Human Action Recognition Based on Random Spectral Regression

GuangFeng Lin[1], Hong Zhu[2], YinDi Fan[3], and CaiXia Fan[1]

[1] Information Science Department, Xi'an University of Technology, Xi'an, China
lgf78103@ xaut.edu.cn, fcx_1981@yahoo.com.cn
[2] Electronic Information Engineering Department,
Xi'an University of Technology, Xi'an, China
zhuhong@ xaut.edu.cn
[3] Shaanxi College of Communication Technology, Xi'an, China
nezimi5584@126.com

Abstract. For solving the uncertain parameter selection, the highly spatio-temporal complexity and the difficulty of effectively extracting feature in manifold learning algorithm processing higher-dimension of human action sequence, human action recognition algorithm based on random spectral regression (RSPR) is presented. The algorithm has three steps. Firstly, according to uniform distribution of human action data in the manifold and the classification labels of human action, the weight matrix is built. This method overcomes the neighborhood parameter selection of the manifold learning algorithm. Secondly, by spectral regression, the spatial manifold based on frame is approximated, and the manifold mapping of unlabeled sample is obtained. At last, the feature of the temporal series is extracted in the spatial manifold based on frame, and then in Gaussian process classification the feature of human action is classified. The experiment has three parts. When RSPR tests the recognition of human action by leave-one-out crossvalidation in Weizmann database, the recognition rate reach 93.2%; comparing RSPR with locality preserving projection (LPP) and neighborhood preserving embedding (NPE), through extracting the statistical feature of temporal sequences RSPR shows better performance; in the test of walk action influenced RSPR displays better adaptability.

Keywords: random spectral regression, manifold learning, locality preserving projection, neighborhood preserving embedding, human action recognition.

1 Introduction

In machine and computer vision research, human action recognition can be applied in many fields, such as video surveillance, human-computer interaction and robot vision. In human action recognition the first problem is feature description, which include two broad categories [1]: the global representation (contour or silhouette [2],grid-based representations[3] and space-time volumes[4]) and the local representation (space-time interesting point detector[5], local descriptors[6], local grid-based

H. Deng et al. (Eds.): AICI 2011, Part III, LNAI 7004, pp. 451–461, 2011.

representation[7] and correlations between local descriptors[8]). The above representations mostly involve spatial, temporal features and spatio-temporal features, but no matter what kind of features forming feature vector, the dimensions of which are very large. Now there are many methods of linear dimensionality reduction (principal component analysis, singular value decomposition, independent component analysis and linear discriminant analysis) and nonlinear dimensionality reduction (nonlinear subspace method of kernel mapping, manifold learning) [9]. During the process of the research it is recognized that the non-linear manifold is the perception of the foundation, because highly dimensional information typically exists in a non-linear low-dimensional manifold and to a large extent cognitive processes understands things through the nonlinear low-dimensional manifold [10]. In [11] by local preserving projection the manifold of human action sequence is analyzed, in [12] the manifolds distinguishing feature of the transform is learned, in [13] in spatio-temporal data the manifold discriminant of human action is well identified. However, there are three questions, firstly the traditional manifold building graphic models, using $\varepsilon-\mathrm{ball}$ nearest neighbor or K-nearest neighbor of Euclidean distance constructing neighborhood to calculate the weights, the selection of parameter ε or K greatly affects on the efficiency of the algorithm, moreover the samples of European distance nearest neighbor often do not belong to the same class [14]; secondly in the traditional manifold of human action frames the dimensions of the sample are greater than the number of samples, which involves a possible dense matrix, and increases the complexity of the algorithm [9]; thirdly, the time series of human action have strongly associated, between the different human action frames under some conditions (run and skip) have the greater similarity, whereas the low-dimensional manifold only extracts the local structural feature, and could not obtain the optimal classification feature. For the above three questions random spectral regression algorithm is presented to extract the low-dimensional manifold of human action space, then the time series features of space manifold is obtained, lastly the Gaussian process classifier for classification of human action is used. In the article the study object is the whole human silhouette sequences, and the study mainly focuses on the manifold description of the feature and the classification efficiency of the manifold methods.

This article has five sections. In section 1 the main work associated with this article is narrated; in section 2 the manifold learning algorithm based on RSPR on human actions frame is proposed; in section 3 the time series features are extracted, and Gaussian processes classifier is established; in section 4 the contrast experiment is carried out among LPP, NPE and RSPR; in section 5 the conclusion is obtained.

