Time-Based Identification of Human Ankle Impedance via Microsoft Kinect

Mauricio E. Segura, Enrique Coronado, Antonio Cardenas
Facultad de Ingeniería
Universidad Autónoma de San Luis Potosí,
San Luis Potosí, México, 78290
Email: {mauricio.segura, luis.coronado, antonio.cardenas}@uaslp.mx

Davide Piovesan
Biomedical Engineering Program
Gannon University
Erie, Pennsylvania, 16541
Email: piovesan001@gannon.ed

***Abstract*—This paper presents an affordable platform to estimate human ankle mechanical impedance. This platform uses Microsoft Kinect version 2 as the motion capture system of choice and a hybrid algorithm to estimate the biomechanical parameters. The algorithm is based on the combination of an Extended Kalman filter and a Genetic Algorithms. The information provided by Kinect together with the ankle biomechanical parameters can be utilized to estimate the dynamic behavior of the recovery from falls. To prove the precision of the 3D measurements obtained with Kinect a comparison with a visual system, based on two industrial cameras was performed. Both systems were calibrated tracking the end-effector position of an industrial robot. The hold and release (H&R) experimental paradigm was used to estimate the ankle mechanics on seven subjects. The results show that Kinect v2 is a reliable motion capture device to study the neuro-mechanical response of recovering from falling.**

I. Introduction

Motion Capture has become a vital tool to study human performance. Examples of applications in rehabilitation analysis, gait and running performance analysis are shown in [1], [2].

Because of its speed, accuracy and flexibility, vision based systems are a regular option chosen for motion analysis. This systems are composed of a well suited cameras arrangement, visual markers (passive or active) and a central unit that process the information and displays it. However, this systems require well controlled environments or professional illumination systems, increasing the overall cost. A detailed description of the advantages and disadvantages of the main types of motion capture systems are presented in [3].

The rapid development of new technologies presents new alternatives to classical motion capture systems. These new methods, could offer appropriate accuracy to be used in the clinical field with the advantage of being cheap and portable.

Microsoft Kinect is a low-cost motion capture used recently in applications in the field of robotics and biomechanics [4]–[7]. Two version of this device have been commercialized so far. The sensing system of version one (v1) consists of the combination of an infrared laser emitter and an infrared camera providing a color and depth image (RGB-D). This system works on the principle of structured light using patters of dots. The dots are recorded by the infrared camera and compared with a known pattern to create a depth map. The position of the body can be inferred from the depth map using machine learning algorithms. Kinect version 2 (v2) for Xbox One uses infrared time-of-flight to measure depth. This provide greater depth accuracy and camera resolution than the original Kinect sensor. The main limitation of both version is the limited distance at which the target can be acquired and poor performance with ambient light [8], [9].

Studies of the validity of the Microsoft Kinect for different biomechanical tasks reported good results [10]–[13]. Recently Yeung et al. in [14] evaluated the Kinect as a clinical assessment tool of body sway, making a comparison of the Kinect accuracy against a Vicon motion capture system. The reported body center of gravity (CoG) positional error was of 4.38 mm on average.

Postural stability affects the capacity to compensate the movement of the body's CoG, caused by unexpected perturbations. Stability problems in the human body can hinder the performance of daily activities and make an individuals more prone to falls. The estimation of human ankle's mechanical impedance is an important tool used to gain insight on the interaction between biomechanics and the neural correlates proper of the control of postural stability [15]. The Hold and Release (H & R) paradigm [16] is a technique designed to perturb the quiet standing of an individual with the scope of exciting a neuro-mechanical response against falling. This technique have been used to calculate ankle's mechanical impedance in [17], [18] using low-end web-cameras as a motion capture system. While the apparatus proved very effective the scene preparation and camera calibration were time consuming.

