

Multimodal Resources for Human-Robot Communication Modelling

Stavroula–Evita Fotinea, Eleni Efthimiou, Maria Koutsombogera, Athanasia–Lida Dimou,
Theodore Goulas, Kyriaki Vasilaki

Institute for Language and Speech Processing - ILSP/ATHENA RC
Artemidos 6 & Epidavrou, Maroussi, 15125 Athens, Greece
{evita, eleni_e, mkouts, eleni_e, ndimou, tgoulas}@ilsp.gr, kiki-vas@hotmail.com

Abstract

This paper reports on work related to the modelling of Human-Robot Communication on the basis of multimodal and multisensory human behaviour analysis. A primary focus in this framework of analysis is the definition of semantics of human actions in interaction, their capture and their representation in terms of behavioural patterns that, in turn, feed a multimodal human-robot communication system. Semantic analysis encompasses both oral and sign languages, as well as both verbal and non-verbal communicative signals to achieve an effective, natural interaction between elderly users with slight walking and cognitive inability and an assistive robotic platform.

Keywords: multisensory data acquisition, multimodal semantics, multimodal annotation scheme, sign language, multimodal HRI model extraction, multimodal human-robot communication, natural HRI

1. Introduction

Cognitive and robotic architectures may be able to provide advanced interactive capabilities with humans and influence the usability and functionality of a resulting system, contributing to its quality of services. Such systems may not only integrate multiple advanced cognitive abilities, but also employ methods that are extendible to various other robotic and non-robotic applications required in assisting humans with mobility disabilities. In this work, the driving concept envisions cognitive robotic assistants that act (a) proactively by realizing an autonomous and context-specific monitoring of human activities and by subsequently reasoning on meaningful user behavioural patterns, as well as (b) adaptively and interactively, by analyzing multi-sensory and physiological signals related to gait and postural stability, and by performing adaptive compliance control for optimal physical support and active fall prevention.

A primary focus in this work has been the analysis of available data and the definition of the semantics of human actions in interaction between humans. At a second stage, their capture and representation in terms of behavioural patterns have been exploited to feed a multimodal human-robot communication system. Semantic analysis, in this sense, encompasses both oral and sign languages, as well as both verbal and non-verbal communicative signals. Towards defining the system's multimodal HRI model, we discuss the acquisition of the multimodal sensory corpus which served as a primary source of data retrieval, analysis and testing of the developed mobility assistive robot prototypes. We further focus on the definition of the representation scheme of audio-gestural signals which provided the ground truth dataset for the training and evaluation of visual and audio recognition algorithms, and ensured the usability of the acquired dataset in defining the adopted multimodal human-robot communication model.

Acquisition, annotation and analysis of the here discussed multimodal-multisensory dataset as well as work on definition of the HRI communication model based on this specific dataset have taken place in the

framework of the MOBOT project¹, which aims at developing intelligent active mobility assistance robots for indoor environments that provide user-centred, context-adaptive and natural support.

Specifically, a multimodal action recognition system able to monitor, analyze and predict user actions with a high level of accuracy and detail is currently under development. While a major research effort focuses on enhancing computer vision techniques with modalities such as range sensor images, haptic information as well as command-level speech and gesture recognition, data-driven multimodal human behaviour analysis has been conducted to extract behavioural patterns of elderly people in their everyday life communication with other humans (i.e. carers) and also in their interaction with a robotic rollator. Findings are constantly enriching a multimodal human-robot communication system involving both verbal and nonverbal communication which is conceptually and systemically synthesized into mobility assistance models taking into consideration safety critical requirements.

Pursuing the definition of a model of multimodal human-robot-interaction (HRI) as close to natural human behaviour as possible, in the following sections we discuss the design of the multimodal MOBOT corpus and the description of the annotation scheme applied to the project's acquired dataset of verbal and gestural communicative signals (Fotinea et al., 2014).

2. Dataset Acquisition

Given the MOBOT characteristics of human action recognition and context-aware robot control, the data acquisition method had the goal to promote the natural interaction between (elderly) users and mobility aids along with the assistance or not, in some cases, of carers. To obtain the necessary multi-sensory data required for machine learning purposes and for the definition of the project's human-robot interaction model, several recording scenarios, sessions and measurements were implemented and tested so that a maximum set of actions and movements would be well represented in the recordings and hence provide significant input from

¹www.mobot-project.eu

different modalities.

