Walking-age estimator based on gait parameters

Tomohito Kuroda
Department of Computer Science
Tokyo Metropolitan University
Hino, Japan

Shogo Okamoto
Department of Computer Science
Tokyo Metropolitan University
Hino, Japan

Yasuhiro Akiyama
Dept. Mechanical Engineering and Robotics
Shinshu University
Ueda, Japan

***Abstract*—Human gait motions differ depending on age. We estimated peoples' age using kernel regression analysis with reported height and weight and representative gait parameters as explanatory variables. The samples were Japanese women aged in their 20s-70s represented in an open database. The estimated ages exhibited a correlation coefficient of 0.65 with the actual ages, and the mean absolute error was 17.7 years. Estimation errors were significantly greater for older adults than for young people. Given that the method involves gait parameters that can be measured using wearable devices, such as inertial measurement units, this technique provides an accessible way for people to independently monitor their physical health in daily life.**

***Index Terms*—gait parameters, age estimation, kernel regression**

I. Introduction

Human gait motions differ depending on age [1]–[3]. For example, Ko et al. [1] showed that the values of several gait parameters change with age. Specifically, gait speed and stride length decrease with age, whereas step width increases. For therapists and training experts, it would be helpful to know patients' or trainees' biological age as well as the age group to which their gait patterns belong. Furthermore, understanding the age group with which one's gait characteristics are associated may help individuals maintain a healthy lifestyle.

The majority of previous studies used camera images to estimate the ages of passers-by. Begg et al. [4] used a support vector machine to classify young and older adults according to gait features extracted from images. They aimed to classify walkers' ages into a few categories. By contrast, Makihara et al. [5] utilized frequency spectra computed from movies or successive pictures and Gaussian process regression to estimate the ages of people aged in the range of 2 years to 94 years. In these studies [4], [5], walking people were monitored on cameras installed in the environment. Although camera images provide a variety of information, in some situations a camera may not be available or suitable. For example, when collecting a single person's daily life log, continuous measurement by camera is not expected.

As shown in Fig. 1, we used kernel regression analysis to estimate walkers' ages based on their gait parameters including mediolateral and anterior velocity, stride length, step width, swing duration, and minimum foot clearance, as well as their height and weight. The gait parameters were semantically well-defined, and many were measured using a wearable device equipped with an inertial measurement unit. We also used people's self-reported parameters, such as their height and weight. Such a system could help people monitor their physical health in their daily lives, as opposed to being utilized for security or commercial surveillance purposes.

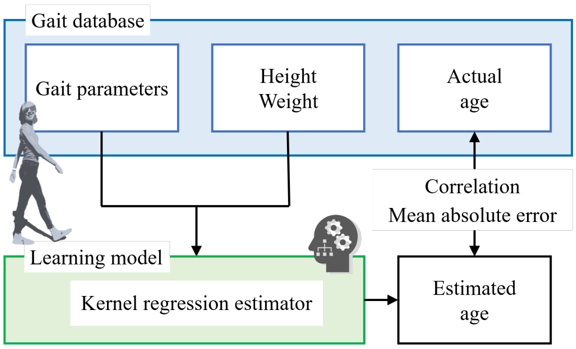

Fig. 1. Scheme of this study. We estimated age based on walkers' height, weight, and gait parameters using kernel regression analysis.

II. Methods

A. Data from an open gait database

We used an open gait database [6], which provides three-dimensional time-series coordinates of body segments during normal gait. Participants walked barefoot for 10-meter in a straight line, and approximately one gait cycle in the middle of their walk was recorded. In total, the positions of body segments of the ten gait cycles for individual participants were recorded by an optical system. The database has been used, for example, for computing the gait stability during normal walking [2], [7], [8]. We used five gait cycles starting with right heel contact for 75 women aged 20–75. In total, we used 375 samples collected from 75 healthy women.

