2011 IEEE International Conference on Rehabilitation Robotics
Rehab Week Zurich, ETH Zurich Science City, Switzerland, June 29 - July 1, 2011

An Embedded Human Motion Capture System for An Assistive Walking Robot

Cong Zong, Xavier Clady and Mohamed Chetouani
Université Pierre et Marie Curie
Institut des Systèmes Intelligents et de Robotique
Paris, France
Email: {cong.zong, clady, chetouani}@isir.upmc.fr

***Abstract*—An embedded 3D body motion capture system for an assistive walking robot is presented in this paper. A 3D camera and infrared sensors are installed on a wheeled walker. We compare the positions of the human articular joints computed with our embedded system and the ones obtained with an other accurate system using embodied markers, the Codamotion. The obtained results valid our approach.**

I. INTRODUCTION

Elderly or disabled people use walker to stabilise their body balance while walking. In the last years, some new cognitive or robotic devices [1], [2], [3], [4] have been proposed in order to restore ambulation to patients suffering from mild cognitive, visual and physical impairments. This type of intelligent mobility assistance can provide high levels of care and safety for users. The PAM-AID (Personal Adaptive Mobility AID) project [1] built a mobility aid for the infirm blind which provides both a physical support for walking and navigational intelligence. The main objective is in navigation to prevent collisions: it provides obstacle detection to a visually impaired person using bumpers, sonars and infra-red based sensors. Force and torque sensors are mounted on the handles to estimate the users intent, to share the control of the motorized walker with the user [2] and to adapt the behaviour of the robot to the user's characteristics [3]. Integrating a controlled two Degrees-Of-Freedom (DOF) mechanism for the handles, Monimad [4] allows mobility rehabilitation and assistance, safe walking and safe sit-to-stand transfer. For these systems, the six-axis force-torque sensor is the main user interface. In this paper, we propose to embed on an advanced version

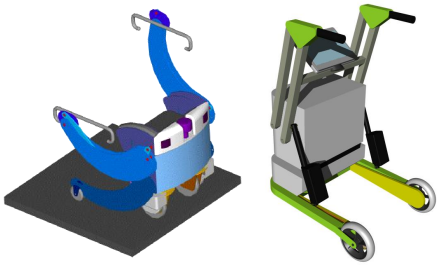


Fig. 1: Two versions of the RobuWalker

of Monimad, named RobuWalker (cf. Fig. 1), a new user monitoring system to perceive his movements and to recognize his intentions and actions. It consists in a three-dimensional (3D) human body motion capture system using a SwissRanger 3D camera and two infrared distance sensors (cf. Fig. 9) that could be embedded on the robot. Providing a fine analysis of the human motion, this system could help the robot to detect abnormal situations during the assisted human gait and therefore control the mobility assistance system accordingly in order to prevent unexpected and unstable motions [5], [6].

As illustrated in figure 2, the 3D camera provides a 3D points cloud from the top body (approximately from the knees to the neck). The infrared sensors are dedicated to the feet movements capture. A 3D human body model is fitted on the data provided by the sensors. In order to test this approach, we have embedded this sensor system on a non-motorized wheeled walker. The obtained results are compared with a 3D body reconstruction using Codamotion [7], an accurate motion capture system (but using embodied markers - cf. Fig. 9). The

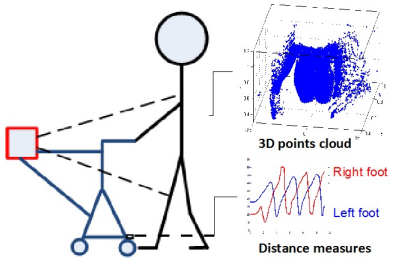


Fig. 2: Scheme of the embedded sensors system and the provided data

rest of the paper is organized as follows: Section II presents some related works on human motion capture system and describes 3D human body model used in this study. Section III explains the proposed method. The experimental setup and the results are shown in Section IV. This section presents a comparison between the results obtained from our system and a complete observation with Codamotion. Finally, Section V concludes our paper and succinctly describes some ongoing and future works.

