Surveys of Multimodal Data Resources, Annotation Schemes and Tools

Malene Wegener Knudsen, Laila Dybkjær, and Niels Ole Bernsen
Natural Interactive Systems Laboratory, Denmark

This paper presents groundwork on data resources (corpora), corpus annotation schemes and annotation tools which is being carried out in the European Natural Interaction and MultiModality (NIMM) working group in the EU/US ISLE (International Standards for Language Engineering) project. The NIMM working group began its work in early 2000 and has now produced three public (http://isle.nis.sdu.dk) surveys. These reports address NIMM data resources, annotation schemes, and annotation tools, respectively.

The report on data resources surveys 31 facial resources and 19 gesture resources. In several cases speech is included together with facial expression and/or gesture.

The report on corpus annotation schemes includes 8 descriptions of annotation schemes for facial expression and speech and 15 descriptions of schemes for gesture and speech.

The report on corpus coding tools describes 12 annotation tools most of which support speech annotation combined with gesture annotation, facial expression annotation, or both.

The first step towards producing the surveys was to establish selection criteria for which data resources, annotation schemes or tools to include in order to make the surveys as useful and valuable as possible to the intended readership. Among the criteria were: public accessibility (free or for a fee), in use by a fair number of people or of potential relevance to many people, and, whenever possible, evaluation results on usability provided by people from academia or industry who actually used the resource/coding scheme/tool.

Secondly, the European NIMM partners produced a common description template for each report, taking into account the contents and structure of existing templates, such as the LDC and ELRA templates for corpus description. The templates prescribe a rather detailed description of each resource/coding scheme/tool, including general information, such as usage, examples, references to additional information, and reference to contact person. The common templates help making descriptions easily comparable and ensure availability of important items of information throughout.

Thirdly, to ensure descriptive completeness and accuracy, the descriptions were made in close contact with the developer(s), whenever possible. Similarly, most, survey descriptions have been validated by the developers.

The survey process also enabled the group to gather information on user profiles, user needs and best practice, and address those issues in the reports. Potential NIMM users work with annotation in different areas, from different perspectives, and with an unlimited number of different annotation purposes. User profiles include, e.g., people working in spoken and written dialogue systems development and evaluation, multimodal systems, embodied agents, face and gesture annotation, and research in prosody, linguistics, psychology, anthropology, human behaviour and human factors. There is a felt, and growing, need for NIMM resources, coding schemes, and tools in academia and industry. For instance, the game industry and the interface agents community need resources for developing natural and human-like characters, and educational institutions need resources for learning and teaching purposes. Best practice and standards are fairly developed in some areas but more or less absent in others. For instance, MPEG-4 and FACS set the current best practice standard for facial expression annotation.

Malene Wegener Knudsen (mwk@nis.sdu.dk) Tel.: +45 63 15 73 04
Laila Dybkjær (laila@nis.sdu.dk) Tel.: +45 65503553
Niels Ole Bernsen (nob@nis.sdu.dk) Tel.: +45 65503544

Natural Interactive Systems Laboratory,
University of Southern Denmark
Main Campus: Odense University
Forskerparken 10, DK-5230 Odense M, Denmark