한국지능시스템학회 논문지 2009, Vol. 19, No. 1, pp. 90-95

자료별 분류분석(DDA)에 의한 특징추출

Datawise Discriminant Analysis For Feature Extraction

박명수 · 최진영
Myoung Soo Park and Jin Young Choi

서울대학교 전기컴퓨터공학부

요 약

본 논문은 선형차원감소(Linear Dimensionality Reduction)을 위해 널리 이용되고 있는 특징추출 알고리듬인 선형판별분석(Linear Discriminant Analysis)의 문제점을 해결할 수 있는 새로운 특징추출 알고리듬을 제안한다. 선형판별분석에 포함되는 평균-자료 간 거리 및 평균-평균 간의 거리에 기반한 분산행렬은 역행렬 연산, 계수의 제한 등으로 인하여 계산상의 문제와 추출되는 특징의 수가 제한되는 한계를 가지고 있다. 또한 자료의 집단이 단일 모드의 정규 분포로부터 얻어진 것으로 가정되며 그렇지 않은 경우에 대해서는 적절한 결과를 얻을 수 없다. 본 논문에서는 자료-자료 간의 거리에 기반하고 적절하게 가중치가 추가된 새로운 행렬을 정의하였으며, 이에 기반하여 특징을 추출하는 방법을 제안하였다. 그럼으로써 앞서 선형판별분석의 여러 문제를 해결하고자 시도하였다. 제안된 방법의 성능을 실험을 통해 확인하였다.

키워드 : 선형차원감소, 선형판별분석, 자료별판별분석, 자료-자료간 거리. 비정규 분포 자료.

Abstract

This paper presents a new feature extraction algorithm which can deal with the problems of linear discriminant analysis, widely used for linear dimensionality reduction. The scatter matrices included in linear discriminant analysis are defined by the distances between each datum and its class mean, and those between class means and mean of whole data. Use of these scatter matrices can cause computational problems and the limitation on the number of features. In addition, these definition assumes that the data distribution is unimodal and normal, for the cases not satisfying this assumption the appropriate features are not achieved. In this paper we define a new scatter matrix which is based on the differently weighted distances between individual data, and presents a feature extraction algorithm using this scatter matrix. With this new method, the mentioned problems of linear discriminant analysis can be avoided, and the features appropriate for discriminating data can be achieved. The performance of this new method is shown by experiments.

Key Words : Linear dimentionality reduction, linear discriminant analysis, data discriminant analysis, distance between individual data, non-normally distributed data

1. 서 론

큰 입력차원을 가지는 자료를 분류하는 경우, 입력을 낮은 차원의 특징 공간(feature space)으로 투영(project)하여 차원을 감소시킬 필요가 있다. 선형차원감소(LDR: Linear Dimensionlity Reduction)는 이것을 목적으로 하는데, 주성분분석(Principal Component Analysis)[1], 선형판별분석(LDA: Linear Discriminant Analysis)[2] 등이 여기에 속한다. 이 가운데 LDA는, 최종적으로 분류해야 하는 자료집단(class)에 대한 정보를 이용하기에 감소된 차원 내에 분류를 위한 정보를 PCA에 비해 보다 많이 유지할 수 있다는 장점이 있다.

