2008 IEEE International Conference on
Robotics and Automation
Pasadena, CA, USA, May 19-23, 2008

Natural Hand Posture Recognition Based on Zernike Moments and Hierarchical Classifier

Lizhong Gu and Jianbo Su

Abstract—**View-independence and user-independence are two fundamental requirements for hand posture recognition during natural human-robot interaction. However only a few research concerns on the two issues simultaneously. The difficulty for natural gesture-based human-robot interaction lies in that appearances of the same hand posture vary with different users from different viewing directions. In this paper, we propose a systematic feature selection approach based on Zernike moments and Isomap dimensionality reduction. A hierarchical classifier based on multivariate decision tree and piecewise linearization is developed to deal with the irregular distribution of the same hand postures. The proposed method is compared with other commonly used ones in hand posture recognition. Experimental results indicate that the proposed method can effectively identify different hand postures, irrespective of viewing directions and users.**

I. Introduction

TO coexist with human in the real world, humanoid robot should interact with people naturally and acquire skills from human demonstration [1][2]. As gesture is a natural and intuitive communication channel between human and robot, hand gesture recognition has been studied extensively [3][4][5]. Analysis of hand posture contributes to further dynamic gesture recognition. This work mainly concerns on natural hand posture recognition during human-robot interaction. Many approaches have been proposed for vision-based hand posture recognition. Those approaches can be divided into two main categories: 3D hand model-based methods and appearance-based approaches [6]. Approaches based on 3D model obtain hand posture features by recovering parameters of hand joint angles and palm positions from images. Since the accurate 3D information of hand with multi-degrees of freedom cannot be estimated accurately, colored markers or gloves are often used for simplifying image processing, which makes human-robot interaction not so natural and direct. Instead, appearance-based approaches could extract hand posture features directly from images and are computationally efficient. However, appearance of hand posture contains insufficient source information due to projective transformation from 3D space to 2D image plane, which essentially has weak discrimination compared with the 3D model methods. Hence, the set of recognized hand postures is relatively small. As for the command gestures during human-robot interaction, there are not many gestures to be recognized. Therefore the appearance-based approach is adopted here for hand posture recognition.

A general humanoid robot head system has been developed to study multi-modal human-robot interaction. We want to build a natural gesture-based interface between human and robot, which is the starting point of our research. Most of current researches focus on the recognition of frontal view hand gesture. However, a practical human-robot interaction system should be:

--User-independent. Although there are more or less differences among users' hand shape and finger length and thickness, the robot should recognize the same posture made by different users.

--View-independent. For a natural interaction between human and robot, user should not be constrained to make posture in fixed places. The robot should be able to recognize hand posture from different viewing directions.

There are two ways to address the above two issues: the first is to find invariant features for hand posture; the other is to find a more robust and efficient classifier.

State-of-the-art feature extraction methods select a few invariant moments or Fourier coefficients as hand posture features. Few researches consider features representation capability and what features are equivalent to characterize hand postures. There is lack of a systematic analysis on feature selection for hand postures issued by various users from different viewing directions. Different kinds of hand postures might cause various recognition rates since incomplete features set are utilized to represent hand postures. Chalechale et al. studies how to select appropriate features and hand postures, which have relatively high recognition rates based on seven hu moments [5]. However, considering user preference for hand postures in natural human-robot interaction, various kinds of postures should be recognized. We propose a systematic approach to feature selection based on Zernike moments and Isomap dimensionality reduction. The features are determined according to the reconstruction effect of hand postures. Therefore a hand posture is fully characterized by a set of Zernike moments, which form a high dimensional feature vector. As for the perspective projection, appearances of hand posture are nonlinear distorted at various viewpoints. More over, the length and thickness of fingers and palm of various users are totally different. Non-linear structure lies among samples of natural hand postures. Thus Isomap [9] is utilized to discover meaningful

Manuscript received September 14, 2007. This work was supported by National Natural Science Foundation of China under Grant (60428303).

Lizhong Gu and Jianbo Su are with the Dept. of Automation, Shanghai Jiaotong University, Shanghai, 200240, China. (phone: 86-21-34204274; e-mail: {gulizhong, jbsu}@sjtu.edu.cn).

978-1-4244-1647-9/08/$25.00 ©2008 IEEE.

low-dimensional structures hidden in the high-dimensional feature space.

