Available online at www.sciencedirect.com
ScienceDirect
Procedia Computer Science 151 (2019) 194–201
Procedia Computer Science
www.elsevier.com/locate/procedia

The 10th International Conference on Ambient Systems, Networks and Technologies (ANT)
April 29 – May 2, 2019, Leuven, Belgium

An Open Vibration Platform to Evaluate Postural Control using a Simple Reinforcement Learning Agent

Virgile Lafontaine[a], Kevin Bouchard[a]*, Jean-Michel Gagnon[b], Mathieu Dallaire[b], Sébastien Gaboury[a], Rubens A. da Silva[b], Louis-David Beaulieu[b]

[a]*Laboratoire d'Intelligence Ambiante pour la Reconnaissance d'Activités (LIARA)*
[b]*Laboratoire de recherche en biomécanique & neurophysiologique en réadaptation neuro-musculo-squelettique (BIONR)*
Université du Québec à Chicoutimi, 555 boulevard de l'Université, Chicoutimi G7H2B1, Canada

Abstract

In this paper, our research team proposes an inexpensive open vibration platform built from easily available electronic components to be used as a tool by physiotherapists in order to help them in their evaluation of the postural control of individuals at risk of postural imbalance which could lead to falls. The platform has been thought to be easily reproducible and all the code necessary to make it work is made available on the researchers' websites. In addition, a simple reinforcement learning agent has been developed and tested to automatically calibrate the vibration motors for optimal stimulation. Finally, we present in this paper pilot experiments done on 7 healthy participants (40.8 years old) to validate the proper functioning of the platform.

Keywords: Sensors; Vibration motors; Postural control; Reinforcement learning; Technology for health;

1. Introduction

The aging process of the human body is natural, but can introduce early changes on the biological system including psychologic and physiologic alterations. The physiological factors are known to affect the physical capacities and the

* Corresponding author. Tel.: +1-418-545-5011 5063.
E-mail address: Kevin.Bouchard@uqac.ca

1877-0509
10.1016/j.procs.2019.04.029

functional independence of the geriatric population [1]. One of the most important physical changes is in the postural control. It is why it has been widely studied in the literature in the past decades from balance and falls risk prevention. For instance, it has been documented that postural imbalance in post-stroke patient may have detrimental effects on their functionality and autonomy, and may ultimately lead to a decrease in their quality of life as well as to an increase in their risk of falls [2]. In the general elderly population, over time, the cutaneous mechanoreceptors located under the feet progressively lose their sensitivity [3]. These mechanoreceptors are important contributors to the postural control. In fact, they desensitize and become less likely to detect small pressure variation while the person is standing or walking. This phenomenon, known as peripheral neuropathy, is exacerbated by several conditions or diseases associated with aging. For instance, the persons suffering from chronic type 2 diabetes are especially afflicted by this condition and they are more at risk to fall during their daily life activities [4]. Testing the postural control of the elders and the population at risk of falling could, therefore, contribute to early detection of the degradation of the mechanoreceptors. Moreover, it could help in reducing the dreadful consequences of falls, both in human lives and costs [5].

Overall, postural control is quantified by technological tools in the scientific literature [6], [7]. One valid, reliable and precise method is to securely place the patient standing on a force platform to track his center of pressure (CoP) [8][9]. This evaluation method can be used to detect postural anomalies from the patterns of CoP traced by the force platform. Once calibrated, a force platform produces real-time data regarding the movement of CoP in anteroposterior and mediolateral directions. The data can be used to compare a baseline, obtained by testing normal subjects, to the patient being evaluated. If, for instance, the center of pressure moves more than the average healthy patient, then the clinician can deduce that there might be a problem in postural control. The force platforms exist in several variations that all need their independent evaluation, but in general, the scientific literature tends to demonstrate that the use of these tools is an effective method to assess a postural anomaly control so that to establish better clinical diagnostic for rehabilitation prescription when the patients are of concern [10].

On the other hand, one of the limitations of relying simply on the force platform is that the person simply assumes a standing posture which may not shows the important information unless the postural control is starting to severely deteriorate. In many cases, for early detection, it is better to stimulate imbalance than to measure one's ability to keep a standing position. Unfortunately, there is no known clinical tool to stimulate imbalance on a force platform. In the scientific literature, there are a few projects that used various method to stimulate imbalance. For instance, Toosizadeh et al. [11] installed mechanical vibrators on the calves of patient to produce imbalance, and Roll et al. [12] used small vibrators under the feet to test the sensitivity of the cutaneous mechanoreceptors. The approaches in the literature generally suffer either from being impossible to reproduce due to lack of technical details on the material involved, or from using proprietary platform that is often expensive and does not provide an API.

