

## 잡음 민감성이 개선된 변형 퍼지 주성분 분석 기법

김성훈\*, 허경용\*\*, 우영운\*\*\*

### A Variant of Improved Robust Fuzzy PCA

Seong Hoon Kim\*, Gyeongyong Heo\*\*, Young Woon Woo\*\*\*

#### 요약

주성분 분석(PCA)은 차원 축소와 특징 추출을 위해 널리 사용되는 기법 중의 하나이지만 자승 오류의 사용으로 인해 잡음에 민감한 단점이 있다. 이러한 잡음 민감성을 개선하기 위해 다양한 방법이 소개되었고 그 중 improved robust fuzzy PCA(RF-PCA2)는 퍼지 소속도를 이용한 반복적 최적화 기법으로 다른 방법에 비해 우수한 성능을 보였다. 하지만 RF-PCA2 역시 국부적인 최적해에 빠질 수 있으며 그 원인 중 하나는 RF-PCA2 알고리즘이 소속도를 균일한 값으로 초기화시키기 때문이다. 또한 퍼지 소속도를 사용하고 있지만 여전히 목적함수가 자승 오류 최소화에 기초하고 있다는 사실도 그 원인이 된다. 이 논문에서는 RF-PCA2의 이러한 문제점을 개선한 RF-PCA3를 제안한다. 제안하는 알고리즘은 RF-PCA2의 목적 함수를 바탕으로 하고 있다. 여기에 PCA의 목적 함수를 추가하고 초기 소속도 값을 데이터의 분포로부터 계산함으로써 전역 최적해에 가까운 해를 얻을 수 있는 가능성을 높여준다. 이러한 사실들은 실험 결과를 통해 확인할 수 있다.

▶ Keyword : 주성분 분석, 잡음 민감성, 국부 최적해

#### Abstract

Principal component analysis (PCA) is a well-known method for dimensionality reduction and feature extraction. Although PCA has been applied in many areas successfully, it is sensitive to outliers due to the use of sum-square-error. Several variants of PCA have been proposed to resolve the noise sensitivity and, among the variants, improved robust fuzzy PCA (RF-PCA2) demonstrated promising results. RF-PCA2, however, still can fall into a local optimum due to equal initial membership values for all data points. Another reason comes from the fact that RF-PCA2 is based on sum-square-error although fuzzy memberships are incorporated. In this paper, a variant of RF-PCA2 called RF-PCA3 is proposed. The proposed algorithm is based on the objective function of RF-PCA2. RF-PCA3 augments RF-PCA2 with the objective function of PCA and initial membership calculation using data distribution, which make RF-PCA3 to have more chance to converge on a better solution than that of RF-PCA2. RF-PCA3 outperforms RF-PCA2, which is demonstrated by experimental results.

▶ Keyword : Principal Component Analysis, Noise Sensitivity, Locally Optimal Solution

• 제1저자 : 김성훈 교신저자 : 우영운

• 투고일 : 2010. 09. 14, 심사일 : 2010. 10. 21, 게재확정일 : 2010. 11. 09.

\* 경북대학교 컴퓨터정보학부(Division of Computer Information, Kyungpook National Univ.) \*\* 동의대학교 영상미디어센터(Visual Media Center, Dong-Eui Univ.) \*\*\* 동의대학교 멀티미디어공학과(Dept. of Multimedia Eng., Dong-Eui Univ.)

