

선형 강도 교정을 이용한 라만 스펙트럼 인식

박준규, 백성준*, 박아론
전남대학교 전자컴퓨터공학부

The identification of Raman spectra by using linear intensity calibration

Jun-Kyu Park, Sung-June Baek*, Aaron Park

School of Electronics and Computer Engineering, Chonnam National University

요약 라만 스펙트럼은 측정 장비 및 환경 조건에 따라 동일한 물질이라도 스펙트럼의 강도 차이를 보인다. 이는 라만 분광의 패턴 인식적인 접근에 제약을 주기 때문에 장비간의 호환성 및 라만 데이터베이스의 재사용을 위해 반드시 해결해야 하는 문제다. 이를 위해 이전의 주요 연구들에서는 측정 장비 간에 전달 함수를 가정하고 이를 구한 후 직접적인 스펙트럼의 교정을 수행하였다. 하지만 이 방식은 강도 왜곡을 발생시키는 다른 조건들에 대해서는 대처 할 수 없는 방법이다. 따라서 본 논문에서는 다양한 측정 조건에 보다 유연하게 대응 할 수 있는 선형 강도 교정을 이용한 분류 방법을 제안하였다. 제안한 방법의 성능 평가를 위해 실험에서는 14033종의 화학 물질에서 측정된 라만 라이브러리를 실험물질에 대한 판별 지표로 사용하였으며, 3개의 라만 분광기로부터 측정된 10종의 화학 물질 라만 스펙트럼을 실험 데이터로 사용하였다. 실험결과에 따르면 제안한 방법을 사용하였을 때 강도 왜곡된 스펙트럼에 대해 100%의 판별 성능을 보였으며, 판별된 스펙트럼에 대해서도 이전보다 높은 상관점수를 보여 사용자가 화학 물질을 판별하는 데 유용한 도구로 사용될 수 있음을 확인하였다.

Abstract Raman spectra exhibit differences in intensity depending on the measuring equipment and environmental conditions even for the same material. This restricts the pattern recognition approach of Raman spectroscopy and is an issue that must be solved for the sake of its practical application, so as to enable the reusability of the Raman database and interoperability between Raman devices. To this end, previous studies assumed the existence of a transfer function between the measurement devices to obtain a direct spectral correction. However, this method cannot cope with other conditions that cause various intensity distortions. Therefore, we propose a classification method using linear intensity calibration which can deal with various measurement conditions more flexibly. In order to evaluate the performance of the proposed method, a Raman library containing 14033 chemical substances was used for identification. Ten kinds of chemical Raman spectra measured using three different Raman spectrometers were used as the experimental data. The experimental results show that the proposed method achieves 100% discrimination performance against the intensity-distorted spectra and shows a high correlation score for the identified material, thus making it a useful tool for the identification of chemical substances.

Keywords : Chemical identification, Linear intensity calibration, Machine learning, Pattern recognition, Raman spectroscopy

1. 서론

최근 다양한 분야에서 라만 스펙트럼을 이용한 인식

시스템이 사용됐으며, 각 필드의 요구에 따라 그 판별 능력이 점차 증가되고 있다[1-3]. 하지만 라만 스펙트럼은 측정 조건에 따라 동일물질에 대한 스펙트럼일지라도 피

이 논문은 2017년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (no.2017R1A2B4012450).

*Corresponding Author : Sung-June Baek(Chonnam National Univ.)

Tel: +82-62-530-1795 email: tozero@chonnam.ac.kr

Received December 6, 2017

Revised January 15, 2018

Accepted March 9, 2018

Published March 31, 2018

크 강도에서 차이가 발생한다[4]. 이러한 특성은 라만 스펙트럼을 이용한 패턴인식적인 접근에 제한을 주며, 오랜 시간동안 축적해온 라만 데이터베이스의 재사용성, 효율성, 라이브러리 간의 합병 및 장비 간의 호환성을 줄이는 주요 원인이 되어왔다[5].

이를 해결하기 위해 다양한 연구가 행해졌는데 대표적인 것으로는 DS(direct standardization)[6], PDS(piecewise direct standardization)[7], SWD(standardization in the wavelet domain)[8]와 같은 교정 방법과 NIST에서 제안한 SRM을 이용한 교정 방법[9] 등이 있다. 이 방법들은 라만 장비 간에 스펙트럼을 왜곡시키는 전달 함수가 있다는 것을 가정하고, 이를 이용하여 스펙트럼의 강도 차이를 보정한다. 따라서 한번 보정함수를 얻게 되면, 상관계수와 같은 단순한 분류알고리즘을 이용하여 라만 스펙트럼을 판별할 수 있는 장점이 있고 훌륭한 판별 결과를 보여준다. 하지만 이 방식들은 사용해야하는 모든 장비를 컨트롤해야하는 번거로움이 있고, 장비 간의 차이 이외에 측정 온도 등을 포함한 여타 조건의 변화에는 대응하기 어렵다는 단점을 지니고 있다[10].

