

インタラクション・コーパスの閲覧・分析環境の開発

来嶋 宏幸 角 康之 西田 豊明
京都大学

1. はじめに

近年、会話分析や社会的なインタラクションを研究するにあたって、会話シーンの音声・ビデオのコーパスを利用したアプローチが広がってきている。しかし、従来の研究の多くは会話参加者や状況が固定的なシーンのみを対象としてきた。それに対し、我々の研究グループは、より開放的な状況で会話の参加者が動的に変化するような状況を対象としたインタラクションのダイナミクスの理解に興味がある。従来扱われてきたデータの分析・ラベリングには Anvil[1]の様なビデオ・アノテーションツールが利用されて来た。しかしながら、我々が扱うようなデータの分析では単一の映像データを分析するのではなく多視点の同期した映像データを分析するための環境が必要となる。そこで我々は自分達の扱うデータをより効率的に分析するためのシステムである iCorpusStudio を試作した。本稿ではこの iCorpusStudio を用いたインタラクション・コーパスの分析の手法を提案する。

2. インタラクション・コーパスとその分析

我々の研究グループでは、複数のセンサ群を協調的に用いて複数人のインタラクションを記録する方法論の開発を進めてきた[2]。文献[2]では、展示見学会に来訪した人々のインタラクションデータを大量に蓄積し、それに構造を与えた物をインタラクション・コーパスと呼んでいる。インタラクション・コーパス構築の試みとして、赤外線 ID センサを用いて LED タグ固有の ID を識別し、それを用いて実時間で自動的にビデオデータにインデクスを与え、それを解釈する手法を提案した。具体的には、得られたインデクス付のデータがあるルールに沿って生のデータから段階的に解釈することによって、抽象度の高いインタラクションの解釈を行う試みがなされた。そのようにして

Development of a System for Analysing and Annotating Interaction Corpora: Hiroyuki Kijima(kijima@ii.ist.i.kyoto-u.ac.jp), Yasuyuki Sumi, Toyooki Nishida; Kyoto University

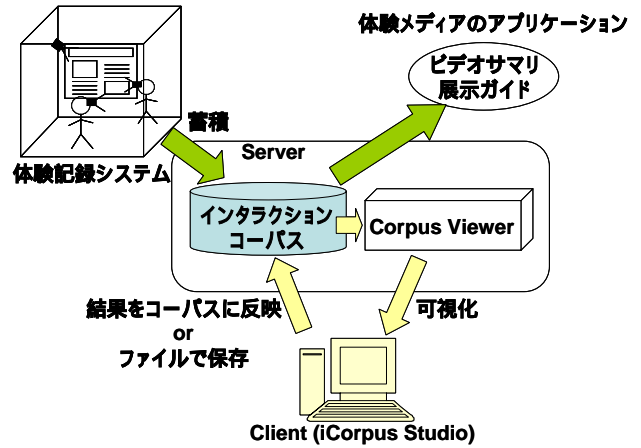


図1 システム全体の構成

得られたインタラクション・コーパスを核に据えた体験メディア(ビデオサマリ、展示ガイド)の構築を進めている(図1)。また、大量のデータを分析し、人のインタラクション・パターンを理解することもインタラクション・コーパス構築の動機の一つである。

インタラクション・コーパスの特徴として、開放的な状況の中に複数のシーンが混在しており、その中の各シーンが複数視点で記録されているということが挙げられる。しかし、複数のシーンの中から注目すべきシーンを特定することや、たくさんある視点データのどれが今注目しているシーンの構成員であるかは自明ではない。我々の研究グループではこれらの問題を解決するために、コーパスを可視化し分析するためのツール(Corpus Viewer)を開発してきた[3]。このツールでは利用者は注目すべきシーンを容易に見つけることが出来、その場面を構成している複数の人間や環境のビデオデータを同期して表示することを可能としている。しかしながら、分析作業そのものには従来のツールを用いてきたため、複数のシーンを構成するような時空間的に大量のデータの分析には不十分であった。

