

차분진화 알고리즘을 이용한 지역 Linear Discriminant Analysis Classifier 기반 패턴 분류 규칙 설계

Design of Pattern Classification Rule based on Local Linear Discriminant Analysis Classifier by using Differential Evolutionary Algorithm

노석범* · 황은진* · 안태천**

Seok-Beom Roh*, Eun-Jin Hwang* and Tae-Chon Ahn**

* 원광대학교 전자 및 제어 공학부

요 약

본 논문에서는 전형적인 Linear Discriminant Analysis를 확장시켜 전체 입력공간을 다수의 지역공간으로 분할하고 분할된 공간에 Local Linear Discriminant Analysis 기반으로 하여 패턴 분류 규칙을 설계하는 새로운 방법을 제안한다. 전체 입력공간을 여러 개의 지역공간으로 분할하기 위한 방법으로 unsupervised clustering의 대표적인 방법인 k-Means 클러스터링 기법과 최적화 알고리즘인 차분 진화 연산 알고리즘을 사용한다. 제안된 알고리즘의 성능 평가를 위해 기존의 패턴 분류기와 비교 결과를 제시한다.

키워드 : 선형 판별식 분석, 지역 선형 판별식 분석, 차분 진화 알고리즘, 패턴 분류 규칙

Abstract

In this paper, we proposed a new design methodology of a pattern classification rule based on the local linear discriminant analysis expanded from the generic linear discriminant analysis which is used in the local area divided from the whole input space. There are two ways such as k-Means clustering method and the differential evolutionary algorithm to partition the whole input space into the several local areas. K-Means clustering method is the one of the unsupervised clustering methods and the differential evolutionary algorithm is the one of the optimization algorithms. In addition, the experimental application covers a comparative analysis including several previously commonly encountered methods.

Key Words : Linear Discriminant Analysis, Local Linear Discriminant Analysis, Differential Evolutionary Algorithm, Pattern Classification Rule

1. 서 론

최근 연구되고 있는 다양한 패턴 분류기들 중 통계적 분류기는 Linear Discriminant Function, Quadratic Discriminant Function, Nearest Neighbors Classifier (NNC), Parzen window 분류기 등의 분야에서 활발하게 연구되어지고 있다. 통계적 기법을 이용한 패턴 분류를 위한 다양한 접근 방법들 중에서 Linear Discriminant Analysis (LDA) 는 수십 년 동안 패턴 인식 [1, 2, 3], 정보 검색 [4, 5], 얼굴 인식 [6, 7], 컴퓨터 비전 [8]과 같은 응용분야에서 성공적인 결과를 보이고 있다.

Between-class scatter matrix와 within-class scatter matrix의 비율을 최대화하는 LDA를 이용하여 선형 분류기를 설계함으로써 서로 다른 클래스 사이의 분류 성능을 개선할 수 있다. 전형적인 LDA를 다양화하여 다른 특성을 가지게 된 변형된 LDA 기법들이 제안되어 왔다[8]. Fukunaga와 Mantock [9]은 non-parametric between-class scatter matrix를 사용하는 non-parametric discriminant analysis 기법을 제안하였다. Mahalanobis distance를 기반으로 하는 feature reduction 기법이 Brunzell과 Eriksson [10]에 의해 제안되어졌다. 또한, 모든 데이터를 동일한 가중치를 주고 between-class scatter matrix와 within-class scatter matrix를 계산하는 방법을 대신하여, 각 데이터의 가중치를 지역적 관점에서 차등을 두어 계산하는 local fisher discriminant analysis 방법이 제안되었다[11].

본 논문에서는 local fisher discriminant analysis 방법을 수정하여 local area의 정의하기 위하여 clustering 기법을 이용하였다. 이 clustering 기법은 k-Means algorithm과 합리적인 구조와 빠른 탐색 속도 가 장점 [12]인 차분진화 알고리즘을 이용한 데이터 분석 방법

접수일자: 2011년 11월 19일

심사(수정)일자: 2012년 2월 10일

게재확정일자 : 2012년 2월 11일

* 교신저자

감사의 글 : 이 논문은 2010년도 원광대학교 교비지원에 의해 수행됨.

을 사용하였다. 끝으로 제안된 분류기의 성능을 검증하기 위해 다양한 Machine learning classification 문제에 제안된 분류기를 적용한 분류 성능을 기존 패턴 분류기의 성능과 비교하였다.