2 Related Research

Classic manifold learning algorithm is divided into two types, which is global and local. Multidimensional Scaling (MDS) [15] is the global method of the linear dimensionality reduction, in which the structural features are measured by Euclidean distance matrix, but do not represent the nonlinear relationship of the manifold between samples. In Isometric mapping (ISOMAP) [16] the structural feature is measured by geodesic distance, however, the cost of the algorithm is large because of estimating the shortest path on the global neighborhood graph, moreover the noise and empty of the sample influence the efficiency of the algorithm greatly. Locally Linear Embedding (LLE) [17] is the local algorithm of the non-linear dimensionality reduction, which gets neighborhood graph by the linear reconstruction of the sample, then obtains the global linear construction by the local fitting, but the question is the parameter effects the reconstruction and the non-neighborhood data mapping. The learning method of the manifold can find the geometry features of the manifold, and then reflect the nature of the potential contact data. The actual application has two aspects that are neighborhood selection and extension of algorithm generalization. To a certain extent Linear Extension of Graph Embedding (LGE) [18] has solved the extension of algorithm generalization, moreover derived different linear dimensionality reduction algorithms from the different weight matrices selection, such as Locality Preserving Projection (LPP) [19] and Neighborhood Preserving Embedding (NPE) [20]. However neighborhood selection question is not solved all along.

In the ordinary construction of the graphic, ε–bal neighbors or K nearest neighbors is used for local neighborhood, between which the weight is obtained by binary number, Gaussian kernel and L2 reconstruction. In [21] L1 reconstruction of the sample data is put forward, and the reconstruction parameter is regarded as the weight of the construction graphic. L1 method has the broader locality than $\varepsilon - \text{ball}$ neighbor or K-nearest neighbor, but its cost is large when reconstructing data in the whole sample.

Spectral regression [22] is the regularization method of the subspace learning, and effectively avoids computational complexity increase, which is produced by the dense matrix decomposition. In face database the efficiency of algorithm is proved, which extends the application scope of the manifold learning algorithm.

To analyze time series data, in [23] one-dimensional sequence the features of the data is described, and forms 13 eigenvector, which is clustered, then the higher accuracy of the clustering is obtained. In [24] human action sequences data is described by the feature, in human action database the better recognition rate is obtained by Gaussian process classifier.

3 Manifold Learning Algorithm Based on Random Spectral Regression

Establishing neighborhood graph of human actions frame is to preserve and extract the frame structure of human action spatial information, which includes the intra-class and inter-class features of human silhouette, and the sequence feature can describe the main discrimination feature of human action, although it might not be the optimal discrimination feature, the method may reduce the spatial dimension of the frame. The article studies the main content that is how to use the random graph to extract the classification information of the manifold structure, and in the spectral regression framework how to obtain the valid feature description and reduce the dimensions of human action frame.

The current manifold learning on face recognition application is more, but on the classification of human action sequence is fewer, and applies only to the dimension reduction of data, does not go into the nature of the data sets. Between human action sequence and face images three differences exist:

(a) Face feature can be clearly represented by the description of the single image, while human actions require the multiple-frame images to describe clearly.

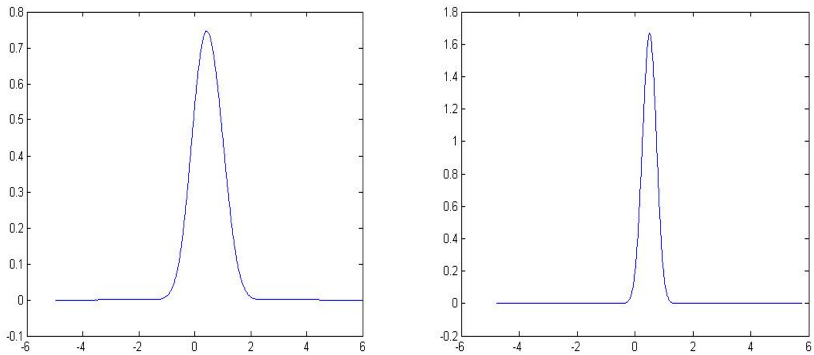
(b) Between face images, the temporal correlation of the data is not strong, from the view of the data distribution that face image is independent, so the independent feature is similar obviously in the same class, and the independent feature is different widely in the different class. Human action images have the strong temporal correlation, from the view of the data distribution human action frames is not independent, so the correlative feature is similar obviously in the different frame of the same class, and the correlative feature is different widely in the different frame of the different class.

