

# Advances in Intelligent Mobility Assistance Robot integrating Multimodal Sensory Processing

Xanthi S. Papageorgiou<sup>1</sup>, Costas S. Tzafestas<sup>1</sup>, Petros Maragos<sup>1</sup>, Georgios Pavlakos<sup>1</sup>,  
Georgia Chalvatzaki<sup>1</sup>, George Moustiris<sup>1</sup>, Iasonas Kokkinos<sup>2</sup>, Angelika Peer<sup>3</sup>,  
Bartłomiej Stanczyk<sup>4</sup>, Evita-Stavroula Fotinea<sup>5</sup>, and Eleni Efthimiou<sup>5</sup>

<sup>1</sup> Inst. of Communication & Computer Systems,  
National Technical Univ. of Athens, Greece  
{xpapag, gmoustri}@mail.ntua.gr,  
{ktzaf, maragos}@cs.ntua.gr,  
geopavlakos@gmail.com,  
gchal@central.ntua.gr

<sup>2</sup> INRIA Ecole Centrale Paris, France  
iasonas.kokkinos@ecp.fr

<sup>3</sup> Technische Universität München, Munich, Germany  
Angelika.Peer@tum.de

<sup>4</sup> ACCREA Engineering, Lublin, Poland  
b.stanczyk@accrea.com

<sup>5</sup> Institute for Language and Speech Processing,  
ATHENA RC, Athens, Greece  
{evita, elenie}@ilsp.gr

**Abstract.** Mobility disabilities are prevalent in our ageing society and impede activities important for the independent living of elderly people and their quality of life. The goal of this work is to support human mobility and thus enforce fitness and vitality by developing intelligent robotic platforms designed to provide user-centred and natural support for ambulating in indoor environments. We envision the design of cognitive mobile robotic systems that can monitor and understand specific forms of human activity, in order to deduce what the human needs are, in terms of mobility. The goal is to provide user and context adaptive active support and ambulation assistance to elderly users, and generally to individuals with specific forms of moderate to mild walking impairment.

To achieve such targets, a reliable multimodal action recognition system needs to be developed, that can monitor, analyse and predict the user actions with a high level of accuracy and detail. Different modalities need to be combined into an integrated action recognition system. This paper reports current advances regarding the development and implementation of the first walking assistance robot prototype, which consists of a sensorized and actuated rollator platform. The main thrust of our approach is based on the enhancement of computer vision techniques with modalities that are broadly used in robotics, such as range images and haptic data, as well as on the integration of machine learning and pattern recognition approaches regarding specific verbal and non-verbal (gestural) commands in the envisaged (physical and non-physical) human-robot interaction context.

# 1 Introduction

## 1.1 Motivation

Mobility problems, particularly concerning the elderly population, constitute a major issue in our society. According to recent reports, approximately 20% of people aged 70 years or older, and 50% of people aged 85 and over, report difficulties in basic activities of daily living. Mobility disabilities are common and impede many activities important to independent living, [1], [2]. A significant proportion of older people have serious mobility problems. Furthermore, current demographics show that the elderly population (aged over 65) in industrialized countries shows a constant increase, [3].

Mobility is a crucial activity especially in the elderly since it promotes physical exercise, independence and self-esteem. Robotics seems to fit naturally to the role of assistance since it can incorporate features such as posture support and stability, walking assistance, navigation in indoor and outdoor environments, health monitoring etc. Our motivation in this approach stems from the fact that for an efficient and intelligent robotic assistant, a variety of multimodal interaction and cognitive control functionalities must be embedded, so that the robot can autonomously reason about how to provide optimal support to the user whenever and wherever needed.

## 1.2 Related Work

Several intelligent robot mobility aids are known, which are divided into two large categories: robotic wheelchairs and robotic walkers, [4]. They are designed to accommodate a normal walking pattern with opposite arm and leg moving together. Most robotic walkers are robotized variations of the typical Rollator frame, which is a standard walking frame attached to wheels, mainly used where balance -rather than weight bearing- is the major problem.