2. Local Fisher Linear Discriminant Analysis

Fisher가 제안한 Linear Discriminant Analysis (LDA)는 Principle Component Analysis (PCA)와 같이 입력 공간의 차원감소 (Dimension Reduction) 기법들 중의 하나이다. 일반적인 LDA 기법에서는 모든 데이터들이 아래와 같은 선형식을 이용하여 원래 입력공간의 차원보다 아주 작은 차원을 가지는 feature space로 투영된다.

$$y = w^T x \quad (1)$$

여기서 $y \in R$ 는 원래의 입력공간에 존재하는 데이터들을 feature space로 투영한 feature vector이며, $w \in R^1$ 는 변환 벡터이다. 이 1은 입력변수의 개수, 다시 말하면 입력공간의 차원을 의미한다.

LDA 기법은 (2)와 같이 정의된 목적 함수를 최대화시키는 최적의 w 을 구하는 것이다.

$$J = \frac{w^T S_B w}{w^T S_W w} \quad (2)$$

여기서, S_B 는 between-class scatter matrix를 의미하며, (3)과 같이 구하며, S_W 는 within-class scatter matrix를 나타내며, (4)를 통해 구한다.

$$S_B = (m_1 - m_2)(m_1 - m_2)^T \quad (3-1)$$

$$m_i = \frac{1}{n_i} \sum_{x \in D_i} x \quad (3-2)$$

여기서, D_i 는 i 번째 클래스에 속한 데이터들의 집합을 의미하며, n_i 는 집합 D_i 에 속한 원소들의 개수를 의미한다.

$$S_i = \sum_{x \in D_i} (x - m_i)(x - m_i)^T \quad (4-1)$$

$$S_W = S_1 + S_2 \quad (4-2)$$

목적함수 (2)를 최대화시키는 w 는 (5)와 같이 구할 수 있다.

$$S_B w = \lambda S_W w \quad (5-1)$$

$$w = (S_W)^{-1} (m_1 - m_2) \quad (5-2)$$

위에 언급한 전형적인 LDA는 입력공간의 모든 영역에 분포하는 모든 데이터를 대상으로 하여 (3)과 (4)를 이용하여 between-class scatter matrix와 within-class scatter matrix를 계산하였다.

2.1 Local Fisher Discriminant Analysis (LFDA) [11]

전형적인 Fisher Discriminant Analysis는 하나의 클래스 안에 여러 개의 데이터 클러스터들이 존재하는 패턴 분류 문제에 적용할 경우 좋지 못한 패턴 분류 성능을 보인다고 알려져 있다[11]. 이를 극복하기 위하여 Local Fisher Discriminant Analysis가 M. Sugiyama에 의해 제안되었다. Sugiyama에 의해 제안된 LFDA는 아래와 같이 정의된 local within-class scatter matrix

$\widetilde{S}^{(w)}$ 와 local between-class scatter matrix $\widetilde{S}^{(b)}$ 을 이용하여 변환 벡터를 결정한다.

$$\widetilde{S}^{(w)} = \frac{1}{2} \sum_{i,j=1}^n \widetilde{W}_{i,j}^{(w)} (\mathbf{x}_i - \mathbf{x}_j)^T (\mathbf{x}_i - \mathbf{x}_j) \quad (6)$$

$$\text{여기서, } \widetilde{W}_{i,j}^{(w)} = \begin{cases} A_{i,j}/n_l & \text{if } y_i = y_j = l \\ 0 & \text{if } y_i \neq y_j \end{cases}$$

$$\widetilde{S}^{(b)} = \frac{1}{2} \sum_{i,j=1}^n \widetilde{W}_{i,j}^{(b)} (\mathbf{x}_i - \mathbf{x}_j)^T (\mathbf{x}_i - \mathbf{x}_j) \quad (7)$$

$$\text{여기서, } \widetilde{W}_{i,j}^{(b)} = \begin{cases} A_{i,j}(1/n - 1/n_l) & \text{if } y_i = y_j = l \\ 1/n & \text{if } y_i \neq y_j \end{cases}$$

하중 행렬 $A_{i,j}$ 는 다양한 함수들을 이용하여 정의할 수 있으며, 대표적으로 아래 함수를 이용하여 정의한다.

$$A_{i,j} = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma^2}\right) \quad (8)$$

여기서, σ 는 함수의 기울기를 결정하는 파라미터이다.

2.2 Local Linear Discriminant Analysis (LLDA)

앞에 보인 M. Sugiyama에 의해 제안된 Local Fisher Discriminant Analysis는 데이터들 사이의 지역적 관계를 고려하여 변환 벡터를 결정한다. LFDA에 의해 결정된 변환 벡터는 전체 입력공간에 대하여 하나의 선형식을 정의한다.