## I. 서론

주성분 분석(principal component analysis, PCA)은 차원 축소를 위해 널리 사용되는 방법 중 하나로, 자승 오차(squared error) 또는 재건 오류(reconstruction error)의 합을 최소화하는 직교 변환이다[1]. PCA는 여러 영역에 성공적으로 사용되었지만, 자승 오차 합을 사용함으로써 잡음에 민감하다는 문제점이 있다. PCA의 잡음 민감성을 줄이기 위해 여러 가지 방법이 제안되었고, 이들은 부분 집합을 이용하는 방법과 퍼지 소속도(membership)를 이용하는 방법의 두 가지로 크게 나누어볼 수 있다. 부분 집합을 이용하는 방법은 하나 이상의 데이터 부분 집합에 대해 PCA를 수행하고 이들 결과를 이용하여 전체 데이터의 주성분을 구해낸다[2][3][4][5]. 퍼지 소속도를 이용하는 방법은 전형적인 데이터 포인트에만 큰 소속도 값을 할당함으로써 잡음 민감성을 줄인다[6][7][8]. 이들 PCA의 변형 중 improved robust fuzzy PCA (RF-PCA2)는 소속도와 주성분을 반복적으로 최적화하는 과정을 통해 잡음에 민감하지 않은 주성분을 구해내는 방법으로 다른 방법들에 비해 우수한 성능을 보여주었다. 하지만 RF-PCA2 역시 국부 최적해만을 보장하는 문제점이 있다[9].

RF-PCA2의 국부 최적해 문제는 여러 가지 원인이 있으며 RF-PCA2의 초기 소속도 값이 모든 데이터 포인트에 동일하게 주어지는 것이 그 중 하나이다. 모든 데이터 포인트에 동일한 소속도 값을 할당함으로써 RF-PCA2는 PCA를 이용하여 초기화하는 효과를 가져오며 이는 잡음에 민감한 국부 최적해에 빠지도록 하는 원인이 된다. 또 다른 이유로는 RF-PCA2의 목적함수가 소속도를 포함하고 있지만 여전히 자승 오류의 합을 사용한다는 점이다. 이러한 문제점을 해결하기 위해 이 논문에서는 RF-PCA2를 개선한 새로운 방법을 제안한다. 제안하는 방법은 먼저 초기화 문제를 해결하기 위해 초기 소속도 값을 추정하여 사용한다. 잡음 또는 특이값(outlier)을 판별하기 위한 많은 방법이 소개되었고 성공적으로 사용되고 있지만[10] 이 논문에서는 초기 소속도 값으로 특이값에 작은 값을 할당해줄 수 있는 간단한 방법으로 충분하다. 이 과정에서 특이값이 아닌 포인트에 작은 소속도 값이 할당될 수 있지만 이는 반복적인 최적화 과정을 통해 수정될 수 있다. 중요한 점은 RF-PCA2에서와는 다른 출발점을 가지도록 함으로써 보다 나은 해에 수렴할 수 있는 기회를 증가시키도록 한다는 점이다.

자승 오류의 합을 사용하는 한계점은 완전히 극복될 수는

없지만 PCA의 목적함수에 해당하는 항을 첨가함으로써 완화시킬 수 있다. PCA와 RF-PCA2는 기본적으로 자승 오류의 합이 최소화되도록 한다. 제안하는 방법에서는 자승 오류의 합이 증가하는 것을 일부 허용함으로써 RF-PCA2에 비해 나은 결과를 얻을 수 있다. 이러한 사실들은 실험 결과를 통해 확인할 수 있다.

이 논문의 구성은 다음과 같다. 다음 장에서는 RF-PCA2를 간략히 소개하고 3장에서는 RF-PCA2의 목적함수 수정과 초기화 방법을 도입한 RF-PCA3를 제안한다. 실험 결과는 4장에서 보이며 결론 및 향후 연구 방향은 5장에서 언급한다.

