



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0064971
(43) 공개일자 2019년06월11일

(51) 국제특허분류(Int. Cl.)
G06N 99/00 (2019.01) G06F 17/16 (2006.01)
(52) CPC특허분류
G06N 20/00 (2019.01)
G06F 17/16 (2013.01)
(21) 출원번호 10-2017-0164350
(22) 출원일자 2017년12월01일
심사청구일자 없음

(71) 출원인
서강대학교산학협력단
서울특별시 마포구 백범로 35 (신수동, 서강대학교)
(72) 발명자
양지훈
서울특별시 은평구 백련산로 2길 19, 백련산힐스테이트 1차 105동 603호
오병화
서울특별시 영등포구 영등포동 2가 29-80, 604호
(74) 대리인
이지연

전체 청구항 수 : 총 3 항

(54) 발명의 명칭 그래프 기반 준지도 학습에서의 빠른 낮은 계수 표현을 기반으로 한 그래프 구축 방법 및 시스템

(57) 요약

본 발명에 따른 그래프 기반 준지도 학습에서의 그래프 구축 방법은, (a) 데이터셋(X), 최근접 이웃수 (k) 를 입력받는 단계; (b) 데이터셋(X)를 사용하여 k-NN 그래프 생성후, L을 계산하는 단계; (c) Skinny SVD 를 사용하여 X 를 분해하는 단계; (d) SVD 를 사용하여 βL 을 구하는 단계; (e) 일정 조건을 만족하는 동안, J, W, Q를 갱신하는 단계; (f) Y_1, Y_2 를 계산하여 최적해를 구하는 단계; 를 구비하여, 빠른 낮은 계수 표현 알고리즘을 기반으로 하여 그래프 구축한다. 본 발명에 따른 그래프 구축 방법은, 빠른 낮은 계수 표현을 기반으로 하여, 기반 최적화 목표에 추가 제약 조건을 도입하고 이를 최적화시킴으로써, 더 좋은 해를 빠르게 찾아낼 수 있게 된다.

대표도 - 도1

Input: 데이터셋 X , 최근접 이웃 수 k , 매개 변수 λ, a
Output: 최적 해 W^*, Q^*, Z^*, E^*
1 :
 $W = Q = J = Y_1 = Y_2 = 0, \mu = 10^{-1}, \rho = 1.1, \epsilon = 10^{-8}, k = 5$
2: X 를 사용하여 k-NN 그래프 생성 후, L 을 계산
3: Skinny SVD를 사용하여 $X = U_r S_r V_r^T$ 으로 분해
4: SVD를 사용하여 다음을 수행: $\beta L = U_2 \Sigma_2 U_2^T$
5: while $\|V_r^T - w - Q\|_\infty > \epsilon$ do
6: J 를 식 (14)에 따라 갱신(다른 변수 값 고정)
7: W 를 식 (19)에 따라 갱신(다른 변수 값 고정)
8: Q 를 식 (20)에 따라 갱신(다른 변수 값 고정)
9: $Y_1 = Y_1 + \mu(V_r^T - W - Q), Y_2 = Y_2 + \mu(W - J)$
10: $\mu = \min(\rho\mu, \mu_{\max})$
11: end while
12: $W^* = W, Q^* = Q, J^* = J, Z^* = V_r W^*, E = X - X V_r W^*$

이 발명을 지원한 국가연구개발사업

과제고유번호 1711055428

부처명 과학기술정보통신부

연구관리전문기관 정보통신기술진흥센터

연구사업명 정보통신기술인력양성

연구과제명 글로벌 핀테크 벤처 기업 및 창의적 인재 육성을 위한 스마트 핀테크 플랫폼 기술 개발

기 여 율 1/1

주관기관 서강대학교 산학협력단

연구기간 2015.06.01 ~ 2018.12.31

명세서

청구범위

청구항 1

레이블된 데이터와 레이블되지 않은 데이터를 모두 사용하여 모델을 학습하는 준지도 학습에서의 그래프 구축 방법에 있어서,

- (a) 데이터셋(X), 최근접 이웃수 (k) 를 입력받는 단계;
- (b) 데이터셋(X)를 사용하여 k-NN 그래프 생성후, L을 계산하는 단계;
- (c) Skinny SVD 를 사용하여 X 를 분해하는 단계;
- (d) SVD 를 사용하여 βL 을 구하는 단계;
- (e) 일정 조건을 만족하는 동안, J, W, Q를 갱신하는 단계;
- (f) Y_1, Y_2 를 계산하여 최적해를 구하는 단계;

를 구비하여, 빠른 Graph-Regularized Low-Rank Representation (FaGLRR) 알고리즘을 기반으로 하여 그래프 구축하는 것을 특징으로 하는 준지도 학습에서의 그래프 구축 방법.

청구항 2

레이블된 데이터와 레이블되지 않은 데이터를 모두 사용하여 모델을 학습하는 준지도 학습에서의 그래프 구축 방법에 있어서,

- (a) 데이터셋(X), 최근접 이웃수 (N) 를 입력받는 단계;
- (b) W, Q, J, Y1, Y2 를 0으로 설정하는 단계;
- (c) 데이터셋(X)를 SMA를 수행하여 A를 계산하는 단계;
- (d) Skinny SVD 를 사용하여 X 를 분해하는 단계;
- (e) W를 갱신하는 경우에만 SVD 를 수행하는 단계;
- (f) 일정 조건을 만족하는 동안, J, W, Q를 갱신하는 단계;
- (f) Y_1, Y_2 를 계산하여 최적해를 구하는 단계;

를 구비하여, 빠른 Manifold Low-Rank Representation (FaMLRR) 알고리즘을 기반으로 하여 그래프 구축하는 것을 특징으로 하는 준지도 학습에서의 그래프 구축 방법.

청구항 3

청구항 제1항 및 제2항 중 어느 한 항에 있어서, Z^* 이 계산되면, Z^* 의 각 열 z_i^* 는 x_i 를 재구축하는 기여도이

$$z_i^* = z_i^* / \|z_i^*\|_2$$

므로, 모든 열을 로 정규화한 후, 문턱값(Threshold) θ 보다 작은 값을 0 으로 만들고, 그래프 행렬 w 를 수학적 25과 같이 계산하는 것을 특징으로 하는 준지도 학습에서의 그래프 구축 방법

[수학적 25]

$$w = \frac{\widehat{Z}^* + (\widehat{Z}^*)^T}{2}, \text{ for } \widehat{Z}_{ij}^* = \begin{cases} z_{ij}^* & \text{if } z_{ij}^* > \theta \\ 0 & \text{otherwise} \end{cases}$$

발명의 설명

기술 분야

[0001] 본 발명은 그래프 기반 준지도 학습에서의 그래프 구축 방법에 관한 것으로서, 더욱 구체적으로는 빠른 낮은 계수 표현을 기반으로 하여, 기반 최적화 목표에 추가 제약 조건을 도입하고 이를 최적화시킴으로써, 더 좋은 해를 빠르게 찾아낼 수 있도록 한 그래프 기반 준지도 학습에서의 그래프 구축 방법에 관한 것이다.

