Advanced human motion trajectories comparison using Dynamic Path Warping approach

TOMASZ HACHAJ
Pedagogical University of Cracow
Institute of Computer Science
ul. Podchorazych 2, 30-084 Cracow
POLAND
tomekhachaj@o2.pl

MARCIN PIEKARCZYK
Pedagogical University of Cracow
Institute of Computer Science
ul. Podchorazych 2, 30-084 Cracow
POLAND
marcin.piekarczyk@up.krakow.pl

Abstract: This paper describes and evaluates advanced aligning and comparison method dedicated for human motion trajectory analysis. It utilizes Dynamic Time Warping approach and can be applied for relatively long (30 seconds or longer) and complex motion paths. In contrary to other human motion analysis techniques we do not have restriction on motion direction, we use only kinematic data and we are able to compare any foot trajectory no matter how many rotations take place during the motion. As the result an algorithm outputs set of vectors along motion path that corresponds to beginning and end positions of footsteps. The left and right foot is analyzed separately. The difference of two motion paths can be expressed in any DTW-based feature namely minimal, maximal, median, mean and normalized DTW based distance. We have evaluated our method on karate kata dataset that contains four types of motion sequences performed by two black belt Shorin-Ryu karate masters with more than 20 years of experience. The evaluation of our method assured us that our approach can be easily applied for aligning and comparison of any other motion class described by two dimensional motion trajectories. The method can be applied for example in sport or physical therapy exercises data evaluation and it is invariant to body proportion and motion execution speed.

Key–Words: Signal processing, Human motion analysis, Path analysis, Dynamic Time Warping, Karate kata.

1 Introduction

This paper is an extension of an article published in conference proceedings [11]. We have extend it significantly by adding extensive analysis of proposed approach on karate kata dataset (in original paper only one MoCap pair was analyzed and discussed). We have also extended state of the art section. In this article we revise all figures and introduce many new ones that help reader to follow the idea of our signal processing method. We have also evaluated additional DTW-based motion parameters namely minimal, maximal, median and mean DTW distances (previously we only took into account normalized DTW distance). We have also added explanation of the MoCap aligning procedure.

Nowadays not only professional but also amateur athletes might have access to computer aided systems that optimize they sport performance giving them tips or advices concerning training plan, diet and other sport activities aspects. Thanks to motion capture (MoCap) systems which becomes an affordable purchase even a medium sport club can have its own or cooperate with motion capture laboratory. MoCap systems allows to perform calculation of various motion parameters and comparing those parameters between two or more persons.

Dynamic time warping (DTW) technique is excellent tool for single and multidimensional signal aligning. Among many important application this method can be used with various human motion data, for example for motion editing, such as refining motions to meet new timing constraints or modifying the acting of animated characters [12] or multi-modal motion sequences aligning [18]. Another application of DTW is for classification purposes, for example in [9] it is applied for 2-3 seconds long karate actions classification or in [6] [3] [15] [1] [5] [4] for recognition of other classes of motions. DTW finds its application also in continues motion segmentation [13] [7] or gait classification [16]. The minimal distance methods can be applied together with DTW in order to detect abnormal behavior [2]. DTW can also be applied to perform automatic evaluation of spatial and temporal errors in sport motion [14] or physical therapy exercises [17] while using wearable motion capture sensors. In [19] authors align time series in order to compare motion capture data of two subjects performing

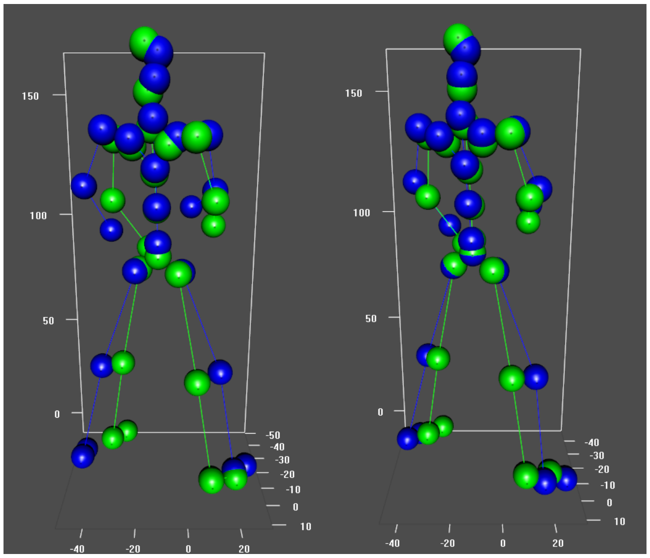


Figure 1: This figure presents the rendering of two example frames from our MoCap recordings in the skeletal form. Each sphere represents body joint position. Lines links body joints in hierarchy. The left image shows two persons (green and blue) from two recordings before aligning them to each other. The right one shows the same two persons however the green person is aligned to blue one.

similar walking action in certain direction, and alignment of two people with similar facial expressions.

In this paper we will propose and evaluate a novel DTW-based method that allows to align and compare motion trajectories of relatively long (30 seconds or longer) and complex human motion paths. In contrary to other human motion analysis techniques we do not have restriction on motion direction, we use only kinematic data and we are able to compare any foot trajectory no matter how many rotations take place during the motion. The output of our algorithm are vectors defined along motion trajectory that corresponds to beginning and end positions of steps. The left and right foot is analyzed separately. The difference of two motion paths can be expressed in any DTW-based features. The validation of the method was performed on MoCap recordings of two Shorin-Ryu karate masters. They performed karate kata which are sets of precisely defined karate movements. During kata all motions have to be performed with correct dynamic and in defined sequence of actions, however each athlete does not have a strict requirements how fast a particular kata should be performed. The kata has to demonstrate the correct actions patterns however it does not need to be done as fast as it is possible. Due to this there are large difference in timing between participants. Those factors makes this class of action excellent example for algorithm testing.

2 Material and methods

In this section we will present dataset and our methodology description. Both source code and data set we used in this paper can be downloaded from [10].

2.1 Dataset

The dataset is consisted of four types of karate kata performed by two black belt Shorin-Ryu karate masters with more than 20 years of experience. Those kata are Pinian Shodan, Pinian Nidan, Fukyugata Ichi and Fukyugata Ni. Recording was made with Shadow 2.0 MoCap system that has 17 IMU (inertial measurement units) with 3-axis accelerometer, gyroscope, and magnetometer. The tracking frequency was set to $100Hz$ with 0.5 degree static accuracy and 2 degrees dynamic accuracy. In Figure 1 we present example visualization of one MoCap frame acquired by our system. Persons we have recorded were male and female having different height, body proportions, weight and motion ranges. In Table 1 we present number of MoCap frames acquired for each kata both for the first and the second person.