본 논문에서는 입력공간의 영역을 데이터 분포 분석을 통하여 여러 개의 지역 공간 (local area)으로 나누고, 나누어진 지역공간에 적합한 (9)과 같은 다수의 선형식을 얻고자 한다.

$$y^j = u_j(\mathbf{x})(\mathbf{w}^j)^T \mathbf{x} \quad (9)$$

여기서, y_j 는 j 번째 지역공간에서 정의된 feature 벡터를 의미하며, $u_j(\mathbf{x})$ 는 전체 입력 공간 중에서 j 번째 입력공간을 정의하는 함수이며, \mathbf{w}^j 는 j 번째 지역공간에서 정의된 변환 벡터이다.

전형적인 LDA와 마찬가지로 Local LDA의 목적함수는 (10)과 같이 정의 할 수 있다.

$$J_j = \frac{(\mathbf{w}^j)^T S_B^j (\mathbf{w}^j)}{(\mathbf{w}^j)^T S_W^j (\mathbf{w}^j)} \quad (10)$$

여기서, J_j 는 전체 입력 공간 중에서 j 번째 지역공간에서 정의된 목적함수를 의미하며, S_B^j 는 그 지역공간에서 정의된 between-class scatter matrix를 의미하며, S_W^j 은 지역공간에서 정의된 within-class scatter matrix를 말한다.

$$S_B^j = (m_1^j - m_2^j)(m_1^j - m_2^j)^T \quad (11-1)$$

$$m_i^j = \frac{1}{n_{ij}} \sum_{\mathbf{x}_k \in D_i} u_j(\mathbf{x}_k) \mathbf{x}_k \quad (11-2)$$

$$u_j(\mathbf{x}_k) = \begin{cases} 1, \mathbf{x}_k \in C_j \\ 0, \mathbf{x}_k \notin C_j \end{cases} \quad (11-3)$$

여기서, m_i^j 는 j 번째 지역공간에서 클래스 i 에 속한 데이터들의 평균 벡터를 의미하며, (11-2)와 같이 정의된다. n_{ij} 는 i 번째 클래스에 속한 데이터들의 집합인 D_i 의 원소들 중에서 j 번째 지역공간에 속한 요소들의 개수를 의미한다. C_j 는 j 번째 지역공간에 속한 데이터들의 집합을 의미한다.

$$S_{ij} = \sum_{\mathbf{x} \in D_j, \mathbf{x} \in C_j} (\mathbf{x} - \mathbf{m}_i^j)(\mathbf{x} - \mathbf{m}_i^j)^T \quad (12-1)$$

$$S_W^j = S_{1j} + S_{2j} \quad (12-2)$$

각각의 지역 공간 마다 정의된 목적함수를 최대화 하기위한 최적의 \mathbf{w}_j 는 (13)과 같이 구할 수 있다.

$$S_B^j \mathbf{w}_j = \lambda S_W^j \mathbf{w}_j \quad (13-1)$$

$$\mathbf{w}_j = (S_W^j)^{-1} (\mathbf{m}_1^j - \mathbf{m}_2^j) \quad (13-2)$$

본 논문에서는 지역공간을 정의하기 위하여 K-Means Clustering 알고리즘을 이용하여 입력 공간 분할을 수행하였다.

K-Means Clustering 알고리즘에 의해 분할된 지역 공간에 속한 데이터들의 집합 C_j 는 (14)과 같이 정의 할 수 있다.

$$C_j = \left\{ \mathbf{x}_i \mid \underset{k}{\operatorname{argmin}} \|\mathbf{x}_i - \mathbf{v}_k\| = j \right\} \quad (14)$$

여기서, \mathbf{v}_k 는 K-means clustering 방법에 의해 얻어진 지역공간의 중심 벡터를 의미한다.

본 논문에서는 clustering 알고리즘방법 뿐만 아니라 최적화 방법들 중의 하나인 차분진화 알고리즘을 이용하여 입력공간의 분할을 수행하였다. 다양한 최적화 알고리즘들 중에서 간단한 구조와 빠른 탐색 속도를 가진 차분 진화 연산 알고리즘[10]을 사용하여 지역공간의 중심 벡터 (prototype)의 위치를 결정한다.