배경 기술

[0002] 준지도 학습(Semi-Supervised Learning)은 기계 학습의 한 분야로서 레이블 된(Labeled) 데이터와 레이블 되지 않은(Unlabeled) 데이터 모두를 사용하여 분류 모델을 학습한다. 이 방법은 지도 학습에 비해 예측 성능을 높일 수 있다.

[0003] 그리고 높은 예측 성능으로 인해 최근 주목받고 있는 그래프 기반(Graph-Based) 준지도 학습은, 그래프 구축(Graph Construction)과 레이블 추론(Label Inference)의 두 단계로 구성된다. 그러므로 데이터에 내포된 특징적 구조를 잘 표현하는 좋은 그래프를 구축하는 것은 매우 중요하다. 왜냐하면 입력 그래프가 달라지면 추론의 결과도 달라지기 때문이다.

[0004] 그래프 구축에서 가장 널리 사용되는 방법은 k-최근접 이웃(k-Nearest Neighbors, k-NN) 방법인데, 최근 다른 방법들도 제안되었다. 데이터의 다양체 구조를 반영하기 위해 여러 방법들이 국부 선형 임베딩(Locally Linear Embedding, LLE)을 활용하며, 척도 학습(Metric Learning)을 활용하기도 한다.

[0005] 최근에는 낮은 계수 표현(Low-Rank Representation, LRR) 기반 그래프 구축 방법이 각광받고 있다. 이렇게 구축한 그래프는 군집화(Clustering), 그래프 기반 준지도 학습 등의 다양한 응용에 사용된다. LRR 기반 방법들은 표현 행렬(Representation Matrix)의 Trace Norm을 최소화하기 위하여 반복 계산을 수행하는데, 이 반복 계산에 특이값 분해(Singular Value Decomposition)를 수반한다. 이는 SVD가 $n \times n$ 행렬에 대해 $O(n^3)$ 의 시간 복잡도를 요구하므로 (n 은 인스턴스 수) 계산적으로 매우 비효율적이다.

[0006] 그럼에도, LRR 기반 방법은 높은 정확도 때문에 널리 적용되고 있다. 이를 해결하기 위해 Fixed-Rank Representation(FRR)등이 제안되었으나, LRR과 다른 해를 도출한다. 한편, Fast LRR Solver(FaLRR)는 속도를 향상시키면서도 LRR과 동일한 해를 도출한다. 그러나 LRR은 2010년에 발표되었기에, Graph-Regularized Low-Rank Representation(GLRR), Non-Negative Low-Rank And Sparse(NNLRs) Graphs, Manifold Low-Rank Representation(MLRR) 등의 발전된 방법들보다 분류에 사용할 때 성능이 떨어진다.

[0007] 그러므로 본 발명에서는 GLRR보다 훨씬 빠르면서도 동일한 해를 도출하는 Fast GLRR(FaGLRR)과 동일한 해를 도출하는 방법을 찾아내기는 어렵지만 최소화하는 목표가 추가될 경우 더 좋은 결과를 나타내는 Fast MLRR(FaMLRR)을 제안한다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) 한국특허공개공보 10-2016-0076927호

발명의 내용

해결하려는 과제

[0009] 기술한 문제점을 해결하기 위한 본 발명의 목적은 GLRR보다 훨씬 빠르면서도 동일한 해를 도출하는 Fast GLRR(FaGLRR)과 동일한 해를 도출하는 방법을 찾아내기는 어렵지만 최소화하는 목표가 추가될 경우 더 좋은 결과를 나타내는 Fast MLRR(FaMLRR)을 제공하는 것이다.

과제의 해결 수단

[0010] 기술한 기술적 과제를 달성하기 위한 본 발명의 제1 특징에 따른 그래프 기반 준지도 학습에서의 그래프 구축 방법은, (a) 데이터셋(X), 최근접 이웃수 (k) 를 입력받는 단계; (b) 데이터셋(X)를 사용하여 k-NN 그래프 생성 후, L을 계산하는 단계; (c) Skinny SVD 를 사용하여 X 를 분해하는 단계; (d) SVD 를 사용하여 βL 을 구하는 단계; (e) 일정 조건을 만족하는 동안, J, W, Q를 갱신하는 단계; (f) Y_1, Y_2 를 계산하여 최적해를 구하는 단계; 를 구비하여, 빠른 낮은 계수 표현 알고리즘을 기반으로 하여 그래프 구축한다.

[0011] 본 발명의 제2 특징에 따른 그래프 기반 준지도 학습에서의 그래프 구축 방법은, (a) 데이터셋(X), 최근접 이웃수 (N) 를 입력받는 단계; (b) W, Q, J, Y_1, Y_2 를 0으로 설정하는 단계; (c) 데이터셋(X)를 SMA를 수행하여 A를 계산하는 단계; (d) Skinny SVD 를 사용하여 X 를 분해하는 단계; (e) W를 갱신하는 경우에만 SVD 를 수행하는 단계; (f) 일정 조건을 만족하는 동안, J, W, Q를 갱신하는 단계; (f) Y_1, Y_2 를 계산하여 최적해를 구하는 단계; 를 구비하여, 빠른 낮은 계수 표현 알고리즘을 기반으로 하여 그래프 구축한다.

발명의 효과

[0012] 낮은 계수 표현(Low-Rank Representation, LRR) 기반 방법은 얼굴 클러스터링, 객체 검출 등의 여러 실제 응용에 널리 사용되고 있다. 이 방법은 그래프 기반 준지도 학습에서 그래프 구축에 사용할 경우 높은 예측 정확도를 확보할 수 있어 많이 사용된다. 그러나 LRR 문제를 해결하기 위해서는 알고리즘의 매 반복마다 데이터 수 크기의 정방행렬에 대해 특이값 분해를 수행하여야 하므로 계산 비효율적이다. 이를 해결하기 위해 속도를 향상시킨 발전된 LRR 방법을 제안한다. 이는 최근 발표된 Fast LRR(FaLRR)을 기반으로 하며, FaLRR이 속도는 빠르지만 실제로 분류 문제에서 성능이 낮은 것을 해결하기 위해 기반 최적화 목표에 추가 제약 조건을 도입하고 이를 최적화하는 방법을 제안한다. 실험을 통하여 제안 방법은 LRR보다 더 좋은 해를 빠르게 찾아냄을 확인할 수 있다. 또한, 동일한 해를 도출하는 방법을 찾아내기는 어렵지만 최소화하는 목표가 추가될 경우 더 좋은 결과를 나타내는 Fast MLRR(FaMLRR)을 제안한다.

도면의 간단한 설명

[0013] 도 1은 본 발명의 제1 실시예에 따른 FaGLRR 알고리즘을 도시한 표이다.

도 2는 본 발명의 제2 실시예에 따른 FaMLRR 알고리즘을 도시한 표이다.

도 3은 본 발명에 따른 그래프 구축 방법들에 따른 실험 결과 표이다.