In the recordings, elderly human individuals of varying age, gender, motor and cognitive abilities performed a variety of assistance requiring tasks in dialogues involving human carers and a passive rollator².

The recorded scenarios were based on a use case list that included actions that a typical end user may need to perform in real life situations and that was adapted to the recording environment. Critical situations or potential barriers due to insecurity, cognitive or physical impairment, as well as the risk of falling were taken into account. The envisioned goal was to incite a variety of actions, movements, gestures and functions within the restricted recording setting, in order to obtain sufficient representative data to set the foundation for the technological development of the assistive device. Previous experience on the design of multimodal corpora [Matthes et al., 2012], and multisensory acquisition data [Wallraven et al., 2011] was proven a valuable background asset, in both scientific and methodological aspects.

The recruitment strategy included the selection of participants in the recordings following an assessment at various levels to know whether they meet the inclusion criteria (e.g. patient's charts, personal interview, consent, testing in the conditions of the experiment). This enabled the identification of limiting factors and helped tune more adequately the test levels to adjust with the patients' ability to endure testing. The experimentation followed structural and technical specifications which also account for limitations and possible inconveniences. In this respect, the physical setup took into consideration the geography of the space (i.e. large enough recording rooms, no obstacles that distract the subjects), lighting conditions (e.g. not subject to weather conditions such as sunlight, rather use of artificial lights or respective solutions to overcome reflections), and location of the equipment (i.e. tripods installation supporting all kinds of high sensitivity sensors).

Inconveniences and limitations in the setting and the experimental process became more obvious in practice and/or after several iterations. Hence, the experimental design had to allow a certain degree of flexibility to avoid complexity and make resulting data as profitable as possible. Aspects of tasks within a scenario (i.e. number and order of activities, duration of tasks, etc.) were at times modified and conventions were introduced to ensure the completion of tasks from all participants and the comparability of the recordings.

The recordings consisted in six different scenarios including gestures and other activities close to real life situations such as obstacle avoidance, interaction with other persons, simple everyday life operations (open/close a door, switch on/off the light etc.) that needed to be reproduced by the aged impaired participants during the recordings and were captured by a wide range of sensors. Each scenario covered a specific set of actions/activities enabling the acquisition of respective data, i.e. walking for gait modelling or audio and gestural commands for HRI modelling. Each scenario was realized in three variants, which actually represent the following different types of interaction the patient may have with the rollator

and/or carer:

- a. The informant being supported by the rollator;
- b. The rollator being in following mode;
- c. The informant being supported by a human carer.

The dataset was acquired by means of a sensorized passive rollator comprising multimodal input from laser range finder sensors, force/torque sensors, RGB and RGB-D cameras and an 8-microphone MEMS³ array mounted on the horizontal bar of the MOBOT rollator in a linear configuration (with a 4cm uniform spacing) in front of the user. For the recording of the scenarios four High Definition (HD) cameras with sensitive sensors were used. Three of them were mounted on tripods to record the informants' full bodies in a set-up which could recover possible occlusions covering optical gaps and providing further information of motion, posture, and details of manoeuvring and possible patient-carer human interaction. The fourth HD camera (GoPro) was mounted on the passive rollator, on top of the upper Kinect camera to record closely and at a constant distance the patient's torso, arms and all his/her movements, in some cases including also the head (cf. Fig.1).

In order to end up with a corpus of properly annotated video data, the acquired video files (from HD cameras, Kinect, GoPro) were rendered, and the Kinect raw files have been synchronised with the rest of the visual data, providing a synchronisation scheme between external HD data and the ROS (Robot Operating System)⁴ bag related multi-modal multi-sensorial data. The streams were rendered independently and together in a single stream "picture in picture" (PiP) to provide all related information accumulated (cf. Fig. 2).

The data acquisition process involved a total number of 18 patient-subjects (12 female, 6 male), their age ranging 74-87 years old. Participants' metadata consist of information about gender, age, height, weight and knee height of each subject, as well as their cognitive and mobility score as determined with the aid of the diagnostic tool MMSE (Folstein et al., 1975), and their subsequent classification into a cognitive and mobility category (Woodford & George, 2007).

