

平成 14 年度

学士学位論文

手話単語映像データベースからの腕の動き抽出

Extraction of the movement of the arm
from the sign language word image database

1 0 3 0 2 8 2 田畑 博紀

指導教員 岡田 守

2003 年 2 月 12 日

高知工科大学 情報システム工学科

要旨

手話単語映像データベースからの腕の動き抽出

本稿では、動画計測ソフトウェアを用いた手話の腕の動き抽出を行う。独立行政法人通信総合研究所のホームページより公開されている手話の動画を使用し、腕・手の関節を対象にして、マーカー付けする。そして、マーカー追跡から位置座標を求める。計測結果から手話単語の動作をデータ化する。

キーワード 動画計測ソフトウェア, 抽出, マーカー追跡

Abstract

Extraction of the movement of the arm from the sign language word image database

In this paper, the movement extraction of the arm of the sign language which motion picture measurement software is performed. The one with the marker has it by using the motion picture of the sign language that it is introduced to the public from the homepage of the independence administrative law person communication synthetic laboratory for the joint of the arm, hand. Then, a position coordinate is found from the marker chase. The movement of the sign language word is made data from a result of measurement.

key words motion picture measurement software, extraction, marker chase

目次

研究背景	1
研究目的	2
研究内容	3
3.1 動画計測ソフトウェア『Move-tr/2D』	5
3.1.1 概要	5
3.1.2 自動認識の原理	5
3.2 実験で使用した動画	6
3.3 手話動作の抽出方法	7
3.3.1 連続画像の取り込み	7
3.3.2 条件設定	7
3.3.3 マーカー付け及び追跡計測	7
3.3.4 結果表示	8
3.4 位置座標の変換	9
3.4.1 位置座標の変換(1)	9
3.4.2 位置座標の変換(2)	10
3.4.3 位置座標の変換(3)	10
実験結果	11
4.1 手話単語・挨拶	12
4.2 手話単語・曖昧	13
考察	14
まとめ	15
謝辞	16
参考文献	17
付録	18

目次

図 1.1 手話人口と聴覚障害者数の比較.....	1
図 3.1: しきい値の変化による 2 値化後の比較.....	6
図 3.2: 計測ソフト『Move-tr/2D』操作画面.....	8
図 3.3: 位置座標の変換処理.....	9
図 3.4: 左グループへの両肩距離の加算処理.....	10
図 4.1: 手話動作・挨拶(x 座標)	12
図 4.2: 手話動作・挨拶(y 座標)	12
図 4.3: 手話動作・曖昧(x 座標)	13
図 4.4: 手話動作・曖昧(y 座標)	13
付図 1: 手話単語(アドバイス)のワイヤーフレーム(1)	18
付図 1: 手話単語(アドバイス)のワイヤーフレーム(2)	19
付図 1: 手話単語(アドバイス)のワイヤーフレーム(3)	20
付図 1: 手話単語(アドバイス)のワイヤーフレーム(4)	21
付図 1: 手話単語(アドバイス)のワイヤーフレーム(5)	22
付図 1: 手話単語(アドバイス)のワイヤーフレーム(6)	23
付図 1: 手話単語(アドバイス)のワイヤーフレーム(7)	24

表目次

表 3.1: 使用する手話単語の一覧.....	3
表 3.2: マーカー箇所の一覧.....	4
表 3.2: 手話の動作特徴.....	6
付表 1: 手話動作(アドバイス)における動作のグループ分け.....	24

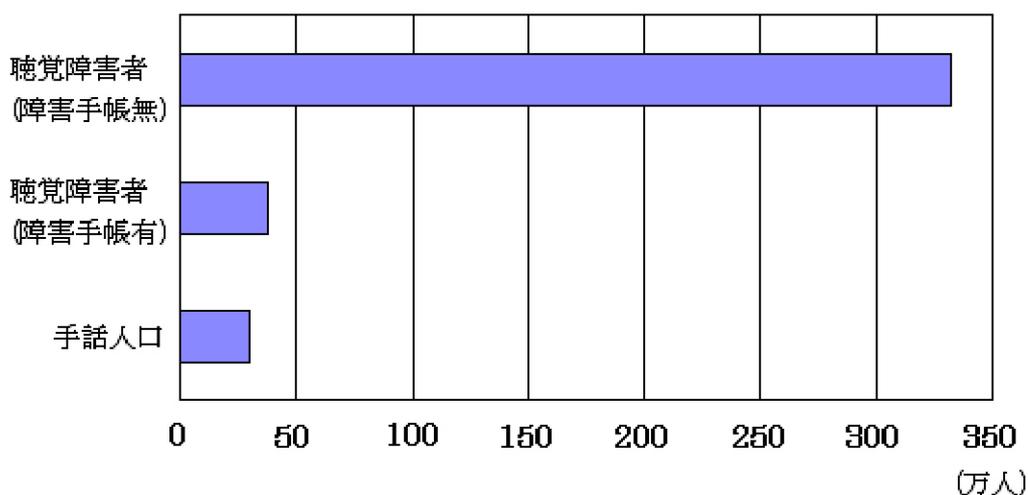
第1章

研究背景

聴覚障害者と健聴者とのコミュニケーションの手段として考えられる主なものには、筆談やジェスチャーなどが挙げられる。筆談の伝える媒体は文字であり、相手に情報を正確に伝えることができる。しかし、慣れない者同士であれば、その関係に隔たりを感じてしまうだろう。また、ジェスチャーは動作で情報を即時に伝えることができるので間延びすることなく会話を進めることができる。しかし、情報の伝え方に個人差があり、確実に伝わるとは考えにくい。

手話は聴覚障害者と健聴者とのコミュニケーション手段として会話と同じように感情なども交ぜて表現できる。そのため手話を使うには十分な訓練が必要となることから使える人は少ない。現在の日本の手話人口は約 30 万人である。それに対して聴覚障害者数は全国に 38 万人、障害手帳を交付されない 70db 未満の聴覚障害者を含めると約 400 万人とされている。70db 未満とは通常の会話レベルの音域であり、一般にこの音域未満の音が聞こえない障害を聴覚障害と指す。図 1.1 には手話人口と聴覚障害者との比較を示す。図より、手話人口は聴覚障害者の 1 割程度しかないのが現状である。手話をコンピュータで認識することにより、聴覚障害者と健聴者のコミュニケーションを円滑に進めることが期待できる。

図 1.1 : 手話人口と聴覚障害者数の比較



第2章

研究目的

動作の認識・生成を行う場合にはその動きを示すデータが必要になる。例えば、手話動作では腕や手の動きをデータ化してデータベース化することにより、認識のテンプレートとして利用することが期待できる。また、CG アニメーションと結びつけることにより、手話のアニメーションが期待できる。

本研究では手話単語映像データベースの映像を対象に手話動画から手話者の各関節の動きに着目し、その動きをデータ化することを目的とする。

第3章

研究内容

本研究では手話単語の動画(表 3.1)を対象に腕の動き抽出を行い，抽出結果から汎用性のある手話動作データへの変換を提案する。

まず，手話動作で重要とされる腕と手の関節の合計 44 箇所へマウスによるポイント入力によるマーカー付けを行った。手話における意思伝達は，その 8 割が腕の動きによるものである。そのことから，初期の実験ではマーカーの位置を腕部位の 6 箇所とした。しかし，手話動作の中で指の特徴的な動きが目立ち，指しか動いていないシーンも存在するため手部位も含めることとした。手の細かい動きを捉えるにはすべての関節をマーカーとする必要があるため，手部位の合計 38 箇所を新たに追加する事とした。次に，動画計測ソフトを使用してマーカー追跡を行い，腕の動きを抽出した。そして汎用性のある手話動作データへの変換として 3 段階の位置座標の変換を提案した。

表 3.1：使用する手話単語の一覧

手話単語名	動画時間(s)	シーン数	距離スケール(cm/画素)
挨拶	1.100	33	0.197
曖昧	1.267	38	0.199
諦める	1.100	33	0.194
明日	1.000	30	0.195
あなた	0.833	25	0.195
おいしい	1.000	30	0.193
アドバイス	1.267	38	0.197
笑う	1.000	30	0.194