도 4는 ORL (d = 1024, n = 400) 에 따른 결과 그래프이며, 도 5는 Extended Yale B (d = 1024, n = 1000) 에 따른 결과 그래프이며, 도 6은 COIL100 (d = 1024, n = 7200) 에 따른 결과 그래프이며, 도 7은 Isolet (d = 617, n = 7797)에 따른 결과 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0014] [낮은 계수 표현]

[0015] $X = [x_1, x_2, \dots, x_n] \in R^{d \times n}$ 를 데이터셋(각 열은 인스턴스를 나타냄), d 를 특징 수, n 을 인스턴스 수라 하면, 각 인스턴스는 사전(Dictionary) $A = [a_1, a_2, \dots, a_m] \in R^{d \times m}$ 의 기저들의 선형 결합으로

표현된다. $X = AZ$ 이고, 이때 $Z = [z_1, z_2, \dots, z_n] \in R^{m \times n}$ 는 계수 행렬(Coefficient Matrix)으로, 각 x_i 는 이를 대표하는 새로운 표현 z_i 로 나타난다. 그리고 사전 A 는 Overcomplete(필요 기저보다 더 많은 벡터로 구성됨)하므로 $X = AZ$ 를 만족하는 여러 해가 존재할 수 있다. 여기서, Z 의 낮은 계수 조건을 만족시키기 위해 다음 수학적 1을 최적화한다.

수학식 1

[0016]

$$\min_Z \|Z\|_*, s.t. X = AZ$$

[0017]

여기서 $\|Z\|_*$ 는 Z의 Trace Norm으로, Z의 계수(Rank) 대신 최소화하는 목표이다. 데이터의 잡음, 변형, 결측값이 있을 경우를 고려한다면, 최소화시키기 위해 제안된 목표 식은 수학식 2와 같다.

수학식 2

[0018]

$$\min_{Z,E} \|Z\|_* + \lambda \|E\|_{2,1}, s.t. X = AZ + E$$

[0019]

여기서, $\|E\|_{2,1} = \sum_{j=1}^n \sqrt{\sum_{i=1}^d E_{ij}^2}$ 는 E의 $\ell_{2,1}$ Norm 이고, $\lambda > 0$ 는 저계수 항과 오류 최소화 항 사이의 영향력을 조정하는 매개 변수이다. 최적 해 Z^* 의 Exact 및 Inexact Augmented Lagrange Multiplier(ALM) 방법으로 구할 수 있다.

[0020]

FaLRR은 수학식 2의 문제에 대하여 동일한 해를 도출한다. 그 대신, Z 보다 더 작거나 같은 크기의 행렬 $W \in R^{r \times n}$ (r 은 X의 계수)에 대해 최적화를 수행한다. W^* 가 다음 목표식에 대한 최적 해이고, $Z^* = V_r W^*$, $E^* = X - X V_r W^*$ 으로 정의할 때, 수학식 3은 수학식 2의 Z^* 와 동일한 해를 도출한다. 단 수학식 2에서의 $A=X$ 의 경우이다.

수학식 3

[0021]

$$\min_W \|W\|_* + \lambda \|S_r(V_r^T - W)\|_{2,1}$$

[0022]

여기서, $V \in R^{n \times r}$ 는 Skinny SVD를 사용하여 $X = U_r S_r V_r^T$ ($U_r \in R^{d \times r}$, $S_r \in R^{r \times r}$)로 분해

함으로써 도출되는 열간 직교 행렬이다. $U_r^T U_r = V_r^T V_r = I$ 이므로,

$\|Z^*\|_* = \|W^*\|_*$, $\|E^*\|_{2,1} = \|S_r(V_r^T - W)\|_{2,1}$ 도 항상 성립한다. r 이 d, n 중의 작은 수보다 항상 작거나 같으므로, $n \gg d$ 일 경우 LRR의 반복 SVD 계산의 시간 복잡도는 $O(tn^3)$ 으로 FaLRR의 $O(tnr^2)$ 이 계산적으로 훨씬 효율적이다.

[0023]

GLRR은 수학식 2에 그래프 정규화항을 추가한 다음 목표식을 최적화한다. 단 수학식 2에서 $A=X$ 의 경우이다.

수학식 4

[0024]

$$\min_{Z,E} \|Z\|_* + \lambda \|E\|_{2,1} + \beta \text{tr}(ZLZ^T), s.t. X = XZ + E$$

[0025]

여기서, L 의 Graph Laplacian 이며, $\beta > 0$ 는 λ 와 같은 각 항의 영향력을 조정하는 매개 변수이다.

$$\text{tr}(ZLZ^T) = \sum_{i,j=1}^n w_{ij} \|z_i - z_j\|_2^2$$

이며, $w \in R^{n \times n}$

는 k -NN 으로 계산한 X 의 그래프의 가중치 행렬이다. 이 항은 곧, k -NN 그래프에서 두 인스턴스 x_i 와 x_j 가 가깝다면 새 표현 공간의 z_i 와 z_j 또한 가까워야 함을 의미한다.

[0026]

NNLRS(Non-Negative Low-Rank and Sparse) 그래프 구축 방법의 최적화 목표는 다음과 같다. 이 목표는 입력 데이터에 대하여 낮은 계수이면서도 더 적은 유효 성분으로 표현되는(즉 0인 성분이 많은) Z 의 해를 찾는다.

수학식 5

[0027]

$$\min_{Z,E} \|Z\|_* + \lambda \|E\|_{2,1}, s.t. X = AZ + E, Z \geq 0$$

[0028]

여기서 $Z \geq 0$ 의 조건은 이후 해 Z^* 를 구한 후 이를 사용하여 무방향 가중치 그래프를 구축할 때, 가중치 행렬 W 의 각 원소가 0보다 크거나 같아야 하는 조건을 만족시키기 위해서 제약 조건으로 추가한 것이다.

[0029]

MLRR은 수학식 5에 다양체 정규화(Manifold Regularization) 항을 추가한 수학식 6을 최적화한다.

수학식 6

[0030]

$$\min_{Z,E} \|Z\|_* + \lambda \|E\|_{2,1} + \beta \text{tr}(ZLZ^T), s.t. X = XZ + E, Z \geq 0$$

[0031]

여기서, $\text{tr}(ZLZ^T) = \|Z - ZW\|_2^2 = \sum_{i=1}^n \|z_i - \sum_j W_{ji} z_j\|_2^2$ 이다. 가중치 행렬 W 는 수학식 6을 최적화하기 전에, Sparse Manifold Adaption(SMA) 방법을 통해 사전에 계산되어진다. 즉, 수학식 6은 새로운 데이터 표현형 Z 의 각 인스턴스가 사전에 구한 Sparse Manifold을 통해 얼마나 정확하게 복원될 수 있는지를 평가하는 항을 추가한 것이다.

