

Forensic Classification of Latent Fingerprints Applying Laser-induced Plasma Spectroscopy Combined with Chemometric Methods

Jun-Ho Yang and Jai-Ick Yoh[†]

*eXtreme Energy Laboratory, Department of Aerospace Engineering, Seoul National University,
1, Gwanak-ro, Gwanak-gu, Seoul 08826, Korea*

(Received February 21, 2020; Revised March 18, 2020; Accepted March 25, 2020)

An innovative method for separating overlapping latent fingerprints, using laser-induced plasma spectroscopy (LIPS) combined with multivariate analysis, is reported in the current study. LIPS provides the capabilities of real-time analysis and high-speed scanning, as well as data regarding the chemical components of overlapping fingerprints. These spectra provide valuable chemical information for the forensic classification and reconstruction of overlapping latent fingerprints, by applying appropriate multivariate analysis. This study utilizes principal-component analysis (PCA) and partial-least-squares (PLS) techniques for the basis classification of four types of fingerprints from the LIPS spectra. The proposed method is successfully demonstrated through a classification example of four distinct latent fingerprints, using discrimination such as soft independent modeling of class analogy (SIMCA) and partial-least-squares discriminant analysis (PLS-DA). This demonstration develops an accuracy of more than 85% and is proven to be sufficiently robust. In addition, by laser-scanning analysis at a spatial interval of 125 μm , the overlapping fingerprints were separated as two-dimensional forms.

Keywords: LIPS (laser-induced plasma spectroscopy), Latent fingerprint, Overlapping fingerprint, PCA (principal component analysis), Multivariate analysis

OCIS codes: (020.1335) Atom optics; (140.3440) Laser-induced breakdown; (300.6210) Spectroscopy, atomic; (300.6365) Spectroscopy, laser induced breakdown

케모메트릭 방법과 결합된 레이저 유도 플라즈마 분광법을 적용한 유류 지문의 법의학적 분류 연구

양준호 · 여재익[†]

서울대학교 기계항공공학부, 고에너지 응용 연구실
Ⓞ 08826 서울특별시 관악구 관악로 1

(2020년 2월 21일 받음, 2020년 3월 18일 수정본 받음, 2020년 3월 25일 게재 확정)

본 논문에서는 다변량 분석법과 결합된 레이저 유도 플라즈마 분광법을 사용하여 겹친 유류 지문을 분리하는 혁신적인 방법을 연구하였다. LIPS는 겹친 유류 지문의 화학 성분에 대한 데이터뿐만 아니라 실시간 분석 및 고속 스캐닝이 가능한 분광법이다. 레이저 유도 플라즈마 분광법을 통해 도출된 스펙트럼은 적절한 다변량 분석이 적용되어 법의학적 분류와 겹친 유류 지문의 재구성에 유용한 화학적 성분을 제공한다. 본 연구에서는 LIPS 스펙트럼에서 4가지의 유류 지문을 분류하기 위하여, 주성분 분석 방식과 부분 최소 제곱 회귀 분석을 사용하였다. 제안된 방법은 SIMCA 및 PLS-DA와 같은 구별 방식을 사용하여 4개의 유류 지문의 분류를 성공적으로 입증하였다. 본 연구의 결과는 대략 85% 이상의 정확도를 가졌으며, external validation 실험에서도 분류의 가능성을 보였다. 최종적으로, 125 μm 의 공간 간격으로 레이저 스캐닝 분석을 통한 겹친 유류 지문의 2차원 형태의 분리가 가능함을 입증하였다.

Keywords: 레이저 유도 플라즈마 분광법, 유류 지문, 겹친 지문, 주성분 분석법, 다변량 분석법

OCIS codes: (020.1335) Atom optics; (140.3440) Laser-induced breakdown; (300.6210) Spectroscopy, atomic; (300.6365) Spectroscopy, laser induced breakdown

[†]E-mail: jjyoh@snu.ac.kr, ORCID: 0000-0002-5622-8368

Color versions of one or more of the figures in this paper are available online.

I. 서 론

모든 사람들의 지문은 독특하고 구별되는 모양을 가지고 있다. 이러한 패턴은 일생동안 변화하지 않으며, 동일한 두 개의 지문을 찾는 것은 사실상 불가능하며, 그 확률은 백만분의 일과 같다^[1]. 추가적으로 식습관, 행동 습관, 물리적인 환경, 성장 배경 등의 화학적 정보가 지문에 묻어있는 유류에 존재하게 되며, 이를 추출할 수 있다^[2-6]. 결과적으로 한 사람의 지문은 사람을 인식하는 하나의 소재로서 활용될 수 있으며, 유류 지문을 범죄 현장에서 검출하는 것은 범죄 현장에서 범인을 식별하는 것에 매우 중요한 역할을 하게 된다.

그러나 실제 범죄 현장에서 완벽한 유류 지문의 형태를 도출하는 것은 매우 어려운 일이다. 지문은 손상을 입은 형태로 존재하거나 겹친 형태로 존재하는 것이 다반사다. 그렇기에, 겹친 유류 지문을 명백하게 분리하고 식별할 수 있게끔 하는 것은 추가적인 노력이 필요하다. 겹친 유류 지문을 분리하려던 다양한 선행연구가 존재한다. 선행 연구의 주류는 지문의 벡터 방향을 인식하고, 이를 연결하여 겹친 유류 지문을 각각의 벡터 연결선으로 이루어진 두 개의 유류 지문으로 분리하는 방식이다. 최종적으로는 그러한 알고리즘을 구축하여, 어떠한 겹친 유류 지문이라도 분리할 수 있게끔 하는 방식을 만들었다^[7-10]. 또다른 겹친 유류 지문을 분리하는 방식은 화학적인 방식이 있다. Continuous wave 초록색 빔을 발진하는 레이저를 이용하여 분광 스펙트럼을 검출하여 분리하는 방식^[11]과, MALDI MSI (matrix assisted laser desorption ionization mass spectrometry imaging)를 이용하여 다변량 분석 방식과 융합하여, 지문의 내생적인 그리고 외생적인 성분을 검출함으로써 분리를 수행하였다^[12]. 또한 나노 입자를 유류 지문에 주입하여 분자 분석법을 이용하여 유류 지문을 검출한 사례도 존재한다^[13].

