



Title	衛星通信における手話の動画像伝送
Author(s)	青木, 由直; Aoki, Yoshinao; 棚橋, 真 他
Citation	北海道大學工學部研究報告, 167, 77-85
Issue Date	1994-01-14
Doc URL	<a href="https://hdl.handle.net/2115/42400">https://hdl.handle.net/2115/42400</a>
Type	departmental bulletin paper
File Information	167_77-86.pdf



## 衛星通信における手話の動画像伝送

青木 由直 棚橋 真  
徐 軍 村林 知明

(平成5年8月31日受理)

### Transmission of Animation of Sign Language by a Communication Satellite

Yoshinao AOKI, Shin TANAHASHI, Jun Xu and Tomoaki MURABAYASHI

(Received August 31, 1993)

#### Abstract

We propose a communication method to transmit sign language image by a communication satellite, where animations corresponding to the sentences of sign language are transmitted. Since a transmission line of a large capacity is necessary for animation transmission of sign language images, we adopt a method of an intelligent communication transmission to reduce capacity of image data, where a computer graphics (CG) system reconstruct sign language image after receiving textual data or parameters describing the hands and fingers motions. We propose a method to draw sign language images according to the textual data or parameters and develop a CG system to construct animation of sign language images by combining the CG-generated sign language images. An experiment using a communication satellite was done to verify the effectiveness of the developed system for sending animation of sign language images, resulting in showing the potentiality of the developed system.

#### 1. はじめに

手話は、聴覚障害者にとって、非常に重要なコミュニケーション手段である。しかし、手話で通信を行うためには、動画像伝送が必要であり、大容量の伝送路が必要となるため、手話に対応しうる通信手段は少ない。回線容量がそれほど大きく取れないような場合、テレビ会議などの目的には十分な画質の画像伝送が可能な場合でも、手話画像においては、細かい手の動きなどが読み取れず、認識が困難であるため、不十分な場合がある。そこで、本研究では、知的画像伝送と呼ばれる伝送方式を用い、手話画像の伝送を試みた。

手話の知的通信では、手話画像を一度、意味情報の形に変換し、受信側で画像再構築を行う。そのため、きわめて小さな回線容量での画像伝送が可能であり、また、音声言語・意味情報間の変換を行うことで、健聴者・聴覚障害者間の通信(通訳)を実現することが可能である。

知的通信を行うためには、意味情報の抽出・意味情報の伝送・画像の再構築の3段階が必要と

なる。音声言語と手話の間の自動翻訳については、さまざまな研究が行われており<sup>1)</sup>、その中で、手話画像の生成については、モデルを用い、コンピュータ・グラフィックスによって表現する方法と、あらかじめ光ディスクなどの媒体に手話画像を記録しておき、表現したい内容に応じてその場面を検索する方法が報告されている<sup>2)</sup>。しかし、コンピュータ・グラフィックスを用いた方法では、表情の表現が困難であり、一方、記憶媒体にあらかじめ手話画像を記録しておく方式では、検索に時間がかかること、単語と単語がうまくつながらない、などの問題点があった。しかし、近年、コンピュータ・グラフィックスによる顔画像の表現について、多くの研究が行われ、表情の表現が実現されつつある。そこで、本研究では、表情を含めたコンピュータ・グラフィックスによる手話画像生成システムを作成した。また、このシステムを用いた、衛星通信による画像伝送実験について報告する。

## 2. 手話の要素および記述モデル

手話は、音声言語との関連から、一般的に3種類に分類され<sup>3)</sup>、それぞれ表1に示される特徴を持っている。

1. 伝統的手話	語順が音声言語と異なる 助詞を使わない 口話を併用しない
2. 同時法的手話	音声言語を手指で表現したもの (語順は音声言語と一致する) 助詞を使用する(指文字)
3. 中間型手話	前二者の中間的な手話 口話との併用が多い 語順は日本語とほぼ一致

表1 手話の種類

いずれの手話も、文法(語順)などは異なるが、手話単語は共通である。本研究では、音声言語からの変換が容易なことから、同時法的手話を対象にしているが、語順や口話への対応を行うことにより、中間型手話ならびに伝統的手話にも応用することが可能である。

手話は、視覚言語と呼ばれるように、視覚によって情報を伝える。これらは、身振りと、表情・口の形などによって表現される。以下に、手話表現に用いられる要素を示す(表2)。

