CLAWAR 2013
Proceedings of the Sixteenth International Conference on
Climbing and Walking Robots, Sydney, Australia, 14 – 17 July 2013

A mobile 3D vision-based embedded system for robust estimation and analysis of human locomotion

Cong Zong and Philippe Bidaud
Université Pierre et Marie Curie, Institut des Systèmes Intelligents et de Robotique,
Paris, France
E-mail: cong.zong@isir.upmc.fr
philippe.bidaud@isir.upmc.fr

Xavier Clady
Université Pierre et Marie Curie, Centre de Recherche Institut de la vision,
Paris, France
E-mail: xavier.clady@upmc.fr

In this paper, we propose and experiment a system for human motion capture which can be incorporated into a mobility aid device such as a walker. This measurement system uses two Kinect 3D vision sensors whose data are mapped after treatment on a physical model of virtual human. This filtering method ensures data consistency as well as the observability of the whole body movement. It is beyond a basis for the analysis of human motor activity and more specifically here of locomotion (gait parameters, posture, joint efforts, energy). We detail the entire process for rebuilding the movement and discuss its performance from the experimental results.

Keywords: Physically-Based Motion capture, human motion analysis, multiple-camera tracking

1. Introduction

The analysis of human locomotion and its changes has led to numerous works by the past. This motor activity can be defined as the combination in time and space of more or less complex movements of different body parts leading to the displacement of the subject. To propel in upright position, the man uses a discrete cycle of unipodal and bipodal supports through the movement of the musculoskeletal system in which it seeks to maintain the controllability of the dynamic equilibrium. Methods for observation and analysis of this function have diversified progressively with technology development.

Most of the techniques used today are based on laboratory devices such as locometers for determining the temporal spatial parameters of gait (walking speed, step length, step duration, the duration of the course and not time propulsion), motion capture systems and posturographic platforms to access respectively the kinematics and the dynamics of the whole body movement. These means are implemented in particular to explore the range of disorders affecting locomotion-temporal spatial parameters of gait, average posture changing and the movements of all body parts.

The development of "holter" technologies for qualitative and quantitative assessment of walking and its disorders is today a challenge for the diagnosis and monitoring of patients at risk.

This paper proposes and explores the performance of a mobile system for capturing and analyzing human motion during walking. The latter consists of two Kinect sensors installed on a walker device. A set of algorithms have been developed for extracting data of the characteristic points of the anatomic envelope. These data are then repositioned on a physical model of virtual human to reproduce the observed motion and be analyzed to find anomalies for instance.

2. Embedded vision-based motion capture system

2.1. *General description*

Two Kinect sensors (see figure 1) $C1$ and $C2$ are used respectively to capture the relative motion of the upper body parts and the feet trajectories with respect to the walker. Wiimote sensors are installed on the wheels of the walker to determine the movement of the walker and then the absolute motion of the person (odometry system[1]).

2.2. *Data extraction*

2.2.1. *Capture of the upper body parts*

The sensor $C1$ is used to capture the upper body part motions. The data extraction method developed in this aim is based on an algorithm combining two steps: 1) Segmentation of the body system; 2) Modeling of the body parts from the sensor data.

A) Segmentation

The objective of the segmentation step is the extraction of the 3D data for each member of the human body. 3D data on the torso, arms and at the

