Selection of pose configuration parameters of motion capture data based on Dynamic Time Warping

Adam Switonski[*,†], Henryk Josinski[†], Hafed Zghidi[†] and Konrad Wojciechowski[*]

[*]Polish-Japanese Institute of Information Technology, Aleja Legionow 2, 41-902 Bytom, Poland
[†]Silesian Univercity of Technology, ul Akademicka 16, 44-100 Gliwice, Poland

Abstract. The paper deals with the problem of motion data classification based on dynamic time warping transform. In the preliminary stage the selection of pose configuration parameters by hill climbing procedure is carried out. To construct dissimilarity matrices Euler angles and unit quaternion distance functions are taken into consideration and compared. To examine proposed approach gait database containing data of 30 different humans is used. The obtained results are satisfactory. The classification has over than 90% accuracy and joint selection allows to reduce dimensionality of pose configuration space noticeably.

Keywords: motion capture, dynamic time warping, attribute selection, gait identification, classification of motion data
PACS: 80

INTRODUCTION

Model based motion capture data contains sequences of poses represented by tree like structure of a kinematical chain. The pose configuration parameters are related to rotations performed by subsequent joints in respect to given reference pose. The rotations are specified in local coordinate systems. Thus in most obvious way dissimilarity between two poses P_1 and P_2 is determined by aggregated total distance of corresponding skeletal joints:

$$d(P_1,P_2) = \sum_{joint} d_{rotation}(P_1(joint),P_2(joint)) \tag{1}$$

Therefore the crucial challenge is the proper choice of distance function $d_{rotation}$ responsible for assessment of dissimilarity between two rotations. By default rotations are coded by three Euler angles. The data contains basic rotations performed around axes of local coordinate system. In such a case any classical distance functions of a vector space can be utilized, as for instance Euclidean or Manhattan metrics. However much more efficient and compact representation of rotations are given by unit quaternions. They are a natural extension of complex numbers with three-dimensional imaginary part:

$$q = a + i \cdot b + j \cdot c + k \cdot d \tag{2}$$

Quaternions are unit length which means they are located only on a hypersphere S^3. Thus to compare rotations, it is sufficient to calculate geodesic distance between related quaternions which is reflected by angle between vectors formed by their components. The scalar product $< q_1, q_2 >$ can be used to accomplish the task:

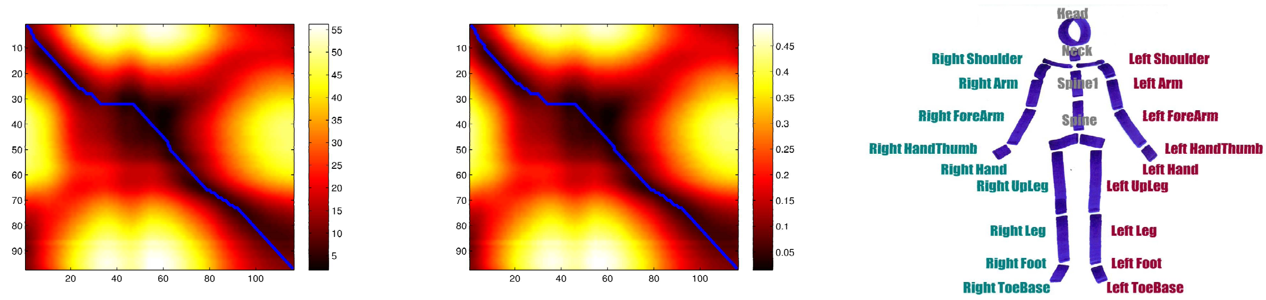
$$d_{geodesic}(q_1,q_2) = a\frac{1}{\pi} \cdot arccos(< q_1, q_2 >) \tag{3}$$

In other approach instead of raw angle its cosine can be determined which transforms geodesic distance in nonlinear way:

$$d_{cosine}(q_1,q_2) = \frac{1 - < q_1, q_2 >}{2} \tag{4}$$

What is more, some joints can have greater, other weaker and still others can have no impact on final pose comparison, thus equation (1) can be extended with weighted average.

$$d(P_1,P_2) = \sum_{joint} w_{joint} \cdot d_{rotation}(P_1(joint),P_2(joint)) \tag{5}$$



Hip Joint, Euclidean   Hip Joint, Quaternion Geodesic   Skeleton model

FIGURE 1. Similarity matrices with DTW paths and applied skeleton model

In case when weights w_{joint} are binary, which means they can take only 0,1 values, it leads to a selection of pose configuration parameters.

The main contribution of the paper relates to proposed method of determination of the above described weights w_{joint} in respect to specified classification problem. It is based on Dynamic Time Warping Transform and chosen heuristics search strategies. The method is applied to a gait human identification challenge problem and it allows to obtain joint subsets with most individual traits.