手話の要素：手や指の位置(絶対位置・相対位置) 手や指の形(指文字・単語) 手や指の動き(方向→対象・速さ→意味の強弱) 表情(感情・疑問) 口の形・動き
---

表2 手話の構成要素

基本的に、手話は、手や指の形や動きで意味を表現する。なお、ここでは、一つの意味を表現する手話を、手話単語と呼ぶ。

手話単語には、形と位置で意味を表現する場合と、動作を伴う場合がある。前者は、手の形と、その指す方向・位置、場合によっては左右の手の位置関係によって意味が定まる。例えば、「わたし」という手話単語は、自分を指さすが、「あなた」を表現する場合は、手の形は同じだが、向きを相手の方に向ける。一方、後者の場合、手の形と動作およびその方向によって意味が定まる。動作を伴う場合、「歩く」や「泳ぐ」などのように、実際の動作を疑似的に表現する場合と、上下関係を表現する場合（例えば、「兄」は中指を立てて、手を上にあげるが、「弟」の場合は、逆に下げる）、時間を表現する（一般に、手話では、自分の前方への移動を示すことで、未来の意味を現し、逆に、自分の背後を示すことで、過去を表現する）場合などがある。

助詞や、あるいは固有名詞などの表現には、指文字が使われることがある。指文字は、指の形で五十音を表現するものである。

また、特に感情をこめる場合、ゆっくり表現したり、表情を大きくして意味を強調する。

表情は、このような感情表現のほか、疑問を表す場合に用いられる。また、速度変化について、「美しい」などの形容詞は、ゆっくりと動作を行うことによって強調するが、「速い」「遅い」という単語のように、手話単語自身の速度が早い（あるいは遅い）場合もある。

口の動きは、特に指文字の表現などにおいて、おもに手話を補う為に用いられる。

手話画像の伝送を行うためには、これらの情報を相手側に伝える必要がある。これまで、手話の表現を記述するためのパラメータとしては、「手の動作位置、手の形、手の動き」<sup>4)</sup>、またはさらに「掌の向き」を加えたモデル<sup>5)</sup>が提案されているが、表情などへの対応が不十分であった。そこで、本論文では、それにさらに表情および口の形を含めたモデルを提案する（表3）。

手の動作位置 手の形	各関節の角度で記述
手の動き	角度の変化
表情	AU

表3 手話の記述モデル

プログラム内では、「手の動作位置、手の形、手の動き」については関節角、およびその変化で記述し、また、表情については、アクション・ユニットと呼ばれる要素に分解して扱っている。

これは、心理学の分野で研究されている表記法 FACS (Facial Action Coding System) に用いられているもので、人間の顔の動きを44の AU (Action Unit) と呼ばれる基本的な動きに分解する（表4）。あらゆる表情は、これらの AU の組合せで表現できるとされている。

本研究では、顔の表情を、これらのパラメータの度合いによって記述する。

AU No.	AU 名	AU No.	AU 名
1	眉の内側を上げる	20	唇両端を横に引く
2	眉の外側を上げる	23	唇を強く閉じる
4	眉を下げる	24	唇を押さえつける
5	上瞼を上げる	25	顎を下げずに唇を開く
6	頬を持ち上げる	26	顎を下げて唇を開く
7	瞼を緊張させる	27	口を大きく開く
8	唇を互いに接近させる	28	唇を吸い込む
9	鼻にしわをよせる	29	下顎を突き出す
10	上唇を上げる	30	顎を左右にずらす
11	鼻唇溝を深める	32	唇を噛む
12	唇両端を引き上げる	35	頬を吸い込む
13	唇を鋭く引き上げる (頬を膨らませる)	41	上瞼を(力なく)下げる
14	えくぼを作る	42	薄目
15	唇両端を下げる	43	閉眼
16	下唇を下げる	44	細目にする
17	下顎(おとがい)を上げる	45	まばたく
18	唇をすぼめる	46	ウインクする

表4 AU のリスト (一部)

### 3. 知的画像伝送

知的通信方式では、まず、原画像の情報を、共有知識および意味情報に分解する。共有知識は、通信を通じて変化しない情報であり、あらかじめ、双方に与えておく。変化する部分が意味情報であり、これを伝送することによって通信を行う(図1)。