레이저 유도 플라즈마 분광법(laser-induced plasma spectroscopy, LIPS)은 상대적으로 높은 에너지의 레이저를 사용하여 플라즈마를 만들고, 원자 성분을 분석하는 분광 분석법이다. 상대적으로 높은 에너지의 레이저가 분석하고자 하는 샘플에 조사되면, 레이저 삭마 현상이 일어나게 되고 뒤이어 플라즈마가 생성된다. 플라즈마 내부에서 샘플의 원자들은 증발 이후, 원자화되고 이온화되어 있다. 원자와 이온은 들뜬 상태로 존재하다가 일정 시간 이후 다시 바닥 상태로 떨어지면서 원자 성분에 따른 방출광을 가지게 된다. 이 에너지는 원소의 종류와 여기 상태에 따라 고유 파장으로 방출된다. 또한 시료의 모든 원소는 각 원소의 특정 신호가 고체, 액체 및 기체 상태에서 독립적으로 분석되는 것과 동시에 측정이 가능하다. 레이저 유도 플라즈마 분광법은 여타 다른 레이저 분광법과 다르게 광학 셋업이 단순하며, 실시간 분석이 가능하며, 원거리 분석이 가능하고, 특정한 화학적 전처리가 필요하지 않다는 장점이 존재한다. 결론적으로 레이저 유도 플라즈마 분광법은 실제로 연소 계측, 우주 탐사, 그리고 법의학 영역까지 다양한 분야에서 활용되고 있다^[13-16].

유류 지문을 분석하기 위한 다양한 레이저 유도 플라즈마 분광을 이용한 선행연구가 존재한다. 120 fs의 레이저 펄스와 Si wafer를 이용하여 지문을 추출하는 연구가 존재하는데, 이 연구는 레이저 유도 플라즈마 분광법을 이용한 최초의 지문 연구였으며, 가로 5 mm, 세로 1 mm의 2차원 평면 형태로 지문의 모양을 복원하는 연구를 수행하였다^[17]. 다른 연구에서는 이동형 LIPS 장비를 이용하여 유류 지문을 검출하는 시도를 하였다. 5 mJ 정도의 낮은 에너지로 플라즈마를 발생시켜, 이중 펄스 방식으로 삭마시에 필요한 최소 에너지로 지문의 손상도를 줄이고 분석하고자 하였다^[18]. 폭약을 지니는 유류 지문을 분석하고, 2차원 형태로 복원하는 연구도 존재하는데, 본 방식은 OC-LIPS라 하여 optical catapulting 기법이 사용되었으며, TNT, DNT, MNT와 같은 폭약성분을 분류하는 것을 기반으로 연구가 수행되었다^[19,20]. 이러한 다양한 유류 지문 연구에도 불구하고, LIPS는 현재까지의 하나의 지문을 분석하는 것에 머물러 있다.

최근, 데이터 사이언스 이론의 발전과 함께 다변량 분석법은 LIPS 방식과 융합되어 사용되어 왔다. 다변량 chemometric 방식을 사용하여, 토양의 구성 성분을 규명하고 구성 성분때 따른 분류를 수행하였다. 다양한 다변량 분석법이 사용되었으며, partial least-squares regression (PLSR), soft independent modeling of class analogy (SIMCA), partial least squares discriminant analysis (PLS-DA), linear discrimination analysis (LDA), principal component regression (PCR)과 같은 분석 방식들이 토양의 분류를 위해 사용되었다^[21-24]. 다변량 분석법은 LIPS와 통합되어 고고학적 분석 연구에도 사용되었는데, 그림의 작성 연대나 동선의 성분 구성과 제작 시기를 추정하는 연구 등에 활용되었다^[25-27].

본 연구는 겹친 유류 지문을 분리하는 방식으로 데이터 사이언스의 일환인 다변량 분석법을 차용하여, LIPS와 결합하는 시도를 하였다. 본 연구는 각 유류 지문의 화학적 성분을 분석하고, 생활 습관에 따른 화학적 성분의 차이가 있다는 것을 규명하였고, 이를 다변량 분석법을 이용하여 분류하는 연구를 수행하였다. 이러한 분류의 정확도를 검증하기 위하여, 분류는 cross-validation 실험과 external-validation 실험을 각각 수행하였다. 본 연구에서 제안된 방식은 출처가 다른 겹친 두 유류 지문을 화학적으로 구분하여, 최종적으로는 2차원 평면의 형태로 분리할 수 있는 것을 확인하였다.

II. 실험 재료 및 방법

2.1. 유류 지문의 준비 방식

모든 유류 지문은 20대와 30대 아시아인에게서 추출하였으며, 총 4명에게서 추출하였다. 그리고 유류 지문에는 사람마다 가지고 있는 고유한 화학적 특성이 있다는 것을 가정하였다. 또한 이러한 가정을 기반으로 화학적 성분에 따른 분류를 통하여 2차원 화학적 분포도 형태로 출처가 다른 겹친 유류 지문을 분류할 수 있다는 것을 확인하려 하였다. 그러므로 서로 다른 4명의 유류 지문을 동일한 방식으로 추출

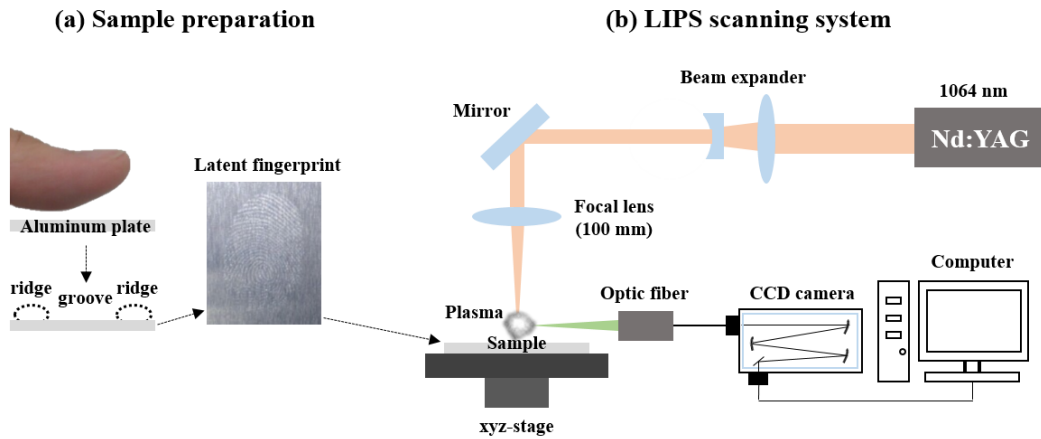


Fig. 1. (a) Sequence of preparation of a fingerprint on an aluminum plate. (b) LIPS experimental setup.

하여, 화학적 성분 차이가 존재하는지 혹은 분류가 가능한지 확인하였다. 최우선적으로 외부 환경에 대한 영향을 최대한 배제하기 위하여, 알코올 솜으로 지문을 닦고 5분 뒤에 지문에서 자연스럽게 나온 유류를 알루미늄 판 위에 묻힌다. 10초 정도 누르는 상태를 유지하면, 그림 1에서 확인할 수 있듯이 알루미늄 판 위에 선명히 지문이 남게 된다. 선명한 지문을 시각화하기 위하여, 형광 빛을 사용하여 지문을 그림 1(a)처럼 확인하였다. 겹친 지문의 경우에는 각자 다른 유류 지문을 추출하여 완성하였으며, 동일한 절차대로 알코올 솜으로 검지 손가락을 세척하고 10분 뒤에 알루미늄 판 위에 지문을 묻혔고, 겹친 지문을 완성하기 위하여, 두 번째 사람은 10초 뒤에 동일한 방식으로 겹친 형태의 지문 모양을 완성하였다. 알루미늄 판의 경우, 지문을 묻히기 전까지는 비닐을 진공포장하여 어떠한 외부 요인도 접촉하지 않도록 설정하였다.

