Hand Gesture Recognition Based on Online PCA with Adaptive Subspace

MINGHAI YAO ,XINYU QU,QINLONG GU
College of Information Engineering
Zhejiang University of Technology
Hangzhou 310014,CHINA
ymh@zjut.edu.cn

Abstract: - The learning method for hand gesture recognition that compute a space of eigenvectors by Principal Component Analysis(PCA) traditionally require a batch computation step, in which the only way to update the subspace is to rebuild the subspace by the scratch when it comes to new samples. In this paper, we introduce a new approach to gesture recognition based on online PCA algorithm with adaptive subspace, which allows for complete incremental learning. We propose to use different subspace updating strategy for new sample according to the degree of difference between new sample and learned sample, which can improve the adaptability in different situations, and also reduce the time of calculation and storage space. The experimental results show that the proposed method can recognize the unknown hand gesture, realizing online hand gesture accumulation and updating, and improving the recognition performance of system.

Key-Words: - Online Learning , Online PCA , Adaptive Subspace , Hand Gesture Recognition

1 Introduction

Hand gesture recognition is an important method for human-machine interaction, which is widely used in variety of areas such as gesture language service, intelligent human-machine interface, interfaces of virtual reality navigation and manipulation, robot demonstrating learning, digital art and entertainment. In recent years, hand gesture recognition based on computer vision has aroused wide interest among researchers. Compared with data glove or electromagnetic wave methods, it is more natural and flexible using computer vision in hand gesture recognition.

The process of hand gesture recognition based on vision can be commonly divided into three phases, that is segmentation, representation and recognition, among which the recognition phase is composed of classifier learning part and recognition part. Training classifier with samples(learning) followed by recognition of hand gestures using trained classifier achieved a recognition phase. Yet there are limitations and disadvantages with this recognition architecture, which is, 1) since it is off-line and batch learning for classifier, the classifier cannot be updated when it comes to new sample unless learning repeatedly with previous learned samples in an off-line mode; 2) the adaptability and robustness of the recognition system is bad in new situations since it is unable to detect novel hand gestures.

Because of the bad real-time performance, expansibility and robustness, researchers have proposed several online learning architectures and algorithms. Stephan Kirstein et al.[9] have proposed online vector quantization algorithm and incremental learning vector quantization algorithm for online learning of objects using Gabor hierarchical features. The method learnt and recognized 50 objects online in 3 hours. J. Luo et al.[10] have applied incremental SVM algorithm for mobile robot to online learning of indoor scenes. Reference [11]-[13] have focused on online PCA algorithm for online learning of visual objects, behaviors and scenes.

In 2003, D. Skocaj et al.[14] proposed a weighted and robust incremental PCA method for subspace learning, which aroused a good many researchers to pay much attention to it. Peter M. Roth et al.[15] have applied this method for online learning of unknown hand held objects. H. V. Neto et al.[16] have applied this method for novelty detection on mobile robot. In this paper, we proposed the adaptive subspace method to improve the subspace updating strategy in [14], and then apply the improved method named online PCA with adaptive subspace in hand gesture recognition.

The organization of this paper is as follows: a brief introduction of system architecture is given in Section II. Section III describes the theory and flow of online PCA with adaptive subspace algorithm. In Section IV, experiments are designed for the comparison of proposed method and off-line PCA, and results are presented to validate the advantages of the proposed method. Finally, conclusions are made in Section V.

2 Architecture for Online Learning of Hand Gestures

Online learning and recognition of hand gestures is mainly composed of preprocess, initialization of classifier, online PCA learning for classifier. The proposed architecture is shown in Fig. 1.

In order to assure the effectivity and accuracy of PCA algorithm, preprocesses are needed for the input video stream, including video stream reading, localization, tracking and segmentation of hand gestures. Finally, the normalization of the image sequences is done after the segmentation.

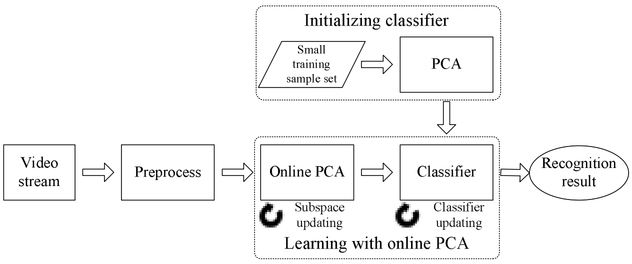

Fig. 1 Architecture for online learning and recognition of hand gestures

Before online learning, a few hand gesture samples are needed for off-line PCA to obtain an initial subspace and classifier. Since the subspace and classifier are not accurate, more hand gesture samples are learned online through real-time collection. With the online PCA, the subspace and the classifier are updated online then a better performance is achieved.

