

P4

日本手話の多次元データベースの構築

長嶋祐二¹、原大介²、酒向慎司³、渡辺桂子¹、堀内靖雄⁴、菊澤律子⁵、加藤直人⁶、市川熹¹

(¹工学院大学、²豊田工業大学、³名古屋工業大学、⁴千葉大学、⁵国立民族学博物館、⁶NHK 放送技術研究所 [日本])

要旨

ろう者が用いる手話言語は、音声言語からは独立した自然な対話言語である。国立情報学研究所において音声資源コンソーシアムが設置され、日本語の話し言葉を対象とする研究は、工学と言語学の分野にまたがって大きく進展してきた。一方、工学と言語学の分野において、手話言語の研究は遅れている。その理由のひとつとして、研究者が利用できるデータベースの無いことが挙げられる。

そこで、本研究は日本手話の汎用データベースを構築する方法を論じる。我々の目標は、工学、認知科学、言語学その他の多様な分野で利用可能な学際的データベースを構築することである。この計画では、様々な分野の研究者に利用可能なデータベースを構築するために、4つの作業を必要とする。

第一に、言語学と工学の分野での利用に適した日本手話データを収集する必要がある。この収集において、手話単語の種類や手話文の種類を考慮し、そして情報提供者を適切に選定しなければならない。

第二に、言語学や工学といった学術分野で十分に利用を可能にするためには、入力形式、空間的・時間的解像度、電子データのフォーマット、そして媒体への保存方法について適切かを論じる。現時点では、3次元動作データ、多視点映像、および深度映像を含めることを考えている。3次元動作データは、世界最高精度の光学式モーションキャプチャー技術を用いて収録する。

第三に、様々な入力形式のデータを同一のデータとして研究に利用可能なデータベースとするかの方法を検討しなければならない。

最後に、データベースが効率的に機能するためには、3種類の異なる形式のデータを有機的に結びつけることのできる新しいタグづけ法を開発することも必要である。

2017年時点でのスペックは表1の通りである。

表1：技術スペック

メディアの種類	内容
モーションキャプチャー	
カメラの種類	VICON T160(V16)
フレームレート	120fps
解像度	4,704×3,456 ピクセル
カメラの数	42
再帰反射マーカースの数	112
映像	
カメラの種類	SONY PWX-X200
解像度	1920×1080 ピクセル
フォーマット	MPEG-4 AVC/H.264
フレームレート	60fps
カメラの数	3
スーパースロー映像	
カメラの種類	SONY PWX-Z90
解像度	1920×1080 ピクセル
フォーマット	XAVC
フレームレート	120fps
カメラの数	1
深度映像カメラ	
カメラの種類	Kinect2
解像度	512×424 ドット
測定できる深度	0.5～4.5m
水平視野	70 度
垂直視野	60 度
フレームレート	30fps
カメラの数	1