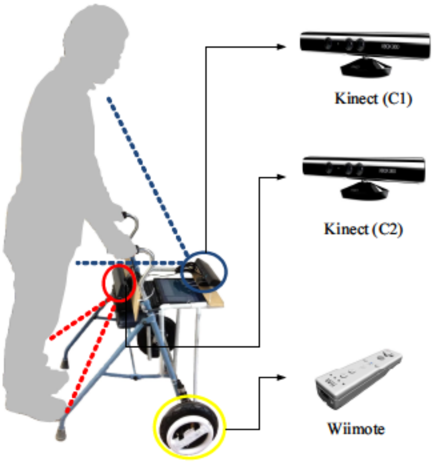
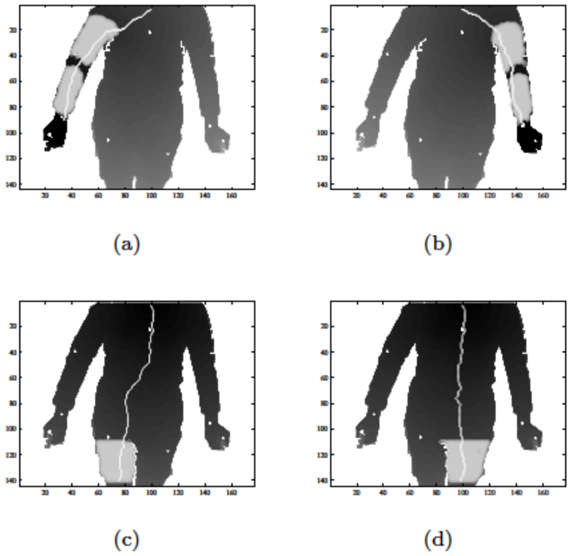


Figure 1. Plate-forme de capture

upper thighs are recorded by the sensor (C1). They are extracted after subtracting the background images using a threshold applied to the depth map. Then, knowing the user anthropomorphic parameters,[2] 3D data of the human body are distributed over the different members, torso, arms, forearms and thighs based on the geodesic distance between different the parts of body during movement. Four geodesic paths are defined between ends of the four members from an origin located near the head (see figure 2).

B) Modeling of the human body parts

After the segmentation phase, we associate a cloud of 3D points to each member in order to estimate the 3D coordinates of the anatomic joints. For that, a cylinder is fitted to the 3D data of each member using iterative Gauss-Newton method. The length of the cylinder is calculated from the anthropomorphic table associated with the human virtual model. The positions of shoulders, wrists and hips joints are detected from the positions of the extreme points of the cylinder axes. The positions of the knee joints are computed from the estimated length of the thighs and the estimated positions of the hip and the directions of the axes of the corresponding cylinders. The elbow positions are inferred as the midpoint between the point in the center of the upper face of the cylinder corresponding to the forearm and of the lower face of the cylinder corresponding to the arm. The



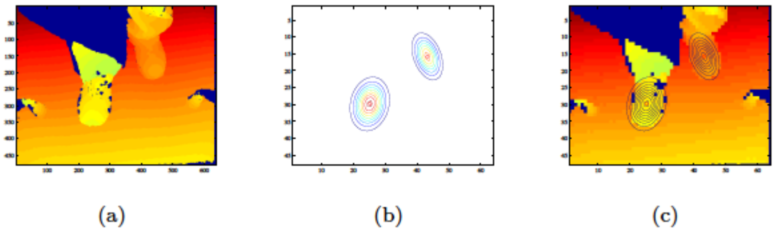
(a) (b) (c) (d)

Figure 2. Different members of the body: (a) the forearm and the right arm (b) the forearm and the left arm (c) the right thigh (d) the left thigh

positions of these anatomic joint points are stacked in the vector $P_h(k)$ at each acquisition period k.

2.3. *Capturing foot positions*

The position of the feet is provided by the sensor data Kinect (C2) directed towards the ground. It is obtained by approximating data of the feet as Gaussian distributions, the centers are considered as a characteristic point of the foot.