DYNAMIC TIME WARPING

Dynamic time warping (DTW) originally applied to spoken word recognition [1], but it was also successfully used in motion data analysis [2], [3], is a general technique to synchronize time series data. It speeds motions up and down to obtain most similar corresponding poses. As the result DTW allows to compare gait data efficiently because it takes into consideration possible local shifts between gait phases, which are removed by synchronization process.

DTW requires to calculate a dissimilarity matrix containing distances between every pair of poses of analyzed motions as presented in Fig. 1 The matrix is used to determine monotonic path with minimal aggregated total distance across underlying pose dissimilarity values, connecting edge points corresponding to start and end of compared time series. To perform the computations, dynamic programming can be applied. The path is monotonic, which means that moving backward in a time domain is not allowed. Thus the cost to reach the specified poses i and j of the first and second motions respectively can be determined on the basis of costs reaching to possible previous poses with indexes (i-1,j). (i,j-1), and (i-1, j-1). As the result it is sufficient to carry on the calculations in the proper order for subsequent rows or columns. The final cost of the path found corresponds to motions dissimilarity and the shape of the path to synchronization performed. The shape can be reconstructed on the basis of backtracing the nearest neighbors with minimal cost.

In Fig. 1 there are visualized example dissimilarity matrices and determined DTW paths for two randomly chosen gaits of different actors based on Euclidean metric applied to Euler angles and quaternion geodesic distance from equation (3), in respect to left hip and all joints of applied skeleton model from Fig. 1.

To classify motion data, the path cost with nearest neighbors classification scheme is mainly used [3, 4]. However, it is also possible to utilize the shape of the DTW path. More irregular path usually means greater number of shifts between gait phases which can be adequate to less similar motions.

DTW has O(n·m) computation complexity, where n and m correspond to the lengths of analyzed time series. To decrease it, fast DTW approaches approximating DTW transform may be applied. For instance in [1] and [5] some constraints on constructed dissimilarity matrix are used and in [6] multilevel method is proposed.

COLLECTED GAIT DATABASE

To validate proposed method motion capture gait database was collected. The data were acquired by Human Motion Laboratory of PJIIT `http://hml.pjwstk.edu.pl` equipped with Vicon software and hardware. It contains data of 436 gaits of 30 different males. It is database described in [7] and [8], extended by data of 5 additional humans.

TABLE 1. DTW classification based on complete set of joints in respect to different number k of nearest neighbors and rotation (Euclidean and Manhattan, geodesic (3) and cosine (4)) distance functions.

k	Euler angles		Quaternions	
	Euclidean	Manhattan	Geodeisc	Cosine
1	87,64	85,96	90,45	84,27
3	84,27	82,58	87,08	82,02
5	82,58	83,15	86,52	82,58

To detect main gait cycle containing two middle adjacent steps, tracking of distances between feet is carried out, as described for instance in [7].

The default skeleton model of Vicon Blade software was applied, as presented in Fig.1. It consists of 22 joints described by Euler angles triplets, global rotation and three-dimensional translation vector. In total it gives 72-dimensional pose space. However, the translation attributes are not taken into consideration by classification, they are removed in preprocessing stage.

To minimize the influence of the strict markers location and calibration process of the motion capture measurements on the final recognition, the acquisition is divided into sessions. Markers are attached on a human body and the calibration is carried out independently in every session.

ATTRIBUTE SELECTION

The pose configuration space contains 23 different rotations corresponding to skeleton segments and its global rotation, coded by Euler angles or unit quaternions. In considered selection stage the subsets of those rotations are discovered in respect to accuracy of subsequent classification. In general there are two major challenges of selection tasks. The first one is a search strategy. In most cases exhaustive search which takes into consideration all possible combinations of subsets is unworkable, because the problem is NP-complete. For motion data with 23 different rotations there are more than eight millions of different subsets.

In most basic search approach called ranking every parameter is examined separately, which allows to construct ranking of them. In a proper selection only top ranked attributes are taken into consideration. The main drawback of such an approach occurs in case of attributes which are discriminative only if considered with others.

One of the most often used search techniques in selection tasks is a greedy hill climbing procedure [4] with forward or backward directions. In such a case the search starts with empty or complete subsets and in successive iterations the best first attributes are added or removed depending on the direction. It is also possible to utilize any other heuristic search to explore attributes spaces as for instance genetic search [4] or extended invasive weed optimization [9].

The second challenge of selection is the way subsets are evaluated. In the most obvious case usually called wrapper approach [4], it is sufficient to perform classification experiment which takes into consideration only selected rotations. The overall assessment corresponds to obtained accuracy of classification.

RESULTS AND CONCLUSIONS

To perform classification collected gait database is split into two subsets of approximately equal size - training and testing ones, which contain the data of different sessions.

The obtained classification results in respect to considered complete set of joints with uniform weights as it is presented by equation (1) are shown in Table 1. Noticeably best accuracy is given by quaternion geodesic distance measure, Euclidean and Manhattan metrics are very similar and quaternion cosine is the worst one. The geodesic measure discriminates more stronger the distance between more similar rotations in comparing to cosine distance, which may enhance individual features. In all cases 1NN (nearest neighbor) classifier gives the greatest precision and it is chosen to further computations.