In this paper, a new open vibration platform to stimulate cutaneous mechanoreceptors of the feet during postural control is proposed. As shown by the literature, the assumption our team makes is that it is possible to disturb the postural equilibrium by producing targeted vibration under the feet [12], [13]. Based on this hypothesis, the team built and developed an inexpensive and open vibration platform that could easily be reproduced by professional and other potential end-users. The platform uses widely available electronic components and is designed to be used in conjunction with a force platform. The cost of the platform in its current state is 124.67 CAD$ (see Section 3, Table 1). Moreover, the code to control the platform will be made available to the community on the researchers' websites. The main goal in this paper is to demonstrate the feasibility of such an approach and to provide everything needed for an easy reproduction. In addition, as a secondary objective, the platform was designed to automatically calibrate to reproduce the optimal test sequence that would be performed by a physiotherapist. To do so, a simple reinforcement learning agent was programmed and trained with human interaction. This agent, using the feedback from normal subjects, is able to automatically associate the vibration motors to the testing sequence. Consequently, it would not be necessary for a physician to modify the Arduino code to associate the motors to the testing sequence.

2. Related Work

While there are several studies using force platform to evaluate postural imbalance [6], [7], very few worked on the stimulation of imbalance through sensory manipulation. One example of those works is from Toosizadeh *et al.* [11]. They used vibration on the calves to induce body tilt on participants to compare the reaction among three profiles (young, healthy elders, and elders at risk of postural imbalance). They were able to observe, due to the vibration, a notable difference on balance performance among the high fall risk participants. The vibration served as a tool to measure reduced sensitivity in peripheral nervous system in those participants and therefore proved to be useful for research on aging and on assessing the risks of fall. The team of Najafi *et al.* [3] also worked on using electrical stimulation on foot plantar, but through the use of widely available transcutaneous electrical nerve stimulator (TENS). In their study they aimed to evaluate if the electrical stimulation could improve postural balance of patients with moderate to severe peripheral sensory neuropathy. They exploited Food and Drug Administration (FDA) cleared TENS units named named SENSUS® from NeuroMetrix Inc. (Waltham, MA, USA). The study was separated into two groups; one received working TENS, and the other received a placebo (a TENS producing no stimulation). They were able to observe significant planter sensibility increase through the end of the trials for the group with the device suggesting that electrical stimulation could not only be used as a tool to assess postural imbalance, but even to improve it. The inconvenient of using the TENS is that electrodes need to be replaced, and it requires two units (one per foot) turning it into a costly solution. Additionally, since the devices are not open, it is not easy to get full control over them. Finally, the team of Kavounoudias *et al.* [13] used a similar platform to the one proposed in this paper. In their paper, they defined the conditions upon which the stimulation of feet soles could modify the center of pressure of the standing person. The objective of the paper was to demonstrate that two types of proprioceptive information are jointly used to ensure the standing position. The first information is tactile and the second originates from muscle spindles. Therefore, an unstable ground under a foot will imply that it detects this information through its cutaneous sensors, information that the central nervous system will treat to finely adjust gait pattern and postural control. The method that was used in their paper is the motivation behind our work. Unfortunately, the platform is not reproducible since the paper lack the technical details to do so.

3. The Vibration Platform

The goal of the vibration platform is to enable a physiotherapist to easily test the postural control of the person in the standing position by stimulating the sensors under the feet through vibration and thus forcing the patient to modify his stance to restore his equilibrium. Such a testing platform can act as a tool to help the physiotherapist in his decision process and to carry out his job efficiently. Consequently, the platform must possess properties that enable the stimulation without creating discomfort. As seen in the literature [13], the vibration platform must be equipped with motors able to vibrate at a certain minimal vibration (50Hz ± 20Hz). A higher vibration frequency seems to correlate with more stimulation. Still, there are other constraints to consider such as the size of the motors. They have to be in direct contact with the feet soles. The team opted for the ROB-08449 which are 10mm in diameter and .5mm thick. They operate at 2.3～3.6V and require 60mA. There are 6 vibration motors disposed as shown on Figure 2. As a side note, the motors' connections are too fragile for repeated trampling over. For this reason, the connections are coated with heavy duty contact glue which is cheap, flexible and very durable.