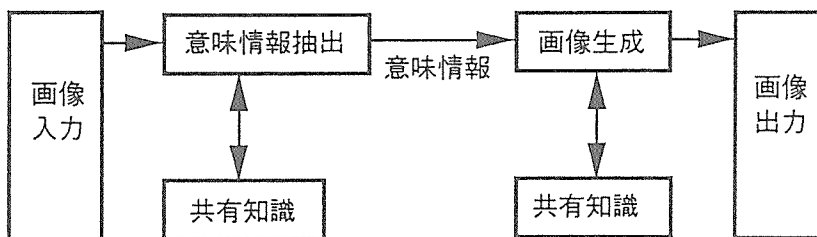


図1 知的画像伝送

顔画像の知的伝送の場合、共有知識は、背景、顔の形状モデル、顔の濃淡(色)データになる。それに対して、意味情報は、AUの変化を示すパラメータである。受信側では、AUの変化に応じて、顔の形状を変形させ、これに顔の原画像をマッピングすることにより、顔画像を再構成する。

手話の場合は、顔画像のほかに手や腕の画像を伝送する。したがって、共有知識としては、骨格構造および表面の濃淡情報が含まれる。意味情報としては、関節角または手話単語を用いる。受信側では、関節角をもとに手や腕の位置と形を計算し、表面の濃淡情報にしたがって画像を生成する(図2)。なお、手話の速度変化に関しては、時間情報を別に送ることで処理するものとする。

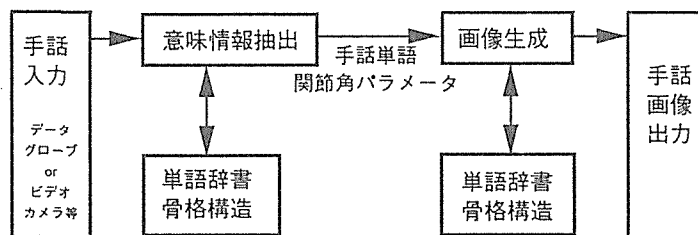


図2 知的手話画像伝送

#### 4. 手話の内部表現・手話単語辞書

手話画像の生成は、基本的には関節角およびAUの各パラメータによって行われる。しかし、手話による会話は、その大部分が手話単語によって構成される。したがって、これらの手話単語を辞書として用いることで、関節角情報すべてを送る必要はなくなる。そこで、手話について、手話単語を識別できる場合とできない場合に分類し、識別できる場合は手話単語を示すテキストを用い、識別できない場合にのみ関節角情報を伝送するものとする。また、表情についても、感情表現のための辞書を用い、疑問あるいは驚きといった感情表現の情報を伝送に用いることにする（表5）。

手話単語	手話単語を識別できる場合 → 手話単語を伝送 手話単語を識別できない場合 → 手や指の位置・形・動きを 関節角データとして伝送
表情・口の動き	表情 → 喜怒哀楽、疑問などの感情情報 口の形・動き → 特徴点の位置・動き

表5 意味情報の分類

#### 5. 手話単語辞書

手話単語を伝送する場合、受信側では、手話単語辞書により、手話単語から関節角データへの変換を行う。手話単語は、意味を現すテキストで表現し、一つの意味を表す手話（手話単語）の関節角を離散的に記録している<sup>6)</sup>。なお、手話の表現に関しては、相当量の手話が、手話動作の始点と終点が意味を伝達する上で重要な要素を有するという結果が得られている<sup>7)</sup>。したがって、単語辞書には、手話単語の始点と終点、ならびに、その中間点を数点記録している（ただし、手の位置及び形状だけで表現する手話単語については、その位置を表す関節角のみ記述する）。また、辞書には、手話単語の表現時間に関する情報も含まれる。

表情については、疑問や喜び・驚きなど、基本的な感情について、AUの情報辞書に格納されている。一般に、会話を行う場合、表情（感情）は、言葉そのものに依存するものではなく、分脈によって生じることが多い。したがって、手話単語辞書とは独立したものである。しかし、分脈を解析し、感情の推移を推定して表情を生成するのは現段階ではまだ困難であるので、ここでは、言葉そのものによって規定される表情についてのみ対応する。例をあげると、疑問文などの