2.2 Recordings aligning procedure and direction vector

In most cases two MoCap recordings of the same motion cannot be compared before some initial preprocessing. This is because depending on initial position the person is facing, the motion path might be rotated towards the other. It can be observed in Figures 4-10 (a) that present trajectories of right foot of two persons (blue and green) before aligning. The aligning procedure goes as follows: at first we center the initial hips position of both recordings so that $x_0 = 0$ and $z_0 = 0$. Next we will rotate the input recording around Y axis by the angle α calculated with Algorithm 1.

After determining the α angle we perform rotation of hips joint around Y axis. Because rotation operation is performed on root joint of hierarchical kinematic model the rotation parameter has to be changed only on this single joint. The hierarchical model can be easily recalculated to direct model, which we use later on.
The output data is than processed with heuristic method described in paper [8], which corrects the motion direction under assumption that at least one foot of an observed person remains at the ground. After above translation and rotation the aligning procedure

Table 1: This table presents number of MoCap frames acquired for each kata both for the first and the second person.

Kata	Person 1 frames count	Person 2 frames count
Pinian Shodan	3107	3824
Pinian Nidan	2838	3269
Fukyugata Ichi	3006	2794
Fukyugata Ni	3116	3797

Algorithm 1: Recordings aligning.

Data: input.Root.RotationXYZ - Euler rotation angles of root joint of input MoCap; input.LeftThigh.X - X coordinate of LeftThigh joint of input MoCap; ref.RightThigh.Z - Z coordinate of RightThigh joint of reference MoCap.

Result: angle α - after rotation around Y axis by this angle input and reference MoCap became aligned.

```
Function optimize.angle(x)
begin
    /* recalculate Euler angles to quaternion */
    q1 ← euler2quaternion(input.Root.RotationXYZ)
    /* calculate quaternion from axis angle */
    q2 ← quaterionFromAxisAngle([0,1,0],x)
    q3 ← q2 * q1
    /* recalculate quaternion to Euler angles */
    input.Root.RotationXYZ ← quaternion2euler(q3)
    /* define vectors in X-Z plain */
    v1 ← [input.LeftThigh.X,0, input.LeftThigh.Z] - [input.RightThigh.X,0, input.RightThigh.Z]
    v2 ← [ref.LeftThigh.X,0, ref.LeftThigh.Z] - [ref.RightThigh.X,0, ref.RightThigh.Z]
    /* find Euclidean distance between vectors */
    return euc.dist(v1, v2)
/* run simplex optimization */
α ← simplex(optimize.angle(x = 0))
```

is completed - see Figure 3.
Those trajectories, although are clearly visually similar, cannot be yet successfully compared by DTW. That is because typical Euclidean distance measure we use will not deal well with curvature mapping. Because of that we will create different motion primitives. At first we will determine which foot moves at certain moment of time and which remains on the ground – see Algorithm 2.

After finding lr vector we can determine when motion of left and right foot begins – see Algorithm 3.

Because Algorithm 3 is based on heuristic from Algorithm 2, vectors returned by it might not be identical with real steps taken be a person. The temporal index of beginning and end of motion are replaced by spatial parameters. An example visualization of those vectors for input and reference data is presented in Figure 4. That image shows only data for right foot. The set of vectors that starts at the beginning of foot motion and ends when this motion is finished is used for DTW with Euclidean distance function. In DTW right and left foot vectors are processed separately.

3 Results

In order to check the correctness of our approach we have tested in on dataset described in section 2.1. We wanted to check if algorithm correctly aligns motions of left and right foot of persons from our dataset. We have implemented our algorithm in R language using RMoCap package [4]. The experiment can be reproduced by using our source code and available data. In this section we will present the detailed evaluation results of our approach. Figure 2 presents aligning error plot which is a result of the DTW aligning procedure described in Section 2.2. There are four results for each kata because aligning is made in the same manner no matter if we analyze left or right foot. In Table II we present measurements taken on DTW-based paths computed between input and reference MoCap. Those are maximal, minimal, median, mean and normalized DTW distance.

DTW approach generates distance matrices which together with the warping paths supplies us with valuable information about path matching process between template and reference data. In Figures 3-9 we present plots that visualize results of our algorithm. Each figure contains four subplots: MoCap data before (a) and after aligning (b), vectors that corresponds to motions directions (c) and color-coded DTW cost matrix (d). The analysis of left and right foot motion of each kata was made separately.

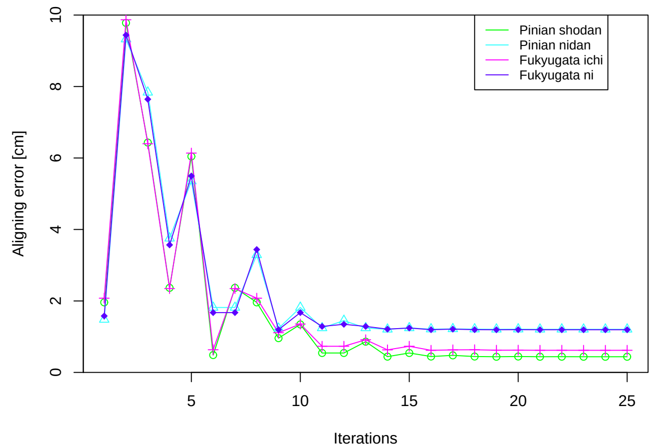


Figure 2: This figure presents aligning error plot which is a result of the DTW aligning procedure described in Section 2.2.

4 Discussion

All kata we have tested were correctly aligned according to our expectations. Values of aligning procedure error presented in Figure 2 indicate that simplex-based approach introduced in Section 2.2 coverages giving a stable solution. In case of Pinian Nidan and Fukyugata Ni kata error we do not observe large relative change of error rate. This is because input and reference MoCap were nearly aligned to each other in the beginning. This was because before each MoCap acquisition our IMU-based system was calibrated and in those two cases participants begun kata just instantly after calibration ended, so they did not have time to turn in other direction. In case of Fukyugata Ichi and Pinian Shodan kata the relative drop in error rate is clearly observable. An effect of DTW aligning procedure is also visible in Figure 3-10 in subplots (a) and (b). By comparing (a) and (b) we can clearly see that after aligning initial body position of experiment participants the motion trajectories became parallel. The subplot (c) and (d) will be discussed jointly together with results from Table II.