The results of subsequent stages containing data of joints added by hill climbing selection are shown in Table 2. The search is carried out only on the basis of the training data, which is split once again into two equal parts considered to be training and testing sets of the hill climbing validation. The column CCR1 just gives information about recognition accuracy in respect to hill climbing validation and CCR2 to the proper testing set.

TABLE 2. Forward hill climbing results

Segments	Euler angles Euclidean			Quaternion Geodesic		
	Segment	CCR1	CCR2	Segment	CCR1	CCR2
1	LeftUpLeg	64,44	69,66	RightUpLeg	65,56	65,17
2	RightUpLeg	77,78	74,72	RightLeg	77,78	74,16
3	RightForeArm	83,33	75,84	LeftUpLeg	80,00	78,09
4	LeftLeg	83,33	81,46	RightForeArm	82,22	80,90
5	Spine	84,44	83,15	LeftShoulder	85,56	75,84
6	LeftShoulder	86,67	84,83	Spine1	87,78	83,71
7	Spine1	90,00	87,08	root	90,00	83,71
8	RightHandThumb	90,00	87,08	LeftFoot	91,11	90,45
9	LeftHandThumb	90,00	87,08	Spine	93,33	90,45
10	RightHand	90,00	80,34	RightHandThumb	93,33	90,45
11	RightLeg	92,22	80,90	LeftHandThumb	93,33	90,45
12	RightToeBase	90,00	80,90	LeftToeBase	92,22	89,33
13	RightFoot	91,11	80,34	LeftLeg	94,44	89,89
14	LeftToeBase	90,00	83,71	RightFoot	92,22	89,89
15	LeftForeArm	91,11	80,90	RightArm	94,44	88,76
16	RightArm	87,78	79,21	RightShoulder	94,44	88,20
17	LeftFoot	91,11	83,71	RightToeBase	93,33	87,08
18	Head	88,89	84,27	LeftForeArm	94,44	88,76
19	root	88,89	86,52	Head	91,11	88,76
20	RightShoulder	87,78	88,20	RightHand	92,22	88,76
21	LeftArm	82,22	87,64	LeftHand	87,78	89,33
22	LeftHand	85,56	88,20	LeftArm	90,00	91,01

The joint selection allows to reduce considered pose configuration parameters without loss of the precision of classification. In case of default stop criteria of the hill climbing procedure which terminates calculations after first non improving node and quaternion geodesic distance function, pose configuration space is diminished to nine joints. For the Euclidean metric seven joints are selected which implies only 0.5% worse recognition accuracy.

The results differ strongly depending on the training and testing sets, as represented by CCR1 and CCR2 columns. The Pearson correlation coefficients have values of 88% and 50% for quaternion geodesic and Euclidean distance function respectively. It is caused by an overfitting phenomena of the selection process. The performance is adjusted to the hill climbing validation data and it becomes much worse in case of unseen data. This is a reason why the results of of CCR1 are about 3% better in comparing to CCR2 on average. What is more there are some combinations of joints which improve the classification noticeably for a hill climbing validation. In such a case the recognition rate is even raised to 94.44%.

ACKNOWLEDGMENTS

The project was supported by The Polish National Science Centre on the basis of decision number DEC-2011/01/B/ST6/06988.

REFERENCES

1. H. Sakoe, and S. Chuba, *IEEE Transactions on Acoustics, Speech and Signal Processing* **8**, 43–49 (1978).
2. M. Kulbacki, and M. Bak, *Advances in Soft Computing* **17**, 217–226 (2002).
3. A. Switonski, A. Polanski, and K. Wojciechowski, "Human identification based on the reduced kinematic data of the gait," in *7th IEEE International Symposium on Image and Signal Processing and Analysis*, Dubrovnik, 2011.
4. I. Witten, E. Frank, and M. Hall, *Practical Machine Learning Tools and Techniques*, Morgan Kaufmann, Boston, Massachusetts, 2011, ISBN 978-0-12-374856-0.
5. F. Itakura, *IEEE Transactions on Acoustics, Speech and Signal Processing* **ASSP-23**, 52–72 (1975).
6. S. Salvador, and P. Chan, *Intelligent Data Analysis* **11**, 561–580 (2007).
7. A. Switonski, H. Josinski, A. Michalczuk, P. Pruszowski, and K. Wojciechowski, *Lecture Notes in Computer Science* **8397**, 535–544 (2014).
8. A. Switonski, A. Polanski, and K. Wojciechowski, *Lecture Notes in Computer Science* **6915**, 531–542 (2011).
9. H. Josinski, D. Kostrzewa, A. Michalczuk, and A. Switonski, *The Scientific World Journal* **2014** (2014), article ID 831691.