This paper presents an affordable platform to estimate human ankles mechanical impedance using the (H & R) paradigm and Microsoft Kinect v2. A Kinect sensor v2 was used for measure the 3D position of the ankle and the approximate location of the center of gravity in human subjects. The advantage of using this sensor for biomechanical analysis is the possibility to obtain detailed 3D measurements without a time consuming process of camera calibration. Moreover, the device is non intrusive, portable and affordable ($200 USD) for low-budget research projects.

To prove the precision of the 3D measurements obtained with the Kinect a comparison with the visual system described in [18] was performed. The visual system was based on two industrial cameras calibrated with the aid of an industrial robot.

The impedance parameters of the ankle were estimated together with the system states using a hybrid time-based identification algorithm. The identification technique entails a Genetics Algorithm that uses an Extended Kalman Filter as cost function [20].

II. Muscle-tendon mathematical model

The neuro-mechanical model responsible for maintaining stability during bipedal quiet standing can be represented mathematically by nonlinear state-space equations, representing the interaction between muscle-tendon units (MTU) mechanical parameters (e.g. stiffness and damping) and the inertia of the whole body.

For this a second-order Kelvin-Voigt (KV) model is usually preferred for its low number of parameters and simplicity of the control [23]. In this representation the body is modeled as a single mass positioned in the center of gravity (CoG), rotating at a fixed distance l from the ankle. The rotational axis of the ankle is assumed fixed in space. The skeletal model is actuated by an equivalent MTU model. This equivalent model represents the combined action of all active MTUs around the ankle joint.

Figure 1 shows the representation of this biomechanical model as a simple inverted pendulum . For small oscillations, this model is expressed by the following equation [18]:

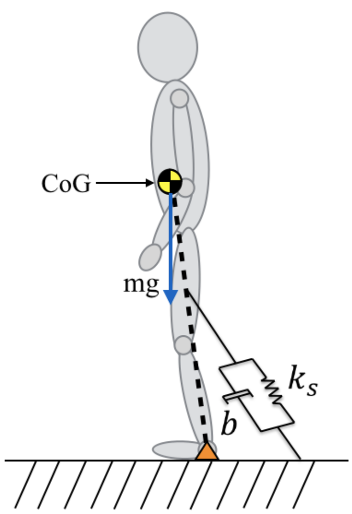

Fig. 1. Graphical representation of the second order muscle-tendon system

$$I_h\ddot{\theta} + b\dot{\theta} + k_s\theta = 0 \tag{1}$$

where θ is the relative angular displacement of the CoG with respect to the vertical, b is the muscles' damping coefficient, m is the mass of the subject, l is the radius of gyration of the subject, $I_h = ml^2$ is the inertia of the whole body with respect to the ankle, and g is the acceleration of gravity. The term $k_s = k - mgl$ is the apparent stiffness of the system given by the difference of the ankle stiffness k and the destabilizing effect of the acceleration of gravity [18]. Equation (1) can be express in state space as follows [17]:

$$\dot{x} = \begin{bmatrix} 1 & 0 \\ k_s/I_h, & -b/I_h \end{bmatrix} \begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix} \tag{2}$$

To estimate the muscle-tendon parameters of the aforementioned model a system able to track the continuous change in the state θ is necessary. The Kinect v2 was used to fulfill such requirement.

III. Hybrid algorithm for parameters estimation

To estimate MTU's parameters from the data obtained by the Kinect a hybrid approach that includes the combination of a Genetic algorithm (GA) and hybrid Extended Kalman Filter (EKF) was used.

Genetic algorithms are an approximation of how Nature performs optimization. This algorithm searches globally a set of parameters that represent possible solutions to a problem. The parameters to find are encoded in chains of binary numbers called chromosomes, each chromosome represents an individual. An initial set of random individuals (populated with a random sequence of 0s and 1s) are initialized in a population. The objective of each GA's iteration is to evolve the population and find a more optimal solution to the problem. To asses how good a possible set of parameters are when compared to others a fitness function is required [19]. The fitness function employed in a GA is problem dependent and the fitness value is the objective to optimize.