LDA은 일반적으로 피셔 기준(Fisher Criteria)에 의하여, 자료들의 집단내분산(within-class scatterness)을 최소화하고 집단간분산(between-class scatterness)을 최대화하는 기저벡터(basis vector)를 찾아내는 문제로 수식화된다[2]. 이 경우 기저벡터는, 집단간분산에 해당하는 행렬 S^b와 집단내분산에 해당하는 행렬 S^w의 역행렬을 곱하여 얻어지는 행렬 $(S^w)^{-1}S^b$의 고유벡터에 해당한다. 일반적으로 이러한 행렬을 구하고 고유치문제를 푸는 것이 LDA를 수행하는 과정인데, 앞서의 행렬에 포함되는 역행렬 연산과 집단내분산 및 집단간분산의 계수(rank)로 인하여, 계산상의 문제가 발생하게 되고 추출될 수 있는 기저벡터의 수가 제한된다. 게다가 피셔 기준을 이용한 수식화과정에서는 각 집단에 속하는 자료들이 서로 다른 단일 모드의 (unimodal) 정규 분포로부터 발생한 것으로 가정된다[2]. 따라서 그렇지 않은 경우로부터 얻어진 자료에 대해서는 추출된 결과가 분류에 적절하지 않을 수 있다. 역행렬의 이용으로 인한 계산상의 문제는 여러 연구자들에 의해 많이 연구되었지만[3, 4] 다른 문제점에 대해서는 상대적으로 제한된 연구만이 이루어졌다.

최근 제안된 선형차원감소 알고리듬인 국소화된 피셔분석(LFDA: Localized Fisher Discriminant Analysis)[5]에서는, 기존의 LDA를 개선하여 각 집단에 속하는 자료들이

접수일자 : 2008년 11월 1일
완료일자 : 2009년 1월 19일
본 연구는 지식경제부 BK21프로그램과 삼성테크윈에서 지원 받았습니다.

여러 모드(multimodal)로 이루어진 정규분포로부터 발생하는 경우에 대해서도 적절한 특징을 찾을 수 있게 하였다. S^b와 S^w가 자료-평균 차이가 아니라 개별 자료-자료 간의 차이로부터 계산될 수 있음을 보이고, 보다 가깝게 위치하여서 분류가 어려운 자료들의 차이에 상대적으로 큰 가중치를 부여함으로써 이러한 목적을 달성하였다. 가중치의 부여로 인하여 S^b와 S^w의 계수가 증가되면서 추출될 수 있는 특징 수의 제한 문제 또한 해결될 수 있다. 그러나 전체 계산에 여전히 기존의 LDA에서와 마찬가지로 역행렬이 이용되고, 이로 인하여 발생하는 문제를 해결하기 위해 LDA에 대해 개발된 여러 기법들이 그대로 적용되고 있다.

이 논문에서는 자료간 차이로부터 행렬을 계산하는 과정에 주목하여, 이를 활용하는 새로운 자료별 분류분석(DDA: Datawise Discriminant Analysis)을 제안하고자 한다. 제안하는 방법은 자료간 차이로부터 계산되고 새로운 가중치를 포함하는 행렬을 정의하고, 이에 대한 고유치문제를 풀어 기저벡터를 구한다. 이러한 방식을 이용함으로써 전체 자료의 분포에 대한 가정이 필요하지 않게 되므로, 자료가 단일 모드의 정규본포를 따르지 않는 경우에 대해서도 적절한 특징을 추출할 수 있다. 각각의 자료간 차이에 따른 가중치로 역행렬 계산을 필요로 하는 S^w을 대신함으로써 LDA 및 LFDA와 달리 역행렬 연산으로 인한 계산상의 문제를 근본적으로 피할 수 있다. 이용되는 행렬의 계수는 LFDA의 경우와 마찬가지로 가 적은 값으로 제한되지 않아, 많은 수의 기저벡터를 추출할 수 있다.

2. 자료별 분류분석

LDA에서 목적으로 하는 기저벡터 w는 아래와 같은 함수를 최대화하는 해로부터 구할 수 있다.

$$J(w) = \frac{w^T S^b w}{w^T S^w w}$$

이 식에서 S^b와 S^w는 각각 집단간분산과 집단내분산과 관련된 것으로 다음과 같이 정의된다.

$$S^w = \sum_{c=1}^{n_c} \sum_{\forall i: y[x(i)]=c} [x(i)-m(c)][x(i)-m(c)]^T$$

$$S^b = \sum_{c=1}^{n_c} N(c)[m(c)-m][m(c)-m]^T$$

여기서 $x(i)$는 i번째 자료, $y[x(i)]$는 i번째 자료의 집단 인덱스, n_c는 집단의 수, $m(c)$와 m은 각각 집단간 평균과 전체 평균으로 다음과 같이 계산되는 값이다.

$$m(c) = \frac{1}{N(c)} \sum_{\forall i: y[x(i)]=c} x(i)$$

$$m = \frac{1}{N} \sum_{i=1}^{N} x(i)$$

이다. $N(c)$와 N은 각각 c번째 집단에 속하는 자료의 수와 전체 자료의 수를 의미한다.