따라서 본 연구에서는 측정 조건에 따라 다양한 방식으로 강도 왜곡된 스펙트럼에 즉각적인 대처가 가능한 선형 강도 교정 방법을 제안하고 이를 라만 스펙트럼 판별에 적용하였다.

2. 연구방법

2.1 화학 물질 라만 스펙트럼

실험에서는 총 10종의 화학물질인 acetone, acetonitrile, benzene, cyclohexane, ethyl alcohol, ethylene glycol, hexane, nitromethane, 2-nitrotoluene, toluene을 사용하였는데 시료는 Sigma-Aldrich(St Louis, MO, USA)사를 통해 구입하였다. 라만 스펙트럼 측정에 사용된 분광기는

Renishaw 2000 Raman microscope system(Renishaw, New Mills, UK)과 inVia Renishaw Raman microscope system(Renishaw, New Mills, UK)이다. Renishaw 2000의 경우 argon ion laser($\lambda = 514.5nm$)을 사용하였고, inVia 라만 시스템은 He-Ne laser($\lambda = 632.8nm$) 및 diode laser($\lambda = 785.0nm$)를 여기원으로 사용하였다. 세 기기 모두 검출 가능한 스펙트럼의 측정 범위는 $201 - 3500 cm^{-1}$ 이고 레이저 출력은 약 1mW이다. 10종의 화학 물질은 각 레이저 소스에 의해 3회씩 측정하여 총 90개의 실험데이터를 구성하였다. 실험데이터의 대조군으로는 사전 준비된 14033개의 라만 라이브러리가 사용되었으며, 라이브러리는 FT-Raman spectrometer를 이용해 측정된 라만 스펙트럼으로 구성되었다. 실험에 사용된 4가지 장비의 하드웨어 스펙을 표 1.에 나타내었으며, 각 장비에서 측정된 Benzene 스펙트럼을 그림 1.에 보였다. 그림에서는 보인 네 스펙트럼은 동일 물질이지만 측정 장비조건에 따라 서로 간에 강도차이가 확인한 것을 확인 할 수 있다.

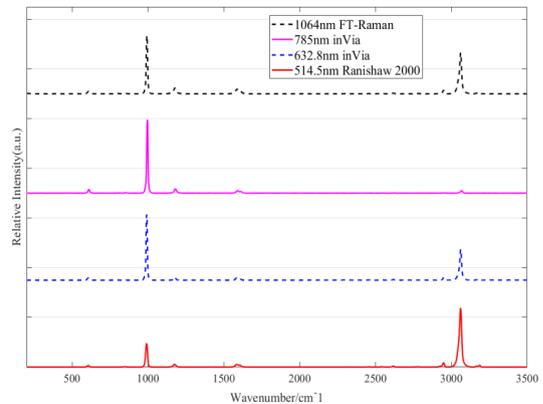


Fig. 1. Benzene spectra measured on the four instrument sets

Table 1. Properties of the Raman spectrometers used in this study

Spectrometer	Laser power(mW)	Excitation wavelength(nm)	Spectral range	System resolution(cm^{-1})
FT-Raman spectrometer	400-600	1064	201-3500	1.93
Renishaw 2000	1.0	514.5	201-3500	4
inVia	1.0	632.8	201-3500	1
inVia	1.0	785.0	201-3500	1

2.2 전처리

측정된 라만 스펙트럼은 배경 잡음 및 가산 잡음을 가지고 있기 때문에 이를 효과적으로 처리하여야 한다. 전처리는 배경 잡음 제거, 가산 잡음 제거 및 정규화로 구성된다. 먼저 배경 잡음제거를 위해 arPLS(asymmetrically reweighted penalized least squares smoothing) 방법을 사용하였다[11]. 이 방법은 최소자승법을 이용한 평활화 방법으로 부분 균형을 이루는 비대칭 가중치를 사용하여 배경잡음이 제거된 스펙트럼을 얻는다. 실험에서는 평활도 파라미터를 $\lambda = 10^5$ 로 설정하였다. 다음으로 가산잡음을 제거하기 위해서는 저역 통과 필터보다 피크를 잘 보존하는 것으로 알려진 무빙 윈도우 필터를 사용하였다[12]. 마지막으로 잡음을 제거한 스펙트럼은 측정된 스펙트럼들 사이의 편차를 최소화하고, 각 상관도를 동일하게 비교하기 위해 각 스펙트럼의 놈(norm)이 1이 되도록 정규화 하였다.