表 3.2 : マーカー箇所の一覧

名称	ID	名称	ID	名称	ID
右親指付根	1	右小指付根	16	左薬指付根	31
右親指関節	2	右小指第 2 関節	17	左薬指第 2 関節	32
右親指先端	3	右小指第 1 関節	18	左薬指第 1 関節	33
右人差し指付根	4	右小指先端	19	左薬指先端	34
右人差し指第 2 関節	5	左親指付根	20	左小指付根	35
右人差し指第 1 関節	6	左親指関節	21	左小指第 2 関節	36
右人差し指先端	7	左親指先端	22	左小指第 1 関節	37
右中指付根	8	左人差し指付根	23	左小指先端	38
右中指第 2 関節	9	左人差し指第 2 関節	24	右肩	39
右中指第 1 関節	10	左人差し指第 1 関節	25	左肩	40
右中指先端	11	左人差し指先端	26	右肘	41
右薬指付根	12	左中指付根	27	左肘	42
右薬指第 2 関節	13	左中指第 2 関節	28	右手首	43
右薬指第 1 関節	14	左中指第 1 関節	29	左手首	44
右薬指先端	15	左中指先端	30		

3.1 動画計測ソフトウェア『Move-tr/2D』

本研究では、手話の動画からの腕の動きを抽出するために動画計測ソフトウェア『Move-tr/2D』[1]を使用した。

3.1.1 概要

動画計測ソフトウェア『Move-tr/2D』は、一定の時間間隔で連続的に取り込まれた画像から、追跡したい箇所を1画像上に最大4096個まで指定してマーカを自動認識し、マーカ面積重心を計測しながら自動追跡を行うプログラムである。マーカの認識は、明るさあるいは色で自動認識されるが、本研究で使用した動画のようにマーカが付加されていない場合はマニュアルによる指定を行う。

このソフトウェアで計測できる画像フォーマットは、AVI形式とBMP形式である。AVI形式の場合、使用するパソコンの「Windows Media Player」で再生可能なファイルであることが条件である。BMP形式の場合、ファイル名の最後に連続番号(1桁から8桁まで対応)のついた、BMPフォーマットの画像ファイルでなければならない。モノクロ画像の場合8ビット、カラー画像の場合24ビットで保存されている必要がある。各画像の時間間隔は一定であり、単位時間は7種類から選択できる。また、画像のサイズが一定でなければ計測できない。

3.1.2 自動認識の原理

このソフトウェアは画像処理の手法を使い、計測マーカの自動認識を行う。モノクロ画像の場合は明るさ(輝度)、カラー画像の場合は色で認識する。

8ビットのモノクロ画像は256段階の明るさ(輝度)を持っている。この明るさに対して一定のレベル(しきい値)で暗い部分と明るい部分に分ける(2値化する)。例えば、図3.1の画像の場合、しきい値(100)で2値化すると白い部分が2箇所出てくるが、しきい値(130)で2値化すると1箇所だけ抽出することができる。

カラー画像の場合、マーカは色で認識される。RGB毎の明るさを、モノクロ画像の場合と同じように2値化し、RGBすべて白になった部分が抽出される。

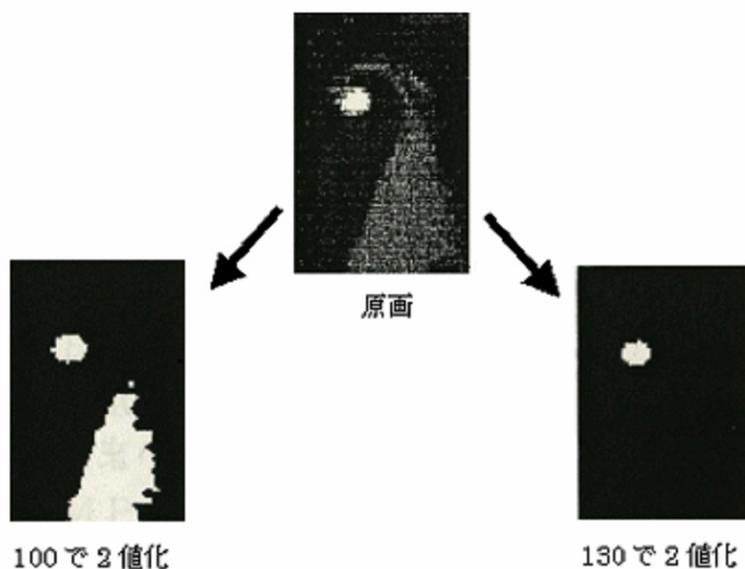


図 3.1 : しきい値の変化による 2 値化後の比較

3.2 実験で使用した動画

本実験で使用した動画は、独立行政法人通信総合研究所のホームページ[2]で公開されている手話動画を提供して頂いたものである。このホームページは手話単語 4114 語が収録されており、手話者に男性と女性の両方が用意されている。また、一つの手話単語にカメラ角度の異なる 3 パターン（前面、側面、斜面）がある。全日本ろうあ連盟発行「私達の手話(1)～(10)」[3]をベースに作成されており、データは 320x240 ピクセルの MPEG1 形式で構成されている。

実験で使用する手話動画には、カメラ角度が前面のものを使用する。これは 3 パターンのうちに前面が腕の動きを捉えるのに適しているためである。実験に使用する手話単語は、手話動作に注目し動作の類似している単語や動きが特徴的（左右、上下、円運動等）である単語を選択した。以下の表 3.2 に選択した単語と特徴について示す。

表 3.2 : 手話の動作特徴

手話単語名	動作する手	動作の特徴
挨拶	両手	左右対称, 上下運動.
曖昧	両手	タイミングをずらして円運動
諦める	右手のみ	手を開閉
明日	右手のみ	人差し指を上向けて前に出す.
アドバイス	両手	両手を合わせた後, 右手を出す.
あなた	右手のみ	人差し指で前方をさす.
おいしい	右手のみ	左から右へあごを撫でる.
笑う	右手のみ	あごを 2 回さわる.

3.3 手話動作の抽出方法

3.3.1 連続画像の取り込み

手話動作の抽出対象である動画ファイルを指定して取り込む。『Move-tr/2D』では、使用できるファイル形式が AVI 形式と BMP 形式であることに注意する。本実験では、AVI 形式で行ったため、以後の説明は AVI 形式に沿ったものとする。

3.3.2 条件設定

計測画像の時間スケールを設定を行う。本実験で使用した動画は、撮影コマ数/単位時間を 30/s として、取り込みインターバルは 2 とする。取り込みインターバルは、動画をコマ分割した場合のスタートシーンから 2 コマずつある同じコマを読み込ませないように 1 コマずつスキップさせるための設定である。

3.3.3 マーカー付け及び追跡計測

計測対象のマーカーを追跡計測する。追跡計測にはマーカーの自動認識による自動計測とマウスによるマニュアルポイント入力によるマニュアル計測がある。そのため、今回の実験での手話動画の抽出には 2 つの方法が考えられた。1 つ目は反射板を付けた被験者に手話の動きを真似てもらい撮影する方法である。2 つ目は実際の手話者による手話動画に対してマーカー付けを行う。1 つ目の方法は手話動作が真似たものであるため、動きの中に誤った動作が含まれてしまう可能性がある。本実験では手話動作の抽出を目的とするため訓練された手話者による手話動画にマーカーを付けてマニュアル計測を行う方法を選択した。しかし、1 つ目の方法がより正確な関節位置のマーカー付けが可能であるため、今後は 1 つ目の方法で被験者に実際の手話者を採用して行うことが必要である。