Many robotic walkers have been developed, generally presenting some of the following functionalities: (i) physical support; (ii) sensorial assistance; (iii) cognitive assistance; (iv) health monitoring and (v) advanced human - machine interface, [5]. These platforms largely fall into two categories: passive and active devices, [6]. While passive mobility aids either steer or brake, but cannot move forward without the human applying forces on them, [7]-[9], active devices are equipped with actuators and thus, their motion and interaction behaviour can be actively controlled, [10]-[12].

The research work reported in this paper aims to further extend and enhance the functionalities of such systems, by focusing on the development of robotic mobility aids for indoor environments that provide intelligent and active walking/mobility assistance (standing up, walking, and sitting down) to elderly people (with and without enough strength to support and stabilize themselves), in particular by supporting safe autonomous proactive control (for instance, by incorporating fall prevention features) and adaptive user-robot interaction, through multimodal sensory processing and intuitive human-robot communication.

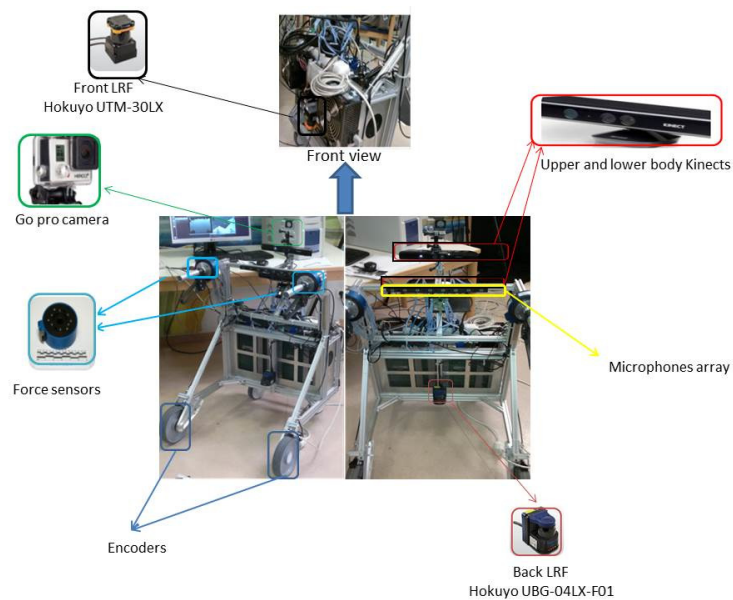
### 1.3 Organization and Overview

This paper is organized as follows. In Section 2, the first prototype of the experimental platform is described. The multimodal information processing and action recognition system is discussed in Section 3. The appropriate analysis and identification of human walking motions and classification of specific mobility weaknesses are discussed in Section 4. The design of a context-aware control architecture for an active mobility aid robot is briefly analysed in Section 5. Finally, conclusion and plans for future work are given in Section 6.

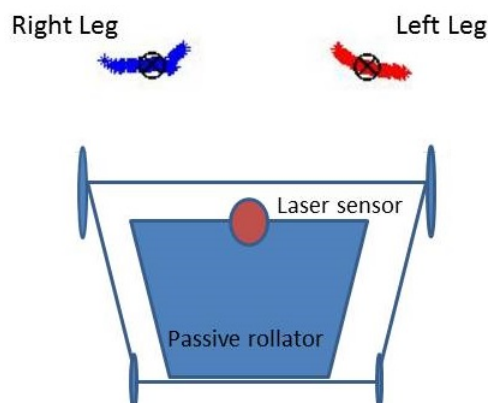
## 2 Experimental Platform

The experimental platform consists of a sensorised passive rollator prototype, as depicted in Fig. 1, which has already been used for the purposes of data collection and recording in a set of predefined typical use-cases and scenarios involving elderly people. The system incorporates multimodal information from laser range finder sensors, force/torque sensors, RGB and RGB-D cameras, encoders, and microphones.

