

研究種目：若手研究（B）
 研究期間：2008～2009
 課題番号：20700458
 研究課題名（和文） 遠隔コミュニケーションのための手話の実写立体映像の評価に関する研究
 研究課題名（英文） A Study on Live Action 3D-Video for Remote Sign Language Communication
 研究代表者
 若月 大輔（WAKATSUKI DAISUKE）
 筑波技術大学・産業技術学部・助教
 研究者番号：50361887

研究成果の概要（和文）：本研究では手話による遠隔コミュニケーション環境を改善するために、手話の実写立体映像に関する基礎的研究を行った。カラー映像と距離映像を組み合わせて手話の実写立体映像をリアルタイム撮影、伝送、表示することができる方法を提案した。試作システムを構築して、通常の映像（2D映像）と実写立体映像（3D映像）の比較実験を行った結果、3D映像は2D映像よりも手話を正確に読み取ることができることが明らかとなった。また、指文字の読み取りやすさも向上する傾向がみられた。

研究成果の概要（英文）：We conducted a basic study on live action 3D video-imaging to improve remote-communication and legibility of sign language. We proposed a method that enables to take, transmit and display a real time live 3D video of sign language by using a color image and a depth image. Using an experimental system with the method, we performed comparison experiments between conventional 2D to live 3D video. The experimental results showed that live 3D video improves legibility of sign language and reduce errors in reading finger spelling.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2008年度	2,100,000	630,000	2,730,000
2009年度	500,000	150,000	650,000
年度			
年度			
年度			
総計	2,600,000	780,000	3,380,000

研究分野：総合領域

科研費の分科・細目：人間医工学・リハビリテーション科学・福祉工学

キーワード：健康・福祉工学，聴覚障害者支援

1. 研究開始当初の背景

聴覚障害者の円滑な社会参加や高等教育就学を支援するために、遠隔から手話や字幕による情報保障を提供することができる遠隔情報保障システムを構築し、遠隔コミュニ

ケーション技術に関する研究開発と実験的支援を進めてきた。本システムの概要を図1に示す。会場の音声と映像をインターネットで送り、手話通訳や字幕、およびコミュニケーションサービスを遠隔から提供すること

ができる。本研究では、より臨場感が高くわかりやすい遠隔情報保障を目指してシステムの改善に取り組んできた。

手話は手や指の型と動き、頭の動作や顔の表情などを用いて空間的に表現される言語であり、手の型や動作の奥行き、前後関係が重要になることが多い。しかし、従来の遠隔情報保障や放送等の手話は平面的に提示されるため、手話の奥行き感がなく読みにくかったり誤解を生んだりする可能性がある。また、遠隔手話通訳では本来相手に対して垂直方向に表現する動作を斜め方向の動作に替える方法もある。しかし、手話を表現する側も読み取る側もその表現方法に適応する必要があり、動作の方向の違いによって意味が変わる可能性もある。

2. 研究の目的

本研究では、遠隔手話通訳や遠隔コミュニケーション時に映像で表示される手話の読みやすさや臨場感を向上させるために、手話の実写立体映像を撮影、伝送できるシステムを構築し基礎的研究を行った。手話の実写立体映像のリアルタイム撮影、伝送方法として、カラー映像と、距離映像とを合成する方法を提案してシステムを試作した。試作システムで撮影した手話の映像について、裸眼立体ディスプレイで立体映像表示した場合と、通常の映像で表示した場合の比較実験を行った。手話および指文字で表現される単語について読み取りの正答率と主観評価から、手話の実写立体映像の特徴と有効性を明らかにした。

3. 研究の方法

(1) 実写立体映像の撮影と伝送



図 1 遠隔情報保障システムの概要

リアルタイムに実写立体映像を撮影する主な方法として、同期させた2台以上のカメラで視差映像を撮影する方法がある。しかし、伝送時に複数の視差映像を1フレームに縮小して合成し、表示時に元のサイズに引き伸ばすため解像度が低下して手話が読みにくくなる可能性がある。

