RESEARCH ARTICLE OPEN ACCESS

Human Action Recognition Based on the Adaptive Weighted Dynamic Time Warping Algorithm

ZhaoLiang Li[1],QingFeng Zhang[2],WenTian Huang[3]
[1,2,3](Department of Computer Science,Jinan University, Guangzhou, China)

--************************----------------------------------

Abstract:

The skeleton information of human action could be extracted by Kinect sensor, and it was also a hot topic to identify the action based on them. So, an action recognition method based on adaptive weighted dynamic time warping algorithm was proposed. In the method, firstly, the skeleton joints' coordinates were obtained from the action by Kinect, and then they were normalized. Secondly, the weights were assigned to joint based on inter-intra class joints' variability and were introduced into the distance computation in the adaptive weighted dynamic time warping algorithm. Finally, voting method was used to recognize the action. Experimental results show that the presented method has a high performance for action recognition in the MSR Action3D database and it is efficient.

Keywords **—Action Recognition,DTW, Kinect, Improved Otsu, Voting algorithm.**

--************************----------------------------------

I. INTRODUCTION

In the pass decades, Human action recognition has become a hot topic in the area of computer vision, mainly related to its huge potential applications value. However, certain complexity of data extraction and analyses of human action make the academic community face many difficulties. The main reasons are as follow.

First, the human body itself is non-rigid and has many degrees of freedom and people have various movements. Second, each person has their own body and attitude characteristics, which have increased the difficulty of movement recognition. Third, some uncertainties (such as: viewing angle, shelter, wearing, etc.) also make the problem more complex.

The method based on Hidden Markov Models (HMM) and Support Vector Machine (SVM) are the common methods to human motion recognition. Zhang Yi et[1] and Yan Yan et[2] put forward their own gesture recognition method on the basis of hidden Markov model respectively. Their methods have high recognition rates to identify human gesture orbit in a certain degree, but we must determine the number of states of the action in advance when we using motion recognition method based on hidden Markov model. However, due to the certain complexity ofhuman action, it is difficult to determine the optimal number of states of each action.Thi-Lan Le et[3] use Kinect collect data, and use the method based on SVM to processthe data. Although their strategy has high recognition rate, it can only identify the relatively simple movement,lying, seat, stand and bend. C.Schuldt et[4]apply SVM to human behaviour recognition by extracting local spatial and temporal characteristics of the video frame. However, the major drawback of using SVM is its high computational complexity. In addition, there are some other methods. F. Ofli et[5] propose a very intuitive method—joint sequence information maximization method, which can be explained qualitatively action feature representation.By carrying out experimental tests in the cross database, the results show that the method has a high recognition rate. But the method still has some limitations when identifying different action planes. Wang Xin et[6] use manifold learning method for dimensionality reduction to train data and then combine the nearest neighbour difference and the improved Hausdorff distance for action recognition. Finally, they get good recognition results. Lin Shui Qiang et[7]] use Limb articulation regular expressions to represent the trajectory. On the basis of this, they establish finite state machine of posture sequence and achieve action identification. Although the accuracy of the method is better, the method should reconsider the rejection and compatibility between the action definition when adding a new set of actions. So the robustness is not very good.

Additionally, because the Dynamic Time Warping (DTW) algorithm[8] has a large comprehensive advantage [9] in terms of identifying action based on time series. It attractsmany attention of the scholars[10-14]. Liu Fei et[10] calculate the similarity of joint angle between the different action sequences to identify action by using improved DTW algorithm, which alleviates the pathological problems of the conventional DTW algorithm and has a better recognition rate. Li Haitao[11]establishes a set of DTW gesture recognition methods based on joint weighted from the joint of the action gesture image starting. The results show that the method has some validities. M.Reyes et[12] construct the feature vectors by the 3D coordinates of human joints, then they use the DTW algorithm to process feature vectors. Although this method improves the recognition rate to some extent by improving the DTW distance function, the robustness of this method is still low when facing the different people height. However, S.Sempena et[13] construct feature vectors through the joint direction which overcomes the problem of people height, but this method also has a very high complexity and is not suitable for using in real-time system.