ように、疑問を表す表情をする事が明かである場合などである。

## 6. 手話画像生成・動画像処理

姿勢や動きは、意味情報から与えられる関節角によって規定される。

手話画像の生成（腕および手の部分）には、多関節構造体の形状モデルを用いている。まず、得られた関節角をもとに位置を決定し、次に表面の形状を決定する。レンダリングにはZバッファアルゴリズムを使用している。骨格の構造を図3に、腕の構造を図4に示す。骨格は、背骨をグローバル座標系の基準とし、腕や肘など、それぞれについて、ローカル座標系を定めて計算を行う。なお、座標の向きは、関節から次の関節へ方向をz軸とし、そのときの上方向をy軸、右方向をx軸としている。データ構造は、図5のようになっている。先頭が名前で、順に、長さ、オフセット、角度が記述されている。オフセット、角度はともに、ローカル座標を基準にしている。角度は、それぞれ、x軸、y軸、z軸まわりの回転角である。中カッコの中に括弧のある名前は、連結する部位の名前を表す。

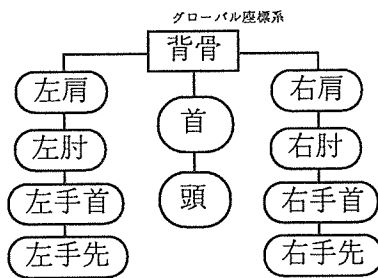


図3 骨格構造

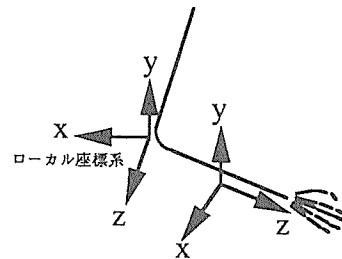


図4 腕の構造

```
tekubir 100    0 0 -100    0 0 0
{
    trko0
    trku0
    trna0
    trhi0
    troy0
}
```

図5 データ構造

生成した手の部分の画像を図6に示す。図6 aは、ワイヤフレームで表現したもので、図6 bはレンダリングを行ったものである。

次に、顔画像の生成について述べる。顔画像を生成するためには、顔の表面形状モデルおよび顔画像を用いる。まず、顔の表面形状モデルを補正して、顔の画像にあわせ、合成する。合成する際、立体形状と、顔の各構成要素を合わせるため、現システムでは、いくつかの特徴点を指定するようにしている。立体形状の補正・顔画像の合成を行った後、指定されたAUの値によって、画像を変形させ、表情を生成し、最後に他の部分と合成する。また、顔の表面構造に顔画像を張りつけているので、ある程度の顔の動きを表現することができる。例を図7に示す。

