# Affordable, automatic quantitative fall risk assessment based on clinical balance scales and Kinect data.

P. Colagiorgio, F. Romano, F. Sardi, M. Moraschini, A. Sozzi, M. Bejor, G. Ricevuti, A. Buizza - *IEEE Member* and S. Ramat- *IEEE Member*

***Abstract*— The problem of a correct fall risk assessment is becoming more and more critical with the ageing of the population. In spite of the available approaches allowing a quantitative analysis of the human movement control system's performance, the clinical assessment and diagnostic approach to fall risk assessment still relies mostly on non-quantitative exams, such as clinical scales. This work documents our current effort to develop a novel method to assess balance control abilities through a system implementing an automatic evaluation of exercises drawn from balance assessment scales. Our aim is to overcome the classical limits characterizing these scales i.e. limited granularity and inter-/intra-examiner reliability, to obtain objective scores and more detailed information allowing to predict fall risk. We used Microsoft Kinect to record subjects' movements while performing challenging exercises drawn from clinical balance scales. We then computed a set of parameters quantifying the execution of the exercises and fed them to a supervised classifier to perform a classification based on the clinical score. We obtained a good accuracy (~82%) and especially a high sensitivity (~83%).**

## I. Introduction

The classical technique for fall risk assessment in a clinical setting is currently based on testing a patient's balance skills through Balance Evaluation Scales. Clinical balance assessment scales are composed of a set of exercises (items) aimed at assessing patients' motor and balance control capabilities. Clinicians evaluate visually the patient's performances for every item, summarizing their judgment with an integer score, chosen following an evaluation guideline allowing for only a limited set of values. Different balance scales may be found in the literature, differing in the number and type of proposed exercises, evaluation guidelines and granularity.

An alternative is provided by the EquiTest system, which automatically evaluates the contribution of the three sensory systems that are most relevant for posture control (vision, vestibular and proprioception) by studying the sway path of the Center of Pressure (COP) in response to combined stimuli delivered through a moving platform and a visual display [1]. Such system, though, is expensive and is thus available in only a few specialized centres and laboratories throughout the world. Moreover, several studies on geriatric fall risk have proposed a quantitative evaluation based on wearable inertial sensors [2].

The aim of our work was to develop an automatic system for balance assessment scales evaluation. We chose thirteen exercises for evaluation, drawn from the Tinetti Test [3] and adding exercises from the Berg Balance Scale (BBS) [4] and BESTest [5]. To train and test the system we recorded 66 elderly subjects, some of which with motor difficulties (largely caused by ageing or traumas), and 13 control subjects.

Movements were recorded using a Microsoft Kinect device (placed 2m in front of the subject), a low-cost 6-DoF marker less tracking system, which despite its lacks the precision of a multi-camera video system with markers ant its time resolution is limited by the 30 fps rate, is inexpensive, portable and simple to setup and use. Nonetheless published results shown [6]–[8], and our own data are promising and suggest that the Microsoft Kinect can be used to validly assess postural control in clinical settings.

## II. Evaluation Techniques

The chosen items were first divided into two categories: *static*, i.e. exercises that are meant to inspect postural control during stance, and *dynamic*, i.e. exercises that require the execution of a movement. The proposed automatic evaluation system performs different analyses for each item depending on its category.

The maximum score of the Tinetti Test is 28, and fall risk is identified with scores lower than 25. Our maximum test score is instead 33, thus, scaling the scores of the Tinetti Test, we considered patients as being at risk of falling when they achieved values lower than 29.