The part of online PCA learning is the key point of this system. The new mean vector, new eigenvalues, new eigenvectors and new subspace are obtained through real-time online learning of hand gesture samples after preprocess. New hand gesture sample is projected onto the previous computed subspace, and then recognized or detected of a novel hand gesture according to the coefficient vector. As the system run as the time goes, more samples are learned and classifier is updated continuously, higher recognition rate is achieved, through online learning, computing and recognition.

3 Online PCA with Adaptive Subspace Method

3.1 The Proposed Method

Assume that the subspace and the classifier were already built from n hand gesture samples. Some variables are defined as follows: the current subspace $\mathbf{U}^{(n)}$, the current sample mean vector $\overline{\mathbf{x}}^{(n)}$, the current coefficient vectors $\mathbf{A}^{(n)}$, a new sample $\mathbf{x}$, its reconstruction $\hat{\mathbf{x}}$ and projection $\mathbf{a}$, the updated subspace $\mathbf{U}^{(n+1)}$, the updated mean vector $\overline{\mathbf{x}}^{(n+1)}$ and the updated coefficient vectors $\mathbf{A}^{(n+1)}$.

At step n+1, when the new sample $\mathbf{x}$ comes, the new subspace can be obtained by calculating PCA through low dimensional coefficient vectors instead of high dimensional reconstructions, since coefficient vectors and reconstructed images encompass the same visual variability, i.e., they are just represented in different coordinate frames[14]. Since this method is computationally very efficient, the need for real-time application is fulfilled.

At step n+1, the new sample $\mathbf{x}$ is projected onto the current subspace $\mathbf{U}^{(n)}$ and the projection is obtained

$$\mathbf{a} = \mathbf{U}^{(n)T}\left(\mathbf{x} - \overline{\mathbf{x}}^{(n)}\right) \tag{1}$$

Reconstruct the sample

$$\hat{\mathbf{x}} = \mathbf{U}^{(n)}\mathbf{a} + \overline{\mathbf{x}}^{(n)} \tag{2}$$

Compute the residual vector

$$\mathbf{r} = \mathbf{x} - \hat{\mathbf{x}} \tag{3}$$

And its Euclidean norm $\|\mathbf{r}\|$ which is orthogonal to $\mathbf{U}^{(n)}$, and the new basis is obtained

$$\breve{\mathbf{U}} = \begin{pmatrix} \mathbf{U}^{(n)} & \dfrac{\mathbf{r}}{\|\mathbf{r}\|} \end{pmatrix} \tag{4}$$

Build a new coefficient vector matrix

$$\breve{\mathbf{A}} = \begin{pmatrix} \mathbf{A}^{(n)} & \mathbf{a} \\ \mathbf{0} & \|\mathbf{r}\| \end{pmatrix} \tag{5}$$

In which $\mathbf{A}^{(n)}$ is the coefficient vectors of previous n samples. The mean vector $\breve{\boldsymbol{\mu}}$ and subspace $\breve{\mathbf{U}}'$ is obtained by calculating PCA of $\breve{\mathbf{A}}$. The updated coefficient vectors of step n+1 is obtained by

$$\mathbf{A}^{(n+1)} = \breve{\mathbf{U}}'^{T}\left(\breve{\mathbf{A}} - \breve{\boldsymbol{\mu}}\mathbf{1}\right) \tag{6}$$

The updated mean vector is

$$\overline{\mathbf{x}}^{(n+1)} = \overline{\mathbf{x}}^{(n)} + \breve{\mathbf{U}}\breve{\boldsymbol{\mu}} \tag{7}$$

And the updated subspace is obtained by rotating the new basis $\breve{\mathbf{U}}$

$$\mathbf{U}^{(n+1)} = \breve{\mathbf{U}}\breve{\mathbf{U}}' \tag{8}$$

The updated $\mathbf{A}^{(n+1)}$, $\overline{\mathbf{x}}^{(n+1)}$ and $\mathbf{U}^{(n+1)}$ then become the current state for next updating loop, by which the system online learning of hand gestures.