First of all two laser sensors are used on the experimental platform. The first sensor (Hokuyo UTM-30LX) is placed at the front of the platform facing towards the motion direction, to provide full scanning of the walking area. For the purpose of human legs' detection and tracking, a second laser range finder (Hokuyo UBG-04LX-F01 rapid LRF) sensor is mounted at the back of the experimental platform facing the user legs, scanning a horizontal plane at the lower limbs at a height of approximately 40cm from



**Fig. 1.** Experimental platform: first sensorized prototype used for data acquisition and recordings



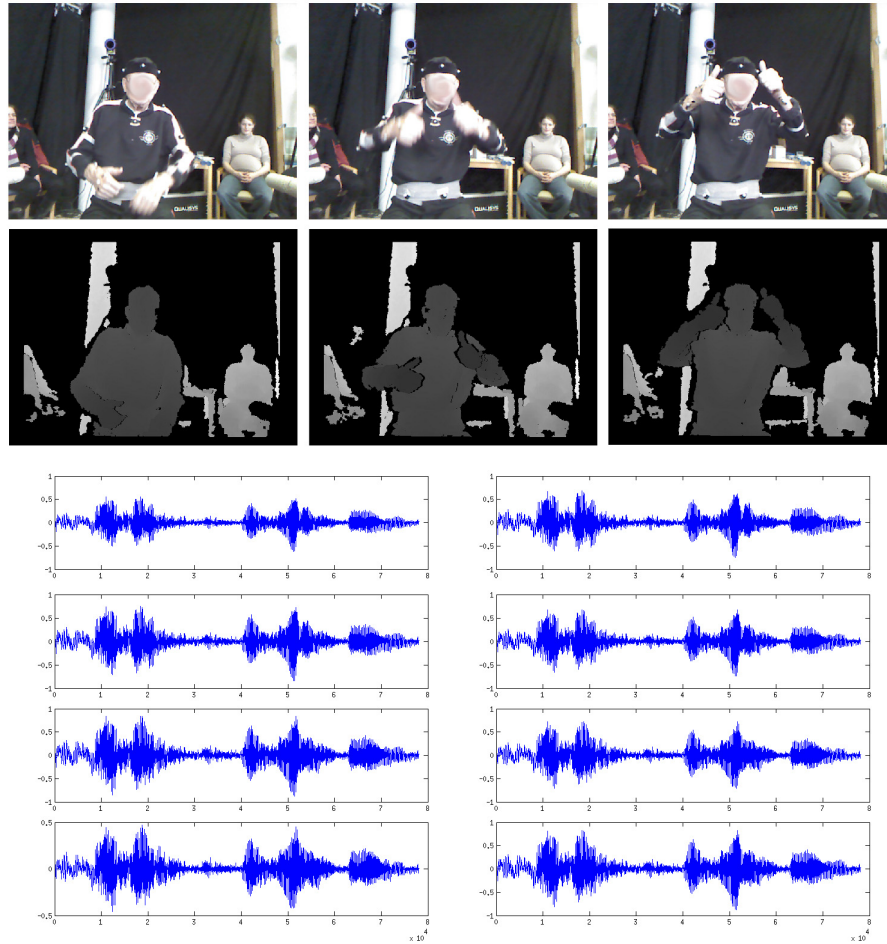
**Fig. 2.** A snapshot of the detected user's legs along the legs' centers

the ground, in order to take measurements from the users' gaiting performance. An example of the detected subject's legs is depicted in Fig. 2. A rectangular area (a search window) in front of the laser sensor is defined, inside of which we search for the potential user legs. For the detection process, a simple background extraction is used, in which we discard outliers that do not satisfy certain constraints defined for the potential legs, such as the minimum distance between consecutive points in each scanning frame and also the jump distance between neighboring laser groups. Subsequently, a K-means clustering algorithm is implemented to detect the legs' clusters.

Furthermore, two 6 DOF HR3 force/torque sensors (JR3 45E15) are placed at the handles of the experimental rollator in order to measure the applied forces between the patient and the rollator. In addition to the above sensors, an HD small, high performance and versatile video camera (GoPro) is also mounted on the rollator. This device is able to record the patient's head, torso, arms and all his/her movements and it provides necessary information for the subject's condition.

