

최대 부피 Simplex 기반의 Isomap을 위한 랜드마크 추출

지 준 화[†]

퍼듀대학교 토목공학과

Maximum Simplex Volume based Landmark Selection for Isomap

Junhwa Chi[†]

School of Civil Engineering, Purdue University

Abstract : Since traditional linear feature extraction methods are unable to handle nonlinear characteristics often exhibited in hyperspectral imagery, nonlinear feature extraction, also known as manifold learning, is receiving increased attention in hyperspectral remote sensing society as well as other community. A most widely used manifold Isomap is generally promising good results in classification and spectral unmixing tasks, but significantly high computational overhead is problematic, especially for large scale remotely sensed data. A small subset of distinguishing points, referred to as landmarks, is proposed as a solution. This study proposes a new robust and controllable landmark selection method based on the maximum volume of the simplex spanned by landmarks. The experiments are conducted to compare classification accuracies with standard deviation according to sampling methods, the number of landmarks, and processing time. The proposed method could employ both classification accuracy and computational efficiency.

Key Words : Manifold learning, feature extraction, classification, landmark, Isometric feature mapping

요약 : 초분광 영상에 내재된 비선형 현상을 다루기 위해서는 과거에 주로 사용되었던 선형 피쳐 추출 방법은 적합하지 않았다. 따라서 최근 Manifold learning이라 불리는 비선형 피쳐 추출 방법이 초분광 원격 탐사 분야를 비롯 여러 분야에서 관심이 증가되고 있다. Manifold learning 방법 중 널리 이용되는 Isomap은 분류와 분광 혼합 분석 등의 분야에서 좋은 결과를 보여주지만, 지나치게 복잡하고 높은 계산량, 특히 원격탐사 자료와 같이 자료의 크기가 큰 경우 문제가 된다. 따라서 자료의 일부분을 이용하는 랜드마크 기법이 해결책으로 제안되었다. 본 연구에서는 좀 더 통제가 가능한 랜드마크 추출을 위해 자료를 구성하는 최대 부피를 지닌 Simplex를 이용하여 랜드마크를 선택하는 방법을 제안한다. 초분광 영상을 이용하여 랜드마크의 개수, 선택 방법에 따른 분류 정확도와 편차, 그리고 처리 시간을 비교하였고, 그 결과 제안된 랜드마크 선택 기법은 분류 정확도, 처리시간 모두에서 효율적인 결과를 보여주었다.

접수일(2013년 9월 24일), 수정일(1차 : 2013년 10월 15일), 게재확정일(2013년 10월 17일).

[†] 교신저자: 지준화(junhwa.chi@gmail.com)

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited

1. 서론

초분광 원격탐사 기술은 지상 대상물의 분광 정보를 연속된 파장에 걸쳐 기록하므로 지난 십여년간 원격탐사 분야에서 큰 주목을 받아왔다. 초분광 영상은 다중분광 영상에서는 불가능한 미세한 분광 정보 차이를 구분할 수 있지만, 상관 관계가 매우 높은 연속된 밴드들과 많은 밴드수는 일반적으로 처리와 분석에 있어 어려움을 낳았다. 특히, Bi-directional Reflectance Distribution Function(BRDF), 다중 산란(Multiple scattering), 비선형적 쇠퇴(Nonlinear attenuation) 현상 등 다양한 원인에 의해, 특정 밴드간, 사물에 대해서 비선형적(Nonlinear) 관계가 Fig. 1과 같이 나타난다(Bachmann *et al.*, 2006; Goodin *et al.*, 2004).

따라서 과거에 전통적으로 사용되던 Principal Component Analysis(PCA)와 같은 선형 기반의 피쳐 추출 기법(DR: Dimensionality Reduction or Feature Extraction)은 이러한 비선형 현상을 잘 표현 할 수가 없었다. 최근 이런 문제점을 다루기 위하여 기계 학습(Machine learning) 연구 그룹에서는 Manifold Learning(ML)이라고 불리는 비선형 DR 기법이 제안되었고, 초분광 원격탐사 분야에서도 널리 이용되고 있다(Bachmann *et al.*, 2005; Bachmann *et al.*, 2006; Li *et al.*, 2012; Mathieu *et al.*, 2009).