本実験で行ったマニュアル計測では、マウスでのマーカー付けを行わなければならないため、44 箇所のターゲットのマーカー付けを全シーンに対して行う。

3.3.4 結果表示

本研究ではマーカー追跡による抽出結果から以下のデータを使用する。全マーカーのシーン毎の座標データを得た。また、計測ソフト『Move-tr/2D』では、計測結果を画像、マーカー、フレームの有無を指定して表示できる(図3.2)。計測で得られたマーカー座標の結果のひとつとしてワイヤーフレームを作成した。付録には一例として手話動作(アドバイス)全シーンのワイヤーフレームを載せている。



図 3.2 : 計測ソフト『Move-tr/2D』操作画面

3.4 位置座標の変換

抽出結果である全マーカーの座標データは計測領域（画像サイズ）内での座標値であるため、計測領域に影響されるという問題点がある。そこでマーカーの座標データに変換処理を行い、それぞれ独立した座標値への変換を行う。また、体格の個人差に対応した座標値のスケーリング変換を提案する。

3.4.1 位置座標の変換(1)

位置座標の変換(1)では、抽出した各マーカーの座標データを計測領域にとられない座標データへの変換を行った。

両肩の座標値を基準点として、左右の腕、手をグループ化する。肩以外のマーカーの座標データは同グループである肩の座標値を基準に修正を行った。修正後は、肩座標は(0,0)となり、マーカー座標は肩の座標を引いた値へと変換されている。

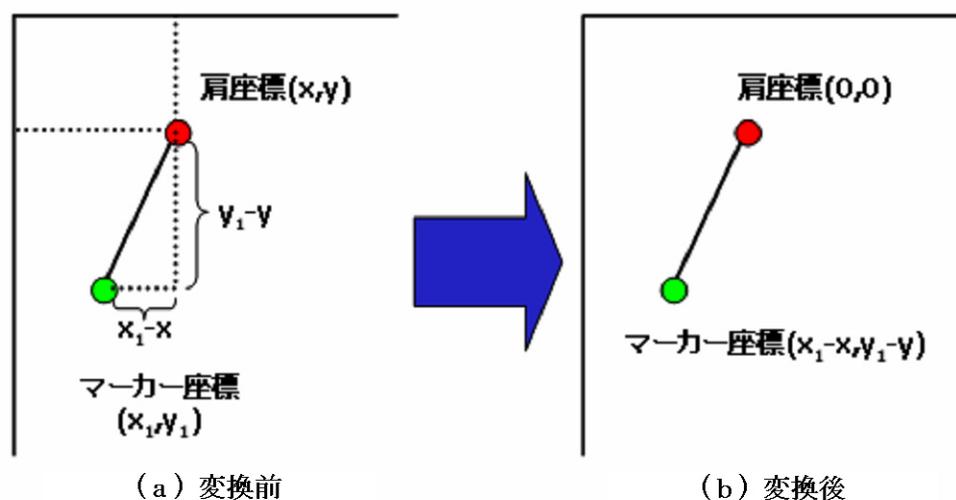


図 3.3 : 位置座標の変換処理

3.4.2 位置座標の変換(2)

位置座標の変換(1)によって左右の肩を基準点として座標値を修正しグループ化した。各マーカーはそれぞれの肩を基準点として位置座標としているため、両グループ間の位置関係の情報がない。そこで左グループのマーカーには右肩を基準とする座標値へと変換した。左グループ全てのマーカーに対して両肩距離の加算を行った。

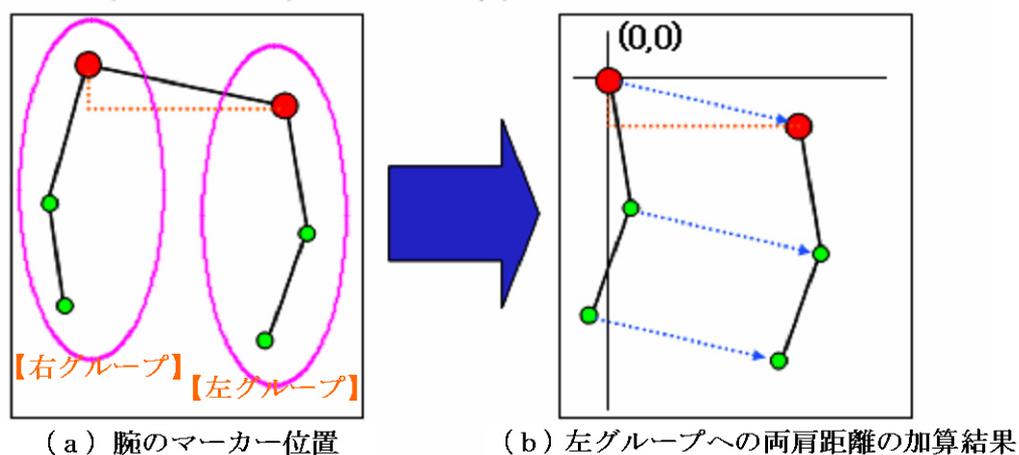


図 3.4 : 左グループへの両肩距離の加算処理

3.4.3 位置座標の変換(3)

位置座標の変換(3)では、位置座標の変換(2)の結果を用いて異なる体格に合わせたデータへの変換を行う。

1 コマ目の肩幅を基本単位として、対象となる体格の肩幅との比率をとり、得られた比率で全シーンのマーカーの座標値を求めた。

第4章

実験結果

本実験では、動画計測での抽出結果に前章の位置座標の変換(1)～(3)の処理を行い手話動作のデータ化とした。以下には実験から得た動作データを手話単語毎にグラフにした。手話動作の意思伝達で8割を占める腕部位の中から移動量の多い肘と手首についての x, y 座標値を示す。

4.1 手話単語・挨拶

図4.1: 挨拶 (x座標)