In the proposed approach, detailed in [20], the unknown MTU's parameters of the model are the output of the GA. The parameters are "mutated" at each iteration of the GA until the response of a second-order biomechanical state model estimated using a hybrid EKF match closely with the measured outputs from the Kinect.

The Extended Kalman Filter is a recursive filter used broadly for state estimation of non linear systems. Hybrid EKFs are often used in applications where a continuous-time dynamics is modeled but the system measurements are obtained at discrete instants of time.

The fitness function used within the GA is the sum of the least squares residuals (RSS) between the measurement of the Kinect and i hybrid EKF estimation. This fitness function is denoted as:

$$fitness = \frac{1}{\sum_0^n (\mathbf{z}_i - H\hat{\mathbf{x}}_i)^2} \tag{3}$$

where i is the number of sample, $\mathbf{z}_i$ the noisy measurements and $H\hat{\mathbf{x}}_i$ the EKF estimated value.

IV. EXPERIMENTAL SETUP

A. Vision systems set-up

A vision system composed of two industrial uEye cameras with resolution of 640×480 were used to benchmark the 3D measurements of the Kinect. The cameras and Kinect were situated approximately $1.8m$ and $2.3m$ from the center of the industrial robot workspace, respectively.

1) Camera-based system calibration: the calibration process is required to estimate the parameters of a camera model that relates points in 3D space to their projections in camera space. One of the most used model, the pinhole model, gives the following equation:

$$\tilde{q} = P\tilde{Q}$$

$$\begin{bmatrix} wx_i \\ wy_i \\ w \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & 1 \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix} \tag{4}$$

where P is the 3D to 2D projection matrix, $\tilde{q} = [\omega x_i, \omega y_i, \omega]^T$ is the point in the image, $\tilde{Q} = [X_i, Y_i, Z_i, 1]^T$ is the point in the world reference frame and ω is a scale factor [21]. For convenience equation (4) can be expressed as:

$$\begin{bmatrix} \mathbf{X}_i & \mathbf{0} & -xc_iX_i & -xc_iY_i & -xc_iZ_i \\ \mathbf{0} & \mathbf{X}_i & -yc_iX_i & -yc_iY_i & -yc_iZ_i \end{bmatrix} \mathbf{p} = \begin{bmatrix} xc_i \\ yc_i \end{bmatrix} \tag{5}$$

To calculate the parameters vector $\mathbf{p} = [p_{11}, p_{12} \ \cdots \ p_{33}]^T$ at least $m = 6$ corresponding points $(i = 1, 2, ..., 6)$ are required. If more than 6 points are used to compute $\mathbf{p}$, an over constrained system is obtained and a minimization process is required.

For precise calibration of the visual system an industrial robot FANUC LR Mate 200iC with a visual feature attached to the end-effector was used (figure 2). During the calibration process the robot follows a pre-planned trajectory of 30 points spread in approximately $500 \times 500 \times 500mm^3$. For each calibration point the cameras obtain 2D coordinates of the visual feature in each camera space. Then, the camera parameters are obtained from equation (5). The section of space chosen for calibration was appropriately chosen so to encompass the CoG locations of all the subjects during the hold and release experiments.