そこで、本研究では、カラー映像と距離映像を用いて実写立体映像を伝送する方法を提案した。視差映像を伝送する方法と異なり、カラー映像と距離映像の解像度をそろえる必要がない。したがって、カラー映像を高解像度にして手話や表情を読み取りやすくすることが可能である。なお、カラー映像も距離映像もいずれも映像であるため、立体映像伝送用の特殊な機器を追加することなく従来の遠隔情報保障システム(図1)に導入することが可能である。図2のように距離映像を分割してカラー映像に合成して伝送する方法を提案した。

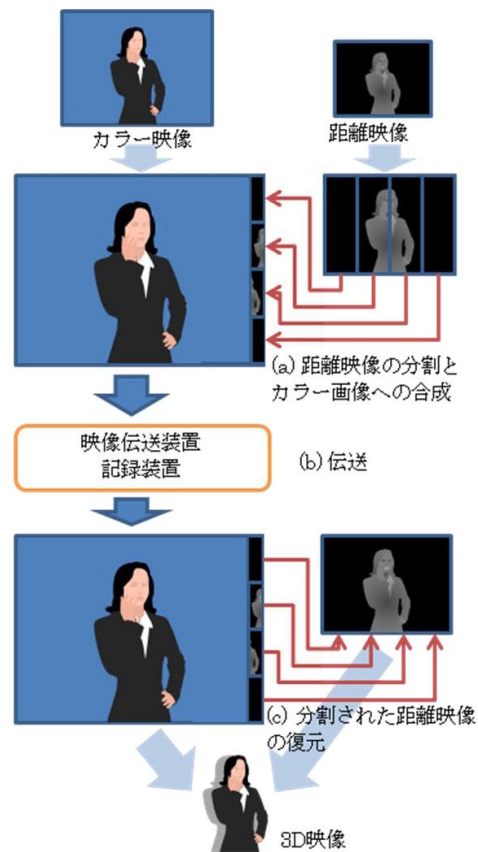


図 2 立体映像の伝送方法

カラー映像と距離映像から立体映像を生成する際は、距離映像の各画素をメッシュ化して2.5次元モデルを生成し、カラー映像をテクスチャマッピングする方法で実装した。生成したCGのモデルを、視差をつけてレンダリングすることで、様々な立体ディスプレイに対応することが可能である。

なお、本試作システムで表示される実写立体映像が自然な立体映像を提示することが

できる条件についても調査を行った。

(2) 手話の実写立体映像の有効性の調査実験

試作システムを用いて、実写立体映像で提示した手話の有効性を確認するために、実写立体映像(3D映像)と、通常の映像(2D映像)について、手話単語の模倣実験と、指文字の読み取り実験を実施した。

日常生活で主に手話を活用している聴覚障害者が、椅子に座って口話なしで表現する手話と指文字を正面から2Dと3Dで撮影したものを刺激映像に使用した。実際の遠隔情報保障では、テレビ会議システム等を用いてインターネット回線を通して手話通訳を行う。今回の実験ではネットワークによる伝送エラーや遅延の影響を除外するために、実写立体映像をレコーダ(XDCAM EXレコーダPMW-EX30, Sony社)に記録(1440×1080/29.97PのMPEG-2形式)し、動画編集ソフト(Adobe Premiere Pro CS4, Adobe社)を用いて編集、リサイズした映像を用いた。また、実際の遠隔情報保障では十分なネットワーク帯域が得られない場合に伝送する画像サイズを小さくすることが多い。標準の画像サイズ800×600と、縮小したサイズ256×192を作成した。

作成した映像を20インチの裸眼立体ディスプレイ(通常画面解像度1600×1200、立体表示時の画面解像度800×600)に表示した。被験者は聴覚障害者18名(18歳から21歳の大学生、男性10名、女性8名)とし、最適な立体映像が得られる距離(約1m)で座った状態で映像を観察させた。

映像の種類は画像サイズ800×600の2D、3D映像、256×192の2D、3D映像の4種類である。これらについて手話単語の模倣実験と、指文字の読み取り実験を実施した。

