Opening the Personal Gate between Technology and Health Care
M. Giacomini et al. (Eds.)

doi:10.3233/SHTI260118

Developing an AI-Trained Movement Screening Tool, Based on Skeleton Avatar Technique, to Evaluate and Promote Sustainable Physical Functioning in Daily Life

SOFIA BACKÅBERG[a,1], GUNILLA ELMGREN FRYKBERG[b], KATARINA ERIKSSON ÖSTH[c], WELF LÖWE[d] , CECILIA FAGERSTRÖM[a,e], JOAKIM NIKLASSON[f], ALISA LINCKE [d]

[a] *Department of Health and Caring Science, Linnaeus University*
[b] *Department of Medical Sciences, Uppsala University*
[c] *Department of Family Health, Region Kronoberg, Sweden*
[d] *Department of Computer Science and Media Technology, Linnaeus University*
[e] *Department of Research, Region Kalmar Län, Sweden*
[f] *Department of Applied Health Technology, Blekinge Institute of Technology*

ORCiD ID: Sofia Backåberg https://orcid.org/0000-0001-9687-7242, Gunilla Frykberg https://orcid.org/0000-0003-2926-810X, Welf Löwe https://orcid.org/0000-0002-7565-3714, Cecilia Fagerström https://orcid.org/0000-0002-4257-282x, Joakim Niklasson https://orcid.org/0000-0001-8585-2218, Alisa Lincke https://orcid.org/0000-0001-9062-1609

Abstract. Maintaining mobility is vital for older adults. However, standardized functional tests often overlook crucial qualitative aspects, and expert assessments (EA) are costly and lack standardization. This project aims to develop an AI-based movement screening tool (SAT-Movement Analysis) utilizing the low-cost Skeleton Avatar Technique (SAT) and standardized Observational Movement Analysis (OMA) to detect deviations in daily movement. The initial phase automated expert assessments to establish a reliable foundation for machine learning. Five participants (ages 35–57) performed Sit-To-Stand, Stand-To-Sit, and One-Leg Stance, assessed by three physiotherapists using a modified IRAF protocol. Results demonstrated correspondence between automatically aggregated expert scores and consensus scores across all aggregation levels (Pearson's $r = 0.90$–0.97, ICC = 0.91–0.98, = 0.78–1.00). These findings motivate continued development of an AI-trained screening tool providing accurate movement quality feedback based on 2D smartphone video, supporting early detection and personalized intervention.

Keywords. machine learning, movement analysis, physical functioning, skeleton avatar technique

[1] Corresponding Author: Sofia Backåberg, sofia.backaberg@lnu.se, Linnaeus University, Växjö, Sweden

1. Introduction

Maintaining mobility and balance is vital for older adults' well-being, independence, and physical activity. As the aging population grows, global efforts must focus on promoting healthy aging and enabling early detection of functional impairments to support prevention and fall risk assessment. Optimizing everyday movements, performing daily activities as energy-efficiently as possible, is key to sustainable health.

Standardized Functional Tests (FTs) quantitatively assess functional ability but require controlled settings and trained personnel, limiting their use in community care. Moreover, they often miss qualitative aspects of movement. Expert Assessment (EA) methods address this gap by evaluating movement quality, but they are time-consuming, costly, and lack standardization, reducing reproducibility.

This project aims to develop and test a prototype for AI-based movement screening (SAT-Movement Analysis) to detect deviations in daily movement quality by using low-cost Skeleton Avatar Technique (SAT) recordings and standardized Observational Movement Analysis (OMA) expert assessments. This paper presents the initial phase, focusing on automating and supporting expert assessments as a foundation for machine learning within the SAT framework.

2. Background

Sustainable human movements are defined to be efficiently and gently performed, optimized to the individual's motivation and goals, prerequisites and abilities, everyday activities, and the environment [1]. Movement awareness is key to changing movement behaviour and reflecting on one's performance through video feedback can have lasting impact [2]. Close-related projects show that video feedback, used with rehabilitation staff, empowers older adults and encourages active participation in their rehabilitation [3].

2.1. The Skeleton Avatar Technique (SAT) and its Validation

The Skeleton Avatar Technique (SAT) is a pipeline integrating hardware, software, and AI to record human movement via sensors (originally 3D), estimate joint positions, and translate kinematic data into movement quality scores using machine learning [4]. In a pilot study with 54 adults (65+), we demonstrated the feasibility of an AI-trained screening tool based on SAT [4], using Kinect 3D recordings and deep learning models (RNNs, CNNs) to predict functional tests and expert assessment scores. The 3D SAT showed high predictive accuracy (cross-validated MAE < 10%). To improve accessibility, a 2D SAT version [5] using mobile-based MoveNet pose estimation was developed, with only slightly reduced accuracy - MAE increased by 4% for physical activity and 6% for physical function. SAT, based on smartphone video of a 10-second One-Leg Stand test, effectively detects physical activity levels [4–5] and predicts balance [4]. We therefore aim to explore how SAT can support and partially automate expert assessment scoring.

