

Study on 3D Object Recognition Based on KPCA-SVM

YaJuan WU, XingMing WANG and FuHua SHANG

School of Computer & Information Technology, Northeast Petroleum University, DaQing, China

Email: wyjjsj@163.com, xmwang_dqpi@163.com, shangfh@163.com

Abstract. Only these features which characterize the difference of similar views have value to recognize in 3D object recognition system based on view. And Principal Component Analysis (PCA) and Kernel PCA (KPCA) can extract features and reduce dimensionality. In this contribution, we use the PCA and KPCA to extract features firstly, and then classify the 3D objects with SVM (Support Vector Machine). In this experiment, we select the Columbia Object Image Library (COIL-100) and compare the results of KPCA-SVM with PCA-SVM and SVM. The results clearly demonstrate that the feature extraction can greatly reduce the dimensionality of feature space without degrading the classifiers' performance. Among these methods, the best performance is achieved by SVM with KPCA.

Keywords: 3D object recognition; SVM; KPCA; PCA

1. Introduction

The 3D object recognition is an important area in Computer Vision and Pattern Recognition and mainly includes two key steps: Object Detection and Object Recognition. Object detection is used to determine whether there potentially exist objects in one scene and its location. Object Recognition decided to object ID and its position [1].

The algorithms of 3D object recognition based on views or appearances, recently become a hot research topic [2-3]. Although the view-based method is difficult to overcome the overlap of objects, but if the geometrical models are very difficult or impossible to obtain, then we can recognize the objects with this method [4].

In recognition system based on view, the dimensions of the extracted features may be several hundreds [5-6]. After obtaining the kinds of features of 3D object from 2D images, the 3D object recognition is a classification problem according features from the perspective of pattern recognition. Cover theorem points out that one complex pattern recognition classification problem will be more easily classified linearly in high dimensional space than lower dimensional space [7]. Therefore, SVM firstly maps the low dimensional data of the training samples into the higher dimensional feature space, and then constructs a Classification Hyperplane (CH) in the higher dimensional feature space to realize the classification function.

Through reducing the feature dimensions properly, on the one hand, we can eliminate redundancy, speed up the processing and improve efficiency. On the other hand, we can reduce the complexity of classifiers and the classification error rate. To improve the performance of the classifier, feature extraction is another effective way which follows the feature selection, and can obtain new low dimensional features through transforming the inputting higher features. PCA [8] is a familiar feature extraction method. Through calculating the eigenvector of sample covariance matrix, PCA maps the input space into lower dimensional space linearly. That is to say it uses less comprehensive index instead of primary more index, and the new features are not related. PCA [9] has many nonlinear forms, and the kernel PCA is one of them.

To reduce the dimensionality of feature space selected in 3D object recognition system based on view, this contribution will explore the possibility that we use the nonlinear feature extraction method and combine

KPCA feature extraction with SVM classifier organically. Feature extraction maps the input higher dimensional space into a new lower dimensional feature space, and then the input will be classified by SVM. Experiment shows that the feature extraction can reduce the dimensionality of feature space and improve system identification effectively without reducing the performance of classifiers.

2. SVM

Support Vector Machine (SVM) [10-11] is based on statistics theory, and has strict theory and mathematical foundation, and is similar to the Structural Risk Minimization principle (SRM). Through training, the classifier has very good generalization ability with it. And it has shown many special advantages in solving nonlinear, small samples and higher dimensional pattern recognition problems.

The mechanism of support vector machine can be simply described as: looking for the required split plane and make the points of the training set away from the split plane it can. That is to say, we will look for a split plane and make its blank areas on both sides largest. The basic idea of SVM is that it maps the input vector X which is prior chosen into a higher dimensional feature space Z , and constructs an optimization classification hyperplanes in this space.

