Estimation of the Center of Mass with Kinect and Wii Balance Board
Alejandro González, Mitsuhiro Hayashibe, Philippe Fraisse

To cite this version:

HAL Id: lirmm-00734629
https://hal-lirmm.ccsd.cnrs.fr/lirmm-00734629
Submitted on 13 Nov 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Estimation of the Center of Mass with Kinect and Wii balance board

Alejandro González        Mitsuhiro Hayashibe        Philippe Fraisse

Abstract—Center of mass (CoM) trajectory is important during standing and walking since it can be used as an index for stability and fall prediction. Unfortunately current methods for CoM estimation require the use of specialized equipment (such as motion capture and force platforms) in controlled environments. This paper aims at applying the statically equivalent serial chain (SESC) method to obtain CoM position using widely available and portable hardware; a Microsoft’s Kinect and a Nintendo’s Wii balance board. During identification, CoM is approximated by CoP measurements and the virtual chain is created for able-bodied subjects. The result demonstrates that the SESC method can be applied outside the laboratory environment using a Kinect. Cross-validation of the identified model was performed to evaluate the accuracy of the method. Results obtained of five subjects are shown and discussed.

I. INTRODUCTION

Human motion stability while walking and standing is closely related to the motion of the center of mass (CoM). During quiet standing, static stability is achieved by maintaining the CoM’s projection inside the support polygon. Walking humanoid robots have achieved dynamic stability by controlling their center of mass in order to create a suitable center of pressure (CoP) or zero moment point (ZMP) [1], [2], [3] trajectory. The close relationship between ZMP/CoP and CoM can be observed in simplified models such as the cart-table model [4] and its equivalent, the linear inverted pendulum model (LIPM) [3].

Several techniques have been developed for CoM estimation in humans. Those most widely used are based on body segment data (center of mass position and its moments of inertia [5]). A subject’s parameters may be estimated using regression equations and a set of anthropometric data which have been obtained from experiments involving living subjects and/or cadaver studies. An example of such an anthropometric table was developed by de Leva [6]. One disadvantage of using this method is that large estimation errors can be made due to a mismatch between the subject and the used data [5], [6].

When a subject-specific estimation of body segment parameters is desired, a series of elaborate and costly processes may be followed. They usually involve the analysis of medical images such as magnetic resonance imaging (MRI) or computer tomography (CT) [5]. Venture et al. [7] estimated segment properties (mass and inertia) by using ground reaction forces and video motion capture (MoCap) system while solving the inverse dynamics of their human model.

Other methods also make use of data obtained by video MoCap and force platforms. For example, a clear relationship has been observed between the oscillations of the CoP and those of the CoM during quiet standing [8], [9], [10]. CoM can be determined from force plate measurements by the double integral method [11] where CoP is obtained from ground reaction forces. The CoM’s acceleration is estimated and then integrated to give position. This method can only be used when walking over an instrumented surface (or when wearing instrumented shoes); special care should be taken to find initial CoM position and velocity [11]. Betker et al. [12] used a genetic algorithm to determine the parameters of a sum of sines function for the CoM estimation.

The statically equivalent serial chain (SESC) developed by Espiau and Boulic [13] creates a virtual chain with an end-effector located on the subject’s CoM. The parameters of each link in the SESC of a human subject may be obtained after identification, using CoP and segment orientation data. Once this is accomplished, only limb orientations are required to estimate CoM position during the subject’s motions. This estimation can be made in real-time. The technique developed previously requires the data to be gathered in controlled environment [14].

In general, CoM estimation makes use of expensive equipment with a large set up time (typically, video MoCap requires a large number of cameras and a long marker placement process). Portable tools should be pursued to conveniently assess a patient’s balance and track his improvement during home rehabilitation. Off-the-shelf products are being exploited to this effect; such as the Wii balance board as a possible force plate substitute [15], [16], or the Kinect for 3D reconstruction and as a rudimentary MoCap system capable of performing angular measurements [17].

The work presented here is focused on locating a subject’s CoM by building a statically equivalent serial chain (SESC) with portable devices. The Kinect camera was used to obtain limb orientations to drive the SESC, while the Wii balance board is used during identification for reading the subject’s CoP position. After identification, CoM can be estimated from angular measurements alone. Cross-validation was performed by comparing CoM measurements to estimated the CoM projection in the ground plane.

II. STATICALLY EQUIVALENT SERIAL CHAIN (SESC)

A. SESC model

It is possible to locate the center of mass of any linkage by means of a serial chain. In this chain, each link is represented by a set of constants determined from the geometric configuration and mass distribution of the original link. Cotton et
The frame’s origin measured from the origin of link’s CoM (A) on the link’s frame. The total mass is given by \( \Sigma m \) mass attached to it, represented by \( m_1 \) on the link’s frame. The total mass is given by \( \Sigma m = M \).