평균에 대한 위의 식을 S^b에 대입하면 아래와 같이 정리되는데

$$S^b = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij}^b [x(i)-x(j)][x(i)-x(j)]^T$$

여기서, 계수 a_{ij}^b는

$$a_{ij}^b = \begin{cases} 1/N - 1/N_c & \text{if } y[x(i)] = y[x(j)] = c \\ 1/N & \text{if } y[x(i)] \neq y[x(j)] \end{cases}$$

로 정의된다. 이를 이용하면, S^b를 포함하는 $J(w)$의 분자항은

$$w^T S^b w = w^T \left[\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij}^b [x(i)-x(j)][x(i)-x(j)]^T \right] w$$

$$= \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij}^b [w^T [x(i)-x(j)][x(i)-x(j)]^T w]$$

로 표현된다. $d_{ij}^b = w^T[x(i)-x(j)]$는 두 자료의 차이를 기저벡터 w위로 투영시킨 값이라고 하면,

$$w^T S^b w = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij}^b (d_{ij}^b)^2$$

로 표현된다. 따라서 이 값을 최대화하는 것은, 각 자료 간의 차이의 가중치 a_{ij}^b에 의한 자료-자료간 차이의 가중평균을 최대로 포함하는 w를 찾는 것을 의미한다. 따라서 가중치를 조정하면 자료-자료간 차이 가운데 특정한 것들이 상대적으로 강조되도록 할 수 있다. LFDA는 거리에 반비례하는 값을 가중치에 곱하여, 자료-자료간의 국소적인 차이가 강조되도록 하였다. S^w에 대해서도 마찬가지로 전개하고, 국소적인 차이를 강조하였다. 그리고 구해진 S^b와 S^w를 LDA와 마찬가지의 방식으로 이용하여 기저벡터를 구하였다.

그런데, 본래의 항에 곱해지는 가중치를 자세히 살펴보면, 소속집단이 같은 자료들의 차이에 대한 가중치는 음의 부호를 가져 그러한 자료들 간의 차이를 적게 포함하는 기저벡터를 선호되게 하며, 소속집단이 다른 자료들의 차이에 대한 가중치는 양의 부호를 가져 그러한 자료들 간의 차이를 많이 포함하는 기저벡터가 선호되게 하는 것이라 볼 수 있다. 또한 소속집단이 같은 자료들의 차이에 대한 가중치는, S^b가 자료집단의 평균-평균간 차이에서 계산되는 것으로 정리되게 한다. 여기서 평균이 나타나는 것은 피셔기준으로부터 비롯하는데, 각각의 집단에 속하는 자료가 단일모드의 정규 분포를 가지지 않는 경우에 이용하는 것은 적절하지 않다. 그러한 경우에는 평균-평균간 차이 대신 개별 자료간 차이를 강조하는 것이 보다 적절할 것이다. 이를 위해서 같은 집단에 속하는 자료간 차이에 대한 가중치를 0으로 하는 새로운 형태의 S_{new}^b를 다음과 같이 정의할 수 있다.

$$S_{new}^b = \sum_{i=1}^{N} \sum_{i=1}^{N} k_{ij}^b [x(i)-x(j)][x(i)-x(j)]^T$$

여기서 가중치는 아래와 같다.