Basing on results from Table II we can conclude that median distance is a reliable indicator of motion similarity in the term of DTW approach however it cannot be discussed alone without knowing minimal and maximal distances values. It is obvious that normalized distance is always higher then minimal but in the same time it is sufficiently smaller than the maximal. The mean and normalized distance is

Algorithm 2: Finds which foot is in motion.

Data: RightFoot.a, LeftFoot.a - vectors containing acceleration vectors of right and left foot.
Result: Vector lr that contains 1 where right foot is moving, -1 when left foot is moving and 0 in other cases.

```
/* Perform Gaussian smoothing of foot acceleration trajectory */
ar ← smooth(norm(RightFoot.a));
al ← smooth(norm(LeftFoot.a));
for a in 1:length(ar)-1 do
    if ar[a] > al[a] and ar[a + 1] > al[a + 1] then
        lr[a] ← 1;
    else
        if ar[a] < al[a] and ar[a + 1] < al[a + 1] then
            lr[a] ← −1;
```

Algorithm 3: Finds two vectors sets for left and right foot.

Data: lr - output vector from Algorithm 2.
Result: Lists list.right and list.left that contains vectors of foot translations.

```
start ← 0;
end ← 0;
list.right ← list();
list.left ← list();
for a in 2:length(lr)-1 do
    if prev ≠ lr[a] then
        if start = 0 then
            start ← a;
        else
            end ← a − 1;
            if prev = 1 then
                list.right[length(list.right) + 1] ← [start, end];
                start ← 0;
                end ← 0;
            else
                list.left[length(list.left) + 1] ← [start, end];
                start ← 0;
                end ← 0;
    prev = lr[a];
```

Table 2: This table presents various DTW measurements (in centimeters) taken on DTW based paths computed between input and reference MoCap.

	min	**max**	**median**	**mean**	**normalized**
Pinian Shodan left	7.574	218.848	35.062	64.074	39.84
Pinian Shodan right	18.97	117.10	58.88	53.28	44.16
Pinian Nidan left	7.872	119.527	57.740	62.685	45.97
Pinian Nidan right	25.27	159.14	55.69	66.57	52.31
Fukyugata Ichi left	20.75	141.11	72.60	76.07	49.22
Fukyugata Ichi right	4.857	109.575	30.629	40.783	24.47
Fukyugata Ni left	25.39	86.84	73.71	65.73	45.97
Fukyugata Ni right	24.38	70.41	49.73	45.70	26.66

not much useful because of high variance in set of calculated distances.

In ideal case when both motion paths (input and reference) would have been partitioned in the same number of vectors and those vectors have similar magnitude and direction the warping path is a straight line with slope coefficient equals 1; however as can be seen our case this situation never happened. Also there might be some differences in number of vectors that represents the motion path. That is because our approach is based on heuristic that only estimates real footsteps. This event is clearly visible in analysis of right foot motion in Piniadn shodan kata. There are seven vectors in green path and nine vectors in blue path. All detected vectors in blue path beside those with indexes 4, 5 and 6 have equivalents in green path. This situation happened because our heuristic detected footsteps in certain direction in one recording only. This also resulted in high maximal DTW distance of Piniadn shodan in Table 2. This event however does not affect the normalized DTW distance and median DTW distance that is relatively low comparing to other results we have obtained. In all other cases our method worked as expected: we were able to correctly align complex motion paths that contained several rapid rotations and not alternate leg movements, which make this dataset very difficult to align and analyze. In other cases number of vectors that represent trajectories of input and reference data might differs between those sets however they can be nicely aligned with relatively low DTW median distance as can be seen in Table 2. The differences are caused mainly by different body proportions between recorded people and small differences in kata execution. The temporal factor of those recordings seems not to affect correctness of solution.

5 Conclusion

The additional analysis of our method that was initially proposed in [xx] assured us that our approach can be easily applied for aligning and comparison of any other motion class described by two – dimensional motion trajectories. It most cases it gives accurate results however some inaccuracies are inevitable especially in real, not artificially created dataset. Our method can successfully evaluate motion paths that were generated by people with different body proportion and motion speeds. Proposed approach can be applied for example in sport or physical therapy exercises data evaluation.

Acknowledgements: This work has been supported by the National Science Centre, Poland, under project number 2015/17/D/ST6/04051.

References:

[1] K. Adistambha, C. H. Ritz, and I. S. Burnett. Motion classification using dynamic time warping. *2008 IEEE 10th Workshop on Multimedia Signal Processing*, pages 622–627, 2008.

[2] A. Aristidou, D. Cohen-Or, J.K. Hodgins, and A. Shamir. Self-similarity analysis for motion capture cleaning. *Computer Graphics Forum*, 37(2):297–309, 2018.

[3] M. J. Black and A. D. Jepson. Recognizing temporal trajectories using the condensation algorithm. *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*, pages 16–21, 1998.

[4] Jaron Blackburn and Eraldo Ribeiro. Human motion recognition using isomap and dynamic

time warping. *Proceedings of the 2Nd Conference on Human Motion: Understanding, Modeling, Capture and Animation*, pages 285–298, 2007.

[5] Y. Chen, G. Chen, K. Chen, and B. C. Ooi. Efficient processing of warping time series join of motion capture data. *2009 IEEE 25th International Conference on Data Engineering*, pages 1048–1059, 2009.

[6] O. Cigdem, T. De Laet, and J. De Schutter. Classical and subsequence dynamic time warping for recognition of rigid body motion trajectories. *2013 9th Asian Control Conference (ASCC)*, pages 1–6, 2013.

[7] A. Gupta, J. He, J. Martinez, J. J. Little, and R. J. Woodham. Efficient video-based retrieval of human motion with flexible alignment. *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1–9, 2016.

[8] Tomasz Hachaj and Marek Ogiela. Heuristic method for calculation of human body translation using data from inertial motion capture costume. *International Journal of Electrical and Electronic Engineering & Telecommunications*, 7:26–29, 2018.

[9] Tomasz Hachaj, Marek R. Ogiela, and Katarzyna Koptyra. Application of assistive computer vision methods to oyama karate techniques recognition. *Symmetry*, 7:1670–1698, 2015.

[10] Tomasz Hachaj, Marcin Piekarczyk, and Marek R Ogiela. *R language source code and example data for this paper*, 2018. `https://github.com/browarsoftware/DTWHumanMotionPathAnalysis`.