手話を動画像として表示するためには、毎秒30枚程度のアニメーションにする必要がある。与

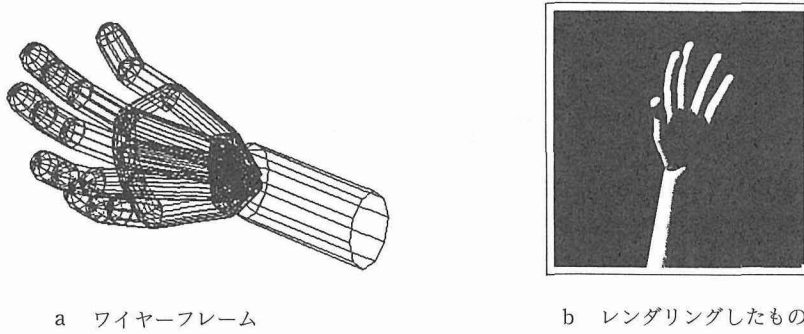


図6 手話画像（手の部分の拡大）

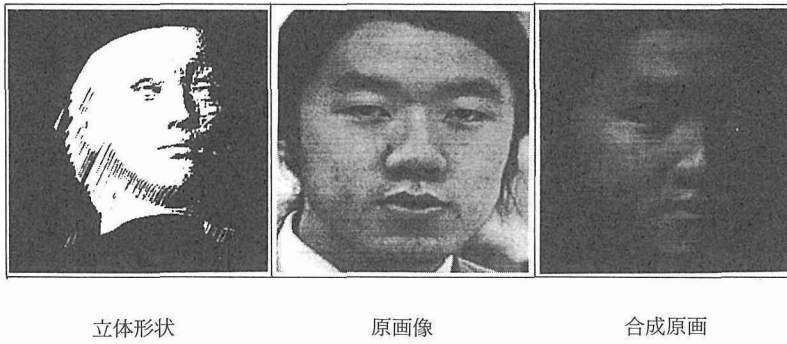


図7 顔画像の合成例

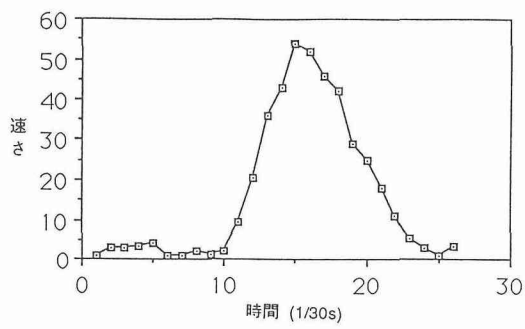


図8 腕の移動速度

えられる関節角は、離散的なので、時間情報をもとに補間を行い、各時間のフレームのデータを作成する。意味情報から与えられる時間情報には、始点から終点までの時間が記述されている。また、手の加速度を考慮し、動作の始点および終点ではゆっくりと、中点付近で最大の速度になるように加速・減速制御を行う。実際の動きは、図8のようになるので、ここでは、速度変化を直線的に近似している。

## 7. 衛星通信による画像伝送

衛星通信を使用して、画像の折り返し伝送実験を行った。そのときの実験システムの構成を図9に示す。送信側では、意味情報を表すテキストを直接キーボードから入力する。入力されたテキストデータには、表情の情報が付加され、伝送用のパケットが生成される。パケットの構造を図10にしめす。伝送すべき内容は、手話単語、または関節角と時間情報、それに表情を表すテキストデータである。伝送プロトコルは、HDLCに準拠している。フラグ、アドレスおよびコントロールフィールド、フレームチェックシーケンスは、制御用のデータである。伝送する意味情報はインフォメーションフィールドの部分に記述される。ここに含まれる情報は、手の形（姿勢）に関する情報と、表情を表す情報である。手の形に関する情報には、先頭に、手話単語か否かを判別するためのフラグが付加され、受信側では、その内容に応じて処理が行なわれる。

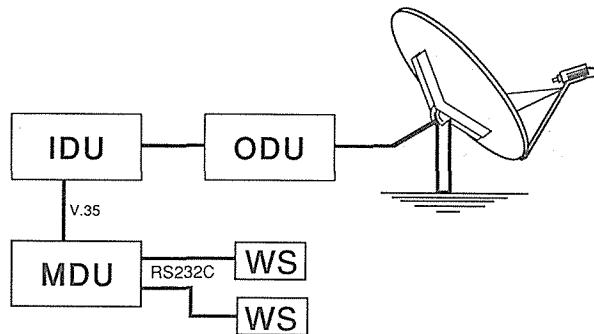


図9 衛星通信実験

### 1 フレームの構成

フラグ	アドレス フィールド	コントロール フィールド	インフォメーション フィールド	フレームチェック シーケンス	フラグ
-----	---------------	-----------------	--------------------	-------------------	-----

### インフォメーションフィールドの構成

手話単語 or 関節角+時間情報	表情データ
---------------------	-------

図10 パケットの構成

生成されたデータは、RS232Cの回線を通してMDUへ入力され、IDUのV.35シリアルポートへ伝えられる。IDUは、このデータを符号化し、中間周波数としてODUへと送りだす。ODUでは、周波数変換を行い、パラボラアンテナより送信が行われる。受信の際は、この逆の経路をたどり、ワークステーションで手話画像を生成する。

## 8. ま と め

知的通信の技術を用いて、表情を含む手話の動画像伝送システムを作成した。手話画像に表情をつけることにより、理解しやすい手話画像を生成できた。

衛星通信により、動画像の伝送実験を行った。その結果、64kbps という比較的遅い回線でも、なめらかな動画像を伝送することができることが示された。

### 参 考 文 献

- 1) 原島博：顔の表情の人工現実感, JSPE-57-08 '91-08-1335
- 2) 川崎, 他：ビデオディスクを利用した手話生成システムの開発, 情報処理学会, 第29回全国大会, 4L-10, (昭和59年後期)
- 3) 中野善達：手話の考察, 福村出版
- 4) W. C. Stokoe, etc: A Dictionary of American Sign Language on Linguistic Principles, Callaudet College Press (1965).
- 5) L. A. Friedman Ed.: On the Other Hand-New Perspective on American Sign Language usage, Academic Press (1977).
- 6) 徐, 青木, 鄭：手話画像知的通信のための単語辞書の構成法, 電気関係学会北海道支部連合大会 H3. 345
- 7) 橋本, 世木, 小川：連続画像生成による手話情報伝送パラメータに関する基礎的研究, 電子情報通信学会創立70周年記念総合全国大会, 2643 (1987)