Development of Operation Estimation Method Based on Tracking Records Captured by Kinect

Bin Wang[1,2], Fuchun Sun[1,2], Huaping Liu[1,2(✉)], Xuan Guo[1,2],
Sota Yoshii[3], Naoyuki Fujiwara[3], Weihang Wu[3], and Guangyu Zhao[4]

[1] Department of Computer Science and Technology,
Tsinghua University, Beijing, China
hpliu@tsinghua.edu.cn
[2] State Key Laboratory of Intelligent Technology and Systems,
Tsinghua University, Beijing, China
[3] Yokohama R&D Center, Mitsubishi Heavy Industries, Tokyo, Japan
[4] Little Wheel Robot Co., Beijing, China

Abstract. In order to evaluate workers' progress in the factory, we developed a model of operation estimation method based on tracking records captured by Kinect. In this paper, we use Kinect sensors to capture the human motion process (such as broadcast gymnastics), extract the skeleton frame sequence from training data as a template, then improve the DTW algorithm to match the skeleton frame in the test data and in the training data, to estimate the percentage of the action's completion. The improved DTW algorithm achieves state-of-the-art performances on our dataset, higher than 90%.

Keywords: Working progress completion estimation · DTW · Kinect

1 Introduction

In order to enhance competitiveness and improve work efficiency, the factory's manager need to get the worker's progress in real-time. In the traditional method, the wearable device (such as a microphone, PDA and other devices) were used, they were cumbersome and ineffective compared to Kinect, which can capture the human skeleton, the skeleton data is more easily identified for body gesture recognition than RGB data, and is not prone to be interfered by the back ground (Fig. 1).

In this paper, we used Kinect to record some person's whole working progress as xef file, extracted body skeleton as feature from it, labeled one of them as training template, and used other person's data as testing data (Fig. 2).

In the Sect. 2, the improved Dynamic Time Warping algorithm (DTW) was explained, In the Sect. 3, we will show how DTW could be employed to identify similar to query subsequences in the long data streams. In the Sect. 3, the evaluation result was shown. In the Sect. 4, we summarized.
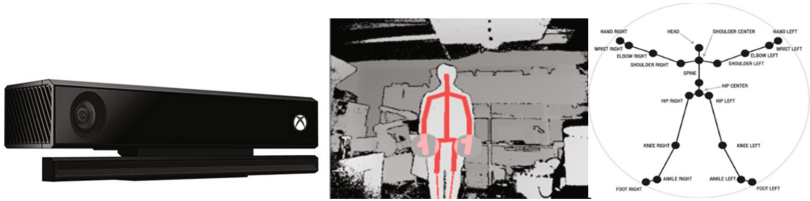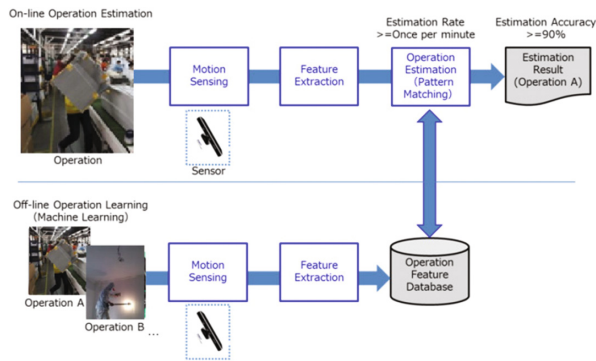
Fig. 1. Kinect v2 and skeleton frame

Fig. 2. System architecture

2   Development of Operation Estimation Method

2.1   Introduce of Dynamic Time Warping Algorithm (DTW)

The Dynamic Time Warping algorithm (DTW) is a well-known algorithm in many areas: handwriting and online signature matching [1, 2], sign language recognition [3] and gestures recognition [3, 4], data mining and time series clustering (time series databases search) [5–10], computer vision and computer animation [11], surveillance [12], protein sequence alignment and chemical engineering [13], music and signal processing [11, 14, 15].

DTW algorithm has earned its popularity by being extremely efficient as the time-series similarity measure which minimizes the effects of shifting and distortion in time by allowing "elastic" transformation of time series in order to detect similar shapes with different phases. Given two time series $X = (x_1, x_2, \ldots x_N), N \in N$ and $Y = (y_1, y_2, \ldots y_M), M \in N$ represented by the sequences of values (or curves represented by the sequences of vertices) DTW yields optimal solution in the O (MN) time which could be improved further through different techniques such as multi-scaling [14, 15]. The only restriction placed on the data sequences is that they should be sampled at equidistant points in time (this problem can be resolved by re-sampling). If sequences are taking values from some feature space Φ than in order to compare two different
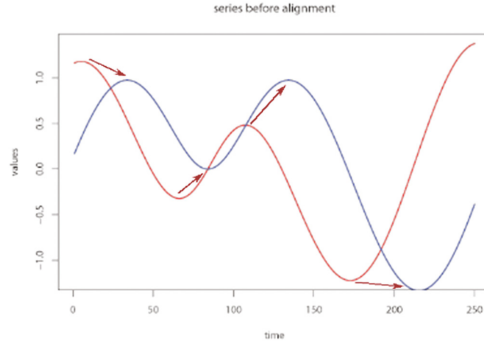
Fig. 3. Raw time series, arrows show the desirable points of alignment