### *A. Static Exercises*

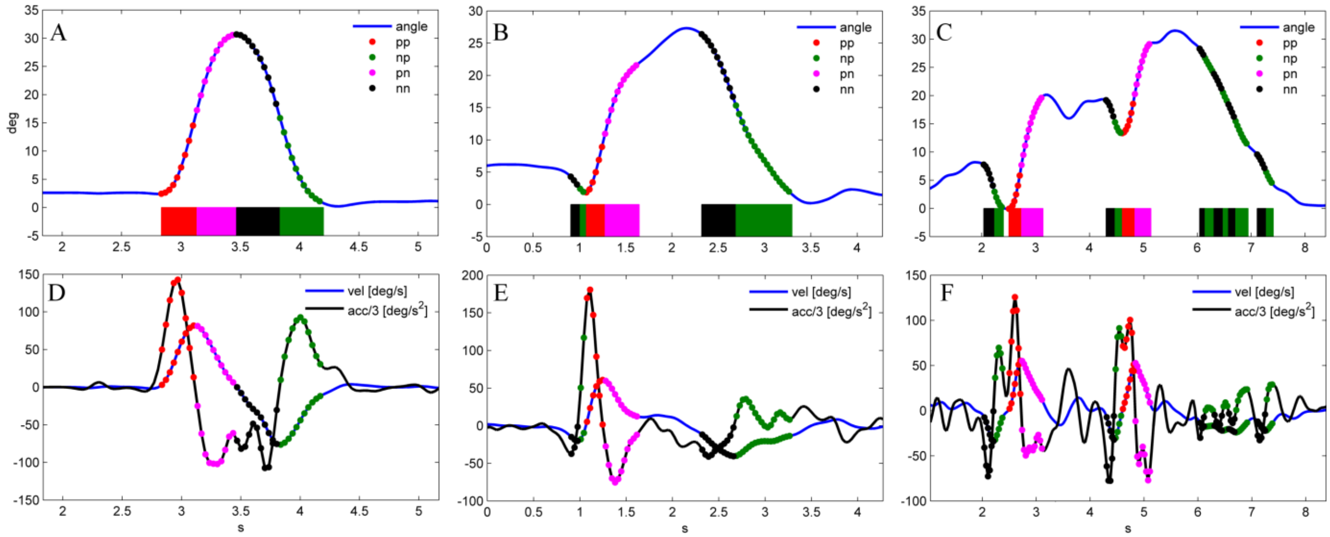
In order to assess postural control capabilities, during the execution of a static exercise patients are requested to maintain a given position for 30 seconds. Analysed items were: sitting (ST), standing with eyes open (SEO), Romberg test (standing with eyes closed SEC) and sternal nudge (NG, three gentle nudges on the patient's sternum to see how he responds). The last three items were also executed while

Research supported by the Italian Ministry of Research PRIN grant 2010R277FT_004 to S. Ramat.

P. Colagiorgio, F. Romano, M. Moraschini, A. Buizza and S. Ramat (corresponding author, phone: +39 0382985745, fax: +39 0382985373) are with the Department of Electrical, Computer and Biomedical Engineering, University of Pavia, Pavia, 27100 Italy (e-mails: paolo.colagiorgio@unipv.it , fausto.romano01@universitadipavia.it, marco.moraschini01@universitadipavia.it , angelo.buizza@unipv.it , stefano.ramat@unipv.it)

F. Sardi and G. Ricevuti are with the Department of Internal Medicine and Therapy, Gerontology and Geriatrics, University of Pavia (e-mails: sardi.francesca@gmail.com , giovanni.ricevuti@unipv.it)

M. Bejor and A. Sozzi are with the Department of Clinical-Surgical, Diagnostic and Paediatric Sciences, University of Pavia (e-mail: maurizio.bejor@unipv.it , arisozzi@yahoo.it).




standing on a soft foam cushion (as in the BESTest), studied to alter proprioceptive information (SEOF, SECF, NGF).

Figure 1. A-B-C) Chest pitch angle during SU (blue solid line); points represent kinetems segmentation based on the sign of the first and second derivative (blue and black lines respectively in D-E-F; acceleration is scaled). A/D show the exercise for a healthy subject scored 1 by PHMM. B/E and C/F show the same movement for two elderly subjects scored 0.47 and 0 respectively. Particularly elderly subject in C/F needed two attempts to stand up.

In the same category we considered also the standing unsupported one foot in front (HT) and the extension of the reaching forward while standing (RF), both drawn from BBS.

In a research setting these exercises are typically evaluated by studying the subject's Centre of Pressure (CoP) sway path.

Without a force plate, however, no CoP data is available, and therefore we considered studying the sway path of the ground projection of the subject's Centre of Mass (CoM) provided by the Kinect. Yet, because of the susceptibility of the instrument to background noise (i.e. the presence of other subjects, or objects, near the one being tracked) we had to approximate the CoM with the ground projection of the chest joint provided by the Kinect, i.e. a point below the sternum on the body midline.