위에서 설명한 local linear discriminant analysis 기법을 이용하여 패턴 분류를 하기 위하여 아래와 같은 규칙 베이스 패턴 분류기를 설계한다.

$$\begin{aligned} R_1 : IF \mathbf{x} \in C_1, \text{ then } y = \mathbf{w}_1^T \mathbf{x} \\ \text{and } \hat{L} = \begin{cases} 1, & \text{When } |y - \tilde{m}_1^1| < |y - \tilde{m}_2^1| \\ 2, & \text{ELSE} \end{cases} \\ \\ R_2 : IF \mathbf{x} \in C_2, \text{ then } y = \mathbf{w}_2^T \mathbf{x} \\ \text{and } \hat{L} = \begin{cases} 1, & \text{When } |y - \tilde{m}_1^2| < |y - \tilde{m}_2^2| \\ 2, & \text{ELSE} \end{cases} \\ \\ \vdots \\ R_c : IF \mathbf{x} \in C_p, \text{ then } y = \mathbf{w}_p^T \mathbf{x} \\ \text{and } \hat{L} = \begin{cases} 1, & \text{When } |y - \tilde{m}_1^p| < |y - \tilde{m}_2^p| \\ 2, & \text{ELSE} \end{cases} \end{aligned} \quad (15)$$

여기서, p는 지역공간의 개수를 의미하고, \hat{L} 은 입력 데이터 \mathbf{x} 가 속 할 것으로 추정되는 class의 label을 의미하며, \tilde{m}_1^j 는 j번째 지역공간에서 선형식 $y = \mathbf{w}_j^T \mathbf{x}$ 에 의해 변환된 y값들의 중심 값을 의미하며 (16)과 같이 나타낸다.

$$\begin{aligned} \tilde{m}_1^j &= \frac{1}{n_{1j}} \sum_{\mathbf{x} \in C_j} y = \frac{1}{n_{1j}} \sum_{\mathbf{x} \in C_j} \mathbf{w}_j^T \mathbf{x} \\ &= \mathbf{w}_j^T \left(\frac{1}{n_{1j}} \sum_{\mathbf{x} \in C_j} \mathbf{x} \right) = \mathbf{w}_j^T \cdot \mathbf{m}_1^j \end{aligned} \quad (16)$$

3. 차분진화 연산 알고리즘을 이용한 지역 공간 분할

차분진화 알고리즘(differential evolution algorithm: DE)은 Price와 Storn에 의해 벡터 차분(vector differential)을 사용하여 Chebychev 다항곡선의 내삽문제(polynomial fitting problem)를 해결하는 과정에서 개발되었다. 차분진화 알고리즘은 현재 널리 사용되고 있는 통계적 임의탐색법인 유전 알고리즘(genetic algorithm: GA)과 유사한 알고리즘이다[13]. 유전 알고리즘과 차분진화 알고리즘은 초기 개체군을 이루는 개체를 교배(crossover), 돌연변이(mutation), 선택(selection) 과정을 거쳐 적합도(fitness)가 개선되는 과정을 거쳐 우수한 성능을 지닌 개체들을 추출해 낸다는 공통점이 있다. 그러나 이런 연산 과정에서 GA 알고리즘이 개체들의 표현형(phenotype)을 유전형(genotype)으로 바꾸는 코딩(coding)이 필요한 반면, 차분진화 알고리즘은 코딩 과정이 필요 없으며 개체를 벡터로 표현하기 때문에 이들의 산술적 연산을 통해 새로운 개체를 생성한다는 이점이 있다. 또한 유전 알고리즘의 단점이었던 너무 복잡한 구조와 연산 대신 간단하고 합리적인 구조와 빠른 탐색 속도가 장점이다[12].

차분진화 알고리즘에서는 초기 개체군을 이루는 모든 개체들을 벡터로 표현하는데, 이들을 파라미터 벡터라고 표현한다. 집단 크기는 진화 과정에서 변하지 않는다.

$$v_{jG} \in R^D \quad j = 1, 2, \dots, NP \quad (17)$$

여기서 NP는 파라미터 벡터의 개수이고, 이 값은 최적화 과정에서 일정하며, 설계자가 문제에 따라 정의할 수 있는 샘플링 개수이다. G는 파라미터 벡터가 몇 번째 세대인가를 표현하며 알고리즘에서는 최적화 횟수(Generation)를 의미한다. 파라미터 벡터의 차원은 목적함수를 이루는 설계변수의 개수 D와 같으며, 초기 파라미터는 벡터 가용 영역(feasible region)에서 임의로 선택하되, 일반적으로 균등 확률 분포를 따르도록 한다.