sequences $X, Y \in \Phi$ one needs to use the local distance measure which is defined to be a function:

$$d : \Phi \times \Phi \rightarrow R \geq 0 \qquad (1)$$

Intuitively d has a small value when sequences are similar and large value if they are different. Since the Dynamic Programming algorithm lies in the core of DTW it is common to call this distance function the "cost function" and the task of optimal alignment of the sequences becoming the task of arranging all sequence points by minimizing the cost function (or distance). Algorithm starts by building the distance matrix $C \in R^{N \times M}$ representing all pairwise distances between X and Y (Fig. 4). This distance matrix called the
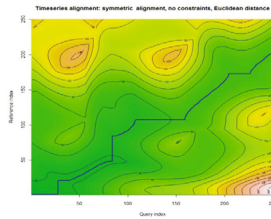

Fig. 4. Time series alignment, cost matrix heat map.

local cost matrix for the alignment of two sequences X and Y:

$$C_1 \in R^{N \times M} : c_{i,j} = k x_i - y_i k, \, i \in [1 : N], \, j \in [1 : M]$$

Once the local cost matrix built, the algorithm finds the alignment path which runs through the low-cost areas - "valleys" on the cost matrix, Fig. 5. This alignment path (or warping path, or warping function) defines the correspondence of an element $x_i \in X$ to $y_i \in Y$ following the boundary condition which assigned first and last elements of X
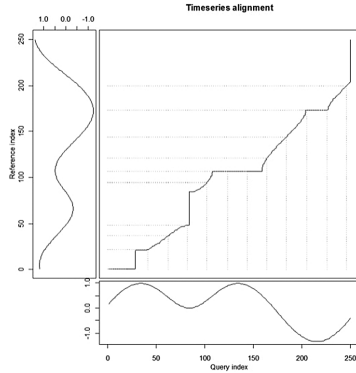
Fig. 5. The optimal warping path aligning time series from the Fig. 3.

and Y to each other, Fig. 6. Formally speaking, the alignment path built by DTW is a sequence of points $p = (p_1, p_2, \ldots, p_k)$ with $pl = (pi, pj) \in [1 : N] \times [1 : M]$ for $I \in [1 : k]$.

2.2   The Improved DTW

We optimized the DTW algorithm according to the order of the human motion. In the whole action routines, every section is in a certain sequence, will not be disorder. For example, the Sect. 3 can only happen before the Sect. 4, and not happen before the Sect. 2. Therefore, according to the sequence of section, we can fix the error results, the evaluation accuracy were improved more than 10%.
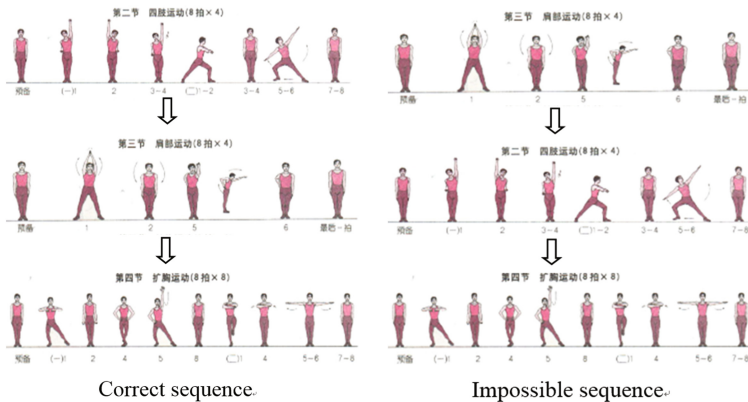

Correct sequence   Impossible sequence

Fig. 6. Correct sequence and impossible sequence

3   Experiment Conclusion

3.1   Data and Evaluation Formula

6 persons (Fig. 7) did Chinese gymnastic (Fig. 8) as data set. Each section has n activities, which are labeled as 1, 2 …, n.