[0032]

[Fast GLRR]

[0033]

본 발명에서는 GLRR 을 빠르게 계산하는 방법을 제안한다. 이를 위해 Z 보다 작거나 같은 크기의 W 에 대한 최적화 목표를 새로 정의하고, 이 식이 수학식 4와 동일한 해를 도출함을 보인다.

[0034]

정리 : W^* 가 다음 문제에 대한 최적 해이며, $Z^* = V_r W^*, E^* = X - X V_r W^*$ 일때 (Z^*, E^*) 는 수학식 4의 최적해이다.

수학식 7

[0035]

$$\min_W \|W\|_* + \lambda \|S_r(V_r^T - W)\|_{2,1} + \beta \text{tr}(WLW^T)$$

[0036]

증명 : (Z^*, E^*) 가 최적해가 되기 위해서는 (Z, E) 의 어떤 가능한 값에 대해서 다음 수학식 8의 부등식이 항상 성립하여야 한다.

수학식 8

[0037]

$$\begin{aligned} \|Z\|_* + \lambda \|E\|_{2,1} + \beta \text{tr}(ZLZ^T) \\ \geq \|Z^*\|_* + \lambda \|E^*\|_{2,1} + \beta \text{tr}(Z^*L(Z^*)^T) \end{aligned}$$

[0038]

이를 위하여는 수학식 9와 같은 부등식을 전개할 수 있다.

수학식 9

[0039]

$$\begin{aligned} \|Z\|_* + \lambda \|E\|_{2,1} + \beta \text{tr}(ZLZ^T) \\ \geq \|V_r^T Z\|_* + \lambda \|S_r(V_r^T - V_r^T Z)\|_{2,1} + \beta \text{tr}(V_r^T ZLZ^T V_r) \\ \geq \|W^*\|_* + \lambda \|S_r(V_r^T - W^*)\|_{2,1} + \beta \text{tr}(W^*L(W^*)^T) \\ = \|Z^*\|_* + \lambda \|E^*\|_{2,1} + \beta \text{tr}(Z^*L(Z^*)^T) \end{aligned}$$

[0040]

우선, Trace Norm 의 성질에 의하여, $\|Z\|_* \geq \|V_r^T Z\|_*$ 이다. 그리고,

$$\|E\|_{2,1} = \|S_r(V_r^T - V_r^T Z)\|_{2,1} \quad \text{이므로,} \quad \ell_{2,1} \text{ Norm 의 성질에 의하여}$$

$$\|E\|_{2,1} = \|S_r(V_r^T - V_r^T Z)\|_{2,1} \quad \text{이다. 또한,} \quad V_r^T V_r = I \quad \text{이므로}$$

$$\text{tr}(ZLZ^T) = \text{tr}(V_r V_r^T ZLZ^T) = \text{tr}(V_r^T ZLZ^T V_r) \quad \text{이다. 따라서, 수학식 9의 첫번째 부등식}$$

이 성립한다.

[0041]

또한, $Z^* = V_r W^*$ 에서 $W^* = V_r^T Z^*$ 가 유도되며, Z 는 최적해 또는 그 이전 해이므로 수학식 9의 둘째 줄의 각 Norm은 셋째줄의 각 Norm 보다 크거나 같다. 따라서, 두번째 부등식도 성립한다.

[0042]

마지막으로 $E^* = X - X V_r W^* = U_r S_r(V_r^T - W^*)$ 에서, 각각의 Norm 의 정의에 따라

$$\|X^*\|_* = \|W^*\|_*, \quad \|E^*\|_{2,1} = \|S_r(V_r^T - W^*)\|_{2,1} \quad \text{가 성립하고, Trace 의 성질에 의}$$

$$\text{tr}(W^*L(W^*)^T) = \text{tr}(V_r^T V_r W^*L(W^*)^T) =$$

하여 $\text{tr}(V_r W^*L(V_r W^*)^T) = \text{tr}(Z^*L(Z^*)^T)$ 이므로, 마지막 등식도 성립한다. 따라

서 수학적식 8의 조건이 항상 성립한다.

[0043] 그러므로, 수학적식 9의 셋째줄은 수학적식 7의 최적값이며, 수학적식 9의 넷째줄과 동일한다. 그리고, 이는 수학적식 4의 최적값이다.

[0044] 일단 수학적식 7을 최적화하기 위해, 새로운 변수 $Q \in R^{r \times n}$ 를 도입하여 수학적식 10과 같이 새로 쓸 수 있다.

수학적식 10

[0045]
$$\min_{W, Q} \|W\|_* + \lambda \|S_r Q\|_{2,1} + \beta \text{tr}(WLW^T), s.t. W + Q = V_r^T$$

[0046] 그리고 보조 변수 $J \in R^{r \times n}$ 를 추가로 도입하여 다시 쓰면, 수학적식 11이 된다.

수학적식 11

[0047]
$$\min_{W, Q, J} \|J\|_* + \lambda \|S_r Q\|_{2,1} + \beta \text{tr}(WLW^T), s.t. W + Q = V_r^T, W = J$$

[0048] 따라서 해당하는 Augmented Lagrangian 함수는 다음 수학적식 12와 같이 정의하며, Inexact ALM을 사용하여 최소화할 수 있다.

수학적식 12

[0049]
$$\begin{aligned} L(W, Q, J, Y_1, Y_2, \mu) &= \|J\|_* + \lambda \|S_r Q\|_{2,1} + \beta \text{tr}(WLW^T) \\ &+ \langle Y_1, V_r^T - W - Q \rangle + \langle Y_2, W - J \rangle \\ &+ \frac{\mu}{2} \left(\|V_r^T - W - Q\|_F^2 + \|W - J\|_F^2 \right) \\ &= \|J\|_* + \lambda \|S_r Q\|_{2,1} + \beta \text{tr}(WLW^T) \\ &\quad - \frac{1}{2\mu} \left(\|Y_1\|_F^2 + \|Y_2\|_F^2 \right) \\ &+ \frac{\mu}{2} \left(\left\| V_r^T - W - \frac{Y_1}{\mu} \right\|_F^2 + \left\| W - J + \frac{Y_2}{\mu} \right\|_F^2 \right), \end{aligned}$$

[0050] 여기서 Y_1 과 Y_2 는 Lagrange 승수, $\mu > 0$ 는 페널티 매개 변수(Penalty Parameter)이다. 이 함수는 각 변수를 각각 최소화하는 방법으로 최소화할 수 있다. 매 반복마다, 각 변수에 대해 다른 변수들의 값을 고정시킨 상태에서 변경을 수행한다. 이를 위해 수학적식 12를 부분 문제들로 나눈 후 갱신 규칙을 유도한다.

[0051] 먼저 J에 대한 항을 모아서 만든 최소화 목표식은, 수학적식 13이 된다.

수학식 13

$$\min_J \left\| J \right\|_* + \frac{\mu}{2} \left\| J - \left(W + \frac{Y_2}{\mu} \right) \right\|_F^2$$

[0052]

[0053] 이 식은 Singular Value Thresholding 연산자(아래 수학식 14에서 Θ)를 사용하여 해를 구할 수 있다. 즉 J의 갱신 규칙은 수학식 14가 된다.