The homogeneous transformation between links is given by

\[
T_i = \begin{bmatrix} A_i & d_i \\ 0 & 1 \end{bmatrix}
\]

where \( A_i \) is a 3-by-3 matrix containing the relative orientation of the frame and vector \( d_i \) determines the position of the frame’s origin measured from the origin of link \( i - 1 \). The chain’s CoM (\( C_M \)) can be determined as follows:

\[
\begin{aligned}
C_M &= \begin{bmatrix} m_1 \end{bmatrix} T_1 \begin{bmatrix} c_1 \\ 1 \end{bmatrix} + \begin{bmatrix} m_2 \end{bmatrix} T_1 T_2 \begin{bmatrix} c_2 \\ 1 \end{bmatrix} \\
&+ \begin{bmatrix} m_3 \end{bmatrix} T_1 T_2 T_3 \begin{bmatrix} c_3 \\ 1 \end{bmatrix}
\end{aligned}
\]

After performing the necessary operations, (2) can be rewritten as

\[
C_M = d_1 + A_1 r_1 + A_1 A_2 r_2 + A_1 A_2 A_3 r_3
\]

where

\[
\begin{aligned}
r_1 &= \frac{1}{M} (m_1 c_1 + m_2 d_2 + m_3 d_3) \\
r_2 &= \frac{1}{M} (m_2 c_2),
\end{aligned}
\]

When dealing with revolute and spherical joints \( r_i \) is a constant vector which describes the static properties of the SESC. Equation (3) may also be written as the matrix multiplication

\[
C_M = \begin{bmatrix} I & A_1^* & \ldots & A_n^* \end{bmatrix} \begin{bmatrix} d_1 \\ r_1 \\ \vdots \\ r_n \end{bmatrix} = BR
\]

where \( A_i^* \) represents the absolute orientation of link \( i \). In our example \( A_1^* = A_1, A_2^* = A_1 A_2 \) and \( A_3^* = A_1 A_2 A_3 \). Matrix \( B \) is then of size 3-by-3(\( n+1 \)) and is composed of the individual rotation matrices. Vector \( R \) contains the segment properties of the SESC chain. It has been observed in [14], that (5) is akin to the direct geometric model of a serial chain with links described by \( r_i \). It is interesting to note that the sizes of \( R \) and \( B \) may vary depending on assumptions made regarding the original chain. For example, when dealing with planar mechanisms \( r_i \) is generally a 2-by-1 vector, reducing the overall size of the SESC.

**B. Identification**

Vector \( R \) may be obtained from a complete knowledge of the geometric and static parameters of the original chain [18]. These parameters may not always be exactly determined, such is the case when dealing with a human subject. It is possible to estimate the values of \( R \) by means of the Moore-Penrose pseudoinverse when a set of measurements in which both the CoM position and link orientations are known.

\[
\hat{R} = B^+ C_M
\]

An obvious difficulty is the need for a good CoM measurement. When not all three components of \( C_M \) are known, an estimation may be performed using the available data. To illustrate this, consider rewriting (5) as follows,

\[
\begin{bmatrix} C_{Mx} \\ C_{My} \\ C_{Mz} \end{bmatrix} = \begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix} R
\]

Any one row or combination of rows from (7) may be used for identification. To reduce identification errors, a large number of measurements (m) should be used. In the most general case, a set of \( (n+1) \) independent measurements which provide the \( x \), \( y \) and \( z \) components of the CoM are needed in order for \( B \) to be invertible. If, for example, only one component is known, a set of \( 3(n+1) \) measurements is required.

Human CoM ground projection may be approximated by the CoP in stable poses. This measurement provides us with 2 components of the CoM position and creates a linear system which may be solved, in the least squares sense, by the use of the pseudoinverse.

\[
\hat{R} = \begin{bmatrix} B_{x,1} & \vdots & B_{x,m} \\ B_{y,1} & \vdots & B_{y,m} \end{bmatrix}^+ \begin{bmatrix} C_{Mx,1} \\ \vdots \\ C_{Mx,m} \\ C_{My,1} & \vdots & C_{My,m} \end{bmatrix} = D^+ \begin{bmatrix} C_{Mx,1} \\ \vdots \\ C_{Mx,m} \\ C_{My,1} \\ \vdots \\ C_{My,m} \end{bmatrix}
\]
Consider the linear system

\[ Y = WX + \rho \]  \hspace{1cm} \text{(9)}

where \( Y \) is a vector of measurements, \( W \) is the configuration matrix of size \( r \times c \) and \( \rho \) represents the residual error. A least squares estimate aims to find a vector \( \hat{X} = W^+Y \) such that the Euclidean norm of \( \rho \) is minimized. The linear system described by (9) has \( r \) number of measurements, and \( c \) parameters to be estimated.