II. RELATED WORKS

This section briefly describes related works on motion capture and motivates our hardware and software choices.

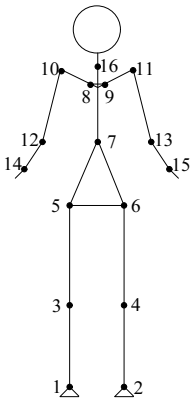
A. Human Motion Capture System

Such systems capture the pose changes of a human body during motion, based on motion sensor technologies [8]. In general, they can be divided into: 1) the non-visual systems [9], [10] (e.g. inertial sensor, magnetic sensor and electromyogram, etc.); 2) the visual systems [11], [12] (e.g. marker and markerless based); 3) a hybrid system [13]. Among visual marker based systems, the Codamotion [7] could minimize the uncertainty of movement of a subject, because of the uniqueness of markers: it has been commonly used as ground truth (as in this study IV) to evaluate the 3D motion measurements due to its high accuracy. However, the markers (and most of the non-visual systems) should be embodied, which is inconsistent with a realistic use of the robot.

For this reason, we have focused our work on a markerless system. In this domain, various vision based approaches have been investigated in the literature (see recent surveys [11], [12]). The most accurate methods estimate the human pose by minimizing distance errors between a 3D model body model and the data extracted from monocular or multiple view images or videos; these errors could be expressed in the image or 3D space. But they are often designed for static cameras: a motion based segmentation preprocessing step or a stable and precise calibration between multiple cameras is often required. Furthermore they could be cpu time and memory resource consuming. In order to override these potential limitations, we propose to use a 3D camera which provides directly 3D data. Because the robot is size constrained (it should be used in cluttered locations like home), the low relative distance between the user and the camera limits the user observation to his upper-body part; we complete it by using two infrared distance sensors as explained in section III-A.

B. 3D Human Body Model Description

A 3D human body model Human36 is employed in our approach. It derives from the Human36 model, which is based on the anatomical model from the Humanoid Motion Analysis and Simulation (HuMAnS) toolbox [14] and on the dynamic model implemented in Arboris. The HuMAnS toolbox developed at the INRIA in Grenoble offers tools for the modeling, the control and the analysis of humanoid motion, being that of a robot or a human. For example, Barthélemy et al. [15] use this toolbox to reconstruct the joint trajectories corresponding to the sit-to-stand motion in order to compare the inertial forces computed from the reconstructed motion with the recorded ground reaction forces. This model has 36 rotational degrees of freedom corresponding to 16 joints (see Fig. 3): ankle (2 DOF), knee (1 DOF), hip (3 DOF), thorax (3 DOF), sternoclavicular (2 DOF), shoulder (3 DOF), elbow (2 DOF), wrist (2 DOF) and head (3 DOF). There are also 6 DOF (three rotations and three translations) from the global coordinate system to the local coordinate system at the hip. Our system is a motion capture system under partial



Tags number	Tags name	DOF
1,2	Ankle	2
3,4	Knee	1
5,6	Hip	3
7	Thorax	3
8,9	Sterno-clavicular	2
10,11	Shoulder	3
12,13	Elbow	2
14,15	Wrist	2
16	Head	3

Fig. 3: Articular notations of Human36

observation. Because of the limited field of view of 3D camera, not all parts of body are captured, such as head, shank and foot. In our 3D model body, we fix their associated degree-of-freedom. Furthermore, the motion of human body is limited by the walker. For example, the hands should always hold the handles of walker. Therefore, in order to avoid instabilities and non-natural poses, we predefine joint limits for all of the articulated angles $\theta_1, ..., \theta_n$ of dynamic model:

$$\theta_i = \begin{cases} \theta_i^{max} & if\ \theta_i \geq \theta_i^{max} \\ \theta_i^{min} & if\ \theta_i \leq \theta_i^{min} \\ \theta_i & otherwise \end{cases} \quad (1)$$

where θ_i^{max} and θ_i^{min} are predefined joint limits for θ_i.