Most existing classification methods for hand posture belong to distance-based classifier [4][5][10][14]. The distance-based method calculates the distances between the input hand posture and the templates according to a certain criteria. The template that has the smallest distance to the input hand posture is chosen as the recognized result. However, these hand postures from different users and viewing directions cannot be simply classified by observing their distances. Samples of hand posture might be irregularly distributed in the feature space and distance-based method is not suitable for this case. Thus a complex discriminating boundary is needed to discriminate those hand postures. Multivariate decision tree (MDT) is taken advantage of to classify those hand postures through hierarchical discriminants. The MDT splits samples using a linear combination of relevant attributes at non-leaf nodes and is of a hierarchical classifier. In some cases, MDT has large fragments and needs further splitting at the child node. Generally speaking, the decision tree with smaller size tends to generalize better for the unseen samples [18]. It is better to divide hand posture clusters into two distinct subsets that have no intersections through piecewise linear discriminants. Thus the tree size will be reduced. Therefore multivariate piecewise linear decision tree (MPLDT) is proposed to discriminate hand postures. Hence, multivariate decision tree is utilized to find the most relevant features that are helpful to classification, while piecewise linear discriminants are used to divide a set into two distinct ones at the non-terminal tree node to reduce tree size and improve classification accuracy.

The remainder of this paper is organized as follows. Section II reviews some related works on hand posture recognition. Section III systematically analyzes feature representation for hand posture. Section IV proposes multivariate piecewise linear decision tree for hand posture recognition. Experiment results are reported in section V, followed by conclusions in section VI.

II. Related Works

A limited number of researches concern on either view-independent or user-independent hand posture recognition. These researches focus on seeking for invariant features to represent hand posture. There are two means to overcome the influence of viewing direction on hand posture appearance. The first way is to adopt two or more cameras to reconstruct the 3-D scene of the hand and obtain view-invariant hand posture representation. By modeling the user's arm as a 3D line and using arm orientation information to unwarp hand posture [12], the frontal view of hand posture can be obtained. This approach implies that the major axis of hand is aligned with that of arm. However, this prerequisite is not always met in practice since arm orientation is not equivalent to that of hand. Three cameras are mounted on the ceiling in the intelligent sweet home environment [13]. A 2-layered hand-posture database structure representing both 2D appearance features and 3D features is proposed to recognize view-invariant hand posture. Yet a data glove is needed to get the users' finger angles. Another way attempts to find invariant features for hand posture appearances from mono camera, such as invariant moments and Curvature Scale Space (CSS). Seven hu's invariant moments are used to represent hand posture [5]. Hu's moments are invariant under translation, scaling and in-plane rotation when recognizing the same object. This method is not suitable for those postures issued by various users from different viewing directions. CSS images are proposed to represent the shapes of contours of hand postures [14]. Nearest neighbor techniques are adopted to match the input CSS features with the stored templates. This approach relies upon fine contour extraction and is computationally complex.

A person-independent hand posture classification system is presented in [15]. Compound jets are introduced as multiple feature types and the elastic bunch graph matching method is employed to model variances in object appearance and complex backgrounds. This method is view-dependent.

III. Feature Extraction

Since appearances of the same hand posture vary due to different users from different viewing directions, it is ideal to be able to extract invariant features to represent those postures so that the same posture could be clustered as compact as possible in the feature space. There are several feature representations for hand posture in current appearance-based approaches. Two widely used invariant features are Fourier descriptors [4] and region moments [5]. Fourier coefficients are easily affected by local shape variation. Here, Zernike moments are used to represent hand postures for its computational efficiency and global feature description ability.

A. Zernike Moments

For a digital image f, the Zernike moment [16] of order n with repetition m is defined as follows:

$$A_{nm} = \frac{n+1}{\pi}\sum_{x}\sum_{y} f(x,y)V_{nm}^{*}(\rho,\theta), \qquad x^2+y^2 \le 1. \tag{1}$$

where $n-|m| = even, \quad |m| \le n$. The symbol * denotes the complex conjugate. The image pixel is defined over the polar coordinate space inside a unit circle. The radius and angle on a certain pixel are calculated as follows:

$$\begin{aligned} &\rho = \sqrt{x^2+y^2} \\ &\theta = \mathrm{atan}(y/x) \end{aligned} \quad . \tag{2}$$

The Zernike polynomials $V_{nm}(x,y)$ form a complete orthogonal set over the interior of the unit circle.