ML은 Isometric feature mapping(Isomap)과 같은 전역적(Global-ML) 방법(Tenenbaum *et al.*, 2000)과 Locally

Linear Embedding(LLE)과 같은 지역적(Local-ML) 접근 방법이 있는데(Roweis and Saul, 2000), 전역적 방법은 자료의 전반적인 기하(Geometry) 구조를 영상 전체에 걸쳐 잘 보존 하는 방법이고, 지역적 접근법은 자료 전체에 걸쳐 산재된 복잡한 분광 구조를 가진 기하를 더 잘 표현할 수 있는 특징이 있다. 가장 널리 사용되는 G-ML 중 하나인 Isomap은 분류(Classification)에 있어 다른 전통적 DR 기법에 비해 분광학적으로 구분이 어려운 클래스들에 대해 높은 정확도를 보여주지만, 최단 경로(Shortest path)와 Multidimensional scaling(MDS) 계산과 관련된 부분에서 병목 현상이 발생하여 매우 높은 물리적 메모리양과 처리 시간을 요구를 한다(Silva and Tenebaum, 2004). 특히, 화소수가 많은 원격탐사 자료의 경우 이러한 문제가 더욱 심각해진다. 따라서 이를 해결하기 위하여 구분되어지는 자료의 일부분(Subset)을 이용하여 나머지 자료의 근사치를 예측하는 랜드마크(Landmark) 방식이 제안되었다(Silva and Tenebaum, 2004).

복잡하지 않은 자료 구조의 경우 충분한 수의 랜드마크를 무작위로 선택하여도 자료를 낮은 차원의 공간에 투영하는데 크게 문제가 없어 일반적으로 사용되고 있지만, 이는 결과에 있어 편차가 심하고, 분광학적으로 중복된 랜드마크가 선정될 수 있어, 영상에 복잡한 분광 구조를 지닌 기하들이 많이 내재되어 있는 경우 이를 표현하는데 어려움이 따른다. 따라서 본 연구에서는 영상

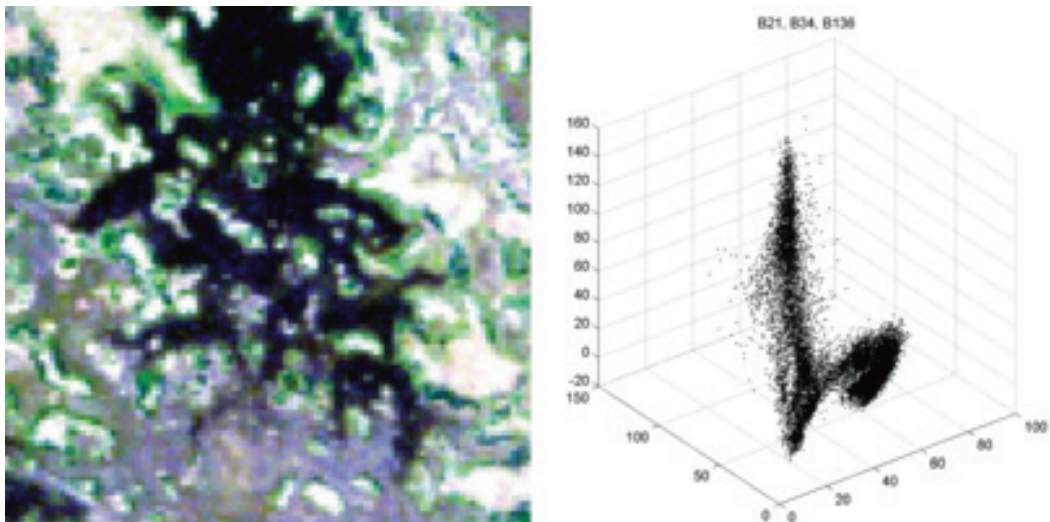


Fig. 1. Effects of nonlinear phenomena on hyperspectral bands.

을 구성하는 분광학적으로 의미있는 화소들을 이용하여 좀 더 통제가 가능한 랜드마크를 선정하는 방법을 제안하고, 이를 초분광 영상의 감독 분류(Supervised classification)와 처리 시간 비교를 통해 그 성능을 검증하고자 한다.

2. Isomap과 Landmark Isomap(L-Isomap)

예를 들어, D 개의 분광 밴드와 N 개의 화소수를 가진 초분광 영상 $\mathbf{X} \in \mathbf{R}^{D \times N}$ 에서 DR은 \mathbf{X} 를 $\mathbf{R}^{d \times N}$ 에 원 자료의 기하 구조를 가능한 충실하게 보존하여, d 개의 낮은 차원 공간에 투영하는 것이 목적이다. Isomap은 다음과 같이 크게 세가지 단계로 구성된다(Tenenbaum *et al.*, 2000): 1) k -최근근(Nearest neighbor) 방법을 이용한 인접 화소 그래프 G 정의; 2) G 로부터 최단 거리 행렬 계산; 3) 새 투영 좌표 계산을 위한 MDS 계산. Table 1은 Isomap의 각 단계별 계산에 필요한 비용을 보여주는데, 최단 거리 행렬과 MDS 계산 과정의 비용이 매우 높은 것을 알 수 있다.

