

いつ誰が何を話したか？を即座に分析 —複数人会話のリアルタイムモニタリング&ブラウジング—

背景・課題：人間同士の自然なコミュニケーションを認識し、状況を理解することでコミュニケーションを分析・支援する技術が求められています。しかし、実環境での複数人の会話において、誰が誰に話しているかといった会話の状況やその発言内容を高い精度で認識することは技術的に極めて難しい課題です。

アプローチ：複数人の会話の状況をカメラやマイクでセンシングし、「いつ、誰が、誰に、何を、どのように話したか」を高速かつ高精度に分析するためのモデリングや学習・認識技術を研究しています。また、種々の分析結果を即座に統合して可視化するモニタリングおよびブラウジングの方法も研究しています。

到達点：これまで「いつ、誰が、誰に話したか」のリアルタイム分析を実現してきました。今回は極めて難しい、実環境の複数人会話リアルタイム音声認識にも挑戦し「何を話したか」までを含む、会話状況のより豊富な情報を即座に提示するデモをお見せします。



会話シーンを
即座に分析し、
モニタリング

いつ、誰が、誰に、何を、
どのように、話したか？

NTT CS研が生み出した世界トップレベルの
オンラインメディア処理技術を統合した
リアルタイム会話分析システム



リアルタイム会話ブラウザ



関連文献

- ・大塚, 荒木, 石塚, 藤本, 大和:「複数人会話シーン分析に向けた実時間マルチモーダルシステムの構築」, 信学技法MVE2008-68, pp. 55-62, 2008
- ・大塚, 荒木:「会話シーン分析のための音声映像技術」, NTT技術ジャーナル, May. 2009, (特集:コミュニケーション科学の融合研究における新潮流)
- ・堀, 須藤, 塚田, 中村:「世界メディアブラウザ」, NTT技術ジャーナル, May. 2009, (特集:コミュニケーション科学の融合研究における新潮流)

連絡先

堀 貴明 (Takaaki Hori), 大塚和弘 (Kazuhiro Otsuka)

メディア情報研究部 信号処理研究グループ, メディア認識グループ