차분 진화는 두 개의 개체 벡터의 차이에 가중치를 곱한 것을 세 번째 개체 벡터에 더해서 교배용 벡터를 발생한다. 발생한 교배 벡터와 교배 대상 벡터가 교배되어 새로운 벡터가 얻어진다. 이 새로운 벡터의 목적함수 값이 교배 대상 벡터의 것 보다 좋으면 이 벡터로 교배 대상 벡터를 대체한다. 이는 유전자 알고리즘에서 사용되는 엘리트 전략과 유사한 효과를 거둘 수 있다. 집단에서 선택된 벡터의 거리와 방향 정보를 사용하여 벡터를 랜덤하게 변화시키는 것이 차분 진화로 하여금 뛰어난 수렴성을 갖도록 해준다.

교배용 벡터는 벡터 $x_{i,G}(i=1,2,\dots,Popsiz)$ 집단에서 랜덤하게 선택된 서로 다른 3개의 벡터로부터 (18)와 같이 만든다.

$$v_{i,G+1} = x_{r_1,G} + F \cdot (x_{r_2,G} - x_{r_3,G}) \quad (18)$$

여기서 r_1, r_2 와 r_3 은 $[0, Popsiz - 1]$ 범위의 서로 다른 정수이고, $0 \leq F \leq 2, i \neq r_1, r_2, r_3$ 이다. F는 선택된 교배 대상 벡터들의 차이($x_{r_2,G} - x_{r_3,G}$)를 제어하는 파라미터로 0과 2사이의 값을 갖는다. $v_{i,G+1}$ 을 발생하기 위해 선택된 벡터 $x_{r_1,G}$ 는 $v_{i,G}$ 와 관련이 없는 집단에서 랜덤하게 선택된 개체이다.

차분 변화를 통해 얻어질 벡터는 (18)의 교배용 벡터와 교배 대상 벡터인 $x_{i,G}$ 를 교배하여 만들어진 것이다. 생성된 벡터의 다양성을 증가시키기 위해 주로 균일 교

배과정을 (19)과 같이 이루어진다.

$$x'_{i,G+1} = (x_{1i,G+1}, x_{2i,G+1}, \dots, x_{di,G+1}) \quad (19)$$

$$x'_{ji,G+1} = (v_{ji,G+1} * x_{ji,G}) \quad j=1, \dots, d \quad (20)$$

여기서 *는 교배 연산자로 균일 교배가 주로 사용된다. 균일 교배는 (21)과 같다.

$$x'_{ji,G+1} = \begin{cases} v_{ji,G+1} & \text{if rand} \leq C_r \\ x_{ji,G} & \text{otherwise} \end{cases} \quad (21)$$

여기서, rand는 0과 1사이의 랜덤 변수이고, C_r 은 교배율 이다. 교배 후 만들어진 개체의 생존 여부는 $x'_{i,G+1}$ 와 $x_{i,G}$ 의 성능 비교를 통해 이루어진다.

이러한 과정은 설계자가 정해준 세대 횟수만큼 반복하며, 결국 목적 함수 값이 작은 파라미터 벡터들로 구성된 개체군이 만들어진다. 이 벡터 중 가장 작은 목적 함수 값이 전역해가 된다.

차분 진화에서는 실수를 직접 사용하기 때문에 개체 평가 시에 유전자형을 표현형으로 변환할 필요가 없다. 다음 세대를 위한 개체 선정이 적합도 비교에 의해서만 이루어지기 때문에 일반적인 진화알고리즘에서 필요한 선택 메커니즘과 최대 또는 최소화 문제에 따른 적합도 함수로의 변환이 필요 없다. 이는 알고리즘의 처리 속도 개선에 유리하다.

차분진화의 실행 단계를 정리하면 다음과 같다.

[Step 1] 초기 집단 구성 : 랜덤 값으로 N개의 개체를 초기화. 각 개체는 n개의 목적변수로 구성, t=0)

[Step 2] 집단내의 모든 개체의 목적함수 평가

[Step 3] 모든 개체 (i = 1, ..., N)에 대해 차분 변화를 위한 개체 a_{r1} , a_{r2} 와 a_{r3} 을 선택하여 교배용 벡터를 만들고 이를 교배 대상 벡터와 교배

[Step 4] 모든 개체의 목적함수 평가

[Step 5] 종료조건을 확인하고 종료조건이 만족되지 않으면 t=t+1로 하고 [Step 3]으로 복귀

그림 1은 전체적인 진화 연산 알고리즘의 순서도를 보인다.