2.3 선형 강도 교정방법

2.3.1 보정 데이터 선택

제안한 교정 방법은 먼저 스펙트럼의 주요 피크를 선택하는 단계로 시작된다. 라만 스펙트럼은 물질에 따라 고유의 스펙트럼 영역에서 독특한 피크가 존재하기 때문에 피크 위치, 모양, 크기 등은 물질을 구분 할 수 있는 중요한 정보이다. 따라서 대조 되는 두 스펙트럼에서 피크의 위치와 상대적인 크기 정보를 획득한 후 이들 피크 점을 중심으로 선형 교정 함수를 구한 후 이를 이용하여 스펙트럼을 최대한 매칭 시킨다. 비록 피크 강도의 차이는 다르지만, 당연히게도 동일 물질의 스펙트럼일 경우 동일한 위치에서 같은 피크가 존재 할 것을 예상 할 수 있으며, 서로 다른 물질의 경우 피크 점이 서로 어긋날 것을 예상할 수 있다. 스펙트럼에서 피크를 검출하기 위해 본 논문에서는 Matlab의 Signal processing toolbox 를 이용하여 스펙트럼에 존재하는 피크들의 위치 정보를 얻었다. 피크의 선택은 피크 높이가 큰 순서로 선택을 하였으며, 선택 개수에 제한을 두기 위해 스펙트럼에서 가장 큰 피크의 크기를 기준으로 일정 비율 이하의 아주 작은 피크들은 선택 후보에서 제외시켰다.

2.3.2 선형 교정

기준 스펙트럼 $f(x)$, 그리고 이것과 비교할 스펙트럼을 $g(x)$ 라 할 때 두 스펙트럼 사이의 선형 교정 함수는

$f(x) \approx w(x)g(x)$ 와 같이 되도록 하는 함수 $w(x)$ 가 된다. 이때 $w(x)$ 는 다음 수식 (1) 로 얻을 수 있다.

$$\min_{w(x)} \int (f(x)/g(x) - w(x))^2 dx \quad (1)$$

제안한 방식에서는 $w(x)$ 를 p 차 다항식으로 표현하였는데 이를 이산화하면 다음과 같다.

$$w_i = w(x_i) = \sum_{k=0}^p a_k x_i^k, \quad (2)$$

보정 데이터로 선택한 피크 영역의 데이터 개수를 N 이라 하고 $b_i = \frac{f_i}{g_i} = \frac{f(x_i)}{g(x_i)}$ 라 할 때, 단순한 표현을 위해 수식 (1)과 식 (2)를 벡터와 행렬로 나타내면 결국 수식(3)과 수식(4)로 표현할 수 있다.

$$w = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^p \\ 1 & x_2 & x_2^2 & \dots & x_2^p \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_N & x_N^2 & \dots & x_N^p \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_p \end{bmatrix} = \mathbf{X}\mathbf{a}, \quad \mathbf{b} = \begin{bmatrix} f_1/g_1 \\ f_2/g_2 \\ \vdots \\ f_N/g_N \end{bmatrix} \quad (3)$$

$$\min_{\mathbf{a}} \|\mathbf{b} - \mathbf{X}\mathbf{a}\|^2 \quad (4)$$

위 문제에 대한 해는 잘 알려져 있으며, 그 해는 아래와 같다.

$$\mathbf{a} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{b} = \mathbf{X}^+ \mathbf{b} \quad (5)$$

이제 이렇게 구한 \mathbf{a} 로부터 두 스펙트럼을 최대한 매칭 할 수 있는 다항식 곡선을 얻을 수가 있고, 이를 이용하면 스펙트럼 $g(x)$ 를 교정 할 수 있다. 제안한 방법의 피크 검출 및 교차 교정 과정을 그림 2.에 나타내었다.

먼저 그림 2.의 (a)에서는 Benzene-Library 스펙트럼에서 피크를 검출하고, 5개의 피크를 크기순서로 선택하였다. 선택된 피크 점을 기반으로 Benzene-514.5nm을 Benzene-Library에 매칭 하였다. 두 스펙트럼은 동일 물질이므로 피크의 위치가 정확히 일치 하는 것을 확인할 수 있으며, 모든 피크들에서 서로의 강도 차이를 줄일 수 있는 1 대 1 교정을 수행하였다. 그 결과 스펙트럼 전반 부영역에서 후반부 영역으로 갈수록 점차 완만하게 낮아

지는 형태의 선형함수를 얻을 수 있었으며, 이를 이용하여 교정된 Benzene-514.5nm 스펙트럼을 얻은 결과, Benzene-Library 스펙트럼의 피크 강도와 유사해졌음을 확인 할 수 있다.