Another set of sensors placed on the rollator consists of two Kinect-for-Windows (KFW) sensors. This set of sensors allows the recording of the human body from a short distance. The first KFW sensor is placed horizontally towards the patient, in order to capture the area of the torso, waist, hips and the upper part of the limbs. The second KFW sensor is facing downwards in complementary direction with the first one, so as to achieve broad coverage and not to interfere with each other. This second sensor is able to capture the lower limb motion information, and therefore to enhance the localisation of limb positions and assist the analysis of gait abnormalities.

In addition to the above, an array of 8-microphone MEMS is also used as audio capturing device, mounted on the horizontal bar of the experimental platform in a linear configuration (with a 4cm uniform spacing) in front of the user (see Fig. 1). An example of a specific audio-gestural command is depicted in Fig. 3 (the captured audio visual data for the command "MOBOT: I want to stand up"). The visual data is obtained from the RGB and depth stream of the Kinect visual sensor while the audio data from the



**Fig. 3.** Multimodal data captured for a specific audio-gestural command (“MOBOT: I want to stand up”, “MOBOT: Ich will aufstehen”). Top: RGB data from Kinect, Middle: Depth data from Kinect, and Bottom: Multichannel waveform outputs from MEMS microphone array.

microphone array mounted on the rollator. A distinct waveform for each channel of the MEMS microphone array is presented along with indicative frames of the visual stream.

### 3 Multimodal Sensory Processing for Human Action Recognition

Different sensory modalities need to be combined into an integrated human action recognition system. The development of robust and effective computer vision techniques are needed in order to achieve the visual processing goals based on multiple cues such as spatio-temporal RGB appearance data as well as depth data from Kinect sensors. The enhancement of computer vision techniques with modalities that are broadly

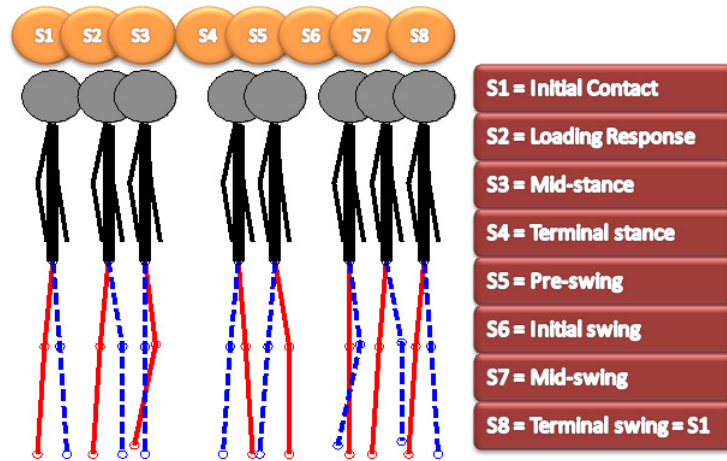
used in robotics, such as range images and haptic data will be introduced. Another major challenge is the integration of recognizing specific verbal and non-verbal (gestural) commands in the considered human-robot interaction context. All the above technical problems and research challenges embody all tasks related to visual, speech, haptic, and physiological data processing for detecting, tracking and recognizing the human actions. Specific objectives include the following:

- detection of human actions rapidly and robustly based on low and high level information,
- classification of human actions by analyzing spatio-temporal video information,
- analysis of walking patterns and detection of abnormalities by tracking human body pose,
- processing of multimodal data (visual, speech, haptic, physiology) originating from different sensors,
- developing of isolated spoken and gestural command recognition capabilities for human-robot communication, and
- ultimately performing human action recognition by multimodal sensory integration.

The approach followed towards this end is divided into several tasks in the current study. The first task focuses on visual processing for human localisation and action classification, based on a set of enhanced visual features and object representations. Another task is related to limb localisation and body pose estimation, combining visual appearance, motion and range data, that is, also exploiting data provided by the Kinects' depth sensors. Results of this task will also be integrated in a system performing isolated gesture recognition, which is currently under development and testing. Preliminary results in this direction are very promising, showing that the combination of multiple modalities enhances specific performance characteristics. Another task focuses on spoken command processing for human-robot communication and is based on HMMs for the recognition of specific spoken commands from distant microphones. Multimodal fusion approaches are also being investigated to achieve human action recognition, integrating cues from various sensorial modalities, including vision (RGB appearance, depth, tracking data) and speech modalities.