## II. RF-PCA2

RF-PCA2는 잡음에 강건한  $k$ 개의 정규직교(orthonormal) 벡터 즉, 주성분(principal component)을 찾아내기 위한 반복 알고리즘이다. 데이터 집합  $X = \{x_1, \dots, x_N\}$ 가 주어진 경우 RF-PCA2의 목적 함수는 식 (1)과 같이 나타낼 수 있다.

$$\begin{aligned} \operatorname{argmin}_W J = & \sum_{i=1}^N u_i \| (x_i - \mu_R) - WW^T(x_i - \mu_R) \|^2 \dots\dots\dots (1) \\ & + \sigma^2 \sum_{i=1}^N (u_i \log u_i - u_i) \\ = & \sum_{i=1}^N u_i e(x_i) + \sigma^2 \sum_{i=1}^N (u_i \log u_i - u_i) \\ \text{s.t.} \quad & W^T W = I \end{aligned}$$

이 때  $u_i$ 는 데이터 포인트  $x_i$ 의 소속도를,  $\mu_R$ 은 가중치 평균을,  $e(x_i)$ 는 재건 오류[11]를,  $W$ 는  $k$ 개 단위 직교 벡터를 열로 가지는 행렬을,  $I$ 는 단위 행렬을 나타내며  $\sigma$ 는 regularization 상수이다.  $k$ 는 재건에 사용될 직교 벡터의 개수를 나타내며 알려진 것으로 가정한다. 식 (1)은 두 개의 항으로 구성되어 있다. 첫 번째 항은 소속도를 가중치로 가지는 재건 오류의 합을 나타내고 두 번째 항은 주성분의 잡음 민감성을 줄이기 위한 regularization 항이다.

먼저 식 (1)을  $u_i$ 에 대해 편미분하면 식 (2)를 얻을 수 있다.

$$\frac{\partial J}{\partial u_i} = e(x_i) + \sigma^2 \log u_i = 0 \dots\dots\dots (2)$$

식 (2)를  $u_i$ 에 대해 정리하면 식 (3)의 소속도에 대한 update equation을 얻을 수 있다.

$$u_i = \exp\left(-\frac{e(x_i)}{\sigma^2}\right) \dots\dots\dots (3)$$

유사하게, 식 (1)을  $\mu_R$ 에 대해 편미분하면 식 (4)를 얻을 수 있다. 이 때  $I - WW^T$ 는 역행렬이 존재하여야 한다.

$$\frac{\partial J}{\partial \mu_R} = \sum_{i=1}^N u_i (-x_i + \mu_R) = 0 \dots\dots\dots (4)$$

식 (4)로부터 가중치 평균에 대한 update equation인 식 (5)를 얻을 수 있다.

$$\mu_R = \frac{\sum_{i=1}^N u_i x_i}{\sum_{i=1}^N u_i} \dots\dots\dots (5)$$

가중치 평균을 이용하여 주어진 데이터 집합은 영의 가중치 평균을 가지도록 만들 수 있고,  $x'_i = x_i - \mu_R$ , 이 경우 식 (1)을 최적화하는  $W$ 를 구하는 것은 식 (6)을 최적화하는 것과 같아진다.

$$J' = \sum_{i=1}^N u_i \|x'_i - WW^T x'_i\|^2 \dots\dots\dots (6)$$

$W^T W = I$ 의 제약 조건 하에서 식 (6)을 최소화하는  $W$ 는 식 (7)로 정의되는 가중치 분산 행렬(weighted covariance matrix)에서  $k$ 개의 최대 고유치를 가지는 고유 벡터로 구성할 수 있다.

$$C_R = \frac{1}{N} \sum_{i=1}^N u_i x'_i x'^T_i \dots\dots\dots (7)$$

$$= \frac{1}{N} \sum_{i=1}^N u_i (x_i - \mu_R)(x_i - \mu_R)^T$$

이에 관한 증명은 참고문헌 [9]을 참조하면 된다.

RF-PCA2 알고리즘은 그림 1과 같이 요약할 수 있으며, 이 때  $\epsilon$ 은 미리 정해지는 상수이고  $t_{\max}$ 는 최대 반복 횟수를 나타낸다.

