논문 2006-43SP-2-10

특징추출을 위한 특이값 분할법의 응용

(The Application of SVD for Feature Extraction)

이 현 승*

(Hyun Seung Lee)

요 약

패턴인식 시스템은 일반적으로 데이터의 전처리, 특징 추출, 학습단계의 과정을 거쳐서 개발되어 진다. 그중에서도 특징 추출 과정은 다차원 공간을 가진 입력 데이터의 복잡도를 줄여서 다음 단계인 학습단계에서 계산 복잡도와 인식률을 향상시키는 역할을 한다. 패턴인식에서 특징 추출 기법으로써 principal component analysis, factor analysis, linear discriminant analysis 같은 방법들이 널리 사용되어져 왔다. 이 논문에서는 singular value decomposition (SVD) 방법이 패턴인식 시스템의 특징 추출과정에 유용하게 사용될 수 있음을 보인다. 특징 추출단계에서 SVD 기법의 유용성을 검증하기 위하여 원격탐사 응용에 적용하였는데, 실험결과는 널리 쓰이는 PCA 에 비해 약 25% 의 인식률의 향상을 가져온다는 것을 알 수 있다.

Abstract

The design of a pattern recognition system generally involves the three aspects: preprocessing, feature extraction, and decision making. Among them, a feature extraction method determines an appropriate subspace of dimensionality in the original feature space of dimensionality so that it can reduce the complexity of the system and help to improve successful recognition rates. Linear transforms, such as principal component analysis, factor analysis, and linear discriminant analysis have been widely used in pattern recognition for feature extraction. This paper shows that singular value decomposition (SVD) can be applied usefully in feature extraction stage of pattern recognition. As an application, a remote sensing problem is applied to verify the usefulness of SVD. The experimental result indicates that the feature extraction using SVD can improve the recognition rate about 25% compared with that of PCA.

Keywords : Pattern Recognition, Feature Extraction, SVD

I. INTRODUCTION

Generally, the design of a pattern recognition system involves the following three aspects: preprocessing, feature extraction, and decision making. It is agree that a well-defined pattern recognition system will lead to a compact pattern representation. The system is operated in two modes: training and classification. In the training mode, the feature extraction module finds the appropriate features for representing the input patterns and the classifier is trained to partition the feature space. In the classification mode, the trained classifier assigns the input pattern to one of the pattern classes under consideration based on the measured features.

In pattern recognition problems, feature extraction methods determine an appropriate subspace of dimensionality in the original feature space of dimensionality. It means that, for input data with high dimensional space, the method can appropriately reduce the dimension to lessen computational burden and improve the accuracy of a classification problem.

Linear transforms, such as principal component analysis, factor analysis, and linear discriminant analysis have been widely used in pattern recognition for feature extraction[1]-[2]. The well known linear feature extractor is the principal component analysis

* 정회원, 삼성전자
(Samsung Electronics)
접수일자: 2005년9월27일, 수정완료일: 2006년2월20일

(PCA). Once a feature selection procedure finds a proper representation, a classifier can be designed using a number of possible approaches. The most straightforward K-nearest neighbor (K-NN) rule can be used as a benchmark for all the other classifiers since it appears to provide a reasonable classification performance in most applications. It assigns patterns to the majority class among K-NN using a performance optimized value for k. K-NN is attractive because no prior knowledge of the distributions is required. This rule relies instead on the training set of objects with known class membership to make decisions on the membership of unknown objects. K-NN classifies an unknown object to the class of its nearest neighbor in the measurement space using, most commonly, Euclidean metrics[3].