From the analysis of this sway path we calculated 12 parameters: medio-lateral (ML) and antero-posterior (AP) standard deviations, area of the circumscribed ellipsis, mean velocity and 8 parameters extracted from diffusion plots as suggested in [9], [10], to evaluate the open- and closed-loop control of posture ($\Delta r^2_c$, $\Delta t_{rc}$, $\Delta x^2_c$ , $\Delta t_{xc}$, $\Delta y^2_c$, $\Delta t_{yc}$, Ks, Kl).

For the RF item we evaluated the maximum amplitude of chest pitch (RF-PA).

### *B. Dynamic exercises*

The dynamic items considered in our test were: stand up from the chair (SU), sit down (SD), place each foot alternately on the stool while standing unsupported (SOS) and make a complete 360 deg turn (TR).

For the evaluation of dynamic exercises we chose to create a model of the movement using an approach derived from neurolinguistics [11], [12]. To this goal we tracked seven body segments (forearms, arms, legs and the chest) and extracted the signals representing the evolution of their orientation in space in terms of angles in the roll and pitch planes. Data were filtered with a low pass filter (3 Hz cut-off frequency).

The detection of the movement was based on an energy criterion [13], i.e. all the samples under a chosen energy threshold were considered as static. Non-static data were segmented according to the sign of the first (velocity A') and second derivative (acceleration A'') and each segment was then labelled as one of four possible kinetemes (i.e. *pp*: A'>0 and A''>0; *pn*: A'>0 and A''<0; *np*: A'<0 and A''>0; *nn*: A'<0 and A''<0). Fig. 1 shows an example of such representation for the chest pitch angle during SU, for a healthy (left) and two elderly subjects.

We thus obtained a sequence of symbols from every actuator (i.e. a pair segment-plane e.g. "right arm on the roll plane") during the execution of the dynamic exercises. These sequences were then used to train, on the 13 healthy subjects, a set of Profile Hidden Markov Models (PHMM), obtaining one model for each actuator. For each exercise we modelled only the actuators involved in the movement. Therefore each patient's exercise may be scored by PHMM based on the probability that the observed sequence of letters may be produced by the corresponding actuator model.

Moreover for the SU exercise we calculated the time to get up and the fluidity of the movement (in terms of number of attempts) from the pitch orientation of the chest, and for the SOS we calculated the velocity of the steps (SOS-V=number of steps/exercise duration).

## III. AUTOMATIC CLASSIFICATION

The set of parameters computed on each item was then fed to a supervised classifier, trained to perform a dichotomy classification (1/0 fall risk or not) based on the clinical score (sum of the scores obtained in each exercise performed by the subject).

Our dataset was composed by two groups: a first control group of 13 healthy subjects (aged 26±5 years, mean±std) acquired in our laboratory, scored with the maximum clinical score; and a second group of 66 elderly subjects (aged 76±10 years) acquired in two clinics participating to the study: the "Santa Margherita Institute" in Pavia and the "Santa Maria alle Fonti Centre" in Salice Terme. The experimental paradigm was approved by the local ethics committee and all enrolled subjects signed an informed consent prior to participating in the study. In both groups, the clinical score was assigned by the clinician managing the test, and 22 patients had a history of at least one previous fall.

The resulting labelled dataset contained 49.3% of subjects at fall risk and 50.7% not at risk.

### *A. Data pre-processing*

In order to limit noise in the training data and to manage possible missing data, the whole dataset was review after the computation of all features.

First, a manual selection of exercises was carried out; three dynamic and two static exercises were removed based on the following considerations:

- ST exercise did not give us any additional information, since all subjects easily did it;
- during SECF and NGF exercises several subjects needed assistance, altering and reducing their sway;
- TR exercise is not correctly acquired by the Kinect device as it identifies a person only in frontal position relative to the camera;
- correct execution of HT exercise is strictly related to the distance between the two feet, which may alter the sway measure, and cannot be monitored using data provided by the Kinect device.

Moreover, we automatically detected and deleted the uninformative features, i.e. the features with same values in more than 95% subjects.