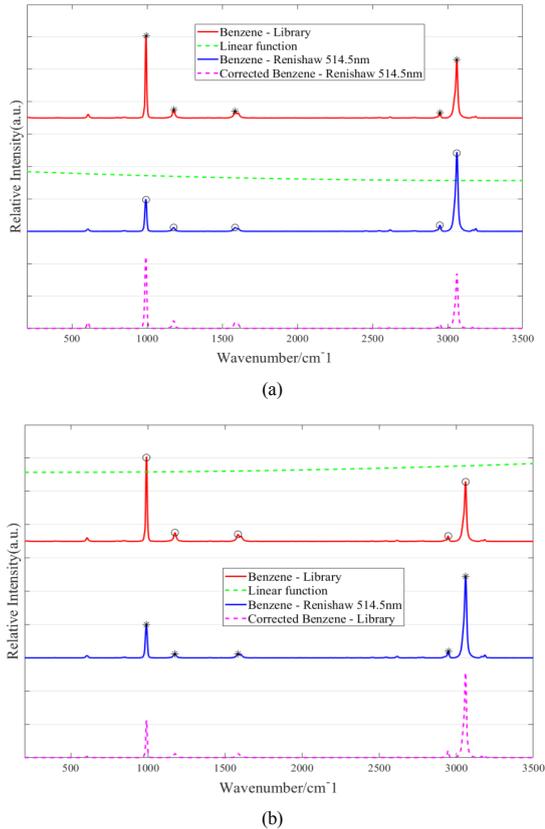


Fig. 2. The process of calibrating
 (a) the Benzene-library to the Benzene-Renishaw 514.5nm
 (b) the Benzene-Renishaw 514.5nm to the Benzene-library

반대로 그림 2.의 (b)에서는 Benzene-514.5nm에서 피크를 선택하고, 이를 기반으로 Benzene-Library를 교정할 수 있는 선형 교정함수를 얻을 수 있었다. 5개의 모 든 점에서 자연스러운 교정이 이뤄 졌으며, 이전의 교정 과는 달리 스펙트럼의 후반부 영역인 $1300 - 3500\text{cm}^{-1}$ 이 완만하게 높아지는 형태를 보이고 있으며 이를 이용 하여 교정된 Benzene-Library를 얻을 수 있었다. 이처럼 두 스펙트럼의 대조에서 교차 교정을 수행하므로, 선형 강도 교정 후 두 스펙트럼의 유사성을 판별하기 위하여 총 2개의 상관 점수를 얻을 수 있으며, 이들의 평균을 최 종 상관도로 결정하였다.

만약 서로 다른 스펙트럼들의 교정을 수행한다면, 교 차 교정에서 서로 제공하는 피크 정보가 상이하기 때문 에 자연스러운 교정이 어렵다는 것을 직관적으로 알 수 있다.

2.4 판별 알고리즘

본 논문에서 사용한 스펙트럼 판별 방법은 상관도 기 반의 알고리즘으로 HQI를 사용하였다[13]. 두 스펙트럼 이 완전 상관할 경우 1로, 겹치는 영역이 전혀 없다면 0 으로 표현된다. 일반적인 라이브러리 기반의 매칭에서는 입력 스펙트럼과 가장 큰 상관도를 갖는 라이브러리의 스펙트럼을 입력 스펙트럼의 신원으로 결정한다. x, y 각각을 정규화한 라만 스펙트럼이라 가정하면, HQI는 아래 수식으로 나타낼 수 있다.

$$HQI = \frac{(x \cdot y)^2}{(x \cdot x)(y \cdot y)} \tag{6}$$

3. 실험 결과 및 고찰

제안한 방식의 성능 평가를 위해, 3가지 장비에서 측 정한 10종의 물질을 14033개 물질의 라이브러리에 대해 판별 실험을 수행 하였으며 그 과정을 그림 3.에 보였다.