J.Wang et[14] extract self-similarity matrix by comparison of the differences between video frames and then carry out the experiments by DTW algorithm and k nearest neighbour algorithm. Although this method has some efficacies, it needs to compare and calculate all the video frames data when extracting features. Thus it has a very higher complexity and is very difficult to reduce the identification time to real-time recognition.

In recent years, with the emergence of some lower cost depth sensors ,such as, Microsoft Kinect, which make us get the 3D human skeleton more easier.And in the behaviour-based skeleton representation, many scholars havestudied [3,12,13,15].

Human skeleton model is mainly composed of joints and joint body.Joints connect all the joint bodies and all the moving of joint body constitute the body's movement behaviour. However, different joint corresponding spatial movement trajectory is different in the same action. So different joints have different effects in a particular action. In an action which only involves arm movement, for example, leg joints are substantially unaffected to this action. In turn, in an action which only involves leg movement, the joints of arm are substantially unaffected. Based on this characteristic, we propose an adaptive weighted DTW algorithm. In this method, we collect skeleton information of human action sequences (3D coordinates of the joint point) through Microsoft Kinect and then construct feature vectors by moving tracks of joints. Moreover, we calculate each joint weights by using the improved Otsu method and then sum all the DTW distancesof joints to get the adaptive weighted DTW distance. Finally, voting algorithm is used to identify the type of action.

The rest of the paper is organized as follows. Section 2 focuses on the data collection and normalization by MS Kinect. Section 3 is a brief introduction of the traditional DTW algorithm. Subsequently, we describe the adaptive weighted dynamic time warping algorithm in section 4. Simultaneously, voting algorithm is briefly described and then the full action recognition procedure are summarized at the end of this section. After that, we conduct a large number of experimental tests and analyses. In the last section, we draw the conclusions of this paper and future work is prospected.

II. DATA COLLECTION AND NORMALIZATION

A. *Joint Data Collection*

We can use Kinect sensor and MATLAB to get 20 individual joint coordinates of action sequence, showing in figure 1a.

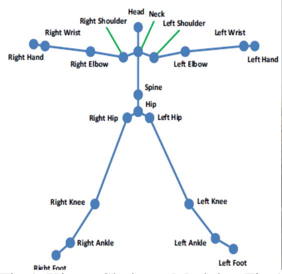


Fig 1aKinect Skeleton Model

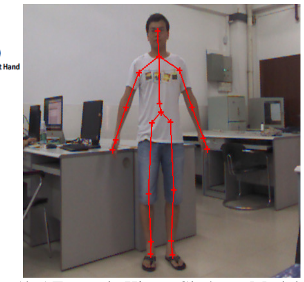


Fig 1b AExample Kinect Skeleton Model

Figure 1

But, MATLAB need to be configured slightly [16] before data collection. Then, we collect a set of joint coordinates, as the figure 1b showing.

B. Data Normalization

After data collection, we can get the initial skeleton information of the human action. However, we need to do some pre-processing for the initial data in the actual Kinect motion recognition process. The main reasons are as follow. First, the height is often different for different people, then the joint coordinatesfor the certain posture are also different. Second, for the same action and same person, if we have the different relative distances to Kinect, then the coordinates for the given posture have also some difference, mainly referring to the Z-axis coordinate here.

Therefore, we use the following method for data pre-processing.

Step 1Translation

We use equation 1 to process the initial coordinatescollected in Section *II-A*,there $(x, y, z)_{hip}$ standing for the hipcoordinate, showing in figure 1a.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} x_{hip} \\ y_{hip} \\ z_{hip} \end{pmatrix} \quad (1)$$

Step 2Scale

Because the distance between left hip and right hip is proportional to the person height in general, so we can use equation 2 to process the result of above step 1.

$$\begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} = k \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \quad (2)$$

there,k is equal to the Euclidean distance between left hip and right hip. That is

$$k = \sqrt{\left(x_{hip\_left} - x_{hip\_right}\right)^2 + \left(y_{hip\_left} - y_{hip\_right}\right)^2 + \left(z_{hip\_left} - z_{hip\_right}\right)^2}$$

Suppose $P_i = \left(x_i^{"}, y_i^{"}, z_i^{"}\right)$ $(i = 1, 2, ..., 20)$represents the result of after scale. Then, if we process a set of continuous action sequences(F frames) by the above method, we can get $R_i = \{P_{if}\}$ ($i = 1, 2, ..., 20$, $f = 1, 2, ..., F$). P_{if} is the preprocessed coordinate of the i th joint in the f th frame. R_i is a continuous sequence of the i th joint in F frames.