1. 소속도  $U_0 = [u_1, \dots, u_N]^T = [1, \dots, 1]^T$ 와 카운터  $t = 0$ 를 초기화한다.
  2. **repeat**
  3.  $t \leftarrow t + 1$
  4. 평균  $\mu_R$ 과 분산 행렬  $C_R$ 을 구한다. (식 (5), (7))
  5.  $C_R$ 의 고유 벡터로  $W_t$ 를 구성한다.
  6. 식 (3)으로 소속도  $U_t$ 를 구한다.
  7. **until**  $J^{t-1} - J^t < \epsilon$  또는  $t > t_{\max}$  8. **return**

그림 1. RF-PCA2 알고리즘  
Fig. 1. RF-PCA2 algorithm

### III. RF-PCA3

비록 RF-PCA2가 이전 방법들에 비해 나은 결과를 보였지만 여전히 국부 최적해에 빠질 수 있다. 이를 개선하기 위해 이 논문에서는 새로운 PCA의 변형 RF-PCA3를 제안하며 이는 RF-PCA2와 두 가지 측면에서 다르다. 첫 번째는 목적함수로 RF-PCA3의 목적함수는 식 (8)과 같이 정의된다.

$$\begin{aligned} \operatorname{argmin}_W J = & \sum_{i=1}^N u_i \| (x_i - \mu_R) - WW^T (x_i - \mu_R) \|^2 \\ & + \sigma^2 \sum_{i=1}^N (u_i \log u_i - u_i) \\ & - \alpha \sum_{i=1}^N \| (x_i - \mu_R) - WW^T (x_i - \mu_R) \|^2 \\ = & \sum_{i=1}^N (u_i - \alpha) e(x_i) + \sigma^2 \sum_{i=1}^N (u_i \log u_i - u_i) \end{aligned}$$

s.t.  $W^T W = I$

..... (8)

식 (8)은 식 (1)에 세 번째 항, PCA의 목적 함수를 첨가한 것이다. 이 항은 음의 값을 가지며 지능 오류의 합이 커지는 것을 일부 허용하는 역할을 하므로 지능 오류의 합을 최소화함으로써 발생하는 잡음 민감성을 완화시키는 역할을 한다. 이 때  $\alpha (> 0)$ 는 첫 번째와 세 번째 항 사이의 비율을 결정하는 역할을 하며 이 논문에서는 실험적으로 결정된 값 0.5를 사용하였다.

RF-PCA3의 update equation은 2장에서와 유사하게 폐형식으로 구할 수 있다. RF-PCA3에서 소속도의 update equation은 식 (3)과 동일하고 평균과 분산 행렬은 식 (9)와 (10)으로 각각 정의된다.

$$\mu_R = \frac{\sum_{i=1}^N (u_i - \alpha) x_i}{\sum_{i=1}^N (u_i - \alpha)} \dots\dots\dots (9)$$

$$C_R = \frac{1}{N} \sum_{i=1}^N (u_i - \alpha) x_i' x_i'^T \dots\dots\dots (10)$$

$$= \frac{1}{N} \sum_{i=1}^N (u_i - \alpha) (x_i - \mu_R)(x_i - \mu_R)^T$$

두 번째 RF-PCA2와 다른 점은 초기화 방법이다. RF-PCA2는 그림 1에서와 같이 모든 데이터 포인트에 대하여 동일한 값으로 소속도를 초기화한다. 하지만 이는 국부 최적해에 빠지게 하는 원인이 되므로 RF-PCA3에서는 데이터 분포를 고려하여 초기값을 계산하는 방법을 사용한다.