[11] Tomasz Hachaj, Marcin Piekarczyk, and Marek R. Ogiela. Signal processing methods in human motion path analysis: a use case for karate kata. *EECS 2018 conference proceedings, in press*, 2019.

[12] Eugene Hsu, Marco da Silva, and Jovan Popović. Comparing the difference between front-leg and back-leg round-house kicks attacking movement abilities in taekwondo. *In Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on Computer animation (SCA '07). Eurographics Association, Aire-la-Ville, Switzerland, Switzerland*, pages 45–52, 2007.

[13] Mingqin Liu, Xiaoguang Zhang, and Guiyun Xu. Continuous motion classification and segmentation based on improved dynamic time warping algorithm. *International Journal of Pattern Recognition and Artificial Intelligence*, 32:59–66, 2018.

[14] Marion Morel, Richard Kulpa, Anthony Sorel, Catherine Achard, and Séverine Dubuisson. Automatic and generic evaluation of spatial and temporal errors in sport motions. *In Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, pages 542–551, 2016.

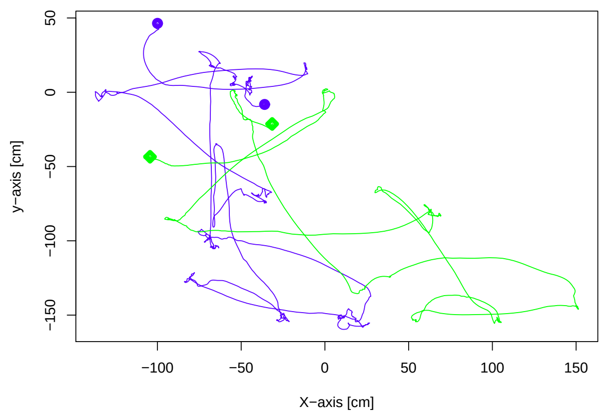
[15] A.S. Soares and Jr A.L. Apolinário. Realtime 3d gesture recognition using dynamic time warping and simplification methods. *Journal of WSCG*, 25:59–66, 2017.

[16] Adam Switonski, Agnieszka Michalczuk, Henryk Josinski, Andrzej Polanski, and Konrad Wojciechowski. Dynamic time warping in gait classification of motion capture data. *World Academy of Science, Engineering and Technology International Journal of Computer and Information Engineering*, 6:1289–1294, 2012.

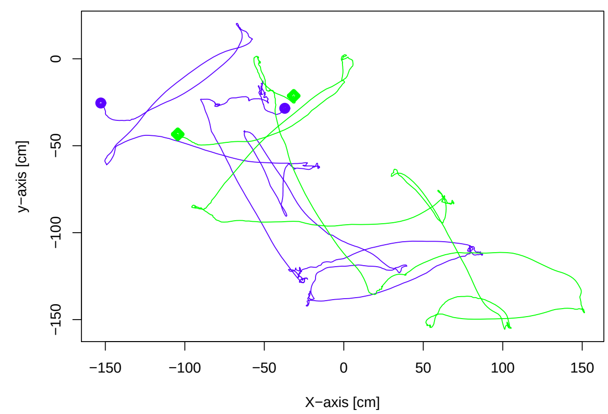
[17] Aras Yurtman and Billur Barshan. Automated evaluation of physical therapy exercises using multi-template dynamic time warping on wearable sensor signals. *Computer Methods and Programs in Biomedicine*, 117:189–207, 2014.

[18] F. Zhou and F. De la Torre. Generalized time warping for multi-modal alignment of human motion. *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1282–1289, 2012.

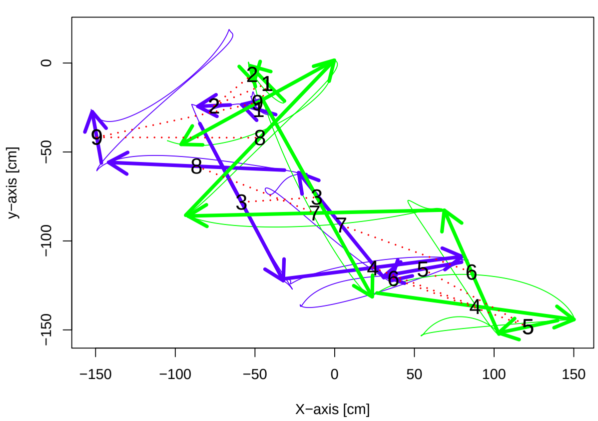
[19] Feng Zhou and Torre Fernando. Canonical time warping for alignment of human behavior. *Advances in Neural Information Processing Systems 22*, pages 2286–2294, 2009.