2.2. 광학 장비 구성

모든 플라즈마 분광 실험은 1064 nm의 파장을 가지는 Q-switched Nd:YAG 레이저(RT-250Ec, Applied Spectra Inc.)를 사용하였다. 레이저 펄스의 duration time은 5 ns로 설정하였다. 본 연구에서는 2차원 화학적 분포도를 검출하는 상황에서 해상도를 극대화하기 위하여, 레이저 삭마 현상을 최소화하기를 원하였다. 그렇기에 본 광학 셋업에서 레이저 삭마가 일어날 수 있는 최소의 레이저 에너지인 10 mJ의 레이저 에너지를 사용했다. 레이저가 발진한 이후 플라즈마 방출광을 검출하기 위하여, 게이트 지연시간은 0.5 μ s, 게이트 폭은 1.05 ms로 설정했다. 분광기의 경우, Applied Spectra Inc.에서 자체 제작된 6채널 CCD 카메라가 탑재된 분광기를 사용하였다. 분광기 해상도의 경우, 자외선부터 가시광선 영역은 0.1 nm, 가시광선부터 근적외선 영역은 0.12 nm으로 구성되어 있다. 레이저 스펙트럼은 198 nm부터 1050 nm까지의 파장을 13000 채널로 나누어 해석하고 있다. 플라즈마를 발진하기 위한 레이저 빔의 경로를 만들기 위해서는 코팅이 되지 않은 100 mm 집광거리를 가지는 quartz lens가 사용되

었다. 2차원 화학적 분포도를 검출하는 작업에서, 레이저 삭마 현상을 최소로 줄이기 위한 광학 셋업을 구축했다. 우선적으로 12.7 mm 직경의 입구 조리개로 Nd:YAG 레이저의 넓은 빔에 초점을 맞추고, 8.6 mm 직경의 렌즈를 사용하여 레이저 빔을 직선 레이저 빔으로 소형화 시켰다. 또한 규칙적인 레이저 스펙트럼을 검출하기 위하여, 지문 샘플과 레이저 사이의 거리는 LIPS 장치 내부의 xyz축으로 이동할 수 있는 스테이지를 통해 15 cm으로 일정하게 유지하였다. 실험 설정은 그림 1(b)와 동일하다. 레이저 spot의 크기는 125 μ m으로 고정하였고, 레이저 스캐닝 실험에서도 이동거리를 125 μ m로 설정하였다. 동일하게 레이저 빔의 irradiance는 6.366×10^{13} W/m²로 설정하였다.

2.3. 다변량 분석법

주성분 분석(PCA) 기술은 레이저 스펙트럼과 같은 고차원, 고통량의 데이터를 저차원의 데이터로 변경하는 기술로서, 데이터가 PCA plot 상에 배치가 될 때, 가장 큰 분산을 가지는 축을 각각의 데이터를 분석 및 분할하기 위한 제 1 주요 성분으로서 설정하고, 할당 주요 성분(PC)은 레이저 스펙트럼 데이터에 대한 분산의 크기에 따라 순차적으로 설정된다^[28]. PCA를 통해 PC 점수 영역에 따른 그룹화는 추가적인 SIMCA 및 PLS-DA 분류의 기초가 된다^[28].

Soft independent modeling of class analogy (SIMCA)는 PCA 분석을 통해 가능한 그룹의 로컬 모델을 식별하고 새로운 관측에 대한 가능한 분류를 하는 방식으로 적용된다^[29]. 먼저, PCA 분석이 분류하고자하는 전체 데이터에 대해 가용 데이터 구조에 따라 전역적인 분석이 수행되며 관찰 그룹을 식별한다. 그런 다음, 각 모델에 대해 지역 모델을 평가하고 기존 클래스의 모델 중에 하나로 분류를 하는 방식이다. 각 분류 클래스에 유지되는 주요 구성 요소의 수는 일반적으로 자유롭게 구성할 수 있지만, 구성 요소가 너무 적거나 많게 되면 분류 클래스에 대한 모델에 포함된 신호 또는 정보가 왜곡되고 신호가 줄어들기 되기 때문에 적절한 수의 주요 구성 요소를 설정하는 것이 중요하다^[29].

Partial least squares discriminant analysis (PLS-DA)은 각 유류 지문 레이저 스펙트럼 데이터가 PLS 구성 요소로서 분할되어 관찰 그룹내의 분류를 선명하게 하는 분류 클래스 간의 최대 분리를 유도할 수 있다. PLS-DA 방식은 응답 변수가 통계 단위의 계층을 나타내는 범주형 변수인 PLS 회귀 분석으로 구성된다^[30]. 따라서 PLS-DA는 다른 응답 변수가 개별 그룹을 정의하는 것을 허용하지 않으므로 분류 클래스 할당과 관련하여 모든 측정 변수가 동일한 역할을 수행한다. 실제로, PLS 구성 요소는 합리적인 변수를 설명하고 응답 변수를 예측하는 두 가지 목표 간에 적절한 절충점을 찾으려고 시도한다^[30]. PLS 기반의 분류는 예측 기능을 구성하는데 유리하며, 이 방법은 사전 정의된 분류 클래스의 단가 재할당됨으로써 SIMCA 분류 방식에 비교하여 더 능동적인 분류 방식으로 평가될 수 있다.

본 연구에서는 SIMCA 및 PLS-DA 방법에 대한 cross-validation 실험과 external-validation 실험이 수행된다^[31]. Cross-validation 실험의 경우, 총 4명의 유류 지문 레이저 스펙트럼을 100개씩 추출한다. 각 지문의 100개의 샘플 레이저 스펙트럼을 2개의 그룹으로 나누어, 70개의 훈련 그룹(training data)과 30개의 테스트 그룹(validation data)으로 구분한다. 그 후, 각 유류 지문의 테스트 그룹에서 3개씩 각 유류 지문의 PCA, PLS 데이터를 추출하여 총 12개의 테스트 샘플(validation data)을 선별한다. Calibration 세트에서 하나의 유류 지문 분류 클래스를 제거하는 것과 각 유류 지문의 분류 클래스에 대한 시퀀스 반복의 차이를 제외하고는 cross-validation 실험과 external-validation 실험을 유사한 순서로 수행하였다. 분석의 정확성을 높이기 위하여, 전체 데이터를 cross-validation, external-validation 실험을 위한 테스트 데이터로 재분할하여 100개의 독립적인 반복을 수행하였다. 현재 연구에서는 각 유류 지문의 PCA, PLS 데이터가 올바르게 할당될 경우를 "Correct-classification", 잘못 할당된 경우에는 "Mis-classification", 할당이 되지 않는 경우에는 "Un-classification"로 명명하였다. 본 연구에서 다변량 분석용으로 사용된 software는 Unscrambler X 10.1과 Matlab 2016a가 사용되었다. 추가적으로 그래픽 그림은 Origin Pro 8.5.1을 사용하여 그렸다.

