

手話映像の切断が読み取りに与える影響の評価

若月大輔¹⁾, 加藤伸子¹⁾, 村上裕史¹⁾, 皆川洋喜¹⁾, 西岡知之¹⁾,
河野純大¹⁾, 内藤一郎¹⁾, 三好茂樹²⁾, 石原保志²⁾

筑波技術大学 産業技術学部 産業情報学科¹⁾

筑波技術大学 障害者高等教育研究支援センター 障害者支援研究部²⁾

要旨: 手話による遠隔コミュニケーションでは、テレビ会議等の映像を通信する手段が用いられる。通信速度が十分ではない場合は、映像の遅延や切断が発生して手話の読み取りが困難になる可能性がある。本報告では、手話映像の切断時間と切断間隔が手話の読み取りに与える影響について、主観評価実験を行った結果について述べる。手話の読み取り、内容の理解、映像切断の許容、ストレスについて評価を行った結果、切断間隔を変更してもほとんど影響がないことがわかった。一方、切断時間については、時間が大きくなると手話を読み取りにくくなり、ストレスを感じるようになった。手話映像の切断と手話の読み取りの関係を明らかにすることによって、遠隔手話コミュニケーションに適した通信方法や環境の設計、およびシステム実装や運用に寄与できると考える。

キーワード: 手話, 手話映像, 切断, 映像通信, 遠隔情報保障

1. はじめに

聴覚障害者の高等教育就学や、社会参加、生活自立を支援するために、図1のように遠隔からネットワークを介して手話通訳や文字通訳を提供する遠隔情報保障システムの開発と、その関連研究を行ってきた[1]。特に、専門教育現場の情報保障の質の向上を目指して、授業の資料や専門用語等のキーワードを手話通訳映像に合成する方法[2-4]、通訳者に効率良く現場の状況を伝える方法[5-7]、および情報保障の臨場感を高める方法[8-10]について研究を重ねてきた。

遠隔情報保障システムにおいて、手話通訳を提供する際の映像の送受信には、主にテレビ会議システムを利用する。テレビ会議システムでは高画質な映像の送受信が可能であるが、安定かつ高速なネットワークが要求される。したがって、現場のネットワーク構成やセキュリティ関係の事情から十分な品質のネットワークを利用できない場合には、画質や解像度、フレームレートを調整し映像の質を下げる必要がある。また、遅延によってリアルタイムなコミュニケーションが困難となる場合もある。このような映像を送受信する際の映像の質の低下や遅延が、手話コミュ