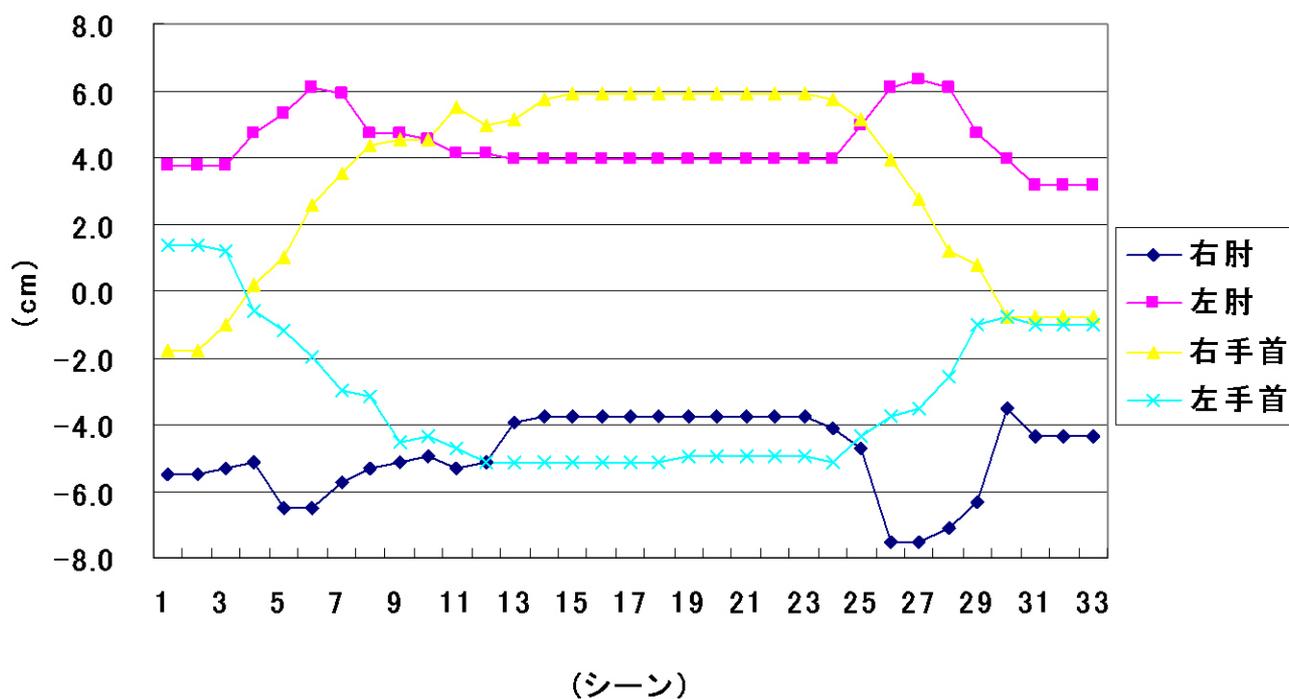
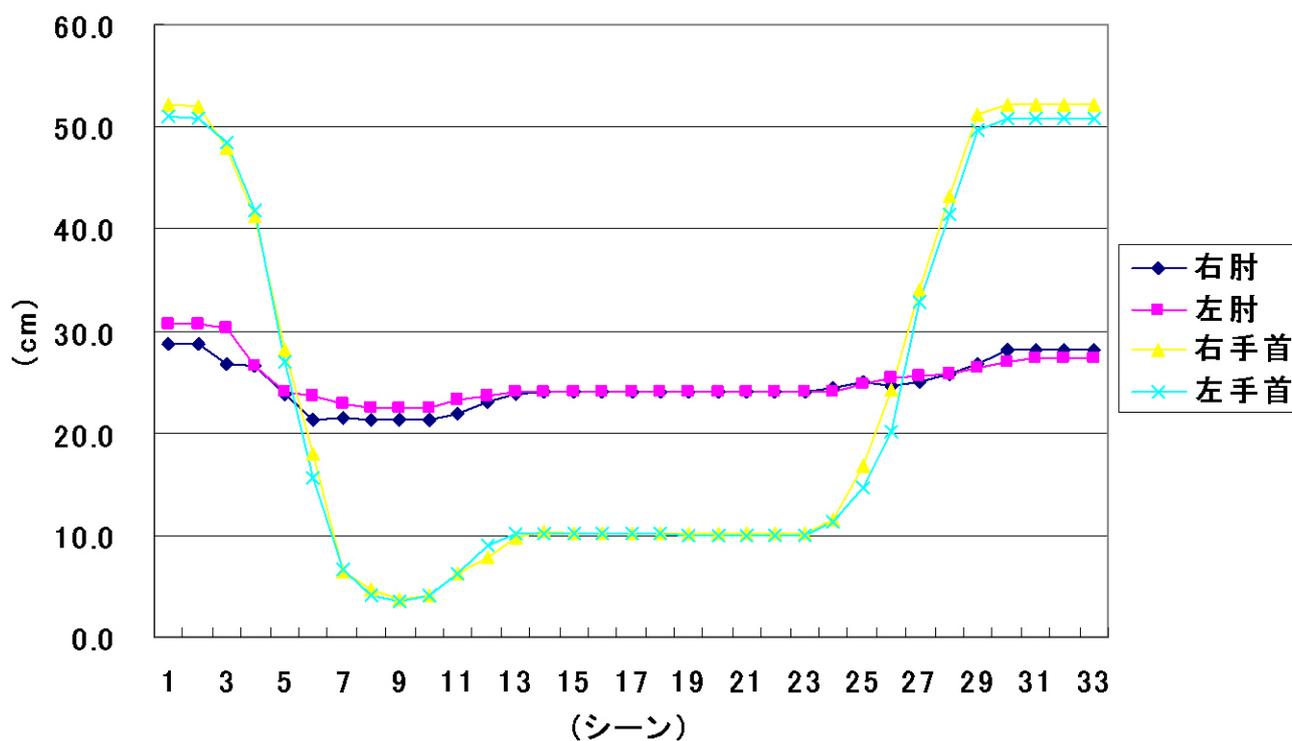


図4.2: 挨拶 (y座標)



4.2 手話単語・曖昧

図4.3: 曖昧(x座標)

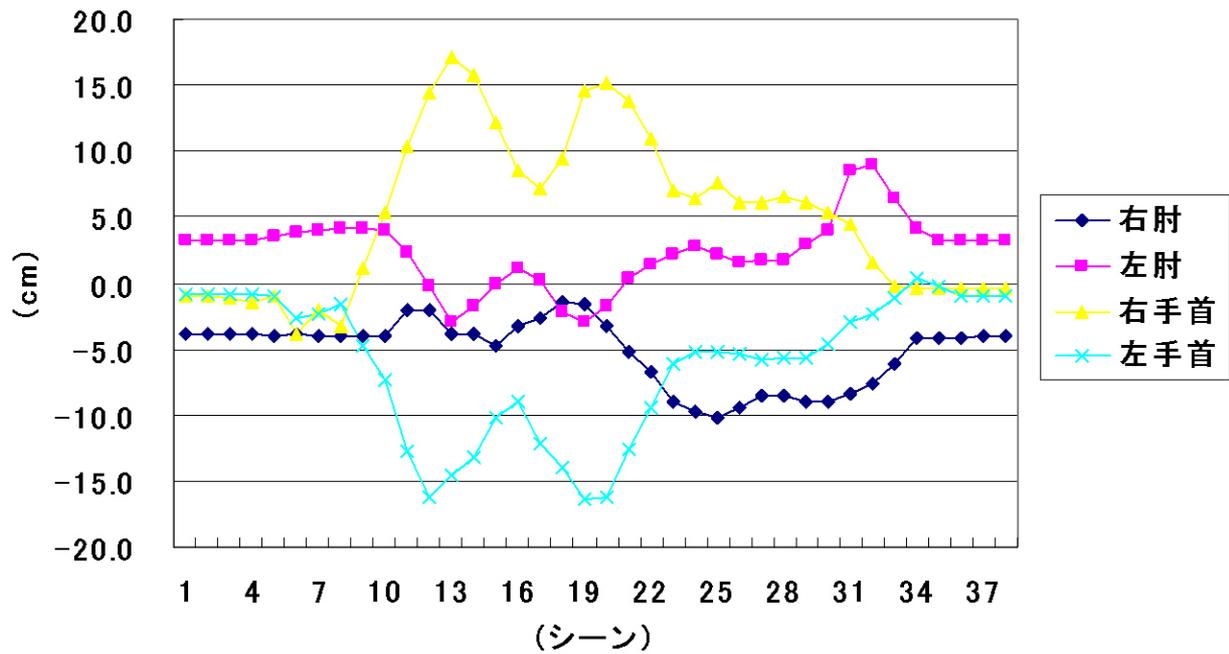
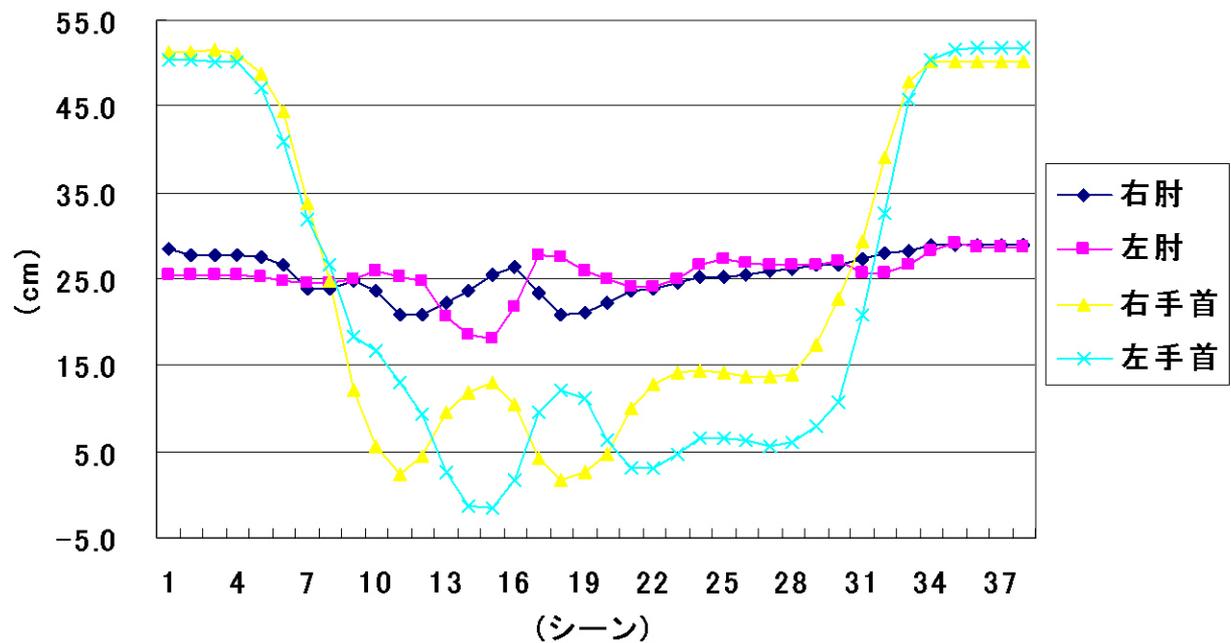