III. OUR EMBEDDED CAPTURE MOTION SYSTEM

In this section, we describe how the human body model is fitted on the sensor data. Firstly, the sensors system is presented. Secondly, we described the developed algorithm.

A. Description of the sensor system

We capture 3D data using SwissRanger SR4000 camera based on the Time-Of-Flight distance measurement principle [16], [17]. It sends short pulses of infrared light and analyse their response. A part of emitted light is reflected from objects in the field of view and returns to the 176x144 image sensor of the camera. The time of arrival is measured independently by each pixel and thus determines the distance between the camera and objects. Each pixel (x, y) corresponds to a 3D point (X, Y, Z). Based on this, three 176x144 maps (X map, Y map and Z map), corresponding to the 3D points cloud, are obtained. In order to estimate 3D human body posture during walking, the 3D camera is installed on the walker (see Fig. 1 and 9).

Our robot will provide mobility assistance for the users in daily life, thus the dimensions of the robot are limited. The camera could not be mounted far from the user. Moreover, the height of users are different. Only a part of head, the chest, the arms and the upper of thighs will be captured for tall people; and chest, arms and thighs will be obtained for short people.

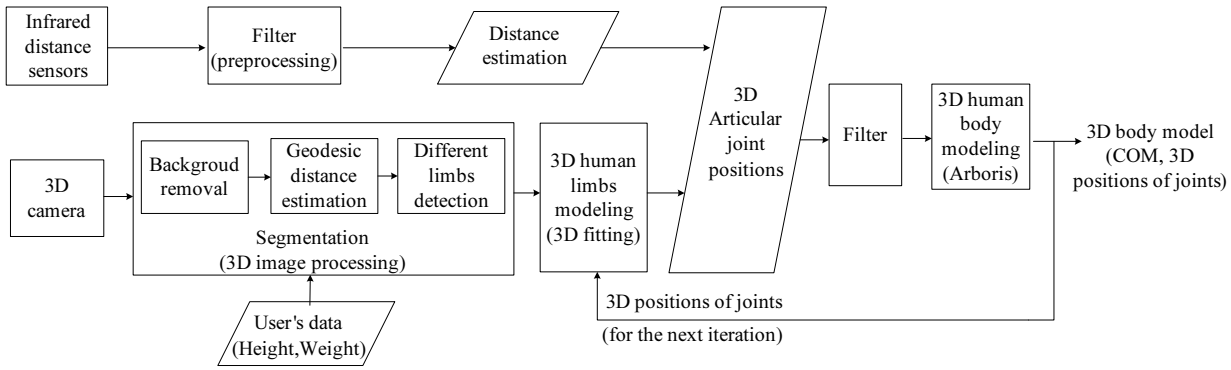


Fig. 4: Scheme of system under partial observation

The head of users could not guaranteed to be observed. We choose two infrared distance sensors (Sharp GP2Y0A02YK 20cm -150 cm) to complement the walking user observation. The two sensors were installed on the walker at 10 cm of the ground. They could measure the distance between the walker and ankle of user.

We define the middle point of two anterior wheels of walker as the origin point, the direction from left to right as the direction of X axis, the direction of walking as the direction of Y axis and the vertical direction as the direction of Z axis (see Fig. 8). The calibration process of this sensor is explained in Section IV.