$$k_{ij}^b = \begin{cases} 0 & \text{if } y[x(i)] = y[x(j)] \\ 1 & \text{if } y[x(i)] \neq y[x(j)] \end{cases}$$

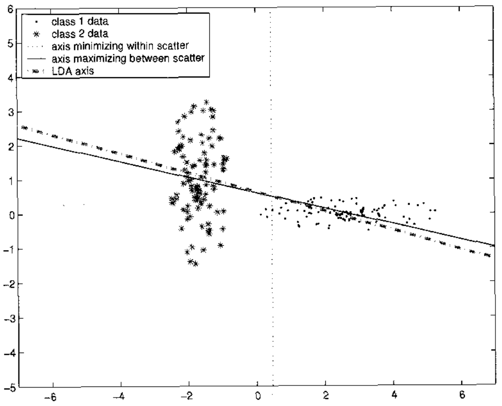


그림 1. S^b의 이용이 유용한 경우: S^b에 의한 기저벡터에 의해 계산되는 특징이 자료를 분류하기에 제일 적합하다.

Fig. 1. The case when the use of S^b is appropriate for discriminating data: The features computed by basis vector from the use of S^b is appropriate for discrimnating data.

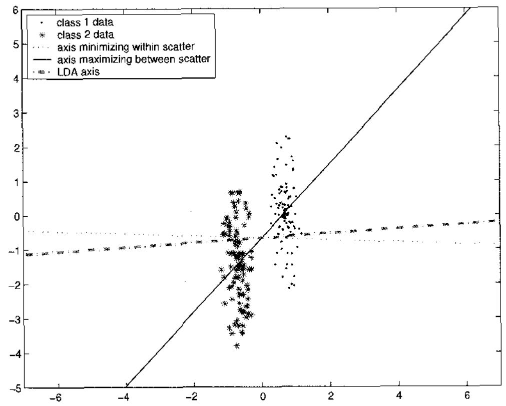


그림 2. S^w의 이용이 유용한 경우: S^w에 의한 기저벡터에 의해 계산되는 특징이 자료를 분류하기에 제일 적합하다.

Fig. 2. The case when the use of S^w is appropriate for discriminating data: The features computed by basis vector from the use of S^w is appropriate for discrimnating data.

다음으로 집단내분산에 해당하는 S^w를 살펴보자, 몇 가지 간단한 경우에 대해 S^w를 최소화하는 경우와 S^b를 최대화하는 경우, 그리고 두 가지를 함께 고려하는 LDA에 의해 추출된 기저벡터의 방향을 그림 1과 2에 나타내었다. 점선은 집단내분산만을 이용하여 추출되는 기저벡터의 방향에 해당하며, 실선과 파선은 각각 집단간분산을 이용하여 추출되는 기저벡터와, 집단내분산과 집단간분산을 모두 이용하는 LDA에 의해 추출되는 기저벡터의 방향에 해당한다. 두 그림에서 각 집단별 자료는 단일 모드의 정규분포를 따르고 있다. 이 경우, 두 가지 종류의 분산을 최대화하는 기저벡터는 자료의 분포에 따라 자료의 분류에 유용한 경우도 있고 그렇지 않은 경우도 있지만, 두 가지를 모두 이용하는 LDA의 경우에는 분류에 적절한 특징을 추출함을 알 수 있다.

보다 상세히, 두 그림을 관찰해보면 S^w가 유용하게 이용되는 경우는, 각 집단에 있는 자료들 중 다른 집단과의 경계에 위치하는 자료들이 다른 집단에서 마찬가지의 위치에 있는 자료들과 최대한 멀리 위치하게 하는 방향이, 집단내분산이 최소가 되는 방향과 일치하는 경우들이라 볼 수 있다. 그렇지 않은 그림 1과 같은 경우에 S^w의 이용은 S^b만을 이용한 경우에 비해 유용하지 않다. 즉 중요한 것은 집단내분산 그 자체라기보다 집단의 경계에 가까운 자료들의 거리가 경계에서 먼 자료에 비해서 최대한 고려되도록 만드는 것이다. 따라서 우리는 앞서 S^b_{new}에, LFDA에서와 마찬가지로 자료-자료 간의 거리가 적은 경우를 강조하도록 가중치를 조정하면, S^w를 포함하지 않고도 좋은 성능을 얻을 수 있을 것임을 예상할 수 있다. 이 예상에 기반하여 수정된 가중치를 이용하여 S_{new}를 다음과 같이 정의할 수 있다.