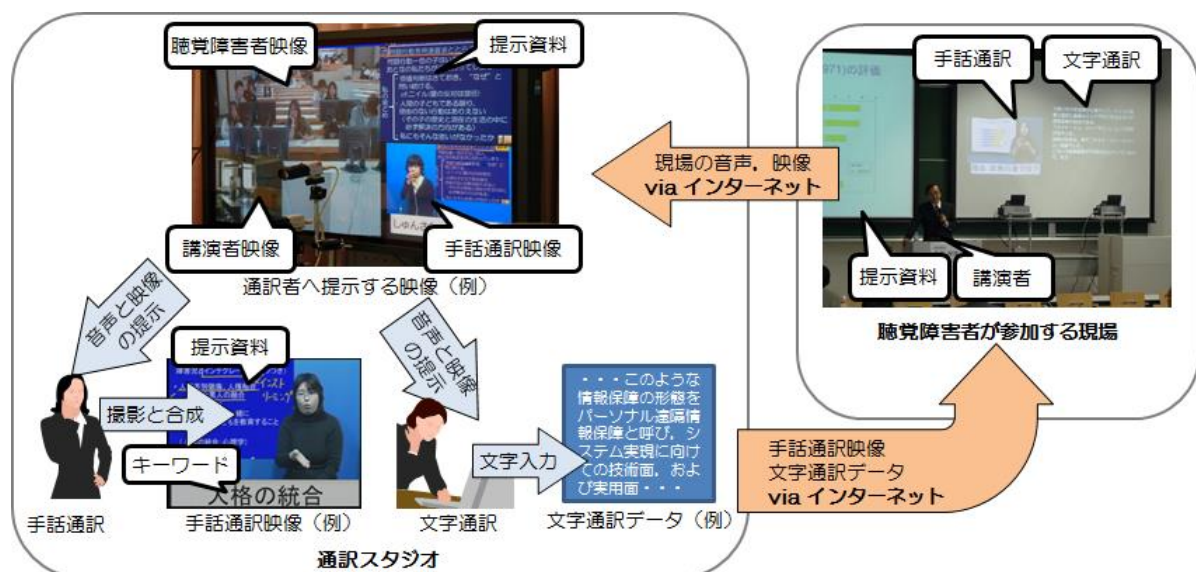


図1: 遠隔情報保障システムの概要

ニケーションに与える影響について研究が行われている。手話映像の画面サイズやフレームレート、解像度の影響を調査した研究[11, 12]や、映像の遅延を対象とした研究[13]が報告されている。

一方、ネットワークが不安定な状態では映像が一時的に切断することもある。これまでに手話映像が切断される状況について評価を行った例はなく、映像の切断が手話の読み取りに与える影響は明らかになっていない。本報告では、手話映像の切断時間と切断間隔について、手話の読み取りについての主観評価を行う。切断時間と切断間隔を変化させたときに、手話映像の読み取りがどのように変化するかを調べ、手話映像の切断の影響を明らかにする。

2. 関連研究

通信インフラが整備されて映像による遠隔手話コミュニケーションが可能になり、手話映像の品質や提示方法に関する研究が行われてきた。テレビ電話を利用した手話の遠隔通信を想定して、手話画像の品質評価を行った研究[11]では、フレームレートが十分に高ければ映像の画質や詳細度は手話の可読性にほとんど影響がないことが報告されている。携帯端末において手話映像を提示する際の画面サイズを評価した研究[12]では、3インチ以下の画面サイズでは手話映像の内容の理解度、見やすさの嗜好度が低下することが報告されている。

手話映像を送受信する場合、映像のエンコードやデコードの処理、および通信に時間を要するため遅延が発生する。手話映像通信を利用した手話対話における遅延の影響を評価した研究[13]では、遅延が対話進行を妨げるかの許容限、遅延を知覚し始める検知限について評価を行い、音声対話と比較して遅延に対して寛容な傾向があることが示されている。このように手話映像によるコミュニケーションについて、適した映像設定や提示方法、対話時の特性が明らかになってきている。

これらの研究では映像通信が切断されないことを前提として評価を行っているが、実際の環境では通信状態が不安定な場合もあり、映像が一時的に切断されることもある。特に無線WAN、LANなどの無線ネットワークは有線と比較して不安定であるため切断が発生する可能性が高い。本研究では手話映像が一時的に切断される状態を想定し、切断時間と切断間隔が手話の読み取りに与える影響について評価を行う。手話映像の切断と手話の読み取りの関係を明らかにすることによって、遠隔手話コミュニケーションに適した通信方法や環境の設計、およびシステム実装や運用に寄与することができると思う。

3. 切断された手話映像の読み取り実験

切断時間と切断間隔が異なる手話映像を被験者に提示して、手話の読みやすさ、内容の理解などに関する主観評価アンケートと、読み取れた内容を確認するための内容確認テスト（ニュース内容に関する4形式の設問4問）を実施した。本報告では内容確認テストの結果は保留し、主観評価アンケートの結果について議論する。ここでは、刺激映像の作成、提示方法、主観評価アンケートについて述べる。