수학식 14

$$J = \Theta_{\frac{1}{\mu}} \left(W + \frac{Y_2}{\mu} \right)$$

[0054]

[0055] 그리고 수학식 12에서 W에 관한 항만 모은 후 W에 대하여 미분한 뒤 0으로 놓으면 W를 최소화하는 수학식 15를 구할 수 있다.

수학식 15

$$W = \frac{1}{2} \left(V_r^T - Q + J + \frac{Y_1}{\mu} - \frac{Y_2}{\mu} \right) (\beta L + \mu I)^{-1}$$

[0056]

[0057] 그러나 μ 가 매 반복마다 변하기에 역행렬도 역시 매번 계산하여야 하므로, 여기서는 다른 방법을 사용한다. 수학식 15로부터 수학식 16을 유도한다.

수학식 16

$$\mu I W + W \beta L = \frac{\mu}{2} \left(V_r^T - Q + J + \frac{Y_1}{\mu} + \frac{Y_2}{\mu} \right) (\beta L + \mu I)^{-1}$$

[0058]

[0059] 여기서, μI 와 βL 은 모두 실수 대칭 행렬이다. 그러므로, SVD를 통해,

$$\mu I = U_1 \Sigma_1 U_1^T, \quad \beta L = U_2 \Sigma_2 U_2^T$$

(U_1, U_2 는 유니타리 행렬)로 분해 가능하며, 수학식 16의 등호

오른쪽 부분을 M으로 두면, 수학식 16은 $U_1 \Sigma_1 U_1^T W + W U_2 \Sigma_2 U_2^T = M$ 이 된다. 그 후 양변의

왼쪽에 U_1^T , 오른쪽에 U_2^T 를 곱하면 수학식 17과 같다.

수학식 17

$$\begin{aligned} \Sigma_1 U_1^T W U_2 + U_1^T W U_2 \Sigma_2 &= U_1^T M U_2 \\ \Sigma_1 \widetilde{W} + \widetilde{W} \Sigma_2 &= T \end{aligned}$$

[0060]

[0061] 여기서, $U_1 = U_1^T = I$ 이고 $\Sigma_1 = \mu I$ 이 된다. 그러므로, 다음 두식, $\widetilde{W} = U_1^T W U_2 = W U_2$, $T = U_1^T W U_2 = M U_2$ 이 성립한다. Σ_2 가 대각 행렬임을 고려할 때, 각 성분 $\widetilde{W}_{ij}$ 는 수학식 18과 같이 계산된다.

수학식 18

$$\widetilde{W}_{ij} = \frac{T_{ij}}{(\mu + \sigma_2^j)}$$

[0062]

[0063] 여기서, $\Sigma_2 = \text{diag}(\sigma_2^1, \sigma_2^2, \dots, \sigma_2^n)$ 이다. 그러면 W 의 갱신식은 수학식 19와 같다.

수학식 19

$$W = \widetilde{W} U_2^T$$

[0064]

[0065] 한편, Q 는 FaLRR에서 제시한 방법에 의해서 갱신하면 된다. 즉, $C \in R^{r \times n}$ 를 $C = V_r^T - W + Y_1 / \mu$ 로 정의하였을 때, Q 의 i 번째 열인 q_i 의 최적해는 수학식 20과 같이 계산된다.

수학식 20

$$q_i = \begin{cases} \left[\frac{\delta \alpha_i c_{1i}}{\delta \alpha_i + s_1^2}, \dots, \frac{\delta \alpha_i c_{ri}}{\delta \alpha_i + s_r^2} \right]^T & \text{if } \| S_r^{-1} c_i \| > \frac{1}{\delta} \\ 0_r & \text{otherwise,} \end{cases}$$

[0066]

[0067] 여기서, $\delta = \mu / \lambda$ 이고, s_j 는 대각 행렬 S_r 의 j 번째 특이값이다. c_i 는 C 의 i 번째 열이며, c_{ji} 는 c_i 의 j

$$a_i \triangleq \sum_{j=1}^r \left(\frac{c_{ji} s_j}{\delta a_i + s_j^2} \right)^2 = \frac{1}{\delta^2}$$

번째 원소를 의미한다. 그리고, a_i 는 $\frac{1}{\delta^2}$ 의 유일한 양수 근이며, 근 찾기 알

고리즘을 사용하여 값을 구할 수 있다.

[0068] FaLRR은 근을 찾기 위해 이분법(Bisection Method)를 사용하고 근의 최대하계를 0, 최소상계를 임의의 양수로 제안한다. 그런데 이렇게 하면 많은 시간이 소요되므로, 최대하계가 정확하지 않아도 동작하며, 이분법보다 빠른 Brent's Method를 사용한다. 그리고 최대하계는 이전 반복에서의 ai를 사용한다.

[0069] Y1, Y2 및 종료 조건 등은 도 1과 같이 FaGLRR 알고리즘과 같이 설정한다.

[0070] [Fast MLRR]

[0071] MLRR의 문제는 Z의 각 열마다 희소도가 높도록 조정(즉, 각 열의 Norm이 최소화 됨)하는 항의 $\beta \|Z\|_1$ 인데, 이 항을 FaLRR 및 FaGLRR에서처럼 W에 대하여, 또는 Z보다 작은 크기의 행렬로 표현하면서 동시에 t수학식 6과 동일한 해를 찾아내기는 어렵다. 하지만 이후 실험 결과에 따르면 어떤 데이터들(ORL, Extended Yale B, COIL20)에 대해서는 $\|Z\|_1$ 을 최소화하는 목표가 추가될 경우 더 좋은 결과를 나타내므로, 이 희소성 목표를 포함하는 빠른 알고리즘 개발을 계속하여 추진할 필요가 있다.

[0072] 즉 Fast MLRR에서 최소화하는 목표는 다음 수학식 21과 같이, MLRR에서 희소도와 관련된 항을 빼고 SMA와 관련된 항만 포함하는 형태로 설정한다. 그리고 일단 $Z \geq 0$ 의 조건을 제외한다.

수학식 21

$$\min_{Z,E} \|Z\|_* + \lambda \|E\|_{2,1} + \beta \text{tr}(ZGZ^T) \text{ s.t. } X = AZ + E$$

[0073]

[0074] 이 식의 목표는 또한 식에서 G 대신 L을 대입하면 GLRR의 수학식 2와 완전히 동일한 형태가 되므로, 수학식 2가 수학식 5와 동일한 목표를 최적화한다는 것을 응용하면, 역시 수학식 21과 동일한 최적 해를 도출하는 W에 대한 목표 식을 수학식 22와 같이 정의할 수 있다. 이를 뒷받침하는 정리 및 증명을 구성하기 위해서는 위의 Fast GLRR 정리에서 G 대신 L을 대입함으로써 정리 및 증명을 다시 기술하면 되는데, 이는 자명하므로 생략한다.

