

Web ブラウザを利用した 手話動作の検索・表示システム

柴田幸男 八村広三郎

立命館大学 理工学部

これまで人間の身体運動を計算機上で処理するため、さまざまなモデル化の手法が数多く考案されてきた。しかしこれらはそれぞれ長所短所があり、どのような身体運動を扱うかによって手法を使い分けることが適当である。本研究では、人間の身体運動のとくに手話の動作に重点を置いたモデル化の手法を提案する。ここでは、ASL(American Sign Language:アメリカ手話)を対象として取り上げる。インターネット上での利用を想定してモデリングと表示に VRML を使用し、手話動作の検索・表示システムの構築を行った。このシステムにより単語単位の検索が可能となった。今後は文単位の検索や日本語手話への対応などを検討している。

Sign Language Display System by using Web Browsers

Yukio Shibata Kozaburo Hachimura

Faculty of Science and Engineering, Ritsumeikan University

Several computerized methods for handling human body movement have been studied, each of which has its own advantage and disadvantage. It is appropriate to choose the best one depending on the purpose of computer processing. This paper describes a system for retrieving and displaying human body motion of sign languages, which is intended to be used for learning sign languages. Lots of schemes for sign language exist, but in this research we deal with American Sign Language (ASL) as the first trial. The human body has been modeled by using VRML, and the arm gesture has been specified by supplying change of joint angles of the body model. Pose of hand has been changed by switching models of hand depending on the word specified. The system is operable on the internet, and users can display body movement of sign language, on a Web browser, by just specifying corresponding words.

1 はじめに

人間の身体運動を計算機上で分析・解析・表示する技術は、ここ数年の計算機の処理能力の向上とグラフィック性能の向上にともない飛躍的に進歩した。そして、スポーツ科学や映画・アニメーション映像などの分野でその実用性が認識されはじめている。現在、このようなさまざまな分野に応用可能な人間の身体動作の入力・記述・記録・検索・表示のための総合的な身体運動処理システムの構築は、本研究室のテーマのひとつとなっている [1-3]。

本研究では手話動作という身体運動にとくに重点をおき、手話動作を計算機上で入力・記述・表示するための手法を考案した。また、近年のパーソナルコンピュータおよび WWW(World Wide Web) の普及とそれに伴った Java や VRML など代表されるインターネット技術の急速な発展により、Web ブラウザをとおして高度なネットワーク型アプリケーションを提供することが可能となった。これらの要素技術をベースとして手話の学習機会の提供を目的とした手話動作の検索・表示システムの開発をおこなった。

2 手話について

2.1 手話に関するこれまでの研究

手話は、手指・顔の表情・口形および身体の動作によるコミュニケーション手段の1つである。近年、この手話を対象とした工学的アプローチの研究が盛んに行われている。聴覚障害者と健聴者との情報格差のひろがりや、手話通訳者の不足などの問題を解消するものとして各方面から期待されている。

この分野での重要な要素技術としてこれまで手話単語の認識が研究されてきた。これは、手話動作や顔の表情、口形などの画像による入力情報から手話単語を判別・認識するものである [4]。そのほか手話の電子辞書の構築にともない手話動作間の類似性の検出や分類方法などの研究もおこなわれてきた [5]。

2.2 本研究で扱う手話について

本研究では、英語をもとに開発されたアメリカ手話 (American Sign Language : 以後 ASL と呼ぶ) を対象とする。ASL は意味が比較的類似した複数の英単語に対してひとつの動作が定義されており、この英単語を手話単語、動作を手話動作と呼ぶ。ASL の手話動作を

規定するパラメータとして Stokoe らが提案した、手の形 (Designation)・手の動き (Signation)・動作位置 (Tablation) の3つを使用するモデルがある [7]。その他、掌の向き (Orientation) を追加した4つのパラメータを用いる Friedman のモデル [8] や、手話の継時的要因に着目した継時モデルなどが提案されている [9]。

