

# Application of dynamic time warping algorithm for pattern similarity of gait

Hyun-Seob Lee\*

Department of Physical Education, Korea University, Seoul, Korea

The purpose of this study was to investigate the effectiveness of dynamic time warping (DTW) in gait research. Participants in this study were consist of 10 males and 10 females. Equipment used for collecting the gait data of participants in this study was three-dimensional (3D) motion analysis system consisted of 8 infrared CCD cameras operated with a sampling frequency of 120 frames/sec. DTW program used in this study was made using the MATLAB and the normal operation of the DTW program was verified by comparison of result manually calculated and output by the DTW program. Flexion angle of the knee joint of both feet obtained by 3D motion analysis system was analyzed by the DTW program and symmetry index (SI) equation. Statistical analysis of the values obtained by DTW was performed by one-sample *t*-test in confi-

dence interval (CI) 99%, 95%, 90%, 85%, and 80% each using the SPSS. The subjects' left and right legs were compared 20 times, and other steps of the same foot were compared 20 times. In this study, DTW showed different results from SI which is generally used to test the similarity of gait. Compared to that of DTW, the threshold figure for similarity evaluation in SI, 10%, is considered too large/high. When the CI threshold figure of 95% was adopted in statistical analysis, DTW demonstrated a lower rate of judging two sequences as similar even in the case of normal gait. This study suggests that DTW can be used for the similarity test of gait research.


**Keywords:** Similarity, Dynamic time warping, Gait, Symmetry index

## INTRODUCTION

In biomechanics, with its main academic focus on human movements, time-series data such as spatiotemporal, kinetics, and kinematics variables is frequently analyzed as it is useful and easy to understand when describing or explaining the human movement. Periodicity is often used as the analysis unit of research in biomechanics since human movement in gait or sports has periodicity with repeats of motion for a certain period of time. Prime examples of motions with periodicity are that of the lower and upper extremity in gait, pitching motion in baseball, and various technical motions in gymnastics. Periodicity is most frequently structured in the form of time-series data. Analysis of the time-series data reveals a specific pattern or a sequence, which is also called a sequence pattern or a signal pattern depending on the field of study. Many researchers analyze such a pattern as part of their study since it is useful in explaining and understanding movements, as well

as in identifying differences with other subjects or similar motions. However, evaluation of pattern similarity is often challenging. For instance, one has to identify where and to what degree the left and the right foot in gait diverge from each other. Analyzing the pattern similarity of time-series data can be carried out by either visually confirming the graph etc. or using mathematical and statistical calculation.

The visual method is intuitive and convenient in grasping the overall tendency, but the objectivity of the study can be compromised. The mathematical methods such as correlation analysis or Euclidean distance method have higher degree of objectivity compare to the visual method, but there are limitations in comparing various patterns. While the correlation analysis can measure changes in the two patterns in comparison, it does not properly reflect the time-shift phenomenon, difference in the time length, or the difference of amplitude between two patterns. The method using the Euclidean distance also has a limitation in comparing two pat-

\*Corresponding author: Hyun-Seob Lee  <https://orcid.org/0000-0001-7735-7081>  
Department of Physical Education, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul 02841, Korea  
E-mail: lhs.phd@gmail.com  
Received: April 5, 2019 / Accepted: May 1, 2019

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

terns with different time frames.

The dynamic time warping (DTW) algorithm was developed to solve the difficulties in analyzing pattern similarity for time-series data. DTW is utilized in various academic fields such as word recognition (Gilbert et al., 2010; Rodriguez-Serrano and Perronin, 2012), on-line handwriting and signature recognition (Efrat et al., 2007; Tappert et al., 1990), off-line recognition of a small gesture and vision-based gesture (Bautista et al., 2016; Santos et al., 2015), data mining (Shokoohi-Yekta et al., 2017), computer animation, animated films, and chemical engineering (Vial et al., 2009). DTW is known as the best algorithm for evaluating the similarity between two time-series data that vary in time frames as well as nonlinear time-series data. In addition, DTW is able to statistically determine the degree of similarity by normalization of the matching cost calculated in DTW. DTW performed in three conditions (monotonicity, continuity, and boundary) has been continuously improved in its efficiency by adding extra conditions using dynamic programming (DP) such as warping window or slope constraint. DTW can be useful for rehabilitation, treatment, and development of exercise programs because it can identify the degree of similarity and where the difference appeared in time-series data. However, DTW has not yet been actively used in gait research. While some previous studies (Patterson et al., 2010; Pulido-Valdeolivas et al., 2018) applied DTW to their research, they lack analysis of the effectiveness or pros and cons of DTW itself in gait research. This study therefore aims to confirm the effectiveness of DTW in gait research.