(a) (b) (c)

Figure 3. Example of images capturing the position of the feet

3. Reconstruction of the human motion

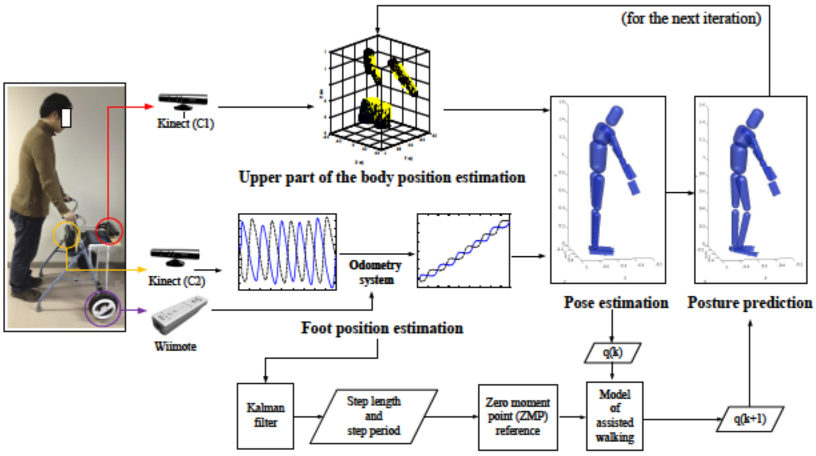


Figure 4. Scheme of system

3.1. *Mapping motion capture data to virtual human model*

To map all of the data collected with the vision system, we use a virtual human model[3] which consists of 19 body segments modeled as rigid bodies (7 for the lower and 12 upper part) and 16 anatomical joints. The virtual human posture is controlled and dynamically animated in a physically consistent way through a Linear Quadratic Program (LQP) controller.[4] At a frequency of 30 Hz, joint key points of the model are driven to the measured positions (see figure 5).

At each period of time, we compute joint motions V_{mesure} such as :

$$V_{mesure} = \frac{P_{mes}(k) - P_{mod}(k)}{dt} \tag{1}$$

and then are introduced into the objective function of a quadratic program LQP. At each new capture, which seeks to resolve:

$$\begin{aligned} \dot{q}^* &= argmin_{\dot{q}}((J\dot{q} - V_{mesure})^T W (J\dot{q} - V_{mesure})) \\ &= argmin_{\dot{q}}(\frac{1}{2}\dot{q}^T Q \dot{q} + \dot{q}^T P + \frac{1}{2}C) \\ under\ the\ constraints\ &: \begin{cases} q + \dot{q}dt \leq q_{max} \\ q + \dot{q}dt \geq q_{min} \end{cases} \end{aligned} \tag{2}$$

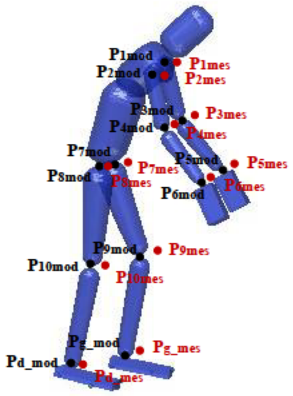


Figure 5. Digital model: the red dots are the measured positions of the points P_I of the joint axis and the black dots are the positions of the analog points on the model P_{mod}

‘ where the C, Q and P matrices are respectively :

$$C = V_{mesure}^{T} W V_{mesure} \tag{3}$$

$$Q = J^T W J \tag{4}$$

$$P = -J^T W V_{mesure} \tag{5}$$

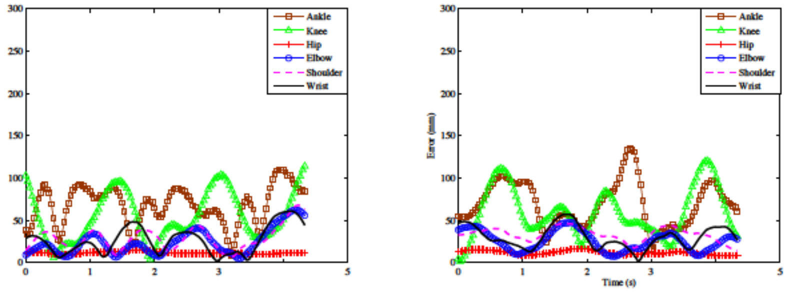
J is the Jacobian matrix joint points of the virtual human model. W is a diagonal matrix of weights associated with different articulations reflecting motor coordination between the different joints. $W = \begin{pmatrix} w_i & 0 & 0 \\ 0 & w_i & 0 \\ 0 & 0 & w_i \end{pmatrix}$. We define $w_{hand_{right}} = w_{hand_{left}} = 10$ to make that the hands stay in contact with the walker handles. $w_i = 1$ for the other joints. These data are fitted to the current positions of joints estimated from a prediction model.