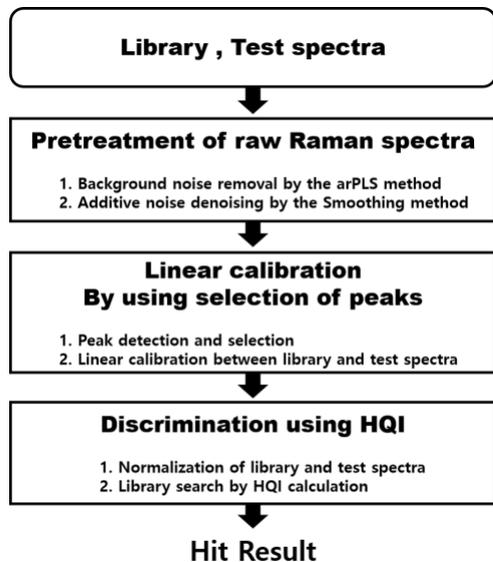


Fig. 3. A flowchart of the identification experiment

제한한 선형 강도 교정 방법의 대표 파라미터는 다항식의 차수이다. 즉, 다항식의 차수가 너무 작으면 스펙트럼의 특징을 적절히 모델링 할 수 없는 문제가 발생하고, 차수가 너무 크면 과적합의 문제가 발생한다. 따라서 먼저 교정 다항식의 차수를 결정하기 위해 차수 변화에 따른 분류성능의 결과를 표 2.에 나타내었다.

이 결과에 따르면 다항식의 차수를 4-7 개로 선택하였을 경우 100% 분류결과를 보인 것을 확인 할 수 있으며, 본 실험에서는 5차 다항식을 사용하였다.

다음으로 선형 피크 교정이 없는 기존의 HQI와 선형 강도 교정을 수행 한 후 HQI 상관도 방식을 수행한 결과를 표 3.에 나타내었다. 표 3.에서 HQI의 판별 결과를

Table 2. Effects of polynomial order for HQI with linear intensity calibration method

The order of polynomial	3	4	5	6	7	8	9	10
Renishaw - 514.5nm	The number of errors	3	0	0	0	0	3	3
inVia - 632.8nm		3	0	0	0	0	0	0
inVia - 785.0nm		3	0	0	0	0	3	3
Total error		9	0	0	0	0	6	6
Classification rate(%)		90	100	100	100	100	93.33	93.33

Table 3. HQI, HQI with linear calibration results for test data sets from 3 different instruments

Instrument	Material	HQI		HQI with Linear calibration	
		Rank	Score	Rank	Score
Renishaw - 514.5nm	Acetone	1	0.9738	1	0.9918
	Acetonitrile	1	0.9726	1	0.9858
	Benzene	1	0.8840	1	0.9987
	Cyclohexane	4	0.9754	1	0.9896
	Ethyl Alcohol	1585	0.8622	1	0.9940
	Ethylene Glycol	4763	0.5660	1	0.9967
	Hexane	1	0.9916	1	0.9904
	Nitromethane	1	0.8605	1	0.9847
	2-Nitrotoluene	1	0.9492	1	0.9983
Toluene	1	0.9278	1	0.9916	
Mean of correlation score		0.8963		0.9922	
Standard deviation of correlation score		0.1259		0.0048	
inVia - 632.8nm	Acetone	1	0.9812	1	0.9682
	Acetonitrile	1	0.9783	1	0.9814
	Benzene	1	0.9677	1	0.9742
	Cyclohexane	1	0.9905	1	0.9775
	Ethyl Alcohol	1	0.9804	1	0.9790
	Ethylene Glycol	1	0.9580	1	0.9900
	Hexane	1	0.9933	1	0.9734
	Nitromethane	1	0.9748	1	0.9767
	2-Nitrotoluene	1	0.9673	1	0.9893
Toluene	1	0.9257	1	0.9575	
Mean of correlation score		0.9717		0.9767	
Standard deviation of correlation score		0.0194		0.0096	
inVia - 785.0nm	Acetone	59	0.5413	1	0.9720
	Acetonitrile	4	0.6387	1	0.9727
	Benzene	13	0.7468	1	0.9552
	Cyclohexane	17	0.5964	1	0.9517
	Ethyl Alcohol	2	0.7958	1	0.9699
	Ethylene Glycol	1	0.9532	1	0.9713
	Hexane	222	0.8030	1	0.9642
	Nitromethane	1	0.7832	1	0.9785
	2-Nitrotoluene	1	0.9088	1	0.9776
Toluene	118	0.6034	1	0.9556	
Mean of correlation score		0.7371		0.9669	
Standard deviation of correlation score		0.1382		0.0097	

살펴보면 Renishaw-514.5nm에서 측정된 10종의 물질 중 3종에서, inVia-785.0nm의 스펙트럼에서는 총 7종에서 분류 오류가 발생한 것을 확인 할 수 있다. 일반적으로 HQI는 두 스펙트럼에서 가장 큰 피크들이 겹칠 경우 나머지 피크들의 형태와 관계없이 높은 상관도를 얻을 수 있는 단점을 보인다. 따라서 피크 강도차이가 발생한 스펙트럼에 대해 매우 취약한 결과를 보이는 것으로 알려져 있고 실험에서 이를 확인할 수 있다.