## MATERIALS AND METHODS

### Participants and equipment

Participants in this study were consist of 10 males and 10 females, and anthropometric data of participants are as shown in Table 1. Sample size was calculated as 17 by Gpower of setting to an effect size of 0.8, significant level  $\alpha = 0.05$ , and statistical power of 0.85. Actual power of this study was calculated as 0.92 due to sample size of 20. The subjects were instructed to refrain from unusual activities such as drinking, excessive exercise, insufficient

**Table 1.** Anthropometric data of participants

Sex	Age (yr)	Height (cm)	Weight (kg)
Male	23.7 ± 1.8	177.1 ± 4.2	68.4 ± 5.9
Female	21.5 ± 1.0	164.6 ± 3.3	54.7 ± 7.1
Total	22.6 ± 1.8	170.8 ± 7.4	61.5 ± 9.5

Values are presented as mean ± standard deviation.

sleep, etc. and it was confirmed by interview at the day of the experiment. The subject carried out self-selected normal walking practice for 5 min on barefoot at the experiment site to prevent abnormal walking due to psychological tension from a new environment.

Equipment used for collecting the gait data of participants in this study was three-dimensional (3D) motion analysis system (MotionAnalysis Corp., Santa Rosa, CA, USA) consisted of 8 infrared CCD cameras operated with a sampling frequency of 120 frames/sec. Collected data by 3D motion analysis system was analyzed by Cortex and OrthTrak provided by MotionAnalysis Corp. Matlab was used for smoothing by cubic spline of curve fitting tool.

### DTW and symmetry index

DTW is a well-known algorithm to find an optimal alignment between two time-series sequences under three conditions with boundary condition, monotonicity condition, continuity condition. Definition of these conditions are following as:

Boundary condition:

$$p_1 = (1, 1) = c(x_1, y_1) \text{ and } p_L = (m, m) = c(x_i, y_i)$$

Monotonicity condition:

$$m_1 \leq m_2 \leq \dots \leq m_L \text{ and } n_1 \leq n_2 \leq \dots \leq n_L$$

Continuity condition:

$$p_{L+1} - p_L = \in \{(1,0), (0,1), (1,1)\}$$

Where  $p$  is the path found by DTW,  $c$  is the local cost in warping path,  $x, y$  is sequence data and  $m, n$  is index number of warping path, and  $L$  is a maximum length of warping path.

DTW finds the optimal warping path by transform two sequences into the matrix. In the process, different of the time length of two sequences are normalized to optimal warping path length. DTW calculates the different value of each element of two sequences and is called local cost. The optimal warping path is the path with the smallest sum of these local costs and the sum of local cost is called global cost. Definition of global cost and optimal warping path are as following.

$$\text{Global cost: } c_G(X, Y) = \sum_{i=1}^L c(x_{iL}, y_{iL})$$

Optimal warping path:

$$DTW(X, Y) = c_{OP}(X, Y) = \min\{c_G(X, Y)\}$$

The optimal warping path mentioned above can be found by the following equation.

$$\gamma(i, j) = d(x_i, y_j) + \min\{\gamma(i-1, j-1), \gamma(i-1, j), \gamma(i, j-1)\}$$

Where  $\gamma$  is distance accumulated of warping path and  $d$  is distance between each element of two sequences.

If the difference between two sequences does not exist, the

warping path will be diagonal of matrix.