2.1 The MOBOT Dataset: Some Qualitative and Quantitative details

As already mentioned above, the MOBOT database was acquired by means of a sensorised passive rollator (Fig.1) comprising multimodal input from (i) laser range finder sensors, (ii) force/torque sensors, (iii) RGB and RGB-D cameras and (iv) microphones. In addition, a motion capture system was used to record human limb movements as well as the rollator and subject's absolute positions in space, which makes it one of the very few available multimodal-multisensorial resources of relatively rich content of activities performed by humans in interpersonal interactions, and by humans in interaction with a device.

Synchronization of multimodal data streams was achieved by recording all the data to a ROS-Robot Operating System bag. For this purpose ROS nodes were programmed for the two laser range finders, the two

²Acquisition of the MOBOT multimodal sensory corpora took place in the rehabilitation centre Agaplesion Bethanien Hospital/ Geriatric Centre at the University of Heidelberg.

³ STMicroelectronics:http://www.st.com/web/catalog/sense_power/FM89.

⁴www.ros.org.

Kinects and the microphone array. The amount of ROS bag data collected altogether reached 1.3 TB. The amount of data collected from the two HD cameras and the GoPro camera mounted on top of the passive rollator is approximately 250 GB (437 files), whereas a rough estimation of video data duration is 10 hours for global video duration and 8 hours for useful video cleaned from pieces of preparations time, meaningless poses etc. A detailed account of the devices used and the procedures followed during the MOBOT dataset acquisition and its post-processing is provided in Fotinea et al. (2014).

The devices used for data capture and the richness of activities included in the recorded scenarios make the MOBOT dataset one of the most up-to-date such resources worldwide among other recently acquired datasets.

In the field of human-action recognition several datasets with rich sets of activities, complex environments representing real-world scenarios have been published. Such datasets that combine video and mocap systems in a systematic way by collecting synchronized and calibrated data include the HumanEva I and II datasets (Sigal et al., 2010). The creation of HumanEva datasets were motivated mainly by the need for having ground truth that can be used for quantitative evaluation and comparison of both 2D and 3D pose estimation and tracking algorithms. Although the HumanEva datasets have been extensively used in establishing the state-of-the art in human action recognition, their application areas remain limited to evaluation of 2D and 3D motion and pose estimation based on video and mocap data only.

There are a number of other multimodal datasets that enhance the standard mocap-video data with additional modalities, such as magnetic sensors or microphones. The TUM Kitchen Dataset (Tenorth et al., 2009), which consists of activities in a kitchen setting (i.e., subjects setting a table in different ways), for example includes also RFID tag and magnetic sensor readings in addition to the multi-view video and mocap data. Similarly, the CMU Multimodal Activity (CMU-MMAC) Dataset (De La Torre et al., 2009) contains multimodal measures captured from subjects performing tasks such as meal preparation and cooking. The set of modalities utilized in

this dataset is rather comprehensive, consisting of video, audio, mocap, internal measurement units (i.e., accelerometers, gyroscopes and magnetometers) and wearable devices (i.e., BodyMedia and eWatch). These two datasets are the first examples of publicly available multimodal datasets with a rich selection of various modalities. Finally, the Berkeley Multimodal Human Action Database (MHAD) is currently the only to-date dataset that systematically combines multiple depth cameras with multi-view video and mocap that are geometrically calibrated and temporally synchronized with other modalities such as accelerometry and sound (Ofli et al., 2013). The specific dataset consists of multi-view video, depth and color data from multiple Kinect cameras, movement dynamics from wearable accelerometers and the accurate mocap data with the skeleton information. In addition, ambient sound during the action performance was recorded and synchronized to reveal discriminative cues for human motion analysis.

