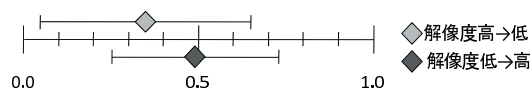


図 8 奥行映像が徐々に変化する場合の違和感について

調査と一緒に記述してもらった判断理由を参照したところ、「画面がちらついたため」、「回転のアニメーションが一瞬ひっかかったため」という奥行映像の解像度変化による違和感とは考えにくい回答があった。そこで、この

2 名の結果を除いた 5 名で集計してみた結果を図 9 に示す。平均値は解像度高→低で約 0.2 倍（解像度：160x120

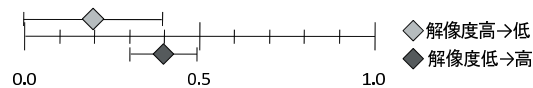


図 9 奥行映像が徐々に変化する場合の違和感について

程度）、解像度低→高で約 0.4 倍（解像度：320x240 程度）であり、標準偏差がそれぞれ約 0.2、約 0.1 であった。

この結果から、奥行映像の解像度がカラー画像の約 0.2 倍（解像度：160x120）以下となった場合にほとんどのユーザが立体感に違和感を感じる可能性が高いことが示唆された。これは、先の実験の 1/4 = 0.25 程度でも立体感への影響がみられなかった結果と矛盾しない。

5.2 試作システムの公開

本研究の立体映像を用いたリアルタイム遠隔情報保障への取り組みを公表するために、遠隔講義での活用、ならびにイベントでの試作システム公開を行った。

平成 20 年 1 月～3 月にかけて筑波技術大学主催で実施された「筑波技術短期大学卒業生対象スキルアップ講座」において本システムも用いて遠隔講義を行った（図 10）。東京（丸ビル）で行われた講義を撮影してテレビ会議システムで筑波技術大学に伝送し、立体映像として学生に提示した。立体映像と通常の 2D 映像で比較したところ、疲労感はあるが立体映像は講師の存在感がある等のコメントが遠隔で受講した学生から寄せられた。

また、平成 20 年 4 月 26、27 日につくばエキスポセンターの「サイエンスシティ つくば再発見・研究機関等紹介コーナー」にて、本システムを一般者向けに公開した（図 11）。ここでは、遠隔ではなく自分を撮影した立体映像を即座に裸眼立体視ディスプレイに映し出すデモンストレーションを実施した。フレームレートは 30fps で、遅延もなく提示することができた。手をのばして立体感を確かめる様子も見られ、ほとんどの体験者が立体感を得ることができたようである。

6 まとめ

本研究では、遠隔手話通訳の可読性や品質を改善するために、手話の立体映像をリアルタイムに撮影、伝送して立体映像を提示するシステムを試作した。映像に奥行映像を分割して付加することによって従来の映像伝送装置などを介して立体映像を遠隔地へ伝送する方法を提案した。

提案法を実装した試作システムを用いて、奥行映像の解像度と立体映像の見え方について調査を行った。その結果、カラー映像（解像度：800x600）に対する奥行き



図 10 筑波技術短期大学卒業生対象スキルアップ講座の様子（東京⇄筑波技術大学，平成 20 年 1 月～3 月）



図 11 サイエンスシティつくば再発見・研究機関等紹介コーナーの様子（つくばエキスポセンター，平成 20 年 4 月 26 日（土）～27 日（日））

映像の解像度が 0.25 倍程度でも自然に立体映像が観察できることが示唆された。また，今回試作システムを実際の遠隔講義へ導入した事例，およびイベントでの一般公開の実施について紹介した。

今後は，手話の実写立体映像の伝送に関する調査を実

施し，遠隔情報保障へ広く活用していくための検討を行っていく予定である。

参考文献

- [1] 加藤，他 10 名：講義資料とキーワードを画面合成した遠隔手話通訳システムー通訳スタジオにおける検討ー，ヒューマンインタフェース学会研究報告集 Vol.9, No.1, pp.23-28, 2007.
- [2] 河野，他 10 名：講義資料とキーワードを画面合成した遠隔手話通訳システムにおける聴覚障害学生への提示方法，ヒューマンインタフェース学会研究報告集，Vol.9, No.1, pp.29-32, 2007.
- [3] 加藤，他 9 名：聴覚障害者の情報保障におけるパソコン要約筆記入力者に対するキーワード提示，ヒューマンインタフェース学会論文誌，Vol.9, No.2, pp.125-134, 2007.
- [4] 西岡，他 9 名：遠隔地リアルタイム字幕提示システムにおける字幕作成者に対するキーワード提示について，WIT2005-88, pp.81-86, 2006.
- [5] 村上，他 8 名：就職面接場面での遠隔情報保障に関する一考察，WIT2006-57, pp.41-46, 2007.
- [6] 加藤，他 8 名：遠隔コミュニケーションのための 3 次元手話映像提示システムの検討，筑波技術大学テクノレポート，Vol.14, pp.13-16, 2007.
- [7] 若月，他 6 名：手話映像の臨場感向上のための立体映像撮影とその表示，筑波技術大学テクノレポート，15:7.11, 2008.
- [8] 沢井，石井，高橋，牧野：時分割ステレオビジョンにおける像の歪みの除去方法について，信学技報 IE, Vol.97, No.282, pp.7-12, 1997.
- [9] 田口，高橋，苗村：ネットワークカメラアレイを用いた実時間全焦点自由視点映像合成システム，信学技報 PRMU, Vol.107, No.539, pp.79-86, 2008.
- [10] 3D 立体放送「3D 立体革命」，BS11, <http://www.bs11.jp/>.
- [11] ステレオビジョンカメラ「bumble-bee」，Point Grey Research 社，<http://www.ptgrey.com/products/stereo.asp>
- [12] Thierry Oggier:(株)日本クラビス：SwissRanger SR3000：光到達時間計測を用いた奥行き計測カメラ，日本バーチャルリアリティ学会誌，Vol.12, No.3, pp.176-177, 2007.
- [13] 若月，河野：手話の 3 次元動画の撮影・表示方法に関する研究，筑波技術大学テクノレポート，14:81.87, 2007.

A Fundamental Study on Live 3-D Video Transmission for Remote Sign-Language Interpretation Services

Daisuke WAKATSUKI, Takeaki SHIONOME, Nobuko KATO, Sumihiro KAWANO,
Hiroshi MURAKAMI, Hiroki MINAGAWA, Tomoyuki NISHIOKA and Ichiro NAITO

Department of Industrial Information, Faculty of Industrial Technology,
Tsukuba University of Technology

Abstract: We have developed a remote communication support system to facilitate the social participation of hearing impaired people. In this paper, we describe a technique to transmit live 3-D video which consists of color images and depth images in our approach. Because one frame of the 3-D video applies an image that attaches subdivided depth images to a color image, the 3-D video can be transmitted by typical teleconference systems. We made a prototype live 3-D video transmission system on a trial basis and performed an experiment for live 3-D video transmission. We also introduced samples of remote lectures and exhibitions using our prototype system.

Key words: Live 3-D video, sign-language interpretation services, remote communication support, hearing impaired