The problem in the above online PCA is that in each step the dimension is increased by one, as a result, computational cost and storage are increased and redundant sample information is produced. Therefore, the concept of adaptive subspace is proposed in this paper to adjust the subspace updating strategy of the above online PCA algorithm.

Each new hand gesture sample can be divided into three situations: novel hand gesture class, already learned class with low similarity to learned samples, and already learned class with high similarity to learned samples. We present two threshold to differentiate three situations: θclass, threshold of reconstruction error for inter-class; θdistance, threshold of distance for inner-class. The updating strategy is divided into three ways according to θclass and θdistance.

1) Novel hand gesture class. If $\|\mathbf{r}\|>\theta_{class}$, it is a novel class. The subspace is updated through online PCA algorithm described above, and the dimension of subspace is increased by one, namely $dim(\mathbf{U}^{(n+1)})=dim(\mathbf{U}^{(n)})+1$;

2) Already learned class with low similarity to learned samples. Define deuclidean_min as the minimum Euclidean distance between new sample and all the learned samples. If $\|\mathbf{r}\|<\theta_{class}$ and deuclidean_min$>\theta_{distance}$, the similarity between new sample and learned samples is low, perform the online PCA described above. The difference compared to 1) is that the subspace dimension is remain unchanged, namely $dim(\mathbf{U}^{(n+1)})=dim(\mathbf{U}^{(n)})$;

3) Already learned class with high similarity to learned samples. If $\|\mathbf{r}\|<\theta_{class}$ and deuclidean_min$<\theta_{distance}$, it is considered that the sample is redundant for class representation. The sample is classified directly without learning.

Each sample is classified through the corresponding sample class of deuclidean_min in situation 2) and 3). Assume the current learned samples $\mathbf{A}=[\mathbf{a}_1,\mathbf{a}_2,...,\mathbf{a}_w]$, classes are $class_1,class_2,...,class_j$, and the projection $\mathbf{a}$ of new sample $\mathbf{x}$, if deuclidean_min= $\|\mathbf{a}-\mathbf{a}_i\|_{Euclidean}$, in which $(i=1,2,...w)$ and $\mathbf{a}_i\in class_s\,(1\le i\le w,1\le s\le j)$, thus $\mathbf{a}\in class_s$, namely sample $\mathbf{x}\in class_s$, recognition complete.

With this subspace updating strategy, the subspace dimension will not increase rapidly, the computational cost and storage are saved, therefore, it is suitable for hand gesture recognition with high demand of real-time performance. The hand gesture classifier is learned through online PCA with adaptive subspace, which let the hand gesture learning perform in a real-time, continuously and incrementally fashion, and the subspace updating strategy can be adjusted adaptively to different situations of hand gestures. The structure of online PCA with adaptive subspace is shown in Fig. 2.

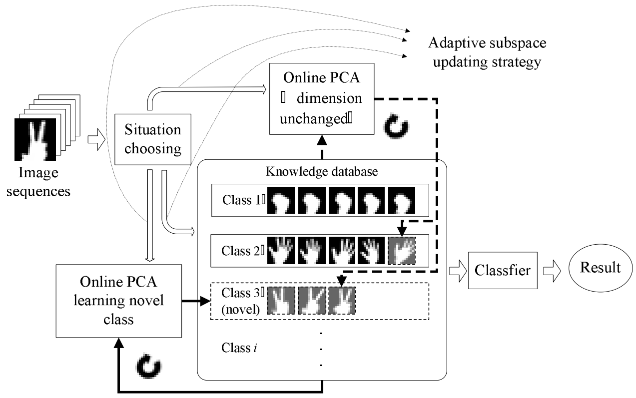

Fig. 2 Structure of online PCA with adaptive subspace

3.2 Algorithm Flow

The flow steps of the proposed algorithm, which is obtained through the combination of the subspace updating strategy and online PCA, is described as follows:

Step1 When the new sample comes, calculate its projection, reconstruction and residual vector by formula(1),(2),(3);

Step2 If $\|\mathbf{r}\|>\theta_{class}$, calculate updated subspace $\mathbf{U}^{(n+1)}$ using online PCA by formula(4)~(8), and $dim(\mathbf{U}^{(n+1)})=dim(\mathbf{U}^{(n)})+1$, with Step6 followed; if $\|\mathbf{r}\|<\theta_{class}$, perform Step3;