III. DYNAMIC TIME WARPING ALGORITHM

Dynamic Time Warping algorithm is often used to find the optimal matching of two sequences of different time series through dynamic programming algorithm and then calculates the distance between two time series. This method can solve the problem of unequal length time series. So in addition to be used for motion recognition, DTW can also be widely used in speech recognition[17,18], handwriting recognition[19,20,21], whose main principles are as follow.

For one joint and its two different time series $R = \{r_i\}$, $T = \{t_j\}$ ($i = 1, 2, ..., L_1$, $j = 1, 2, ..., L_2$), r_i and t_j represent joint coordinates of the corresponding point in time respectively. Then, using cumulative distance equation (3)

$$D(i, j) = min \begin{Bmatrix} D(i, j-1) \\ D(i-, j-1) \\ D(i-1, j) \end{Bmatrix} + d(r_i, t_j) \quad (3)$$

there, $d(r_i, t_j)$ is the distance of r_i and t_j , Euclidean distance is often be used. That is：

$$d(r_i, t_j) = \sqrt{\left(r_{ix} - t_{jx}\right)^2 + \left(r_{iy} - t_{jy}\right)^2 + \left(r_{iz} - t_{jz}\right)^2} ,$$

After that, we can calculate the distance of R and T — $D(L_1, L_2)$ and optimal matching $Path = (Path_R, Path_T)$, there

$$Path_R = \left\{path_R^k\right\}\left(1 \le path_R^k \le L_1\right), \quad k = 1, 2, ..., K ;$$

$Path_T = \{path_T^k\}(1 \le path_T^k \le L_2)$, $k = 1,2,...,K$ 。

Obviously, $D(L_1, L_2)$ is more smaller, R and T is more similar .Additionally, in the calculation of optimal matching— $Path$, the following constraints must be satisfied.

a) Boundary condition：

$(path_R^1, path_T^1) = (1,1)$ and $(path_R^K, path_T^K) = (L_1, L_2)$

b) Monotonicity condition：

$path_R^k \le path_R^{k+1}$ and $path_T^k \le path_T^{k+1}$

c) Continuity condition：

$path_R^{k+1} - path_R^k \le 1$ and $path_T^{k+1} - path_T^k \le 1$

IV. ADAPTIVE WEIGHTED DTW AND VOTING ALGORITHM

In this section, we introduce this paper's core algorithm — adaptive weighted dynamic time warping algorithm firstly. Then providing a brief description of voting algorithm. At last, the main procedure of the action identified issummarized.

A. *Adaptive Weighted Dynamic Time Warping Algorithm*

For a set of test actions and a set of sample movements, we make use of the DTW algorithm described in section *III* to get 20DTW distances D_i ($i = 1,2,...,20$)of joints whosemain reason is that there are 20 joint coordinates for each frame image and each joint coordinate corresponds to a time sequence in a set of action sequences. But how to further calculate the similarity of the test action and sample action by D_i ? We most likely to think is that taking the average of these 20 D_i ,that is:

$$\varphi_0 = \frac{1}{N}\sum_{i=1}^{N} D_i \, (N = 20) \quad (4)$$

here, we call it traditional DTW algorithm.

However, we find that, by careful observation, the amplitude of motion (total movement distance) of different joint is different for different action sequence and has the different contribution rate to the end result. E.g, when our left arm wave, there are only left shoulder, left elbow, left wrist, left hand involved in motion. While the remaining 16 joints almost remain unchanged. In addition, when we take right leg to the right, only the four joints on the right leg involved in motion, the remaining joints don't substantially move. So like this, if we just take the average of D_i by formula (4) and then it is bound to cause a lot of errors. Thus, M.Reyeset et[12] assign different weights to everyjoints and improve the recognition rateto some extent. In this paper, however, we propose a new adaptive weighted method, Adaptive Weighted Dynamic Time Warping or AWDTW, according to the characteristic having different contribution to the final for different joints in the different sequence of actions. The discussion in detail about this algorithm is following.