PCA는 가우시안 분포를 가정하고 있으므로 전형적인 데이터 포인트는 밀집된 영역에 존재하는 포인트로 볼 수 있다. 따라서 이들 전형적인 포인트는 다른 포인트들과 가까운 거리에 존재하며 식 (11)의 다른 포인트들과의 거리 합은 전형성 (typicality)을 나타내는 척도로 사용될 수 있다.

$$d_i = \sum_{j=1}^N \|x_i - x_j\|^2 \dots\dots\dots (11)$$

식 (11)로 구해진 거리의 합을 이용하여 식 (12)는 각 데이터 포인트의 소속도를 계산한다.

$$u_i = \exp\left(-\frac{d'_i}{\sigma_d}\right) \dots\dots\dots (12)$$

이 때  $d'_i$ 는 식 (13)과 같이 계산된다.

$$d'_i = \max(d_i - \bar{d}, 0) \dots\dots\dots (13)$$

식 (13)에서  $\bar{d}$ 는 거리 합의 평균으로 식 (14)로 정의된다.

$$\bar{d} = \frac{1}{N} \sum_{i=1}^N d_i \dots\dots\dots (14)$$

식 (13)은 거리 합이 평균 거리 합보다 작은 경우 거리를 영으로 설정함으로써 소속도 값이 일이 되도록 해준다. 식 (12)에서  $\sigma_d$ 는 소속도 값의 감쇄를 결정하는 상수로 이 논문에서는 식 (15)에서와 같이 거리 합의 표준편차를 사용하였다.

$$\sigma_d^2 = \frac{1}{N} \sum_{i=1}^N (d_i - \bar{d})^2 \dots\dots\dots (15)$$

RF-PCA3 알고리즘은 그림 1과 동일한 구조를 가지지만, 단계 1에서 초기화 방법이 다르고 단계 4에서 식 (9)와 (10)을 사용하는 점이 다르다. 그림 2는 RF-PCA3 알고리즘을 요약한 것이다.

1. 카운터  $t = 0$ 를 초기화하고 식 (12)로 초기 소속도를 계산한다.
2. **repeat**
3.  $t \leftarrow t + 1$
4. 평균  $\mu_R$ 과 분산 행렬  $C_R$ 을 구한다. (식 (9), (10))
5.  $C_R$ 의 고유 벡터로  $W_t$ 를 구성한다.
6. 식 (3)으로 소속도  $U_t$ 를 구한다.
7. **until**  $J^{t-1} - J^t < \epsilon$  또는  $t > t_{\max}$
8. **return**

그림 2. RF-PCA3 알고리즘  
Fig. 2. RF-PCA3 algorithm

### IV. 실험 결과 및 고찰

RF-PCA3가 RF-PCA2에 비해 나은 결과를 보인다는 것을 확인하기 위해 이 장에서는 RF-PCA2와 RF-PCA3를 Matlab으로 구현하여 실험하였다. 그림 3은 실험에 사용한 데이터와 PCA 및 RF-PCA3를 사용하여 찾아낸 첫 번째 주성분을 표시한 것이다. 실험에 사용한 데이터는 가우시안 분포에서 생성하고 여기에 인위적으로 잡음을 첨가하여 구성하였다. 그림 3은 실험에 사용한 데이터에 PCA와 RF-PCA3를 이용하여 구한 첫 번째 주성분을 표시한 것이다.

RF-PCA2는 RF-PCA3와 거의 동일한 주성분을 찾아냈기 때문에 그림 3에서는 표시하지 않았다. 그림 3에서 알 수 있듯이, PCA의 경우 잡음에 의해 주성분의 방향이 잡음 쪽으로 일부 치우친 것을 확인할 수 있다. 하지만 RF-PCA3는 RF-PCA2와 마찬가지로 잡음에 작은 소속도 값을 할당함으로써 잡음에 민감하지 않은 주성분을 찾아낼 수 있다.