반면에, L-Isomap은 랜드마크를 이용하여 이 두 과정의 계산 비용을 줄이는데 목적이 있다. 즉, n 개의 랜드마크를 사용할 경우, 최단거리 행렬과 MDS 계산에 있어, Isomap의 $N \times N$ 거리 행렬 대신 $n \times N$ 크기의 행렬이 필요하다. L-Isomap 역시 세가지 단계로 구성되는데(Silva and Tenebaum, 2004): 1) 샘플링 방법에 의한 n 개의 랜드마크 선택; 2) 각 랜드마크간 거리를 나타내는 $n \times n$ 행렬을 이용하여 MDS 계산; 3) 랜드마크와 나머지 화소간의 $n \times N$ 거리 행렬을 이용한 d 공간에 투영. 따라서 Table 1에서 보듯이 $n \ll N$ 이므로 최단 거리와 MDS 계산 과정의 비용이 크게 줄어들음을 알 수 있다. L-Isomap을 위한 랜드마크의 수는 자료 구조의 복잡도에 따라 다르지만, 이론적으로 필요한 최소 개수는 $d+1$ 이며, 일반적으로 새로운 투영 공간 좌표의 안정성을 유지하기 위하

Table 1. Computational complexity comparison between Isomap and L-Isomap

	Isomap	L-Isomap
k-nearest neighbor graph	$O(N^2)$	$O(N^2)$
Shortest path computation	$O(kN^2 \log N)$	$O(knN \log N)$
MDS computation	$O(N^3)$	$O(n^2 N)$

여 최소 랜드마크 수보다 많이 선택한다.

3. Maximum Simplex Volume (MSV) 랜드마크 선택

양질의 랜드마크는 분광학적으로 중복(Redundancy)을 줄이는 동시에 다양성(Diversity)을 높여야 한다. 분광 혼합 분석(Spectral mixture analysis)에서 가장하듯이 중저해상도 위성 영상의 대부분의 화소는 몇 가지의 순수한 분광 신호들로 표현할 수 있다. 따라서 Endmember라 불리는 순수 화소들로 이루어진 기하들의 Convexity를 이용하는 특성은 매우 자연스럽고 논리적인 개념이므로, Endmember 추출(Extraction) 알고리즘에서 널리 사용되고 있다(Winter, 1999; Chang *et al.*, 2006; Kim, 2012).

본 연구에서는, 영상에서 최대 부피를 구성하는 Simplex를 이용하여 좀 더 통제가 가능하며, 의미있는 분광 정보를 지닌 화소들을 랜드마크로 이용하고자 한다. $n+1$ 개의 랜드마크 $[l_0, l_1, \dots, l_n]$ 로 구성된 Simplex의 부피는 다음 수식1과 2와 같이 계산할 수 있다(Winter, 1999).

$$\mathbf{M} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ l_0 & l_1 & \dots & l_n \end{bmatrix} \quad (1)$$

$$V(\mathbf{M}) = \frac{1}{n!} |\det(\mathbf{M})| \quad (2)$$

수식2에서 \mathbf{M} 의 행렬식(Determinant)이 존재하기 위해서는 \mathbf{M} 은 반드시 정방 행렬(Square matrix)이어야 한다. 하지만 일반적으로 $n > d$ 이므로 수식2는 성립할 수가 없다. 따라서 Simplex의 부피는 Gram determinant를 이용하여 다음과 같이 새롭게 계산 되어야 한다.

$$V(\hat{\mathbf{M}}) = \frac{1}{n!} |\det(\hat{\mathbf{M}})| \quad (3)$$

수식3에서 $\hat{\mathbf{M}} = [l_1 - l_0, l_2 - l_0, \dots, l_n - l_0]$ 은 원점 l_0 과 나머지 꼭지점(Vertex)들로 이루어진 n -simplex의 집합 행렬이다. 임의의 행(열)에 다른 행(열)의 실수 배를 해서 더하여도 행렬식의 결과는 일정하다는 행렬식의 특성에 의해, \mathbf{M} 과 $\hat{\mathbf{M}}$ 은 동일한 행렬식을 가진다. 따라서 수식2와 3은 동일하다. $|\det(\hat{\mathbf{M}})| = \sqrt{|\det(\hat{\mathbf{M}}^T \hat{\mathbf{M}})|}$ 이므로, Gram-Schmidt orthogonalization에 의한 Gram determinant

$\det(\hat{\mathbf{M}})$ 은 다음과 같이 계산된다 (Berger, 1987).