2.2 The Audio-Gestural Commands Dataset

In addition to natural audio-gestural interaction in all action-based scenarios, a specific scenario was designed to gather information of isolated gestural and verbal commands as the ground truth for the definition of the system's human-robot communication model. This set entailed 20 isolated gestural and verbal commands. Each command was performed by every informant in four repetitions; it was also necessary to acquire the same audio and gestural commands within action context, i.e. during the action tasks in the rest of data acquisition scenarios. Thus, informants were instructed to incorporate the predefined audio and gestural commands also within scenarios, while performing the action tasks foreseen. In all cases, the carer was the one who introduced the commands the patient had to perform, also providing additional control data as against patients' actual performance. In addition, the presence of the carer established a second level of interactivity to the human-robot communication model that is of major value, as it sets the grounds for a multi-party communication model.

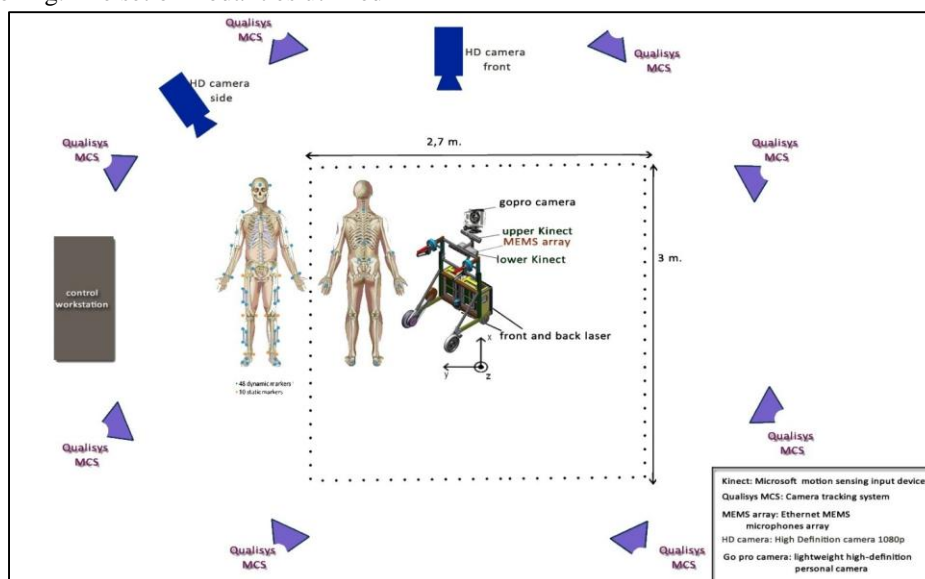


Figure 1: Diagram of the MOBOT data acquisition setting in Bethanien Geriatric Hospital

2.3 Gestural Commands: Selection Procedure

The definition of the gestural commands comprising the predefined MOBOT gesture set was inspired from sign language (SL) semantics. SLs –being articulated in the three-dimensional space by means of a multichannel articulation system– demonstrate a number of gestural semantic markings that relate to prototypical iconic representations of objects in the real world.

In general, SLs present the level of abstraction, common to all linguistic systems; however, sign representations tend to be closer to universal semantics as they portray ground truth visual input, which makes attribution of the “signifier” to the “signified” in an unambiguous and direct manner. More specifically, core semantic qualities that make the identification of items direct and unambiguous are most prevalent in a specific grammatical category within Sign Languages, the so-called Classifiers. These formations incorporate prototypical features of basic aspects of notions, a fact that makes them easily identifiable and implicitly comprehensible even by non-SL speakers. This background justified the decision to let SL universal semantic properties drive the definition of the MOBOT gestural command set, in combination with the crucial decision to have gestural commands incorporated in action performance scenarios, which required to foresee a threshold of consistency as to the acquired data on the basis of as close-to-natural as possible informants’ (re)actions.

Taking into consideration that this task should be feasible for elderly non-native SL informants, the embodiment of the proposed gestural commands was simplified as much as possible on the basis of the hypothesis that embodiment of gestures that intuitively rely on universal semantics not only facilitates data acquisition but also creates a close to natural coding convention through which a patient may easier interact with the robot in real use environments.