$$V_{nm}(x,y) = V_{nm}(\rho\cos\theta, \rho\sin\theta) = R_{nm}(\rho)e^{im\theta}. \tag{3}$$

The real-valued radial polynomials $\{R_{nm}(\rho)\}$ are defined as

$$R_{nm}(\rho)=\sum_{s=0}^{(n-|m|)/2}\frac{(-1)^s(n-s)!}{s!(\frac{n+|m|}{2}-s)!\left(\frac{n-|m|}{2}-s\right)!}\rho^{n-2s}. \quad (4)$$

To compute the Zernike moments of a given image, image coordinates are mapped to the polar coordinate space inside a unit circle and the center of the image is taken as the origin. The magnitude of Zernike moments is rotation invariant. To achieve translation invariance, the object gravity is translated to the image center. As for scale invariance, the computed Zernike moments is normalized [7] using m_{00}, the mass of the object.

$$A'_{nm}=A_{nm}/m_{00}. \quad (5)$$

The following problem is to determine the highest order n of moments that are sufficient to characterize and represent a hand posture. Let $\hat{f}_n$ denote the reconstructed image using moments of order 0 through n. The maximum needed order is measured based on the difference between $\hat{f}_n$ and the original image f [16]. Thus the source feature set consists of Zernike moments of order 0 through n are extracted from the original image f.

$$\begin{aligned}\hat{f}_n(x,y)&=\sum_n\sum_{m<0}A_{nm}V_{nm}(\rho,\theta)+\sum_n\sum_{m\geq 0}A_{nm}V_{nm}(\rho,\theta)\\&=\sum_n\sum_{m>0}A_{n,-m}V_{n,-m}(\rho,\theta)+\sum_n\sum_{m\geq 0}A_{nm}V_{nm}(\rho,\theta)\\&=\sum_n\sum_{m>0}A^*_{nm}V^*_{nm}(\rho,\theta)+\sum_n\sum_{m\geq 0}A_{nm}V_{nm}(\rho,\theta)\end{aligned} \quad . \quad (6)$$

B. Dimensionality Reduction

To efficiently classify hand postures, dimensionality reduction must be performed on the original high-dimensional feature set. Moreover, real-world data lie on low-dimensional manifolds [9]. Based on the effect of reconstructed images of hand postures, the needed moment order is up to 24. The number of complete Zernike moments of order 24 is 169. It is not suitable to classify hand posture samples in such a high-dimensional feature space.

Here supervised Isomap (S-Isomap) [17] is used to produce well-separated classes in a low-dimensional feature space using class labels of samples. The steps of S-Isomap are similar to those of Isomap. The difference is in the first step, where class information is utilized to compute distances between all pairs of points. Thus S-Isomap achieves better classification performance than Isomap. Let x_i denotes original high-dimensional feature of hand posture sample. The algorithm outputs coordinate vectors y_i in a low-dimensional feature space that best preserve the intrinsic geometry of the samples. Thus a mapping function can be learned by a generalized regression network [11] using the corresponding data pairs x_i and y_i. Hand posture classifier is to be trained on the resultant low-dimensional feature space that is a representative of the original high-dimensional feature space. When a new sample to be recognized is entered, it should first be transformed into the reduced feature space using the mapping function.

IV. MULTIVARIATE PIECEWISE LINEAR DECISION TREE

A practical hand posture recognition system should discriminate arbitrary natural hand postures. Although representative features for hand postures have been extracted, samples of a certain hand posture are normally not clustered like a hyper-sphere since they are irregularly distributed in the feature space. Consequently, these scattered samples cannot simply be classified using distance-based methods. Fig. 1 shows that a sample to be recognized is misclassified into pattern 2 by comparing distances between the unknown sample and the center of a certain pattern class. In fact, it belongs to pattern class 1. For this case, there might exist a complex discriminant boundary between two different hand posture clusters in the high-dimensional feature space, which cannot be divided into two distinct subsets by a single hyperplane. We consider using a hierarchical classifier to classify hand postures made from different viewing directions by different users, which can be classified by multiple discriminants.