III. 실험 결과

3.1. 유류 지문들에 대한 레이저 분광 분석과 주성분 분석 결과

LIPS 스펙트럼은 4명의 다른 사람의 유류 지문으로부터 추출되었다. 레이저 스펙트럼의 신호의 신뢰성을 향상시키기 위해, 그림 2에서 사용된 유류 지문 레이저 스펙트럼은 100개의 지문 신호를 평균한 값이다. 지문의 10개의 지점에서 10개씩 데이터를 추출하여, 총 100개의 유류 지문 레이저 스펙트럼을 평균하였다. 그림 2에서 가시화가 용이하게 하기 위하여, 레이저 스펙트럼마다 y축 값을 35,000만큼 offset 하였다. 유류 지문 레이저 스펙트럼은 몇 개의 특정 파장에서

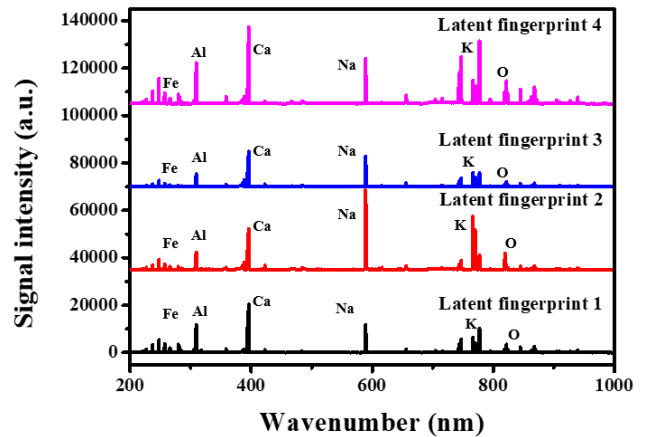


Fig. 2. Plasma emission spectra lines of four types of latent fingerprints.

상대적으로 높고 낮은 신호 세기의 차이를 가지는 것을 확인하였다. 본 그림 2에서 추출한 데이터를 기반으로 후후 다변량 분석법을 수행하였다.

유류 지문 레이저 스펙트럼에서 비교적 높은 신호 세기를 가지는 파장은 대략 238, 239, 274, 308, 309, 393, 394, 396, 588, 589, 747, 766 nm로 이는 Fe, Al, O, Na, K, Ca 등의 신호를 의미한다. 이 신호의 조성은 NIST 원자 스펙트럼 데이터 베이스를 이용하여 분석하였다^[32].

식습관, 주변 환경 및 유전된 가족력과 같은 다양한 요인에 의해, 유류 지문의 화학적 성분에는 차이가 존재한다. 이는 인간의 피부 표면에서 나타나는 유류에서 확인할 수 있다. 따라서 유류 지문의 화학 성분을 LIPS를 사용하여 측정할 때, 화학 성분의 차이에 따른 특정 파장에서 신호 세기의 차이가 존재한다. 그러나 이러한 결과는 육안으로 분명히 구별되기는 어려우며, 유류 지문을 신뢰할 수 있는 데이터로 분류하려면 주성분 분석 혹은 부분 최소 자승법과 같은 기초적인 분류 클래스 생성 방식이 필요하다. 이러한 분석법들은 전체적인 LIPS 스펙트럼 영역을 분석하고, 주요 성분을 분석할 수 있는 비교적 높은 신호 대 잡음비를 가지는 주 방출 파장대를 추출하고, 추출된 주요 성분 스펙트럼에 기초하여 특정값으로서 2차원 형태의 점으로 변화시킨다. 본 연구에서는 주성분 분석과 부분 최소 자승법은 각 유류 지문 레이저 스펙트럼 100개에 대해 수행되었고, 그림 3은 4개의 유류 지문에 대한 PCA, PLS plot 상의 결과를 보여준다. 각각의 유류 지문의 위치는 2차원 평면에서 독립적 위치를 나타내는 것을 확인할 수 있으며, 그림 3(b), 3(d)에서 확인할 수 있듯이 PC1과 PC2를 사용하였을 때 전체 스펙트럼 범위에서 90% 이상(PC 1: 53.80%, PC 2: 35.90%)을 분류할 수 있음을 PLS 분석에서는 전체 스펙트럼 범위에서 95% 가량(Factor 1: 86%, Factor 2: 7%)을 분류할 수 있음을 확인하였다. PLS 모델은 path model이라 불리는 행렬들로 class를 지정하고 모델링을 수행하는 방식이기에, 그림 3(d)는 그림 3(c) 결과의 path model에 대한 결과로 설명될 수 있다.