B. The 3D modeling algorithm

Our system is built on an iterative algorithm with three main steps: 1) segmentation, 2) human limbs modeling and 3) human body modeling with Arboris. The objective of the segmentation step is the extraction of the 3D data related to each human limb. Unlike other approaches in the literature [18], our 3D human body model fitting is decomposed in two steps in order to reduce the cpu-time consuming and to avoid numerical instabilities[1]. Then, the Gauss−Newton algorithm is used to fit each limb by a cylinder. The 3D position of joints could be obtained by computing parameters of cylinders and combining the information of infrared distance sensors. Lastly, the 3D body model will be reconstructed by minimizing the quadratic error between the articular joints on the model and the measured ones. These steps are detailed as follows (see Fig. 4).

1) Segmentation: The 3D data of the chest, the arms and the upper of thighs are recorded by the 3D camera. Based on user's anthropometric data [19], the 3D human body data can be divided into different members: chest, upper arms, lower arms and thighs.

[1]The involved minimisation process in the 3D human body model fitting requires multiple inversions of jacobian matrix. The size of these matrix is proportional to NxM, with N the number of data and M the number of DOF. To decompose the global process in a set of local process reduces considerably the computational complexity, i.e. the time consuming and the numerical instabilities.

a) Depth-based background removal: The 3D data corresponding to the human body could be subtracted from the background using a basic threshold process applied on the Z map. We define a range of depth which corresponds with the region of interest and eliminate any points that do not belong to this range.

b) Geodesic distance: The geodesic distance is widely used to analyze the movement of human body [18], [20]. The advantage is that the geodesic distance between two points of human body will not be changed during motion. Its value is estimated by computing the length of the shortest path between 2 points (called as the starting point and the final point) belong the 3D surface using the Dijkstra's algorithm. Here this geodesic path could help us to detect limbs.

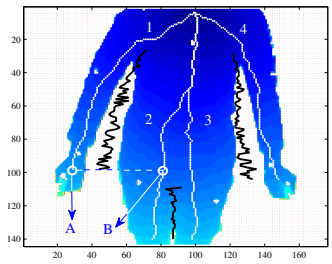


Fig. 5: Four geodesic paths (white curve): from the point near head 1. to the right hand; 2. to the right leg; 3. to the left leg; 4. to the left hand and the curve between limbs (black curve).

Four geodesic paths between the four extremities of the limbs and a point near the head are computed. The projections of these paths on a distance map (computed with the head point as origin point) are shown in the figure 5. The head point corresponds to the nearest point of the upper points in the Z map. The positions of the hand points are computed knowing the geometrical relation between the handles and the camera. The thigh points are selected as the two extrema points of the two modes (corresponding to the two legs) observed in the values of the lower line in the Z map.

c) Different limbs detection: The four geodesic paths are used to attribute the 3D data to the different limbs. The

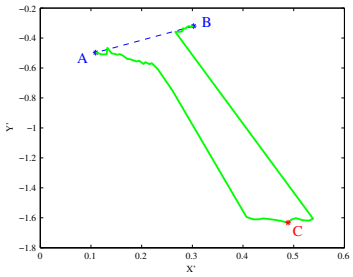


Fig. 6: An example of the projection of 3D curve (green solid curve) between points of path 1 and path 2 on the X'-Y' plane. The red point (C) is the minimum point.

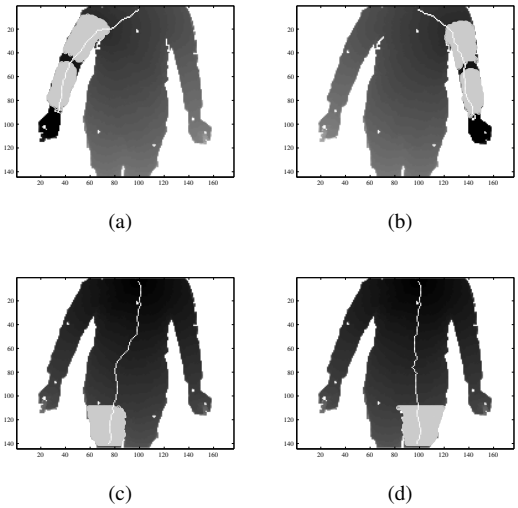


Fig. 7: Different body limbs: (a) right upper arm and lower arm (b) left upper arm and lower arm (c) right thigh (d) left thigh