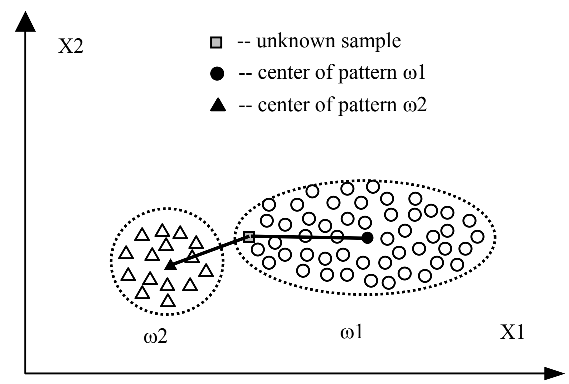

Fig. 1. A 2D illustration of sample misclassification by comparing distances between the unknown sample and the center of a certain pattern class.

Decision tree is chosen as the hand posture classifier for it is essentially a sort of hierarchical classifiers and can save much classification time. There are two types of decision trees, univariate decision tree and multivariate decision tree. Multivariate decision tree separates samples of two classes using multivariate linear tests at non-leaf nodes. Some search methods, such as SBE, SFS and Tabu, are performed to find appropriate feature combinations with respect to some partition-metric criterion [18]. Compared with univariate decision tree, multivariate decision tree has smaller size of tree and higher accuracy.

Although it splits instances using more than one attribute at non-leaf nodes, multivariate decision tree still has large fragments at some non-leaf nodes. Thus it needs subsequent multivariate splitting to partition those fragments at the nodes of deeper layer (Fig. 2). This might lead to larger tree. Statistically, the decision tree with smaller size has more classification accuracy. Thus multivariate piecewise linear decision tree (MPLDT) is proposed to classify hand postures. We hope to use piecewise linear hyper-planes to separate instances into two distinct subsets, which are not interlaced (Fig. 3). Thus the tree size is reduced and the classification accuracy can be improved.

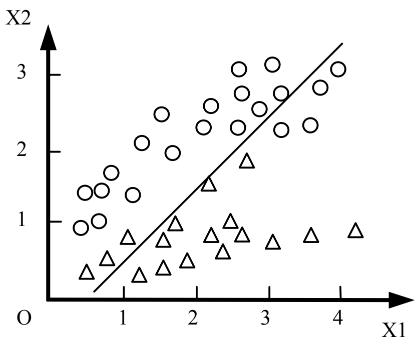

Fig. 2. A 2D illustration of fragments caused by oblique splitting at a certain node. Further child nodes are needed to split those fragments.

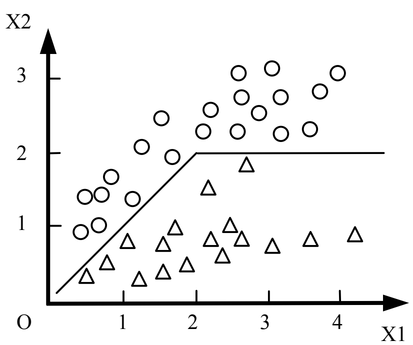

Fig. 3. A 2D illustration of multivariate piecewise linear discrimination. The two clusters are separated into two distinct subsets by multiple discriminants at a certain node.

During the tree construction, piecewise linear discriminants are incorporated into the non-terminal tree node instead of a single oblique hyper-plane. First we group the multi-classes dataset into two super-classes using the similar procedure described in [19]. Each super-class is a combination of several classes and is regarded as one group. Considering various features playing different roles in classification, we do not use all features to construct piecewise linear discriminants. Thus in feature selection stage Tabu search is employed to find most relevant feature combinations since it can overcome local optimality [18]. In the third stage, the piecewise linear discriminants are created in order to divide two groups into two non-intersecting subsets. The piecewise linear boundary is created as perpendicular bisectors of line segments linking centers of two classes [19]. Thus multiple discriminants are obtained and organized as conjunctive and disjunctive discriminants. These linear discriminants are used to partition hand posture clusters into two distinct subsets. To avoid overfitting on the training dataset, the number of obtained discriminants has an upper limit. When the number of remaining samples is less than five times of currently used features, no discriminants will be created. It should be pointed out that piecewise linear classifier (PLC) itself is also of a hierarchical classifier and the structure of PLC is similar to that of decision tree. The difference between PLC and MPLDT is that the former considers all attributes to construct decision tree, whereas the latter uses appropriate combinations of attributes to do so. At a certain non-leaf node, MPLDT finds the most relevant features and then induces conjunctive and disjunctive discriminants from those discriminants obtained by PLC.

The proposed algorithm can be summarized as follows:

1) Grouping the training pattern sets into two super-classes.