(c) The feature extraction of face data only considers the spatial dimension reduction, while the feature extraction of human action data must take account of the spatial and temporal dimension reduction.

Since there are the differences, when building data sets graph the manifold learning should apply the more suitable method for human action recognition. The approximation of the manifold will reflect the nature of the data set.

Human action data is $X_{l\times n} = \{ x_i^c | i \in R \;\; and \;\; 0 < c < m \}$, x_i^c is the *i*th frame image in the *c* class action, have *m* class human action and *l* frame images, *n* is each frame image generating the dimension of row vector. Thru the manifold learning the data of the dimension reduction is $Y_{l\times d}$, d($d << n$) is the reductive dimensions. In the actual observation, video image data and neighbors image of its frame sequences is the most correlative image, moreover the correlation reflects the structure of the video image frame, and each image frame of the similar semantics is such. The structural feature distribution of the data frame regarded as the uniform distribution, the structure relation can be viewed as the manifold approximation of the pair data, and when mapping to the low-dimensional space the structure information is preserved.

Building the undirected graphs of human action image frame, the most factors of the inherent manifold structure character considered are multiple frame, strong temporal correlation, spatio-temporal character and priori classification label of data sample. Specific rules are depending on the data manifolds uniform distribution, the pseudo-random sequence of the uniform distribution may show the differences of the frame feature and the uniform distribution of the manifold between the single image sequences of the intra-class, at the same time reflects the strong temporal correlation. In the multiple image sequences of the intra-class, the same pseudo-random sequence represents the manifold of the same action. Between the multiple image sequences of the different class, the different pseudo-random sequence shows the difference of the inter-class. In comparison with the tradition method, the random spectral regression does not need Euclidean distance information of sample when constructing graph, the neighborhood parameter selection and the weight computation. However, the constructing graph depends on two-type natural information of the data, which are the uniform distribution character of the data manifold and the classification label of the data. Fig. 1 shows LPP and RSPR manifold distribution of human action by kernel density estimator. Because the estimator involves all kinds of the action, the character of the distribution presents the similar normal distribution, but one thing may assure manifold distribution of human action is similar in the different methods, moreover RSPR manifold distribution of human action has smaller variance in Fig. 1, which shows the better manifold approximation.



(a) Manifold distribution of LPP

(b) Manifold distribution of RSPR

Fig. 1. Manifold distribution of human action by kernel density estimator

Through the prior knowledge to get m sample class, the pseudo-random square matrix $A_{m\times m}$ of the uniform distribution is generated by subtraction with borrow [24], and then $B_{l\times m}$ is obtained from $A_{m\times m}$ according to the sample label information. In the same class each row of B matches with each row of A. Using Gram-Schmidt orthogonalization orthogonal matrix Q is obtained by QR decomposition of B, but matrix Q isn't square matrix, which isn't good for the manifold approximation and the linear extension, so the weight matrix W is constructed by formula (1).

$$W = QQ^T \tag{1}$$

When constructing the weights matrix W of the graph, which is combined with spectral regression theory, the mapping matrix a of the linear extension can be solved. $y = [y_1,..., y_m]$ is the mapping of the data from the graph. If the mapping of the adjacent node i and j is away from, they will be punished. Minimizing (2) assures the mapping y_i and y_j of the adjacent node i and j is adjacent.

$$\sum_{i,j}(y_i - y_j)^2 W_{ij} \tag{2}$$

From the above (2) formula (3) is derived.

$$y^* = \underset{y^T Dy=I}{\arg\max}\, y^T Wy = \arg\max \frac{y^T Wy}{y^T Dy} \tag{3}$$

In formula (3) $D_{ii} = \sum_j W_{ij}$. Formula (3) deduces formula (4) in Least Square Method. The optimal y^* can be obtained by decomposition (4).

$$Wy = \lambda Dy \tag{4}$$

The linear extension of y is $y = X^T a$, which is imputed formula (3) to obtain formula (5).

$$XWX^T a = \lambda XDX^T a \tag{5}$$

Because XWX^T is often dense, Cai [18] proposes the regularization method of solving a^* which is the optimal fitting for equation (6) in Least Square Method.

$$a^* = \arg\min_a (\sum_i (a^T x_i - y_i)^2 + h\|a\|^2) \tag{6}$$

When h closes zero, a^* has the stable solution.