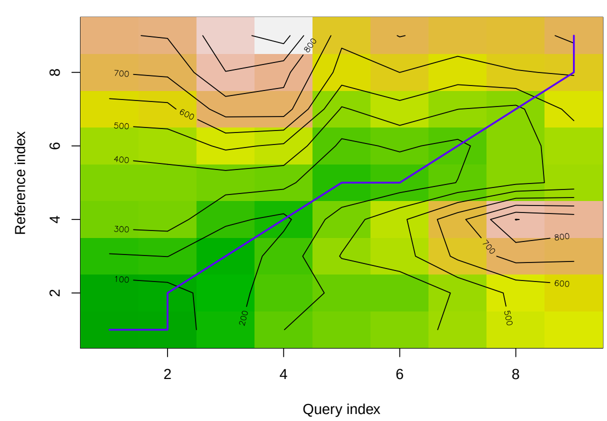
(a) Trajectories before aligning



(b) Trajectories after aligning

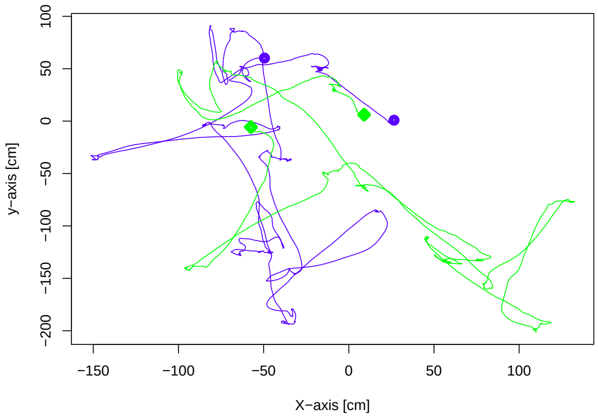


(c) Beginning and end positions of foot during motion

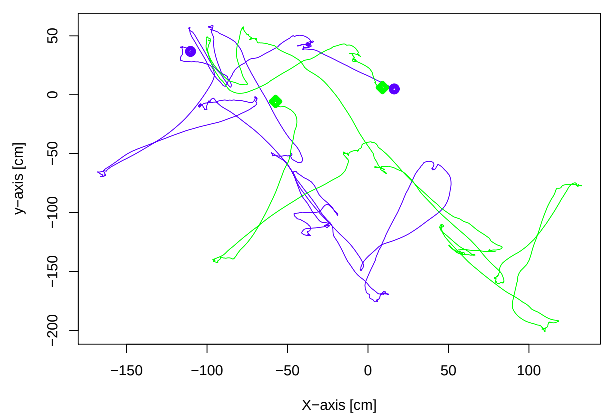


(d) DTW cost matrix

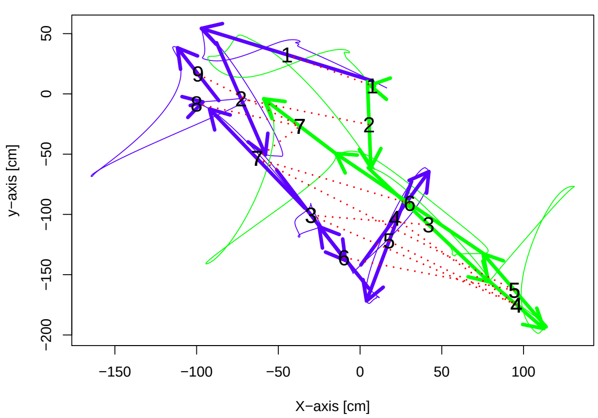
Figure 3: This figure presents results of evaluation of left foot in Pinian Shodan kata. Subplots (a), (b) and (c) present trajectories of left foot of two persons (blue and green) before (a) and after (b) aligning. Motion is projected on the plane where axis are called x and y. Y axis in all figures corresponds to Z axis of MoCap. Blue and green circles are starting and end points of motion. Subplot (c) contains the same motion trajectories like (B), however it also introduces arrows that indicate beginning and end positions of foot during motion. Those arrows might not be identical with real footsteps. Near each arrow is a number that enumerates order of motions. Red dotted lines indicate DTW alignment of blue and green person. Subplot (d) shows color-coded DTW cost matrix. Blue line is a warping path.



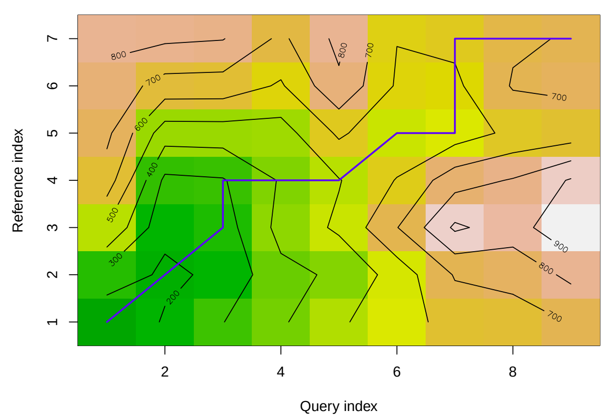
(a) Trajectories before aligning



(b) Trajectories after aligning

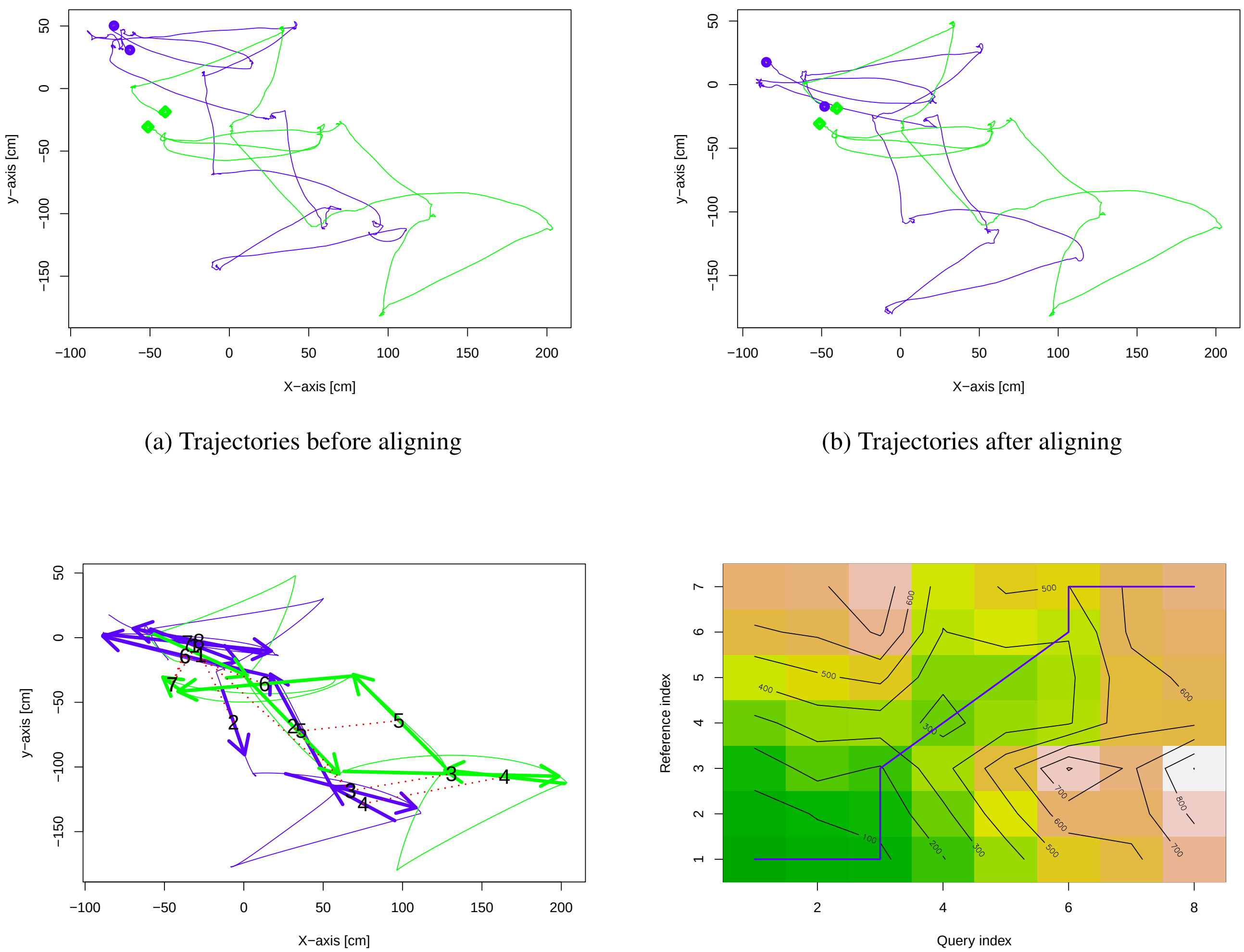


(c) Beginning and end positions of foot during motion



(d) DTW cost matrix

Figure 4: This figure presents results of evaluation of right foot in Pinian Shodan kata. Subplots have the same role as in Figure 3.