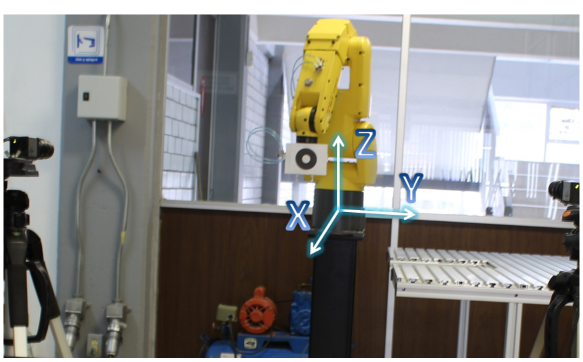

Fig. 2. Calibration of the visual system

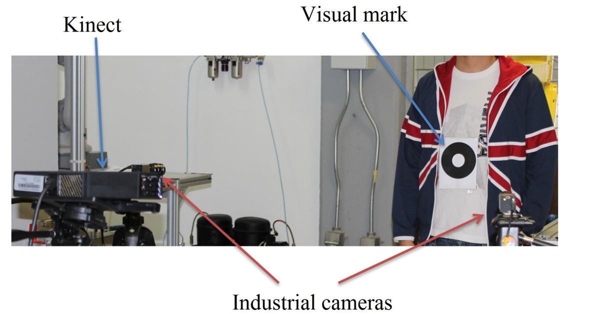

Fig. 3. Locations of the visual system, the visual markers and Kinect

2) Position estimation of a 3D visual marker with visual system and Kinect: A visual marker was attached to each subject to compare the measurements between the camera-system and the Kinect (Figure 3).

Estimation of the 3D coordinates of the visual marker attached to the subject's body was performed using pixels coordinates from the corresponding image point and the current view parameters of the participating cameras. This was accomplished by minimizing the function:

$$F(X_i, Y_i, Z_i) = \sum_{j=1}^{nc} \left\{ \left[x_{ci}^j - h_x(X_i, Y_i, Z_i; \mathbf{p^j}) \right]^2 + \left[y_{ci}^j - h_y(X_i, Y_i, Z_i; \mathbf{p^j}) \right]^2 \right\} \tag{6}$$

where h_x, h_y are the functions representing the pinhole camera, nc is the number of participating cameras, x_i^j, y_i^j are the image coordinates in pixels of the visual marker in the j_{th} camera and X_i, Y_i, Z_i represent the corresponding 3D coordinates.

B. Hold and release experiment

The hold and release paradigm (H&R) was used to elicit a response of the subjects' neuro-mechanical system to a step perturbation, thus inducing a recovery from a fall. The subject stands in a natural posture while a constant pressure to the subject sternum is applied and the subject tries to resist it. Hence, a sudden release triggers the reaction to control the recovering from a fall while the position of the CoG is measured by the Kinect. The description of the paradigm is detailed in [16]. The proposed method of parameter estimation was applied on 7 human volunteers. Table I reports the physiological data of each subject.

C. Software and hardware

Two Graphical User Interfaces (GUIs), one for motion capture using Kinect v2 and a second for MTU's parameter estimation were developed using open source tools. The motion capture interface was developed in C# using the second version of Kinect software development kit (SDK 2.0).

TABLE I
DATA OF THE SUBJECTS

Subject	$m(Kg)$	$l(m)$	$I_h(Kgm^2/rad)$	Age
1	60	0.98	57.62	25
2	66	1.01	67.32	26
3	54	1.06	60.67	24
4	110	1.00	110.0	25
5	76	0.99	74.48	25
6	90	1.13	114.92	26
7	75	1.10	90.75	26

TABLE II
DATA OF THE SUBJECTS

Sub-ject	k_s			b		
	$Test1$	$Test2$	$Test3$	$Test1$	$Test2$	$Test3$
1	1207.60	941.82	1151.04	85.17	52.00	116.11
2	1163.52	1079.58	1148.93	301.74	230.48	271.15
3	867.93	894.44	732.13	181.99	180.66	160.15
4	1879.25	1587.09	1932.24	73.46	24.01	53.84
5	1249.10	1198.10	1278.48	239.10	296.25	276.48
6	1938.44	2438.43	2036.28	598.71	498.87	531.48
7	1093.70	1126.51	1019.32	138.75	223.69	192.58

The motion capture interface was used mainly for body tracking to obtain the angle between the right ankle and an approximation of the subjects' center of mass position in 3D. A PC DELL XPS 15 L521X with Intel i7-3632QM to 2.20 GHz and 16 GB of RAM was used to acquire the images from Kinect and processing the angle estimation algorithm.