Hence, the definition of the 20 gestural commands of the MOBOT dataset is an amalgam of:

1. actually existing lexical signs, i.e.
 - a. **Stop** sign for the Stop command
 - b. **Where** sign for the Where am I command
2. actually existing SL classifiers named after specific commands for the purposes of the recordings, i.e.
 - a. **Door handle sign** for the “I want to go through the door command”. In SL this sign would be used as a reference to any vertical construction with a handle and any accompanied action, “I go through the door”, “I close the door” etc.
 - b. **Drinking Glass Sign** for the “I want to perform a task” command. In SL this sign would be used as a reference to any object that shares the cylindrical properties of a drinking glass and could represent anything from a plastic tube that is held in the vertical axis to a pencil holder or an actual drinking glass which is the most prominent combination.
3. lexical signs that underwent significant modification in order to facilitate as well as ensure their performance by the elderly data acquisition participants, experiencing various mobility impairments and being non familiar with any SL, i.e.

Park sign for the “Park” command. In SL the park sign is actually a classifier that underwent lexicalization. The hand formation of the sign is the same as the one proposed in this dataset (palm facing down) but the movement is significantly different. In SL the hand is placed in the natural space in front of the signer and the movement trajectory is a fast manoeuvre that ends near the chest of the speaker, imitating the manoeuvre performed by a driver when parking a car. As this is a rather complicated sign that would risk being performed with significant variation from one participant to another, it was simplified as follows: palm facing down, the hand almost touches the chest with its side. The hand moves smoothly to the front with fingers pointing out and bringing the arm to the side of the body.

In all cases in which the audio-gestural commands were performed simultaneously with other action tasks, the carer was the one who introduced the commands to the patient.

3. Semantics of Human Actions

Modelling human interactional behavior and actions allows for generalisation and generation of new behaviours and experiences in robots. Language is increasingly viewed in its interaction with perception and action, embracing a clear perspective about the interplay between language and the human conceptual system and enabling computational implementations and integration in a variety of applications. In MOBOT, emphasis is placed on embodied communication via the audiogestural channel based both in the oral and sign language systems. Creating behaviour models that perform inference on human intent, on the basis of which one can reason about and plan robot assistive actions and behaviours, constitutes a major research challenge in oral, as well in sign languages, especially in the light of recent studies, which reveal a relation between the semantic properties of embodied communication devices activated with oral language production and some central elements of SL articulation to be significantly stronger than earlier thought of (Johnston, 2013).

Regarding the Classifier constructions exploited for the definition of the gestural commands of the MOBOT dataset, the adopted deductions were the result of research carried out for the Greek Sign Language (GSL). This research was based on an extensive investigation of the use and function of Classifiers as they occur within the GSL segment of the Dicta-Sign corpus⁵, aiming to provide a solid description of the characteristics and functions of these constructions. At the same time, it investigates simultaneous constructions involving classifiers and examines all lexicalized classifier forms that are noticed in the language.

In general, SL classifiers are usually considered to be morphemes expressed by meaningful hand configurations occurring in combination with predicates that express the motion through space, a change of posture, and the location or existence of a referent. Research studies on classifiers have been carried out for most of the SLs studied to date, while different types of definition, function and categories of these constructions have been