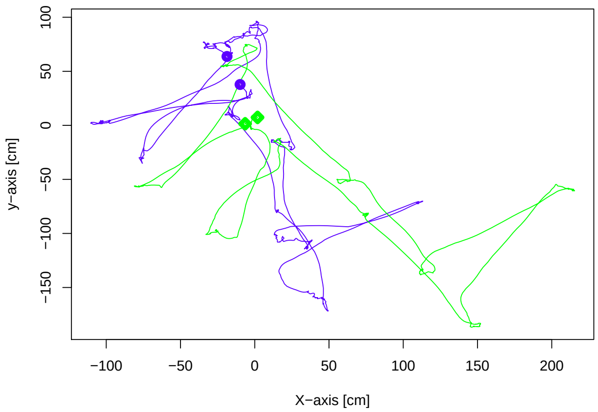
(a) Trajectories before aligning

(b) Trajectories after aligning

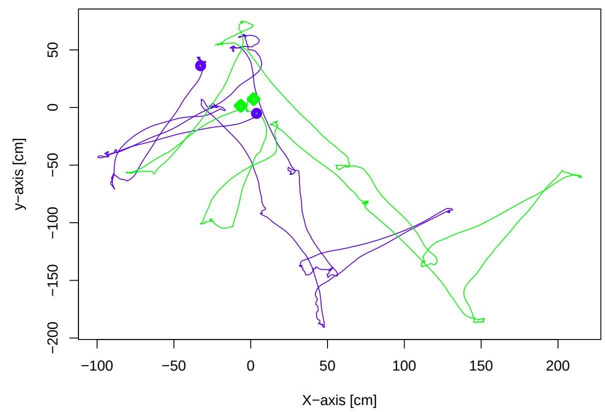
(c) Beginning and end positions of foot during motion

(d) DTW cost matrix

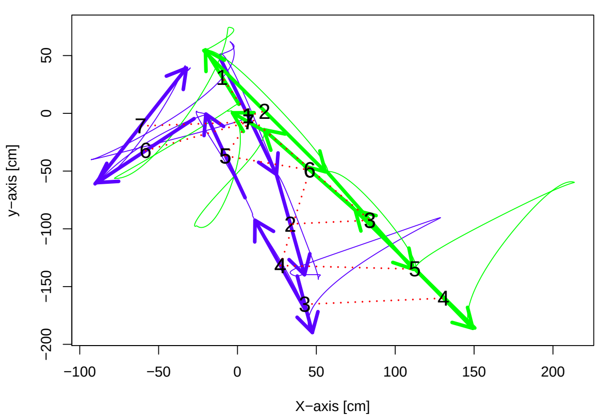
Figure 5: This figure presents results of evaluation of left foot in Pinian Nidan kata. Subplots have the same role as in Figure 3.



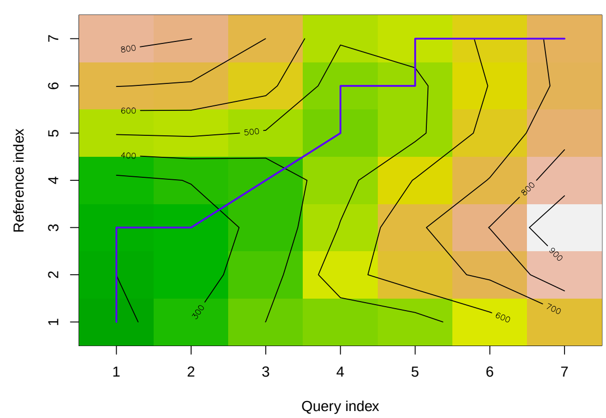
(a) Trajectories before aligning



(b) Trajectories after aligning

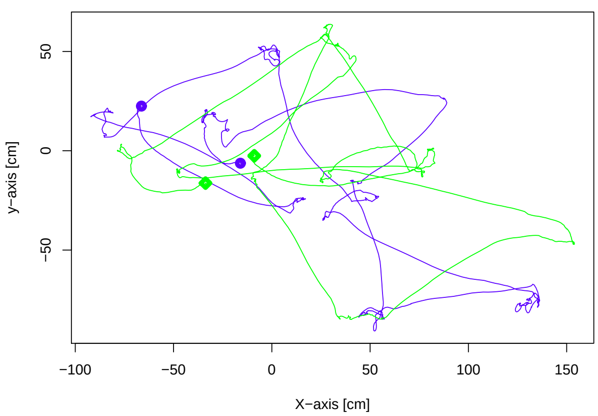


(c) Beginning and end positions of foot during motion

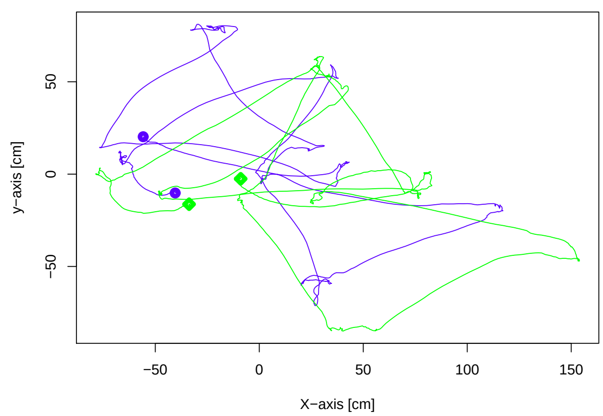


(d) DTW cost matrix

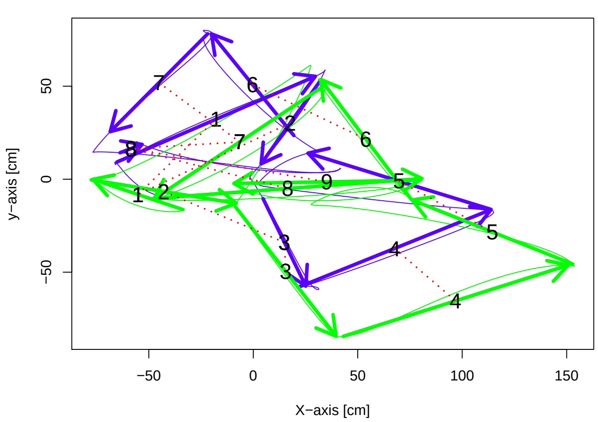
Figure 6: This figure presents results of evaluation of right foot in Pinian Nidan kata. Subplots have the same role as in Figure 3.