An analysis of the numerical sensitivity of this system leads to the use of the condition number of matrix \( W \) as an indicator of the identification quality [19]. It has been shown that a large condition numbers is linked to a greater numerical sensitivity of the solution. The best observation strategies utilize configuration matrices with a condition number close to one. For any given matrix, there exist several condition numbers, depending on the matrix norm used. In the Euclidean norm, it is equal to the quotient of the largest and the smallest singular values (\( \sigma \)) of the matrix.

\[
\text{cond}(W) = \frac{\sigma_{\text{max}}}{\sigma_{\text{min}}}
\]  \hspace{1cm} \text{(10)}

Khalil and Dombre [20] make use of the parameter relative standard deviation (\( \sigma_{\hat{X}_j} \% \)) as a metric for determining the quality of an identified parameter.

\[
\sigma_{\hat{X}_j} \% = 100 \frac{\sigma_{\hat{X}_j}}{|X_j|}
\]  \hspace{1cm} \text{(11)}

where

\[
\sigma_{\rho}^2 = \frac{\|Y - WX\|^2}{r - c},
\]

\[
C_\hat{X} = \hat{\sigma}_\rho^2 (W^T W)^{-1}, \sigma_{\hat{X}_j} = \sqrt{C_\hat{X}(j,j)}
\]  \hspace{1cm} \text{(12)}

\( \sigma_{\rho}^2 \) is an estimator of the standard deviation of the identification error. \( C_\hat{X} \) represents the variance-covariance matrix of the estimation error and \( \sigma_{\hat{X}_j} \) is the standard deviation of the \( j \)-th parameter. Whenever a value of \( \sigma_{\hat{X}_j} \% \) is larger than ten times the smallest one, that parameter is considered to be poorly identified. This is usually the case for small parameters with little influence in \( Y \).

\section*{III. SENSORS}

\subsection*{A. Kinect}

The Kinect sensor is used for detecting three-dimensional position of a subject’s joints and the orientation of her limbs. It works by fitting a skeleton over a depth map obtained with an IR camera by using a projected light pattern [21]. Its motion capture capabilities are used mainly for video games, where an avatar mimics the user’s motions and is even capable of interpreting a number of gestures. As an example, the FAAST project [22] can be used to interact with programs running on a PC using only the Kinect camera. We make use of the middleware provided by OpenNI-PrimeSense\textsuperscript{1} to connect to the Kinect and obtain the subjects' skeleton and joint positions.

The Kinect’s hardware and 3D reconstruction allow us to perform marker-less human tracking in home environments. Stone and Skubik [21] compared tracking results obtained with a Kinect to those obtained using a VICON system and gait reconstruction performed using two webcams. In their work, gait characteristics are evaluated to predict and prevent falls. They found a 2% difference in walking speed, and 1.5% for leg stride time when compared to VICON measurements. Kinect measurements have also been used to improve data obtained by other means, such as RGB cameras [23] or inertial sensors [17].

\subsection*{B. Wii}

Interest on the Wii as a rehabilitation tool has surfaced due to its accessible interface, capable of integrating Virtual Reality (VR) and postural biofeedback methods [15], [16], [24], [25]. Clark et al. [15] studied the validity of conducting postural tests with the Wii balance board. They compared it to standard force plate measurements, and noted its obvious advantages in portability and price. They found that the Wii balance board results were consistent to those obtained with a laboratory force plate, even if the mean displacements were larger for the Wii board and reported this to be “probably due to device specific factors”. The authors warn against the use of the board for fast, high-force measurements but suggest it to be a suitable device for standing balance assessments.

The balance board may be connected via Bluetooth to any capable computer. We communicate with the board by means of the wiiuse\textsuperscript{2} open source project.

\section*{IV. EXPERIMENTAL RESULTS}

\subsection*{A. Experimental setup and Data collection}

Five able-bodied subjects were asked to stand on top of a Wii balance board. They were instructed to hold a series of poses meant to affect the position of their CoM in the sagittal plane. Each pose featured a different configuration of ankle, knee and hip angles (see Fig. 2). Data was recorded simultaneously with a Wii balance board and a Kinect for SESC model identification. Even though we are interested only in the sagittal plane, the camera was placed in such a way as to capture the subjects’ frontal plane. This was done since the skeleton provided by OpenNI-PrimeSense\textsuperscript{3} seems to work best when the subject directly faces the camera.