⁵ <http://www.sign-lang.uni-hamburg.de/dicta-sign/portal/>

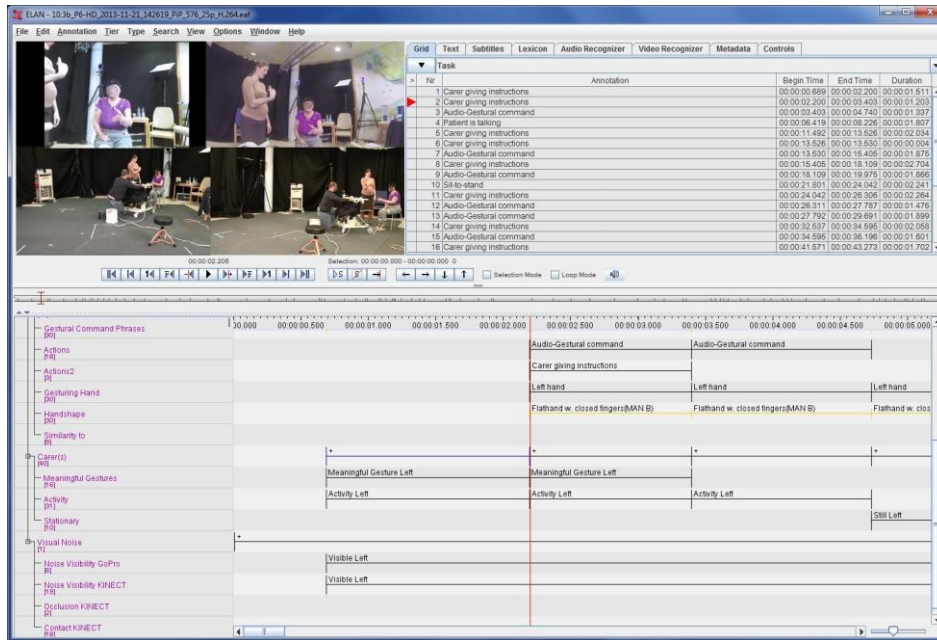


Figure 2: Synchronised data from Scenario 3, Patient 6 (PiP) and annotation sample

reported in the literature (Supalla, 1986, 1990; Engberg-Pedersen, 1993; Zwitserlood, 2012).

The Dicta-Sign corpus consists of semi-spontaneous signing data by native signers in four sign languages: German, British, French and Greek (DGS, BSL, LSF and GSL respectively), the main goals of this data collection being to ensure a high level of naturalness of the data under lab conditions as well as a high degree of parallelisability across the four sign languages. Within the GSL corpus segment, classifiers were recently extensively annotated with respect to their phonological and semantic status. Annotation resulted to an inventory of 16 different hand configurations which are analyzed as classifiers on morpho-phonological as well as on semantic grounds.

On the semantic level, GSL Classifiers were grouped according to what their handshape represents semantically and provided input for the definition of the MOBOT gestural commands set. The available index of GSL Classifier constructions has allowed for significant generalizations and assumptions regarding each semantic category, with direct relevance to embodied communication structures in oral language environments. The following semantic categorizations have provided input for the definition of the MOBOT gestural commands set:

- Entity:** GSL hand configurations which represent nouns belonging to a particular semantic class such as the class of humans, animals, vehicles, etc.
- Handling:** GSL classifiers that are being held/ or moved.
- SASSes:** GSL classifiers which express the size and shape of entities:
 - Static: hand configurations that indicate the size or shape of an entity
 - Tracing: classifiers with a movement of the hand(s) that outlines the shape or size of an entity
- Body part:** GSL classifiers representing part of the body or limbs.

e) **Lexicalized:** lexicalized or “frozen” classifiers that are found in the corpus.

f) **Predicative:** groups of classifiers that belong to the same semantic group and share all other morphemes but handshape, as possible variations of a main non-classifier predicate.

4. The MOBOT Annotation Scheme

The annotation of the audio-visual data was performed in ELAN⁶. Each channel of information was encoded into a distinct annotation tier consisting of several sub-tiers according to the granularity of information needed. The goal of the annotation scheme was to take into consideration the particularities of each scenario variants to provide usable annotation data. The annotation template comprises the following three annotation clusters, each entailing multiple tiers:

- Metadata cluster:** the source of the annotated data, i.e. scenario/variant and a standard account of the duration of the tasks performed.
- Visual Input cluster:** in-depth information about the actually performed actions and gestures of the patient/carer as well as visual noise coming from the recording environment.
- Audio Input cluster:** audio commands & uttered speech vs. non-speech, specifically: (a) timestamps for all audio commands of the patient or the carer(s), (b) translation of verbal commands in English, (c) a speech/non-speech tier of clearly uttered and fairly comprehensible speech versus all parts that contain noise, noise-like audio interventions and non-comprehensible speech.

The annotation scheme adopted provides a modular template according to the needs of each individual scenario variant, and is used for data annotation and corpus evaluation procedures.

⁶ <http://tla.mpi.nl/tools/tla-tools/elan/>, Max Planck Institute for Psycholinguistics, Nijmegen

5. Conclusion

This work reported on strategies employed for the development of multimodal and multisensory corpora targeting the enhancement of the state-of-the-art in the field of HRI, through rich sets of activities in complex environments that represent real-world uses and provide ground truth data for the related algorithms.