## 4 Human Action Analysis

One of the main goals of our work is to perform analysis and identification of human walking motions and classification of specific walking pathologies and associated mobility weaknesses. The idea is to conceptually and systematically synthesize all the modules performing multimodal recognition and inference regarding human behaviour and user intent, onto generative context-aware mobility assistance models. Towards this end, we investigate a completely non-invasive framework for analyzing a normal human walking gait pattern, which is described in this section.



**Fig. 4.** Internal states of normal gait cycle (Left Leg: blue dashed line, Right Leg: red solid line)

#### 4.1 Normal Human Gait Cycle Description

A well-known fact is that the walking patterns are gaits, that is, cyclic patterns with several consecutive phases. These cyclic motions can be modelled using a set of consecutive and repetitive gait phases. A basic requisite of the act of walking is the periodic movement of each foot from one position of support to the next. This element is necessary for any form of bipedal walking to occur, no matter how distorted the pattern may be by an underlying pathology, [13]. This periodic leg movement is the essence of the cyclic nature of human gait.

There are two main phases in the gait cycle, [14,15]: The **stance phase**, when the foot is on the ground, and the **swing phase** when that same foot is no longer in contact with the ground and is swinging through in preparation for the next foot strike. The stance phase may be subdivided into three separate phases: 1. First double support, when both feet are in contact with the ground, 2. Single limb stance, when only one foot is in ground contact and the other foot is swinging forward, 3. Second double support, when both feet are again in ground contact. The same terminology would be applied for both the left and right side of the body. For a normal person, each side is half a cycle behind (or ahead) of the other side. Thus, first double support for the right side is second double support for the left side, and vice versa. In normal gait there is a natural symmetry between the left and right sides, but in pathological gait an asymmetrical pattern very often exists.

Traditionally the gait cycle has been divided into eight events or periods, five during stance phase and three during swing.

**The stance phase events are as follows (see Fig. 4):**

1. Heel strike initiates the gait cycle and represents the point at which the body's centre of gravity is at its lowest position.

2. Foot-flat is the time when the plantar surface of the foot touches the ground.
3. Midstance occurs when the swinging (contralateral) foot passes the stance foot and the body's centre of gravity is at its highest position.
4. Heel-off occurs as the heel loses contact with the ground and pushoff is initiated via the triceps surae muscles, which plantar flex the ankle.
5. Toe-off terminates the stance phase as the foot leaves the ground, [16].

**The swing phase events are as follows:**

6. Acceleration begins as soon as the foot leaves the ground and the subject activates the hip flexor muscles to accelerate the leg forward.
7. Midswing occurs when the foot passes directly beneath the body, coincidental with midstance for the other foot.
8. Deceleration describes the action of the muscles as they slow the leg and stabilize the foot in preparation for the next heel strike.

Thus, there are eight events, but these are sufficiently general to be applied to any type of gait as shown in Table 4.1.

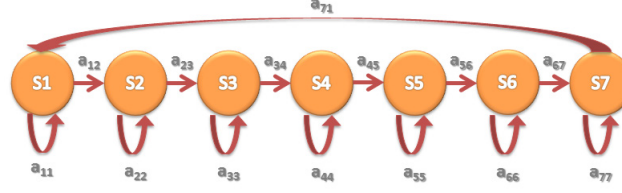
**Table 1.** Gait Cycle Events, [14]

Gait Phase	% Duration
1. Initial contact - <b>IC</b>	(0%)
2. Loading response - <b>LR</b>	(0-10%)
3. Midstance - <b>MS</b>	(10-30%)
4. Terminal stance - <b>TS</b>	(30-50%)
5. Preswing - <b>PW</b>	(50-60%)
6. Initial Swing - <b>IW</b>	(60-70%)
7. Midswing - <b>MW</b>	(70-85%)
8. Terminal swing - <b>TW</b>	(85-100%)