4 Gaussian Process Classification Based on the Feature of Temporal Series

The d dimension data of human action is obtained by the manifold learning of RSPR, each dimension point sequence is Y_l^d, which is the *dth* point sequence of the *lth* action sequence. Y'^d_l is the adjusted data of Y_l^d, and then the statistical features of the temporal series for Y_l^d and Y'^d_l is chosen: trend, seasonality, serial correlation, non-linearity, skewness, kurtosis, self-similarity, chaotic and periodicity [22]. These features form the action feature vector $U_{d\times 13}$ for each action sequence.

The dimensions reduction data set of human action is $U = \{u_{lable}, u_{unlable}\}$, the classification label is V. U and V exists the implicit function $f(u)\sim GP(M,K)$, which defines the mapping of U and V. GP is Gaussian process [25], and $M(u)$ is mean function, $K(u,u')$ is the covariance function. In the article the radial basis function is defined as formula (7).

$$K(u,u') = \sigma \exp[-\frac{1}{2}(\frac{u-u'}{s})^2] \tag{7}$$

Input pair is u and u', σ is the variance of input data, s is the length of input data. In Laplace approximation algorithm the maximum posteriori probability $p(f|U,v)$ and the corresponding f is solved. Gaussian process binary classifier firstly computes the corresponding $p(f^*|U,v,u^*)$ of the implicit function f^*, u^* is the test data set.

$$p(f^* | U,v,u^*) = \int p(f^* | U,u^*,f)\, p(f | U,v)\, df \tag{8}$$

In formula (8) $p(f \mid U,v) = p(v \mid f)p(f \mid U)p(v \mid U)$ is the posteriori probability of the implicit function. $p(f^* \mid U,u^*,f)$ is the posteriori probability of the predict data, thru formula (9) the classification label probability of the test data is computed.

$$p(v^* = +1 \mid U,v,u^*) = \int \Phi(f^*)p(f^* \mid U,v,u^*)df^* \tag{9}$$

In formula (9) $\Phi(f^*)$ is the cumulative Gaussian function, which assures the probability of the classification label within [0, 1].

The multi-class of human action recognition may make the label of one class sample be +1 and the label of other class sample be -1 to construct the classifier. Although the method is not optimal, it is still efficient.

5 Experiment

In the experiment human action database derives from Weizmann action database [26], which has 113 low resolution video sequences(180×140 pixel, 25 frame/second), and includes two parts: one part has 93 video sequences of 10 class action (bend, jack, jump, pjump, run, side, skip, walk, wave1, wave2) which is completed by 9 experimenter, another part is 20 video sequences of 1 class action (walk), which is influenced by several factors: visual angle, occlusion, clothing, accompanying item. Because of the effect of the shadow, color and other noise, the difference exists between the nature silhouette and the extracted silhouette.

The experiment is designed in three parts. Firstly the recognition method based on RSPR is tested. Secondly the performance of RSPR, LPP and NPE is contrasted. At last the robustness of the algorithm is tested.

5.1 Performance Test Based on RSPR

In the performance test of RSPR, in the first part 93 video sequences is used, which have 5687 frames. In these sequences some action are not reduplicate in time series (for instance bend), others are reduplicate (for instance run and walk). In the experiment the overlap action sequence is not segmented, because the redundancy data is eliminated by the statistical features extracted. Video sequences is divided into 10 separated packet, each separated packet has 10 different actions, in these packet there are tow packets belong to the different action of the same person. Classifier is constructed by 9 separated packet learning and 1 separated packet testing. Loop using these data constructs the classifier. Table 1 shows the confusion matrix of the recognition in 10 actions.

In table 1 the result of the experiment shows between pjump and side action the confusion is more, between run and skip action the confusion is such, the observation is the similar state. Finally the average recognition rate reaches 93.2%.

5.2 Contrast and Analysis

In the experiment between RSPR, LPP and NPE the contrast of the performance has three points, which include the time cost of the manifold learning, the error rate of human action recognition and the average time cost. The experimental device is a computer, which has AMD64 bits, 2.3GHz CPU and 2G memory. In LPP and NPE test of the whole data the memory overflows owing to the computation and decomposition dense matrix, so only using one separated packet for the manifold learning explains the above points.

Table 2 shows the time cost of the manifold learning, the error rate of human action recognition and the average time cost in three methods.