For a continuous action sequence $R = \{R_i\}$,there, $R_i = \{P_{if}\}$ ($i = 1,2,...,20$, $f = 1,2,...,F$). F is the number of all frames. Here, we suppose ds_i represents the total distance of ith joint in the all F frames action sequence. Then we can calculate ds_i by formula (5).

$$ds_i = \sum_{f=1}^{F-1} d\left(P_{i(f+1)}, P_{if}\right) \quad (5)$$

There, $i = 1,2,...,20$, $d\left(P_{i(f+1)}, P_{if}\right)$ is the Euclidean distance of $P_{i(f+1)}$ and P_{if} .

In order to extract the main joints involved in the movement,when ds_i is less than the threshold ds^* ,we think it as zero. So, the final $ds_i^{'}$ can be determined with equation (6).

$$ds_i^{'} = \begin{cases} ds_i, ds_i \ge ds^* \\ 0, ds_i < ds^* \end{cases} \quad (6)$$

Then the ith joint weight can be calculate by equation (7).

$$w_i = \frac{1 - e^{-ds_i^{'}}}{\sum_{i=1}^{N}\left(1 - e^{-ds_i^{'}}\right)} (N = 20) \quad (7)$$

Finally, we can get the distance of adaptive weighted DTW,showing at formula (8).

$$\varphi = \sum_{i=1}^{N} w_i D_i (\mathrm{N} = 20) \quad (8)$$

Before calculating the φ ,however, we must determine the optimal ds^* in equation (6). This ds^* must make the joints not related to action to reach minimum contribution for the final result.

Conversely, reach maximum for joints closely related to action. In other words, separating the joints associated with the action and not related to the movement. Thus, we take advantage of the improved OTSU[22]. In this method, to determine the thresholdby minimize the ratio between within-class variance and between-class variance, as the equation (9) showing.

$$\lambda(ds)=\frac{\sigma_w^2(ds)}{\sigma_b^2(ds)} \quad (9)$$

$$ds^*=Arg \underset{0<ds<ds_{max}}{Min} \lambda(ds) \quad (10)$$

There, $\sigma_w^2(ds)$ is the within-class variance of ds_i ($i=1,2,...,20$). $\sigma_b^2(ds)$ is the between-class variance of ds_i ($i=1,2,...,20$). We can effectively extractjoints with the main contributionfor final result by this method. For instance, for an action sequence(A_0) only involved in left arm,its twenty ds_i is showing in Table 1.

Table 1 Move Distance of 20 joints in action A_0 (Uint:m)

Joint	Distance	Joint	Distance
Hip	0	R-Wrist	0.4165
Spine	0.0688	R-Hand	0.3741
Neck	0.2758	L-Hip	0.0694
Head	0.4228	L-Knee	0.3199
L-Shoulder	0.3264	L-Ankle	0.469
L-Elbow	2.9712	L-Foot	0.6745
L-Wrist	5.2199	R-Hip	0.0539
L-Hand	5.6884	R-Knee	0.2878
R- Shoulder	0.2628	R-Ankle	0.5645
R-Elbow	0.2903	R-Foot	0.5157

L-Left, R-Right

We can get the optimal threshold(1.8229)for the data of table 1 by improved OTSU. So, The data in table 1 is divided into two class, $X=\{$ left elbow, left wrist, left hand $\}$ and $Y=\{$others$\}$. As shown in figure 2 about details.

Figure 2 Classification result of Table 1

From figure 2 we can see that the main joints(left elbow, left wrist, left hand) involved in the movement A_0 are extracted out very well.

B. Voting algorithm

For a set of test action sequences, T ,we can calculate the AWDTW distance of T and each sample action sequence using Adaptive Weighted Dynamic Time Warping algorithm showing in section *IV-A*. Then getting a series of distance $\varphi_i(label)$ ($i=1,2,..,n$, $label=1,2,...,LABEL$). n is the number of sample data, $LABEL$ is the number of sample class and $label$ in the $\varphi_i(label)$ is the class number of ith sample class. Moreover, we sort $\varphi_i(label)$ in ascending order. After that, picking up the first p smallest value and counting the number of each label. Suppose the $label$ with maximum number of statistics is $label^*$ and its number of statistics is q .Then,

a) If $q \geq p/2$, identify successful, test action is belong to class $label^*$

b) If $q < p/2$, identify failure, test action is unknown.