2) Finding appropriate feature combinations using the Tabu search technique.

3) In the most relevant feature space, dividing the above two super-classes into two distinct subsets using piecewise linear discriminants.

4) In each subset recursively repeating above procedures until the subset consists of samples from a single class.

V. EXPERIMENTS

To test the effectiveness and practical performance of the proposed approach, we conduct experiments on a general humanoid robot head. The robot head responds to user's posture with its motors conducting pan or tilt movements. Experimental results are compared with other commonly used methods in hand posture recognition systems [4][5], which achieve high recognition rates. These methods adopt different features and classifiers, such as hu moments+Gaussian mixture models (GMM), and Fourier descriptors+maximum likelihood (ML), denoted by GMM and ML respectively. Also Oblique Classifier 1 (OC1) and piecewise linear classifier (PLC) are trained on the collected dataset, denoted by OC1 and PLC. The proposed approach based on Zernike moments and MPLDT is denoted by MPLDT. GMM and ML are of distance-based methods. OC1 is a multivariate decision tree induction system and is freely available on the Internet [20]. OC1, PLC and MPLDT are all known as hierarchical classifiers.

A. Types of Postures

Eleven kinds of hand postures are shown in Fig. 4, where each row shows several samples issued by different users from different viewing directions. The hand posture samples are captured from five persons. All the hand postures are made by right hand.

The hand posture dataset is generated through the following on-line video means. The user's hand region is tracked in a bounding box with the tracking algorithm [21]. The hand region is the largest area in the bounding box. Thus the hand silhouette is obtained through filling the area, which is enclosed by a max contour extracted in the tracked bounding box. When the extracted hand silhouette looks perfect, the silhouette image is saved in the dataset. There are total 3850 hand posture samples in the dataset.

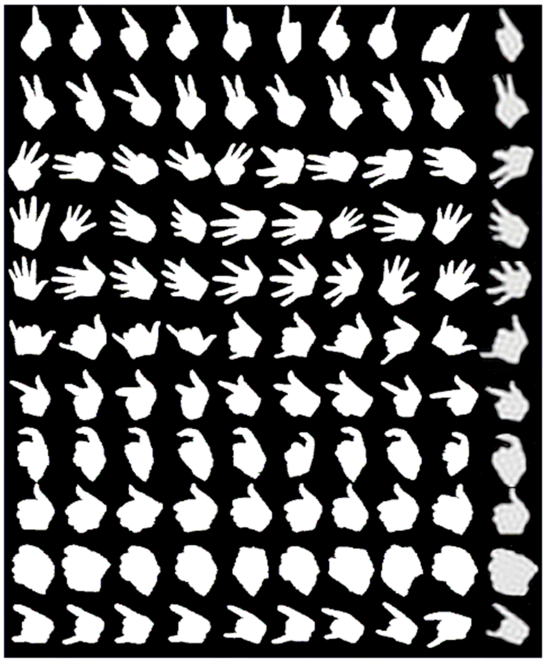

Fig. 4. Samples of eleven hand postures. Each row shows several examples for a posture template. Last column is reconstructed images based on a finite set of Zernike moments.

B. Determination of Representative Features

Before computing Zernike moments, the gravity of hand region is translated to the center of silhouette image. Then the image is resized to a fixed size 61*61 for saving computation time. The Zernike moments are computed over a unit disk, where the hand silhouette is bounded by an appropriate boundary circle. A fast approach is adopted to compute Zernike moments [8]. According to the reconstruction effect of hand postures, the maximum order of Zernike moments is 24. The last column in Fig. 4 shows examples of eleven reconstructed hand postures. There are about 167 Zernike moments in the original feature set for the extracted features start from the second order. The average computation time of those moments is about 50ms on Intel Pentium 4 processor 2.8GHz and 512MB memory.

As for the classification task, pre-processing steps must be taken to discover inherent low-dimensional embedding from the original high-dimensional feature space. Here S-Isomap is adopted before hand posture classification. The inherent dimensionality of the hand posture dataset can be estimated by the elbow point, where the decreasing rate becomes slowly (Fig. 5). The number of features in the resultant low-dimensional feature space is chosen as 7, where residual variance is slightly larger than 0.1.