result of this segmentation is illustrated in the figure 7. Firstly, selecting 2 points belonging to the same line in the Z map and the adjacent geodesic paths, we compute the furthest point between them. In the figure 6, it corresponds to the point C if points A and B are considered to belong to the path 1 and 2 respectively. Beginning from the lower points of the four paths and repeating the same process line after line until the distance between the further point and the line connecting the 2 selected points (the dashed blue line in the figure 6) is smaller than a threshold, four frontier curves (black curves in the figure 5) are obtained. Prolonging horizontally the left and the right curves, we separate the arms from the chest. Prolonging vertically the middle curve (to the left and to the right), we separate the tights and the chest. Finally, according to the anthropometric data, the length of body segments (Standard Human) is known; using the geodesic distance from the hands, we can separate the data belong to the lower arm from the ones belong to the upper arm.

2) 3D human limbs modeling: After segmentation, we have associated a 3D point cloud to each limb. The objective of this step is the modeling of these 3D points for getting certain feature points (the 3D articulation coordinates).Using the iterative Gauss-Newton algorithm, a cylinder is fitted to each limb's cloud. The cylinder is specified by a point (x_0, y_0, z_0) on its axis, a vector (a, b, c) pointing along the axis and its radius r. Using anthropomorphic data, we fix the length of the cylinders. The bases of the finite cylinders are disks: the centers of these circles could be associated to articular joints. For the lower arms and the tights, they are deducted from the positions of the extrema points (respectively near the wrist and the hip) belong to the geodesic paths and the limb's cloud: these points should belong to the disks. The knee articular joint positions is computed from the length of the tights, the positions of the hip articular joints and the axis directions of the corresponding cylinders. The elbow positions are inferred as the middle point between the upper extrema base of the lower arm cylinder and the intersection between this disk and the axis of the upper arm cylinder. Finally the shoulder positions are situated on this last axis using the length of the upper arms.

3) 3D human body modeling with Arboris: From the previous step, approximate positions of 12 articular joints are obtained: shoulder, elbow, wrist, hip, knee and ankle. We could perform simulations and analysis walking motion by using HuMAns toolbox (see Fig. 4). The joints trajectories are then reconstructed by minimizing the quadratic error between the articular joints on the model and the observed ones, using the position of articular joints and jacobian matrix provided by the HuMAns toolbox [14], [15].

IV. EXPERIMENTS AND RESULTS

In this section, we describe the experimental set up and discus about the obtained results.

A. Calibration

Before the experiments, camera and infrared sensors have been calibrated. A 3D pattern has been used to estimate the relative geometric positions and orientations of the camera. This pattern is composed by a 2D grid on the ground and two others parallel grids (built by moving a vertical stick as illustrated in the figure 8). The positions and orientations

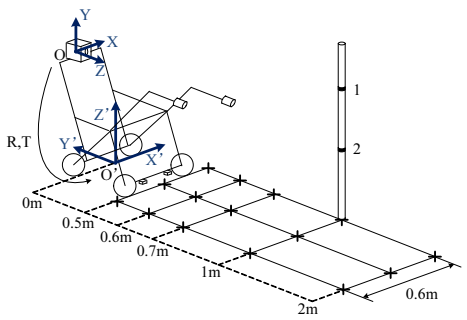


Fig. 8: Illustration of the camera calibration

of the camera are defined respectively by the pairs (R, T). Each point can be represented by its coordinates written as $(x_i, y_i, z_i, 1)^t$ in the camera-related homogeneous coordinate system O. The relation between these coordinates and the coordinates $(x'_i, y'_i, z'_i, 1)^t$ of the same points in the walker-related coordinate system O' is expressed as:

$$\begin{pmatrix} x'_i \\ y'_i \\ z'_i \\ 1 \end{pmatrix} = \begin{pmatrix} R_{3x3} & T_{3x1} \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} x_i \\ y_i \\ z_i \\ 1 \end{pmatrix} \quad (2)$$

From this relation and the known positions of the pattern points expressed in each coordinate system, we can express and resolve a system of linear equations which the unknowns are (R, T). Finally, the geometric positions of the infrared sensors have been evaluated by moving a box on the ground grid: a linear relation between the real distances and the measured distances has been estimated.