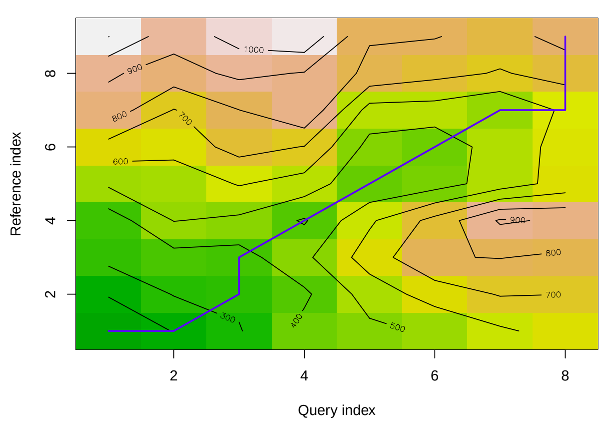
(a) Trajectories before aligning



(b) Trajectories after aligning

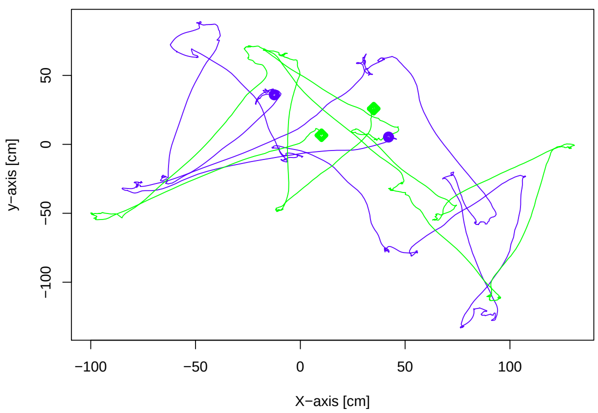


(c) Beginning and end positions of foot during motion

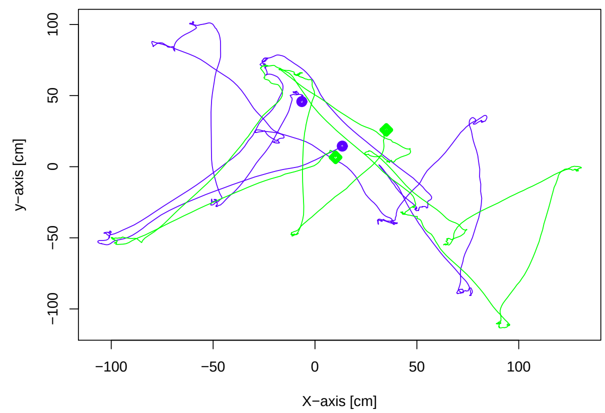


(d) DTW cost matrix

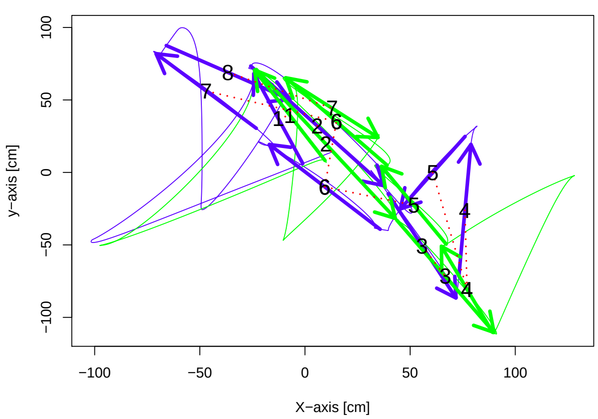
Figure 7: This figure presents results of evaluation of left foot in Fukyugata Ichi kata. Subplots have the same role as in Figure 3.