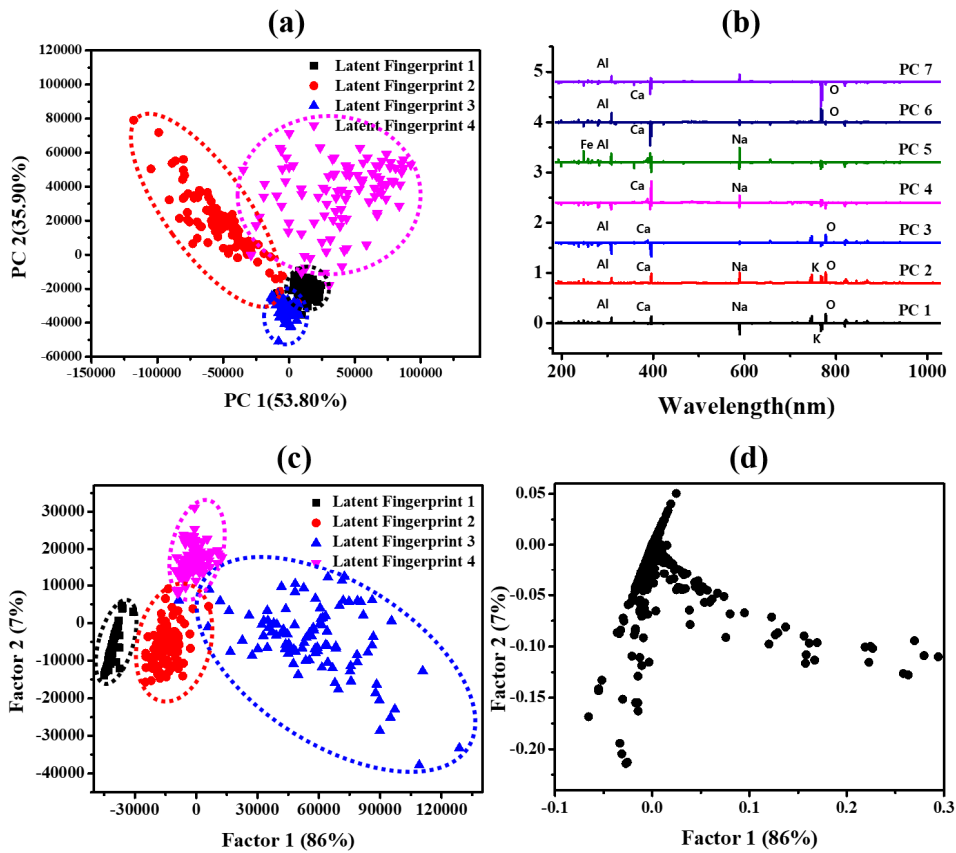


Fig. 3. (a) Results of PCA plot. (b) Results of extracted PCs. (c) Results of PLS method. (d) PLS loading plot.

그림 3(b)의 그래프에서 PCA의 플라즈마 방출광의 분류의 주요 인자들은 Mg, Fe, Na, Ca, K와 같음을 확인할 수 있으며, PC 1~7의 결과는 거의 동일한 파장을 보여주지만 일부 파장의 신호 세기는 각 PC에 의해 반전되어 차이를 유발함을 알 수 있다.

3.2. 다변량 분석 결과

SIMCA 분석을 수행하기 위하여 각 4가지, 100개씩, 총 400개의 유류 지문 레이저 스펙트럼이 사용되었다. Cross-validation 실험의 경우, 하나의 유류 지문마다 100개의 유류 지문 레이저 스펙트럼으로부터 무작위로 선택된 30개의 지문을 테스트 샘플(validation data)로 설정하고, 나머지 70개의 유류 지문 레이저 스펙트럼은 PCA를 통하여 분류(training data)를 수행하였다. SIMCA 분석이 지문 당 30개의 테스트 샘플(validation data)에 대해 수행될 때 결과는 표 1로 설명할 수 있다. External validation 실험의 경우, 하나의 유류 지문 분류 클래스를 calibration set에서 제외하여 각 유류 지문 클래스를 분류하는 것으로 반복하였다. PCA에 의해 분석된 데이터에서는 cross-validation 실험을 하였던 동일한 PCA 데이터를 사용하였으며, PCA 모델간의 거리를 그림 4에서 확인할 수 있다. 이 값은 두 모델이 서로 다른 방식을 의미하며, 모델과 모델 사이의 거리가 클수록 두 분류 모델을 보다 명확하게 식별할 수 있다. SIMCA는 PC1, 2의 두 가지 주성

분을 사용하여 분류하였으며 분류의 정확도가 대략 90% 가량 되기 때문에 분류 시간을 최소화로 설정하고 분류의 정확도를 확보하는 방법으로 2개의 주성분 요소를 사용하였다.

통계적 신뢰성을 도출하기 위하여 SIMCA 분석 결과는 전체 데이터에서 도출하였을 뿐 아니라, 레이저 지문은 각각의 유류 지문의 모든 데이터에서 무작위로 추출하여 100번의 iteration을 수행하였다. Cross-validation와 external-validation 실험의 차이는 테스트 샘플이 처음 100개의 지문 데이터에서 추출되는지 아니면 미리 추출된 70개의 데이터에서 추출되는지 차이이다. 표 1에서 확인할 수 있듯이, cross-validation 실험의 경우, 분류 정확도는 90.36%, 6.18%는 부적절하고, 3.46%는 분류가 되지 않았음을 보여준다. External-validation 실험에서는 84.10%, 13.01%, 2.89%의 시험 평균을 가지는 것을 확인하여, external-validation 실험에서도 80% 이상의 정확도로 유류 지문의 분류가 가능한 것을 확인하였다.

PLS-DA 분석을 위해, SIMCA 분석에서 사용된 동일한 4개 유류 지문 레이저 스펙트럼을 사용하였고, 유류 지문을 분류하기 위해 각 지문에 대해 PLS 회귀분석을 수행하였다. PLS-DA 분석은 SIMCA 분석에 비해 기본적으로 안정적이고 강력하지만 실제 결과는 SIMCA 분석과 크게 다르지 않다. PLS-DA 분석에서도 2가지의 factor를 사용하였고, 역시 분류 정확도가 90% 이상이면서 시간을 최소화할 수 있는 factor의 수를 설정하였다. Cross-validation 실험의 경우 평균

Table 1. Results of SIMCA

(a) Cross-validation				(b) External-validation			
Type	Correct-classification	Mis-classification	Un-classification	Type	Correct-classification	Mis-classification	Un-classification
Fingerprint 1	0.9010	0.0320	0.0670	Fingerprint 1	0.8440	0.1140	0.0420
Fingerprint 2	0.9327	0.0333	0.0300	Fingerprint 2	0.8427	0.1313	0.0260
Fingerprint 3	0.8750	0.1000	0.0250	Fingerprint 3	0.8227	0.1433	0.0340
Fingerprint 4	0.9060	0.0820	0.0120	Fingerprint 4	0.8540	0.1320	0.0140