To install them, the team selected a sheet of polyethylene about 2 cm thick. The vibration motor can be glued to it with ease and since the surface adapt, the contact can be directly on the soles without causing too much discomfort for the patients. Moreover, the surface adapts rapidly to the body temperature and does not incommode the patient by being too cold to touch. The microcontroller chosen to develop the platform is an Arduino Due which possess enough pins to control the vibration motors. The Arduino is a well-known microcontroller widely available and easy to program. It was important for this research project to utilize components widely accessible for anyone to be able to reproduce the platform. For the project, a breadboard 830 with wire kit was utilized with 22AWG standard electrical wire. The 16 motors require 960mA. They are powered by a Recom AC/DC converter (RAC04-3.3SC/W) which outputs 1.2A at 3.3v.

Table 1 shows all the components used in building the platform, their respective unit price and the total average cost. Figure 1 shows the electrical schema for each motor attached to the Arduino. Finally, Figure 2 shows the physical installation of the motors on the polyethylene mat. The wires are hidden under the mat. As a side note, our team encourage any reader to contact us if a piece of the puzzle is missing for reproduction.

Table 1. Detail of the material and prices.

Material	Unit price
Arduino Due	$38.50
Motor (x16)	$2.15
Polyethylene piece	$5.00
Bread board with wires	$9.50
Other wires	$10.00
Transistors x16	$0.30
Voltage converter	$22.47
Total price	$124.67

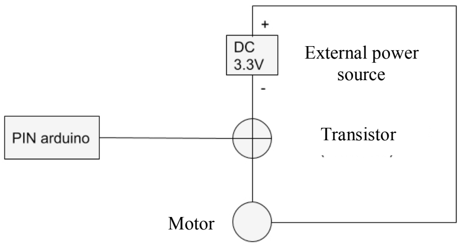

Fig. 1. Electrical schema of one motor.

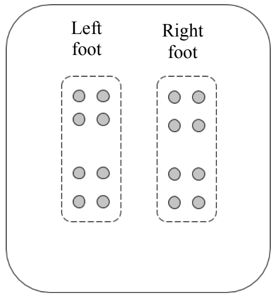

Fig. 2. Disposition of the motors.

3.1. Control software

The platform is controlled through a standard PC. The team developed a small piece of software on the Arduino board to control the vibration motors and to interface the platform with a computer. The software is written in C# with the .NET library. It enables a user to activate each motor or a group of motors for a specified duration. A button can force the disabling of all vibration motors at once in case of an emergency. The software (written in French) is shown on Figure 3. In the current version of the platform, the Arduino must be wired to the PC.