## 4.2 Detection of Gait Cycle Based on Hidden Markov Model

Hidden Markov Models are well suited for gait analysis and recognition because of their statistical properties and their ability to reflect the temporal state-transition nature of gait. An HMM is defined as a doubly embedded stochastic process with an underlying process that is not observable (i.e. it is hidden), but can only be observed through another set of stochastic processes that produce the sequence of observations [17][18]. This means that the states underlying the data generation process are hidden and can only be inferred through observations. HMMs are very common in several applications such as speech recognition [18][19], biological sequence analysis [20], gesture recognition [21], as well as human activity analysis [22].





**Fig. 5.** The topology of the network used for gait analysis and recognition is based on a left-to-right Hidden Markov Model

Based on the analysis of Subsection 4.1, the idea is to build a model that can distinguish between the different gait phases in order to analyze the normal gait cycle. The number of phases used in this study are seven, since the Terminal Swing phase is characterized by heel strike, which is an equivalent trigger as for Initial Contact phase, meaning that these two phases (TW and IC) can be eventually treated as identical. These seven gait phases, thus, correspond to the hidden states of the HMM. As observables, several quantities that represent the motion of the subjects' legs are used. These quantities are estimated using sequential signals from the laser range finder sensor that collects appropriate data (such as relative position w.r.t. the laser, velocities, etc.), while the rollator follows the subject's motion.

Following the HMM notation, the transition probability matrix is defined as

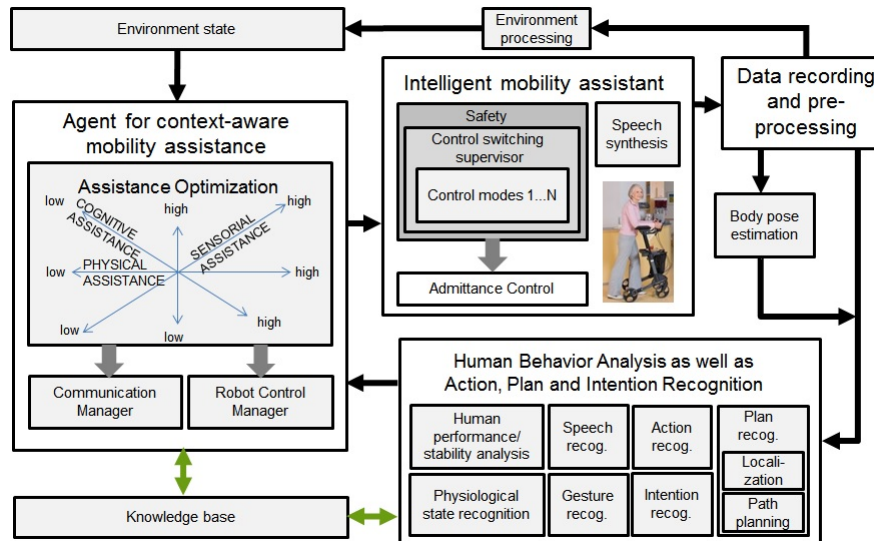
$$A = \{a_{ij}\}, \quad \text{where} \quad a_{ij} = P[s_{t+1} = j | s_t = i], \quad \text{for} \quad 1 \leq i, j \leq N,$$

where  $N$  is the number of states, and the  $(i, j)$  element of the matrix  $A$  represents the transition probability from the  $i^{\text{th}}$  state, at a given time step  $t$ , to the  $j^{\text{th}}$  state at the following time step (where  $t = 1, 2, \dots, T$ , and  $T$  denotes the total time). In the normal gait cycle, the gait phases follow each other sequentially. Thus, this HMM is a left-to-right model. This means that the only feasible transitions from a state  $i$  will be either to remain in the same state or to jump to the following adjacent state, as depicted in Fig. 5. The transition probability matrix, as well as the prior probability vector (i.e. the vector of probabilities  $\pi_i$  of the system being at state  $i$  at the initial time  $t_0$ ), are estimated using the standard and well known Baum-Welch algorithm, [18].