Researchers have applied various global constraints to DTW for reducing the run time of DTW and it is called DP. The most commonly known DPs in literature are the Sakoe-Chiba band and the Itakura (1975) parallelogram.

Traditionally, the symmetry index (SI) was used for evaluating similarity in gait research and equation of SI is as follows.

$$SI = \frac{X - Y}{0.5 \times (X + Y)} \times 100\%$$

Where  $X$ ,  $Y$  is sequence. When  $SI < 10\%$ , the symmetry was considered acceptable (Perttunen et al., 2004). In this study, results of DTW and SI were compared.

### Analysis

DTW program used in this study was made using the MATLAB and the normal operation of the DTW program was verified by comparison of result manually calculated and output by the DTW program. Flexion angle of the knee joint of both feet obtained by 3D motion analysis system was analyzed by the DTW program and SI equation. Values of different between two sequences corresponding to the index of the warping path were calculated for statistical analysis. At this time, the local cost of the warping path was not used as the different values. Because the local cost is a value obtained by taking a square, minus value of deviation disappears in local cost and it is unsuitable as raw data for statistical analysis. Statistical analysis of the different values obtained above was performed by one-sample  $t$ -test in confidence interval (CI) 99%, 95%, 90%, 85%, and 80% each using the SPSS. The subjects' left and right legs were compared 20 times, and other steps of the same foot were compared 20 times. And the DTW analysis results and SI analysis results were compared.

## RESULTS

Table 2 shows result of similarity test by one-sample  $t$ -test for elements not taken a square of warping path includes SI test result. In the male, there was only one case showing different result according to CI when comparison of right and left foot, and showed two cases when a comparison of another step of the same foot. In the female, there were two cases according to each comparison. Tables 3 and 4 show the comparative result between DTW and SI based on the CI of 5%. In the male, DTW and SI showed different results of 21 cases (52.5%) in total and in the female, showed different results of 25 cases (62.5%). All of the different results evaluated to similar in SI but DTW was the opposite.

**Table 2.** Result of similarity test

Similarity test	Confidence interval					SI
	DTW 99%	DTW 95%	DTW 90%	DTW 85%	DTW 80%	
Male						
Similarity	13	10	8	7	7	31
Nonsimilarity	27	30	32	33	33	9
Female						
Similarity	7	5	4	2	2	30
Nonsimilarity	33	35	36	38	38	10

DTW, dynamic time warping; SI, symmetry index.

**Table 3.** Comparison between DTW result and SI result for male (CI: 5%)

SI	DTW			
	Male (right vs. left)		Male (other step of same foot)	
	Similarity	Nonsimilarity	Similarity	Nonsimilarity
Similarity	5	7	5	14
Nonsimilarity	0	8	0	1

DTW, dynamic time warping; SI, symmetry index; CI, confidence interval.

**Table 4.** Comparison between DTW result and SI result for female (CI: 5%)

SI	DTW			
	Female (right vs. left)		Female (other step of same foot)	
	Similarity	Nonsimilarity	Similarity	Nonsimilarity
Similarity	1	12	4	13
Nonsimilarity	0	7	0	3

DTW, dynamic time warping; SI, symmetry index; CI, confidence interval.

Fig. 1 shows the warping path and alignment of two sequences about comparing right and left foot evaluated to similar in DTW and SI. And, Fig. 2 is a representative case evaluated to nonsimilar in DTW and SI. Fig. 3 is a representative case that evaluated to similar in SI but evaluated to nonsimilar in DTW.