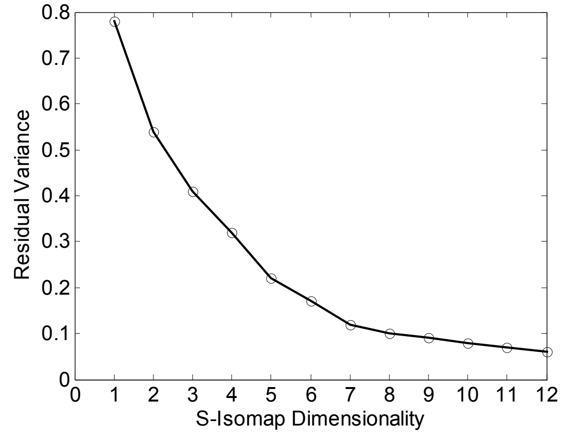

Fig. 5. Residual variance of S-Isomap on the hand posture dataset.

C. Experiment for Natural Hand Posture Recognition

Although the user's hand region can be projected onto the fovea of robot eye through gaze control [22], the inclination between the line of sight and the normal direction of hand palm might be large during nature human-robot interaction. To compare the performance of five methods, we divide viewing directions into seven regions. The frontal view direction is countered as angle zero and left deflection as positive angles. The central region is from -10 to +10 degrees. The angle intervals of left side are in turn from 10 to 30, from 30 to 50 and from 50 to 70 degrees, respectively. The right side is symmetrical to the left side. Each region is denoted by the middle angle of that region. Postures made in a certain region are considered as from the same place.

Average recognition rates of eleven hand postures under different viewing directions by different users are plotted in Fig. 6. We find that hierarchical classifiers are much better than other two distance-based classifiers in terms of classification accuracy. In the neighborhood of frontal view five methods can recognize hand postures with nearly equal recognition rates. However, farther from the frontal view, the distortion of hand posture appearance is larger due to the nonlinear projection of hand shape. The variation of appearance is not of the same scale. Thus the recognition rates with GMM and ML decline more than those with hierarchical classifiers. The reason is that features adopted in GMM and ML is incomplete and not sufficient to characterize such eleven hand postures. Another reason is that GMM and ML essentially belong to distance-based method, which is not suitable to discriminate these irregularly scattered samples of hand postures.

Among the three hierarchical classifiers, MPLDT outperforms the OC1 and PLC. MPLDT has smaller tree size than OC1 since it uses piecewise linear discriminants to split instances at the non-leaf nodes. The discriminants of MPLDT are different from those of OC1. As for PLC, it uses all attributes to split instances in the high-dimensional feature space, whose classification performance is theoretically equal to that of MPLDT. However, PLC does not perform as well as MPLDT since real hand posture samples are more or less polluted by noise and more redundant attributes on the contrary bring in more noise. Only those relevant attributes contributing to classification should be used to split instances at a certain node.

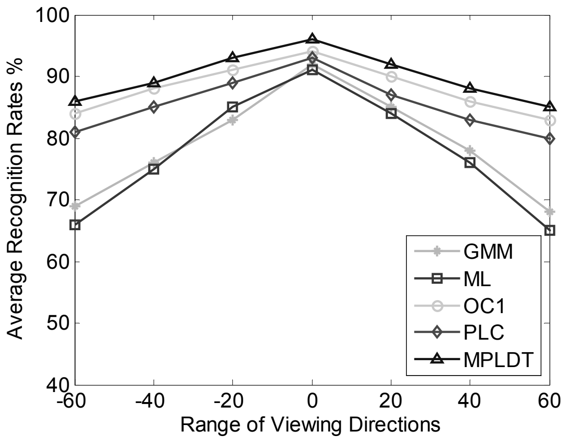

Fig. 6. Average recognition rates for different viewing directions by different users with the eleven hand postures.

VI. CONCLUSIONS

This paper deals with two elementary requirements arising from natural human-robot interaction, that is view-independent and user-independent. A systematically approach is proposed to analyze feature extraction. Representative features, which are capable of characterizing arbitrary hand postures, are selected based on Zernike moments and Isomap dimensionality reduction. A hierarchical classifier, multivariate piecewise linear decision tree, is proposed to deal with the irregular distribution of hand posture samples issued by different users and observed from different directions. Experimental results show the proposed method is insensitive to large variation in viewing direction and users' hand shape. The proposed hierarchical classifier can also be applied to a general classification task.