The MTU's parameter estimation interface was developed in Python. The parameter estimation algorithms were developed using NumPy (Numerical Python library) and DEAP (Distributed Evolutionary Algorithms in Python) [22]. A PC Alienware X51 with Intel i7-3770 3.4 GHz and 8 GB of RAM was used for the processing of the parameter estimation algorithms. A Python script was developed to calibrate the visual system using OpenCV for image processing.

V. RESULTS

A. Comparison of Kinect with a visual system

To calibrate the camera-based system a "leave one out" approach was used. The estimations of the vector $\mathbf{p} = [p_{11}, p_{12} \cdots p_{33}]^T$ were performed using $n-1$ positions of the robot's end-effector and the remaining n end-point position was estimated. Thus the estimated 3D reconstruction was compared to the "real value" imposed by the robot. These approach was repeated for each of the 30 calibration point imposed by the robot. After calibrating the camera-based visual system, the obtained average errors in the 3D reconstruction were 1.109, 0.63 and 0.3140 millimeters for the axes x, y and z, respectively.

After calibration of the visual system, the 3D position of a visual marker attached to the body was acquired using both the visual system and the Kinect, simultaneously. The movement of the marker was 222.5mm along x, 11.0mm along y, and 32.1mm along z. The difference of the position of the marker during the movements estimated by the Kinect and the visual system were of $(6.1, 2.0, 1.3)$ millimetres.

B. Parameter estimation results

The H&R paradigm was applied 3 times to each subject. The estimates of the MTUs parameters for each subject using the hybrid algorithm described in section III on the Kinect data are reported in Table II.

For subject 1, figures 4, 5 and 6 show a comparison between the data obtained by the Kinect and the EKF state-estimate obtained using the MTU's parameters form the final iteration of the GA in the dynamics model.

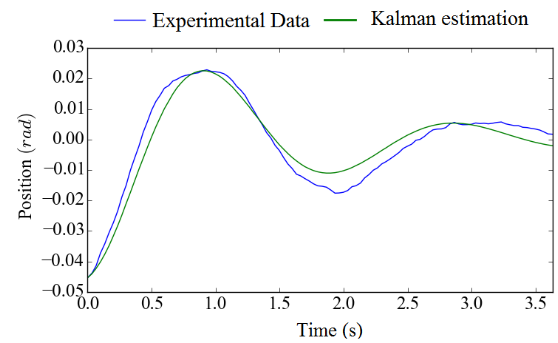

Fig. 4. Trial 1

VI. DISCUSSION AND CONCLUSIONS

This work presented the development of a low-cost motion capture platform based on Microsoft Kinect for the study of the human neuro-mechanical system responsible for maintaining stability during bipedal standing. Following the hold and release paradigm (H&R), experiments where performed on 7 subjects to estimate the impedance parameters proper of the ankle. In this experiments a Kinect v2 was used to track the movements of the subjects' CoG in the ankle joint space. Furthermore, a hybrid algorithm based in EKF and GA was used to estimate the neuro-mechanical properties (stiffness and damping) of the ankle. The information from the joint angular displacements together with the estimation of stiffness and damping at the ankle could be utilized to estimate the dynamic behavior of the recovery from falls.