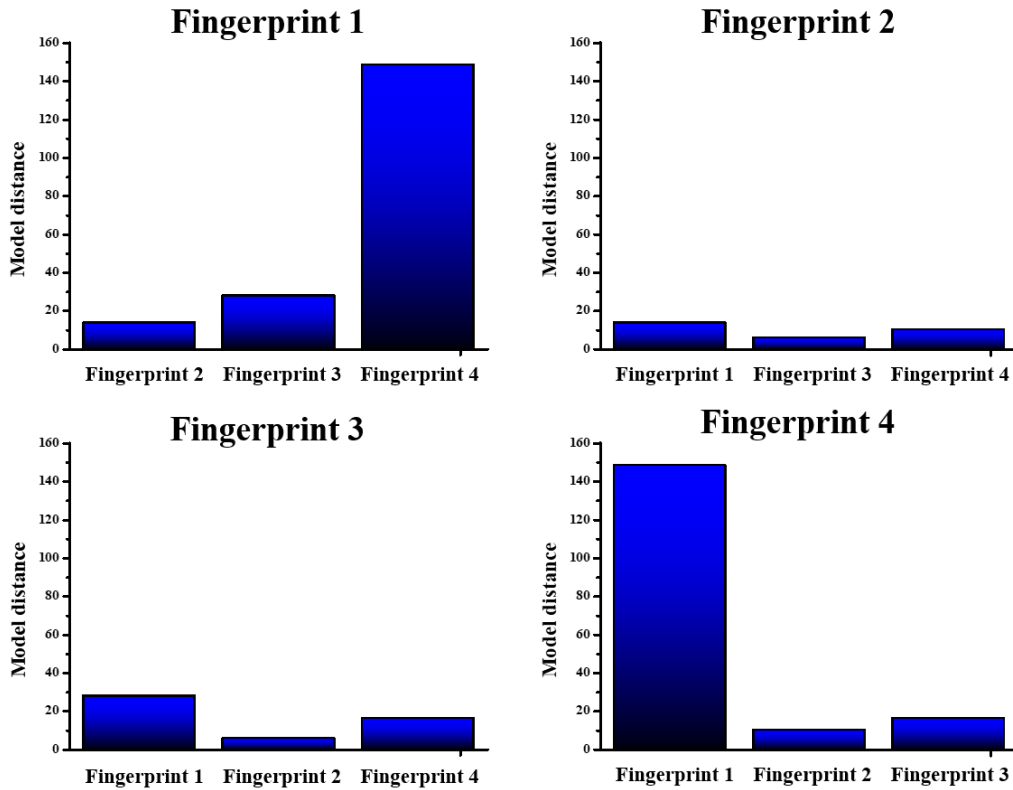


Fig. 4. Model distance obtained from the SIMCA method.

Table 2. Results of PLS-DA

(a) Cross-validation test				(b) External-validation test			
Type	Correct-classification	Mis-classification	Un-classification	Type	Correct-classification	Mis-classification	Un-classification
Fingerprint 1	1.0000	0.000	0.0000	Fingerprint 1	0.9333	0.0333	0.0333
Fingerprint 2	0.8667	0.1230	0.0000	Fingerprint 2	0.7980	0.1667	0.0353
Fingerprint 3	0.9667	0.0210	0.1200	Fingerprint 3	0.8667	0.0987	0.0346
Fingerprint 4	0.9000	0.0920	0.0080	Fingerprint 4	0.8333	0.0667	0.1000

분류 정확도는 93% 이상으로, SIMCA 실험 분석에 비교하였을 때 비교적 정확한 분류가 일어나는 것을 확인하였다. External-validation 실험에서는 85% 정도의 정확도로 SIMCA 보다 약간 향상된 분류 정확도를 확인했다. 그러나 표 2에서 확인할 수 있듯이 미분류의 비율은 PLS-DA 분류 방식이 SIMCA 분류 방식보다 정확도가 높은 것을 확인하였다.

3.3. 겹친 유류 지문 분리

유류 지문의 출처가 다른 경우, PCA, SIMCA 및 PLS-DA와 같은 다변량 분석법을 통해 분류가 수행됨을 확인하였고, 레이저 스펙트럼 경향을 비교하여 2차원 화학적 분포도 형태로 겹친 지문을 분리한 결과를 검출하였다. 다변량 분석 방법을 적용하여 각 지문의 선을 재구성하는 방법은 그림 5와 같다. 우선적으로 각 지문 당 30개의 레이저 스펙트럼 추출을 하고, 그런 다음 총 60개의 지문 스펙트럼에 대하여

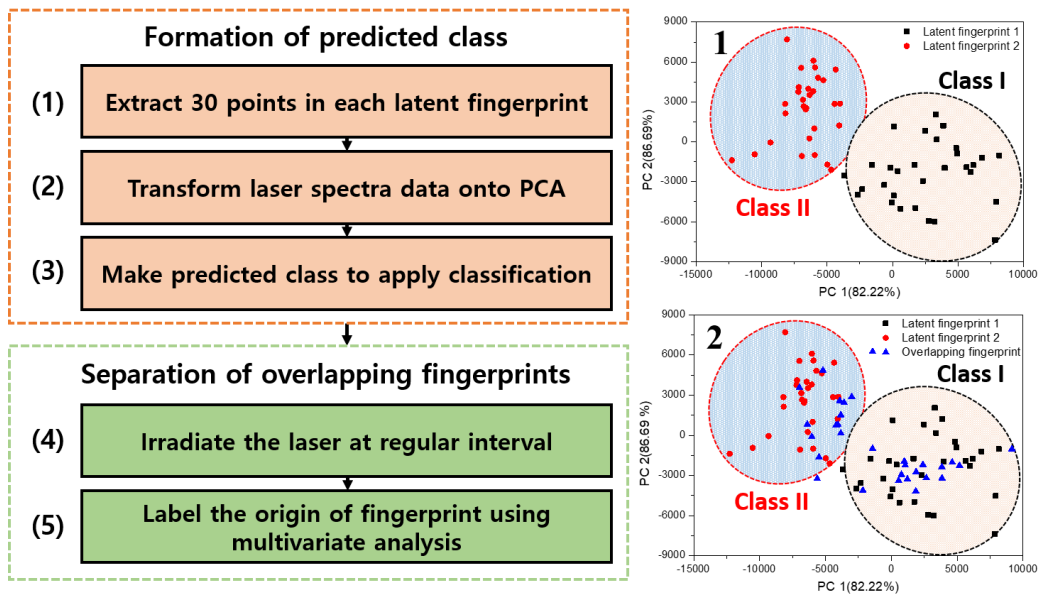


Fig. 5. A process for the separation of overlapping fingerprints.