3.2. *Human posture prediction model*

A biomechanical model of for human walking was proposed by Brubaker[5] to track people from the video stream being processed by a particle filter. It is shown that this approach increases the robustness of the person tracking algorithm, particularly in cases of occlusions. Here we use a model for the human whole-body motion which reflects his dynamic equilibrium when walking to predict the position of the body system at the time of capture and increase the precision of body part fitting process. A Kalman filter is first used to estimate spatio-temporal parameters in the case of assisted walking by a walker: the duration and stride length. This estimation allows to predicts future positions of the feet and therefore the reference trajectory of the Center of Pressure (CoP) from a walking pattern generation using preview control of the Zero-Moment Point (ZMP)[6,7] which coincides with the CoP. It is from this characteristic point that the speed characteristic of the Center of Mass (CoM) and the speed of the foot which is not in contact with the ground (the swing phase) can be calculated. Then, the 3D positions of the joints at the time of capture $k+1$ can thus be predicted using an algorithm which controls the joint positions of virtual mannequin by a LQP program taking into account especially and the constraints imposed by the contacts with the environment (i.e. hands and feet).

4. Experimental results and discussion

In Figure 6, mean values and standard deviations of errors (between a motion capture system Codamotion and the embedded device with two Kinects) are calculated and shown in Table 1. Note that the error values for the upper body parts are around 3 *cm* and values for the lower body parts are around 6 *cm*. As stated in,[8] an approximate location of the joints of the body (with the same order of magnitude noise) can achieve good classification accuracy (80%) for a recognition system gait abnormalities.



(a) right side

(b) left side

Figure 6. The average errors for different joints

Table 1. The mean values and standard deviations of the errors

Joints		Mean (mm)	Standard deviations (mm)
Ankle	Right	67	22
	Left	69	26
Knee	Right	54	28
	Left	59	29
Hip	Right	11	1
	Left	11	1
Elbow	Right	24	15
	Left	26	9
Shoulder	Right	28	13
	Left	33	8
Wrist	Right	27	16
	Left	29	13

Bibliography

1. J. Gutiérrez, F. V. Medina and M. A. Porta-Gándara, Vertically aligned accelerometer for wheeled vehicle odometry, *Mechatronics* 2010.
2. D. A. Winter, *Biomechanics and motor control of human movement* (John Wiley & Sons,Inc., 2009).
3. R. Mozul, P.-B. Wieber, R. Pissard-Gibollet, L. Boissieux and F. Billet, *HuMAnS*, tech. rep., INRIA, http://www.inrialpes.fr/bipop/software/humans/.
4. J. Salini, S. Barthélemy and P. Bidaud, Lqp controller design for generic whole body motion, in *Proceedings of the 12th International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines (CLAWAR)*, (Istanbul, Turkey, 2009).
5. M. A. Brubaker, D. J. Fleet and A. Hertzmann, Physics-based person tracking using the anthropomorphic walker, *International Journal of Computer Vision* 2010.
6. S. Kajita, F. Kanehiro, K. Kaneko, K. Fujiwara, K. Harada, K. Yokoi and H. Hirukawa, Biped walking pattern generation by using preview control of zero-moment point, *IEEE International Conference on Robotics& Automation* 2003.
7. P.-B. Wieber, Trajectory free linear model predictive control for stable walking in the presence of strong perturbations, *IEEE-RAS International Conference on Humanoid Robots* 2006.
8. B. Pogorelc and M. Gams, Discovery of gait anomalies from motion sensor data, in *22nd International Conference on Tools with Artificial Intelligence*, 2010.