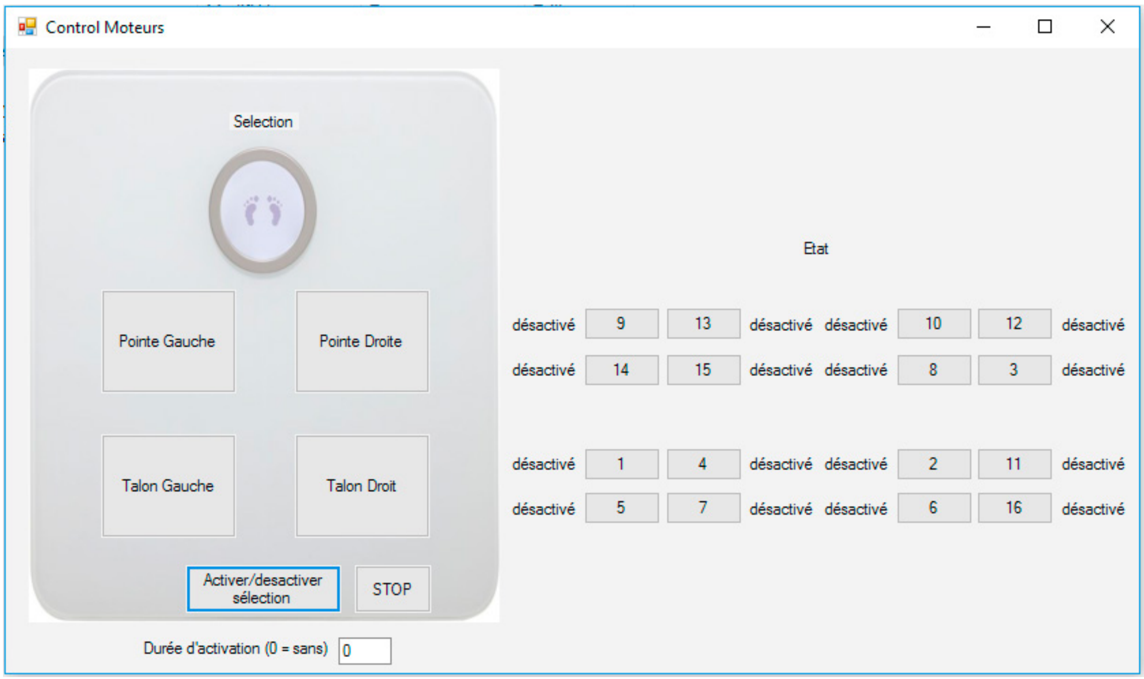

Fig. 3. The software developed to control the vibration motors.

4. A Simple Reinforcement Learning Agent

One of the challenges with the reproduction of the platform would be the motors association and the calibration of the platform to perform the test. Indeed, the physiotherapist would need to delve into the Arduino code to associate the pin to the vibration motors. Moreover, he may not be skilled enough to select the required power to maximize the result. Consequently, with the platform, a simple reinforcement learning agent was developed to automatically perform this aspect of the platform [14]. The reinforcement learning agent is optional and serves only for calibration, but it could simplify the task of clinicians who would venture into reproducing our platform.

4.1. The Reinforcement Problem

In reinforcement learning, the agent learns by evaluating the feedback received from the environment. It observes the state, select an action to perform, and then assesses how the environment has evolved consequently to its action. In our case, the problem is theoretically not difficult. It is episodic, meaning that each learning episodes are independent from each other. The episodes consist of a bounded period for which the agent select an action and observes the reaction from the human standing on the platform. The episodic property implies that the agent does not need to reason about the future. The environment is also fully observable and discrete. Moreover, the problem itself is stationary. This means that an action should always return a similar result within some uncertainty margin. In our case, the activation of a group of motors should always produce the same approximate movement of the center of pressure of the person (plus some margin due to variation of body mass and sensitivity). If the problem was not stationary, the tools could not be relied upon for the evaluation of postural control. The problem can therefore be associated to a single state Markovian Decision Process [15]. Here, the environment possesses a single state and all actions are available directly to the learning agent any time. The most challenging aspect of the reinforcement process is that it has to be done with humans. Therefore, there are no shortcuts available to accelerate the convergence of the learning process. This aspect had to be taken into account during the conception of the learning method. To this end and due to the electrical schema of the platform, in the current version of the learning agent, the motors activation is binary (either on or off), but in later version an analog version with a discrete numerical variable.

4.2. The Reinforcement Method

Let A be the set of all possible actions. An action $a_j \in A$ is defined as the sequence of motors to activate $a_j = [M_1, \dots, M_i], M_i \in \{0,1\}, i \in \{1, \dots, number\ of\ motors\}$. The first step of the reinforcement is to initialize the possible actions and their associated expected rewards. In this stationary problem, an action can only succeed or fail. The expected reward during a learning episode t is therefore defined by the binary function $Q_t(a_j) \in \{0,1\}, \forall t \in \{0,1, \dots, lastEpisode\}, \forall j \in \{0,1, \dots, |A|\}$. The method requires to keep track of the number of times each individual action has been tested since the expected rewards refines for each iteration of the algorithm. The second step is the selection by the agent of an action to perform for an iteration. This step enables the agent to either refine the reward of an action known to be good or to explore a new strategy by selecting actions for which outcome is unknown. In our case, a upper-confidence bounds (UCB) method is implemented [16]. This policy uses the average reward of an action a_j at time t in regards of the number of times it was selected. The selected action a_t is thus given by the equation 1.

$$a_t = \underset{a \in \mathrm{A}}{\operatorname{argmax}} \left[Q_t(a) + c \sqrt{\frac{\ln(t)}{N_t(a)}} \right] \tag{1}$$