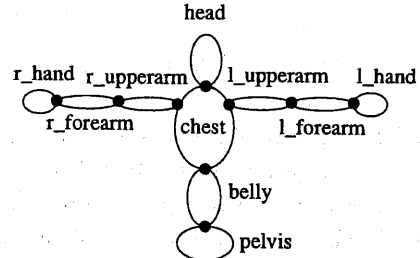
本研究では手話動作を腕の動作と指の動作に2つの要素に分割したモデル化を行った。

3 手話動作のモデル化

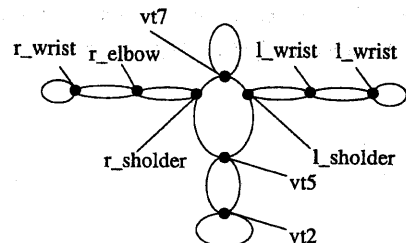
3.1 人体モデルの定義

人体モデルとは、現実世界の人間を抽象化したもので、身体運動を計算機上で分析・解析・表示するうえで最も基礎となるものである。

本研究では、人体モデルをセグメント (骨格) とジョイント (関節) の2つの要素によって構成した。本研究で使用する人体モデルを図1に示す。図中の楕円がセグメント、黒丸がジョイントを表す。それぞれのセグメント、ジョイントには一意に識別するために図中に示すような名称が付けられている。



(a) セグメント名



(b) ジョイント名

図1: 人体モデルの定義

3.1.1 人体モデルの姿勢付け

人体モデルの各ジョイントに回転角度を入力することにより、人体モデルをある特定の姿勢にすることができる(図2)。本研究では人体モデルの定義にはVRMLを使用した[12]。人体モデルの姿勢付けにはVRMLの角度指定方法もちいる。VRMLにおける回転角度は、4つの浮動小数点数からなる配列によって定義されている。図3に示すように、はじめの3つの数値が回転軸であらわし、最後の数値がその軸のまわりの回転角度であらわす。

各ジョイントに回転角度を入力する

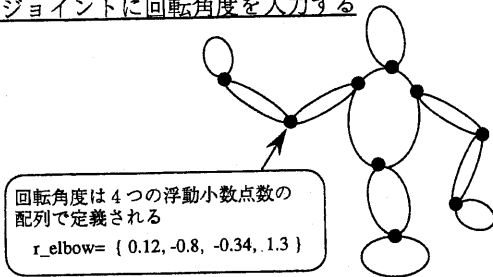


図2: 人体モデルの姿勢付け

回転角度 {x,y,z,r} の定義

1. 回転中心 $[0,0,0]$ と $[x,y,z]$ で回転軸を形成する
 2. 回転軸のまわりに $r(\text{radian})$ 回転する
-

図3: VRMLにおける回転角度の定義

3.2 ASL 手話動作のモデル化

図4にASL手話「charge」の例を示す。本研究では、前述のようにASLの手話の動作を腕の動作と指の動作の2つの動作の組み合わせと考える。図4の例では、「右手のx形を」が指の動作で、「左手の掌にあてがい、直角にこするように降ろす」が腕の動作と考えることができる。

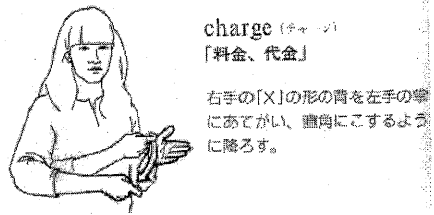


図4: ASL手話「charge」の例(参考文献[10]より)

3.2.1 腕の動作

腕の動作は、図1に示す人体モデルのl_wrist, l_elbow, l_shoulder, r_wrist, r_elbow, r_shoulderの6ヶ所の各ジョイントに時間経過とともに関節角度を指定することによって動作付けをおこなう。

3.2.2 指の動作

ASLにおける指の動作は、図5に示す「ASLアルファベットサイン」と呼ばれる指の形によって指定される。これには各ASLアルファベットサインに対応する指の三次元形状をあらかじめ作成しておき、時間経過とともに指の形状を切り替えることによって動作付けをおこなうようにした(図6)。