図4.4: 曖昧(y座標)



第5章

考察

本研究では、手話動画からの腕の動きのデータ化を行った。実験の結果、位置座標の変換により、データ化に適した肩を基準とした座標データを得ることで計測領域に影響される問題点を改善した。また、肩幅を基本単位とすることから異なる体格への手話動作のスケーリングが可能となった。さらに肩から肘、肘から手首などの距離の比率を個別に設定できれば、より正確な体格にあったデータを求めることができる。そこで、新たに3点間角度と2点間距離をデータに加えることが考えられる。

今回の実験では実際の手話動作であることに注目し、既存の手話動画を使用することにしたが、今後は訓練された手話者を対象に反射板をつけて撮影した画像を使用すべきである。これにより、マーカーをマウスでマニュアル入力する必要がなくなるため、マーカー付け時のわずかなズレを減少させることができ、より正確なデータを得られると考えられる。

第6章

まとめ

実験結果から手話動作のデータ化を進めることができた。今後の課題としては、さらに手話動作のデータを収集してデータベースを完成させることが挙げられる。そして将来的には手話動作の認識や手話 CG への応用が期待される。手話 CG への応用としては、腕の動きと顔表情の同期がとられた手話 CG の生成や、手話単語動画から余分な手話動作への移り変わり部分を取り除き手話単語間の渡りを補うことによる手話文章での CG 生成が期待される。

最後に、本研究において動画を提供して頂いた独立行政法人通信総合研究所に感謝いたします。

謝辞

本研究は、主査である高知工科大学情報システム工学科岡田教授には岡田研究室に所属して以来、終始懇切丁寧なるご指導、ご鞭撻を賜りました。また本論文をまとめるにあたり、主査として数々の貴重な教示を頂きました。深く御礼を申し上げます。

高知工科大学情報システム工学科竹田史章教授、篠森敬三助教授には本研究の副査としてご指導を賜りました事深く御礼を申し上げ、ここに感謝の言葉とさせていただきます。

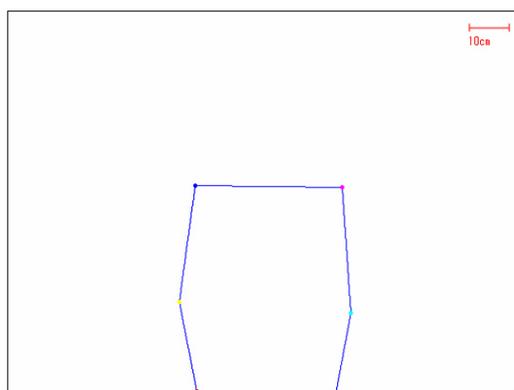
また、本研究や研究室活動に対して御支援や助言を下さりました高知県工科大学情報システム工学科岡田研究室4回生の秋田正氏、麻生太郎氏、鎌倉崇之氏、安友一秋氏、3回生の川島亨介氏、藤村和人氏、山本拓氏に深く感謝いたします。

最後に、本論文をまとめるにあたり、理解を持って協力、励ましてくださった両親と家族に深く感謝いたします。

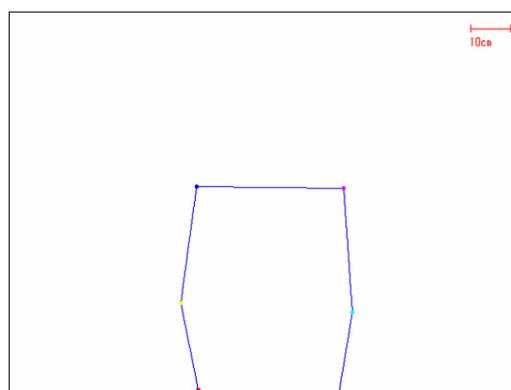
参考文献

- [1] 株式会社 ライブラリー, "「Move-tr/2D」動画計測ソフトウェア User's manual", 2002
- [2] 独立行政法人通信総合研究所, "手話単語映像データベース", <http://jsl-db.crl.go.jp/jsl-db/>, 2002
- [3] 全日本ろうあ連盟, 全日本聾唖連盟手話研究委員会=編, "わたしたちの手話(1)～(10)", 全日本聾唖連盟出版局