$$S_{new}=\sum_{i=1}^{N}\sum_{i=1}^{N}k_{ij}[x(i)-x(j)][x(i)-x(j)]^T$$

여기서 가중치는

$$k_{ij}=\begin{cases}0 & \text{if } y[x(i)]=y[x(j)]\\ f(\|x(i)-x(j)\|) & \text{if } y[x(i)]\neq y[x(j)]\end{cases}$$

이며, $f(\|x(i)-x(j)\|)$는 $\|x(i)-x(j)\|$의 감소함수로 다양하게 정의될 수 있다.

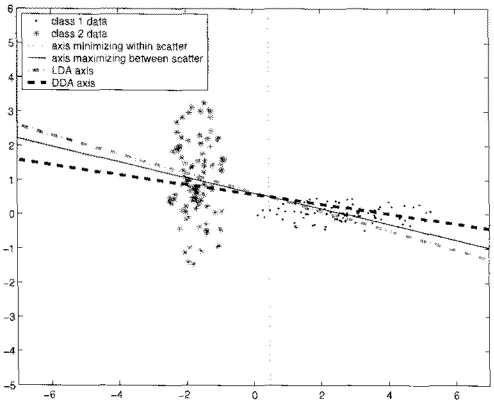
위에서 정의된 S_{new}를 이용하여 정의되는 아래와 같이 새로운 함수 $J_{new}(w)$를 정의하고,

$$J(w)_{new}=w^T S_{new} w$$

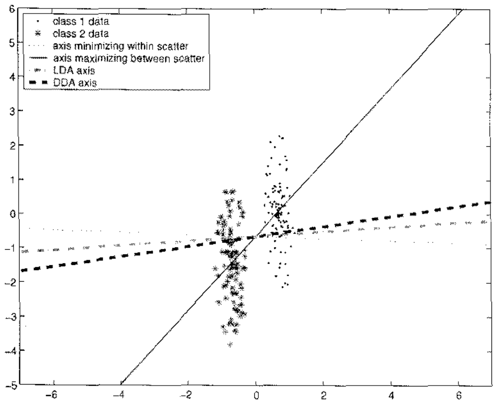
이를 최소화하는 기저벡터를 구하고 그 위의 투영을 통해 차원을 감소시키는 방법을 자료별분류분석(Datawise Discriminant Analysis)이라고 부르고자 한다. LDA와 이에 기반한 알고리듬들과 비교할 때, DDA의 계산은 PCA와 유사하다. 기저벡터를 구하는 과정에서 집단내분산에 해당하는 행렬의 역행렬을 구하거나 일반화된 고유치문제(generalized eigenvalue problem)를 푸는 대신에, 단순한 고유치문제를 푸는 것으로 충분하다. 역행렬 연산으로 인한 계산상의 문제를 피할 수 있다. 또한 S_{new}의 계수는 집단의 수가 아니라 자료의 수와 입력차원에 의해 제한되므로 일반적으로 최대로 얻을 수 있는 특징의 수 또한 LDA에 비해 많다. 이는 적은 입력차원을 가지는 문제에 대하여 좋은 분류성능을 내는 분류기를 설계하는 데에 도움이 될 수 있다. DDA를 앞서 그림 1과 2에서 제시되었던 자료집합에 적용한 아래 결과에서 이러한 사항들을 쉽게 확인할 수 있다.