보다 자세한 분석을 위해 inVia-785.0nm에서 측정된 Toluene 스펙트럼을 분류한 상세 결과를 표 4.에 나타내었다. 표에 따르면 Toluene-785.0nm는 Benzyl alcohol-Library와의 상관도에서 0.7990의 유사성을 보였으며 전체 라이브러리 물질에서 1위를 기록하였다. 이에 비해 Toluene-Library와의 상관도는 0.6034이며 118위를 기록하였다. 이 결과를 기반으로 세 스펙트럼 간의 유사성을 확인하기 위해 각각을 그림 4.에 나타내었다.

Table 4. HQI results of Toluene-inVia 785.0nm

Hit#	Score	Result ID
1	0.7990	BENZYL-ALPHA-13C ALCOHOL, 99 ATOM % 13C
2	0.7757	BENZYL ALCOHOL, 99%
3	0.7579	LITHIUM SULFATE, 99%
⋮	⋮	⋮
118	0.6034	TOLUENE, 99.5+%, A.C.S. SPECTROPHOTOMETRIC GRADE

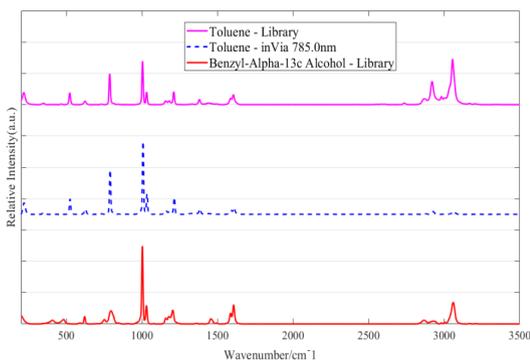


Fig. 4. Three major spectra with miss classification

그림을 살펴보면 Toluene-Library 스펙트럼의 경우 2900–3100 cm^{-1} 영역에서 피크 강도가 우세한 반면에 Toluene-785.0nm의 경우 해당 영역의 피크 강도는

미약하며, 전반부 영역에서는 Toluene-Library의 피크보다 높은 강도의 피크들을 보이고 있다. 두 Toluene 스펙트럼 간에 겹치는 면적은 다소 작기 때문에 낮은 상관도 점수를 얻은 것을 확인 할 수 있다.

이와 대조적으로 Benzyl alcohol-Library 스펙트럼은 2900–3100 cm^{-1} 영역의 피크강도가 다소 낮은 형태를 보이고 있고, Toluene-785.0nm 스펙트럼과 전반부 영역에서 대체적으로 스펙트럼의 면적이 겹치고 있는 것을 확인 할 수 있다. 따라서 Toluene-Library에 비해 Benzyl alcohol-Library이 높은 상관 점수를 획득한 것이다.

다음에는 제안한 선형 교정 방법을 통해 위 세 물질을 교정하는 과정을 그림 5.에 나타내었으며 설명의 간소화를 위해 교차 교정 과정 중 라이브러리 스펙트럼을 입력 스펙트럼에 교정하는 과정은 생략하였다. 먼저 그림 5.의 (a)를 살펴보면 Toluene-Library 스펙트럼에서 강도의 크기 순서로 5개의 피크가 선택되었으며, 입력 스펙트럼은 Toluene-Library와 같은 물질이므로 선택된 5개의 위치에서 아주 유사한 피크가 존재한다. 따라서 우리는 아주 자연스러운 선형 교정 함수를 기대할 수 있고, 획득한 선형 교정 함수는 1500 cm^{-1} 의 스펙트럼 영역을 기점으로 점차 상승하는 형태로서 입력 스펙트럼의 후반부 영역의 피크를 키울 수 있는 형태이다. 이 선형 교정 함수를 기반으로 입력 스펙트럼을 곱한 결과를 살펴보면 Toluene-Library 스펙트럼과 상당히 유사해진 것을 확인할 수 있다. 두 스펙트럼의 상관도는 약 0.9556으로 선형 강도 교정이 없는 HQI에 비해 약 0.3522 정도 상승하였다.