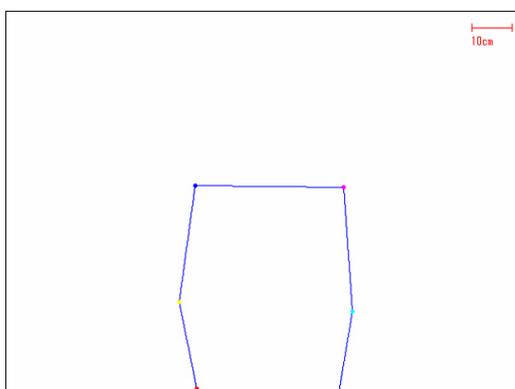
付録



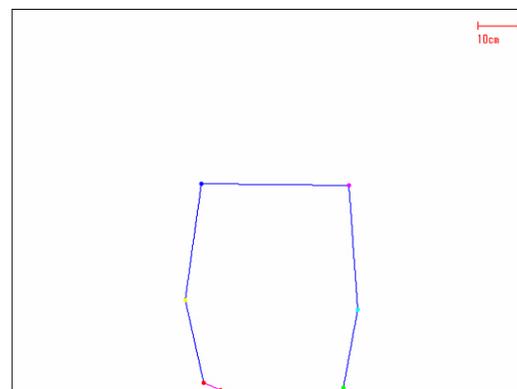
(a) シーン 1



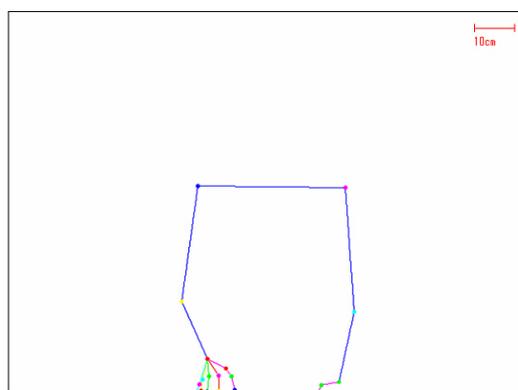
(b) シーン 2



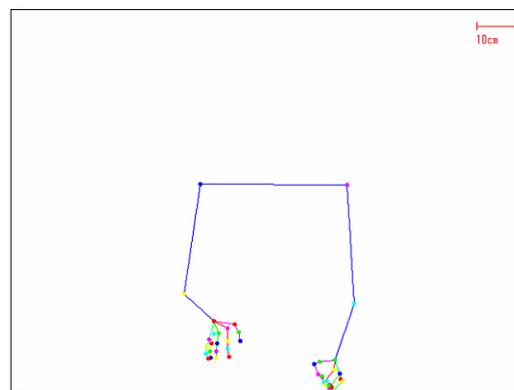
(c) シーン 3



(d) シーン 4

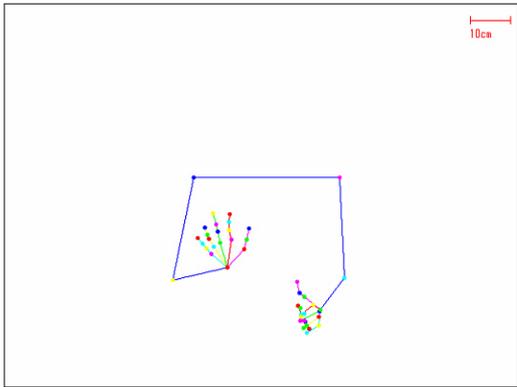


(e) シーン 5

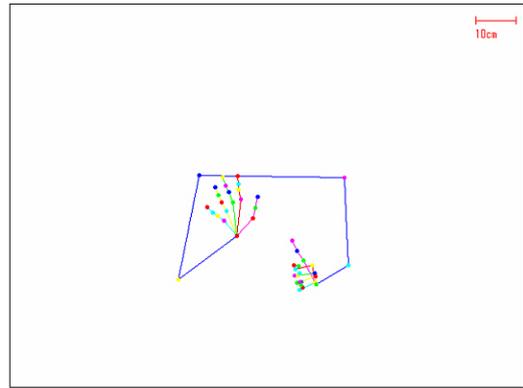


(f) シーン 6

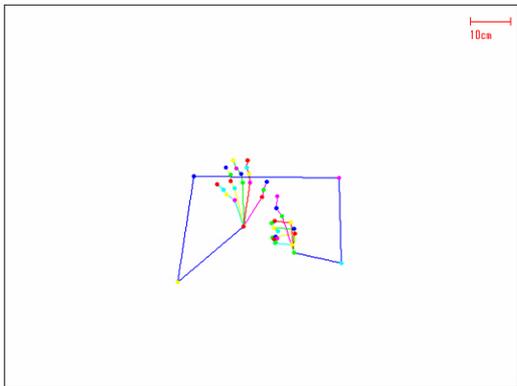
付図 1: 手話単語(アドバイス)のワイヤーフレーム(1)



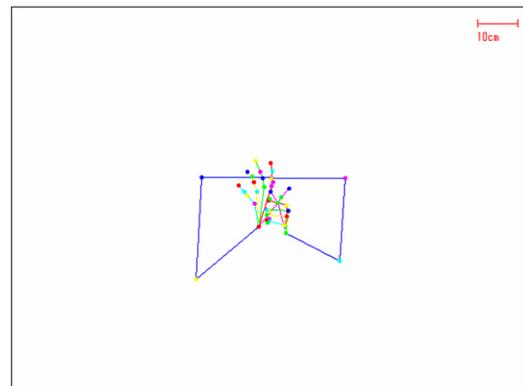
(g) シーン7



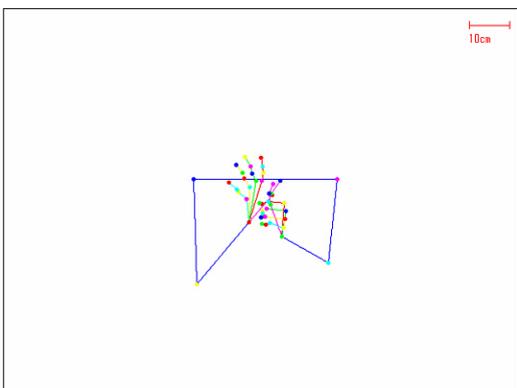
(h) シーン8



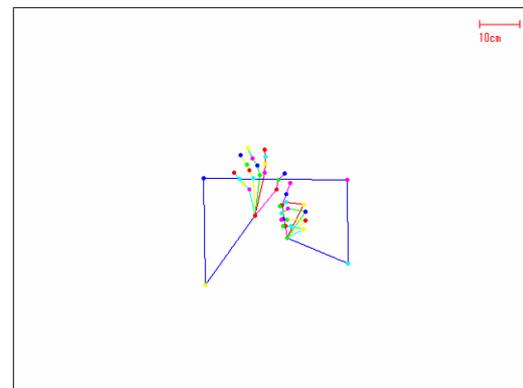
(i) シーン9



(j) シーン10

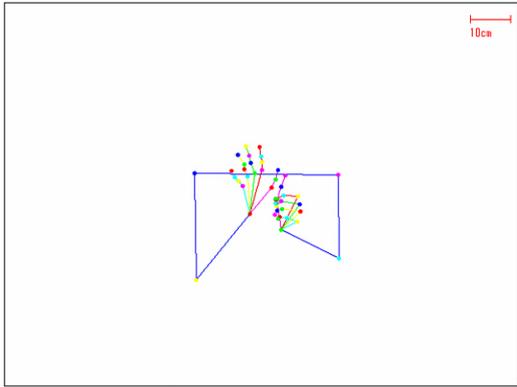


(k) シーン11

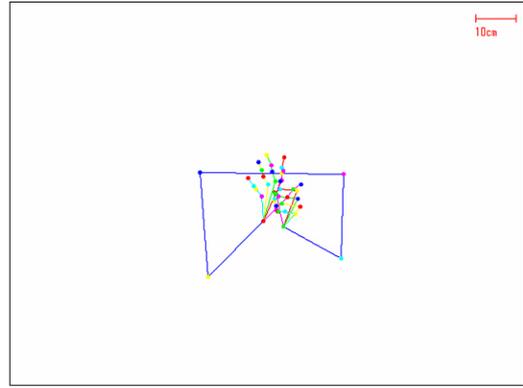


(l) シーン12

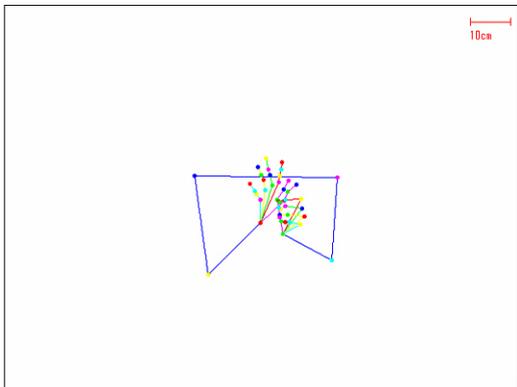
付図1: 手話単語(アドバイス)のワイヤーフレーム(2)