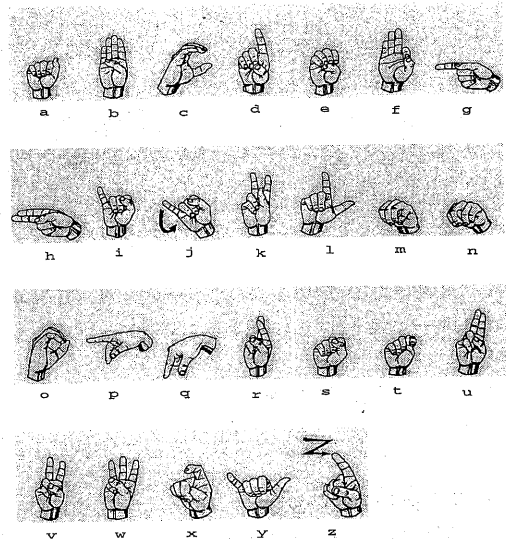


図5: ASLアルファベットサイン

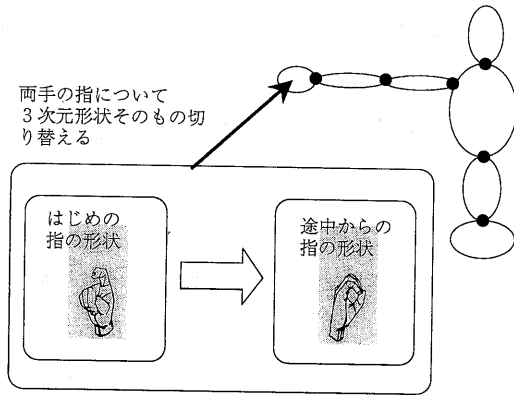


図 6: 指の動作付け

3.2.3 手話動作の記述データ

本研究では、ひとつの手話動作を構成する時間を100単位時間として設定する。単位時間ごとに人体モデルの腕の姿勢と指の形を指定し、ひとつの手話動作を構成する。ただし、すべての単位時間で腕の姿勢と指の形を指定する必要はなく、省略された単位時間での腕の姿勢と指の形は、前後関係の姿勢から補間したものを用いる。指の形の指定は、図5のASLアルファベットサインに整数の番号を与え、その番号によって行う。

ひとつの手話動作に必要なデータは、図7のようなテキスト形式で保存する。一行目に100単位時間に相当する実時間を記述する。以降は単位時間ごとに、腕の姿勢あわすl_wrist, l_elbow, l_sholder, r_wrist, r_elbow, r_sholderの各ジョイントに入力する回転角度と、左手/右手に入力する指の形を整数値で記述する。

4 手話動作の検索・表示システム

ここでは、これまでに述べた手法をもちいて手話動作の検索・表示システムの構築について述べる。

4.1 システムの機能

システムに必要とされるおもな機能を以下に示す。

1. 手話動作の登録機能

手話動作の検索・表示を行うためには、まず手話動作を計算機システムに登録する必要がある。手話動作は人体モデルのキーフレーム法によるアニメーションによって表現するので、各キー

1.1	# 実時間(秒)
0	# 0 単位時間
0 0 1 0	# l_wrist
0 0 1 0	# l_elbow
0 0 1 0	# l_sholder
0 0 1 0	# r_wrist
0 0 1 0	# r_elbow
0 0 1 0	# r_sholder
1	# 左手サイン
1	# 右手サイン
31	# 31 単位時間
-0.65 -0.19 0.72 0.84	# l_wrist
1 0 0 4.35	# l_elbow
-0.81 0.40 0.40 0.67	# l_sholder
-0.64 0.30 -0.70 0.62	# r_wrist
1 0 0 4.19	# r_elbow
-0.69 -0.16 -0.69 0.59	# r_sholder
2	# 左手サイン
2	# 右手サイン
68	# 68 単位時間
-0.59 -0.59 0.54 0.96	# l_wrist
1 0 0 4.35	# l_elbow
-0.81 0.40 0.40 0.67	# l_sholder
0.57 -0.67 0.46 5.24	# r_wrist
1 0 0 4.37	# r_elbow
-0.69 -0.16 -0.69 0.59	# r_sholder
0	# 左手サイン
0	# 右手サイン
94	# 94 単位時間
0 0 1 0	# l_wrist
0 0 1 0	# l_elbow
0 0 1 0	# l_sholder
0 0 1 0	# r_wrist
0 0 1 0	# r_elbow
0 0 1 0	# r_sholder
1	# 左手サイン
1	# 右手サイン

図 7: 手話動作データのファイル保存形式
(手話動作「accident」の例)

フレームにおける人体モデルの姿勢を入力する手段が必要である。計算機システムには図7に示したファイル形式で保存する。

2. 手話動作の検索機能

登録させた手話動作データは、手話単語によって検索できなければならない。また、手話動作どうしの類似度による検索や、逆に手話動作から手話単語を検索する機能があることが望ましい。

3. 手話動作の表示機能

検索された手話動作データは、画面上で実際に人体の動きとしてシミュレートされ、表示されることが必要である。また、画面内の人体モデルを任意の視点から表示したり、表示速度を変更できることが望ましい。