수학식 22

$$\min_W \|W\|_* + \lambda \|S_r(V_r^T - W)\|_{2,1} + \beta \text{tr}(WGW)$$

[0075]

[0076] 따라서 FaGLRR과 유사한 갱신 식을 사용하여 FaMLRR을 수행할 수 있다. 우선 수학식 10과 마찬가지로, Nuclear Norm 부분의 W 만 J로 변경한 후, 수학식 11과 같이 Augmented Lagrangian Function을 정의한다. 그리하면 J는 FaGLRR과 동일하게 수학식 14를 사용하여 갱신할 수 있다. 그리고 W는 다음의 수학식 23을 사용하여 갱신할 수 있다. 이 또한 수학식 15에서 L과 G로 바꾼 것과 동일하다.

수학식 23

$$W = \frac{\mu}{2} \left(V_r^T - Q + J + \frac{Y_1}{\mu} - \frac{Y_2}{\mu} \right) (\beta L + \mu I)^{-1}$$

[0077]

[0078] 또 다른 방법으로 FaGLRR과 마찬가지로, 우선 $\beta G = U_2 \Sigma_2 U_2^T$ 로 분해한 후 $M = \mu/2 \times (V_r^T - Q + J + Y_1/\mu - Y_2/\mu)$ 로 설정하면, $T = M_2, \tilde{W}_{ij} = t_{ij}/(\mu + \sigma_2^j)$

으로 동일하게 정리되므로 $W = \widetilde{W}U_2^T$ 를 사용하여 W 를 갱신한다. 역시 FaGLRR과 마찬가지로, 역행렬 계산이 반복 수행되는 것을 피하기 위하여 후자의 방법을 기본으로 사용하는 것을 권장한다. 정리하자면 도 2와 같은 FaMLRR 알고리즘으로 설명한다.

- [0079] Z^* 가 계산된 이후에는 간단히 그래프를 구축할 수 있다. Z^* 의 각 열 z_i^* 는 x_i 를 재구축하는 기여도이므로, 모든 열을 $z_i^* = z_i^* / \|z_i^*\|_2$ 로 정규화한 후, 문턱값(Threshold) θ 보다 작은 값을 0 으로 만듦으로써, 그래프 행렬 w 를 수학적식 24과 같이 계산한다.

수학적식 24

$$w = \frac{\widehat{Z}^* + (\widehat{Z}^*)^T}{2}, \text{ for } \widehat{Z}_{ij}^* = \begin{cases} z_{ij}^* & \text{if } z_{ij}^* > \theta \\ 0 & \text{otherwise} \end{cases}$$

- [0080]
- [0081] 전술한 방법에 따른 그래프 구축방법에 대하여 실험하여 성능을 분석하였다. 실험에는 2개의 얼굴, 1개의 물체, 1개의 음성 데이터셋 즉 ORL (c=40, d=1,024, n=400), Extended Yale B(c=38, d=1,024, n=2,414), COIL100(c=100, d=1,024, n=7,200), Isolet(c=26, d=617, n=7,797)을 사용한다(c는 클래스 수). 데이터셋들은 각각 다른 기관에서 제공하지만 여기서는 Cai의 MATLAB 기반 데이터셋을 활용한다. Isolet을 제외한 데이터셋은 제공되는 것을 그대로 사용하고, Isolet은 Isolet1~5를 합쳐서 사용한다. 그리고 각 데이터에 대해 레이블된 비율을 10~40%로, 그리고 각 레이블 비율마다 50번의 실험을 수행하여 평균을 측정한다.
- [0082] 레이블 추론 방법으로는 Label Spreading 방법을 사용하며 α 는 0.5로 설정한다. 그리고 LRR 기반 알고리즘마다 데이터셋에 따라 최적 분류 성능을 달성하는 매개 변수 조합이 다르므로, $\lambda \in \{0.5, 1, 2, 3, 5, 10\}$, $\beta = \{0.1, 0.1, 2, 4\}$ 중에서 최적 값을 찾는다. 나머지 매개 변수들은 앞서 설명한 도 1 및 도 2의 알고리즘 코드에서와 동일하게 설정한다. 수학적식 24에서의 θ 값으로는 0.2를 사용한다.
- [0083] 도 3은 본 발명의 바람직한 실시예에 따른 그래프 구축 방법에 따른 실험 결과 표이다.
- [0084] 도 3에서, GLRR과 FaGLRR의 정확도는 거의 동일하다. 이로써 수학적식 5는 수학적식 4와 동일한 해를 도출함이 실험적으로 증명된다. 그리고 준지도 학습에서 GLRR과 FaGLRR은 LRR과 FaLRR보다 높은 분류 정확도를 보이며, 제안한 FaGLRR이 GLRR보다 매우 빠른 것을 확인할 수 있다.
- [0085] FaMLRR은 LRR 및 FaLRR과 비교하여 분류 성능 향상 폭이 적지만, YaleB 데이터셋에서는 희소성 표현을 추가로 고려하는 MLRR과 NNLS를 통하여 구축된 그래프를 사용하였을 때 다른 방법들에 비해 매우 높은 분류 정확도를 나타내는데, 이는 데이터셋에 따라 희소성 표현을 고려하는 것이 더 좋은 데이터 표현을 도출할 가능성이 있음을 보여 준다. 그리고 전체적으로 FaGLRR과 FaMLRR은 다른 것에 비해 매우 빠른 속도로 해를 도출할 수 있음을 확인할 수 있다.
- [0086] 도 4는 ORL (d = 1024, n = 400) 에 따른 결과 그래프이며, 도 5는 Extended Yale B(d = 1024, n = 1000) 에 따른 결과 그래프이며, 도 6은 COIL100 (d = 1024, n = 7200) 에 따른 결과 그래프이며, 도 7은 Isolet (d = 617, n = 7797)에 따른 결과 그래프이다.
- [0087] 도 4 내지 도 7의 꺾은선 그래프를 살펴보면 도 1과 도 2에서는 희소성 표현 관련 목표가 추가된 MLRR과 NNLS를 통한 그래프 구축이 더 높은 분류 정확도를 보인다. 이에 비해 막대그래프를 살펴볼 때는, FaGLRR과 FaMLRR의 수행 시간은 MLRR과 NNLS에 비해 매우 짧고, 기존의 Inexact ALM 기반 LRR 방법인 GLRR보다도 매우 빠른 것을 확인할 수 있다. 도 3과 도 4에서도 FaLRR에 비해 FaGLRR과 GaMLRR은 수행시간이 다소 늘어나지만 더 좋은 그래프를 구축할 수 있음을 알 수 있다.
- [0088] 이 뿐만 아니라 모든 그래프에서 GLRR과 FaGLRR의 꺾은선 그래프가 겹치는 것을 볼 수 있는데, 이로서 본 발명

에 따른 방법인 FaGLRR은 FaMLRR과 동일한 목표를 더욱 시간에 효율적으로 최적화함을 실험을 통해서도 확인할 수 있다.