Step3 Calculate the Euclidean distance deuclidean between new sample and all the learned samples;

Step4 If deuclidean_min$>\theta_{distance}$, calculate updated subspace using online PCA by formula(4)~(8), abandon its last vector element to make $dim(\mathbf{U}^{(n+1)})=dim(\mathbf{U}^{(n)})$; if deuclidean_min$<\theta_{distance}$, perform Step5 directly;

Step5 Output the classification result by the corresponding sample class of deuclidean_min;

Step6 Go back to Step1, the next loop of the algorithm is begun with the input of a new sample.

The algorithm flowchart is shown in Fig. 3.

4. Experimental Result and Analysis

4.1 Experiment design

Experiments are designed to perform the proposed online learning and recognition of hand gestures. The preprocess is done to input image sequence in the first place. In order to eliminate the background skin color distraction and low down the computational cost, the hand is located by skin color combined with motion detection. The hand is tracked by CamShift

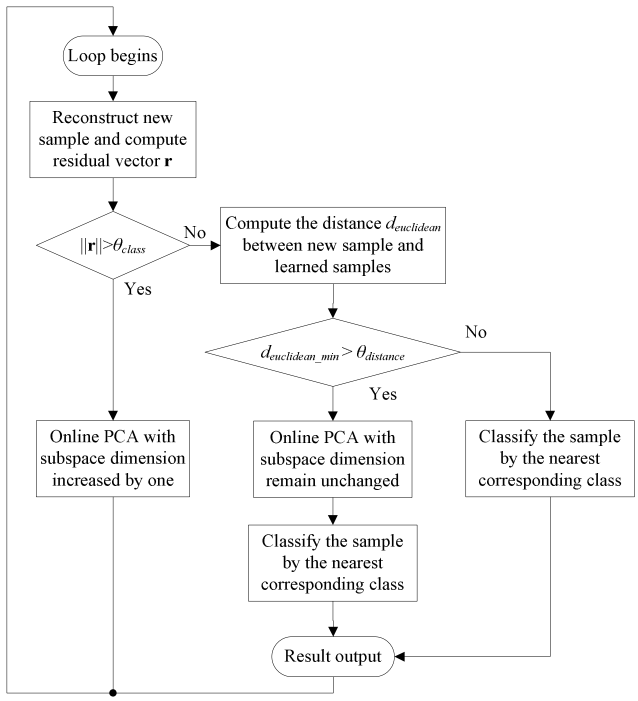

Fig. 3 Flowchart of online PCA with adaptive subspace algorithm

algorithm combined with Kalman filter, which deals with the problem of hand covered by other object and large area skin color distraction. Blur and morphologic process are done to the region that hand is located and tracked. Then the hand is segmented using an adaptive threshold method. The preprocessed hand images are as the sample input for classifier online learning, in which each sample is 50 ×50 pixel in size, namely 2500 in dimension. The Visual C++ 6.0 program environment and OpenCV library is used for developing the online learning program, with a machine of CPU：AMD 3500+ and 1G DDR2 memory.

In the experiments 105 hand gesture samples are used. The classifier is initialized from 20 samples of two classes, with 10 samples in each. The online learning phase is then performed with one sample at a time step after the classifier initialization. The proposed method is compared with off-line PCA in four aspects which are reconstruction error, computation time, recognition rate and dimension scale control.

4.2 Experimental Result and Analysis

The online learning and recognition is begun from the 21st sample. The online learning procedure of samples in one class is shown in figure 4 by steps. The reconstruction comparison with 10 samples in one class is also shown in Fig. 4.

It can be seen in the figure that with the increase of learning step, the reconstruction of proposed method is getting better, the reconstruction error is reduced by steps and the learning result is improved gradually. While the off-line PCA cannot solve the novelty problem, thus when it comes to novel class, samples cannot be reconstructed well, the reconstruction error cannot be reduced and the reconstruction cannot be improved gradually. The sample that actually belonging to novel class is misclassified by similarity measurement among the learned classes, since the actually class is not among the learned classes. In conclusion, the key problem is that the system is divided into two phases of learning and recognition, which cause the novel class unable to be learned meanwhile with recognition and unable to update the knowledge online, as a result, the system perform a bad expansibility, adaptability and robustness.