Fig. 7. Collecting data

Fig. 8. Chinese gymnastic

The evaluation formula:

$$D(m) = |R(m) - C(m)|$$

$$A = 1 - \frac{\sum_{m=1}^{M} D(m)}{M}$$

(a)   m: Number of frames
(b)   C(m): the estimated completion of work
(c)   R(m): the practical completion of work
(d)   A: Accuracy

3.2 Results

The average accuracy of DTW is 86% (Fig. 9). The blue line is ground truth and the red line is test result. X axis: Length of the test job, Y axis: Activities and the complete degree.

Fig. 9. The DTW results (Color figure online)

The average accuracy of improved DTW is 97.63% (Fig. 10).

Fig. 10. The improved DTW results

4   Conclusion and Future Work

In this paper, we improved DTW algorithm to estimate the whole working progress completion, to use gymnastics data of 6 person's to cross test, the average accuracy is more than 90%. The result show when to use one person's data as template, himself data as test data, the result is almost 99%. If to use other person data as test data, or when the skeleton was lost by Kinect, the result is not good enough. In the future research we will resolve above problems.

References

1. Efrat, A., Fan, Q., Venkatasubramanian, S.: Curve matching, time warping, and light fields: new algorithms for computing similarity between curves. J. Math. Imaging Vis. **27**(3), 203–216 (2007). http://dx.doi.org/10.1007/s10851-006-0647-0

2. Tappert, C.C., Suen, C.Y., Wakahara, T.: The state of the art in online handwriting recognition. IEEE Trans. Pattern Anal. Mach. Intell. **12**(8), 787–808 (1990). http://dx.doi.org/10.1109/34.57669

3. Kuzmanic, A., Zanchi, V.: Hand shape classification using DTW and LCSS as similarity measures for vision-based gesture recognition system. In: EUROCON, 2007. The International Conference on Computer as a Tool 2007, pp. 264–269. http://dx.doi.org/10.1109/EURCON.2007.4400350

4. Corradini, A.: Dynamic time warping for off-line recognition of a small gesture vocabulary. In: RATFG-RTS 2001: Proceedings of the IEEE ICCV Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems (RATFG-RTS'01). IEEE Computer Society, Washington, D.C. (2001). http://portal.acm.org/citation.cfm?id=882476.883586

5. Niennattrakul, V., Ratanamahatana, C.A.: On clustering multimedia time series data using k-means and dynamic time warping. In: International Conference on Multimedia and Ubiquitous Engineering, MUE 2007, pp. 733–738 (2007). http://dx.doi.org/10.1109/MUE.2007.165

6. Gu, J., Jin, X.: A simple approximation for dynamic time warping search in large time series database, pp. 841–848 (2006). http://dx.doi.org/10.1007/11875581101

7. Bahlmann, C., Burkhardt, H.: The writer independent online handwriting recognition system frog on hand and cluster generative statistical dynamic time warping. IEEE Trans. Pattern Anal. Mach. Intell. **26**(3), 299–310 (2004). http://dx.doi.org/10.1109/TPAMI.2004.1262308

8. Kahveci, T., Singh, A.: Variable length queries for time series data. In: 17th International Conference on Data Engineering, 2001. Proceedings, pp. 273–282 (2001). http://dx.doi.org/10.1109/ICDE.2001.914838

9. Kahveci, T., Singh, A., Gurel, A.: Similarity searching for multiattribute sequences. In: 14th International Conference on Scientific and Statistical Database Management, 2002. Proceedings, pp. 175–184 (2002). http://dx.doi.org/10.1109/SSDM.2002.1029718

10. Euachongprasit, W., Ratanamahatana, C.: Efficient multimedia time series data retrieval under uniform scaling and normalisation, pp. 506–513 (2008). http://dx.doi.org/10.1007/978-3-540-78646-749

11. Dtw-based motion comparison and retrieval, pp. 211–226 (2007). http://dx.doi.org/10.1007/978-3-540-74048-310

12. Zhang, Z., Huang, K., Tan, T.: Comparison of similarity measures for trajectory clustering in outdoor surveillance scenes. In: ICPR 2006: Proceedings of the 18th International Conference on Pattern Recognition (ICPR 2006), pp. 1135–1138. IEEE Computer Society, Washington, D.C. (2006). http://dx.doi.org/10.1109/ICPR.2006.392

13. Vial, J., Nocairi, H., Sassiat, P., Mallipatu, S., Cognon, G., Thiebaut, D., Teillet, B., Rutledge, D.: Combination of dynamic time warping and multivariate analysis for the comparison of comprehensive two-dimensional gas chromatograms application to plant extracts. J. Chromatogr. A (2008). http://dx.doi.org/10.1016/j.chroma.2008.09.027

14. Muller, M., Mattes, H., Kurth, F.: An efficient multiscale approach to audio synchronization, pp. 192–197 (2006)

15. Dynamic time warping, pp. 69–84 (2007). http://dx.doi.org/10.1007/978-3-540-74048-34