As remote sensing tools, imaging spectrometers, or hyperspectral scanners, are instruments that acquire spectral images in many contiguous spectral bands throughout the spectrum. This system is to permit discrimination among earth surface features that have diagnostic absorption and reflection characteristics[4]. The system has been widely used in remote sensing technology since it can produce data of sufficient spectral resolution for direct identification of the materials. Since hyperspectral remote sensing provides high resolution spectral data and the potential for remote discrimination between subtle differences in ground covers, it can be a good example of pattern classification problems. It has been proven that the high dimensional data space generated by the hyperspectral sensors have the following properties: the volume of a hypercube concentrates in the corners and the volume of a hypersphere or hyperellipsoid concentrates in an outer shell. This property of the high-dimensional space implies that with limited training data, much of the hyperspectral data space is empty. Principally, two solutions exist: (i) provide larger sets of training data or (ii) reduce the dimensionality by extracting pertinent features from the hyperspectral signals. When considering solution (i), one must take into account that as the number of dimensions increase, the sample size of the training data needs to increase exponentially in order to have valuable estimates of multivariate statistics. Consider the impracticality of this scenario when using a device such as the portable spectroradiometer which has 2151 spectral bands. Thus, we must consider solution (ii), creating a need for feature extraction methods that can reduce the data space dimensions without losing the original information that allows for the separation of classes.

The singular value decomposition (SVD) is one of the most important tools of numerical signal processing. It provides robust solution of both overdetermined and underdetermined least-square problems, matrix approximation, and conditioning of ill-conditioned systems. It is employed in a variety of signal processing applications, such as spectrum analysis, filter design, system identification, model order reduction, and estimation[5]. Furthermore, as a pattern classification application, this paper describes that SVD technique can be applied to a powerful feature extraction tool from a example of remote sensing applications.

Ⅱ. SVD Technique as a Feature Extraction

1. Statistical Feature Reduction Method

A well known principal component analysis (PCA) is linear transformation for feature extraction and dimensionality reduction. It computes m largest eigenvectors of the covariance matrix of the input patterns. Using the covariance matrix of the original data, the eigenvalues and eigenvectors are obtained. Then, the principal components are a transform matrix consisted of column eigenvectors. A basic idea of this method finds principal components so that they explain the maximum amount of variance possible by linearly transformed components.

As a feature reduction tool, singular value decomposition (SVD) can be described[5]. For an matrix A, the singular value decomposition of A is given by (i) U is an mxn matrix, ,V are nxn matrices (ii) U,V are column orthonormal (iii) where r = rank(A) and the singular values are $\sigma_1 \geqq \sigma_2 \cdots \geqq \sigma_r > 0$. When SVD is applied to feature vectors, it not only eliminate the noise in the feature vectors, but also reduce the dimension of the feature. The SVD of a matrix can be used to obtain lower rank approximations of the matrix. If we take the first k columns of U, V (denoted U[k], V[k]) and the leading kxk submatrix of (denoted [k]), and define

$$A_k = U_k \Sigma_k V_k^T = \sum_{i=1}^{k} U_i \Sigma_{i,i} V_i^T \tag{1}$$

,where Ak is the best rank k approximation to A. The arbitrary selection of k is not a good choice since the solution cannot be robust. To select the appropriate order of singular values statistically, Akaike (AIC) model selection criteria can be used[6]:

$$AIC(k) = -2\log\left(\frac{\prod_{i=k+1}^{p} l_i^{1/(p-k)}}{\frac{1}{p-k}\sum_{i=k+1}^{p} l_i}\right)^{(p-k)N} + 2k(2p-k) \tag{2}$$

where li is the ith eigenvalue of the sample covariance matrix and N is the data length. The number of signals is determined as the value of k ∈ {0, 1, ..., p-1} for which the AIC is minimized. That is, the number of feature is a value, k, with a minimum value for equation (2).

2. Remote Sensing Application

Purple and yellow nutsedge are two of the most troublesome perennial weeds found in home landscapes, cultivated fields, and waste areas. They have gained this reputation due to their ability to grow in adverse conditions, effectively compete with more desirable vegetation, and remain dormant for extended periods of time. Remote sensing technology is used to detect variation within fields caused by biotic and abiotic stress factors. Improved detection of vegetation not performing to its full genetic potential will allow improved corrective measures to avoid or alleviate yield losses. The potential of remote sensing as a crop management tool is due to the interactions between light and the leaf surface. When light reaches the surface of a leaf, it is either absorbed, reflected, or transmitted by or through the leaf. Leaf reflectance is controlled by the surface properties, internal structure, and biochemical components of the leaf. Therefore, analysis of the reflected light can be used to determine the physiological status of a plant[7].