## DISCUSSION

The purpose of this study was to investigate the effectiveness of DTW in gait research. Although DTW has been used successfully in various fields, it was relatively lacking in gait studies. The stable and periodic movement of the human body has been considered as an important factor in the gait study due to the characteristics of the gait. Therefore, analyzing the similarity of gait pattern is an important part of gait research. For this reason, this study sought to confirm the applicability of DTW for the analysis of similarity of gait. In previous studies using DTW, Gaspar et al. (2017) conducted a study to the validation of sensitivity of DTW

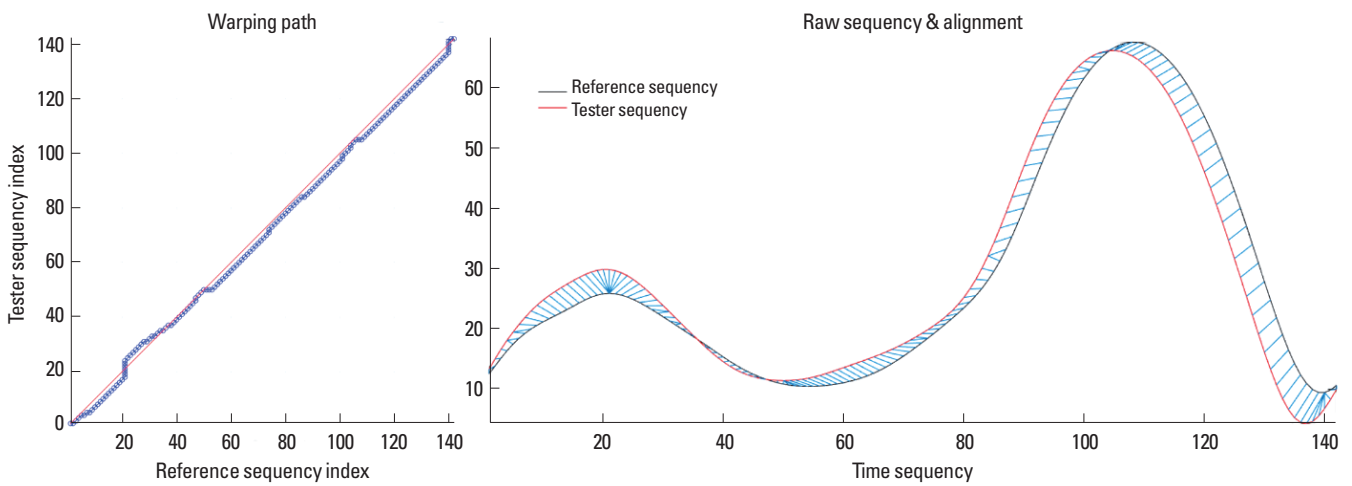


Fig. 1. Warping path, sequence and alignment of right vs. left foot (similarity). Blue circle, warping path; red straight line, diagonal of the matrix; blue line, alignment.