The parameter c modulates the priority given to actions that are rarely selected. A higher value will lead the agent to be exploration driven whereas a small value will lead to quick convergence to the best actions. The function $N_t(a)$ simply returns the number of times the action a has been selected. The third step is the execution of the action by activating the motors with the parameter selected (here 0 or 1). It then observes the environment to obtain a reward $R_{a,t}$. The last step updates the value associated with the selected action using the equation 2.

$$Q_t(a) = \frac{1}{t} \sum_{i=1}^{t} R_{a,i} = Q_{t-1}(a) + \frac{1}{t}\left(R_{a,t-1} - Q_{t-1}\right) \tag{2}$$

5. Evaluation of the platform and the RL agent

In order to evaluate the system and the reinforcement learning agent, the platform was fixed with adhesive over a force platform Biomec 400 v1.1™ force plate (EMG System do Brasil Ltda® - http://www.emgsystem.com.br) [6]. The goal was to verify, using the force platform, if the vibration motors could produce the desired output. Other force platforms exist that may be less expensive than the Biomec and still produce similar results. For example, the TekScan

MatScan® platform was used in the study of [6]. Their study with 3 participants with rheumatoid arthritis, aged between 60 to 80 years old, concluded that the force platform is a useful tool for the measurement of postural stability in clinical and research settings.

5.1. Using Computer vision for Evaluation

The force platform is a proprietary technology that provides in real-time a graphical user interface to show the movement of the center of pressure of the subject. It also provides a detailed trace of the experiments with the raw data captured during the tests. Unfortunately, it is not possible, to our knowledge, to get the raw data in real-time since the platform does not possess an API. To circumvent this limit, the team developed a Python script to visually track the center of pressure in real-time in the proprietary software. The script exploits the well-known OpenCV [17] library for computer vision. The script is very straightforward. It starts by detecting the window showing the tracking in the platform software (the window size is fixed and fairly easy to detect). Then, to detect the center of pressure, the image is first turned in black and white and eroded with a morphological operator in OpenCV. The erosion use a convolution to eliminate the group of black pixels that are too small which translates into the disappearance of the noise and the grid. Thereafter, the opposite operation (dilatation) is performed to enhance the interesting zones in the image. Using an edge detector, the center of pressure can finally be found (the longest edge) and then be used to analyze the displacement of the person on the platform in real-time.

5.2. Tests and results

The testing of the platform was done throughout the duration of this project. Tests were done to make sure the platform would resist long term use (solidity and robustness), the motors were enabled for a long period to make sure the electrical and electronic components would not heat up too much, and the code was tested using the best software engineering practice. To validate the reinforcement learning agent, the lab member simply had to spend a certain time standing on the platform and let the agent evolves. Each episode consists in calibrating (the person has no stimulus and the new center of pressure is established), vibrating for a fixed period, and calculating the displacement. Since the problem is easy and since the association of the motors is known in our experiment, the idea is then simply to verify for a certain desired output if the agent learns the optimal solution. Moreover, it is important that the agent finds the solution quickly, since it is to be used by humans and the learning will be done standing on the platform (impossible to accelerate or simulate normal reactions). For the validation, we fixed each episode to 5 seconds and the maximum learning time for a stimulation to 5 minutes (60 episodes). In the tests presented here, we selected the solution [0000111100001111] that should produce a displacement to the toes (the 8 motors under the heels are activated). For a purely greedy selection (with $\varepsilon = 0.95$), the agent cannot converge to the optimal solution in 200 episodes (there are 65 536 unique actions to test). With the UCB approach (with $\varepsilon = 0.7$), the optimal solution is in general found under 20 episodes and it usually distinguishes the optimal solution from the almost optimal solutions in less than 60 episodes. Figure 4 below shows two examples of learning tests comparing the greedy approach to the UCB approach.

Fig. 4. Left: greedy policy with ε=0.95 and 200 episodes. Right: UCB policy with ε=0.7 and 60 episodes (optimal solution is found).

6. Experiments with human participants

The last step to validate the new vibration platform was to conduct a small set of experiments with humans with normal postural control. The team obtained an ethical certificate (602.604.02) and recruited 7 healthy adults (Table 2 shows their sociodemographic characteristics). Four different vibration conditions were tested: (i) vibration of both heels (expected mean body tilt = forward); (ii) vibration of both forefoot soles (expected mean body tilt = backward);

(iii) vibration of right heel and forefoot (expected mean body tilt = leftward); (iv) vibration of left heel and forefoot (expected mean body tilt = rightward). The expected directions of body tilts are based on reference publications with the gold standard vibration technology. For each condition, five trials were realized in a random order between trials and between participants. Each trial consisted of a 20 sec recording without vibration, followed by 3 seconds of vibration. Rest periods were given between each trial.