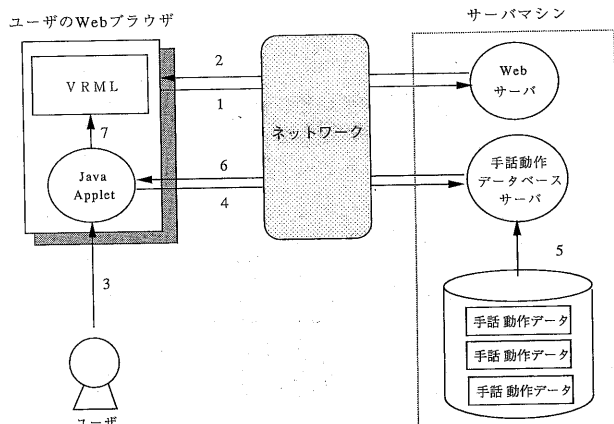


図8: システム構成

4.2 システムの概要

作成した本システムの構成を図8に示す。本システムはWebブラウザで本システムが稼働しているサイトにアクセスすることにより利用できる。

ユーザはJavaアプレット上で、表示したい手話単語を指定する。Javaアプレットは手話単語に対応する手話動作データを手話動作データベースサーバから取得し、その内容を解釈してVRMLの人体モデルによって手話を表示する。

手話動作データベースサーバは、Webサーバと同じマシン上で稼働させユーザ側のJavaアプレットからの手話動作データ要求に対して、図7に示した手話動作データをバイトストリームに変換し、ソケット通信で返送する。

おもな処理の手順を以下にまとめる。

1. Webブラウザでシステムが稼働しているサイトにアクセスする
2. Webサーバから必要なHTMLファイル、VRMLファイル、Javaアプレットなどが転送される
3. ユーザがJavaアプレットのGUI上で手話単語を指定する
4. Javaアプレットが手話動作データベースサーバに手話動作データを要求する
5. 要求を受けた手話動作データベースサーバは手話動作データをファイルからロードする

6. 手話動作データをバイトストリームに変換して、要求元のJavaアプレットに返送する

7. Javaアプレットは、受け取った手話動作データをVRMLデータ形式に変換し、人体モデルで手話動作を表示する

4.3 手話の登録

図7に示すような手話動作データの生成を容易におこなうために手話動作エディタを開発した(図9,10)。

図9は、単位時間(キーフレーム)での各ジョイントの回転角度や指の形の情報を表示するためのものである。

図10の画面中央の人体モデルをマウスで直接ドラッグすることで腕の動作を、画面右下のパネルで指の動作をそれぞれ設定することができる。図10の画面によって設定したジョイント角度などの情報が図9の表示パネルに反映される。

4.4 手話の検索・表示

ユーザがWebブラウザによって本システムが稼働しているサイトにアクセスしたときの表示画面の例を図11に示す。画面右側のウィンドウがWebブラウザであり、その中にVRMLブラウザをもちいて人体モデルを表示する。画面左側のウィンドウはJavaアプレットが表示するものであり、ユーザはこのウィンドウ内に表示されるGUIで操作をおこなう。

Javaアプレットは起動直後に手話動作データベースサーバと通信して、図12のようにサーバに登録されて

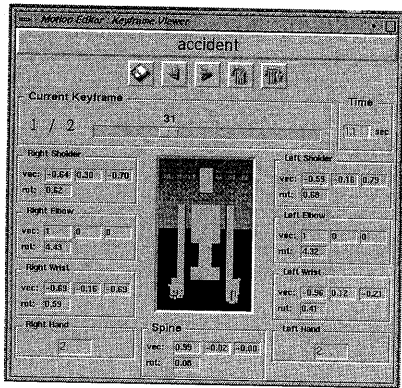


図 9: 手話動作エディタ画面 1

いる手話動作データの単語の一覧リストを表示する。

ユーザがこのリストの単語をクリックすると手話動作データベースサーバから手話動作データをロードし、さらにデータを VRML データ形式に変換する。図 13 に示す表示操作ウインドウで手話動作の表示や、コマ送り表示などの操作をおこない、VRML ブラウザ内の人体モデルに動作を表示される。