Hyperspectral scanner is a instrument that permits discrimination for direct identification of the materials such as purple and yellow nutsedge and has diagnostic absorption and reflection characteristics throughout the spectrum. Hyperspectral remote sensing data, generated by the scanner, cover the full solar reflected portion of the spectrum with high spectral resolution. Because of its dimensionality, hyperspectral data potentially provides the capability to discriminate between nearly any set of classes. It provides rich information on ground cover types and make possible detailed study and monitoring of the Earth. However, data analysis and image classification become difficult due to the high dimensionality resulted by using hundreds of spectral bands. Fig. 1 shows the example of the hyperspectral data of cotton and soybean obtained from

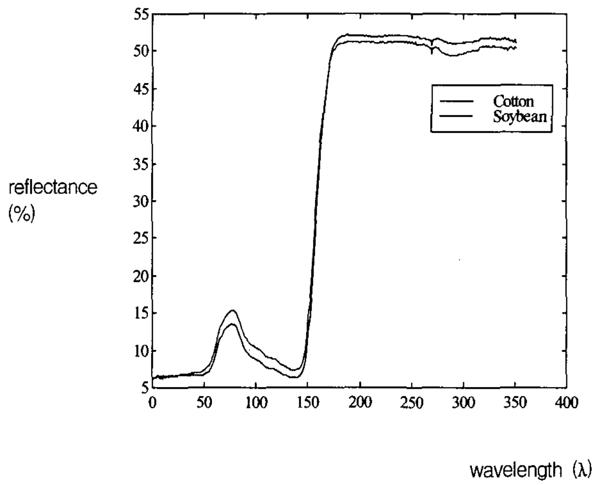

그림 1. cotton과 soybean에 대한 스펙트럼 반사 데이터
Fig. 1. The spectral reflectance data for cotton and soybean.

hyperspectral scanner.

Ⅲ. EXPERIMENTAL RESULTS

A general approach for classification problem is to develop a reliable training set from which predictions of the classes of unknown samples can be accomplished. Then, an evaluation was done with the K-nearest neighbors method. The method utilized for this task is based on the feature extraction and statistical classification. Thus, the process is completed through two phases. For the first phase, the optimal spectral bands features are extracted from the original hyperspectral reflectance curve.

The second phases is called the testing phase during which the classification is implemented based on the extracted features. Ideally, the features extracted should be good at discriminating between classes of interest. In this case, the classes were divided into *cotton or soybean leaf reflectance* measurements where the cotton or soybean was growing without any purple or yellow nutsedge, and cotton and soybean leaf reflectance measurements in which cotton or soybean were growing in the same pot with purple or yellow nutsedge. As a training step, the nearest neighbor analysis is implemented through computing the Euclidean distances between the tested signal and each of other signals in all classes, and then forcing the tested signal into one class in which one signal produces the nearest Euclidean distance with the tested signal, i.e., the tested signal is classified into that class. This analysis is conducted on each spectral band of the hyperspectral reflectance curve. After all training hyperspectral curves are analyzed, the classification accuracy is computed on each spectral band.

The performance of the system was performed using the leave-one-out method. This consisted of leaving one sample out of the system, classifying the sample using the nearest neighbor classifier. This was repeated for the remaining samples thereby providing another system accuracy estimate. The results are plotted as various k-nearest neighbor

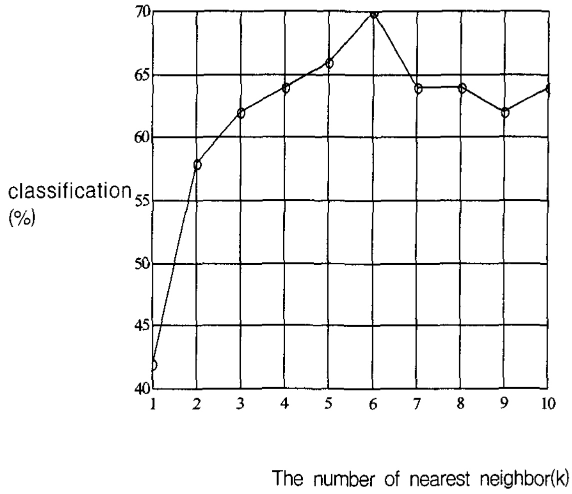

그림 2. PCA 기법을 이용한 분류의 결과
Fig. 2. The classification result for PCA technique.