The data obtained with Kinect v2 in the experiments were less noisy that the data presented in [18] that use web-cams as another low-cost platform of motion capture. This facilitated visualization and calculations of stiffness and damping parameters. Compared to a visual system based on industrial cameras, the Kinect v2 can obtain a reconstruction of the 3D position, within 7 mm. Even though less accurate the Kinect does not require a time consuming preparation phase (cameras calibration and adapting scene illumination) which was necessary as for the previously presented vision systems [17], [18] and some other commercial alternatives. It should be noted that the calibration of the visual system does not need to be performed for every estimation as long as the cameras are not moved from their position. On the other hand, using

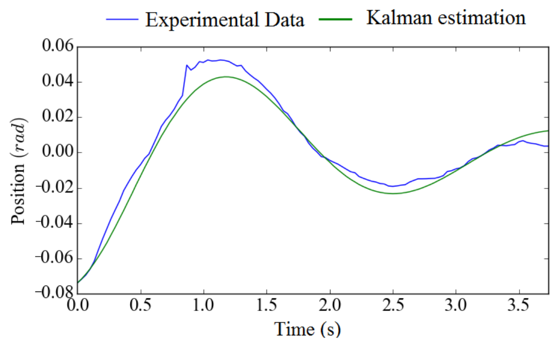

Fig. 5. Trial 2

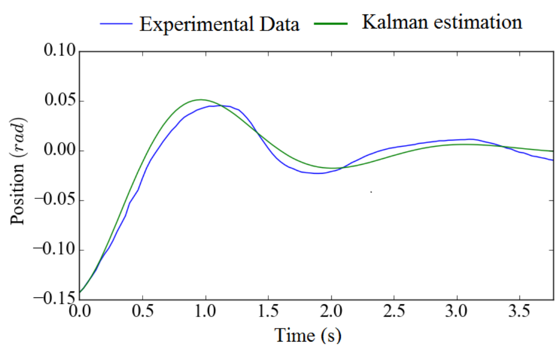

Fig. 6. Trail 3

the Kinect the experiments can be easily performed to other location without the need of a calibration.

The results shown that the Kinect v2 can be a convenient motion capture device to clinically assess the neuromechanical response against falling.

Acknowledgment

This work was partially funded by the PROMEP / 103.5 /13/9353 to support the integration of thematic networks of academic collaboration; by the National Council for Science and Technology of Mexico (CONACYT) using the fund to support the proposal 246 573, and by Gannon University's faculty research grant.

References

[1] S. Noiumkar and S. Tirakoat, "Use of optical motion capture in sports science: A case study of golf swing," in *Informatics and Creative Multimedia (ICICM), 2013 International Conference on*, Sept 2013, pp. 310–313.

[2] S. Corazza, L. Mndermann, E. Gambaretto, G. Ferrigno, and T. Andriacchi, "Markerless motion capture through visual hull, articulated icp and subject specific model generation," *International Journal of Computer Vision*, vol. 87, no. 1-2, pp. 156–169, 2010. [Online]. Available: http://dx.doi.org/10.1007/s11263-009-0284-3

[3] H. Zhou and H. Hu, "Human motion tracking for rehabilitationa survey," *Biomedical Signal Processing and Control*, vol. 3, no. 1, pp. 1 – 18, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1746809407000778

[4] L. V. Calderita, J. P. Bandera, P. Bustos, and A. Skiadopoulos, "Model-based reinforcement of kinect depth data for human motion capture applications," *Sensors*, vol. 13, no. 7, p. 8835, 2013. [Online]. Available: http://www.mdpi.com/1424-8220/13/7/8835

[5] J. Rosado, F. Silva, V. Santos, and Z. Lu, "Reproduction of human arm movements using kinect-based motion capture data," in *Robotics and Biomimetics (ROBIO), 2013 IEEE International Conference on*, Dec 2013, pp. 885–890.

[6] A. Fernández-Baena, A. Susin, and X. Lligadas, "Biomechanical validation of upper-body and lower-body joint movements of kinect motion capture data for rehabilitation treatments," in *Intelligent Networking and Collaborative Systems (INCoS), 2012 4th International Conference on*, Sept 2012, pp. 656–661.

[7] R. El-laithy, J. Huang, and M. Yeh, "Study on the use of microsoft kinect for robotics applications," in *Position Location and Navigation Symposium (PLANS), 2012 IEEE/ION*, April 2012, pp. 1280–1288.