図 14 に示す ASL 手話単語「accident」を入力したときに出力される手話動作の 10 単位時間ごとの表示を図 15 に示す。

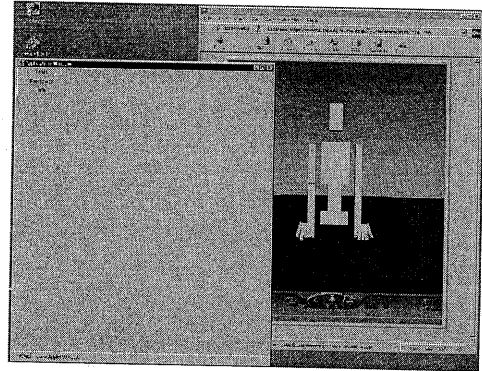


図 11: システムの表示画面

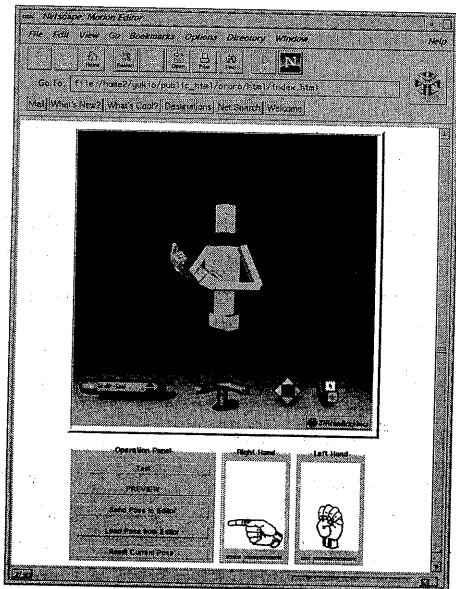


図 10: 手話動作エディタ画面 2

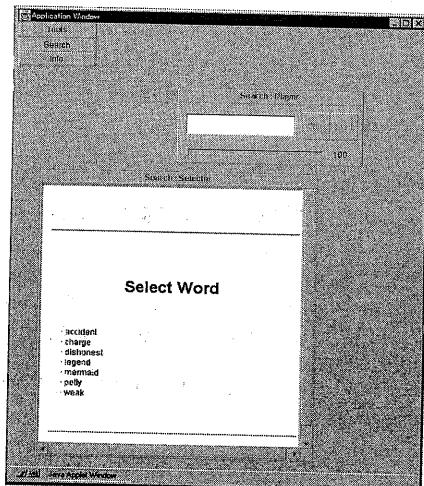


図 12: Java アプレットによる単語指定画面

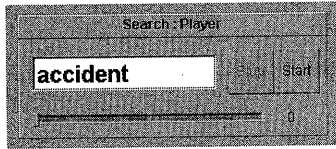


図 13: 表示操作ウインドウ



図 14: ASL 手話「accident」の例 (参考文献より [10])

4.5 システムの実装

本システムは、現在のところ Windows マシン上で VRML ブラウザ CosmoPlayer2.0 のプラグインがインストールされた Netscape4.04 で動作する。また手話動作データベースサーバは perl で記述されているので perl5.004 以降がインストールされているシステムであれば問題なく動作する。

4.6 考察

今後は以下のような点についてさらに検討が必要であると考えている。

(1) 現在のシステムでは手話単語単位でしか手話動作を表示できないが、これをより発展させて文全体を入力してそれを手話動作で翻訳できるように現在検討中である。この場合、連続的に手話動作を表示することになるので、ひとつの手話動作の表示の終了と次の手話動作の開始をいかになめらかに違和感なくつなげるかが課題となる。また ASL の単語だけに限らず、日本語手話の単語にも対応することも考える。

(2) 本研究での手話動作のモデル化には考慮にいれなかった顔の表情や口形なども手話動作において重要な要素なのでこれらも表現できるようにすることがあげられる。それには、本研究で使用した直方体を組み合わせただけの人体モデルでは実現できないので、図 16 のような人体モデルに変更する必要がある [13]。