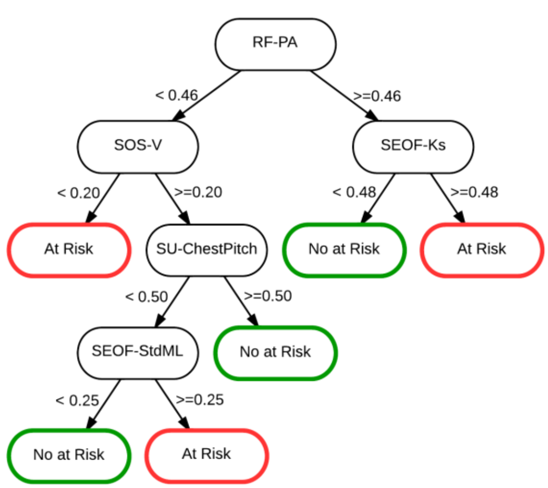



Figure 2. Rules learned by Classifcation Tree using all data; e.g. low values of RF-PA (<0.46) and SOS-V (<0.20) indetify subjects at risk of falling. Features values on braches are normalized between 0 and 1 After the "pruning" process SEOF-MeanVelocity, NG-$\Delta t_{xc}$ were deleted.

All features were normalized to the range [0, 1]. Such process allows us to compute a correct distance between two subjects in the space of parameters, avoiding the problem related to features having different scales.

Few features presented some missing values; if these were more than 20% the feature was discarded, otherwise the missing values were handled by replacing the missing attribute of a subject with the value of the nearest neighbor computed using a distance metric that did not consider the missing values.

### *B. Features Selection*

After the pre-processing step the number of features (80) was still comparable with the number of subjects, thus we used a feature selection algorithm, to discard others uninformative features. We chose a filter method, i.e. Relieff, for evaluating the "discriminate power" of a single feature for our two classes, and considering the interaction between all of them. However, such method did not discriminate between redundant features, then we took only one of the correlated (with $r \geq 0.90$).

The final feature set was thus limited to the N most informative features; where N is equal to 10% of the number of subjects (8 in our dataset).

During evaluation of the classifiers, the selected feature set was computed on each training set, in order to avoid overestimating the classifiers' accuracy. Once the accuracy was estimated, the final features set was selected using all data.

### *C. Classifiers evaluation*

The accuracy, sensitivity and specificity of each classifier was estimated through a non-parametrical approach, i.e the .632 bootstrap technique [14], since the number of our subjects was limited. Consequently, we computed the accuracy confidence interval as the 97.5 and 2.5 percentile, and we tested the medians using a Friedman non-parametric test.

TABLE I. CLASSIFIERS' PERFORMACES

| Classifier | Accuracy C.I. | Sensibility | Specificity |
|---|---|---|---|
| Majority Classifier | 47.9% [-9.3, +9.7]% | 47.7% | 47.8% |
| Classification Tree | 82.3% [-9.3, +7.8]% | 83.1% | 82.4% |
| Linear SVM | 84.3% [-23.3, +10.6]% | 80.2% | 91.3% |
| KNN | 81.2 % [-13.0, +9.47]% | 80.6% | 81.9% |
| Naive Bayes | 82.1% [-13.8, +9.49]% | 80.9% | 81.2% |

## IV. RESULTS

We tested the proposed method's ability to distinguish between control subjects and pathological patients.

To test the ability of the proposed system to separate pathological subjects from controls we applied a set of automatic classifiers as described above. The results of such tests are summarized in Table 1. All tested classifiers resulted statistically different from the "majority classifier". We chose the "Classification Tree" since his classification rules are easily understandable throughout a graphically depiction, as shown in Fig. 2.

The final feature set, ordered by weight based on the Relieff technique was: RF-PA, SOS-V, SEOF-Ks, SU-ChestPitch, SEOF-StdX, SD-ChestPitch, SEOF-MeanVelocity, NG-$\Delta t_{xc}$.

Our results show that the more influent dynamic exercises are SOS and SU, while from static items, the sway during SEOF results as being very informative about stability and the inclination during RF proved a good indicator of fall risk.