REFERENCES

[1] R. A. Brooks, C. Breazeal, M. Marjanovic, B. Scassellati, and M. M. Williamson, "The Cog Project: Building a Humanoid Robot," *Computational for Metaphors, Analogy and Agents*, vol. 1562, pp. 52-87, 1999.

[2] N. Mitsunaga, T. Miyashita, H. Ishiguro, K. Kogure, and N. Hagita, "Robovie-IV: A Communication Robot Interacting with People Daily in an Office," *IEEE Int. Conf. Intelligent Robots and Systems*, pp. 5066-5072, 2006.

[3] M. Stark, M. Kohler, and P. G. ZYKLOP, "Video based gesture recognition for human computer interaction," In: *Modeling - Virtual Worlds - Distributed Graphics*, pp. 1-19, 1995.

[4] A Licsár and T. Szirányi, "Hand Gesture Recognition in Camera-Projector System," *European Conf. Computer Vision*, vol. 3058, pp. 83-93, 2004.

[5] A. Chalechale, F. Safaei, G. Naghdy, and P. Premaratne, "Hand posture analysis for visual-based human-machine interface," in *Proc. the Workshop on Digital Image Computing*, pp. 91-96, 2005.

[6] V. Pavlovic, R. Sharma, and T. S. Huang, "Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 677-695, July 1997.

[7] B. Ye and J. X. Peng, "Invariance analysis of improved Zernike moments," *Journal of Optics A: Pure and Applied Optics*, vol. 4, pp. 606-614, 2002.

[8] S. K. Hwang and W. Y. Kim, "A novel approach to the fast computation of Zernike moments," *Pattern Recognition*, vol. 39, issue. 11, pp. 2065-2076, 2006.

[9] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensional reduction," *Science*, vol. 290, pp. 2319-2323, 2000.

[10] R. Sanchez-Reillo, C. Sanchez-Avila, and A. Gonzalez-Marcos, "Biometric identification through hand geometry measurements," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1168–1171, 2000.

[11] D. F. Specht, "A General Regression Neural Network," *IEEE Trans. Neural Networks*, vol. 2, no. 6, pp. 568-576, 1991.

[12] R. Grzeszczuk, G. Bradski, M. H. Chu, and J. Y. Bouguet, "Stereo Based Gesture Recognition Invariant to 3D Pose and Lighting," *IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, pp. 826-833, Jun. 2000.

[13] H. Jang et al., "View-invariant Hand-posture Recognition Method for Soft-Remocon-system," *IEEE Int. Conf. Intelligent Robots and Systems*, vol. 1, pp. 295-300, 2004.

[14] C. C. Chang and C. Liu, "Modified curvature scale space feature alignment approach for hand posture recognition," *IEEE Int. Conf. Image Processing*, vol. 3, pp. 309 –312, 2003.

[15] J. Triesch and C. von der Malsburg, "A System for Person-Independent Hand Posture Recognition against Complex Backgrounds," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, issue. 12, pp. 1449-1453, Dec. 2001.

[16] A. Khotanzad and Y. H. Hong, "Invariant image recognition by Zernike moments," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 5, pp. 489-497, 1990.

[17] X. Geng, D. C. Zhan, and Z. H. Zhou, "Supervised nonlinear dimensionality reduction for visualization and classification," *IEEE Trans. Systems, Man, and Cybernetics-Part B: Cybernetics*, vol. 35, no. 6, pp. 1098-1107, 2005.

[18] X. B. Li, "Multivariate Decision Trees Using Linear Discriminants and Tabu Search," *IEEE Trans. Systems, Man and Cybernetics*, vol. 33, issue. 2, pp. 194-205, 2003.

[19] A. Kostin, "A simple and fast multi-class piecewise linear pattern classifier," *Pattern Recognition*, vol. 39, pp. 1949-1962, 2006.

[20] S. K. Murthy, S. Kasif, and S. Salzberg, "A system for induction of oblique decision trees," *Journal of Artificial Intelligence Research*, vol. 2, pp. 1–32, 1994.

[21] J. C. Peng, L. Z. Gu, and J. B. Su, "Hand tracking for humanoid robot using Camshift algorithm and Kalman filter," *Journal of Shanghai Jiaotong Uni*versity, vol. 40, pp. 1161-1165, 2006.

[22] L. Z. Gu and J. B. Su, "Gaze Control on Humanoid Robot Head," *IEEE 6th World Congress on Intelligent Control and Automation*, vol. 11, pp. 9144-9148, 2006.