The process of data collection and annotation as presented in this paper provides a useful insight on the acquired multimodal data, allowing researchers to mine deeper into (a) the semantics of human actions during interaction and (b) the correlation of the annotated segments from the same as well from semantically different annotation levels, so that a more detailed representation of the human action- gesture/speech model in this specific environment can be drawn. The adaptation of “expected” human behaviour in the artificial cognitive system promotes the proactive assistance and the contextual mobility support to humans. Thus, the active involvement of end-users together with iterative trials and evaluation of the artificial cognitive system provides continuous consultations and user feedback, and thus guides research and development towards real use cases.

6. Acknowledgements

The research leading to these results has received EU funding under grant agreement n° 600796 (MOBOT project, FP7-ICT).

7. References

- De La Torre, F., Hodgins, J., Montano, J., Valcarcel, S., Forcada, R., Macey, J. (2009). Guide to the Carnegie Mellon University Multimodal Activity (CMU-MMAC) database. Technical Report CMU-RI-TR-08-22, Robotics Institute, Carnegie Mellon University.
- Engberg-Pedersen, E. (1993). Space in Danish Sign Language. *International Studies on Sign language Research and Communication of the Deaf*, vol. 19. Hamburg: Signum.
- Folstein, M.F., Folstein, S.E, McHugh, P.R. (1975). Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*, 12(3), pp. 189-198.
- Fotinea, S-E., Efthimiou, E., Dimou, A-L., Goulas, T., Karioris, P., Peer, A., Maragos, P., Tzafestas, C., Kokkinos, I., Hauer, K., Mombaur, K., Koumpouros, I., Stanzyk, B. (2014). Data Acquisition towards defining a Multimodal Interaction Model for Human-Assistive Robot Communication. In C. Stephanidis and M. Antona (Eds.), *Universal Access in Human-Computer Interaction. Aging and Assistive Environments*, LNCS 8515. Switzerland: Springer International Publishing, pp. 615-626.
- Johnston, T. (2013). Towards a comparative semiotics of pointing actions in signed and spoken languages. *Gesture*, 13(2), pp. 109-142.
- Matthes, S., Hanke, T., Regen, A., Storz, J., Wörseck, S., Efthimiou, E., Dimou, A.-L., Braffort, A., Glauert, J., Safar, E. (2012). Dicta-Sign – Building a Multilingual Sign Language Corpus. In *Proceedings of the 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon (LREC 2012)*, Istanbul, Turkey.
- Ofli, F., Chaudhry, R., Kurillo, G., Vidal, R., Bajcsy, R. (2013). MHAD: A Comprehensive Multimodal Human Action Database. In *Proceedings of Workshop on the Applications of Computer Vision*. Washington: IEEE Computer Society, pp. 53-60.
- Sigal, L., Balan, A., Black, M. (2010). HumanEva: Synchronized video and motion capture dataset and baseline algorithm for evaluation of articulated human motion. *International Journal of Computer Vision*, 87, pp. 4-27.
- Supalla, T. (1986). The classifier system in American Sign Language. In C. Craig (Ed.) *Noun Classes and Categorization: Proceedings of The Symposium on Categorization and Noun Classification*, Amsterdam: John Benjamins, pp. 181-214.
- Supalla, T. (1990). Serial verbs of motion in ASL. In S. Fischer. and P. Siple (Eds.) *Theoretical Issues in Sign Language Research*, Vol. 1 Linguistics, Chicago: Chicago University Press, pp. 127-152.
- Tenorth, M., Bandouch, J., Beetz, M. (2009). The TUM Kitchen Data Set of everyday manipulation activities for motion tracking and action recognition. In *Proceedings of the IEEE 12th International Conference on Computer Vision Workshops*, pp. 1089-1096.
- Wallraven, C., Schultze, M., Mohler, B., Vatakis, A., Pastra, K. (2011). The POETICON enacted scenario corpus - a tool for human and computational experiments on action understanding. In *Proceedings of the 9th IEEE conference on Automatic Face and Gesture Recognition*.
- Woodford, H.J. & George, J. (2007). Cognitive assessment in the elderly: a review of clinical methods. *Quarterly Journal of Medicine*, 100, pp. 469-484.
- Zwitserslood, I. (2012). Classifiers. In R. Pfau, M. Steinbach, B. Woll (Eds.), *Sign Language: an International Handbook*. Berlin: Mouton De Gruyter, pp. 158-186.