$$\det(\hat{\mathbf{M}}) = \begin{bmatrix} \mathbf{a}_1^T \mathbf{a}_1 & \mathbf{a}_1^T \mathbf{a}_2 & \dots & \mathbf{a}_1^T \mathbf{a}_n \\ \mathbf{a}_2^T \mathbf{a}_1 & \mathbf{a}_2^T \mathbf{a}_2 & \dots & \mathbf{a}_2^T \mathbf{a}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_n^T \mathbf{a}_1 & \mathbf{a}_n^T \mathbf{a}_2 & \dots & \mathbf{a}_n^T \mathbf{a}_n \end{bmatrix}^{1/2} \quad (4)$$

여기서 $\mathbf{a}_1 = \mathbf{l}_1 - \mathbf{l}_0$, $\mathbf{a}_2 = \mathbf{l}_2 - \mathbf{l}_0$, ..., $\mathbf{a}_n = \mathbf{l}_n - \mathbf{l}_0$ 이다. 가장 간단하게 Gram determinant를 계산하는 방법은 직교 벡터를 이용하는것으로, Gram-Schmidt orthogonalization 후의 $[\mathbf{a}_0, \mathbf{a}_1, \dots, \mathbf{a}_n]$ 를 $[\hat{\mathbf{a}}_0, \hat{\mathbf{a}}_1, \dots, \hat{\mathbf{a}}_n]$ 이라 가정할 때 $\det(\hat{\mathbf{M}})$ 은 수식5와 같이 계산할 수 있다.

$$\det(\hat{\mathbf{M}}) = \begin{bmatrix} |\hat{\mathbf{a}}_1|^2 & 0 & \dots & 0 \\ 0 & |\hat{\mathbf{a}}_2|^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & |\hat{\mathbf{a}}_n|^2 \end{bmatrix} = |\hat{\mathbf{a}}_1| |\hat{\mathbf{a}}_2| \dots |\hat{\mathbf{a}}_n| \quad (5)$$

따라서 Gram determinant를 사용하면 $\hat{\mathbf{M}}$ 은 항상 정방 행렬이 되므로 영상의 밴드 개수와 랜드마크의 개수와 관계없이 Simplex의 부피를 계산할 수 있다.

본 연구에서 이용한 최대 부피를 가지는 Simplex를 구하기 위해서는, 임의로 지정된 시작점으로부터 모든 화소에 대해 각 랜드마크를 반복하여 교체하는데, 다음 수식6과 같이 더 큰 부피를 가지는 Simplex를 찾을시 해당 랜드마크를 교체하게 된다.

$$\arg \max_{(\mathbf{l}_0, \mathbf{l}_1, \dots, \mathbf{l}_n)} V(\mathbf{l}_0, \mathbf{l}_1, \dots, \mathbf{l}_n) \quad (6)$$

4. 실험 방법 및 결과

본 연구에 사용된 Fig. 2(a)의 초분광 영상은 2001년 5월 Botswana 지역에서 획득된 공간 해상도 30 m의 Hyperion 영상으로, Fig. 2(b)와 Table 2에 표시한 것처럼 9개의 복잡한 식생, 물, 토양 등의 클래스로 구성되었다. 242개의 밴드중 대기 흡수밴드와 신호대 잡음비가 낮은 밴드를 제외한 145개의 밴드가 실험에 이용되었다.

정량적인 실험을 위해 클래스 정보가 있는 자료에 한해 50%는 훈련(Training), 나머지 50%는 검증(Test) 자료로 나누었으며, 결과의 신뢰도를 높이기 위해 같은 방법을 이용하여 무작위로 20개의 세트를 만들어 반복하여 실험하였다. 감독 분류를 위해 분류기(Classifier)에 의한 영향을 최소화하기 위하여 입력 인자가 필요없는 1-최근린(1-Nearest neighbor) 방법을 사용하였고, 랜드마크