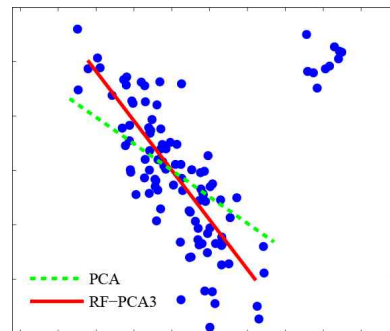
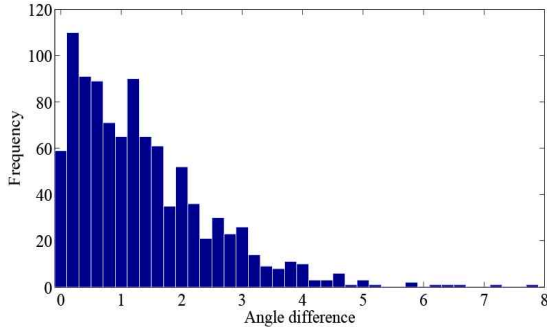
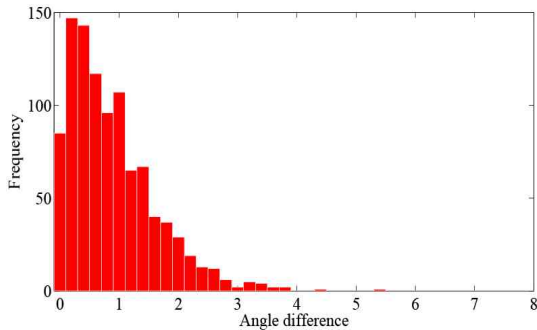


그림 3. PCA와 RF-PCA3를 이용하여 찾아낸 첫 번째 주성분  
Fig. 3. First principal component found by PCA and RF-PCA3



(a) RF-PCA2



(b) RF-PCA3

그림 4. 두 주성분 사이의 각도

Fig. 4. Angles between two principal components

그림 3에서 RF-PCA2의 주성분은 표시하지 않았지만 RF-PCA3의 주성분과는 약간의 차이가 있었다. 따라서 첫 번째 실험은 두 방법이 찾아낸 주성분의 질(quality)을 비교한다. 먼저 잡음이 없는 데이터에서 PCA를 이용하여 첫 번째 주성분  $w_0$ 를 구하고, 잡음이 첨가된 데이터에서 RF-PCA2와 RF-PCA3로 각각 첫 번째 주성분  $w_1$ 과  $w_2$ 를 구한다. 잡음이 첨가된 데이터에서 찾아낸 주성분이 잡음이 없는 경우의 주성분과 유사한 경우 그 성능이 우수한 것으로 생각할 수 있으므로,  $w_0$ 와  $w_1$  그리고  $w_0$ 와  $w_2$  사이의 각도를 계산하여 비교하였다. 이 때 두 가우스 분포의 중심은 각각  $[0 \ 0]^T$ 과  $[4 \ 3]^T$ 로 설정하였다. 그림 4는 그림 3의 데이터를 이용하여 1,000회 반복 실험한 결과를 히스토그램으로 나타낸 것이며 RF-PCA2의 경우 평균 각도는 1.357°, RF-PCA3는 0.886°로 RF-PCA3가 나은 결과를 보였다.

첫 번째 실험에서 RF-PCA2 역시 잡음에 민감하지 않은 결과를 보여주었다. 하지만 잡음의 중심을  $[7 \ 7]^T$ 로 설정한 경우 RF-PCA2는 그림 5에 나타난 것과 같이 국부 최적해에 빠져 잘못된 결과를 보여주었다.

