AMI and AMIDA

Meeting browsers and remote meeting support

Jean Carletta, for the AMI Consortium

http://www.amiproject.org/

Overview

The AMI Consortium develops new technologies to aid groups that hold meetings. The AMI project concentrated on ways of using archives of face-to-face meetings, and AMIDA will contribute aids for people who need to attend a meeting, but can't be together. Edinburgh's main contributions are in project coordination, data collection, speech processing, and language technology.

Our goal is to help users find material in recorded meetings quickly, whether that's through *meeting browsers* that use indexing to help them find their way around a recording, *compressed playback*, or *automatic summarization*. As we move towards *real-time processing*, we will be able to build systems that, for instance, let you catch up on what's happened so far if you're late for a meeting.

Multi-modal A/V Processing

Our first step in providing these technologies is audio-visual processing to find out:

- What has been said during the meeting?
- What events and keywords occur in the meeting?
- Who and where are the persons in the meeting?
- Who in the meeting is acting or speaking?
- How do people act in the meeting?
- What are the participants' emotions in the meeting?
- Where or what is the focus of attention in meetings?

Semantic Content

Given the results of our A/V processing, we need to make sense of a meeting's contents. We are concentrating on constructing good *summaries* of what happened as well as discerning what *topics* the meeting discussed, what *information* was expressed, what *decisions* were made and why — tying all of this back to evidence in the recordings themselves so users of our technologies can make up their own minds.

The AMI Meeting Corpus

To give us the data we need to do this work, we have collected the AMI Meeting Corpus, a multi-modal data set consisting of 100 hours of meeting recordings. Most of the recordings are of groups playing roles in a remote control design team, with other meetings thrown in to help us generalize our work.

To collect the data, we set up three instrumented meeting rooms to collect many different synchronized signals. All 100 hours of the corpus has been transcribed in detail. In addition, much of the corpus has been annotated by hand for:

- dialogue acts, with addressee and some relations
- labelled, hierarchically decomposed topic segments
- abstractive summaries linked to the data extracts that support them
- named entities
- who or what each person is looking at
- some simple head and hand gestures

The annotations are both to help analysts understand the data and as input for *machine learning* to create them for new meetings automatically. This data set is unlike any other, bringing together multi-modality, transcription, and many different kinds of annotation for the first time in one framework that represents how they relate to each other. It was released under a Science Commons "Share-Alike" license in June 2006.

Evaluation

The AMI Meeting Corpus is already being used for international evaluations such as NIST RT07 for speech recongition, CLEAR for A/V tracking, and CLEF for question answering. We carry out consortium-wide evaluations of components that produce automatic versions of all our hand annotations. However, component evaluation is not enough — we need to know that our technologies help. We are currently testing whether our browsers help users answer questions about what happened at a meeting, and whether groups that need to know what happened at past meetings in order to change a design get better results if they use them.

Beam-forming Demo

Recognizing speech automatically in meetings is hard because the participants often talk over one another. A microphone array can make the recognition work better by taking the signals from a number of different microphones and using them to 'listen' to one speaker in a particular place. We show the improvement which can be gained by using a microphone array instead of a single microphone.

Summarization Demo

One way of browsing a recorded meeting quickly is by using a summarizer to extract the important parts and either reading that or using it to navigate through the recording. See what our summarizer chooses as it compresses a meeting into increasingly shorter extracts.