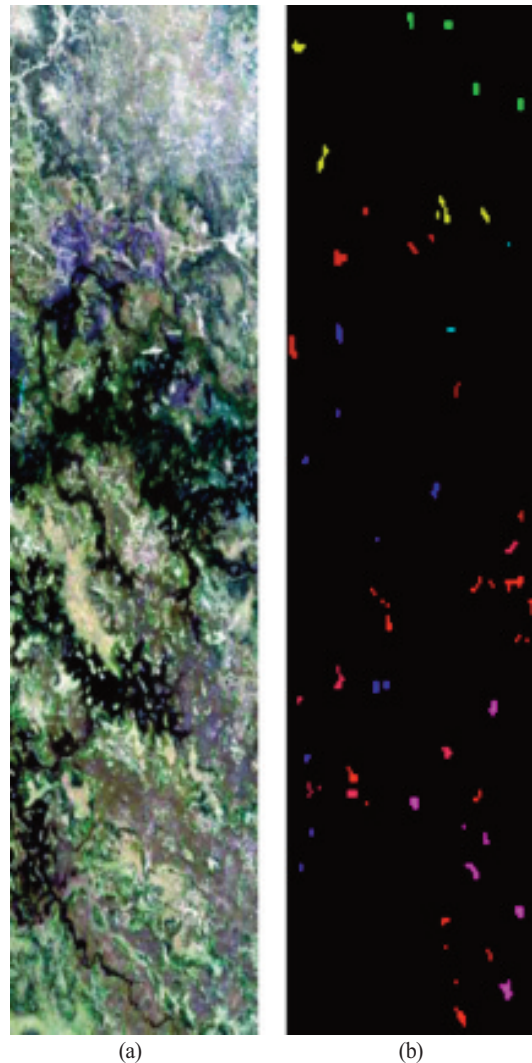
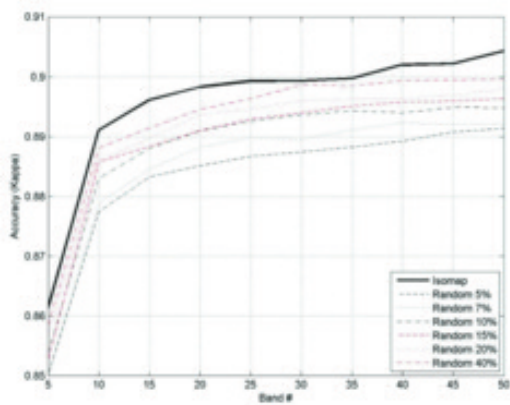


Fig. 2. RGB subset of Botswana (a) and labeled data (b).

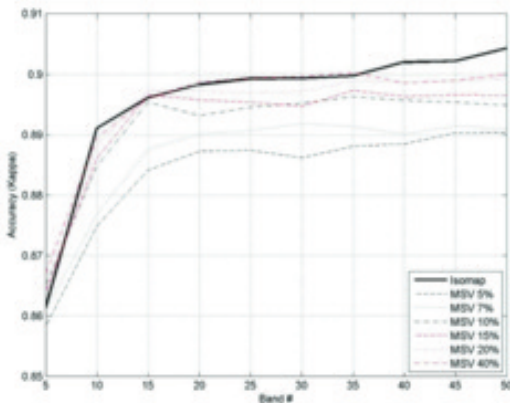
는 무작위(Random) 선택과 본 연구에서 제안한 MSV 샘플

Table 2. Class information with the number of samples

Class Name	No. of Samples
Water	297
Primary Floodplain	437
Riparian	448
Firescar	354
Island Interior	337
Woodlands	357
Savanna	330
Short Mopane	239
Exposed Soils	215
Total number of samples	3014



(a)



(b)

Fig. 3. Mean classification accuracies according to sampling methods and rate.

플링 방법을 사용하여 다음과 같이 비교, 분석 하였다.

사용한 랜드마크의 비율에 따른 각 샘플링 방법의 평균 분류 정확도(Kappa statistics)를 Fig. 3에 표시하였는데, 두 가지 랜드마크 샘플링 방법 모두 랜드마크의 개수가 증가할 수록 분류 정확도 역시 증가였으며, 일정 비율의 랜드마크가 사용 되었을때 결과가 점차 수렴하는 것을 알 수 있다. 특히 10% 이상의 랜드마크 사용 시 본 연구에서 제안한 MSV 샘플링 방법의 정확도가 무작위 추출에 비해 증가하는 것을 알 수 있는데, 이는 MSV에 의해 추출된 랜드마크들이 자료를 표현하기 위해 충분한 분광 정보를 담고 있음을 말해준다. 또한, 분류기를 적절히 훈련하기 위하여 필요한 밴드의 개수 역시 MSV 샘플링 방법이 무작위 선택에 비해 적은것을 알 수 있다. 예를 들어, 10%의 랜드마크를 사용했을 경우, 약 0.895의 분류 정확도를 얻기 위해서는 무작위 추출의 경

우 30개 이상의 입력 밴드가 요구되지만, MSV 방법의 경우 약 15개의 입력 밴드가 필요하였다.