B. Experimental protocol

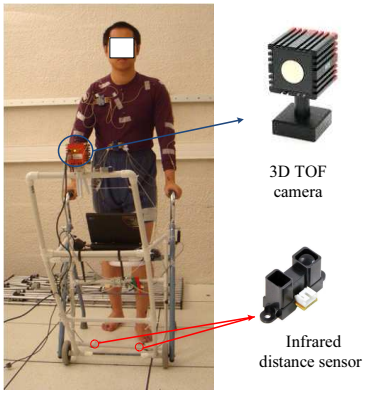
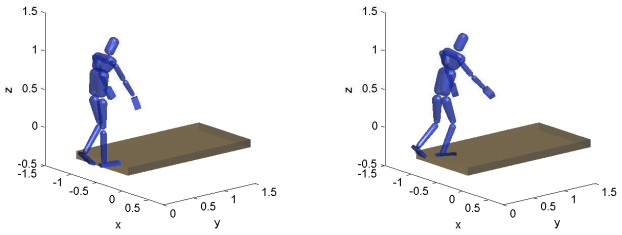


Fig. 9: Photography of the experimental setup

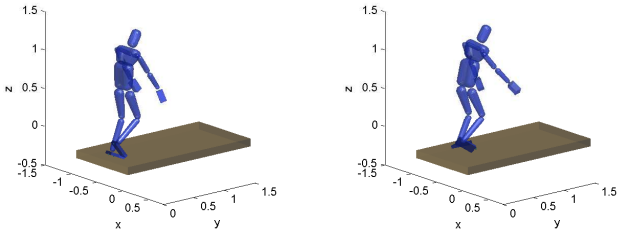
During the experiments, each tested subject has been equipped with 16 Codamotion markers [7]. This system provides us a complete and accurate observation of the subject motion during walking. The HuMAns toolbox and the complete Human36 model have been utilized to analyze walking motion. Subjects were asked to walk straight on a flat floor using our walker system at their most comfortable pace as shown in Fig. 9. Several measurements have been realized; we present here a man who is 29 years old (1.75 m, 80 kg). The collected data are compared with the ones obtained with our motion capture system in the section IV-C.

C. Results and discussion

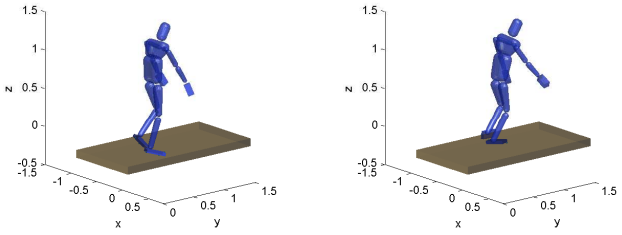
Image sequences in the figure 10 illustrate the obtained results using our embedded system and the Codamotion system. We have computed the distances between them for different articular joints: ankle, knee, hip, elbow, shoulder and wrist are shown in Fig. 11. They are averaged over the right and left sides of the body. We could also compute the trajectory of the center of mass (COM) of human body walking for the two