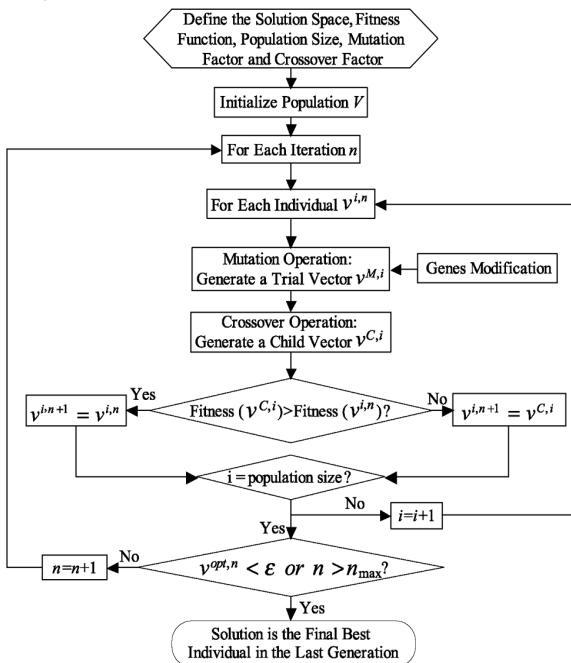


그림 1 차분 진화 알고리즘의 순서도
Fig. 1 Flowchart of Differential Evolution Algorithm

Local Linear Discriminant Analysis의 패턴 분류 성능을 개선하기 위하여 앞에서 설명한 차분 진화 연산 알고리즘을 사용하여 최적의 prototype을 결정한다. 차분 진화 알고리즘을 이용하여 prototype의 위치를 결정하기 위한 차분 진화 알고리즘의 파라미터 벡터들의 구조는 그림 2와 같다[17].

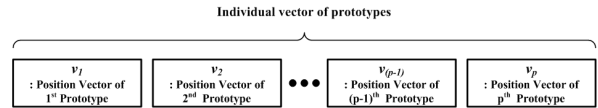


그림 2 초기 파라미터 벡터의 구조
Fig. 2 Construction of initial parameter vectors

차분 진화 알고리즘의 초기화 단계에서는 모든 벡터에서 각 결정 파라미터는 랜덤으로 할당되고, 돌연변이 과정을 거쳐 재생산과 선택과정을 거친다. 그림 2에 보인 prototype의 최적화를 위한 차분 진화 알고리즘의 초기 파라미터 벡터는 표 1에 보인 prototype의 위치 벡터인 $\{v_1, v_2, \dots, v_{p-1}, v_p\}$ 을 표현한다.

4. 실험 및 결과 고찰

본 논문에서는 기존의 Fisher Discriminant Analysis (FDA)와 FDA의 약점을 보완하기 위해 제안된 Local Fisher Discriminant Analysis(LFDA)와 달리 입력공간을 clustering 알고리즘 또는 최적화 알고리즘을 이용하여 분할하고 분할된 지역공간에 local linear discriminant analysis를 적용함으로써 다수개의 변환 벡터를 얻을 수 있어서 하나의 변환 벡터를 얻게 되는 FDA 또는 LFDA와는 다른 특성을 가진다.

본 연구에서는 제안된 분류기의 성능을 평가하기 위하여 제안된 분류기를 기계 학습 데이터 집합(machine learning dataset)에 적용하여 분류기로서의 성능을 평가 및 분석한다. 기계 학습 데이터 집합은 UCI machine learning repository(<http://archive.ics.uci.edu/ml/>)에서 획득한 5가지의 데이터 집합들을 이용하여 기존 논문에서 제안된 분류기 기법들과 비교 평가한다.

제안된 분류기의 성능 평가를 위하여 5 fold cross-validation 기법을 5회 적용하여 실험을 수행하게 된다. 제안된 제어기의 분류 성능 평가를 위한 성능 평가 지수는 식(22)과 같이 패턴 분류기의 오분류 비율(misclassification rate)로 한다.

$$\text{오분류율} = \frac{\sum_{i=1}^N g(L_i - \hat{L}_i)}{N} \times 100 \quad (22)$$

여기서, N은 데이터의 수를 의미한다.

$$g(L_i - \hat{L}_i) = \begin{cases} 1, & L_i \neq \hat{L}_i \\ 0, & L_i = \hat{L}_i \end{cases} \quad (23)$$

제안된 분류기를 설계할 때 필요한 파라미터들을 표 2에 나타내었다. 제안된 패턴 분류기는 MATLAB를 사용하여 구현하여 실험 하였다.