Fig. 4는 각 샘플링 방법들 사이의 20회 반복 실험을 통해서 나타난 오차의 범위를 보여준다. 그 결과, 무작위 샘플링의 경우 오차의 범위가 Isomap이나 MSV 샘플링을 이용한 L-Isomap에 비해 샘플링 비율에 상관없이 약 2배정도 컸다. MSV 샘플링의 경우, 최대 부피를 가지는 Simplex를 이루는 랜드마크를 사용하므로 반복 실험에 관계없이 항상 같은 랜드마크를 이용하게 된다. 따라서 랜드마크 선택 과정에서 편차는 전혀 발생하지 않았고 20회 반복 실험을 위해 제작된 훈련, 검증 자료 세트에서 발생한 편차만 존재 하였다. Fig. 4(a)에서 보듯이 사용된 랜드마크가 전체 자료의 5%로 절대적 개수가 적은 경우 무작위 추출 방법이나 MSV 접근법의 분류 정확도 차이가 크지 않거나 무작위 샘플링의 정확도가 높는데, 이는 MSV 샘플링의 경우 랜드마크 선택시 이상값(Extreme 또는 Outlier)에서부터 우선적으로 랜드마크를 추출할 가능성이 높기 때문이다. 따라서 자료 전체의 미묘한 특징을 표현하기 위한 충분한 분광 정보를 가진 랜드마크가 MSV 샘플링 방법에 의해 선택되지 못했음을 의미한다. 반면에 무작위로 선택된 랜드마크들은 비록 개수는 적지만 자료 전체에서 일정한 분포로 선택되었으므로 상대적으로 자료 전체를 투영하는데 비교적 더 효과적인 역할을 한것으로 해석할 수 있다.

Fig. 5는 특정 수의 입력 피쳐(Feature)에서 랜드마크 비율에 따른 각 방법간 분류 정확도 차이를 보여준다. 랜드마크의 비율이 10% 이하인 경우 분류 정확도가 안정적으로 증가하지 못하고 편차가 비교적 심한것을 볼 수 있다. 이는 영상을 낮은 차원 공간에 투영하기 위해 필요한 랜드마크 수가 부족하여 생긴 문제라 볼 수 있다. Fig. 5(b)에서 보듯이 입력 피쳐의 수가 증가함에 따라 결과적으로 두 샘플링 방법간 분류 정확도는 수렴되지만, Fig. 5(a)에서 보듯이 15개의 입력 피쳐를 사용하고 샘플링 비율이 10% 이상인 경우, MSV 샘플링 방법의 분류 정확도가 Isomap의 정확도에 근접한것을 볼 수 있는데, 이는 사용된 랜드마크들의 개수는 적지만 충분한 분광 정보를 포함하고 있어, 영상을 더 낮은 차원의 공간에 성공적으로 투영할 수 있었고 분류기를 훈련하는데에도 더 큰 기여를 했음을 말해준다.

