

多用途型日本手話言語データベース：KoSign Ver.5(コサイン)

長嶋祐二¹・渡辺桂子¹・市川 薫³・原 大介²・堀内靖雄³・酒向慎司⁴
(¹工学院大学・²豊田工業大学・³千葉大学・⁴名古屋工業大学)

KoSign とは、手話の学際的な研究の推進を図るため、研究者が共通に利用できる汎用的な日本手話のデータベースです。

KoSign Ver.5は、約6,300の手話単語と10テーマの対話を、高精細・高精度で収録した3次元動作と映像で構成されたデータセットです。

KoSign が様々な研究分野で活用されることで、その成果が手話を母語とする人々に向けた新たな情報保障技術の開発などでQOLの向上に寄与していくことを期待しています。

KoSign と周辺ツールの作成は、科学研究費補助金 基盤研究(S)17H06114の補助によって行われ、NII IDRから提供されています。

KoSignの作成動機

手話は言語であるにもかかわらず、音声言語と比べて言語学や工学を含む関連諸分野での研究が進んでいない。

その理由は、言語学や工学など様々な分野の研究者が共通に利用できる汎用的なデータベースが存在しないことが原因の一つと考える。

そこで、研究者が必要とする多様なデータ形式を備え、手話データに求められる信頼性や精度を確保することで、質と量において最高水準のデータベースを目指した。

データの収録方法

どの程度の空間・時間分解能のデータが必要かも分析するため、可能な最高水準の精度の手話動作収録手法とデータ形式を採用し、異種間データの同期収録とした。

(a)3次元動作データ：光学式モーションキャプチャ

手指ならびに非手指動作の詳細な解析のため、手と顔は高精度に、そして全身を合計112マーカーで計測した。

空間分解能：0.5mm、時間分解能：120fps、BVHとC3D形式

(b)映像データ：4KとHDカメラ

各研究機関で利用される手話のデータ形式は、動画映像が多いため、3台構成で多視点画像計測を可能とした。

時間分解能：60fps

(c)深度(距離)データ：Kinect2

安価でかつ比較的高精度に距離を計測可能なToF(Time of Flight)方式センサーで赤外線映像と通常映像とを同時記録した。

時間分解能：最大30fps

(a)~(c)のデータを同期収録している様子を図1に示す。

語彙の選定と言語資料提供者

「日本語-手話辞典」、「日本語の語彙特性」第9巻、「日本語話し言葉コーパス」などを基に、出現頻度の高い語彙、手話で日常的に良く用いられるものを中心に語彙選定を行った。

言語資料提供者の手話者は、手話母語者の家系の手話母語者の男女各1名とした。

NIIからの提供するデータとツール

NII IDRから提供するデータなどは、下記の4種類である。

(1)男女各1名で収録した(a),(b),(c)のデータによる全6,359単語。

(2)男女各1名で収録した10対話の(a),(b),(c)のデータ。

(3)対話などの解析のためのアノテーションツールMATをソースコード付で(自作DLL組み込み可能)。

図2に、アノテーション領域と利用可能なデータ例を示す。

(4)全10対話をMATを用いて単語(表1に単語数)、非手指動作を抽出したアノテーションした結果。

KoSignを用いて

共通のデータを利用することで、研究成果の評価や検証が統一基準で可能となり、信頼性が確保できる。KoSignには、日本手話の主要な語彙が多数収録されており、その動作を3D CGによって立体的に見ることができ、書籍や映像とは異なる新しい形態素辞書のような電子化辞書の役割も期待できる。

手話での空間の使い方を可視化・解析することで、未だ確定していない手話の音素の確定、手話の適格性検証のための手指動作の運動規則導出などが期待できる。MATのC3Dデータからの3次元座標抽出で得られた両手首位置の軌跡を描画した例を図3と図4に示す。

KoSignは身体動作の高精度な3Dデータとして、言語学、人間工学、情報学など幅広い分野で活用されることを期待する。



図1: モーションキャプチャスタジオにおける収録の様子

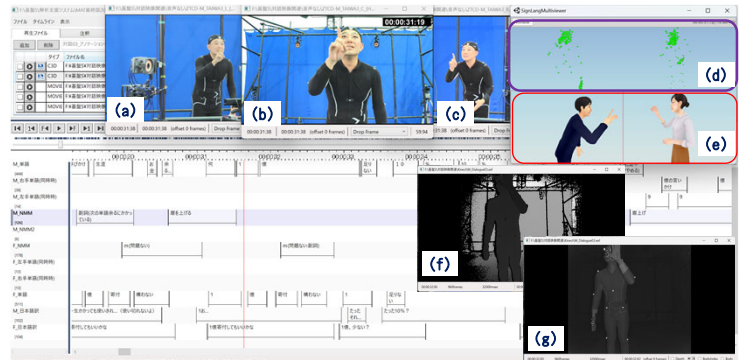


図2: MATを用いた対話3の解析画面と扱えるKoSignのデータの一例
(a)4Kカメラ(左側), (b)4Kカメラ(正面), (c)4Kカメラ(右側), (d)C3D
(e)BVHデータ, (f)Depth(Kinect2), (g)Infrared (Kinect2)
((a)~(g)は同期した同一時間タイミング)

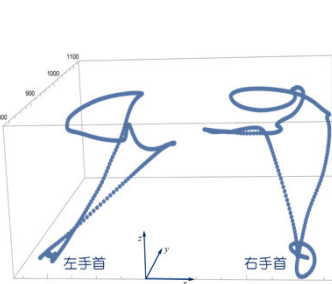


図3: 手話単語{衆議院議員}の
手首の3次元軌跡(C3D)

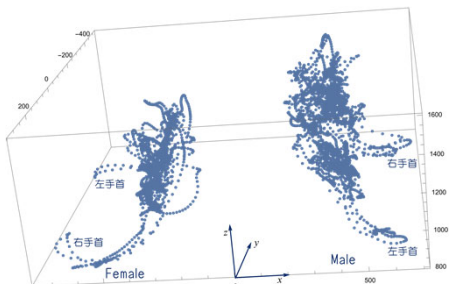


図4: 対話5の話者の1分間の手首の3次元軌跡
120fpsのC3Dデータを30fpsごとに描画

表1: 各対話に含まれる単語数

	Male	Female	合計
対話2	840	579	1,419
対話3	444	511	955
対話4	475	199	674
対話5	110	129	239
対話6	581	582	1,163
対話7	487	656	1,143
対話8	360	522	882
対話9	278	486	764
対話10	583	606	1,189
対話11	483	335	818
合計	4,641	4,605	9,246