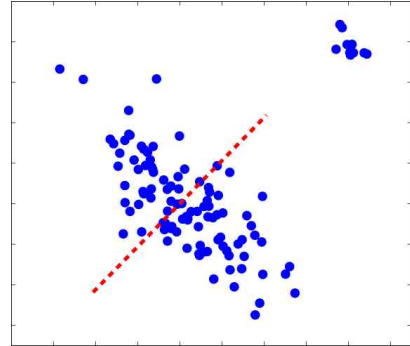
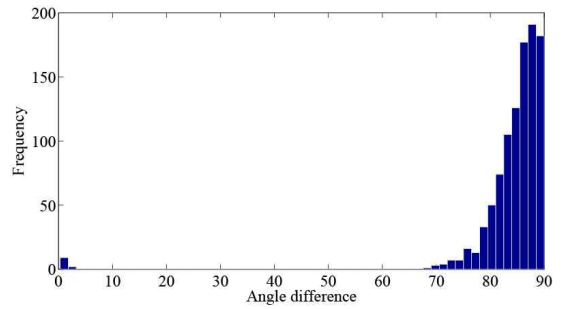
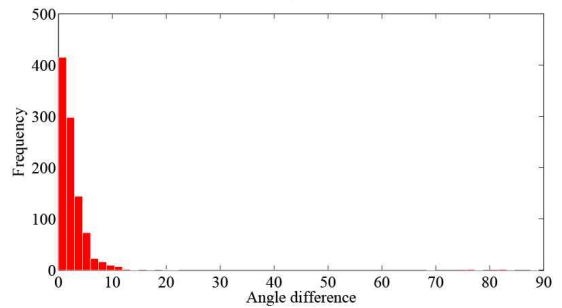


그림 5. RF-PCA2의 국부 최적해  
Fig. 5. Local optimum in RF-PCA2



(a)

RF-PCA2



(b) RF-PCA3

그림 6. 두 주성분 사이의 각도

Fig. 6. Angles between two principal components

하지만 RF-PCA3의 경우 특이값에 작은 소속도 값을 초기값으로 설정함으로써 잘못된 국부 최적해에 빠질 가능성을 줄여준다. 그림 6은 그림 5의 데이터에 대해 RF-PCA2와 RF-PCA3를 적용한 결과를 요약한 것이다. 실험 방법은 앞의 실험과 동일하다. 그림 6에서 알 수 있듯이, RF-PCA2는 대부분이 잘못된 국부 최적해에 빠지는 반면 RF-PCA3는 대부분이 올바른 해에 근접하고 있음을 알 수 있다.

## V. 결론

Improved robust fuzzy PCA(RF-PCA2)는 퍼지 소속도를 이용하여 PCA의 잡음 민감성을 줄여주는 반복 알고리즘(iterative algorithm)으로 이전 알고리즘들에 비해 잡음 민감성이 줄어든다는 사실이 입증되었다. 하지만 RF-PCA2 역시 국부 최적해에 수렴하는 문제점이 있으며 이는 균일한 소속도 값으로 초기화하는 점과 자승 오류를 기반으로 한다는 점에 기인한다. 이 논문에서는 이 두 가지를 보완한 RF-PCA3 알고리즘을 제안하였다. 제안한 알고리즘은 데이터 분포를 활용하여 초기 소속도 값을 결정하고 PCA의 목적함수를 활용하여 RF-PCA2 목적함수를 수정하였다. 제안한 알고리즘의 우수성은 실험 결과를 통해 확인할 수 있다.

비록 RF-PCA3가 이전 알고리즘에 비해 나은 성능을 보이지만 몇 가지 개선의 여지도 남아 있다. 목적함수에 도입된 상수  $\alpha$ 의 결정이 그 첫 번째이다. 이는 자승 오류의 합을 어느 정도 반영할 것인지 결정하는 상수로 이 논문에서는 실험적으로 결정된 값을 사용하였다. 현재 이 상수를 데이터 집합에 따라 자동으로 결정하는 방법에 관해 연구 중에 있다. 또 다른 개선 방향으로는 초기 소속도 결정 방법이 있을 수 있다. 제안한 방법은 PCA의 가정에 근거하고 있고 실험적으로도 우수한 성능을 보여 주었지만  $O(N^2)$ 의 연산을 요구하는 문제가 있다. RF-PCA3는 반복 알고리즘으로 그 자체만으로도 많은 연산을 요구하므로 적은 연산을 요구하는 방법으로서의 전환도 가치 있다 하겠다. 이는 향후 연구 과제로 남아있다.