앞서 설명 했듯이, L-Isomap의 목적은 Isomap이 요구

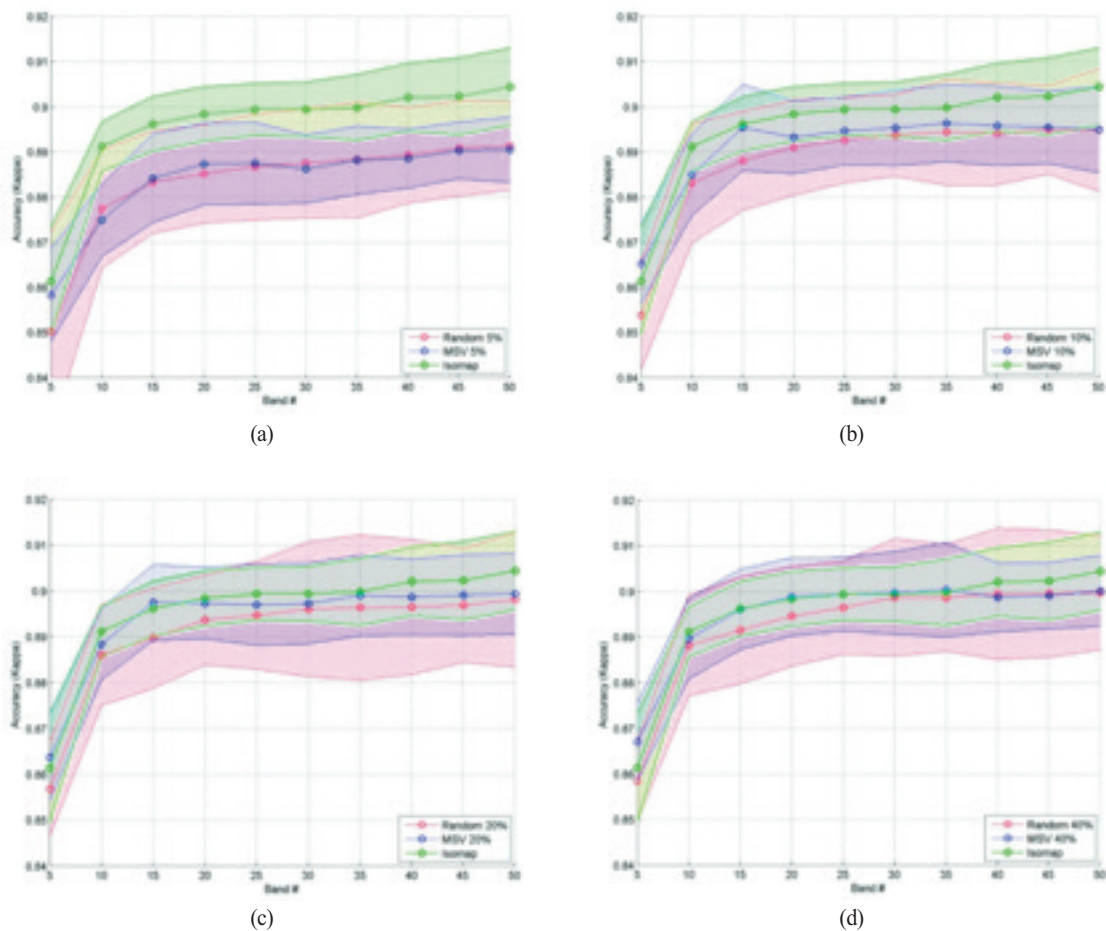


Fig. 4. Kappa statistics with standard deviation according to the ratio of landmarks: (a) 5%, (b) 10%, (c) 20%, and (d) 40%.

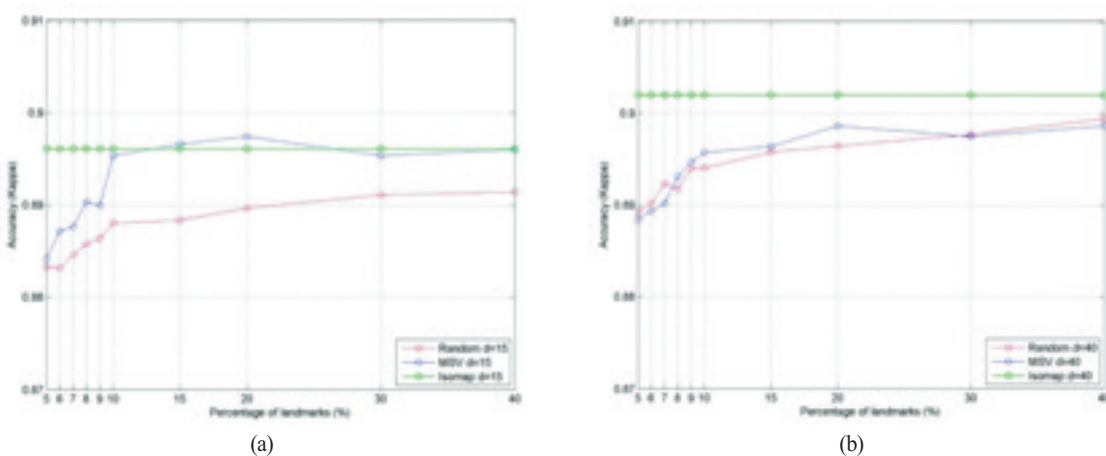


Fig. 5. Mean accuracies as a function of the input dimensionality at: (a) $d=15$, and (b) $d=45$.

하는 높은 메모리 양과 긴 처리 시간을 보완하는데 있다. Fig. 6은 Intel Core i7-3770 시스템에서 Matlab을 통해

구현 했을때, 랜드마크의 수에 따른 L-Isomap의 처리 시간을 표시하였다 (랜드마크의 비율이 100%인 경우

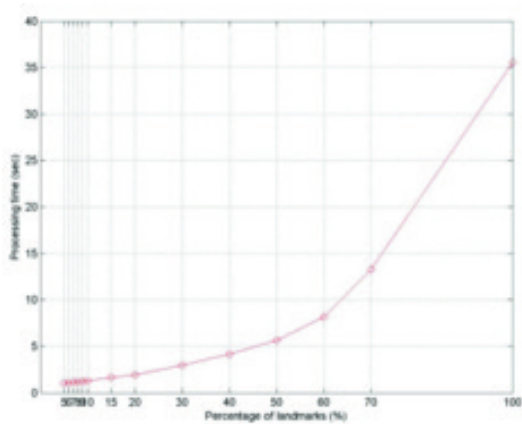


Fig. 6. Processing time comparison as a function of the ratio of landmarks.