Vapnik [12] summed up Support Vector Machine into a quadratic equation problem. So, the solution of the optimal hyperplane becomes a quadratic programming problem:

$$\begin{aligned} \text{Minimize } \Phi(w, b) &= \frac{1}{2} \|w\|^2 \\ \text{s.t. } y_i (w \cdot x_i + b) - 1 &\geq 0 \quad i = 1, 2, \dots, l \end{aligned} \quad (2.1)$$

The optimal classification plane is discussed in the premise of the linear. In linear inseparable case, i.e. some training samples can't meet the type 2.1) of restriction. We introduce a slack variable $\xi_i \geq 0$ and a penalty factor C . Then, the generalized optimal classification plane (separating the sample furthest and making the classification interval maximum) can be expressed as a quadratic programming problem.

$$\begin{aligned} \text{Minimize } \Phi(w, b) &= \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i \\ \text{s.t. } y_i (w \cdot x_i + b) - 1 + \xi_i &\geq 0 \quad i = 1, 2, \dots, l \end{aligned} \quad (2.2)$$

Therefore, C is a constant and control the penalty degree of the wrong separation of the samples. And it realizes the compromise of the rate of wrong samples and the complexity of the algorithms. ξ_i is a slack variable and measure the degree of violating constraint in inseparable case.

For nonlinear problems, according to the way of thinking of the generalized linear discriminant, we can transform them into linear problems in another space (also called feature space) and solve the optimal classification plane there. The transformed feature space always has higher dimension, sometimes even infinite dimension. So, this case will produce some difficult in processing and cause the dimensional disaster problem. Support vector machine can avoid solving in the higher dimension space through a clever way.

Supposed that a nonlinear mapping: $\Phi: R^d \rightarrow H$ maps the samples of input space into a higher dimensional feature space H . When constructing an optimal plane in this feature space H , the training algorithm only used the dot product in space, i.e. $\Phi(x_i) \cdot \Phi(x_j)$, and no single appear. Therefore, if we can find K which can make $K(x_i, x_j)$ equals to $\Phi(x_i) \cdot \Phi(x_j)$, and then the inner product operations in higher dimension space could be realized with the functions in original space. So, we don't even need to know the forms of transformation Φ .

Instead of the dot product of optimal classification plane, we use the inner product to transform a problem of original space into another higher dimension feature space. At present, the corresponding discriminant function could be transformed into

$$\begin{aligned}
f(x) &= \text{sgn}\{(w \cdot x) + b\} \\
&= \text{sgn}\left\{\sum_{i=1}^m a_i^* y_i K(x_i, x)\right\} + b^*
\end{aligned} \tag{2.3}$$

There, the corresponding sample x_i is called Support Vector (SV) when $a_i^* \neq 0$. $K(x, x')$ is called as Kernel function. With different Kernel function, we will get different forms of the classification plane in the sample space.

There are three kinds of kernel functions:

- 1) Polynomial kernel: $K(x, x_i) = [(x \cdot x_i) + 1]^q$.
- 2) Radial Basis Function: $K(x, x_i) = \exp\left(-\frac{|x - x_i|}{2\sigma^2}\right)$.
- 3) Sigmoid kernel: $K(x, x_i) = \tanh(v(x \cdot x_i) + c)$.

Having good form, the RBF kernel already shows good performance in the actual problems. In this experiment, we will select RBF, i.e. $K(x, x_i) = \exp\left(-\frac{|x - x_i|}{2\sigma^2}\right)$, σ is kernel parameter and its value will influence the performance of SVM. So, the solution of σ will become an import part of this experiment.

3. KPCA

Principal Component Analysis (PCA) is put forward by Karl Pearson in 1901[13]. As an analysis method based on statistical optimal principle, PCA operation is a feature transformation which uses the statistical properties of the data set. Subject to no or few loss data, this transform can reduce the dimensionality of data set and reveal the simple structure hidden in complex data.

Kernel PCA (The Kernel principal component analysis, KPCA) is an expansion of PCA with Kernel. In a lot of practical applications [14-15], the performance of KPCA has already been proved better than linear PCA. The basic idea of KPCA is that: the input X is mapped into feature space $V \in F \setminus \{0\}$, $Cv = \lambda v$ through nonlinear function, and then we execute a linear PCA in F to do the results.