B. Parameters used to estimate age

We used height, weight, and seven popular gait parameters: the maximal mediolateral and anterior speeds of the center of mass, step width, stride length, cadence, swing duration, and minimum foot clearance during normal gait. Table I shows those parameters' means and standard deviations. Step width is the distance between the right and left toes in the mediolateral direction. Stride length is the anterior distance between two successive steps taken using the left foot. Cadence is the number of strides per minute. Swing duration is the proportion of the gait cycle for which the swing phase

979-8-3503-4017-4/23/$31.00 ©2023 IEEE

TABLE I
MEANS AND STANDARD DEVIATIONS OF PARTICIPANTS' HEIGHT, WEIGHT, AND GAIT PARAMETERS

Parameters	
Height (cm)	157.53 ± 5.81
Weight (kg)	54.32 ± 7.94
Minimum foot clearance (m)	0.020 ± 0.006
Maximal mediolateral CoM speed (m/s)	0.11 ± 0.029
Maximal anterior CoM speed (m/s)	1.48 ± 0.17
Step width (m)	0.12 ± 0.031
Stride length (m)	1.29 ± 0.11
Cadence (steps/min)	58.56 ± 3.80
Swing duration (%)	85.50 ± 5.02

accounts. Minimum foot clearance is the distance between the foot and the ground when the foot is parallel to the ground.

C. Data analysis

Kernel regression analysis, a type of nonlinear regression analysis, was used to estimate age. In this method, the objective variable is represented by a linear summation of kernel functions. A kernel function is defined by

$$k(\boldsymbol{x}, \boldsymbol{x}') = \exp\left(-\beta \|\boldsymbol{x} - \boldsymbol{x}'\|^2\right) \quad (1)$$

where $\boldsymbol{x}$ and $\boldsymbol{x}'$ are input vectors, β is a pre-determined hyperparameter. The regression coefficients for the kernel functions are determined by minimizing the squared error between observation and estimation of the objective values. Additionally, a regularization parameter is usually introduced to avoid over-training. In this study, we estimated the ages of people represented in the database using Gaussian kernels with their height, weight, and gait parameters as explanatory variables.

The regularization parameter used in the kernel regression analysis was optimized by cross-validation. For this purpose, 375 samples were split into five groups, each of which included only one sample per individual. One of the five datasets was used as a test dataset, and the rest were used as training data. A model was created using the training data, and we estimated ages using the test data. The correlation between the actual age of each person represented in the test data and the age estimated via kernel regression analysis was calculated. This process was repeated five times, with each of the five groups supplying the test dataset. We then determined the normalization parameter such that the mean correlation coefficient among the five groups was maximized.

III. Results

Ages estimated via cross-validation as well as actual ages are shown in Fig. 2. Here, each sample was estimated by the model trained on the dataset excluding itself. The estimated ages exhibited the correlation coefficient of 0.65 with the biological ages of the people in the dataset. The mean absolute error of age was 17.7 years. Estimated ages ranged from 0 – 77 years, and people's actual ages ranged from 20 – 75 years. The maximum mean absolute error was 66.7 years, and the minimum was 0.07 years. These estimation errors exceeded those reported in an earlier study Makihara et al. [5] conducted, potentially because of the quality of the data available for the estimation. In the earlier study [5], video images of walking silhouette were used, whereas the present study merely used several parameters.

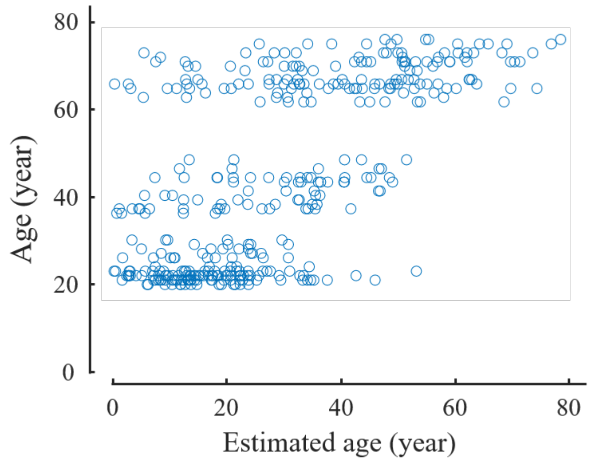

Fig. 2. Scatter plot of walkers' actual and estimated ages.