Initial results obtained by applying this model in normal human gait data are very promising [23], and demonstrate that this human data analysis scheme has the potential to provide the necessary methodological (modeling, inference, and learning) framework for a cognitive behavior-based robot control system. More specifically, the proposed framework has the potential to be used for the recognition of abnormal gait patterns and the subsequent classification of specific walking pathologies, which is needed for the development of a context-aware robot mobility assistant.

## 5 Context-Aware Robot Control

All the subsystems and individual signal processing and control modules developed to achieve some of the aforementioned functionalities need to be integrated seamlessly



**Fig. 6.** Context-aware Mobility Assistant Architecture

into a full-scale robot control system. The current form of the overall functional robot control architecture, encompassing all these submodules, is depicted in Fig. 6. One of the most important characteristics of this system is that it must perform context-aware operations, that is, it must reason and adapt its operational behaviour based on the following information:

- environment state,
- user behaviour, physiological state, and action, plan, and intention recognition,
- user-robot interaction context (verbal or non-verbal, vocal or non-vocal, physical or non-contact, etc.).

Two basic modules are currently being integrated and tested on the platform. Firstly, all the methods and algorithms necessary to enable autonomous sensor-based navigation of the mobile robot assistant in an indoor domestic environment are being developed and integrated. These include: map creation and localisation on static or dynamic maps, global path planning (using metric and/or topological map representations), obstacle detection and on-line collision avoidance, and targeted (goal-oriented) motion control supported by realtime path planning. Secondly, all methods and control algorithms for physical user-robot interaction are being developed. Currently, an adaptive admittance controller is being integrated on the platform, providing haptic collision-avoidance assistance to the user. Indeed, the robot walking assistant is equipped with two haptic interaction points (force/torque sensors on the two handles) and is able to apply different assistance policies when physically supporting a user.

Non-physical user-robot interaction modules are also under development, meaning that the mobility assistant does not necessarily have to be in continuous contact with the user, but could accompany him/her while walking from one place to the other, only

approaching the user to provide assistance when needed. Such user-accompanying assistance behaviors are also being investigated, where the corresponding robot motion controllers that will endow the mobility assistant with such user following functionalities are based on non-contact (mainly, laser rangefinder) sensor data. In more general terms, several models of user assistance can be applied, and special emphasis is needed on schemes that allow independently setting the cognitive and physical dominance levels for the mobility assistant and thus, to take over different roles for workload sharing and decision making.

## 6 Conclusions and Future Work

This paper reports current advances related to the development of a multimodal framework for intelligent robotic mobility aids. The system aims to provide context-adaptive and active walking/mobility assistance (standing up, walking, and sitting down) to elderly people (with and without enough strength to support and stabilize themselves). In particular, the system will support autonomous and proactive control modes (for instance, by incorporating user monitoring and fall prevention features), through multimodal sensory processing empowering intuitive human-robot interaction.

For further research, a reliable multimodal action recognition system needs to be enhanced, that will monitor and analyse human actions and predict user intentions with a high level of accuracy and detail. Different modalities will be combined into an integrated action recognition system. The main thrust of our approach is based on the enhancement of computer vision techniques with modalities that are broadly used in robotics, such as range images and haptic data, as well as the integration of pattern recognition modules focusing on specific verbal and non-verbal (gestural) commands in the considered human-robot interaction context. A context-aware robot control architecture will be further developed to synthesize all these multimodal sensory processing and pattern recognition modules into human-adaptive assistance models, aiming to provide optimal physical and/or cognitive support to the user.

**Acknowledgment.** This work is supported by 7th Framework Program of the European Union, ICT Challenge 2, Cognitive Systems and Robotics, contract “EU-FP7-ICT-2011-9 2.1 - 600796 - MOBOT: Intelligent Active MObility Assistance RoBOT Integrating Multimodal Sensory Processing, Proactive Autonomy and Adaptive Interaction”.