[8] R. A. El-laithy, J. Huang, and M. Yeh, "Study on the use of microsoft kinect for robotics applications," in *Position Location and Navigation Symposium (PLANS), 2012 IEEE/ION*. IEEE, 2012, pp. 1280–1288.

[9] T. Butkiewicz, "Low-cost coastal mapping using kinect v2 time-of-flight cameras," in *Oceans-St. John's, 2014*. IEEE, 2014, pp. 1–9.

[10] B. F. Mentiplay, R. A. Clark, A. Mullins, A. L. Bryant, S. Bartold, and K. Paterson, "Reliability and validity of the microsoft kinect for evaluating static foot posture," *J Foot Ankle Res*, vol. 6, no. 1, p. 14, 2013.

[11] R. A. Clark, Y.-H. Pua, K. Fortin, C. Ritchie, K. E. Webster, L. Denehy, and A. L. Bryant, "Validity of the microsoft kinect for assessment of postural control," *Gait & posture*, vol. 36, no. 3, pp. 372–377, 2012.

[12] R. A. Clark, Y.-H. Pua, A. L. Bryant, and M. A. Hunt, "Validity of the microsoft kinect for providing lateral trunk lean feedback during gait retraining," *Gait & posture*, vol. 38, no. 4, pp. 1064–1066, 2013.

[13] V. Gritsenko, E. Dailey, N. Kyle, M. Taylor, S. Whittacre, and A. K. Swisher, "Feasibility of using low-cost motion capture for automated screening of shoulder motion limitation after breast cancer surgery," *PLoS ONE*, vol. 10, no. 6, p. e0128809, 06 2015. [Online]. Available: http://dx.doi.org/10.1371%2Fjournal.pone.0128809

[14] L. Yeung, K. C. Cheng, C. Fong, W. C. Lee, and K.-Y. Tong, "Evaluation of the microsoft kinect as a clinical assessment tool of body sway," *Gait & posture*, vol. 40, no. 4, pp. 532–538, 2014.

[15] M. M. E. Menz, H. B. and S. R. Lord, "Foot and ankle characteristics associated with impaired balance and functional ability in older people."

[16] S. Bortolami, P. DiZio, E. Rabin, and J. Lackner, "Analysis of human postural responses to recoverable falls," *Experimental brain research*, vol. 151, no. 3, pp. 387–404, 2003.

[17] R. Chavez-Romero, A. Cardenas, and D. Piovesan, "Viscoelastic properties of the ankle during quiet standing via raster images and EKF," *IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, pp.1-5, 13 Dec. 2014

[18] R. Chavez-Romero, A. Cardenas, J. M. Rendon-Mancha, K. M. Vernaza and D. Piovesan, "Inexpensive vision-based system for the direct measurement of ankle stiffness during quiet standing," *ASME. Journal of Medical Devices*, vol. 9, no. 4, 2015.

[19] J. H. Holland, "Genetic algorithms," *Scientific american*, vol. 267, no. 1, pp. 66–72, 1992.

[20] E. Coronado, , R. Chavez-Romero, M. Maya, A. Cardenas, and D. Piovesan, "Combining genetic algorithms and extended Kalman filter to estimate ankle's muscle-tendon parameters," *ASME 2015 Dynamic Systems and Control Conference (Accepted)*, 2015.

[21] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge university press, 2003.

[22] F.-A. Fortin, F.-M. De Rainville, M.-A. Gardner, M. Parizeau, and C. Gagné, "DEAP: Evolutionary algorithms made easy," *Journal of Machine Learning Research*, vol. 13, pp. 2171–2175, jul 2012.

[23] D. Piovesan, A. Pierobon, P. DiZio, and J. R. Lackner, "Measuring multi-joint stiffness during single movements: numerical validation of a novel time-frequency approach," *PloS one*, vol. 7, no. 3, p. e33086, 2012.