4. 研究成果

(1) 提案手法の実写立体映像の調査

提案手法で生成される立体映像では、距離映像の画像サイズが著しく小さい場合、正しい立体感が得られない可能性がある。そこで、自然な立体映像を観察できるカラー映像に対する奥行き映像の画像サイズ比 r を調査した。刺激映像はCG映像(動画)として、 r を0.0~1.0まで0.01ずつ変化させた場合に自然な立体映像に感じる値、および、 r を1.0~0.0まで変化させた場合に不自然に感じる値を調査した。聴覚障害者3名(20歳~21歳の大学生)に対して両調査を2回ずつ実施した結果、全体の平均は $r=0.213$ であった。試作システムの比は約0.22(カラー映像は800×600、距離映像は176×144)であり、ほぼ自然な実写立体映像を提示することが可能であった。

(2) 手話の模倣実験結果と考察

日本語-手話辞典から選び出した前後方向の動きを伴う手話単語を対象とし、画像サイズ800×600の2D、3D映像、256×192の2D、3D映像についてランダムかつ重複がないように20単語ずつ収録した刺激映像を用いた。被験者は各刺激映像の手話単語を模倣して、その正答率を調査した。

被験者全体の正答率の平均は95.1%、標準偏差は4.9%であった。実験条件ごとの正答率の平均値を表1に示す。〈〉内の数値は標準偏差を表す。2Dと3D映像間でt検定を行った結果、有意水準1%で画面サイズ800×600では有意差がなく、画面サイズ256×192、および両サイズの合計では有意差があった。

表1 手話単語模倣の正答率

	2D	3D	df	t
800×600	0.939(0.053)	0.969(0.039)	17	-1.826
256×192	0.928(0.052)	0.967(0.038)	17	-3.289**
Total	0.933(0.052)	0.968(0.038)	35	-3.430**

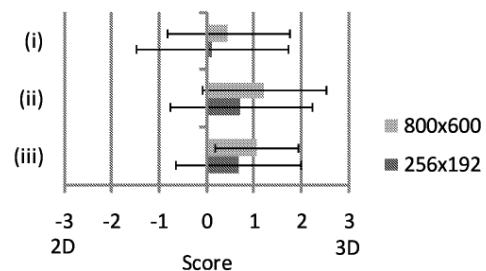
** $p < .01$

((): standard deviation, df: degree of freedom, t: t-value)

手話単語の模倣実験の正答率の平均については全体的に見ると2D映像よりも3D映像の方が高い正答率となった。しかし、画像サイズ別でみると800×600では有意な差は確認できず、256×192で3D映像の正答率が高かった。手話については解像度が低くなると3D映像が有効であることがわかった。

次に、画像サイズ別の2Dと3D映像について主観評価を実施した結果を図3に示す。

(i) 手話の読みやすさ
(ii) 手話の奥行き方向の把握しやすさ
(iii) 両手の前後関係の把握しやすさ
の3項目について両端を「2Dが最良(-3)」, 「3Dが最良(3)」, 中央を「どちらともいえない(0)」とした7段階で評価させた。棒グラフが平均値、エラーバーは標準偏差を表す。



((i): Legibility of sign language, (ii): Motion along depth direction, (iii): Depth relationship between both hands)

図3 手話単語模倣実験の主観評価

主観評価では、(i)~(iii)の3項目について画像サイズが小さい256×192で3D映像の評価が低くなった。解像度低下によって映像が不鮮明になったことが原因と考えられる。しかし、3D映像の正答率には直接影響がなく

不鮮明な映像から手話単語を予想して模倣できたものと考えられる。(i)の読みやすさについては、2Dと3Dでほとんど差がないが、正答率は3Dの方が高かった。2D映像では手話を認識できたと主観的に感じていても、誤解している可能性が高いことが示唆された。

(3) 指文字の読み取り実験結果と考察

歴代のノーベル物理学賞受賞者の名前のカタカナ表記(平均9.68文字)を対象とし、画像サイズ800×600の2D、3D映像、256×192の2D、3D映像についてランダムかつ重複がないように10単語ずつ収録した刺激映像を用いた。被験者は刺激映像の指文字を読み取り、その正答率を調査した。

被験者全体の正答率の平均は62.9%、標準偏差は20.7%であった。実験条件ごとの正答率の平均値を表2に示す。2Dと3D映像間でt検定を行った結果、有意水準5%で画面サイズ800×600、および256×192で有意差があった。両サイズの合計では有意水準1%で有意差があった。