다음으로 그림 5.의 (b)를 보면 입력스펙트럼인 Toluene-785.0nm 를 Benzyl alcohol-Library에 선형 교정한 결과를 확인할 수 있다. 두 스펙트럼은 서로 다른 물질에서 측정되었기 때문에 Benzyl alcohol-Library에서 선택된 5개의 피크 중에 2개의 피크 위치에서 Toluene-785.0nm은 피크가 존재 하지 않고 있다. 따라서 이 두 피크 점이 존재하는 스펙트럼 영역에서는 올바른 교정을 기대하기 어렵다. 교정된 결과를 살펴보면, 스펙트럼 영역 701–900 cm^{-1} 에 존재하는 Toluene-785.0nm의 피크는 전혀 매칭이 이루어지지 않았음을 확인할 수 있다. 결국 Benzyl alcohol-Library는 1000 cm^{-1} 영역의 피크를 제외한 나머지 피크들에서 유사성을 찾기 어려운 것을 확인할 수 있고 Toluene-Library에 비해 낮은 상관도 점수를 예상 할 수 있다. 이렇듯 선형 교정 방법은 라

만 스펙트럼의 고유한 특징인 피크 점을 기반으로 교정을 수하여 강도 차이가 발생한 스펙트럼의 대조에서 물질의 유사도 평가를 용이하게 할 수 있음을 확인하였다.

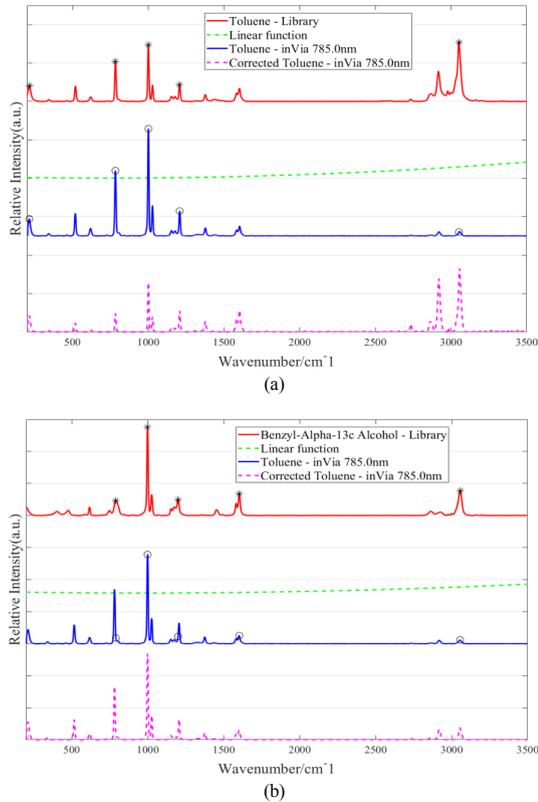


Fig. 5. The process of calibrating
 (a) Toluene-inVia 785.0nm to Toluene-Library
 (b) Toluene-inVia 785.0nm to Benzyl alpha-library

4. 결론

본 논문에서는 측정 장비 혹은 측정 조건에 따른 라만 스펙트럼 강도 차이를 상쇄시키기 위해 선형 교정 방법을 제안하였다. 서로 다른 4가지 기기에서 측정된 실험 데이터를 이용한 실험 결과에 따르면 기존의 HQI 방법이 약 66.7%의 분류 성능을 보인 것에 비해 제안한 방식을 도입한 결과 100%의 분류결과를 보였으며, 상관도의 표준 편차가 낮은 고른 상관도 분포를 보이고 있음을 확인 하였다.

기존 교정 방법들은 장비 간 전달 함수를 미리 모델링하기 때문에 강도차이를 유발하는 측정 온도, 시료의 제

조사, 장비 낙후 등의 문제에 대해 즉각적인 대처가 힘들다. 이에 비해 제안한 방법은 스펙트럼 편별 직전에 피크 중심의 선형 교정을 통해 대조되는 스펙트럼들의 유사성을 최대한 이끌어내는 방식이므로 보다 유연한 방식이라 할 수 있다. 따라서 제안한 방법은 라만 분광에서 요구하는 데이터베이스의 재사용성, 효용성, 라이브러리 간의 합병 및 장비 간의 호환성 등의 문제에 보다 널리 사용될 수 있을 것으로 기대한다.