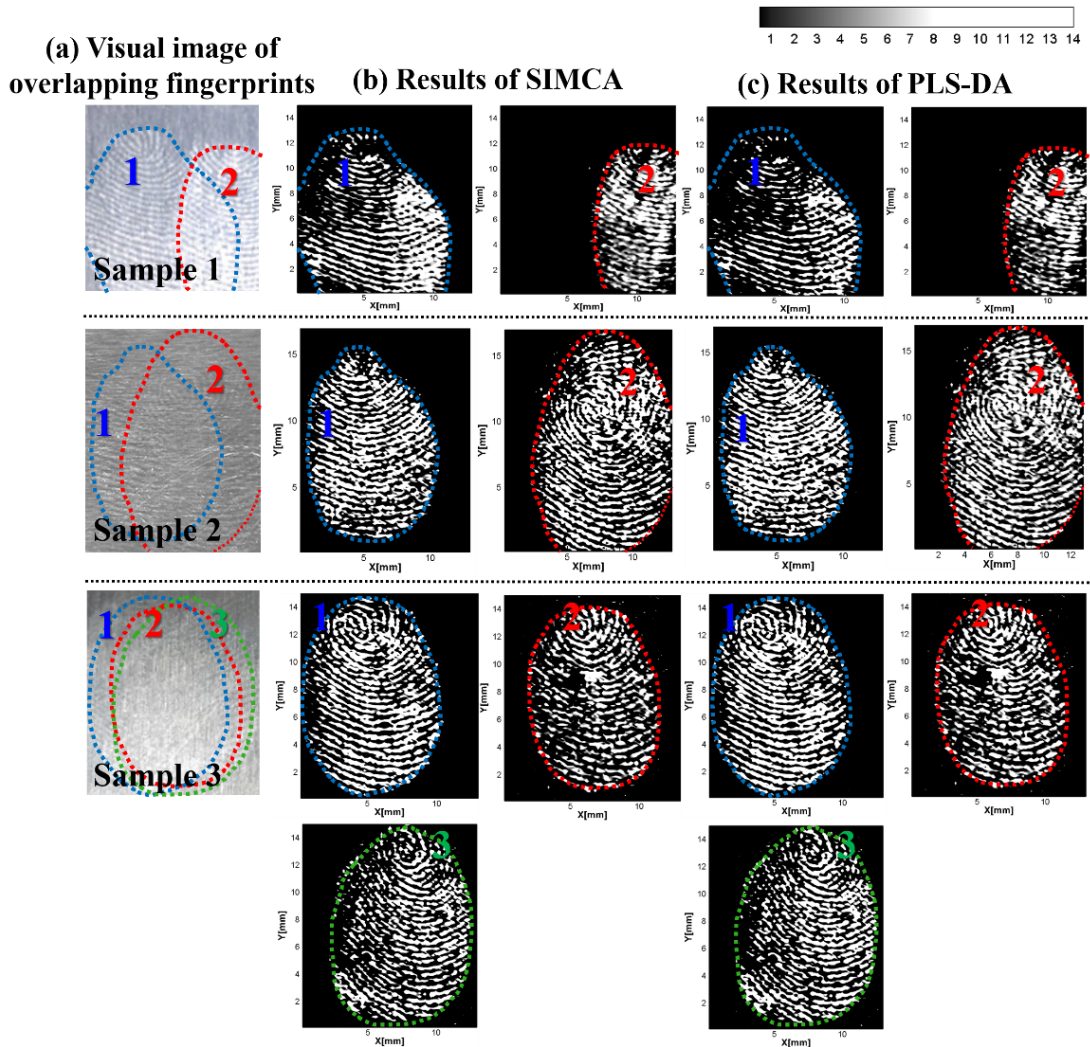


Fig. 6. Separation of overlapping latent fingerprints using multivariate analysis.

PCA와 PLS 분석을 통하여 분류 클래스가 설정된다. 그 후, 레이저는 일정한 간격으로 조사되고 모든 레이저 spot 사이의 간격은 125 μm 로 설정하였다. 각각의 점의 레이저 스펙트럼은 분류되어 분석되고 최종적으로는 그림 6과 같은 겹친 유류 지문을 분리할 수 있다. 이때 기반이 되는 것이 이전 3.2 부분에서 수행한 external-validation 실험을 통하여 검출한 방식이다. 이전에는 4개의 지문에 대한 결과이지만, 2개의 겹친 지문을, 그리고 3개의 겹친 지문을 분류하는 상황에서는 상대적으로 높은 결과의 분류 정확도를 보였을 것으로 추정된다. 그림 6에서 확인할 수 있듯이, sample 1의 결과는 지문의 형태가 완벽하게 복원되지 않는 경우도 존재하지만, sample 2, 3과 같이 지문의 모양이 거의 완벽하게 복원이 되는 경우도 존재한다. 추론하기로는 두 겹친 지문의 화학적 성분의 차이가 명확할수록 지문의 복원이 상대적으로 용이한 것으로 생각된다. 특히나 sample 3의 결과에서는 각 유류 지문의 나선형 꼬리와 같은 부분도 거의 100% 일치하는 형태로 복원이 되었음을 확인할 수 있다. 결과적으로, 실제 겹친 지문의 경우에도, 다변량 분석을 사용하여 각 유류 지문을 분리하고 2차원 화학적 분포도 형태로 분리하는 것이 가능한 것을 입증하였다. 그러나 겹친 영역에 대해서는 복원이 쉽지 않거나 완벽하지 않은 것을 확인할 수 있으며, 이는 추가적으로 겹친 부분에 대한 분류 기준을 만드는 방법으로 추후 연구를 진행해야 할 것이다.

IV. 결 론

본 연구에서는 겹친 유류 지문을 분리하기 위하여, 데이터 사이언스의 일종인 다변량 분석과 LIPS를 결합하는 접근을 검증하였다. LIPS 데이터는 레이저 에너지, 펄스 지속 시간 및 자외선에서 가시광선까지의 범위에서 수집된 매우 신뢰할 수 있는 스펙트럼 검출에 최적화된 빔의 직경을 가진 Q-switched Nd:YAG 레이저(1064 nm)를 사용하여 도출하였다. PCA 및 PLS 방식을 이용하여, 두 개의 겹친 유류 지문 각각에 데이터를 2차원 평면에서 독립적인 영역으로 추출하였고, cross-validation 실험을 통하여서는 90% 이상의 분류 정확도를 external-validation 실험을 통해서는 대략 85% 정도의 정확도를 가지는 분류를 수행하였다. 마지막으로, 겹친 지문의 분리는 LIPS scanning 기술과 다변량 분석 방식이 융합되어 수행되었다. 적용의 용이성과 결과의 정확성에 기초하여, 본 연구에서 제안하는 방법은 실제 범죄 현장에서 겹친 유류 지문을 분리하는데 유용하게 사용될 수 있을 것이다.

감사의 글

본 연구는 서울대학교 차세대 우주추진 연구센터와 연계된 미래창조과학부의 재원으로 한국연구재단의 지원을 받아 수행한 선도연구센터지원사업(NRF-2013R1A5A1073861)의 연구결과입니다.