Table 1. The confusion matrix of the recognition in 10 actions

	Bend	Jack	Jump	Pjump	Run	Side	Skip	Walk	Wave1	Wave2
Bend	**96%**	2%						2%		
Jack	2%	**98%**								
Jump			**98%**		1%		1%			
Pjump				**90%**		8%			2%	
Run			2%		**88%**		8%	2%		
Side				6%		**92%**				2%
Skip			2%		10%		**88%**			
Walk				2%	2%			**96%**		
Wave1				4%					**92%**	4%
Wave2						2%			4%	**94%**

Table 2. The comparison of three method

method	LPP	NPE	RSPR
Learning time	17.21s	40.87s	5.13s
Average error rate	70%	64%	61%
Average recognition time	2.109s	1.910s	1.904s

According to the experiment results, the analysis has three aspects:

1) From the view of the algorithm theory, LPP makes up the basic framework of the manifold learning extension, by Euclidean distance measuring the local neighborhood and heat kernel function computing the weight matrix, in the low dimension space the minimal optimization question is constructed, and the mapping is obtained from the high dimension data to the low dimension data. The difference of NPE lies in L2 reconstruction in high dimension data and using the reconstructed coefficient as the weight matrix. The characteristic of RSPR is the use of the manifold distribution of high dimension data to product the random distribution matrix, and generate the weights matrix combining with the labels of the known data, which obviates the parameter selection of local neighborhood.

2) From the view of the algorithm time complexity, the Time Complexity of the weight matrix computation by LPP is O (n^2×m); the Time Complexity of the weight matrix computation by NPE is O (n^2×m^2); the Time Complexity of the weight matrix computation by RSPR is O (m).
3) From the view of the algorithm space efficiency, in the experiment of LPP and NPE, the Space Complexity is produced by the dense matrix decomposition, which occupies the enormous memory. When the number of the sample exceeds 800, in the computer the exception of the memory overflowing arises, however, the number of the sample is at least 5000 for the learning action pattern. Therefore, the above contrast experiment is carried out in one separated packet, which includes 701 samples, although the error rate of the recognition is high, the performance of RSPR is better than LPP and NPE.

5.3 Algorithm Robust Test

The images of walk action for robust test in Weizmann action data, walk action for robust test includes 20 sequences, and top 10 sequences are walk action with the view angle change, which is obtained from zero degrees to 90 degrees. The angle is defined the angle of camera imaging plane and walk orientation. The rest 10 sequences respectively are walk with bag, walk with briefcase, kneesup walk, limp, arms straight walk, occluded foot walk, walk, fixed block walk, walk with skirt and walk with dog. The table 3 shows the recognition result of the sequence.

In table 3, the result shows RSPR has certain adaptability for the influence of the view angle, but angle of view changes more than 45 degrees does not correctly recognize the action class. In the greater deformation walk (walk with bag, kneesup walk and arms straight walk) or occluded and confused walk (walk with dog) the recognition is incorrect, while under the rest condition the recognition is correct.

Table 3. The recognition results of the robust test

Name	Recognition conclusion	Name	Recognition conclusion
robust_00	Walk	robust_bag	Jack
robust_09	Walk	robust_briefcase	Walk
robust_18	Walk	robust_kneesup	Side
robust_27	Walk	robust_limp	Walk
robust_36	Walk	robust_moonwalk	Jack
robust_45	Walk	robust_nofeet	Walk
robust_54	Side	robust_normwalk	Walk
robust_63	Jack	robust_pole	Walk
robust_72	Side	robust_skirt	Walk
robust_81	Side	robust_dog	Skip

6 Conclusions

In the article, human action recognition algorithm based on RSPR is presented, which solves the parameter selection of the manifold learning, and reduces the time and space complexity combining with spectral regression, puts forward the feasible method to extract the manifold of the large scale video data, and then the discrimination feature of human action is extracted by using the statistic feature of the mapped data, finally in the experiment the feasibility and efficiency of the algorithm is proved. In the next research the manifold distribution of local feature will be studied, and further human action is analyzed and discriminated.

Acknowledgment. The authors would like to thank the anonymous reviewers for their insightful comments, which have helped to improve the quality of this paper. Scientific Research Program Funded by Shaanxi Provincial Education Department (Program No.2010JK718) supported this research.