C. Algorithm Process

For a set of test action sequences, the main action recognition steps are as follow.

Step 1: Collecting the joint coordinates of human action by Kinect sensor;

Step 2: Standardizing the data of step 1 using the method described section *II*;

Step 3: Calculating n AWDTW distance φ_i ($i = 1, 2, \ldots, n$) using the data after standardization

Step 4: Identifying test action using Voting algorithm.

Step 5: Output the result of step 4.

V. EXPERIMENT AND ANALYSIS

We use Windows 7, 64bit operating system、Core I3,cup 2.2Hz processor,4G RAM,KinectSDK 1.8,MATLAB R2015a and VC++ 2005. On below, we take two experiments on ourselves database and MSR Action3Ddatabase.

A. Gesture recognition

When only two upper arm involved in movement, action recognition degenerates to gesture recognition. Thus, we mainly test the performance of our method in gesture recognition in this section.

Firstly, we find 7 persons randomly, everyone only do 8 kinds of actions and repeat 4 timesrespectively, as the table 2 showing. Then, selecting 20 groups as test data for each action randomly and others as sample data.

Table 2 Experiment data about 8 kinds of Gesture

Gesture	Number		Gesture	Number	
	Test	Sample		Test	Sample
LeftHandPullDown	20	8	RightHandPullDown	20	8
LeftHandPushUp	20	8	RightHandPushUp	20	8
LeftHandSwipeRight	20	8	RightHandSwipeLeft	20	8
LeftHandWave	20	8	RightHandWave	20	8

Now, there are 160 groups in test data and 64 groups in sample data. We select five groups from test data to experiment every time and repeat 10 times. The average recognition time is 0.3572s. In addition, we also compare our method with traditional DTW and the algorithm of M.Reyes[12], the detailed resultsare in table 3

Table 3 Performance of traditional DTW, M.Reyes`s and our method on ourselves database

Algorithm	Recognition rate		
	MAX	MIN	Avg
Traditional DTW	86.67%	66.67%	75%
M.Reyes`s	93.33%	73.33%	85.71%
Our method	100.00%	93.33%	96.67%

From table 3 we can see that our method's average recognition rate is superior to traditional DTW`s and M.Reyes`s. In some case.our recognition rate can even reach to 100%.

B. Action recognition

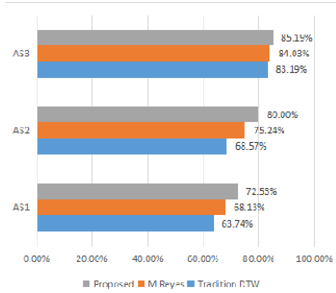
In this part, we have done a lot of experiments on database MSR Action3D[23] and compare our method with related algorithms.In MSR Action3D data sets, there are 20 kinds, 567 groups data in total and each kind of data repeat 2 or 3 times in different surroundings.

First of all, we divide MSR Action3D database into 3 kinds of data sets[24],AS1、AS2、AS3,as showing in table 4.

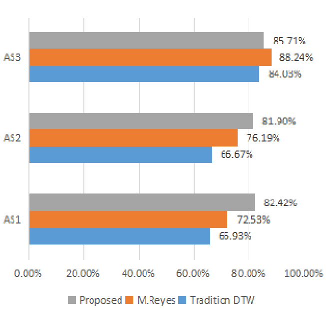
Table 4 Data set AS1, AS2, AS3

AS1	AS2	AS3
Horizontal arm wave	High arm wave	High throw
Hammer	Hand catch	Forward kick
Forward punch	Draw x	Side kick
High throw	Draw tick	Jogging
Hand clap	Draw circle	Tennis swing
Bend	Two hand wave	Tennis serve
Tennis serve	Forward kick	Golf swing
Pickup & throw	Side boxing	Pickup & throw

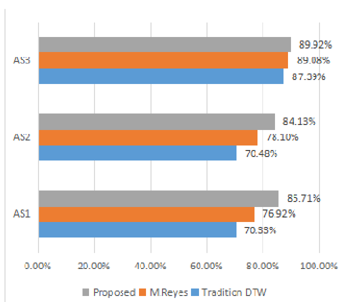
Then, we conduct lots of experiments on data set AS1,AS2,AS3 respectively. But, we must divide AS1,AS2,AS3 into test data and sample data before starting experiment. So, we take one third of data set AS1 as sample data, half of AS2 as sample, two third of AS3 as sample. As like gesture recognition in section *V-A*, we compare our method with traditional DTW and M.Reyes`s algorithm,as the figure 3 showing.