Supposed that the mapped data are centered, i.e.

$$\sum_{\mu=1}^M \Phi(x_\mu) = 0 \tag{3.1}$$

M is the number of input data. And then the covariance matrix of feature space is:

$$C = \frac{1}{M} \sum_{\mu=1}^M \Phi(x_\mu) \Phi(x_\mu)^T \tag{3.2}$$

For this, we should solve $Cv = \lambda v$ for eigenvalues $\lambda \geq 0$ and eigenvectors $V \in F \setminus \{0\}$.

$$(\Phi(x_v) \cdot Cv) = \lambda (\Phi(x_v) \cdot v)$$

Allowing for that all eigenvectors can be expressed as the span of $\Phi(x_1), \Phi(x_2), \dots, \Phi(x_M)$. Then:

$$\begin{aligned}
&\frac{1}{M} \sum_{\mu=1}^M \alpha_\mu \left(\sum_{w=1}^M (\Phi(x_v) \cdot \Phi(x_w)) \Phi(x_w) \Phi(x_\mu) \right) \\
&= \lambda \sum_{\mu=1}^M (\Phi(x_v) \cdot \Phi(x_\mu))
\end{aligned} \tag{3.3}$$

$v = 1, 2, \dots, M$. We define a $M \times M$ dimension matrix K .

$$K_{\mu\nu} := (\Phi(x_\mu) \cdot \Phi(x_\nu)) \tag{3.4}$$

So, the 3.3) could be simplified as

$$M\lambda\alpha = K\alpha \quad (3.5)$$

We solve the eigenvalues and eigenvectors of 3.5). In F , we should standard a required non-zero vector α^k , i.e. $(v^k \cdot v^k) = 1$. So, we can get

$$1 = \sum_{i,j=1}^l \alpha_i^k \alpha_j^k (\Phi(x_i) \cdot \Phi(x_j)) = (\alpha^k \cdot K\alpha^k) = \lambda_k (\alpha^k \cdot \alpha^k)$$

In order to extract the principal components, the mapping between x and $\Phi(x)$ which should be calculated through 3.5) must meet the eigenvector v^k of F .

$$(v^k \cdot \Phi(x)) = \sum_{i=1}^M (\alpha_i^k) K(x_i, x) \quad (3.6)$$

When 3.1) is false, we should correct it

$$\Phi(x_\mu) \rightarrow \Phi(x_\mu) - \frac{1}{M} \sum_{v=1}^M \Phi(x_v) \quad \mu = 1, 2, \dots, M$$

So, the kernel matrix could be revised as

$$K_{\mu\nu} \rightarrow K_{\mu\nu} - \frac{1}{M} \left(\sum_{\omega=1}^M K_{\mu\omega} + \sum_{\omega=1}^M K_{\omega\nu} \right) + \frac{1}{M^2} \sum_{\omega,\tau=1}^M K_{\omega\tau}$$

4. Experimental Results

In this experiment, the graphics library chooses the Columbia object graph library (COIL 100) which contains pictures of 100 every day objects [16]. It is a 3D-object graphics library which is now widely applied in estimating the 3D object recognition algorithms. The images were obtained by placing the objects on a turntable and taking a view every 5° . Through this way, we will get 72 images of each object and 7200 images of 100 objects. The size of the color images was 128×128 . The 0° pose angle views of the images are shown in Figure. 1.

In order to reduce the complexity of the operation, we transform the 128×128 into 16×16 images. Allowing for that each pixel is composed of RGB, the input vector we get has $m=768$ dimensions. After preprocessing the training set, we extract the features with PCA and KPCA, and input the results into support vector machines to train, and then input the pretreated testing set into the classifier to classify.



Figure 1 Images of 100 objects (view angle= 0°). The highlighted images are difficult to recognize. So, in this experiment, they will impact the recognition rate.

To construct the feature space, we only need to select part of images as training samples and the rest as testing samples. In the first experiment, the interval angle of image acquisition is 10° and each object can be got 36 images (see Figure 2), we will get 3600 images of 100 objects as training set and the rest 3600 images as testing set. When the interval angle is 10° , the support vector machine is feasible with recognition ratio up to 95.23% for COIL-100 (see Table I)



Figure 2 Images of one object when the interval angle is 10° and the total is 36 (The first image's angle is 0°). We select these to be training set. And the rest images (the interval angle is also 10° and the first images' angle is 5°) are testing set.

For assessing the feature extraction, we compare our method KPCA-SVM with SVM, PCA-SVM. The optimal parameters of support vector machine are obtained through the searching of parameterized optimization program. In addition, we adjust the interval angle (i.e. 20° , 45° , 60°) of each object of training set, and the results of three different algorithms are shown in Table I.

From table I, we can get good results with the combination of KPCA and SVM. Therefore, we conclude that KPCA can reduce the dimensionality of feature space effectively under not affecting the performance of classifiers. And in this experiment, we got the principal components with KPCA that its dimensions stabilize at around 23.