References

[1] Poppe, R.: A survey on vision-based human action recognition. Image and Vision Computing 28, 976–990 (2010)

[2] Bobick, A.F., Davis, J.W.: The recognition of human movement using temporal templates. IEEE Trans. Pattern Analysis and Machine Intelligence 23, 257–267 (2001)

[3] Kellokumpu, V., Zhao, G., Pietikainen, M.: Human activity recognition using a dynamic texture based method. In: Proceedings of the British Machine Vision Conference (BMVC 2008), Leeds, United Kingdom, pp. 885–894 (2008)

[4] Gorelick, L., Blank, M., Shechtman, E., Irani, M., Basri, R.: Actions as space–time shapes. IEEE Trans. Pattern Analysis and Machine Intelligence 29, 2247–2253 (2007)

[5] Willems, G., Tuytelaars, T., Van Gool, L.: An efficient dense and scale-invariant spatio-temporal interest point detector. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part II. LNCS, vol. 5303, pp. 650–663. Springer, Heidelberg (2008)

[6] Wang, H., Ullah, M.M., Klaser, A., Laptev, I., Schmid, C.: Evaluation of local spatio-temporal features for action recognition. In: Proceedings of the British Machine Vision Conference (BMVC 2009), London, United Kingdom (2009)

[7] Zhao, Z., Elgammal, A.: Human activity recognition from frame's spatiotemporal representation. In: Proceedings of the International Conference on Pattern Recognition (ICPR 2008), Tampa, FL, pp. 1–4 (2008)

[8] Farhadi, A., Tabrizi, M.K.: Learning to recognize activities from the wrong view point. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part I. LNCS, vol. 5302, pp. 154–166. Springer, Heidelberg (2008)

[9] Li, Y.: Researehon Methods of Subspace Feature Extraction Based on Manifold Learning in Face Recognition. Central South University (2009)

[10] Seung, H.S., Lee, D.D.: The manifold ways of perception. Science 290, 2268–2269 (2000)

[11] Wang, L., Suter, D.: Visual learning and recognition of sequential data manifolds with applications to human movement analysis. In: Computer Vision and Image Understanding (CVIU), vol. 110, pp. 153–172 (2008)

[12] Poppe, R., Poel, M.: Discriminative human action recognition using pairwise CSP classifiers. In: Proceedings of the International Conference on Automatic Face and Gesture Recognition (FGR 2008), Amsterdam, The Netherlands, pp. 1–6 (2008)
[13] Jia, K., Yeung, D.-Y.: Human action recognition using local spatio-temporal discriminant embedding. In: Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR 2008), Anchorage, AK, pp. 1–8 (2008)
[14] Yan, S., Wang, H.: Semi-supervised learning by sparse representation. In: SIAM International Conference on Data Mining (SDM 2009), Sparks, Nevada, USA, pp. 792–801 (2009)
[15] Cox, T.F., Cox, M.A.: Multidimensional scaling, pp. 24–139. Chapman & Hall, London (2001)
[16] Tenenbaum, J.B., De Silva, V., Langford, J.C.: A global geometric framework for nonlinear dimensionality reduction. Science 290, 2319–2323 (2000)
[17] Roweis, S.T., Saul, L.K.: Nonlinear dimensionality reduction by locally linear embedding. Science 290, 2323–2326 (2000)
[18] Deng, C., He, X.F., Han, J.W.: Spectral regression for efficient regularized subspace learning. In: IEEE International Conference on Computer Vision, Rio de Janeiro, Brazil, pp. 214–221 (2007)
[19] He, X.F., Niyogi, P.: Locality preserving projections. In: Proc. Of Advances in Neural Information Processing System, Vancouver, Canada (2003)
[20] He, X.F., Deng, C., Yan, S.C., et al.: Neighborhood preserving embedding. In: Proc. of 10th IEEE International Conference on Computer Vision, Beijing, China, pp. 1208–1213 (2005)
[21] Cheng, B., Yang, J., Yan, S., Huang, T.S.: Learning With L1-Graph for Image Analysis. IEEE Transactions on Image Processing 19, 858–866 (2010)
[22] Wang, X., Smith, K.A., Hyndman, R.J.: Characteristic-based clustering for time series data. Data Mining and Knowledge Discovery 13, 335–364 (2006)
[23] Zhou, H., Wang, L., Suter, D.: Human action recognition by feature-reduced Gaussian process classification. Pattern Recognition Letters 30, 1059–1066 (2009)
[24] Pi, X.: A Soft of Pseudorandom Number Sequences With Extremely Long Period. Journal of Unmerical Methods and Computer Applications 4, 286–292 (2001)
[25] Rasmussen, C.E., Williams, C.K.I.: Gaussian Processes for Machine Learning. The MIT Press, Cambridge (2006)
[26] http://www.wisdom.weizmann.ac.il/~vision/SpaceTimeActions.html