表2 指文字読み取りの正答率

	2D	3D	df	t
800×600	0.588(0.197)	0.661(0.197)	17	-2.644*
256×192	0.605(0.236)	0.662(0.201)	17	-2.156*
Total	0.597(0.214)	0.661(0.196)	35	-3.442**

*p<.05, **p<.01

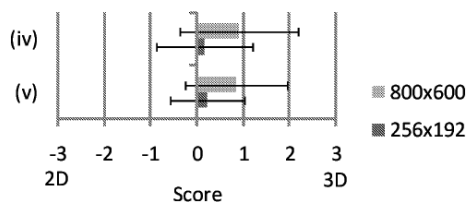
指文字の読み取り実験の正答率の平均を全体的に見ると手話単語の模倣実験と同様に3D映像の方が、正答率が高い結果となった。また、画像サイズごとに見ても3D映像のほうが、正答率が高いことが確かめられた。

次に、画像サイズ別の2Dと3D映像について主観評価を実施した結果を図4に示す。

(iv) 指文字の読みやすさ

(v) 指文字の奥行き方向の動きの把握しやすさ

の2項目について評価させた。



((iv): Legibility of finger spelling, (v): Motion along depth direction)

図4 指文字読み取り実験の主観評価

主観評価では、画像サイズ256×192では(iv)、(v)の両項目で2Dと3D映像の評価の差がほとんどなかった。指文字は主に指先のみで表現するため、読み取りには2Dと3Dの別よりも解像度の影響が強かったと予想できる。しかし、正答率では画像サイズの影響

は見られなかった。指文字はひらがなに対応した手指の型や動きが一意に決まっているため、映像が不鮮明で単語そのものを予想できなくても、ひらがな1文字ずつは類推できたためだと考えられる。

(4) まとめ

本研究では、遠隔情報保障や遠隔コミュニケーション時に映像で送受信される手話の臨場感や読みやすさを向上させるために、手話の実写立体映像に関する基礎的な研究を行った。

カラー映像と距離映像を合成する実写立体映像の撮影、伝送方法を提案した。提案手法を実装した試作システムを用いて手話の実写立体映像(3D映像)と通常映像(2D映像)の比較実験を行った。その結果、3D映像は2D映像よりも手話の正確な読み取りに効果があることが明らかになった。また、3D映像は指文字の読み取りも向上する傾向があることが示唆された。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計3件)

- ① 若月大輔, 加藤伸子, 塩野目剛亮, 河野純大, 村上裕史, 皆川洋喜, 西岡知之, 三好茂樹, 内藤一郎, 遠隔コミュニケーションのための手話の実写立体映像に関する研究, 電子情報通信学会技術報告. WIT, 福祉情報工学, 査読無, Vol. 109, No. 358, pp. 7-12, 2010.
- ② 若月大輔, 塩野目剛亮, 加藤伸子, 河野純大, 村上裕史, 皆川洋喜, 西岡知之, 内藤一郎, 遠隔手話通訳のための実写立体映像の伝送に関する検討, 筑波技術大学テクノレポート, 査読無, Vol. 16, pp. 73-78, 2009.
- ③ 若月大輔, 塩野目剛亮, 加藤伸子, 河野純大, 村上裕史, 皆川洋喜, 西岡知之, 内藤一郎, 立体映像を用いた遠隔手話通訳システムに関する研究, 電子情報通信学会技術報告. WIT, 福祉情報工学, 査読無, Vol. 108, No. 170, pp. 57-60, 2008.

[その他]

ホームページ等

- ① 科学新聞に取り組みの一部が掲載された
(<http://www.sci-news.co.jp/news/to pics/200903/210327.htm>)
- ② 第57回全国ろうあ者大会(in茨城)のバリアフリー展に出展
(<http://57jfd.web.fc2.com/>, 写真:
<http://57jfd.web.fc2.com/data/06pho>)

to_03bf_books_snap.html)

6. 研究組織

(1) 研究代表者

若月 大輔 (WAKATSUKI DAISUKE)
筑波技術大学・産業技術学部・助教
研究者番号：50361887