TABLE I. COMPARISON OF DIFFERENT ALGORITHMS OF DIFFERENT INTERVAL ANGLE

Interval Angle	10°		20°		45°		90°	
	FN	A (%)						
SVM	45	95.23	47	94.38	53	93.20	49	90.12
PCA-SVM	27	94.39	26	91.25	26	91.19	27	89.42
KPCA-SVM	23	95.55	24	96.01	23	95.61	23	95.43

FN: Feature number A: Accuracy

5. Conclusions

In this paper, KPCA is used for feature extraction in 3D object recognition based on the view. As a nonlinear expansion of PCA which solves input features with linear transformation, KPCA transforms it linearly with Kernel. Using RBF Kernel function, we make group experiments with three different algorithms and compare their results with each other. Here, we use KPCA to extract features and SVM to classify. Experiments show that KPCA has strong ability of feature extraction, and the feature extraction method can reduce the dimensionality of feature space of inputting effectively under not reducing the performance of classifiers. Also, it could increase the efficiency of recognition.

Therefore, the 3D object recognition based on view can show strong nonlinear essence, and SVM contains nonlinear kernel mapping and can make good recognition rate. KPCA can extract the nonlinear features from

inputting. Then, we can further increase the efficiency of recognition with the combination of KPCA and SVM.

6. Acknowledgment

This paper is supported by the National Natural Science Foundation of China (Grant NO. 61170132).

7. References

- [1] SUETENS P, FU P, HANSON A J. Computational strategies for object recognition [J] // ACM Computing Surveys, 1992, 24(1): 52- 62.
- [2] ZH L W, LI S W, LI L ZH. 3d object recognition and pose estimation using kernel pca [C]. Proceedings of the Third International Conference on Machine Learning and Cybernetics. Shanghai, China, 2004: 3258-3262.
- [3] JUN H, ZHAO W X, FANG Z S, et al. View-based 3d object recognition using wavelet multi-scale singular value decomposition and support vector machine[C]. Beijing: Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition. 2007: 1428-1432.
- [4] PONTIL M., VERRI A. Support vector machines for 3d object recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(6): 637-646.
- [5] H. Jun, W. Xi-Zhao, Z. Su-Fang, *et.al.* View-based 3d object recognition using wavelet multiscale singular-value decomposition and support vector machine. Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition. Beijing, China, 2007: 1428-1432
- [6] D. Roth, M. H. Yang, N. Ahuja. Learning to recognize three- dimensional objects. Neural Computation, 2002,14(5): 1071-1103
- [7] COSMO D, NICOLA A, PIETRO S Support vector machines for olfactory signals recognition [J]. Sensors and Actuators B, 2003, 88:30-39.
- [8] Jolliffe I J. Principal Component Analysis [M]. New York: Springer, 1986.
- [9] Scholkopf B, Smola A, Muller K R. Kernel principal component analysis[A]. Advances in Kernel Methods-Support Vector Learning[C]. Cambridge; MIT Press,1999.327-352
- [10] JUN H, ZHAO W X, FANG Z S, et al. View-based 3d object recognition using wavelet multi-scale singular value decomposition and support vector machine[C]. Beijing: Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition. 2007:1428-1432.
- [11] VAPNIK V. N. The nature of statistical learning theory [M]. New York: Springer-Verlag, 1995: 180-188.
- [12] Vapnik V N. The Nature of Statistical Learning Theory.Springer,1995
- [13] Pearson,K.(1901).”On Lines and Planes of Closest Fit to Systems of Points in Space”(PDF). Philosophical Magazine 2(6):559-572
- [14] X. Q. Wu, K.Q. Wang, Y. Xu, et al. Differential Feature Analysis for Palmprint Authentication. The 13th International Conference on Computer Analysis of Images and Patterns. 2009
- [15] ZOU Jian-zhang, ZHOU Jing-ye, A Handwritten Digits De-Noising Method Based on the Kernel Method. Computer Engineering & Science.2008,30(20):49-50
- [16] Nene, S. A., Nayar, S. K., Murase,H. (1996) Columbia Object Image Library(COIL-100), Techn. Rep. No. CUCS-006-96, dept. Comp. Science, Columbia University