2.2. Observational Movement Analysis (OMA) and the Instrument for Movement Analysis of Person Transfers (IRAF) protocol

Efforts have been made to standardize the movement analysis process of tasks in clinical practice [6]. To emphasize the qualitative aspects of movements in the perspective of sustainable human movements, a model for structured OMA, has been suggested, in which the so-called OMA-wheel can be used interactively to analyze and reflect on everyday movements [1]. Based on the OMA-wheel structure: the starting position, movement initiation and movement performance, the IRAF protocol [7] has been developed to standardize the OMA expert assessments of movements. The IRAF protocol was used in this project to assess the degree of deviation from an optimal performance for each body segment component of the movement.

3. Methods

For the prototype development, we used:

a) **Video and SAT recordings** of five research group members (ages 35–66) performing Sit-To-Stand, Stand-To-Sit, and One-Leg Stance (right/left), recorded at a 45° angle, see Figure 1. Tasks were first performed in a self-selected manner, and then with varied starting positions (e.g., hunched sitting, weight on heels) to capture movement variation. Stand-To-Sit included a 3-second pause in standing. Each participant completed five trials per task, resulting in 100 recordings.

b) **Instrumental movement analysis** was conducted simultaneously using the MoLab™ three-dimensional movement analysis system with seven sensors measuring joint angles, acceleration, and gyroscope data on lower spine and legs, and the RSscan one-meter pressure mat with sensors registering pressure distribution, center of force line, and stability parameters.

c) **Expert assessments** of the video recordings were performed by three physiotherapists using a modified IRAF protocol, which included weighted components, confidence ratings, and whether assessments were possible or not. Each transfer was first assessed individually, followed by a joint discussion to reach consensus scores.

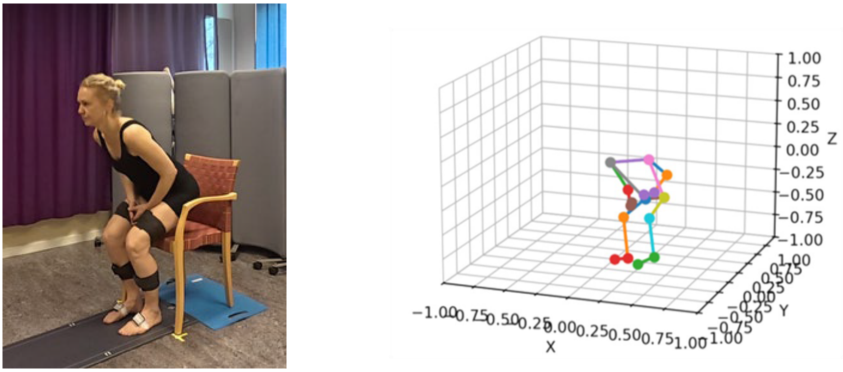

Figure 1. Data collection Sit-To-Stand by video, MoLab™ and RSscan (left), SAT analysis (right)

3.1. Data analysis

3.1.1. OMA and the IRAF protocols

In preparation of the SAT analysis, different aggregated scores were computed. Aggregation of the individual scores was applied **horizontally** integrating the scores of the individual experts to a common score and **vertically** integrating scores of aspects to one score per movement phase, further to one score per complete movement, and finally to one score per person.

More specifically, the data were preprocessed from Excel to CSV and loaded into a Jupyter Notebook. Using *Pandas* and *SciPy* libraries, scores were averaged across the three experts and compared with experts' joined scores. Pearson's correlation assessed linear associations, while the Intraclass Correlation Coefficient (ICC, two-way mixed, absolute agreement) [8] quantified agreement between fixed graders. In addition, the Cohen's linear weighted kappa was used to evaluate internal consistency among graders.

3.1.2. SAT Analysis

In our ongoing analysis, AI models will be trained to map predictor variables (SAT recordings, MoLab, and RSscan) to OMA expert assessments as the response variable. SAT data will undergo pose detection using pretrained models (e.g., MoveNet for 2D or ML Kit Pose Detection for 3D), and all data sources will be preprocessed, temporally aligned (manually, via Dynamic Time Warping, or semi-automatically), and augmented (e.g., mirrored poses). Models will be trained using regression across varying aggregation levels of the response variable and different predictor sets and evaluated using Mean Absolute Error and Mean Percentage Error. Given the small sample size, the analysis aims to explore the feasibility of learning meaningful mappings rather than producing generalizable results.