As described before, both a set of limb orientations and CoM positions are needed for the identification procedure. The orientation of each body segment is calculated using joint positions while CoM is approximated from CoP measurements. All data was filtered using a second order Butterworth low-pass filter with a cut off frequency of 10Hz.

During identification of the SESC’s parameters it is necessary for all positions to be expressed on the same frame.

\textsuperscript{1}http://www.openni.org/

\textsuperscript{2}http://github.com/rpavlik/wiiuse
In order to correctly define CoP position, it was necessary to find the homogeneous transform which relates the Wii board’s local frame to that of the Kinect.

**B. SESC configuration**

A planar, serial chain with four links was chosen to represent the human model. All joints are assumed to be revolute, giving the model $4$ degrees of freedom. Additionally we make the assumption that the links’ centers of mass are located on the straight line connecting the joints (see Fig. 3). This allows us to have only scalar quantities for $r_i$.

The first link on the SESC can represent the position of the virtual chain’s origin in the balance board’s frame. For this model, distance $d_1$ can be seen as representing CoP position when the subject is standing upright. We consider it to be a scalar quantity constrained to the horizontal plane. In the vertical direction, the SESC’s origin is taken to be at the same height as the ankle joint.

Matrix $D$ is then formed as:

$$D = \begin{bmatrix}
1 & \cos(\theta_{1,1}) & \cos(\theta_{2,1}) & \cos(\theta_{3,1}) \\
\vdots & \vdots & \vdots & \vdots \\
1 & \cos(\theta_{1,m}) & \cos(\theta_{2,m}) & \cos(\theta_{3,m})
\end{bmatrix}$$  \hspace{1cm} (13)

**C. Data selection**

In order to have a valid CoM estimation, we should know its projection on the ground plane for a series of configurations. The Wii balance board, strictly speaking, is only able to determine the position of the subject’s CoP. We assume that CoP can be used as an estimate of CoM, only when the body’s acceleration is negligible (e.g., during quiet standing). This explains the preference/need for static poses [14], [18]. The subject’s CoP was observed for each pose. We compute the standard deviation of the CoP and choose a time window that minimizes it. This time window is chosen as a representative period where we can assume a static pose has been performed.

Poses used for identification were chosen manually, placing emphasis in numerical stability. For this reason, we tried to find a set of poses which would give a low condition number of $D$ while having low parameter relative standard deviations for $\hat{R}$. Only poses performed consecutively were considered.

**D. Cross-validation**

Fig. 4 shows the measured CoP and the estimated CoM projection of two subjects during a cross-validation test, that is to say, with poses not used for model identification. The root-mean-squared error ($\text{rmse}$) was calculated and is shown in Table I. The identified values for vector $\hat{R}$, which contains the SESC’s constant parameters, are also given in Table I. Different values of $\hat{R}$ are observed for each individual. This is expected due to the subject-specific morphology appearing in the SESC model. It is encouraging to observe that the maximum height of each subject’s CoM, computed from the identified model, roughly corresponds to half of her body height.

The parameter’s relative standard deviations are also given, since they can show the consistency of the identified parameters. We obtained a series of small, clustered values for this metric indicating confidence on the identified model. Condition number for each of the configuration matrices $D$ was also verified. As previously discussed, the condition number of the configuration matrix can give some insight regarding the sensitivity of the solution and as to the observability of the parameters for the identification poses. Their values range from 8 to 32; with the differences due, most likely, to the variation of poses to identify each subject’s parameters.
TABLE I

RESULTS OF THE PARAMETER ESTIMATION.

<table>
<thead>
<tr>
<th></th>
<th>Sb01</th>
<th>Sb02</th>
<th>Sb03</th>
<th>Sb04</th>
<th>Sb05</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_1$ [mm]</td>
<td>$\hat{R}$</td>
<td>$\sigma_{\hat{R}}$</td>
<td>%</td>
<td>$\hat{R}$</td>
<td>$\sigma_{\hat{R}}$</td>
</tr>
<tr>
<td>$r_1$ [mm]</td>
<td>377.30</td>
<td>2.67</td>
<td>277.44</td>
<td>11.96</td>
<td>338.26</td>
</tr>
<tr>
<td>$r_2$ [mm]</td>
<td>387.96</td>
<td>1.75</td>
<td>339.06</td>
<td>6.54</td>
<td>339.87</td>
</tr>
<tr>
<td>$r_3$ [mm]</td>
<td>110.80</td>
<td>4.02</td>
<td>172.48</td>
<td>6.47</td>
<td>188.33</td>
</tr>
<tr>
<td>cond($\mathbf{D}$)</td>
<td>32.84</td>
<td>16.24</td>
<td>8.97</td>
<td>10.63</td>
<td>19.27</td>
</tr>
<tr>
<td>rmse [mm]</td>
<td>14.12</td>
<td>31.94</td>
<td>13.00</td>
<td>14.82</td>
<td>37.82</td>
</tr>
</tbody>
</table>

Fig. 4. Cross validation of identified SESC models for 2 subjects.