The experiments of recognition rate, computational time and dimension scale control are analyzed below. We take 105 sample training set and 105 sample testing set for off-line PCA in the experiments, and compare the result with proposed method. In the experiments 9 sample points are choose to observe the changing trends of recognition rate, recognition time and subspace dimension with the running of the algorithm. Specially, for off-line PCA, the computation time is obtained by iterated perform PCA algorithm rather than batch PCA, thus, the two compared methods are both in an incremental fashion, which is more appropriate. The experimental data is shown in TABLE I. Recognition rate, computational time and subspace dimension variety comparison curves are shown in Fig. 5, Fig. 6 and Fig. 7, respectively.

Experimental results show that though the recognition rate is low with small sample set, it is increased gradually along with the increase of learned samples online, and it reaches 90.48% with learning of 105 samples in our experiments. The recognition rate of off-line PCA is 91.43% with 105 samples. Though it is a little higher than the proposed method, the expansibility is bad that repeatedly batch learning is needed in order to increase the rate. Along with the increase of learning samples, the computational time of off-line PCA risen in an exponential fashion, while with the proposed method this time is remain gently rising, which bring a better real-time performance specially for online learning and recognition. Along with the increase of learning samples, the subspace dimension of the proposed method increased slowly and far below the number of samples while the subspace dimension of off-line PCA is 105 with 105 samples, which is an advantage feature for saving computational time and storage.

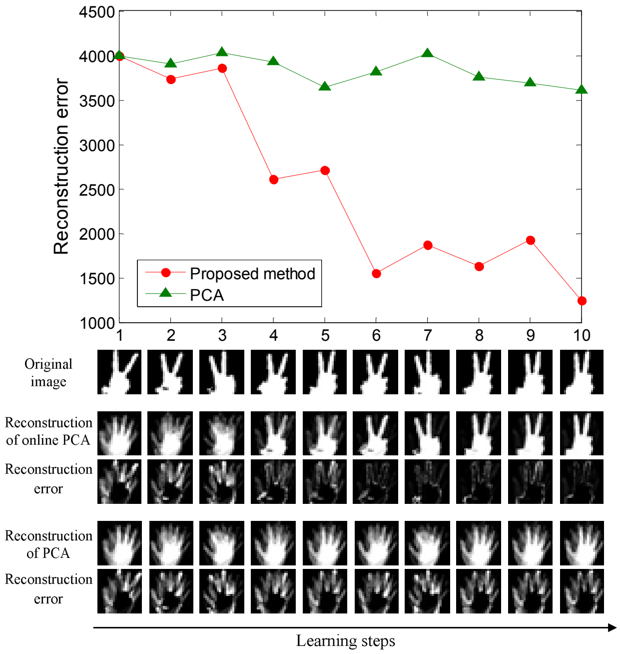

Fig. 4 Analysis of sample reconstruction error

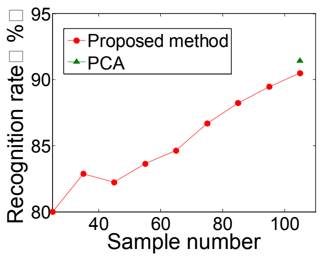

Fig. 5 Comparison of recognition rate

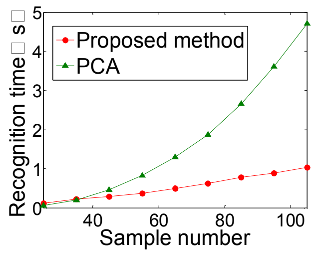

Fig. 6 Comparison of recognition time

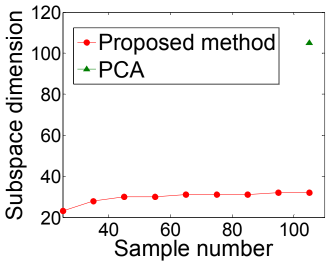

Fig. 7 Comparison of subspace dimension

In addition, there are features that off-line PCA is unable to reach, such as accumulating and updating new knowledge online without affecting learned knowledge, achieving simultaneously online learning and recognition.