4. Results

The preliminary results are shown in Table 1, indicating that the time-consuming consent workshop with all experts is dispensable.

Table 1. Correlation between automatically aggregated expert scores vs jointly aggregated scores.

Measure	Per Aspect	Per Phase	Per Movement	Per Participant
Pearsons's r, p-value, 95% CI*	0.904, 0.000, [0.870, 0.929]	0.925, 0.000, [0.899, 0.945]	0.966, 0.000, [0.953, 0.975]	0.968, 0.006, [0.956, 0.976]
R2	0.816	0.857	0.933	0.936
ICC3, 95% CI	0.977, [0.800, 1.000]	0.912, [0.820, 0.960]	0.942, [0.790, 0.990]	0.938, [0.540, 0.990]
Cohen's kappa	0.843	0.830	0.783	1.000

*Confidence interval (CI)

The strong correspondence between automatically aggregated and consensus scores across all aggregation levels (Pearson's $r = 0.90–0.97$, ICC = 0.91–0.98, $\kappa = 0.78–1.00$)

suggests that automatic aggregation provides reliable variables for use in the next step of machine learning training.

5. Discussion and Conclusion

The promising results for the AI-tool's ability to automatically aggregate the OMA expert scores motivates well the next step in the development of the prototype. This means an important step towards developing a movement screening tool that can provide accurate assessment and feedback regarding mobility, balance and movement quality based on a 2D smartphone video recording. This hold promises that the screening tool will eventually be able to learn new specific movements/skills based on movement experts' assessments. Compared to previous research [9] such a tool adds feasible assessment of movement quality, and offering personalized feedback based on daily life movements performances. The use and value of such a tool is manifold; a) motivating individuals to prevent mobility deterioration, evaluating training efforts and facilitating development of sustainable human movements, b) in primary health care to detect early deterioration and individualizing the support, c) supporting physiotherapy learning and training in structured observational movement analysis, d) to explore correlations between movement deviations and health-related aspects in large clinical studies, and e) assessing a subliminal labelling approach that, if successful, allows to scale these values a)-d) to a great variety of different movements tailored to different ages, (dis-) abilities, goals, etc. Hence, there are large possibilities for extended applications.

References

[1] Backåberg S, Eriksson Östh K, Kimming A, Frykberg GE. Sustainable human movements – a threshold concept with potential to open up new perspectives in physiotherapy. Eur. J. Physiother. 2024;(2),72–80, doi:10.1080/21679169.2024.2343290

[2] Backåberg S, Brunt D, Rask M, Gummesson C. Experiences of using a video-based learning model during a long-term process of movement awareness and learning–a hermeneutical study. Eur. J. Physiother. 2021;(1):41-7, doi:10.1080/21679169.2019.1635639

[3] Strandberg S, Ekstedt M, Fagerström C, Backåberg S. Cocreation of a Video Feedback Tool for Managing Self-Care at Home With Pairs of Older Adults: Remote Experience-Based Co-Design Study. JMIR Form. Res. 2024;8:e57219, doi:10.2196/57219

[4] Backåberg S, Hellström A, Fagerström C, Halling A, Lincke A, Löwe W, Ekstedt M. Evaluation of the skeleton avatar technique for assessment of mobility and balance among older adults. Front. Comput. Sci. 2020;2:601271, doi: 10.3389/2020.601271

[5] Lincke A, Fagerström C, Ekstedt M, Löwe W, Backåberg S. A comparative study of the 2D-and 3D-based skeleton avatar technology for assessing physical activity and functioning among healthy older adults. Health Informatics J. 2023;(4):14604582231214589, doi:10.1177/14604582231214589.

[6] Quinn L, Riley N, Tyrell CM, et al. A Framework for Movement Analysis of Tasks: Recommendations From the Academy of Neurologic Physical Therapy's Movement System Task Force. *Phys Ther*. 2021;101, doi:10.1093/ptj/pzab154

[7] Backåberg, S, Kimming, A, Kraft, P, Reifeldt, K. IRAF – Instrument for Movement Analysis of Person Transfer and Mobility in Daily Living. (2020) Report. OAI:DiVA.org:lnu-97850. Available at: http://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-97850

[8] Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979;86(2):420. doi:10.1037//0033-2909.86.2.420

[9] Milone D, Longo F, Merlino G, De Marchis C, Risitano G, D'Agati L. MocapMe: DeepLabCut-Enhanced Neural Network for Enhanced Markerless Stability in Sit-to-Stand Motion Capture. Sensors (Basel). 2024;(10):3022. doi:10.3390/s24103022