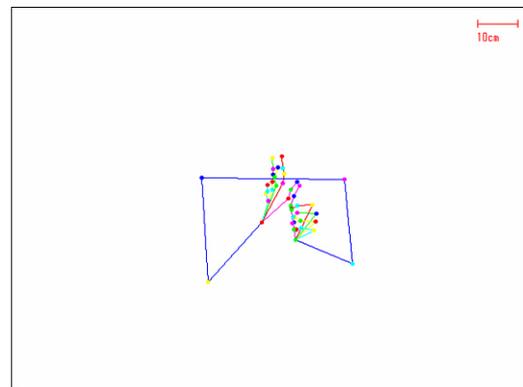
(m) シーン13



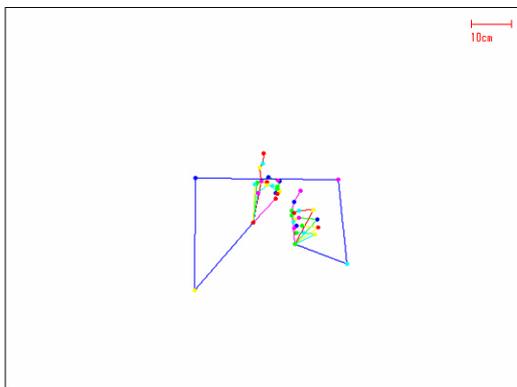
(n) シーン14



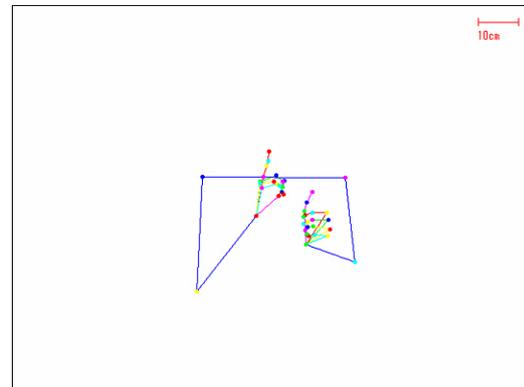
(o) シーン15



(p) シーン16

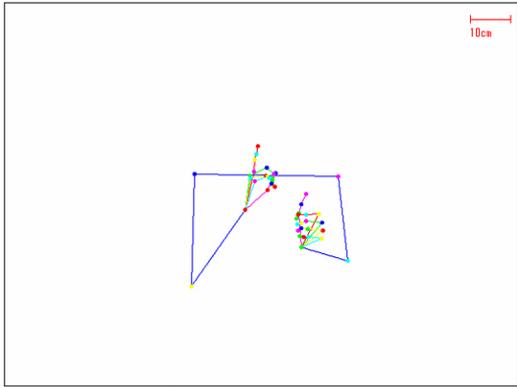


(q) シーン17

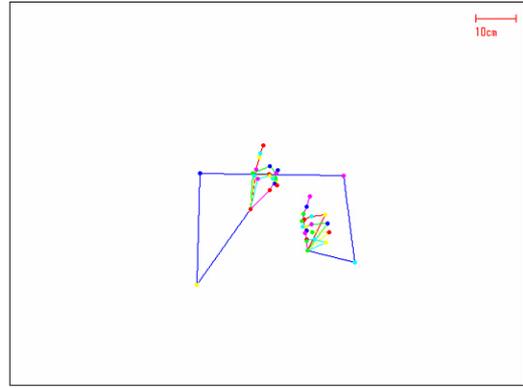


(r) シーン18

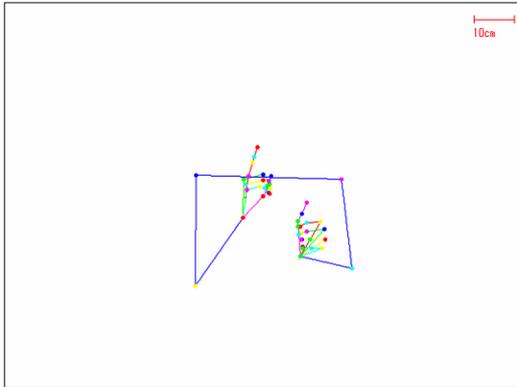
付図1: 手話単語(アドバイス)のワイヤーフレーム(3)



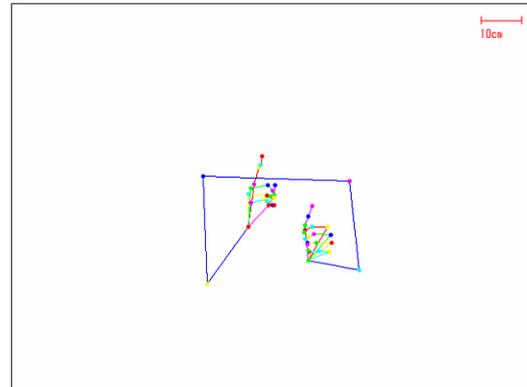
(s) シーン19



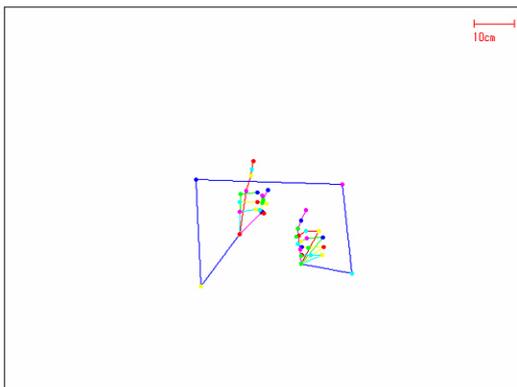
(t) シーン20



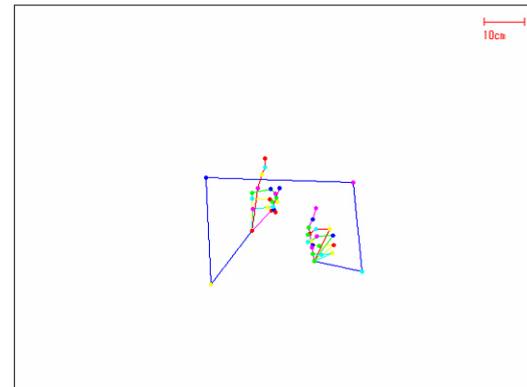
(u) シーン21



(v) シーン22

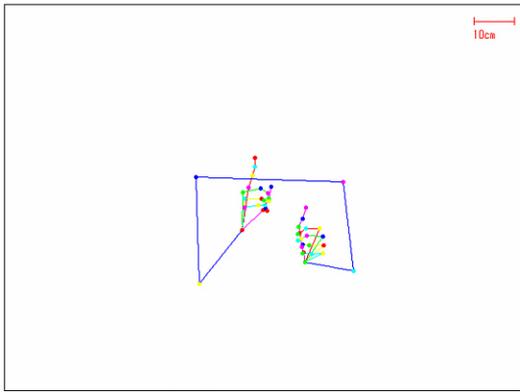


(w) シーン23

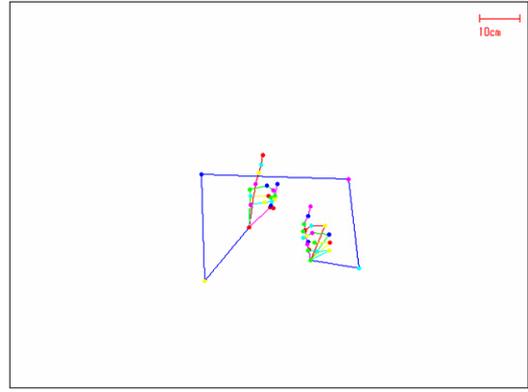


(x) シーン24

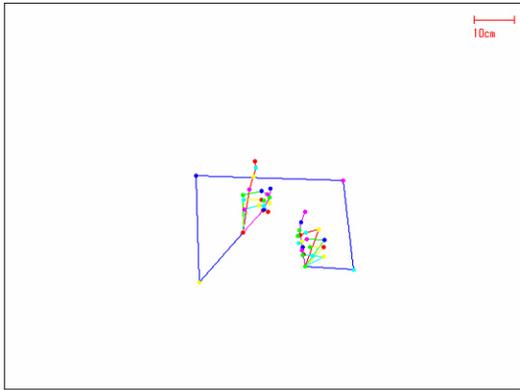
付図1: 手話単語(アドバイス)のワイヤーフレーム(4)