Fig. 5. Estimation of CoM during sit-to-stand motion.

V. DISCUSSION AND FUTURE WORK

A measurement system, such as the one described and used here offers many advantages in cost, portability and reduced set-up time. These advantages originate from the use of off-the-shelf, widely accessible hardware and open source software for performing the measurements. Once the system is identified only limb orientation is needed to estimate CoM position in real-time and for all of the subject’s movements. In Fig. 5 we show tracking of a subject’s CoM performing a sit-to-stand motion, without CoP or ground reaction measurements.

The equipment used here was not developed with this application in mind and certain limitations can be listed. For example, the balance board presents a relatively small sensing area when compared to a conventional force plates. Maintaining balance while avoiding the limits of the board’s working area restricts the subject’s range of motion. The subject’s weight is another limiting factor. Being either extremely light or extremely heavy hinders proper registration of the force sensors, creating uncertainty on the CoP measurement. Kinect also comes with its unique set of limitations. Improper lighting, loose fitting clothes, and large objects which surround the subject can adversely influence the skeleton fitting. Any error in joint position measurements will also affect angular measurements.

Ease of transport and portability are great advantages for the system’s usage for in-home rehabilitation purposes, however the Wii balance board must be accurately registered in the Kinect’s frame when data is collected. In order to provide consistent measurements, repeatability of this localization is desired and should be improved in order to create a reliable system.

The estimation of the SESC’s parameters is essentially a geometric calibration problem. Their value is constant and is related to the subject’s anthropometry; thus once the model is established, it can also be used for estimation during dynamic motions. We made use of static poses during cross-validation to keep the assumption that CoP approximates CoM projection. This is not the case for fast motions or when CoM is accelerated. Further development should be pursued aiming to a more systematic identification procedure. For example, a good set and number of identification poses may be defined using optimization in order to minimize cond($\mathbf{D}$) [20].

A smaller rmse was reported by Cotton et al. during their study in an elderly population using the SESC method,
identified from video MoCap and force plate data [18]. This was likely due to their use of a larger set of poses during identification and the better sensing quality of the high-end equipment. Ideally, to reduce errors, a larger number of distinct poses with clearly differentiated angular values and CoP displacements would be used for identification. A different approach could be implemented in the form of a Kalman filter for on-line estimation of the parameters. It is possible to imagine an interactive, game-like interface that would prompt the subject for new poses until convergence is achieved.

VI. CONCLUSION

Experimental results indicate that it is possible to estimate a subject-specific CoM position using cheap, widely accessible equipment like the Kinect and Wii balance board. Here we validate the CoM estimation performance using the SESC method in two dimensions. In this paper, different postures in the sagittal plane were given for evaluation. However, the proposed method is already suitable for 3D CoM estimation. In the 3D case a larger number of joints and limbs are needed to establish a model with the appropriate motion range. This increases the number of poses required for identification.

With an inexpensive set-up and a fast calibration session, marker-less CoM tracking could be achieved. Typically a recording session lasts 30-45 minutes, including time needed for setting up the equipment. Sessions can take place in any room with enough open space for the subject to move freely. Once the subject’s SESC parameters have been identified, there is no need for CoP measurement. CoM can be estimated from limb orientations alone. It should contribute for in-home rehabilitation, providing a quantitative balance evaluation.

The estimation accuracy obtained with the Kinect and Wii board was lower than that obtained in a controlled environment using video MoCap and force platform reported in [18]. However, qualitatively, the error scale we found was similar to the one obtained in a controlled environment. Nonetheless, the accuracy is always dependant on the dataset used during identification. Further investigation is necessary to perform a strict comparison in equal conditions. We intend to perform synchronous measurements with both MoCap/force plate, and Kinect/Wii board to obtain the same set of postural data.

This work provides a portable tool for CoM estimation. The flexibility obtained by leaving the laboratory shows potential for use in home rehabilitation, motion stability assessment and balance training, quantitatively tracking the subject’s progress. Also a marker-less alert system for fall prevention could be implemented.

REFERENCES