## 참고문헌

- [1] I. T. Jolliffe, *Principal Component Analysis*, 2nd Edition, Springer, 2002.
- [2] P. Rousseeuw, "Multivariate estimation with high breakdown point," *Mathematical Statistics and Applications B*, pp. 283 - 297, Dec. 1985.
- [3] C. D. Lu, T. Y. Zhang, X. Z. Du, and C. P. Li, "A robust kernel PCA algorithm," *Proceedings of the 3rd International Conference on Machine Learning and Cybernetics*, pp. 3084 - 3087, Aug. 2004.
- [4] C. Lu, T. Zhang, R. Zhang, and C. Zhang, "Adaptive robust kernel PCA algorithm," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. VI 621 - 624, Apr. 2003.
- [5] Seok-Woo Jang, Moon-Haeng Huh, Gye-Young Kim, "Effective Handwriting Verification through DTW and PCA," *Journal of the Korea society of computer and information*, Vol. 14, No. 7, pp. 25 - 32, Jul. 2009.
- [6] G. Heo, P. Gader, and H. Frigui, "RKF-PCA: Robust kernel fuzzy PCA," *Neural Networks*, Vol. 22, No. 5-6, pp. 642 - 650, Aug. 2009.
- [7] T. N. Yang and S. D. Wang, "Fuzzy auto-associative neural networks for principal component extraction of noisy data," *IEEE Transaction on Neural Networks*, Vol. 11, No. 3, pp. 808 - 810, Mar. 2000.
- [8] T. R. Cundari, C. Sarbu, and H. F. Pop, "Robust fuzzy principal component analysis (FPCA). A comparative study concerning interaction of carbonhydrogen bonds with molybdenum-oxo bonds," *Journal of Chemical Information and Computer Sciences*, Vol. 42, No. 6, pp. 1363 - 1369, Dec. 2002.
- [9] Gyeongyong Heo, Young Woon Woo and Seong Hoon Kim, "An Improved Robust Fuzzy Principal Component Analysis," *The Journal of the Korean Institute of Maritime Information and Communication Sciences*, Vol. 14, No. 5, pp. 1093-1102, May 2010.
- [10] V. J. Hodge and J. Austin, "A Survey of Outlier Detection Methodologies," *Artificial Intelligence Review*, vol. 22, No. 2, pp. 85-126, Oct. 2004.
- [11] G. Heo and P. Gader, "Fuzzy SVM for noisy data: A robust membership calculation method," *Proceedings of the 2009 IEEE International Conference on Fuzzy Systems*, pp. 431 - 436, Aug. 2009.

저 자 소 개



**김 성 훈**  
 1996년 2월 :  
 연세대학교 본대학원 전자공학과  
 (공학박사)  
 1996년 3월~2006년 2월 :  
 영동대학교 컴퓨터공학과 부교수  
 2006년 3월~현재 :  
 경북대학교 컴퓨터정보학부 부교수  
 관심분야 : 인공지능, 패턴인식, 지능  
 형콘텐츠  
 E-mail : shkim1454@knu.ac.kr



**허 경 응**  
 1996년 8월 : 연세대학교 본대학원  
 전자공학과 (공학석사)  
 2009년 12월 :  
 Department of Computer and  
 Information Science and  
 Engineering, University of  
 Florida (공학박사)  
 관심분야 : Machine Learning,  
 Pattern Recognition,  
 Image Processing  
 E-mail : hgycap@hotmail.com



**우 영 운**  
 1991년 8월 :  
 연세대학교 본대학원 전자공학과(공학  
 석사)  
 1997년 8월 :  
 연세대학교 본대학원 전자공학과(공학  
 박사)  
 1997년 9월~현재 :  
 동의대학교 멀티미디어공학과 교수  
 관심분야 : 인공지능, 패턴인식, 퍼지  
 이론, 의료정보  
 E-mail : ywwoo@deu.ac.kr