References

1. E. E. Hueske, *Firearms and Fingerprints* (Essentials of Forensic Science Series 5), Suzanne Bell, ed. (Facts and File, NY, USA, 2008).
2. N. E. Archer, Y. Charles, J. A. Elliott, and S. Jickells, "Changes in the lipid composition of latent fingerprint residue with time after deposition on a surface," *Forensic Sci. Int.* **154**, 224-239 (2005).
3. R. S. Croxton, M. G. Baron, D. Butler, T. Kent, and V. G. Sears, "Variation in amino acid and lipid composition of latent fingerprints," *Forensic Sci. Int.* **199**, 93-102 (2010).
4. J. S. Day, H. G. M. Edwards, S. A. Dobrowski, and A. M. Voice, "The detection of drugs of abuse in fingerprints using Raman spectroscopy I: latent fingerprints," *Spectrochim. Acta, Part A* **60**, 563-568 (2004).
5. K. M. Antoine, S. Mortazavi, A. D. Miller, and L. M. Miller, "Chemical differences are observed in children's versus adults' latent fingerprints as a function of time," *J. Forensic Sci.* **55**, 513-518 (2010).
6. C. Weyermann, C. Roux, and C. Champod, "Initial results on the composition of fingerprints and its evolution as a function of time by GC/MS analysis," *J. Forensic Sci.* **56**, 102-108 (2011).
7. Q. Zhao and A. K. Jain, "Model based separation of overlapping latent fingerprints," *IEEE Trans. Inf. Forensics Security* **7**, 904-918 (2012).
8. M. Schott, R. Merkel, and J. Dittmann, "Sequence detection of overlapping latent fingerprints using a short-term aging feature," in *Proc. IEEE International Workshop on Information Forensics and Security (WIFS)* (Tenerife, Spain, Dec. 2012), pp. 85-90.
9. J. Feng, Y. Shi, and J. Zhou, "Robust and efficient algorithms for separating latent overlapped fingerprints," *IEEE Trans. Inf. Forensics Security* **7**, 1498-1510 (2012).
10. A. Nakamura, H. Okuda, T. Nagaoka, N. Akiba, K. Kurosawa, K. Kuroki, F. Ichikawa, A. Torao, and T. Sota, "Portable hyperspectral imager with continuous wave green laser for identification and detection of untreated latent fingerprints on walls," *Forensic Sci. Int.* **254**, 100-105 (2015).
11. R. Bradshaw, W. Rao, R. Wolstenholme, M. R. Clench, S. Bleay, and S. Francese, "Separation of overlapping fingermarks by matrix assisted laser desorption ionisation mass spectrometry imaging," *Forensic Sci. Int.* **222**, 318-326 (2012).
12. H.-W. Tang, W. Lu, C.-M. Che, and K.-M. Ng, "Gold nanoparticles and imaging mass spectrometry: Double imaging of latent fingerprints," *Anal. Chem.* **82**, 1589-1593 (2010).
13. S. H. Lee, H. Do, and J. J. Yoh, "Simultaneous optical ignition and spectroscopy of a two-phase spray flame," *Combust. Flame* **165**, 334-345 (2016).
14. J.-J. Choi, S.-J. Choi, and J. J. Yoh, "Standoff detection of geological samples of metal, rock, and soil at low pressures using laser-induced breakdown spectroscopy," *Appl. Spectrosc.*

- 70, 1411-1419 (2016).
15. K.-J. Lee, S.-J. Choi, and J. J. Yoh, "Stand-off laser-induced breakdown spectroscopy of aluminum and geochemical reference materials at pressure below 1 torr," *Spectrochim. Acta, Part B* **101**, 335-341 (2014).
 16. J.-H. Yang, S.-J. Choi, and J. J. Yoh, "Towards reconstruction of overlapping fingerprints using plasma spectroscopy," *Spectrochim. Acta, Part B* **134**, 25-32 (2017).
 17. M. T. Taschuk, Y. Y. Tsui, and R. Fedosejevs, "Detection and mapping of latent fingerprints by laser-induced breakdown spectroscopy," *Appl. Spectrosc.* **60**, 1322-1327 (2006).
 18. Y. Godwal, M. T. Taschuk, S. L. Lui, Y. Y. Tsui, and R. Fedosejevs, "Development of laser-induced breakdown spectroscopy for microanalysis applications," *Laser Part. Beams* **26**, 95-104 (2008).
 19. M. Abdelhamid, F. J. Fortes, M. A. Harith, and J. J. Laserna, "Analysis of explosive residues in human fingerprints using optical catapulting-laser-induced breakdown spectroscopy," *J. Anal. At. Spectrom.* **26**, 1445-1450 (2011).
 20. M. Abdelhamid, F. J. Fortes, J. J. Laserna, and M. A. Harith, "Optical catapulting laser induced breakdown spectroscopy (OC-LIBS) and conventional LIBS: a comparative study," *AIP Conf. Proc.* **1380**, 55-59 (2011).
 21. J. Feng, Z. Wang, L. West, Z. Li, and W. Ni, "A PLS model based on dominant factor for coal analysis using laser-induced breakdown spectroscopy," *Anal. Bioanal. Chem.* **400**, 3261-3271 (2011).
 22. A. D. S. Augusto, É. F. Batista, and E. R. P. Filho, "Direct chemical inspection of eye shadow and lipstick solid samples using laser-induced breakdown spectroscopy (LIBS) and chemometrics: proposition of classification models," *Anal. Methods* **8**, 5851-5860 (2016).
 23. M. J. C. Pontes, J. Cortez, R. K. H. Galvão, C. Pasquini, M. C. U. Araújo, R. M. Coelho, M. K. Chiba, M. F. de Abreu, and B. E. Madari, "Classification of Brazilian soils by using LIBS and variable selection in the wavelet domain," *Anal. Chim. Acta* **642**, 12-18 (2009).
 24. T. Zhang, L. Liang, K. Wang, H. Tang, X. Yang, Y. Duan, and H. Li, "A novel approach for the quantitative analysis of multiple elements in steel based on laser-induced breakdown spectroscopy (LIBS) and random forest regression (RFR)," *J. Anal. At. Spectrom.* **29**, 2323-2329 (2014).
 25. S. Awasthi, R. Kumar, G. K. Rai, and A. K. Rai, "Study of archaeological coins of different dynasties using libs coupled with multivariate analysis," *Opt. Lasers Eng.* **79**, 29-38 (2016).
 26. S. Moncayo, M. Kociánová, J. Hulík, J. Plavčan, M. Horňáčková, M. Suchoňová, P. Veis, and J. O. Cáceres, "Discrimination of copper alloys with archaeological interest using LIBS and chemometric methods," in *Proc. The 23rd Annual Conference of Doctoral Students - WDS 2014* (Prague, Česká republika, Jun. 2014), pp. 131-135.
 27. B. G. Oztoprak, M. A. Simmaz, and F. Tülek, "Composition analysis of medieval ceramics by laser-induced breakdown spectroscopy (LIBS)," *Appl. Phys. A* **122**, 557 (2016).
 28. S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemom. Intell. Lab. Syst.* **2**, 37-52 (1987).
 29. G. R. Flaten, B. Grung, and O. M. Kvalheim, "A method for validation of reference sets in SIMCA modelling," *Chemom. Intell. Lab. Syst.* **72**, 101-109 (2004).
 30. R. G. Brereton and G. R. Lloyd, "Partial least squares discriminant analysis: taking the magic away," *J. Chemom.* **28**, 213-225 (2014).
 31. N. C. Dingari, I. Barman, A. K. Myakalwar, S. P. Tewari, and M. K. Gundawar, "Incorporation of support vector machines in the LIBS toolbox for sensitive and robust classification amidst unexpected sample and system variability," *Anal. Chem.* **84**, 2686-2694 (2012).
 32. P. L. Smith, C. Heise, J. R. Esmond, and R. L. Kurucz, *Atomic spectral line database* (Harvard-Smithsonian Center for Astrophysics, 2016), <http://cfa-www.harvard.edu/amp/data/kur23/sekur.html> (Accessed: March 2020).