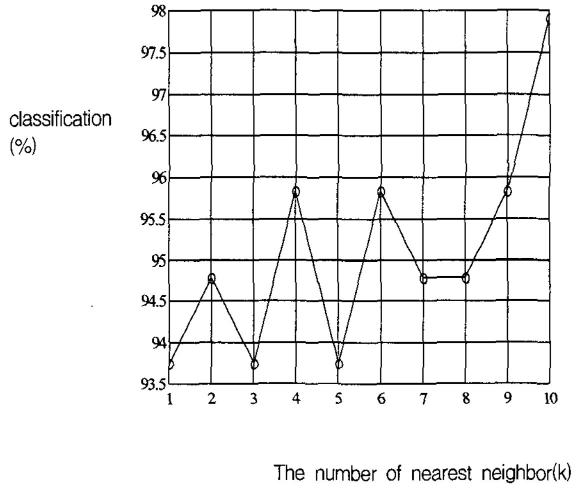

그림 3. SVD 기법을 이용한 분류의 결과
Fig. 3. The classification result for SVD technique.

classifiers. The results are plotted as various k-nearest neighbor classifiers. The result of the classification Using PCA and SVD are shown in Fig. 2 and Fig. 3, respectively.

Also, the classification accuracies for each technique are summarized in Table 1. Each accuracy is calculated as an average of sum of the percentage. The table indicates that the classification accuracy of SVD feature extraction is superior to that of PCA feature extraction technique. That is, SVD technique performed in feature extraction technique of spectral reflectance data has 25% higher accuray than PCA. Furthermore, 'K-nearest neighbor' in the table means

표 1. 특징 추출 기법들에 대한 분류 정확도 측정
Table 1. The classification accuracy for feature extraction techniques.

	Accuracy (%)	95% Confidence Interval (%)
K-nearest neighbor	89.83	89.83±3.35
PCA	70.00	70.00±10.63
SVD	95.82	95.82±7.75

that a k-nearest neighbor classifier without any feature extraction tool is applied to the test data. The experimental result also shows that a classifier with the SVD feature extraction produces the higher classification accuracy about 5% more than that of the classifier without a feature extraction method.

Ⅳ. CONCLUSION

The SVD technique is usually used for spectrum analysis, filter design, system identification, and estimation. In this paper, the application area of the technique is extended to pattern classification problem. Since SVD is a useful tool for model order reduction, it expects a good effect on feature extraction stage of classification problem. The experiment is performed with the soybean and cotton data obtained from a hyperspectral scanner which produce data with high dimensional space and used in remote sensing industry. The result shows that the classification using the SVD technique produces the higher accuracy compared with that of the PCA technique.

References

[1] Todd K. Moon and W. C. Stirling, Mathematical Methods and Algorithms for Signal Processing, Prentice-Hall, Inc., 2000.
[2] John R. Jensen, Introductory Digital Image Processing: A Remote Sensing Perspective, Prentice-Hall, Inc., 1996.
[3] A. K. Jain, R. Duin, and J. Mao, "Statistical Pattern Recognition: A Review," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 1, January 2000, pp. 4-37.
[4] Thomas M. Lillesand,, Remote Sensing and Image Interpretation, John Wiley & Sons, Inc., 1994.
[5] W. M. Steedly, R. L. Moses, and C. J. Ying, "Statistical Analysis of SVD-Based Prony Techniques," IEEE Proceedings of the 25th Asilomar Conference on Signals, Systems, and Computers, Nov. 1991.
[6] M. Wax and T. Kailath, "Detection of Signals by Information Theoretic Criteria," IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 33, No. 2, April 1985.
[7] R.A. Schowengerdt, Remote Sensing: Models and Methods for Image Processing, Academic Press, San Diego, 1997.

저 자 소 개

이 현 승(정회원)
1996년 성균관대학교 전기전자공학과 학사.
1999년 Mississippi State Univ. 전자공학과 석사.
2003년 Mississippi State Univ. 전자공학과 박사.
2002년~현재 삼성전자 책임연구원
<주관심분야 : 영상처리, 통계신호처리, 동영상 압축>