Test 1



Test 2



Test 3

Figure 3 Performance of our method、M.Reyes`s and traditional DTW algorithm on MSR Action3D

Obviously in figure 3, our algorithm issuperior to traditional DTW and M.Reyes`s methodexcept Test 2. Especially in data set AS1 and AS2, the difference is more obvious. Additionally, we can see that the recognition rates of all three means are showing a rising trend from Test 1 to Test3respectively. Our method's result trend from Test1 to Test3 is in figure 4

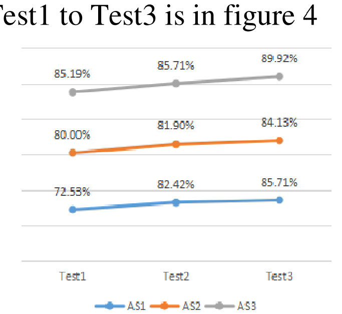


Figure 4 Result trend of our method

The main reason is that the number of sample data is rising from Test 1 to Test 3. On the contrary, average recognition time is also increase 0.1s to 0.3s respectively.

So far, by comparing above three algorithms, we can find that the reason why the traditional DTW algorithm's test result is low is that it takes all joints as an equal treatment in the calculation process. For different actions, however, the amplitude of each joint is different, so it's flawed to take all joints as an equal treatment

In addition, the reason why the overall result of M.Reyes is not as good as the result of this paper is that it uses all sample data in the calculation of joint weights whose are global weight. But, the proposed method can extract joints mainly involved in the movement for a specific action and adaptively calculate the joint weights. So, our method increase the recognition rate in a certain degree.

Finally, by comparing the experimental results of section *V-A* and *V-B*,the average recognition rate of our algorithm in section *V-A* is significantly better than the result of section *V-B*. There are two reasons. On the one hand, the action in MSR Action3D databaseis more complex than the movement in our database. On the other hand, the amount ofMSR Action3D are more higher than ourselves'.Overall, the average recognition rate of the proposed method almost reach to 85% to 90%.

VI. CONCLUSION

In this paper, based on the characteristic having different contribution to the final result for different joints of human in different action sequences, we propose a novel action recognition method based on adaptive weighted dynamic time warping algorithm.

The experimental results demonstrate that the average result of our method can reach to more than 85% on the database MSR Action3D and ourselves'. The advantage of this approach is to highlight the joint weights with significant changing and to suppress withinsignificant. Moreover, it also reduces the influence of noise in the motion sequence.The next step will be further optimize the approach of the weight given, compared with the weight of different distance measures to bring the effect.

ACKNOWLEDGMENT

This paper is supported by National Natural Science Foundation(No.U1431227,11273014,11403008); Guangdong Natural Science Foundation(No. 2014A030313374); Guangdong Provincial Science and Technology Education Innovation Project(No.2013KJCX0020）

REFERENCES

[1] Zhand Yi,Zhang Shuo,Luo Yuan,Xu Xiaodong,"Gesture track recognition based on Kinect depthimage information and its applications",Application Research of Computers,2012,29(9): 3547-3550.

[2] Yan Yan,Liu Rong,Huang Lu,Chen Ting,"The Study of gesture recognition based on hidden Markov model",Journal of HuangZhong Normal University(Nat.Sci.),2012,46(5):555-559.

[3]Thi-Lan Le, Minh-Quoc Nguyen,Thi-Thanh-Mai Nguyen,"Human posture recognition using human skeleton provided by Kinect",International Conference on Computing, Management and Telecommunications,Ho Chi Minh City,Vietnam:IEEE Press,2013,340-345.

[4] C.Schuldt,I.Laptev,B.Caputo,"Recognizing Human Actions:A Local SVM Approach",In Proceedings of the 17th IEEE International Conference on Pattern Recognition, Cambridge,UK,2004:32-36.

[5] F.Ofli,R.Chaudhry,G.Kurillo,R.Vidal and R.Bajcsy,"Sequence of the Most Informative Joints (SMIJ):A New Representation for Human Skeletal Action Recognition",IEEE Computer Society Workshops on Computer Vision and Pattern Recognition,Providence,RI,2012:8-13.