그림 3의 (a)와 (b)는 앞서 제시하였던 자료집합에 대해 DDA가 추출하는 기저벡터의 방향을 보여주고 있다. 두 자료집합 모두에 대해 DDA는 LDA를 이용한 것과 유사한 결과를 보여, 자료를 분류하는데 있어 적절한 특징을 추출할 수 있음을 확인할 수 있다. 이러한 결과가 앞서 언급된 LDA의 여러 가지 문제점을 가지지 않으면서 얻어지는 것이라는 점에 주목할 필요가 있다. (c)의 경우는 비정규적으로 분포하는 자료들의 집합에 대해 DDA가 추출하는 기저벡터의 방향을 보여주고 있다. 이 자료집합에서 첫 번째 자료집합의 자료는 두 개의 모드(mode)를 가지는 정규분포를 따르며, 이는 앞서 말한 단일 모드의 정규분포를 벗어나는 것이다. 이 경우에 집단간분산과 집단내분산을 이용하여 추

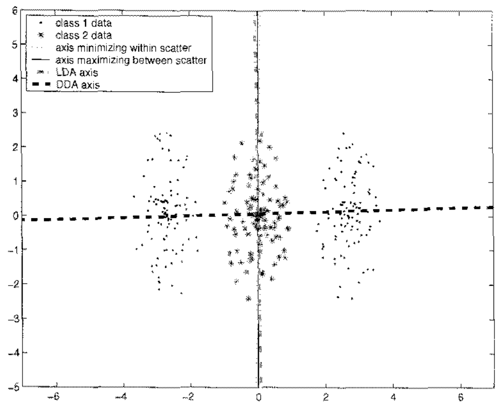
출되는 기저벡터는 분류에 적절한 기저벡터와 거의 수직방향을 지시하게 된다. 그 결과 LDA는 그림에서 보는 바와 같이 자료를 분류하기에 적절한 특징을 계산할 수 있는 기저벡터를 찾아내는데 실패하는 반면에, DDA는 적절한 기저벡터를 찾아낼 수 있음을 확인할 수 있다.



(a) 그림 1의 경우에 대한 기저벡터
(a) The basis vector for the case in Fig.1.



(b) 그림 2의 경우에 대한 기저벡터
(b) The basis vector for the case in Fig.2.



(c) 자료가 비정규적으로 분포하는 경우에 대한 기저벡터
(c) The basis vector for the non-normally distributed data.

그림 3. DDA에 의한 기저벡터: DDA에 의해 계산된 특징이 자료를 분류하기에 적합하다.
Fig. 3. The basis vector found by DDA: The features computed by DDA is appropriate for discrimnating data.

정리하자면, DDA는 LDA와 유사한 성능을 가지면서도 LDA의 여러 가지 문제점은 가지지 않는다는 장점이 있다. S_{new}의 계수(rank)가 집단의 수에 의해 제한되지 않으므로 집단의 수가 적은 경우에도 추출할 수 있는 기저벡터의 수, 즉 특징의 수가 제한되지 않는다. S_{new}는 피셔기준으로부터 유도된 LDA와 달리 분산과 평균이 아니라 자료간 차이에 집중하므로 각 집단의 자료가 단일 모드의 정규분포로부터 얻어졌다고 가정할 필요가 없고, 그렇지 않은 경우에 있어서도 정상적으로 동작한다. S^{w}의 기능을 대신할 수 있는 가중치를 S^{w} 대신 삽입함으로써 S^{w}의 역함수를 계산하는 것으로부터 비롯하는 여러 문제점이 발생하지 않는다. 기저벡터들은 일반적인 고유치문제가 아니라, PCA와 마찬가지로 단순고유치문제를 풀어 얻을 수 있다.