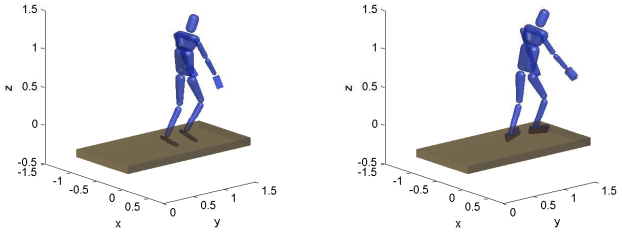
(a) 1 s



(b) 2 s



(c) 3 s



(d) 4 s

Fig. 10: Illustration of the results obtained by our embedded system (left) and the Codamotion system (right)

systems respectively using the HuMAnS toolbox, as shown in Fig. 12. Each wave of the COM represents a gait cycle: its rhythm and its complexity seem to be similar in both systems. It is an important result because they are potentially useful information to analyze the gait [21], [22]. The error values relating to the upper part of the body (elbow, shoulder and wrist) are around 50mm which are quite similar to other results reported in the literature. Other markerless, vision based and not embedded motion capture systems [12] achieve errors in the range of $33.9-85$ mm with most around $40-50$ mm. The error values relating to the body lower part are higher. It is

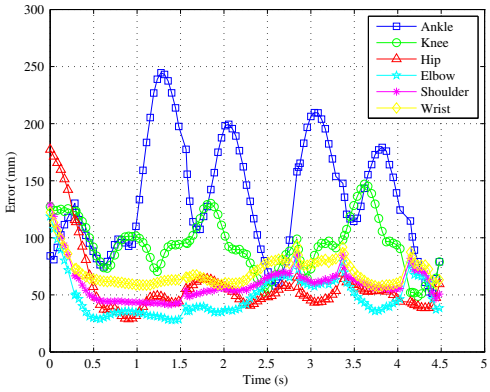


Fig. 11: The average errors for different joints.

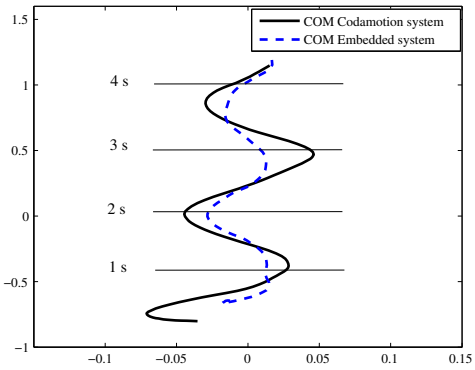


Fig. 12: The trajectory of the center of mass (COM) during walking

essentially due to the few and imprecise data provided by the infrared sensors. In future works, we would test a rangefinder laser (to replace the infrared sensor) and an other camera with a larger field of view (e.g., Kinect) in order to improve the system accuracy. Furthermore we plan to implement a predictive process using a walking motion model in a tracking scheme [23]. We would expect to measure the gap between the predicted motion and the observed motion in order to detect potential abnormal motions.

V. CONCLUSIONS AND FUTURE WORKS

In this paper, we have presented an embedded 3D motion capture system using a 3D TOF camera and two infrared distance sensors. This system could help our intelligent mobility assistive robot to detect abnormal situations during human walking and therefore control the mobility assistance system accordingly in order to prevent risks. We have computed the joint trajectories corresponding to the walking motion by using HuMAns toolbox. We compare the results obtained with our system and the Codamotion system. Even if some local errors are relatively high, the rhythm and the complexity of the COM trajectory are quite similar in both systems. Some improvements are proposed to increase the accuracy of our system. Currently, we embed our motion capture system on a non-robotic walker. We are planning to integrate the described system in our robotic walker. The information provided by our system would be combined with other ones (force-torque sensor, electrocardio-monitoring system mounted in the handles of the robot,...).

VI. ACKNOWLEDGMENTS

This work has been supported by French National Research Agency (ANR) through TecSan program (project MIRAS n°ANR-08-TECS-009).