References

- [1] K. Kong, C. Kendall, N. Stone, I. Notinger, "Raman spectroscopy for medical diagnostics From in-vitro biofluid assays to in-vivo cancer detection", *Adv. Drug Delivery Rev.*, vol. 89, pp. 121-134, March, 2015. DOI: <https://doi.org/10.1016/j.addr.2015.03.009>
- [2] E. L. Izake, "Forensic and homeland security applications of modern portable Raman spectroscopy", *Forensic Sci. Int.*, vol. 202, pp. 1-8, April, 2010. DOI: <https://doi.org/10.1016/j.forsciint.2010.03.020>
- [3] Joonki Hwang, Namhyun Choi, "Fast and sensitive recognition of various explosive compounds using Raman spectroscopy and principal component analysis," *J MOL STRUCT.*, vol. 1038, pp. 130-136, April 2013. DOI: <https://doi.org/10.1016/j.molstruc.2013.01.079>
- [4] J. D. Rodriguez, B. J. Westenberger, L. F. Buhse, J. F. Kauffman, "Standardization of Raman spectra for transfer of spectral libraries across different instruments", *Analyst.*, vol. 136, pp. 4232-4240, July, 2011. DOI: <https://doi.org/10.1039/c1an15636e>
- [5] J.K. Park, A. Park, S. K. Yang, S.J. Baek, et al., "Raman spectrum identification based on the correlation score using the weighted segmental hit quality index", *Analyst.*, vol. 142, pp. 380-388, Dec, 2017. DOI: <https://doi.org/10.1039/C6AN02315K>
- [6] Y. Wang, D. J. Veltkamp and B. R. Kowalski, "Multivariate Instrument Standardization", *Anal. Chem.*, vol. 63, pp. 2750-2756, 1991. DOI: <https://doi.org/10.1021/ac00023a016>
- [7] Y. Wang, M. J. Lysaght, B. R. Kowalski, "Improvement of Multivariate Calibration through Instrument Standardization", *Anal. Chem.*, vol. 64, pp. 562-564, 1991. DOI: <https://doi.org/10.1021/ac00029a021>
- [8] B. Walczak, E. Bouveresse, D. L. Massart, "Standardization of near-infrared spectra in the wavelet domain", *Chemom. Intell. Lab. Syst.*, vol. 76, pp. 41-51, Nov, 1997. DOI: [https://doi.org/10.1016/S0169-7439\(96\)00075-5](https://doi.org/10.1016/S0169-7439(96)00075-5)
- [9] E. S. Etz, S. J. Choquette, W. S. Hurst, "Development and Certification of NIST Standard Reference Materials for Relative Raman Intensity Calibration", *Microchim. Acta.*, vol. 149, pp. 175-184, Feb, 2005. DOI: <https://doi.org/10.1007/s00604-004-0315-2>

[10] J. D. Rodriguez, C. M. Gryniewicz-Ruzicka, J. F. Kauffman, et al. "Transferring Raman Spectral Libraries and Chemometric-based Methods Between Different Instruments and Platforms", *American Pharmaceutical Review*, vol. 16, No1 pp. 9-18, Feb. 2013.

[11] S.-J. Baek, A. Park, Y.-J. Ahn, J. Choo, "Baseline correction using asymmetrically reweighted penalized least squares smoothing", *Analyst*, vol. 140, pp. 250-257, Feb. 2015.
DOI: <https://doi.org/10.1039/C4AN01061B>

[12] S.-J. Baek, A. Park, A. Shen, J. Hu, "A background elimination method based on linear programming for Raman spectra", *J. Raman Spectrosc.*, vol. 42, pp. 1987-1993, Feb. 2011.
DOI: <https://doi.org/10.1002/jrs.2957>

[13] Galactic Industries Corporation: Salem, "Spectral ID Users Guide", *Galactic Industries Corporation*, pp. 71, 1998.

박 아 룬(Aaron Park)

[정회원]



- 2006년 2월 : 한국대학교 한국대학원 한국학과 (한국학석사)
- 2008년 2월 : 한국대학교 한국대학원 한국학과 (한국학박사)
- 2012년 2월 : 한국연구소 책임연구원

<관심분야>

디지털 신호처리, 패턴인식, 바이오 응용 패턴 인식 특징 추출/선택

박 준 규(Jun-Kyu Park)

[정회원]



- 2009년 2월 : 전남대학교 전자컴퓨터공학부(공학사)
- 2017년 2월 : 전남대학교 전자공학과(공학박사)
- 2017년 10월 ~ 현재 : 전남대학교 수학과 Postdoc

<관심분야>

디지털 신호처리, 패턴인식

백 성 준(Sung-June Baek)

[정회원]



- 1989년 2월 : 서울대학교 전자공학과(공학사)
- 1992년 2월 : 서울대학교 전자공학과(공학석사)
- 1999년 2월 : 서울대학교 전자공학과(공학박사)
- 2002년 3월 ~ 현재 : 전남대학교 전자공학과 교수

<관심분야>

의료, 통신, 음성관련 디지털 신호처리