3. 1 刺激映像

3. 1. 1 映像の内容

映像の内容は2012年9月22日～11月24日に放送されたNHK週間手話ニュースの映像を用いた。特定の1名の女性キャスターのみの映像を抜粋し、音声とテロップ、字幕を省いた手話のみの映像を9種類作成した。各映像のサイズは1280×720、フレームレートが29.97fpsであり、長さは1分程度である。手話表現の表出時間を調査したところ、手話の1単語あたりの表出時間が平均で約620ms/単語、1文あたりの表出時間が平均で約8740ms/文であった。以後、9種類の手話映像をV(v1～v9)と記述する。

3. 1. 2 切断時間と切断間隔

テレビ会議システムなどのリアルタイムに映像を通信する方法では、データが受信できなくなると映像が停止し、受信を再開したときに、そこからのデータをもとに映像を再生する。つまり、一時的な切断時に映像が暗転することなく停止され、再接続されるまでの間の映像は損なわれる。図2に示すように、映像が損なわれる切断が発生している時間を切断時間と呼び、切断が発生する間隔を切断間隔と呼ぶ。刺激映像に対して任意の切断時間と切断間隔を与えて再生することができる専用のソフトウェアを開発し、実験に使用した。

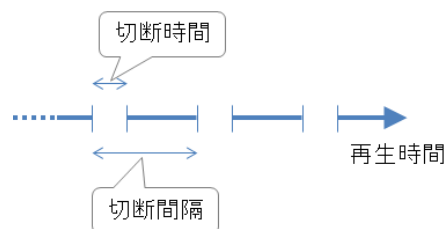


図2：切断時間と切断間隔

映像の1単語あたりの手話の表出時間を考えると、約620msの切断時間で手話の1単語程度が欠損する。実験では1.2単語が提示されなくなる可能性がある750msを中心として、500ms増減させた時間について評価を行うことに

した。つまり 250ms, 750ms, 1250ms の 3 条件の切断時間を評価対象とした。この切断時間の条件をそれぞれ D(d1,d2,d3)と記述する。

一方, 1 文あたりの手話の表出時間については, 約 8740ms の切断間隔で 1 文につき 1 回の切断が発生することになる。これを考慮して切断間隔については 9000ms を中心として, 切断時間と同様 4500ms 増減させた 4500ms, 9000ms, 13500ms の 3 条件の切断間隔を評価対象とした。この切断間隔の条件をそれぞれ I(i1,i2,i3)と記述する。

1 名の被験者に対して提示する切断の条件は, 切断時間 D の d1~d3 3 種類, 切断間隔 I の i1~i3 3 種類を重複なく組み合わせた 9 種類である。これらの 9 種類の切断条件それぞれについて, 手話映像 V の v1~v9 を対応させたものを刺激映像として提示した。

3. 2 実験環境と手順

3. 2. 1 実験の流れ

最初に被験者に対して実験の説明を行い, 5 分間の練習映像を見せる。練習映像は実験に用いる刺激映像以外の内容であり, キャスターは同じ映像である。次に, 切断条件 D と I と手話映像 V を組み合わせた約 1 分間の刺激映像を提示し, 内容確認テストと主観評価アンケートをそれぞれ 3 分間で回答してもらう。回答後 3 分間の休憩を与えて合計 10 分で 1 回の評価が終わる。切断条件 D と I を変更した刺激映像について繰り返し評価を行っていく。1 名の被験者に評価してもらう刺激映像は 9 種類である。つまり, 最初の実験説明と練習映像の提示も含めて, 実験全体に要する時間は約 100 分間であった。

3. 2. 2 被験者

被験者は聴覚障害を持つ 22~23 歳の大学生 9 名で, 性別は男性 6 名, 女性 3 名である。手話歴は 2 年から 16 年の幅があり平均は 7 年, 両耳の聴力の平均は約 109 dB であった。

3. 2. 3 映像の提示方法

刺激映像はノート型計算機 (Sony Vaio S (VPCS149FJ), ディスプレイ: 13.3 型ワイド, 1366×768 画素) に専用のソフトウェアで表示させた。被験者は約 600 ミリメートル離れた距離から映像を観察する。