4 Conclusion

In this paper, we proposed a method for on-line visual learning and recognition of hand gestures using online PCA algorithm with adaptive subspace. With our approach it is possible to use the same mode for training and learning stage. This is very important for the recognition of the new-added gesture. The subspace updating strategy was chosen automatically corresponding to the three different sample situation measured by similarity between new sample and learned samples, which improved the adaptability. Finally, We performed extensive experimental testing which showed that the method had good adaptability and robustness in the real-time hand gesture knowledge updating, adding and accumulating. In the future, we will add the way of human vision in learning and recognition of hand gestures.

References:

[1] Cristina Manresa, Javier Varona, Ramon Mas and Francisco J. Perales, Real-Time Hand Tracking and Gesture Recognition for Human-Computer Interaction, *Electronic Letters on Computer Vision and Image Analysis*, vol.5, no.3, 2000, pp.96-104.

[2] Hongmo Je, Jiman Kim, and Daijin Kim, Hand Gesture Recognition to Understand Musical Conducting Action, *16th IEEE International Conference on Robot & Human Interactive Communication. Jeju*, Korea, 2007, pp. 163-168.

[3] Yikai Fang, Kongqiao Wang, Jian Cheng, and Hanqing Lu, A Real-Time Hand Gesture Recognition Method, *IEEE International Conference on Multimedia and Expo*, Beijing, China, 2007, pp. 995-998H.

[4] H. S. Yoon, J. Soh, Y. J. Bae, and H. S. Yang, Hand Gesture Recognition Using Combined Features of Location, Angle and Velocity,

Pattern Recognition, vol.34, 2001, pp. 1491-1501.

[5] Nguyen Dang Binh, Enokida Shuichi, Toshiaki Ejima, A New Approach Dedicated to Real-Time Hand Gesture Recognition, 2006.

[6] T. Kirishima, K. Sato, and K. Chihara, Real-Time Gesture Recognition by Learning and Selective Control of Visual Interest Points, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.27, no.3, 2005, pp.351-364.

[7] Ren Haibing, et al, Hand Gesture Segmentation and Recognition with Complex Backgrounds, *Acta Automatica Sinica*, vol.28, no.2, 2002, pp.256-261.

[8]Yang Quan, Wang Min, Hand Gesture Recognition Algorithm Based on Euclidean Distance, *Control and Automation*, 2007, pp.265-266.

[9] Stephan Kirstein, Heiko Wersing, Edgar Korner, A Biologically Motivated Visual Memory Architecture for Online Learning of Objects, *Neural Networks*, vol.21, 2008, pp.65-77.

[10] J. Luo, A. Pronobis, B.Caputo, and P. Jensfelt, Incremental Learning for Place Recognition in Dynamic Environments, *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*. San Diego, CA, USA, 2007, pp.721-728.

[11]Seiichi Ozawa, Shaoning Pang, and Nikola Kasabov, A Modified Incremental Principal Component Analysis for On-line Learning of Feature Space and Classifier, *Trends in Artificial Intelligence*, 2004, vol.231-240.

[12]Xianhua Jiang Yuichi Motai, Learning by Observation of Robotic Tasks using On-line PCA-based Eigen Behavior, *Proceedings of the 2005 IEEE International Symposium on Computational Intelligence in Robotics and Automation*. Espoo, Finland, 2005,vol.391-396.

[13]M. Artac, M. Jogan, and A. Leonardis, Incremental PCA for On-Line Visual Learning And Recognition, *16th International Conference on Pattern Recognition*, Quebec, Canada, vol.3, 2002, pp.781-784.

[14]D. Skocaj and A. Leonardis, Weighted and Robust Incremental Method for Subspace Learning, *Proceedings of the IEEE International Conference on Computer Vision*, Nice, France, 2003, pp.1494-1501.

[15]Peter M. Roth, Michael Donoser, Horst Bischof. On-line Learning of Unknown Hand Held Objects via Tracking, *Proceedings of the Second Int. Cognitive Vision Workshop*. Graz, Austria, 2006.

[16]H. V. Neto and U. Nehmzow, Incremental PCA: An Alternative Approach for Novelty Detection, *TAROS 2005: Towards Autonomous Robotic Systems*, London, UK, 2005, pp.227-233.

[17]Peng Juanchun , Gu Lizhong , Su Jianbo, The Hand Tracking for Humanoid Robot Using Camshift Algorithm and Kalman Filter, *Journal of Shanghai Jiaotong University*, vol.40, no.7, 2006, pp.1161-1165.