[0089]

이상에서 본 발명에 대하여 그 바람직한 실시예를 중심으로 설명하였으나, 이는 단지 예시일 뿐 본 발명을 한정하는 것이 아니며, 본 발명이 속하는 분야의 통상의 지식을 가진 자라면 본 발명의 본질적인 특성을 벗어나지 않는 범위에서 이상에 예시되지 않은 여러 가지의 변형과 응용이 가능함을 알 수 있을 것이다. 그리고, 이러한 변형과 응용에 관계된 차이점들은 첨부된 청구 범위에서 규정하는 본 발명의 범위에 포함되는 것으로 해석되어야 할 것이다.

도면

도면1

Input: 데이터셋 X , 최근접 이웃 수 k , 매개 변수 λ , a

Output: 최적 해 W^*, Q^*, Z^*, E^*

1 :

$W = Q = J = Y_1 = Y_2 = 0, \mu = 10^{-1}, \rho = 1.1, \epsilon = 10^{-8}, k = 5$

2: X 를 사용하여 k-NN 그래프 생성 후, L 을 계산

3: Skinny SVD를 사용하여 $X = U_r S_r V_r^T$ 으로 분해

4: SVD를 사용하여 다음을 수행: $\beta L = U_2 \Sigma_2 U_2^T$

5: while $\|V_r^T - W - Q\|_\infty > \epsilon$ do

6: J 를 식 (14)에 따라 갱신(다른 변수 값 고정)

7: W 를 식 (19)에 따라 갱신(다른 변수 값 고정)

8: Q 를 식 (20)에 따라 갱신(다른 변수 값 고정)

9: $Y_1 = Y_1 + \mu(V_r^T - W - Q), Y_2 = Y_2 + \mu(W - J)$

10: $\mu = \min(\rho\mu, \mu_{\max})$

11: end while

12: $W^* = W, Q^* = Q, J^* = J, Z^* = V_r W^*, E = X - X V_r W^*$

도면2

Input: 데이터셋 X , 매개 변수 λ, β , 최근접 이웃의 수 N_i 및 매개 변수 γ (Sparse Manifold Adaption(SMA) [3]을 위해)

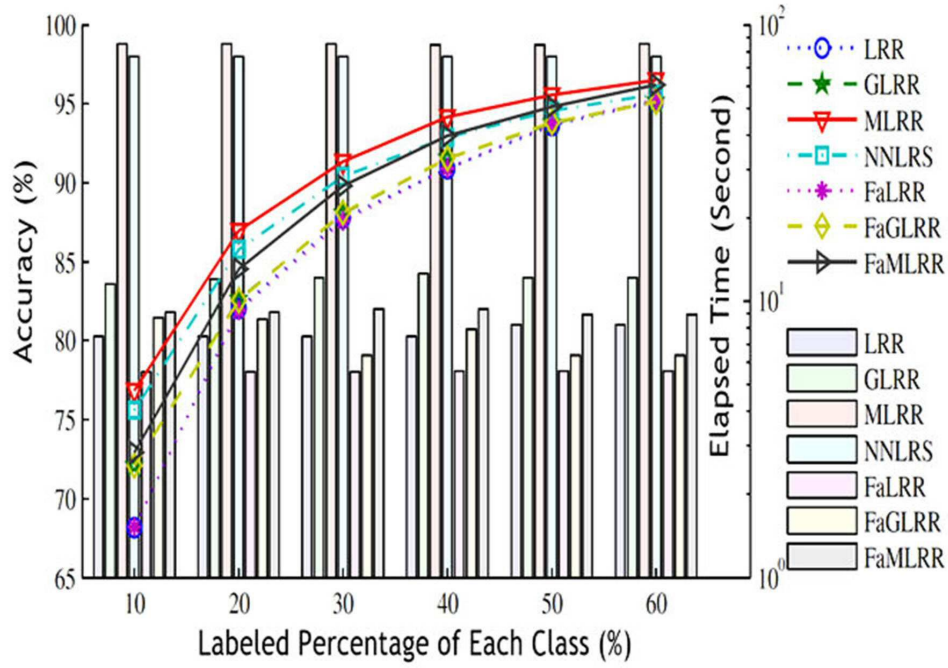
Output: 최적 해 W^*, Q^*, J^*, Z^*, E^*

- 1: $W = Q = J = Y_1 = Y_2 = 0, \mu = 10^{-6}$ (또는 10^{-1}),
 $\mu_{\max} = 10^{10}, \rho = 1.1, \epsilon = 10^{-8}, \gamma = 1, N_i = 20$
- 2: 데이터 X 를 SMA를 수행하여 A 를 계산,
 $G = (I - A)(I - A)^T$
- 3: Skinny SVD를 사용하여 $X = U_r S_r V_r^T$ 으로 분해
- 4: W 의 갱신에 식 (17)을 사용할 경우에만 다음의 SVD
수행: $\beta G = \Sigma_2 U_2^T$
- 5: while $\|V_r^T - W - Q\|_{\infty} > \epsilon$ do
- 6: J 를 식 (14)에 따라 갱신(다른 변수 값 고정)
- 7: W 를 식 (20)에 따라 갱신(다른 변수 값 고정)
- 8: Q 를 식 (21)에 따라 갱신(다른 변수 값 고정)
- 9: $Y_1 = Y_1 + \mu(V_r^T - W - Q), Y_2 = Y_2 + \mu(W - J)$
- 10: $\mu = \min(\rho\mu, \mu_{\max})$
- 11: end while
- 12: $W^* = W, Q^* = Q, J^* = J, Z^* = V_r W^*, E = X - X V_r W^*$