표 1 실험에 이용된 자료집합의 특성들
Table 1. The properties of data sets used in experiment.

	class	input	data
CAR	4	6	1728
CMC	3	9	1473
GLASS	6	9	214
HARBERMAN	2	3	306
IONOSPHERE	2	34	351
IRIS	3	4	150
LENSES	3	4	24
LIVER	2	6	345
PIMA	2	8	768
SONAR	2	60	208
THYROID	3	5	215
TICTACTOE	2	9	958
WINE	3	13	178
WBCancer	2	9	699

3. 실험 및 결과

앞서 제안된 DDA가 LDA와 유사한 방식으로 분류에 적절한 특징을 추출하게 됨을 설명하였다. 이제 실제 문제에 대해 DDA를 적용하여 그 성능을 평가하고자 한다. DDA와 다른 특징추출 알고리듬의 성능에 대한 비교평가를 위하여 UCI Machine Learning Repository에 포함된 여러 자료집합을 이용하였다. 이용된 자료집합은 다양한 입력차원, 자료수, 집단 수를 가지고 있으며, 다양한 응용분야로부터 얻어진 것이다. 이러한 자료집합에 대한 성능으로부터 다양한 실제 응용시의 성능을 예상할 수 있다. 이용된 자료집합의 특성을 표 1에 제시하였다.

표 2. DDA와 다른 특징추출 알고리듬에 의해 추출된 특징을 이용한 1-NN 분류기의 분류성능: 표의 첫 번째 열은 자료집합의 이름이며, 첫 번째 행은 특징추출방법을 말한다. 숫자는 해당하는 행의 자료집합에 대하여 해당하는 열의 특징추출방법에 의해 추출된 특징으로부터 설계되는 1-NN 분류기의 % 분류정확도(Accuracy)에 해당한다.

Table 2. The classification accuracies of 1-NN classifiers which are constructed with input features extracted by different methods: The first column indicates the data set, the first row indicates the feature extraction methods. Each digit indicates the classification percent accuracy of 1-NN classifier with input features extracted by corresponding feature extraction methods, for corresponding data set.

	DDA	LDA	Fisherface
CAR	90.28	88.77	88.95
CMC	48.13	44.87	45.49
GLASS	75.23	74.30	75.23
HARBERMAN	67.97	63.73	65.71
IONOSPHERE	90.03	73.22	77.78
IRIS	96.67	93.33	96.00
LENSES	91.67	70.83	70.83
LIVER	65.80	58.26	58.55
PIMA	68.49	59.11	63.80
SONAR	85.10	57.69	62.50
THYROID	95.81	89.30	93.95
TICTACTOE	97.18	95.51	98.23
WINE	76.97	74.62	75.84
WBCancer	96.14	95.42	96.14

평가는 다음과 같은 방법으로 수행되었다. 우선 각각의 자료집합을 학습자료와 평가자료로 구분짓고, 학습자료에 DDA와 다른 특징추출 알고리듬들을 적용하여 특징을 추출하고, 간단한 1-NN(Nearest Neighborhood) 분류기를 설계한다. 설계된 분류기의 평가자료에 대한 분류성능을 구하여 평가지표로 삼는다. 이 과정에서, 자료집합을 동일한 크기의 k개의 자료로 구분하고, 그 각각을 평가자료로 하고 나머지를 학습자료로 하여 얻어진 분류성능을 평균한 것을, 전체 자료집합에 대한 분류성능으로 평가하는 방식을 k-fold 교차검증이라고 하는데, 여기서는 이의 극단적인 경우인 k=N(N은 자료집합의 모든 자료의 수)에 해당하는 leave-one-out 방법을 이용한다. 비교를 위해서 널리 이용되는 선형차원감소를 위한 인 특징추출 방법인 LDA와, LDA의 계산상의 문제를 해결하기 위하여 PCA를 전처리에 이용한 후 LDA를 사용하는 Fisherface 기법을 실험에 사용하고 성능을 구하여 DDA의 성능과 비교하였다.