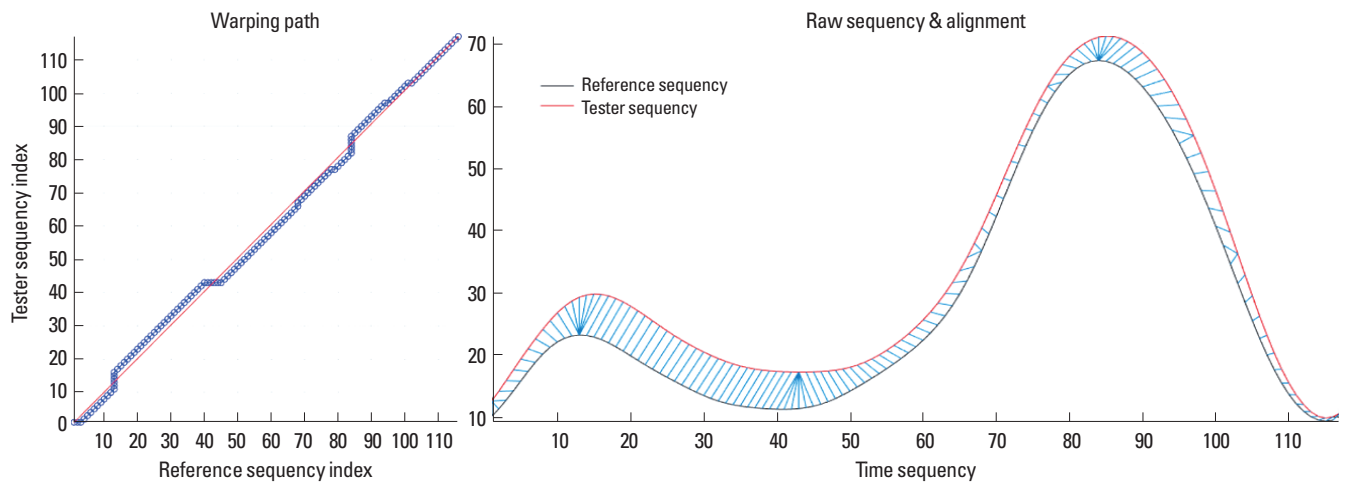


Fig. 2. Warping path, sequence and alignment of right vs. left foot (nonsimilarity). Blue circle, warping path; red straight line, diagonal of the matrix; blue line, alignment.

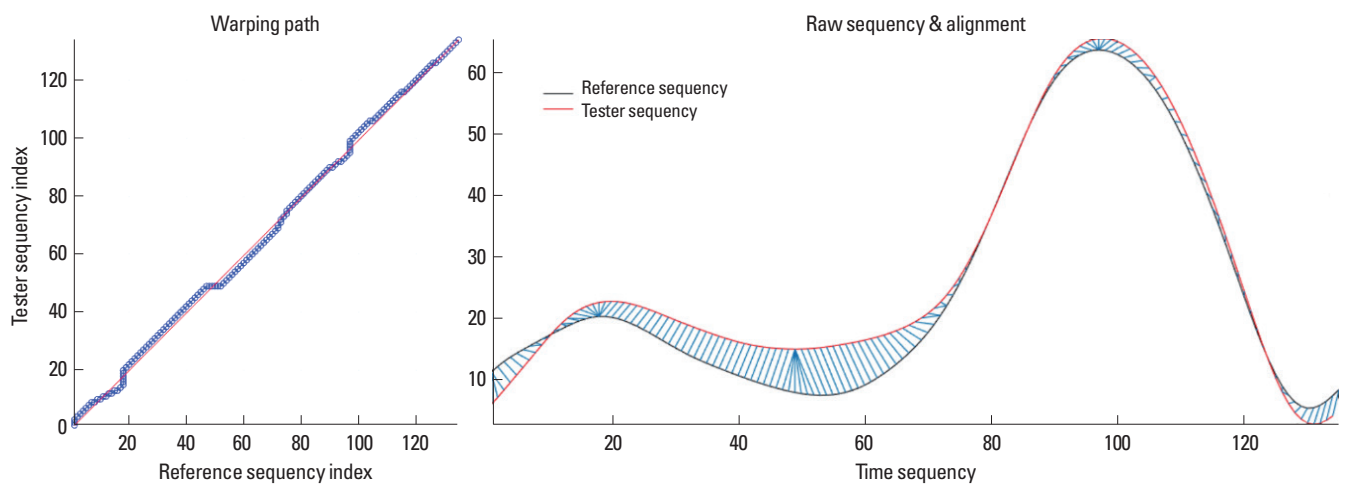


Fig. 3. Warping path, sequence and alignment of right vs. left foot (different result each other). Blue circle, warping path; red straight line, diagonal of the matrix; blue line: alignment.

by comparing to other established metric such as the Pearson correlation coefficient, root-mean-square-error and so on. And, proposed that DTW can be used in both direct and indirect quantitative validation. Barth et al. (2015) used DTW to observe the pattern change of gait. However, there have been no studies on the pros and cons of using DTW as a means of similarity testing. In this study, DTW showed different results from SI which is generally used to test the similarity of gait (Tables 3, 4). Compared to that of DTW, the threshold figure for similarity evaluation in SI, 10%, is considered too large/high. In the case of SI, because only a difference of values between the two sequences is only used as the criterion for judgment similarity, SI is considered that ability for tracking the overall pattern change and the ability to cope with the time-shift phenomenon is lower than that of the DTW.