3. 2. 4 刺激映像の提示順序

切断条件 D と I を組み合わせた 9 種類の d1i1~d3i3 と手話映像 V の v1~v9 が, 必ずすべての順番で現れるように, 1 名の被験者に提示する 9 種類の刺激映像の順番を次の条件を満たした上で無作為に決定した。

- 必ず 9 種類すべての切断条件を試してもらう
- 必ず 9 種類すべての手話映像を見もらう
- 切断条件 D と I の組は被験者間で同じ順番に現れない
- 手話映像 V は被験者間で同じ順番に現れない

これによって, 切断条件 D と I, 手話映像 V の順番の影響は避ける事ができると考える。

3. 2. 5 主観評価アンケート

主観評価アンケートでは次の 4 設問について評価してもらった。各設問は両端を「そう思う」「そう思わない」とし, 中央を「どちらとも言えない」とした 5 段階で回答してもらった。

- Q1. 手話を読める
- Q2. 内容を理解できる
- Q3. 映像の切断を許容できる
- Q4. ストレスがたまる

これらの設問の他にも被験者が感じたことを自由に記述してもらった。なお, 実験説明のときに「内容や手話表現そのものの難易度ではなく映像の切断とその間隔によってどう感じたか」について回答するように説明した。

4. 実験の結果と考察

主観評価アンケートに回答してもらった各設問について, 切断時間 D と切断間隔 I について平均と標準偏差を示し, D と I の 2 要因について分散分析で検定を行った結果について述べる。なお分析には統計解析ソフトウェアの R[14]を用いた。

4. 1 結果

被験者に回答してもらったアンケートの各設問の結果を図 3~6 に示す。各図の(a)が切断時間 D について, (b)が切断間隔 I についての結果である。なお, 棒グラフが評価の平均値, エラーバーが標準偏差を表す。

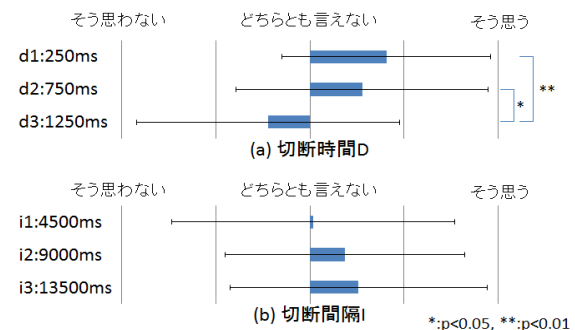


図 3 : Q1 手話を読める (Mean, SD)

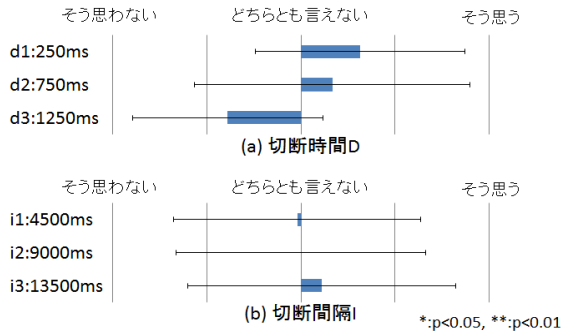


図4：Q2内容を理解できる (Mean, SD)

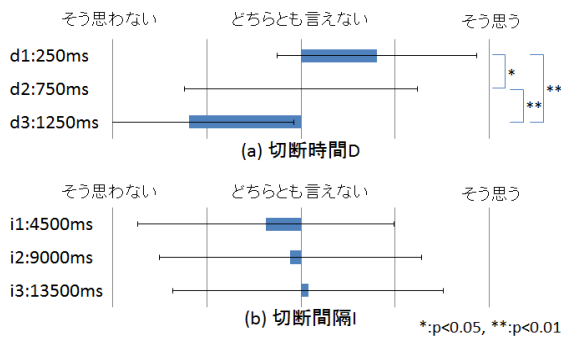


図5：Q3映像の切断を許容できる (Mean, SD)

