

# 계산 분광기 기반 약물 분류 시뮬레이션 연구

\*김철순<sup>1</sup>, 이지오<sup>2</sup>, 이흥노<sup>2</sup>

<sup>1</sup>한국전자기술연구원 스마트전자부품연구센터

<sup>2</sup>광주과학기술원 인공지능대학원

e-mail : csk0315@keti.re.kr, jioh0826@gm.gist.ac.kr, heungno@gist.ac.kr

## Illegal Drugs Classification based on Computational Spectrometer: A Simulation Study

\*Cheolsun Kim<sup>1</sup>, Jioh Lee<sup>2</sup>, Heung-No Lee<sup>2</sup>

<sup>1</sup>Korea Electronics Technology Institute (KETI)

<sup>2</sup>Gwangju Institute of Science and Technology (GIST)

### Abstract

A spectrometer is a device to measure light intensities with respect to wavelength. Recently, there has been a great attention to reduce the size and cost of spectrometers for practical applications. In this paper, we propose a computational spectrometer for classifying illegal drugs based on sparse representation classification. The simulation results show that the proposed method efficiently classifies illegal drugs.

### I. 서론

분광기는 물질이 방출하거나 흡수하는 빛의 세기를 파장에 따라 측정해 제공해주는 장치이다. 동작하는 파장 영역에 따라 다양한 응용분야에 활용이 가능하다. 그러나, 일반적으로 고해상도 및 넓은 파장 범위에 동작하는 분광기는 값이 비싸고 크기가 커 연구소나 산업체 등 한정된 분야에서 쓰인다.

응용 분야 확대를 위해 크기가 작고, 값이 저렴하며, 해상도가 높은 분광기를 개발하기 위한 연구가 지속되어 왔다. 최근에는 구동장치 없이 필터 배열과

CMOS 센서가 결합된 칩 형태의 분광기가 개발되고 있으며, 최신 신호처리 기술을 활용한 계산 분광기가 제안되고 있다.

본 논문에서는 사회적으로 문제가 되고 있는 약물을 분류하기 위한 계산 분광기와 분류 기법을 소개한다. 시뮬레이션을 통해 분류 기법이 동작함을 보였으며, 이러한 기술은 불법 약물 검사, 위조 판별 등 실시간 현장 감지에 활용 될 수 있을 것으로 예상된다.

### II. 본론

#### 2.1 계산 분광기

계산 분광기는 수치최적화기법, 딥러닝 등 신호처리 방법을 활용하여 입력 광원에 대한 고해상도 분광 신호를 제공하는 장치이다. 신호를 압축하여 측정하고 측정된 신호로부터 원 신호를 찾는 역행렬 문제를 신호처리 기법으로 풀어내는 방식으로 동작한다. 적은 수의 센서(필터)로도 넓은 범위의 파장 정보를 고해상도로 복원 할 수 있어 분광기의 소형화에 장점이 있다. 다층박막[1], 퀴텀닷[2], 나노와이어[3] 등을 활용한 다양한 구조의 분광기가 제안되었으며, 신호처리 기술과 결합하여 우수한 분광 성능을 보여주었다.

본 논문에서는 다층박막필터 배열 기반의 분광기[1]의 투과 특성을 약물 분류를 위한

시물레이션에 활용하였다.

### 2.2 분광 데이터

약물 분류를 위한 분광 데이터로 네덜란드 암스테르담 지역 경찰이 압수하여 측정된 분광 정보를 활용하였다[4]. 동작하는 파장 영역이 서로 다른 다섯 종류의 분광기로 측정된 데이터 중 SCiO 분광기로 측정된 740 - 999 nm 범위의 근적외선 영역의 데이터를 활용하였다.

암페타민, 코카인, 케타민 그리고 MDMA(엑스터시) 등 네 종류의 약물을 분류하는데 각 85, 282, 90, 95 개의 데이터를 사용하였다.

### 2.3 분류 기법

네 종류의 약물을 분류하는 방법으로 희소표현 분류 기술[5]을 활용하였다. 딥러닝 방식의 분류기를 활용할 수 있으나 분류기 학습을 위한 데이터가 충분하지 않아 비교적 간단한 희소표현기반 분류기를 사용하였다. 희소표현기반 분류법은 선형 분류법으로 측정 신호가 들어오면 미리 정의된 딕셔너리 행들의 선형 합으로 측정 신호를 표현하는 희소 신호를 찾는다. 찾은 희소 신호를 이용하여 재구성된 측정값과, 실제 측정값의 오차를 계산하고 오차가 작은 값을 갖는 클래스를 예측 값으로 분류한다. 구체적인 분류 과정은 그림 1과 같다.