On the other hand, when the CI threshold figure of 95% was adopted in statistical analysis, DTW demonstrated a lower rate of judging two sequences as similar even in the case of normal gait (Table 2). Gaspar et al. (2017) argued that DTW algorithm can be used for quantitative validity and that it has the analytical capability for data with time-shift. Gaspar's assertion regarding time-shift was confirmed in Figs. 1-3 presented in this study. Taken together, these results suggest that DTW can be used for the similarity test of gait research. However, since the warping path of the DTW can vary depending on the DP application method, it is important to find the appropriate DP area for the gait study. Also, since human walking is carried out with a certain degree of deviation, a walking deviation study for normal gait should be performed.

## CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

## ACKNOWLEDGMENTS

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2018S1A5B5A07073671).

## REFERENCES

- Barth J, Oberndorfer C, Pasluosta C, Schüle S, Gassner H, Reinfelder S, Kugler P, Schuldhuis D, Winkler J, Klucken J, Eskofier BM. Stride segmentation during free walk movements using multi-dimensional sub-sequence dynamic time warping on inertial sensor data. *Sensors* (Basel) 2015;15:6419-6440.
- Bautista MÁ, Hernández-Vela A, Escalera S, Igual L, Pujol O, Moya J, Violant V, Anguera MT. A gesture recognition system for detecting behavioral patterns of ADHD. *IEEE Trans Cybern* 2016;46:136-147.
- Efrat A, Fan Q, Venkatasubramanian S. Curve matching, time warping, and light fields: new algorithms for computing similarity between curves. *J Math Imaging Vis* 2007;27:203-216.
- Gaspar M, Welke B, Seehaus F, Hurschler C, Schwarze M. Dynamic Time Warping compared to established methods for validation of musculoskeletal models. *J Biomech* 2017;55:156-161.
- Gilbert JM, Rybchenko SI, Hofe R, Ell SR, Fagan MJ, Moore RK, Green P. Isolated word recognition of silent speech using magnetic implants and sensors. *Med Eng Phys* 2010;32:1189-1197.
- Itakura F. Minimum prediction residual principle applied to speech recognition. *IEEE Trans Acoust Speech Signal Process* 1975;23:67-72.
- Patterson KK, Gage WH, Brooks D, Black SE, McIlroy WE. Evaluation of gait symmetry after stroke: a comparison of current methods and recommendations for standardization. *Gait Posture* 2010;31:241-246.
- Perttunen JR, Anttila E, Södergård J, Merikanto J, Komi PV. Gait asymmetry in patients with limb length discrepancy. *Scand J Med Sci Sports* 2004;14:49-56.
- Pulido-Valdeolivas J, Gómez-Andrés D, Martín-Gonzalo JA, Rodríguez-Andonaegui I, López-López J, Pascual-Pascual SI, Rausell E. Gait phenotypes in paediatric hereditary spastic paraplegia revealed by dynamic time warping analysis and random forests. *PLoS One* 2018;13:e0192345.
- Rodríguez-Serrano JA, Perronnin F. A model-based sequence similarity with application to handwritten word spotting. *IEEE Trans Pattern Anal Mach Intell* 2012;34:2108-2120.
- Santos DG, Fernandes BJ, Bezerra BL. HAGR-D: a novel approach for gesture recognition with depth maps. *Sensors* (Basel) 2015;15:28646-28664.
- Shokoohi-Yekta M, Hu B, Jin H, Wang J, Keogh E. Generalizing DTW to the multi-dimensional case requires an adaptive approach. *Data Min Knowl Discov* 2017;31:1-31.
- Tappert CC, Suen CY, Wakahara T. The state of the art in online handwriting recognition. *IEEE Trans Pattern Anal Mach Intell* 1990;12:787-808.
- Vial J, Noçairi H, Sassi P, Mallipatu S, Cognon G, Thiébaud D, Teillet B, Rutledge DN. Combination of dynamic time warping and multivariate analysis for the comparison of comprehensive two-dimensional gas chromatograms: application to plant extracts. *J Chromatogr A* 2009;1216:2866-2872.