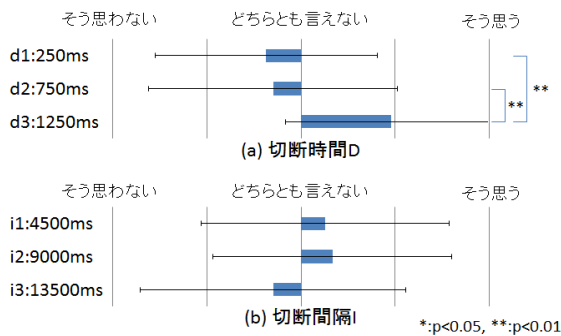


図6：Q4ストレスがたまる (Mean, SD)

切断時間Dと切断間隔Iの2要因について、対応ありの二元配置分散分析を行った結果を表1～4に示す。Q1, Q3, Q4については切断時間Dの要因の効果が有意であったため、ボンフェローニの方法で多重比較を行った。有意な差がある群間について、結果を図3, 図5, 図6に記した。

Q2については交互作用の効果が有意であったため、水準別の各要因の効果を吟味する。各水準別の平均を折れ線グラフで表した結果を図7に示す。切断時間がd1:250msとd3:1250msの場合は切断間隔が大きくなるに従って内容を理解できる傾向があるが、d2:750msについては逆の傾向を示した。したがって、Q2については切断時間Dの効果は有意であるが、この効果は図7で示した有意な交互作用によって限定されている。

表1：Q1手話を読める (ANOVA)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
D	2	23.88	11.94	7.50	0.0050 **
I	2	3.28	1.64	1.79	0.2000
D:I	4	3.46	0.86	0.95	0.4500
Residuals	32	29.21	0.91		

*:p<0.05, **:p<0.01

表2：Q2内容を理解できる (ANOVA)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
D	2	29.73	14.86	6.40	0.0091 **
I	2	1.06	0.53	0.81	0.4630
D:I	4	10.79	2.70	2.79	0.0431 *
Residuals	32	30.99	0.97		

*:p<0.05, **:p<0.01

表3：Q3映像の切断を許容できる (ANOVA)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
D	2	53.03	26.52	15.23	0.0002 **
I	2	2.55	1.27	0.81	0.4620
D:I	4	2.71	0.68	0.81	0.5280
Residuals	32	25.93	0.84		

*:p<0.05, **:p<0.01

表4：Q4ストレスがたまる (ANOVA)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
D	2	30.32	15.16	6.80	0.0073 **
I	2	6.40	3.20	2.70	0.0975
D:I	4	3.82	0.96	1.27	0.3030
Residuals	32	29.21	0.91		

*:p<0.05, **:p<0.01

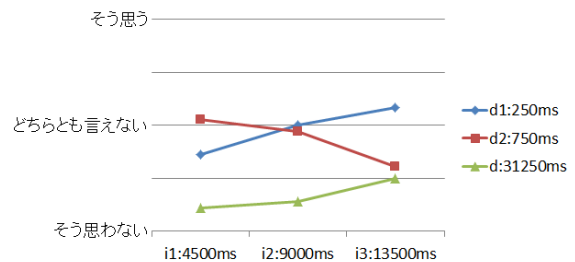


図7：Q2内容を理解できる (群別の平均値)

4. 2 考察

4. 2. 1 手話の読み取り

切断間隔を変化させた場合については、手話の読み取りに大きな困難を感じていないことがわかった (図3(b))。一般的には切断間隔が短いほど切断回数が増加し、手話が

読みにくくなると考えられる。しかし、切断間隔を変化させても評価に差はみられなかった(表1)。つまり、切断間隔が一番短い4500msで想定される1文に2回程度の切断が発生しても、手話の読み取りに影響がないことがわかった。