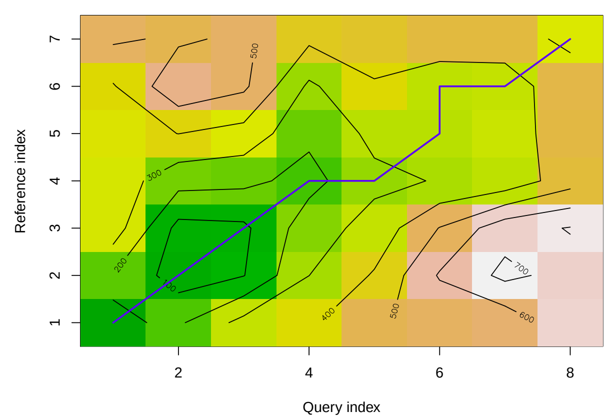
(a) Trajectories before aligning



(b) Trajectories after aligning

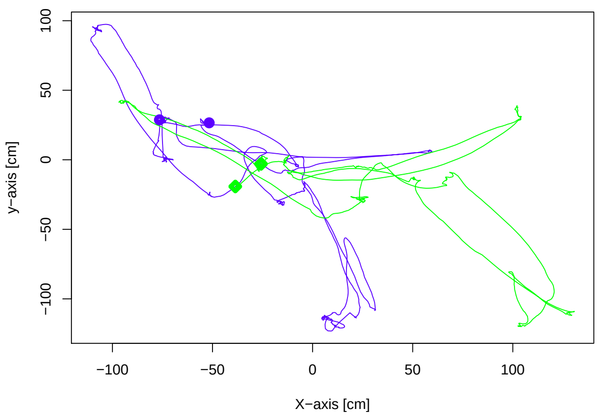


(c) Beginning and end positions of foot during motion

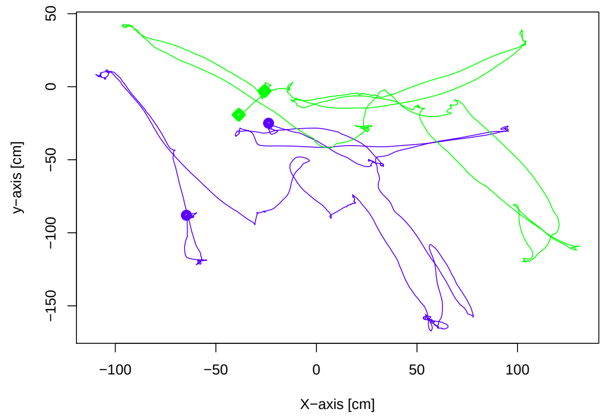


(d) DTW cost matrix

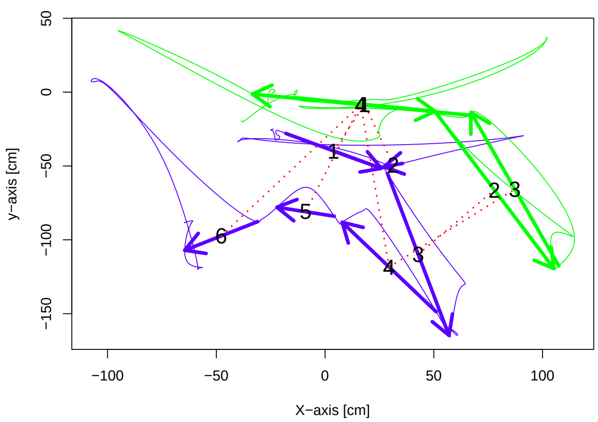
Figure 8: This figure presents results of evaluation of right foot in Fukyugata Ichi kata. Subplots have the same role as in Figure 3.



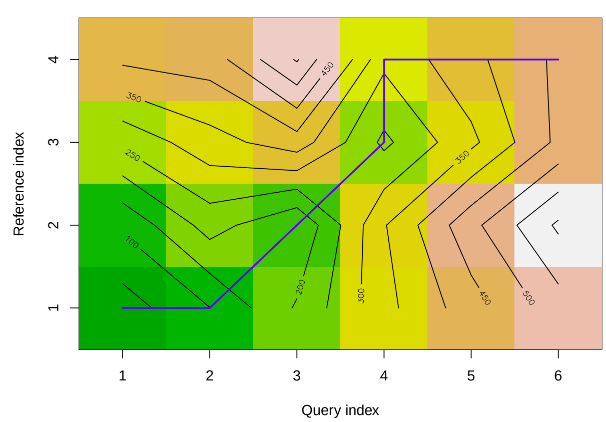
(a) Trajectories before aligning



(b) Trajectories after aligning

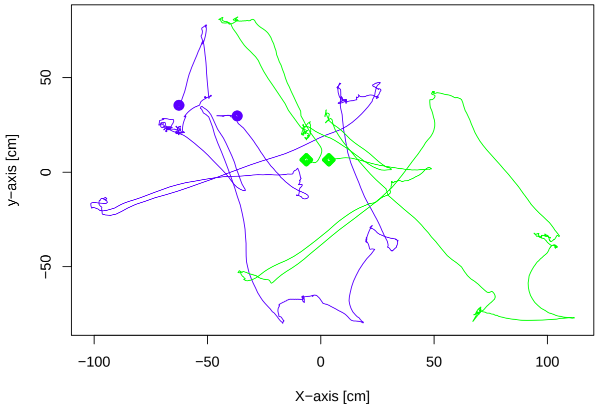


(c) Beginning and end positions of foot during motion

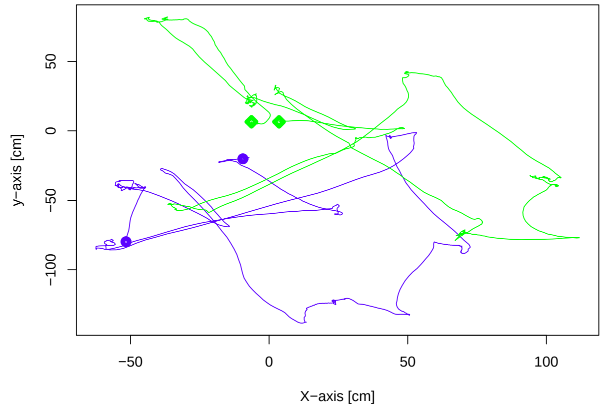


(d) DTW cost matrix

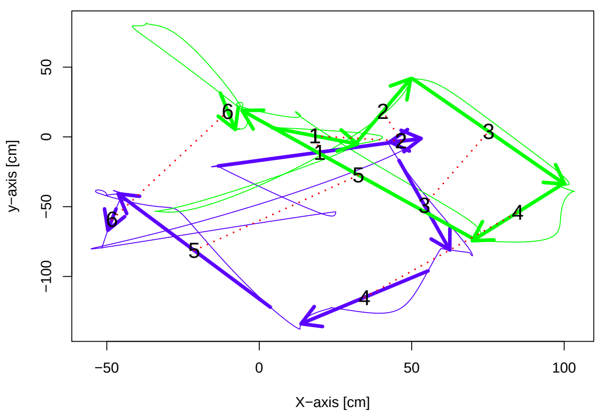
Figure 9: This figure presents results of evaluation of left foot in Fukyugata Ni kata. Subplots have the same role as in Figure 3.



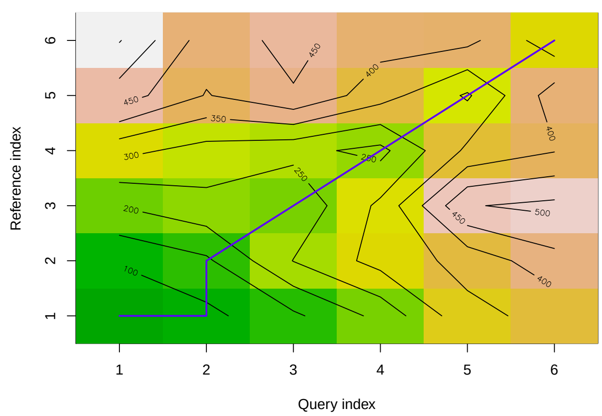
(a) Trajectories before aligning



(b) Trajectories after aligning



(c) Beginning and end positions of foot during motion



(d) DTW cost matrix

Figure 10: This figure presents results of evaluation of right foot in Fukyugata Ni kata. Subplots have the same role as in Figure 3.