표 1. 제안된 분류기의 설계 파라미터
Table 1. Design parameters of the proposed classifier

Number of local areas (p)	2 and 3
Number of Individuals	100
Number of Generation	50
Crossover rate	1.0

제안된 분류기의 패턴 분류 성능을 비교 평가하기 위해 기존 분류기의 benchmark data로 이용되는 기계 학습 데이터 집합에 적용하였다. 다양한 종류의 기계 학습 데이터 집합들 중에 5개의 기계 학습 데이터 집합을 선정하여 실험을 수행 하였다. 선정된 데이터 집합에 대한 일반적인 설명을 표 2에 나열하였다.

전체 입력 공간은 p개의 지역공간으로 분할하였으며, 차분진화 알고리즘의 개체 수는 100개, 진화 세대는 50 세대를 사용하였고, 교배율은 일반적인 유전자 알고리즘과 달리 차분 진화 알고리즘에서 많이 사용하는 1.0으로 설정하여 실험하였다. 본 실험에서는 실험 조건을 간단하게 하기 위하여 차분진화 알고리즘의 설계 파라미터와 제시된 데이터 집합의 분류 문제 사이의 연관관계가 없다고 가정하였다. 이러한 가정으로 인해 동일한 설계 파라미터를 가진 차분 진화 알고리즘을 제시된 5개의 machine learning 데이터 집합의 패턴 분류 문제에 적용하였다.

표 3은 5개의 기계 학습 데이터 집합의 경우에 제안된 분류기와 다른 잘 알려진 분류기들과의 성능을 비교한 것이다. 표 3에 나타난 바와 같이 제안된 패턴 분류기중 차분 진화 연산 알고리즘을 이용하여 지역분할을 수행한 패턴 분류기가 모든 데이터 집합에서 가장 우수한 패턴 분류 성능을 보인다.

표 2. 실험에 사용된 machine learning dataset
Table 2. Machine learning dataset used in the experiment

Dataset	Number of features	Number of patterns
Australian	14	690
Liver	6	345
German	24	1000
Heart	13	270
Diabetes	8	768

표 3. 5 fold cross-validation을 이용한 제안된 분류기와 다른 분류기와의 분류 성능 비교

Table 3. Comparison of the classification error of the proposed classifier and other methods using 5 fold cross-validation

	Australian	Liver	German	Heart	Diabetes
Proposed Classifier (K-Means Clustering)	18.49±2.42 (p=2)	30.14±2.04 (p=3)	29.44±0.67 (p=2)	26.37±1.03 (p=2)	27.11±0.36 (p=3)
Proposed Classifier (Differential Evolution)	11.97±0.33 (p=2)	24.35±1.30 (p=3)	24.20±0.49 (p=2)	12.30±1.06 (p=2)	20.49±0.46 (p=2)
Generic LDA	14.06±0.27	33.80±0.53	28.24±0.35	16.07±0.42	24.40±0.57
LVQ1 [14]	33.3	32.5	30.1	36.0	26.4
LVQ2 [14]	34.0	32.7	28.5	34.0	25.8
LVQ3 [14]	31.1	33.6	28.7	34.0	26.0
SVM1 [14]	35.0	31.9	30.0	39.2	24.3
SVM2 [14]	36.1	42.0	29.2	34.5	24.1
SVM3 [14]	44.5	40.8	30.0	44.5	34.9
LFDA [11]	NA	NA	29.9±2.8	21.9±3.7	32.0±2.5

표 4는 다른 잘 알려진 분류기들과의 성능을 비교하기 위하여 10 fold cross-validation을 수행한 실험결과를 보인 것이다.

표 4. 10 fold cross-validation을 이용한 제안된 분류기와 다른 분류기와의 분류 성능 비교

Table 4. Comparison of the classification error of the proposed classifier and other methods using 10 fold cross-validation

	Liver	Sonar	Diabetes
Proposed Classifier (Differential Evolution)	22.03±6.28	13.57±9.24	18.10±4.43
LBDA[15]	32.41	13.58	23.84
LFDA[15]	38.38	12.20	31.79
AMPSONN[16]	34.75	NA	24.95

*LBDA: Linear Boundary Discriminant Analysis. AMPSONN: Adaptive Michigan PSO based Nearest Neighbors Classifier

5. 결론

본 논문에서는 전형적인 Linear Discriminant Analysis를 변경하여 여러 개의 지역공간들로 분할된

공간에서 Local Linear Discriminant Analysis를 기반으로 하는 패턴 분류 규칙을 만들고, 그 규칙을 최적화하기 위하여 차분 진화 알고리즘을 이용한 전체 입력 공간 분할 방법을 제안 하였다.