切断時間の変化については手話の読み取りに差が見られた。250msと750msの切断時間に対して1250msの評価が低い結果となった(図3(a))。手話映像の1単語に要する表出時間は約620msであるため、1単語を欠損する程度では可読性に影響がなく、1単語より多く欠損される場合は読みにくく感じる事がわかった。

4. 2. 2 内容の理解

手話映像の内容の理解については、切断時間に対して有意な差があるが、切断時間と切断間隔の交互作用の効果があるため、切断時間や切断間隔について差を論じることはできない。

切断時間が250msと1250msの場合は切断間隔が大きくなると内容を理解できると感じる傾向がある(図7)。また、どの切断間隔においても1250msよりも250msのほうがより評価が高い。しかし、切断時間が750msの場合は切断間隔が短い4500msで一番評価が高く、より評価が上がると思われる9000msと13500msで評価が下がっていく逆の傾向が見られた。今後は被験者の手話歴や聴力などの特性、手話映像内の手話表現やトピックとの相性、および本報告で保留した内容確認テストの結果を考慮して解析を行うことによって原因を明らかにできるのではないかと考える。

4. 2. 3 手話映像切断の許容

切断間隔の変化について映像の切断の許容に対する評価は、明確な判断結果にはなっていない(図5(b))。また、手話の読み取りと同様に切断間隔について評価の差がないことがわかった(表3)。

切断時間の変化については250ms, 750ms, 1250msの間で相互に評価の差がみられた。つまり、切断時間が短くなるほど切断を許容できると感じる事が明らかになった(図5(a))。切断時間が250msと750msの間に映像切断の許容について差があるが、手話の読み取りについては差がない。これは1つの単語が欠損することは許容しにくい、手話表現の前後の文脈を利用して読み取ることができたためであると推察する。自由記述でも「映像が切断していても、切断していた部分の手話を予想できる。」という旨の意見がよせられた。

4. 2. 4 切断ともなうストレス

切断間隔の変化に対するストレスの評価については、手話の読み取りや映像切断の許容と同様に差がなかった(表4)。また、評価の平均から明確にストレスを感じていないことがわかった(図6(b))。

切断時間については250msと750msに対して、1250msのほうがよりストレスを感じており(図6(a))、手話映像の手話のうち1単語より多く欠損される場合にストレスをより感じていることがわかった。また、手話の読み取りとほぼ同様な結果であることから、手話が読み取れないことがストレスの一部になっていると予想できる。

5. まとめ

本報告では、手話映像でのコミュニケーションにおいて映像が切断される状況を想定して、手話映像の切断時間と切断間隔が手話の読み取りに与える影響について主観評価を行った結果について述べた。

手話の読み取り、内容の理解、映像切断の許容、ストレスについて評価を行った結果、切断間隔を変更してもほとんど影響がないことがわかった。一方、切断時間の変更については、手話の読み取りやストレスに影響を与えることがわかった。手話映像中に表現される手話単語が切断により1単語より多く欠損されると、読み取りにくく感じる事が明らかになった。

本報告では、内容の理解の主観評価について、切断時間と切断間隔の影響を明らかにすることができなかった。今後は、被験者の特性、手話映像の内容との相性、および本報告で保留した内容確認テストの結果を考慮することによって解析を行いたい。