Table 2. Participants' sociodemographic characteristics.

Participants	Age (years) mean ± SD	Age (years) range	Gender (males/females)	Height (m)	Weight (kg)	Body mass index (m/kg^2)
7	40.8 ± 14.2	28-62	3/4	1.66 ± 0.09	67.0 ± 13.2	24.4 ± 3.9

SD = standard deviation

Variables of body tilts induced by the vibration conditions were based on center of pressure X and Y coordinates measured at a 100 Hz sampling rate during the 30 s recordings with the force platform. The mean CoP location before vibration was calculated for the whole 20 s period before each trial. All further displacements of the CoP during vibration were expressed relative to this mean baseline CoP position, which was automatically re-centered at [0;0] coordinates. Then, a linear envelope was created by filtering CoP variables at 5 Hz (thus having 5 CoP coordinates per second, for a total of 15 data points for the 3s of vibration). The average CoP position across the five trials was calculated for each CoP variables, and the group mean was computed. Because of the preliminary nature of the study, CoP data were only compared qualitatively to the expected postural reactions described in the literature [13][12].

All procedures were well tolerated, and no adverse event occurred. Figure 5 shows that vibration of the heels induced the expected forward body tilts, with minimal mediolateral displacements. For the three other conditions, while the postural reactions were initially in the appropriate direction (i.e. posterior for the vibration of forefoot soles; leftward for the vibration of right foot; rightward for the vibration of left foot) (1), the CoP then diverged mostly anteriorly. These observations suggest that the prototype does induce postural reactions, but that are not yet completely satisfactory to support the validity of our portative vibration platform. We think that the heels condition led to better results because participants were instructed to position their heels above the appropriate vibrators, but this caused the other vibrators to be less appropriately positioned under the forefoot for the participants having smaller feet. Also, the referenced approach used a greater number and larger vibrators, and the vibration frequency was higher than in our work [13]. Altogether, these technological issues could have undermined the strength of the cutaneous feedback transmitted to the postural control centers.

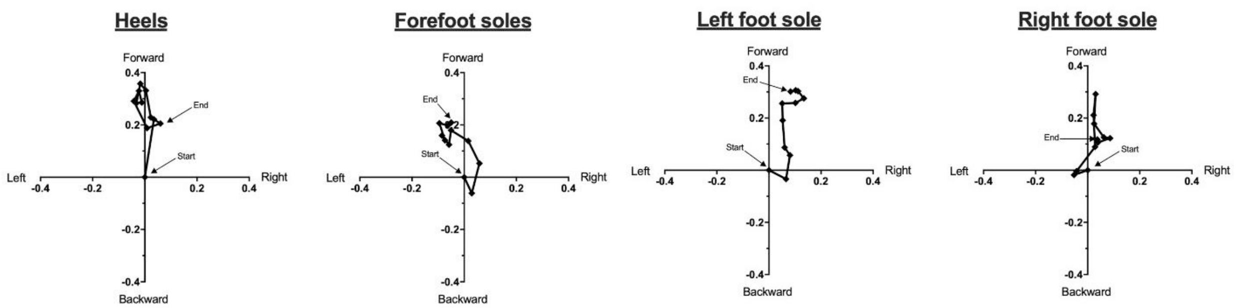

Fig. 5. Effects as visualized on the force platform.

7. Conclusion and future work

In this paper, we presented a new open vibration platform that could be used in conjunction with a force platform in order to evaluate postural control in a quick and non-invasive manner. The platform is inexpensive and easily reproducible by most persons. The experiments presented show that the platform behave accordingly to our hypothesis and that it enables to stimulate the sensory system under the feet to induce body tilt. Nevertheless, the vibration motors were not as strong as expected and the team will be looking and trying different model to further improve the reaction. Additionally, in the current version, the motors could either be on or off, but if the motors were stronger, it may have been better to have an analog output. In the next version of the platform, the motors will be controlled with analog pins, and, accordingly, the reinforcement learning agent will be updated to learn on discrete numerical variables (i.e.: [0.0-0.2, 0.2-0.4, …]). In addition, to further confirm that the platform is effective, the team is recruiting persons to conduct a clinical trial. The trial should involve a baseline group of 30 healthy aged persons and a group of 30 with postural imbalance characteristics. Finally, as a side note, one concern that we have is to keep the price and the complexity of the platform minimal to allow the most persons to be able to reproduce it.