REFERENCES

[1] G. Lacey and K. M. Dawson-Howe, "Personal adaptive mobility aid (pam-aid) for the infirm and elderly blind," in *Association for the Advancement of Artificial Intelligence*, 1996.
[2] B. Graf and R. Schraft, "Behavior-based path modification for shared control of robotic walking aids," *ICORR*, 2007.
[3] H. Yu, M. Spenko, and S. Dubowsky, "An adaptive shared control system for an intelligent mobility aid for the elderly," *Autonomous Robots*, 2003.
[4] V. Pasqui, L. Saint-Bauzel, and O. Sigaud, "Characterization of a least effort user-centered trajectory for sit-to-stand assistance user-centered trajectory for sit-to-stand assistance," in *International Union of Theoretical and Applied Mechanics 2010*.
[5] J. Glover, D. Holstius, M. K. Montgomery, A. Powers, J. Wu, S. Kiesler, J. Matthews, and S. Thrun, "A robotically augmented walker for older adults," Tech. Rep., 2003.
[6] T. Hirotomi, Y. Hosomi, and H. Yano, "Brake control assist on a four-castered walker for old people," in *ICCHP*, 2008.
[7] D. Mitchelson, "Codamotion, http://www.codamotion.com," *Charnwood Dynamics Limited*.
[8] H. Zhou and H. Hu, "Human motion tracking for rehabilitation – a survey," *Biomedical Signal Processing and Control*, 2008.
[9] B. Najafi, K. Aminian, F. Loew, Y. Blanc, and P. Robert, "Measurement of stand-sit and sit-stand transitions using a miniature gyroscope and its application in fall risk evaluation in the elderly," *Ieee Transactions on Biomedical Engineering*, 2002.
[10] C. Mavroidis, J. Nikitczuk, B. Weinberg, G. Danaher, K. Jensen, P. Pelletier, J. Prugnarola, R. Stuart, R. Arango, M. Leahey, R. Pavone, A. Provo, and D. Yasevac, "Smart portable rehabilitation devices," *Journal of NeuroEngineering and Rehabilitation 2005*.
[11] R. Poppe, "Vision-based human motion analysis: An overview," *Computer Vision and Image Understanding*, 2007.
[12] L. Sigal and M. J. Black, "Guest editorial: State of the art in image- and video-based human pose and motion estimation," *IJCV*, 2010.
[13] Y. Tao and H. Hu, "Buiding a visual tracking system for home-based rehabilitation," in *Chinese Automation and Computing Society*, 2003.
[14] P.-B. Wieber, F. Billet, L. Boissieux, and R. Pissard-Gibollet, "The humans toolbox, a homogeneous framework for motion capture, analysis and simulation," *3D analysis of human movement*, 2006.
[15] S. Barthélemy, C. Salaun, and P. Bidaud, "Dynamic simulation and control of sit-to-stand motion," in *CLAWAR*, 2006.
[16] MESA-imaging, "Sr4000 user manual," 2005.
[17] D. Anderson, H. Herman, and A. Kelly, "Experimental characterization of commercial flash ladar devices," in *ICST*, 2005.
[18] N. D. Thang, T.-S. Kim, Y.-K. Lee, and S. Lee, "Estimation of 3-d human body posture via co-registration of 3-d human model and sequential stereo information," *Applied Intelligence*, 2010.
[19] D. A. Winter, *Biomechanics and motor control of human movement*. John Wiley & Sons,Inc., 2009.
[20] P. Correa, F. Marqués, X. Marichal, and B. Macq, "3d posture estimation using geodesic distance maps," *Multimedia Tools Applications*, 2008.
[21] C. Zong, M. Chetouani, and A. Tapus, "Automatic gait characterization for a mobility assistance system." ICARCV, 2010.
[22] L. Decker, F. Cignetti, and N. Stergiou, "Complexity and human gait," *Revista Andaluza de Medicina del Deporte*, 2010.
[23] M. A. Brubaker, D. J. Fleet, and A. Hertzmann, "Physics-based person tracking using the anthropomorphic walker," *IJCV*, 2010.