また、分析する動機の一つとしてインタラクション解釈の誤りの検証が挙げられる。解釈の誤りは得られたデータ自体の誤りやインタラクション

解釈のルールに起因する物がある。これらを修正・検証するための環境も必要とされている。

今回我々が試作した iCorpusStudio では、前述の Corpus Viewer と組み合わせて利用することによって複数の同期したビデオデータを見るだけでなく、赤外線 ID センサや発話検出、心拍数などの生体データといった大量のコーパスデータを入力としてアノテーション候補を提示し、映像データを見ながらそれらを手軽に分析・編集し、結果を複数利用者間で共有出来る環境を提供する。またコーパスデータそのものの修正を共有することによりインタラクション解釈の誤りを減らすことが出来る。

3. iCorpusStudio

3.1 システム構成

システム全体の構成図を図1に示す。体験記録システムでの人や環境側のセンサやカメラのデータは実時間でサーバに蓄積される。本システムでは可視化されたコーパスデータから、分析に必要な範囲をユーザが選択することによって必要なファイルをダウンロードし作業を行う。作業終了後にアノテーションデータをデータベースに登録することで、システムの他の利用者と結果の共有を可能とした。また、クライアント側に XML ファイルを出力する機能も実装した。

3.2 動作

本システムの動作画面を図2に示す。システムの上部には分析対象となるシーンをとらえる複数視点カメラの同期した映像が表示される。下部のパネルではアノテーションデータが横軸に時間を取って表示されている。アノテーションの時間軸は映像の時間と連動しており、見たい時間帯をアノテーション部分から選択することが可能である。逆に動画を再生するとパネル上の時間を示すラインも同期して移動する。下部左側にはアノテーションデータの追加や削除、編集といった作業が行える環境を提供している。また、限られた画面を有効に利用するために、作業中でないアノテーションの行を最小化するような機能を実装した。下部のパネルでは、コーパスデータを入力としたアノテーションデータの候補と、ユーザが定義したアノテーションを表示している。蓄積されたコーパスデータをアノテーションの候補として表示することで、自動で取得されたデータを実際の映像

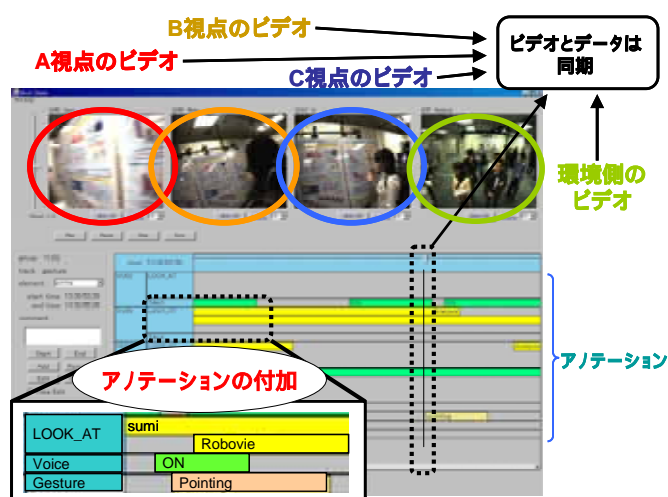


図2 システムの動作画面

を見て比較検討することが可能となった。アノテーションの定義については Anvil などで見られるような、ユーザが XML ファイルである程度自由に定義しそれを読み込むという形で実装した。

4. おわりに

本稿では、インタラクション・コーパスの分析を支援するシステムを試作した。本システムを用いることで多視点の映像データを協調的に閲覧しながらコーパスデータを分析することが可能となった。またそれによってインタラクション解釈の誤りを修正出来ることが示された。今後、本システムをインタラクションの解釈やその結果を利用したアプリケーションのプロトタイプ環境に拡張したいと考えている。

謝辞

本稿で紹介したコーパス環境導入にご協力頂いた ATR の岩澤昭一郎氏、山本哲史氏に感謝する。

参考文献

- [1] Michael Kipp: Anvil - A Generic Annotation Tool for Multimodal Dialogue, In Proceedings of the 7th European Conference on Speech Communication and Technology (Eurospeech), pp. 1367-1370, Aalborg, September 2001.
- [2] 角康之、伊藤禎宣、松口哲也、Sidney Fels、間瀬健二: 協調的なインタラクションの記録と解釈; 情報処理学会論文誌、Vol.44、No.11、pp.2628-2637、2003
- [3] 角康之、岩澤昭一郎、間瀬健二: インタラクション・コーパス分析ツールの試作; 情処研報、ヒューマンインタフェース、Vol.HI104-7、2003.