얻어진 결과는 표 2에 주어진 바와 같다. 대부분의 경우에 대해 DDA가 LDA와 비슷하거나 더 나은 성능을 보임을 확인할 수 있다. 따라서 LDA의 이용이 필요한 경우, 특히 LDA의 계산상 문제 혹은 특징 수의 제한이 발생할 가능성이 있거나 자료가 정규분포를 따르지 않는 많은 경우에 대하여, LDA를 대신하여 DDA를 이용하여 비슷하거나 나은 성능을 얻을 수 있음을 확인할 수 있다. 다만, 이상의 결과는 특징들의 성능을 비교하기 위해 단순한 1-NN 분류기를 이용한 결과이므로, 실제 보다 정교한 형태의 다른 분류기에서 좋은 성능을 낼 것인지에 대해서는 확인할 수 없다. 이에 대해서는 추가적인 연구를 통해 확인할 필요가 있다.

4. 결 론

이 논문에서는 입력자료의 선형차원감소를 위해 기존에 널리 이용되어온 LDA를 대신할 수 있는 새로운 알고리듬인 DDA를 제안하였다. LDA의 핵심이 되는 집단간-집단내 분산에 해당하는 행렬들을 대신할 수 있는 새로운 행렬을 제안하였다. 집단간분산의 계산에 이용되는 집단의 평균간 차이를 대신하여 개별 자료-자료간 차이를 이용하였고, 이에 자료-자료간 차이의 크기에 의존하는 가중치를 부여함으로써 집단내분산을 대신하였다. 이와 같이 정의된 새로운 행렬에 대해 고유치문제를 풀어 특징을 추출하였다. 이와 같이 수행되는 DDA는 LDA가 가지는 계산상의 문제와 특징 수의 제한을 가지지 않으면서도, LDA와 달리 단일 모드의 정규분포를 가지지 않는 자료집단의 구분에 있어서도 정상적으로 동작할 수 있다. 이상의 알고리듬에 대한 실험을 통하여 그 성능을 확인하였다.

참 고 문 헌

[1] I. T. Jolliffe, *Principal Component Analysis*, Springer-Verlag, 1986.

[2] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Morgan-Kaufmann, 1990.

[3] P. N. Belhumeur, J. P. Hespanha, D. J. Kriegman, “Eigenfaces vs. fisherfaces: recognition using class specific linear projection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, 1997.

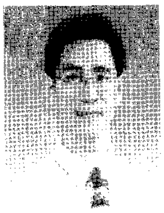
[4] H. Yu, J. Yang, "A direct LDA algorithm for high dimensional data with application to face recognition," *Pattern Recognition*, Vol.34, pp. 2067-2070, 2001,

[5] M. Sugiyama, "Dimensionality reduction of multimodal labeled data by local Fisher discriminant analysis," *Journal of Machine Learning Research*, vol.8, pp.1027-1061, 2007.

저 자 소 개

박명수(Park, Myoung Soo)
1998년 : 서울대 전기공학부 학사
2000년 : 서울대학교 대학원 전기컴퓨터 공학부 석사
2006년 : 서울대학교 대학원 전기컴퓨터 공학부 박사
2006년~현재 : 서울대학교 자동화시스템 공동연구소 인지지능기술 연구센터 연구원.

관심분야 : 기계학습, 영상처리, 신경회로망
Phone : +82-2-872-7283
E-mail : mspark@neuro.snu.ac.kr



최진영(Choi, Jin Young)
1982년 : 서울대 제어계측공학과 학사
1984년 : 서울대 제어계측공학과 석사
1993년 : 서울대 제어계측공학과 박사
1984~1994년 : 한국전자통신연구소 연구원
1998~1999년 : University of California, Riverside 객원교수
1994년~2004년 : 서울대 전기공학부 부교수
2004년~현재 : 서울대 전기공학부 교수

관심분야 : 영상감시(visual surveillance), 신경회로망, 적응제어
Phone : +82-2-880-8372
E-mail : jychoi@neuro.snu.ac.kr