## References

1. Parkinson Disease Foundation. Statistics for parkinson’s disease (2010)
2. Stroke Center. Stroke statistics (2010)
3. USCensus. The elderly population (2010)
4. Machiel Van der Loos, H.F., Reinkensmeyer, D.J.: Rehabilitation and health care robotics. In: Siciliano, B., Khatib, O. (eds.) Springer Handbook of Robotics, pp. 1223–1251. Springer, Heidelberg (2008)

5. Martins, M.M., Santos, C.P., Frizera-Neto, A., Ceres, R.: Assistive mobility devices focusing on smart walkers: Classification and review. *Robotics and Autonomous Systems* 60(4), 548–562 (2012)
6. Ko, C.-H., Agrawal, S.K.: Walk-assist robot: A novel approach to gain selection of a braking controller using differential flatness. In: *American Control Conference (ACC)*, pp. 2799–2804 (2010)
7. Hirata, Y., Hara, A., Kosuge, K.: Motion control of passive intelligent walker using servo brakes. *IEEE Transactions on Robotics* 23(5), 981–990 (2007)
8. Kulyukin, V., Kutiyawala, A., LoPresti, E., Matthews, J., Simpson, R.: Iwalker: Toward a rollator-mounted wayfinding system for the elderly. In: *2008 IEEE International Conference on RFID*, pp. 303–311 (2008)
9. Wasson, G., Sheth, P., Alwan, M., Granata, K., Ledoux, A.: Cunjun Huang. User intent in a shared control framework for pedestrian mobility aids. In: *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)*, vol. 3, pp. 2962–2967 (2003)
10. Lee, G., Jung, E.-J., Ohnuma, T., Chong, N.Y., Yi, B.-J.: Jaist robotic walker control based on a two-layered kalman filter. In: *2011 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3682–3687 (2011)
11. Graf, B., Hans, M., Schraft, R.D.: Care-o-bot ii—development of a next generation robotic home assistant. *Autonomous Robots* 16(2), 193–205 (2004)
12. D’Angelo, L.T., Loercher, A., Lueth, T.C.: A new electrically driven walking frame for both passive and active mobility support. In: *2011 IEEE Conference on Automation Science and Engineering (CASE)*, pp. 816–821 (2011)
13. Ralston, H.J., Todd, F., Inman, V.T.: *Human walking*. Williams & Wilkins, Baltimore (1981)
14. Perry, J.: *Gait Analysis: Normal and Pathological Function*. Slack Incorporated (1992)
15. O’Connor, J.C., Vaughan, C.L., Davis, B.L.: *Dynamics of Human Gait*. Human Kinetics Publishers (1992)
16. Cochran, G.V.B.: *A primer of orthopaedic biomechanics*. Churchill Livingstone, New York (1982)
17. Rabiner, L., Juang, B.: An introduction to hidden markov models. *IEEE ASSP Magazine* 3(1), 4–16 (1986)
18. Rabiner, L.R.: Readings in speech recognition. chapter A tutorial on hidden Markov models and selected applications in speech recognition, pp. 267–296. Morgan Kaufmann Publishers Inc., San Francisco (1990)
19. Katsamanis, A., Papandreou, G., Maragos, P.: Audiovisual-to-articulatory speech inversion using active appearance models for the face and hidden markov models for the dynamics. In: *ICASSP*, pp. 2237–2240. IEEE (2008)
20. Yoon, B.-J.: Hidden markov models and their applications in biological sequence analysis. *Curr. Genomics* 10(6), 402–415 (2009)
21. Theodorakis, S., Katsamanis, A., Maragos, P.: Product-hmms for automatic sign language recognition. In: *ICASSP*, pp. 1601–1604. IEEE (2009)
22. Turaga, P., Chellappa, R., Subrahmanian, V.S., Udrea, O.: Machine recognition of human activities: A survey. *IEEE Transactions on Circuits and Systems for Video Technology* 18(11), 1473–1488 (2008)
23. Papageorgiou, X., Chalvatzaki, G., Tzafestas, C., Maragos, P.: Hidden markov modeling of human normal gait using laser range finder for a mobility assistance robot. In: *Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA 2014)* (2014)