## V. CONCLUSION

In this work we proposed an automatic and low-cost system to identify fall risk in elderly people. The Kinect tracking system proved to be very useful and rather reliable in measuring human movements, in both static and dynamic conditions. Using the clinicians' scores to train the classifier, we obtained a good accuracy (~82%) and especially a high sensibility (~83%). Such result is an encouraging indicator that our analysis is able to grasp the balance and movement characteristics that are relevant to the assessment of fall risk. Nonetheless, the use of a classification based on the clinician's scores is clearly contradictory to the goal of the proposed approach, attempting to overcome the limitation of subjective scoring. For this reason, more assertive data is needed to continue the validation of our approach.

As a next step, in order to overcome the variability of subjective scoring, we will develop our research along two lines. On one hand we will extend the available dataset to allow testing the system against the real fall history of each subject, on the other we will develop unsupervised classifier approaches to evaluate similarities in patients clusters and their relevance in predicting a patient's fall history.


## ACKNOWLEDGMENT

The authors would like to thank Dr. Nora Camerino for her help with the data acquisition in Salice Terme.


## REFERENCES


[1] L. Nashner, F. Black, and C. 3d Wall, "Adaptation to altered support and visual conditions during stance: patients with vestibular deficits," *J. Neurosci.*, vol. 2, no. 5, pp. 536–544, May 1982.

[2] J. Howcroft, J. Kofman, and E. D. Lemaire, "Review of fall risk assessment in geriatric populations using inertial sensors.," *J. Neuroeng. Rehabil.*, vol. 10, no. 1, p. 91, Jan. 2013.

[3] M. E. Tinetti, "Performance-oriented assessment of mobility problems in elderly patients.," *J. Am. Geriatr. Soc.*, vol. 34, no. 2, pp. 119–26, Feb. 1986.

[4] K. O. Berg, S. L. Wood-Dauphinee, J. I. Williams, and B. Maki, "Measuring balance in the elderly: validation of an instrument.," *Can. J. Public Health*, vol. 83 Suppl 2, pp. S7–11, 1992.

[5] F. B. Horak, D. M. Wrisley, and J. Frank, "The Balance Evaluation Systems Test (BESTest) to differentiate balance deficits.," *Phys. Ther.*, vol. 89, no. 5, pp. 484–98, May 2009.

[6] E. E. Stone and M. Skubic, "Unobtrusive, continuous, in-home gait measurement using the Microsoft Kinect.," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 10, pp. 2925–32, Oct. 2013.

[7] 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. 372–7, Jul. 2012.

[8] Y.-J. Chang, S.-F. Chen, and J.-D. Huang, "A Kinect-based system for physical rehabilitation: a pilot study for young adults with motor disabilities.," *Res. Dev. Disabil.*, vol. 32, no. 6, pp. 2566–70, Jan. 2011.

[9] L. Chiari, a Cappello, D. Lenzi, and U. Della Croce, "An improved technique for the extraction of stochastic parameters from stabilograms.," *Gait Posture*, vol. 12, no. 3, pp. 225–34, Dec. 2000.

[10] J. J. Collins and C. J. De Luca, "Open-loop and closed-loop control of posture: a random-walk analysis of center-of-pressure trajectories.," *Exp. brain Res.*, vol. 95, no. 2, pp. 308–18, Jan. 1993.

[11] A. Turcato and S. Ramat, "A computational framework for the standardization of motion analysis exploiting wearable inertial sensors," *... Med. Biol. Soc. EMBC, ...*, pp. 4963–4966, 2011.

[12] G. Guerra-Filho and Y. Aloimonos, "The syntax of human actions and interactions," *J. Neurolinguistics*, vol. 25, no. 5, pp. 500–514, Sep. 2012.

[13] D. Curone, G. M. Bertolotti, A. Cristiani, E. L. Secco, and G. Magenes, "A real-time and self-calibrating algorithm based on triaxial accelerometer signals for the detection of human posture and activity.," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 4, pp. 1098–105, Jul. 2010.

[14] B. Efron and R. Tibshirani, "Improvements on cross-validation: the 632+ bootstrap method," *J. Am. Stat. ...*, vol. 92, no. 438, pp. 548–560, 1997.