IV. Discussion

As shown in Fig. 2, estimation errors largely depended on the actual ages, and errors increased with participants' actual age. The mean absolute error for young adults (aged 20—30 years) was 9.52 years. Those for the middle-aged (aged 35-50 years) and older adults (60—75 years) groups were 15.82 years and 26.90 years, respectively. An F-test for equal variance of errors was conducted for the young and older adult groups, and the p-value was less than 0.001, rejecting equal variance ($F(149, 149) = 6.31$). Similarly, the error variance for the older adult group was greater than that for the middle-aged group ($F(149, 74) = 2.05$, $p < 0.001$). These results indicate that the estimation errors increased with participants' actual age. According to Schrager et al. [9], the variability of walking velocity and step length increases with age, which is believed to be owing to differences in individual physical abilities. Such differences may amplify with age.

V. Conclusion

We estimated walking motion age based on easily interpreted gait parameters in combination with self-reported height and weight, whereas previous studies largely used temporally successive images of a walker. Many of the gait parameters can be estimated by a body-equipped inertial measurement unit, and the model in this study will be useful for suggesting the user's health status based on his or her gait in daily life. The correlation coefficient between the estimated and actual ages was 0.65, and the mean absolute error was 17.73 years. The higher the actual age, the larger the estimation error. In the future, we will select appropriate parameters or explanatory variables to improve the estimation accuracy.

REFERENCES

[1] S. Ko, J. M. Hausdorff, and L. Ferrucci, "Age-associated differences in the gait pattern changes of older adults during fast-speed and fatigue conditions: results from the baltimore longitudinal study of ageing," *Age and Ageing*, vol. 39, pp. 668–694, 2010.

[2] T. Yamaguchi and K. Masani, "Effects of age on dynamic balance measures and their correlation during walking across the adult lifespan," *Scientific Reports*, vol. 12, p. 14301, 2022.

[3] T. Kuroda, S. Okamoto, and Y. Akiyama, "Anterior and mediolateral dynamic gait stabilities attributed to different gait parameters in different age groups," *Journal of Biomechanical Science and Engineering*, 2023.

[4] R. K. Begg, M. Palaniswami, and B. Owen, "Support vector machines for automated gait classification," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 5, pp. 828–838, 2005.

[5] Y. Makihara, M. Okumura, H. Iwama, and Y. Yagi, "Gait-based age estimation using a whole-generation gait database," 10 2011, pp. 1–6.

[6] Y. Kobayashi, N. Hida, K. Nakajima, M. Fujimoto, and M. Mochimaru, "AIST gait database 2019," 2019. [Online]. Available: https://unit.aist.go.jp/harc/ExPART/GDB2019.html

[7] T. Inagaki, Y. Akiyama, S. Okamoto, T. Mayumi, and Y. Yamada, "Relationship between gait stability indices and gait parameters comprising joint angles using walking data of 288 people," *Nagoya Journal of Medical Science*, vol. 85, no. 2, pp. 211–222, 2023.

[8] T. Iwasaki, S. Okamoto, Y. Akiyama, and Y. Yamada, "Gait stability index built by kinematic information consistent with the margin of stability along the mediolateral direction," *IEEE Access*, vol. 10, pp. 52 832–52 839, 2022.

[9] M. A. Schrager, V. E. Kelly, R. Price, L. Ferrucci, and A. Shumway-Cook, "The effects of age on medio-lateral stability during normal and narrow base walking," *Gait & Posture*, vol. 28, pp. 466–471, 2008.