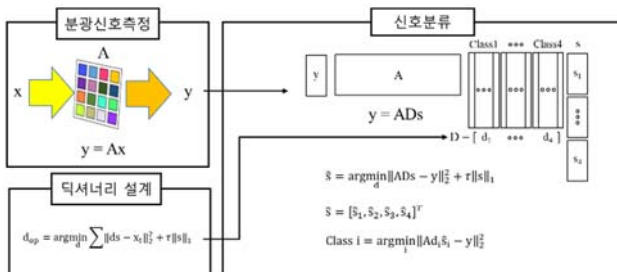


그림 1. 희소표현기반 신호 분류 과정

## III. 시물레이션 결과

그림 1과 같이 측정 신호 y 는 분광기의 투과 특성 A와 분광 신호 x의 행렬곱으로 표현된다. 측정 신호는 분류기의 입력으로 들어가며 4종류의 클래스 중 하나의 클래스로 분류된다.

신호 분류에 앞서 분류기의 딕셔너리를 설계해야한다. 딕셔너리 설계를 위해 딕셔너리 러닝(dictionary learning) [6] 방법을 활용하였다. 위 방법은 주어진 데이터 세트를 희소표현하는 최적의

행렬을 구하는 방법이다.

표 1. 데이터 세트

약물 종류	Train / Test Size
암페타민	65 / 20
코카인	262 / 20
케타민	70 / 20
MDMA	75 / 20

표 1과 같이 약물의 종류에 따라 학습 데이터 세트와 테스트 데이터 세트로 데이터 세트 크기에 맞게 무작위로 나누었다. 각 학습 데이터를 딕셔너리 러닝에 활용하여 최적의 희소화 행렬 d를 구하였으며, 4개의 희소화 행렬을 하나의 딕셔너리 D 로 구성하였다.

약물 종류 별 20개의 데이터 총 80개를 희소표현 기반 신호 분류 기법으로 테스트하였다. 테스트 결과는 표 2와 같다.

표 2. 테스트 결과

실제	예측			
	암페타민	코카인	케타민	MDMA
암페타민	17	3	0	0
코카인	0	18	0	2
케타민	0	0	20	0
MDMA	0	1	0	19

약물 분류 예측 정확도(accuracy)는 92.5%로 나타났다. 각 클래스의 정밀도(precision)는 암페타민 100%, 코카인 81.8%, 케타민 100%, 90.5% 로 나타났다.

## IV. 결론 및 향후 연구 방향

본 논문에서는 계산 분광기와 희소표현기반 분류 기술을 활용하여 4종류의 약물을 92.5% 정확도로 분류하였다. 시물레이션 결과라는 한계가 있지만 사회적으로 문제가 되고있는 불법 약물 분류를 위해 소형의 필터 기반 계산 분광기와 희소표현 분류 방법이 활용될 수 있음을 보였다.

향후, 근적외선 영역에서 동작할 수 있는 소형 계산 분광기를 개발하고 실시간으로 현장 감지 및 분류가 가능한 분광 정보 기반의 신호처리 기술을 개발할 예정이다.

## Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korean government (MSIP) (NRF-2021R1A2B5B03002118)

## 참고문헌

- [1] C. Kim, W.-B. Lee, S. K. Lee, Y. T. Lee, and H.-N. Lee, "Fabrication of 2D thin-film filter-array for compressive sensing spectroscopy," *Opt. Lasers Eng.*, vol. 115, pp. 53-58, 2019.
- [2] J. Bao and M. G. Bawendi, "A colloidal quantum dot spectrometer," *Nature*, vol. 523, no. 7558, pp. 67-70, 2015.
- [3] Z. Yang *et al.*, "Single-nanowire spectrometers," *Science*, vol. 365, no. 6457, pp. 1017-1020, 2019.
- [4] R. F. Kranenburg *et al.*, "Dataset of near-infrared spectral data of illicit-drugs and forensic casework samples analyzed by five portable spectrometers operating in different wavelength ranges," *Data Brief*, vol. 45, p. 108660, 2022.
- [5] K. Huang and S. Aviyente, "Sparse representation for signal classification," *Adv. Neural Inf. Process. Syst.*, vol. 19, 2006.
- [6] I. Tošić and P. Frossard, "Dictionary learning," *IEEE Signal Process. Mag.*, vol. 28, no. 2, pp. 27-38, 2011.