도면3

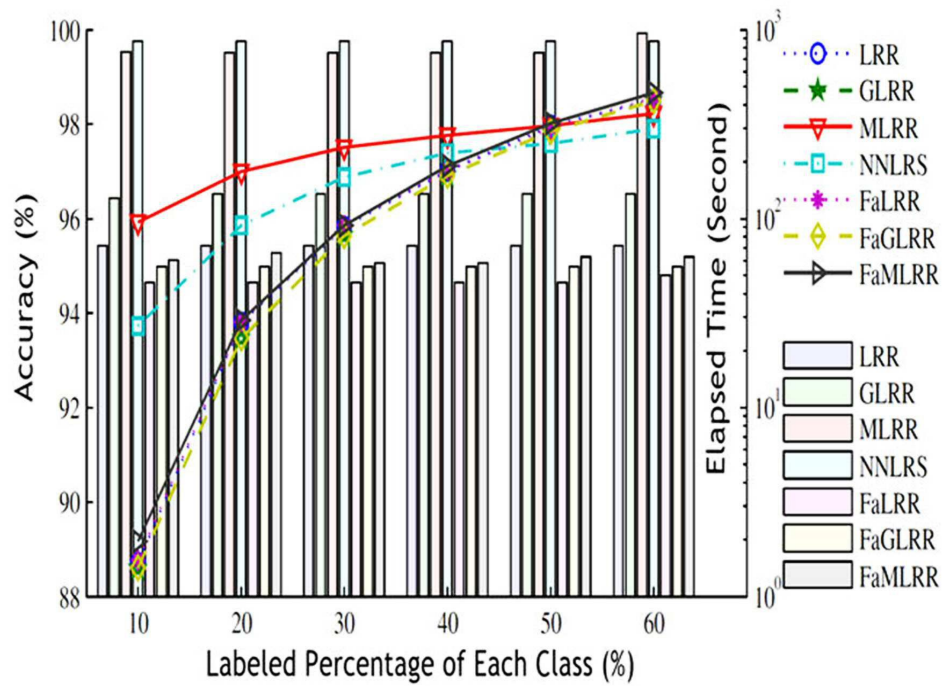
Dataset	%Label	LRR		GLRR		FaLRR		FaGLRR		FaMLRR	
		ACC	TIME	ACC	TIME	ACC	TIME	ACC	TIME	ACC	TIME
ORL	10%	0.681	7.5	0.721	11.5	0.682	5.6	0.721	8.7	0.7291	9.133
	20%	0.820	7.5	0.826	12.0	0.819	5.6	0.826	8.6	0.8456	9.133
	30%	0.878	7.5	0.881	12.2	0.876	5.6	0.881	6.4	0.8979	9.348
	40%	0.909	7.5	0.915	12.5	0.909	5.6	0.915	7.9	0.9289	9.348
Yale_B	10%	0.837	171.4	0.842	1550.8	0.837	139.9	0.842	209.8	0.8537	232.0
	20%	0.911	216.2	0.912	1550.8	0.911	129.4	0.912	209.8	0.9197	242.7
	30%	0.946	216.2	0.946	1728.4	0.946	129.4	0.946	194.3	0.9513	242.7
	40%	0.962	216.2	0.962	1728.4	0.962	129.4	0.962	194.3	0.9661	242.7
COIL100	10%	0.668	479.0	0.729	38362.7	0.668	444.6	0.729	1419.5	0.6920	1321.4
	20%	0.727	479.0	0.789	49181.5	0.727	444.6	0.789	1472.0	0.7619	1362.1
	30%	0.763	743.7	0.819	54072.3	0.764	490.5	0.820	1531.1	0.8026	1362.1
	40%	0.791	743.7	0.840	54072.3	0.791	490.5	0.840	1531.5	0.8290	1022.6
Isolet	10%	0.806	246.2	0.895	46853.7	0.810	236.0	0.895	896.2	0.8904	810.7
	20%	0.875	246.2	0.915	52896.0	0.877	236.0	0.915	793.8	0.9122	810.7
	30%	0.898	246.2	0.923	52896.0	0.899	236.0	0.923	793.8	0.9219	849.4
	40%	0.909	246.2	0.927	58836.0	0.910	236.0	0.927	818.5	0.9267	849.4

도면4



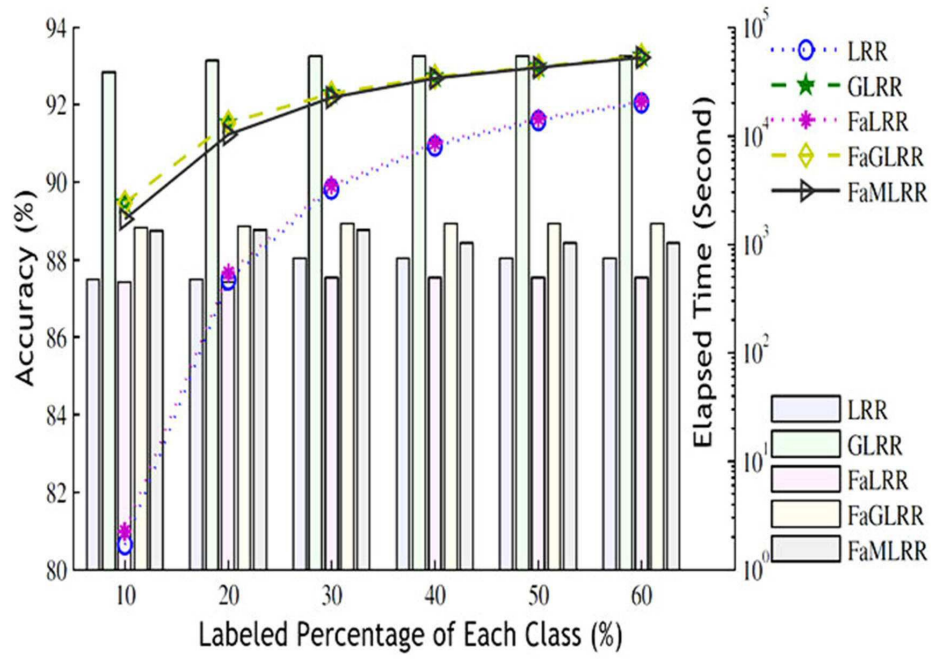
ORL, ($d = 1024, n = 400$)

도면5



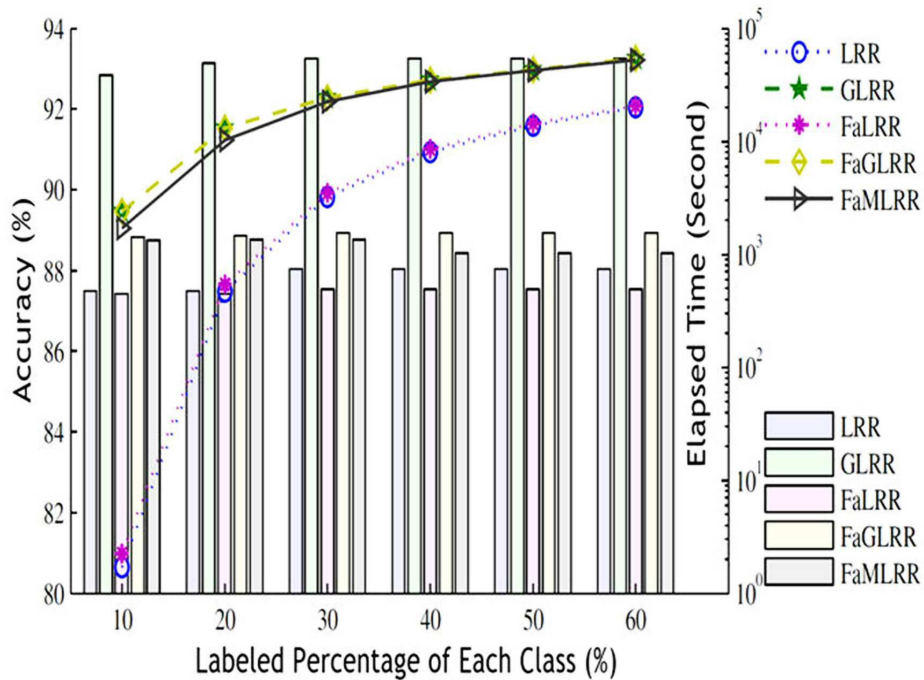
Extended Yale B ($d = 1024, n = 1000$)

도면6



COIL100($d = 1024, n = 7200$)

도면7



Isolet($d = 617, n = 7797$)