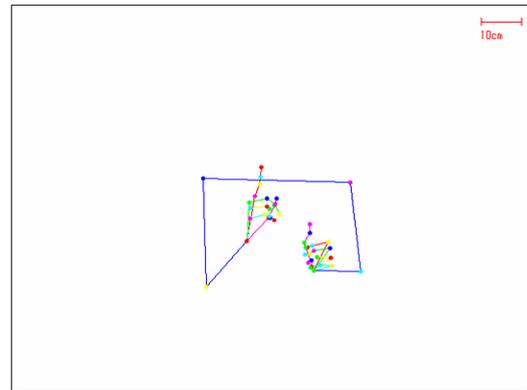
(y) シーン 2 5



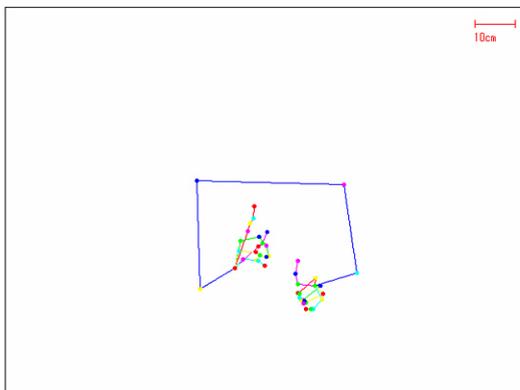
(z) シーン 2 6



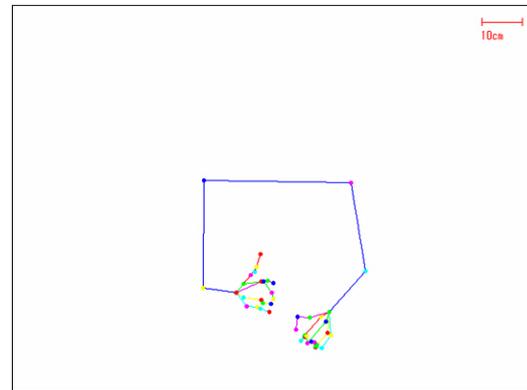
(a') シーン 2 7



(b') シーン 2 8

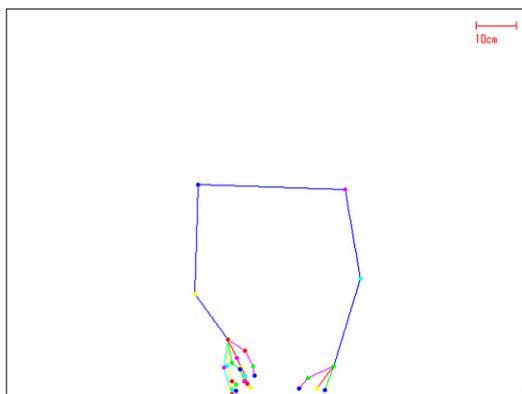


(c') シーン 2 9

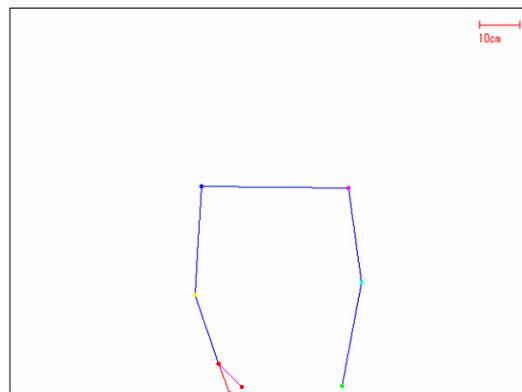


(d') シーン 3 0

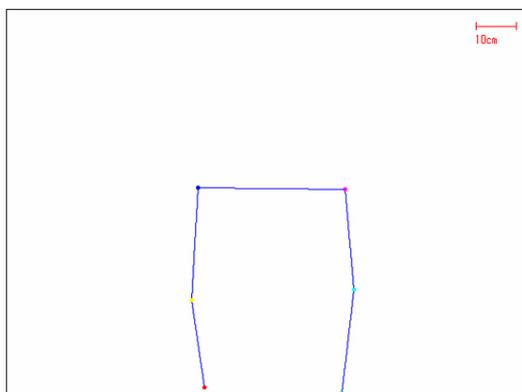
付図 1: 手話単語(アドバイスのワイヤーフレーム(5)



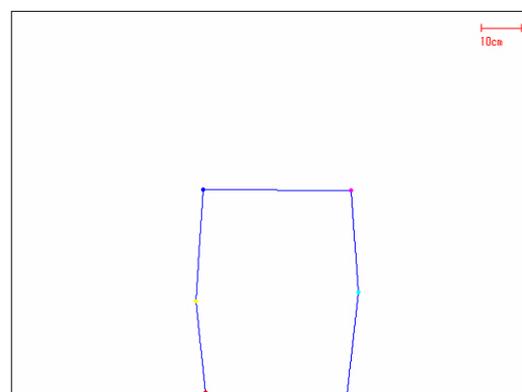
(e') シーン 3 1



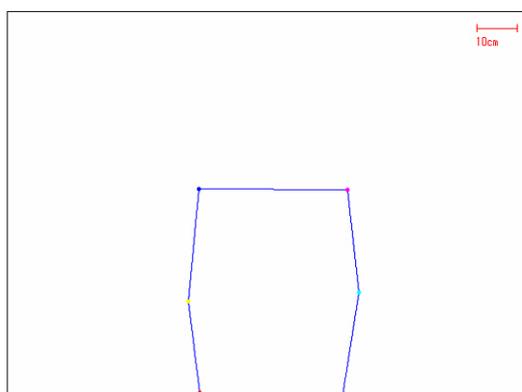
(f') シーン 3 2



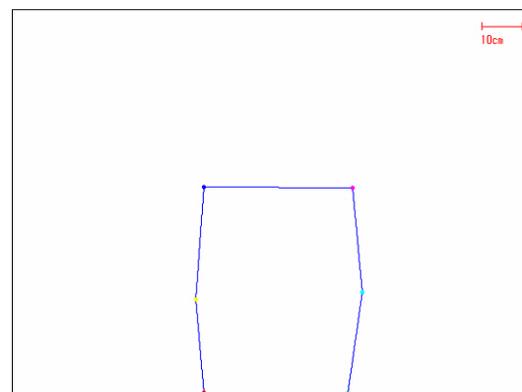
(g') シーン 3 3



(h') シーン 3 4

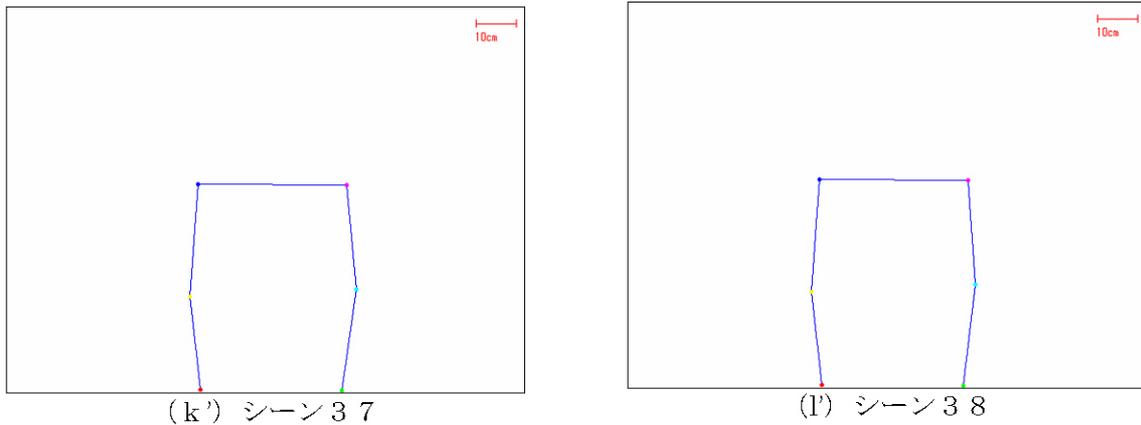


(i') シーン 3 5



(j') シーン 3 6

付図 1: 手話単語(アドバイス)のワイヤーフレーム(6)



付図 1: 手話単語(アドバイス)のワイヤーフレーム(7)

このワイヤーフレームは手話単語のアドバイスを意味する。ワイヤーフレームの動きを見てみると、シーン 4 から両腕に動きが見られる。以後の一連の動作を 4 つの動作群に分けることができる。このグループ分けは表 6 に示す。手話単語(アドバイス)ふたつの動作パターンにより構成されている。シーン数で見みても手話への移り変わりの動作では 4, 5 シーンと少なく両手揃って胸の前あたりと下げ降ろした位置を移動する共通点がある。

付表 1: 手話動作(アドバイス)における動作のグループ分け

グループ名	シーン数	手話動作
第 1 グループ	4 ~ 9	手話の初動作までの動作
第 2 グループ	1 0 ~ 1 6	開いた右手と握る左手で叩き合わせる動作。同じ動作を 2 回繰り返す。
第 3 グループ	1 7 ~ 2 8	右手を前方へ出す動作。
第 4 グループ	2 9 ~ 3 3	手話の動作が終わり戻りの動作