제안된 패턴 분류기 설계 방법에 따라 설계된 local linear discriminant analysis 기반 패턴 분류 규칙의 분류 성능을 평가하기 위하여 다양한 기계학습 데이터 집합에 적용하였다.

제안된 패턴 분류기를 기계학습 데이터 집합에 적용한 실험 결과에 따르면, 5가지의 기계학습 데이터 집합 모두에서 우수한 패턴 분류 성능을 보이는 것을 알 수 있었다.

참 고 문 헌

[1] C.M. Bishop, Pattern Recognition and Machine Learning, Springer, Berlin, 2006.

[2] R. Q. Duda, P. E. Hart, D. G. Stork, Pattern Classification, second ed., Wiley, New York, 2001.

[3] K. Fukunaga, Introduction to Statistical Pattern Recognition, second ed., Academic Press, New York, 1990.

[4] G. Kowalski, Information Retrieval Systems: Theory and Implementation, Kluwer Academic Publishers, Dordrecht, 1997.

[5] W. B. Frakes, R. Baeza-Yates, Information Retrieval: Data Structures and Algorithms, Prentice-Hall PTR, 1992.

[6] Z. Jin, J. Y. Yang, Z. S. Hu, Z. Lou, "Face recognition based on the uncorrelated discriminant transformation," Pattern Recognition, vol. 34, pp. 1405 - 1416, 2001.

[7] D. L. Swets, J. Weng, "Using discriminant eigenfeatures for image retrieval," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18 pp. 831 - 836, 1996.

[8] X. Li, W. Hu, H. Wang, and Z. Zhang, "Linear discriminant analysis using rotational invariant L1 Norm," Neurocomputing, pp. 2571-2579, 2010.

[9] K. Fukunaga, J.M. Mantock, "Nonparametric discriminant analysis," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 5, pp. 671 - 678, 1983.

[10] H. Brunzell, J. Eriksson, "Feature reduction for classification of multidimensional data," Pattern Recognition, vol. 33, pp. 1741 - 1748, 2000.

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

[12] R. Storn, K. V. Price, "Differential Evolution—a fast and efficient heuristic for global optimization over continuous spaces," Journal of Global Optimization, vol. 11, pp. 341-359, 1997.

[13] Z. Michalewicz, "Genetic Algorithms + Data Structures = Evolution Programs," Springer-Verlag, Berlin Heidelberg, 1996.

[14] R. Parades, E. Vidal, "Learning prototypes and distance: A prototype reduction technique based on nearest neighbor error minimization," Pattern Recognition, vol. 39, pp. 180-188, 2006.

[15] J. H. Na, M. S. Park, and J. Y. Choi, "Linear boundary discriminant analysis," Pattern Recognition, vol. 43, pp. 929 - 936, 2010.

[16] A. Cervantes, I. M. Galván, and P. Isasi, "AMPSO: A New Particle Swarm Method for Nearest Neighborhood Classification," IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART B: CYBERNETICS, vol. 39, no. 5, pp. 1081-1091, 2009.

[17] 노석범, 안태천, "차분 진화 알고리즘을 이용한 Nearest Prototype Classifier 설계," 한국지능시스템학회 논문지, Vol. 21, No. 4, pp. 487-492, 2011.

저 자 소 개



노 석 범 (Seok-Beom Roh)

1994년 원광대학교 제어계측공학과 졸업.
1996년 동 대학원 컴퓨터공학과 석사.
2006년 동 대학원 제어계측공학과 박사.
관심분야 : 퍼지 이론, 신경 회로망, Bio-inspired optimization algorithm, Pattern Recognition
E-mail : nado@wku.ac.kr



황 은 진 (Eun-Jin Hwang)

2011년 원광대학교 법학과 졸업.
2011년~현재 동 대학원 제어계측공학과 석사과정.
관심분야 : 퍼지 이론 및 응용, 지능 회로망, Pattern Recognition, 태양광
E-mail : 10azalea@hanmail.net



안 태 천 (Tae-Chon Ahn)

1978년 연세대학교 전기공학과 졸업
1980년 동 대학원 전기공학과 석사.
1986년 동 대학원 전기공학과 박사.
1981년~현재 원광대학교 전기전자 및 정보공학부 교수.
관심분야 : 퍼지 이론 및 응용, 신경망, Bio-inspired optimization algorithm,

Intelligent Pattern Recognition, 태양광.

Phone : 063-850-6344

E-mail : tcahn@wku.ac.kr, tcahn1024304@naver.com