[6] Wang Xin,Wo Bohai,Guan Qiu,Chen Shengyong," Human action recognition based on manifold learning", Journal of Image and Graphics,2014,19(6):914-923.

[7] Lin Yongqiang,Wu Yadong,Yu Fang,YangYonghua,"Posture Sequence Finite-State Machine Method for Motion Recognition",Journal of Computer-Aided Design & Computer Graphics, 2014,26(9):1403-1411.

[8] S.Salvado,and P.Chan,"Fast DTW:Toward Accurate Dynamic Time Warping in Linear Time and Space",KDD Workshop in Mining Temporal and Sequential Data,2004:70-80.

[9] Carmona J M,Climent J,"A performance evaluation of HMM and DTW for gesture recognition",Lecture Notes in Computer Science:Progress in Pattern Recognition,Image Analysis,Computer Vision,and Applications, 2012, 7441: 236-243.

[10] Liu Fei,Huang Kuangrong,Ding Yngsheng,Liu Huan," Human Action Recognition Method Based on Depth Images", Computer Engineering, 2014,40(8):168-172.

[11] Li Haitao," Action Recognition Based on Constraintive DTW", Computer Simulation, 2014,31(11):227-230.

[12] M.Reyes,G.Dominguez and S.Escalera,"Feature Weighting in Dynamic Time Warping for Gesture Recognition in Depth Data",IEEE International Conference on Computer Vision,Barcelona, Spain, 2011:1182-1188.

[13] S.Sempena,N.U.Maulidevi and P.R Aryan,"Human action recognition using Dynamic Time Warping" International Conference on Electrical Engineering and Informatics,Bandung, Indonesia, 2011:1-5.

[14] J.Wang,H.Zheng,"View-robust action recognition based on temporal self-similarities and dynamic time warping",International Conference on Computer Science and Automation Engineering,Zhangjiajie,China,2012:498-502.

[15] Zhu Guanggang,Cao Lin," Human Motion Recognition Based on Skeletal Information of Kinect Sensor", Computer Simulation, 2014,31(12):29-333.

[16] Image Acquisition Toolbox.Using the Kinect for Windows Adaptor[Online].[2015-3].http://www.mathworks.com/help/pdf_doc/imaq/imaq_ug.pdf.

[17] Xianglilan Zhang,Jiping Sun,Zhigang Luo,Ming Li,"Confidence index dynamic time warping for language-independent embedded speech recognition",IEEE International Conference on Acoustics,Speech and Signal Processing,Vancouver,BC,2013:8066-8070.

[18] Zhang Jun,Li Xuebin," A Method of Isolated Word Recognition Based on DTW" , Computer Simulation,2009,10(26):348-364.

[19] Yu-Liang Hsu,Cheng-Ling Chu,Yi-Ju Tsai,Jeen-Shing Wang,"An Inertial Pen With Dynamic Time Warping Recognizer for Handwriting and Gesture Recognition", Sensors Journal,IEEE,2015, 15(1):154-163.

[20] Gan Huang,Dingguo Zhang,Xidian Zheng,Xiangyang Zhu,"An EMG-based handwriting recognition through dynamic time warping",Annual International Conference of the IEEE Engineering in Medicine and Biology Society,Buenos,Aires,2010:4902-4905.

[21] Yu Qiao and Makoto Yasuhara,"Affine invariant dynamic time warping and its applications to online rotated handwriting recognition",Proceedings of the IEEE International Conference on Pattern Recognition.2006,2:905-908.

[22] Mengxing Huang,Wenjiao Yu,Donghai Zhu,"An Improved Image Segmentation Algorithm Based on the Otsu Method",ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel & Distributed Computing, Kyoto,Japan,2012: 135-139

[23]MSRAction3D[Online].http://research.microsoft.com/en-us/um/people/zliu/ActionRecoRsrc/default.htm.

[24] Li Wangqing,Zhang Zhengyou,Liu Zicheng,"Action Recognition Based on a Bag of 3D Points",Proc.of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops,San Francisco,USA,2010: 9-14 .

Author's Profile

ZhaoLiang Li(1990-),male,master, major in Image processing and pattern recognition in JiNan University from 2014 years.