Acknowledgements

The authors are very thankful to the Fondation de l'UQAC for the grant that enabled them to conduct this research. They also want to thank the Centre Intersectoriel en Santé Durable (CISD) for the support provided toward a second phase which will include broader clinical trials in physiotherapy.

References

[1] K. L. Roberts and H. A. Allen, "Perception and cognition in the ageing brain: a brief review of the short-and long-term links between perceptual and cognitive decline," *Front. Aging Neurosci.*, vol. 8, p. 39, 2016.

[2] L. Schmidt, G. Kerkhoff, and K. S. Utz, "Sensory stimulation in post-stroke postural imbalance: A novel treatment approach?," *Clin. Neurophysiol.*, vol. 127, no. 1, pp. 21–22, 2016.

[3] B. Najafi, T. K. Talal, G. S. Grewal, R. Menzies, D. G. Armstrong, and L. A. Lavery, "Using plantar electrical stimulation to improve postural balance and plantar sensation among patients with diabetic peripheral neuropathy: a randomized double blinded study," *J. Diabetes Sci. Technol.*, vol. 11, no. 4, pp. 693–701, 2017.

[4] B. Timar, R. Timar, L. Gaiț\ua, C. Oancea, C. Levai, and D. Lungeanu, "The impact of diabetic neuropathy on balance and on the risk of falls in patients with type 2 diabetes mellitus: a cross-sectional study," *PLoS One*, vol. 11, no. 4, p. e0154654, 2016.

[5] I. Melzer, N. Benjuya, and J. Kaplanski, "Postural stability in the elderly: a comparison between fallers and non-fallers," *Age Ageing*, vol. 33, no. 6, pp. 602–607, 2004.

[6] A. Brenton-Rule *et al.*, "Reliability of the TekScan MatScan®system for the measurement of postural stability in older people with rheumatoid arthritis," *J. Foot Ankle Res.*, vol. 5, no. 1, p. 21, 2012.

[7] M. R. de Oliveira, R. A. da Silva, J. B. Dascal, and D. C. Teixeira, "Effect of different types of exercise on postural balance in elderly women: a randomized controlled trial," *Arch. Gerontol. Geriatr.*, vol. 59, no. 3, pp. 506–514, 2014.

[8] M. R. Oliveira *et al.*, "One-legged stance sway of older adults with and without falls," *PLoS One*, vol. 13, no. 9, p. e0203887, 2018.

[9] U.-S. D. T. Nguyen *et al.*, "Correlations of clinical and laboratory measures of balance in older men and women," *Arthritis Care Res. (Hoboken).*, vol. 64, no. 12, pp. 1895–1902, 2012.

[10] C. K. Rhea *et al.*, "Noise and complexity in human postural control: interpreting the different estimations of entropy," *PLoS One*, vol. 6, no. 3, p. e17696, 2011.

[11] N. Toosizadeh, H. Ehsani, M. Miramonte, and J. Mohler, "Proprioceptive impairments in high fall risk older adults: the effect of mechanical calf vibration on postural balance," *Biomed. Eng. Online*, vol. 17, no. 1, p. 51, 2018.

[12] R. Roll, A. Kavounoudias, and J.-P. Roll, "Cutaneous afferents from human plantar sole contribute to body posture awareness," *Neuroreport*, vol. 13, no. 15, pp. 1957–1961, 2002.

[13] A. Kavounoudias, R. Roll, and J.-P. Roll, "Foot sole and ankle muscle inputs contribute jointly to human erect posture regulation," *J. Physiol.*, vol. 532, no. 3, pp. 869–878, 2001.

[14] M. L. Littman, "Reinforcement learning improves behaviour from evaluative feedback," *Nature*, vol. 521, no. 7553, p. 445, 2015.

[15] M. L. Puterman, *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.

[16] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Mach. Learn.*, vol. 47, no. 2–3, pp. 235–256, 2002.

[17] K. Pulli, A. Baksheev, K. Kornyakov, and V. Eruhimov, "Real-time computer vision with OpenCV," *Commun. ACM*, vol. 55, no. 6, pp. 61–69, 2012.