謝辞

本研究はJSPS 科研費 24700885 の助成を受けたものです。

参考文献

- [1] 若月大輔. 聴覚障害者のための遠隔情報保障システムとその応用. 日本生活支援工学会誌. 2011;11(1):p.8-13.
- [2] 加藤伸子, 河野純大, 若月大輔, 西岡知之, 他. 講義保障における聴覚障害学生と手話通訳者へのキーワード提示の効果に関する検討. ヒューマンインタフェースシンポジウム2007 論文集. 2007; : p.451-456.
- [3] 加藤伸子, 河野純大, 村上裕史, 白澤麻弓, 他. 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—通訳スタジオにおける検討—. ヒューマンインタフェース学会研究報告集. 2007; 9(1) : p.23-28.
- [4] 河野純大, 加藤伸子, 村上裕史, 白澤麻弓, 他. 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. 2007 ; 9(1) : p.29-32.
- [5] 加藤伸子, 河野純大, 若月大輔, 塩野目剛亮, 他. 講義の情報保障におけるキーワード提示タイミングに関する基礎的検討. 電子情報通信学会技術研究報告 (WIT). 2008 ; 108(170) : p.51-56.
- [6] 三好茂樹, 河野純大, 西岡知之, 加藤伸子, 他. 遠隔講義保障におけるリアルタイム字幕作成者を支援するための映像提示手法に関する基礎的研究. 電子情報通信学会論文誌 D, 2008 ; 91(9) : p.2236-2246.
- [7] 加藤伸子, 河野純大, 三好茂樹, 西岡知之, 他. 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, 2007 ; 9(2) : p.125-134.
- [8] 若月大輔, 加藤伸子, 塩野目剛亮, 河野純大, 他. 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研究報告 (WIT), 2010 ; 109(358) : p.7-12.
- [9] 若月大輔, 加藤伸子, 河野純大, 村上裕史, 他. 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, 2008 ; 15 : p.7-10.
- [10] 若月大輔, 塩野目剛亮, 加藤伸子, 河野純大, 他. 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研究報告 (WIT). 2008 ; 109(170) : p.57-60.
- [11] 中園薫, 米原裕貴, 長嶋祐二, 市川熹. 手話動画像の評価実験—画像サイズ, フレームレート, 量子化幅等の影響—. 電子情報通信学会技術研究報告 (WIT). 2005 ; 105(67) : p.19-24.
- [12] 塩野目剛亮, 鎌田一雄, 山本英雄. 手話の知覚と画面大きさの関係に関する一検討. 電子情報通信学会技術研究報告 (HCS). 2004 ; 103(742) : p.29-34.
- [13] 長嶋祐二, 住田英之, 中園薫. 手話通信における遅延の影響と遅延検知限の評価. 電子情報通信学会技術研究報告 (WIT). 2006 ; 105(508) : p.61-66.
- [14] The R Project for Statistical Computing (cited 2013-8-23), <http://www.r-project.org/>.

Subjective Evaluations of the Readability of a Sign Language Video with Interruptions

WAKATSUKI Daisuke¹⁾, KATO Nobuko¹⁾, MURAKAMI Hiroshi¹⁾, MINAGAWA Hiroki¹⁾, NISHIOKA Tomoyuki¹⁾,
KAWANO Sumihiro¹⁾, NAITO Ichiro¹⁾, MIYOSHI Shigeki²⁾, ISHIHARA Yasushi²⁾

¹⁾Department of Industrial Information, Faculty of Industrial Technology, Tsukuba University of Technology

²⁾Division of Research on Support for the Hearing and Visually Impaired,
Research and Support Center on Higher Education for the Hearing and Visually Impaired,
Tsukuba University of Technology

Abstract: Sign language video communication is often provided by a television conference system. A communication network with a low transmission rate causes interruptions in the videos and makes it more difficult to read the sign language. In this paper, we describe the effects of interruption time and interval in a sign language video. Interruption interval had no effect on subjective evaluations of the readability of the sign language, the level of understanding of the video content, tolerance for the video interruptions, and stress caused by the video interruptions. On the other hand, a sign language video with a long interruption time was difficult to read and stressed the users. Our aim was to contribute to the development of a communication system suitable for remote sign language communication by determining the degree of readability of a sign language video with interruptions.

Keywords: Sign language, Sign language video, Interruption of video, Video communication, Remote communication support