Isomap과 같다). 그래프에서 보듯이 랜드마크의 개수가 증가할 수록 처리 시간이 선형적이 아닌 기하급수적으로 증가함을 볼 수 있다. 물론 MSV 샘플링 방법의 경우 무작위 선택에 비해 $O(mN)$ 의 추가 계산이 더 요구되지만, 비슷한 결과를 얻기 위해 필요한 랜드마크의 수와 입력 피처의 수가 더 적었으며, 반복 실험에 의한 편차도 발생하지 않았다.

5. 결론

본 연구에서는 비선형 DR 방법인 Isomap의 높은 계산 비용을 줄이기 위한 L-Isomap을 위한 랜드마크 선택 방법을 제안하였다. 근사치를 이용한 L-Isomap 방법은 이론적으로 Isomap의 좌표를 완벽히 복원할 수는 없지만, 본 연구에서 제안된 MSV 샘플링 방법을 통하여 어느정도 보완하여 Isomap에 준하는 분류 결과를 얻을 수 있었다. 제안된 랜드마크 추출 방법은 자료 전체를 구성할 수 있는 최대 부피를 지닌 Simplex의 화소들을 랜드마크로 사용하는 방법으로 크게 두가지 기여를 하였다. 첫째, 무작위 추출에 비해 좀 더 통제가 가능한 랜드마크를 추출할 수 있어 반복 실험간 편차를 줄일 수 있다. 둘째, 같은 수의 랜드마크를 사용하였을 경우, 적은 입력 피처를 사용했을 경우, 무작위 추출에 비해 높은 분류 정확도를 보이며, Fig. 3에서 보듯이 충분한 수의 랜드마크가 선택되었을 경우 Isomap에 준하는 분류 정확도를 보여주었다.

랜드마크수에 따라 일반적으로 분류 정확도는 증가하지만, Fig. 4에서 보듯이 랜드마크의 개수가 적더라도 높은 분류 정확도를 보여주는 경우도 생긴다. 이는 랜드마크의 수보다는 선택된 랜드마크들이 자료를 충분히 표현할 수 있는 분광 정보를 지녔는지가 중요하다고 볼 수 있다. 본 연구에서 제시한 Simplex의 부피를 이용한 방법은 영상의 분광 정보만을 이용하였다. 하지만, 영상 자료의 가장 큰 특징 중 하나인 공간 정보까지 활용하여 분광학적, 공간적 잉여를 동시에 제거한다며 더 효율적인 랜드마크를 개발할 수 있을 것으로 기대된다.

References

- Bachmann, C.M., T.L. Ainsworth, and R.A. Fusina, 2006. Improved manifold coordinate representations of large-scale hyperspectral scenes, *IEEE Transactions on Geoscience and Remote Sensing*, 44(10): 2786-2803.
- Bachmann, C.M., T.L. Ainsworth, and R.A. Fusina, 2005. Exploiting manifold geometry in hyperspectral imagery, *IEEE Transactions on Geoscience and Remote Sensing*, 43(3): 441-454.
- Berger, M., 1987. *Geometry I and II*, Springer, 19.
- Chang, C., C. Wu, W. Liu, and Y. Ouyang, 2006. A new growing method for simplex-based endmember extraction algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, 44(10): 2804-2819.
- Goodin, D.G., J. Gao, and G.M. Henebry, 2004. The effect of solar illumination angle and sensor view angle on observed patterns of spatial structure in tallgrass prairie. *IEEE Transactions on Geoscience and Remote Sensing*, 42(1): 154-165.
- Kim, K., 2012. A study on fast extraction of endmembers from hyperspectral image data, *Korean Journal of Remote Sensing*, 28(4): 347-355.
- Li, W., S. Prasad, J.E. Fowler, and L.M. Bruce, 2012.

- Locality preserving dimensionality reduction and classification for hyperspectral image. *IEEE Transactions on Geoscience and Remote Sensing*, 50(4): 1185-1198.
- Mathieu, F., C. Jocelyn, and B.J. Atli, 2009. Kernel principal component analysis for the classification of hyperspectral remote sensing over urban areas. *EURASIP Journal on Advanced in Signal Processing*, 11.
- Roweis S., and L. Saul, 2000. Nonlinear dimensionality reduction by local linear embedding, *Science*, 290(5500): 2323-2326.
- Silva V., and J. Tenenbaum, 2004. Sparse multidimensional scaling using landmark points, Technical report, Stanford Univ.
- Tenenbaum, J., V. Silva, and J. Langford, 2000. A global geometric framework for nonlinear dimensionality reduction, *Science*, 290(5500): 2319-2323.
- Winter, M.E., 1999. N-FINDR: an algorithm for fast autonomous spectral endmember determination in hyperspectral data, *Proc. of SPIE's International Symposium on Optical Science, Engineering, and Instrumentation*, pp. 266-275.