(3) 手話動作の記述は図 7 に示したように各ジョイン

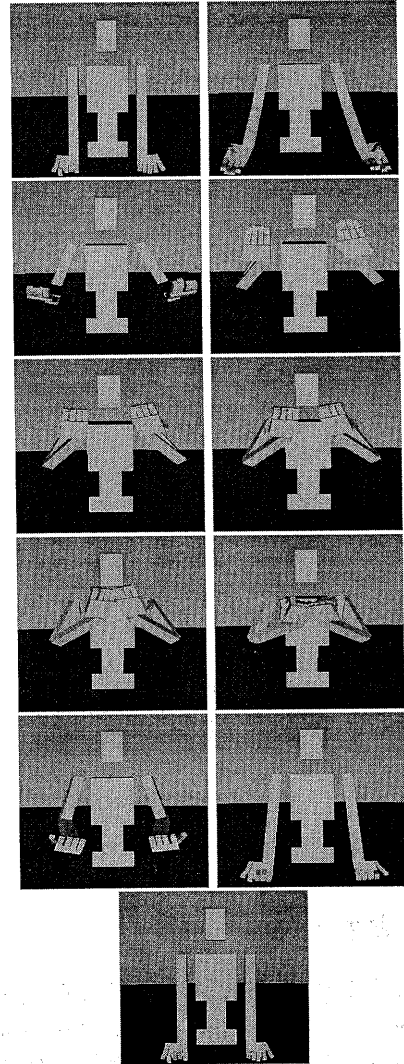


図 15: 手話単語「accident」の表示

トにおける回転角度で記述した。しかし回転角度だけの記述では、人体モデルのセグメントサイズが変更されると表示される動作が異なる場合がある。人体モデルが変更されても動作内容の違いに影響のすくない記述方法が必要である。

(4) 本システムにおける手話動作データは、現状ではすべてシステム管理者が入力しなければならない。これはかなり負担になる。ユーザも Web ブラウザから手話動作データの入力ができる機能が必要とおもわれる。

(5) 現段階では計算機に登録した手話動作データも少ない。また、手話動作の表示もただ画面に表示される単語を指定するだけである。「検索」という意味はまだまだ機能的には不十分である。今後、単語の意味的な類似・手話動作の類似によって検索する機能が望まれる。



図 16: 人体モデルの例

5 おわりに

本研究では、身体運動のなかでもとくに手話の動作を対象とした検索・表示システムについての開発研究をおこなった。手話動作の登録・表示・検索の作業が、インターネットに接続されたマシンと Web ブラウザさえあれば、世界中のどこからでも行えるというネットワークの機能を活かしたシステムとして実現できた。また VRML という新しい技術が複雑な身体運動の表示に充分対応できる機能と性能をもっていることがわかった。

参考文献

- [1] 平松尚子, 八村広三郎: 動作生成ライブラリを利用した身体運動の記述と表示, 情報処理学会人文科学とコンピュータ学会報告, pp.55-60, 1997.
- [2] 平松尚子, 八村広三郎: マルチモーダルコミュニケーションのための身体運動の記述と表示, 電気関係学会関西支部連合大会研究報告, S51, 1997.
- [3] 吉田康行, 松岡洋介, 八村広三郎: 舞踊譜 Labanotation に基づく身体運動の処理, 情報処理学会人文科学とコンピュータ学会報告, 38-6, pp.61-68, 1998.
- [4] S.Lu, S.Igi, H. Matsuo, Y. Nagashima: Towards a Dialogue System Based on Recognition and Synthesis of Japanese Sign Language, Proc. of Gesture Workshop '97, Bielefeld, German, pp.259-271, 1997
- [5] 安達久博: 手話動作特徴に基づく手話単語の分類方法, 第 5 回国語研究所国際シンポジウム論文 p.136-143, 1997.
- [6] 鎌田一雄: 福祉と感性コミュニケーション, 文部省科学研究費補助金重点領域研究, 感性情報処理の情報学・心理学研究, 感性コミュニケーション 9 3 年度第一回研究会資料, pp.5-14, 1993.
- [7] W.C.Stokoe, D.D.Caterline and C.Gronberg: A Dictionary of American Sign Language Of Linguistic Principles, Gallaudet College Press, 1965.
- [8] L.A.Friedman Ed., On the Other Hand - New Perspectives on American Sign Language, Academic Press, 1977.
- [9] S.K.Liddell and R.E.Johnson: American Sign Language: The Phonological Base, Sign Language Studies, 64, pp.195-277, 1989.
- [10] レオナード・G・レーン: やさしいアメリカ手話, リヨン社, 1991
- [11] 手話コミュニケーション研究会: 新・手話辞典, 中央法規, 1992
- [12] Rikk Carey, Gavin Bell, The Annotated VRML2.0 Reference Manual, Addison Wesley Developers Press, 1997.
- [13] Christian Babski: Virtual Humanoids In Vrml, <http://ligwww.epfl.ch/babski/vrml.html>