A comparison of different strategies in Principle Component Analysis (PCA) algorithm for clustering human tooth surface using Laser-Induced Breakdown Spectroscopy (LIBS)

N L P Trisnawati¹, A Krisandi¹, I G A Widagda¹, I E Suprihatin² and H Suyanto¹*

¹Physics Department, Faculty of Mathematics and Natural Sciences, Udayana University, Bukit Jimbaran, Badung, Bali, Indonesia 80361
²Chemistry Dept, Faculty of Mathematics and Natural Sciences, Udayana University, Bukit Jimbaran, Badung, Bali, Indonesia 80361

Corresponding author email: hery@unud.ac.id

Abstract. The aim of this work was to observe homogeneity of human tooth surface using classification technique by laser-induced breakdown spectroscopy (LIBS) coupled with principle component analysis (PCA) algorithm. The human tooth was irradiated by 110 mJ Nd-YaG laser (1064 nm) under Helium gas with flow rate of 50 ml/s to produce plasma. Photon emission of the plasma was captured by ocean optic spectrometer HR 2500+ and displayed spectra of intensity as a function of wavelength. The spectra data were analysed by different strategies in PCA algorithm for classifying human tooth surface. The spectra data were split into three ranges that were a full spectra window, FW (200–850 nm), long special spectra window, LSW (380 – 660 nm) and short special spectra window, SSW (550 – 600 nm). These selected suitable input variables using spectral windows can reduce the influence of over fitting phenomena on classification results. Prior to PCA analysing, data were treated by different strategies of pre-processing namely linear baseline correction, area normalisation, and no pre-processing. The results showed that the short special spectral window (SSW) using pre-processing of area normalization could either clustering and distinguishing parts of human tooth surface clearly. Conclusion dentin surface has highest homogeneity of all.

1. Introduction
Laser-induced breakdown spectroscopy (LIBS) is an atomic spectroscopy method to analyse elements qualitatively and quantitatively. It can analyse solid, liquid and gas samples. It is fast, without sample preparation, non-destructive test and yield spectra in real time [1]. In LIBS, laser is focused on sample surface to produce plasma. This plasma atomizes, ionizes, and subsequently excites the interrogated sample. The photon emitted from plasma is collected, generating spectra as intensities vs wavelength [1]. LIBS is well-established technique in many different applications, such as geological [2], industrial [3], and biological samples [4,5,6,7,8]. In term of biological or organic samples, LIBS would have problem in analysing as organic samples including human teeth have spectra similarities to each other until about 90 %. For this reason, LIBS spectra data are better analysed by multivariate data analysis (MVDA)[4]. One of Multivariate algorithm used to classify various sample is principle component analysis (PCA) [9,10,11,12]. These all researches used all LIBS spectra data start form 200 till 850 nm.
This method can influence an overfitting phenomenon on classification results [17]. PCA is an unsupervised MVDA algorithm to classify sample and for reducing the amount of data when correlation was present like LIBS data [13]. Therefore, in this research, the LIBS spectra data were divided into three ranges that were full spectral window (FW), long special spectral window (LSW) and short special spectral window (SSW). It also used differences pre-processing data which were linear baseline correction, area normalization, and no pre-processing (or original data). The aim of this research, therefore, to classify parts of human tooth surface for observing elements quantitative homogeneity on each tooth surface. All data were processed by using unscrambler X.10.5.1 software.

2. Experimental Procedure

2.1 Material
This experiment used a human premolar tooth. The sample was collected from Indonesian people live in Surabaya, East Java. The tooth was from 35 years old girl as shown in figure 1.a,b. The tooth was cut cross sectioned horizontally to analyse its dentin as figure 1b,

![Figure 1. Photograph of Premolar human tooth. a). enamel (buccal and apical). b) dentin](image)

2.2 Experimental Set-up and procedure
Commercial LIBS set-up is presented schematically in figure 2. LIBS consists of Nd-YAG laser (model CFR 200, 7 ns, 1064 nm, 0.7 nm, 200 mJ), spectrometer HR 2500’ (14,336 CCD pixels, resolution 0,1 nm, spectra range 200-980 nm), sample compartment and computer with AddLIBS software [15,16]. Laser was operated on Q-switch mode, frequency of 5 Hz, energy of 110 mJ focused through 10 cm of focal lens on tooth surface of apical, buccal and dentin, which produced plasma. Photon emission from plasma captured by spectrometer with accumulation of 3, delay time detection of 0.5 μs was then displayed as intensities spectra as a function of wavelength on computer. Intensity and wavelength correlated with concentration and kind of elements, respectively. From each parts of apical, buccal and dentin took 3 points as data. The distance between each point is about 2 mm part. All experiments were done under gas of Helium with flow rate of 50 ml/s.
3. Results and Discussions
Before clustering parts of human tooth surface, we took qualitative data from each part of apical enamel, buccal enamel and dentin. Figure 3 presents the typical spectra of dentin, apical and buccal of premolar tooth. Data showed that major elements found in apical, buccal and dentin were the same. They were Ca, F, Si, Na, Sr, K, Mg, Al, and P. From this findings, they can not be distinguished qualitatively neither quantitively as they only had slight intensity difference. Furthermore, intensity will slightly change due to position changing despite taken from same apical, buccal and dentin areas. In order to characterize and identify homogeneity on each area, clustering was done with PCA algorithm. To avoid over fitting phenomena due to many spectra analysed simultaneously, the spectra data were divided into 3 spectral windows named full spectral window (FW), long special spectral window (LSW), and short special spectral window (SSW) as shown in figure 4. Figure 4 is result from PCA processed of spectra data on loading part. Loading part showed elements having high PCA coefficient and affecting on PCA score. Therefore, the range of wavelength on special
window was obtained by this result. SSW wavelength range is a region that has highest coefficient of PCA and LSW wavelength range are found in almost all region that has coefficient of PCA. Based on figure 4, the wavelength range of LSW and SSW are 380 – 660 nm, and 550 -660 nm, respectively. FW wavelength range is all range data taken from LIBS data, 200 – 850 nm. This selected suitable input variables by using spectral windows can reduce the influence of over-fitting phenomena on classification results because analytically useless teeth spectra are eliminated [17].

**Figure 4.** Loading part from PCA using full window

Before PCA processing, pre-processing was conducted for obtaining optimum condition enable to cluster and distinguish each part clearly. Pre-processing used in this experiment were linear baseline correction, no pre-processing, and area normalization, as shown in figure 5.

**Figure 5 a.** PCA score of FW, LSW, and SSW data taken with pre-processing linear baseline correction
Figure 5 a-c showed PCA score on FW spectra data (200 – 850 nm), LSW spectra data (380 – 660 nm), and SSW spectra data (550 – 660 nm) with pre-processing of linear baseline correction, no pre-processing (or original data) and pre-processing of area normalization. The data were collected from three parts of tooth, dentin D, apical (upper enamel) A, and buccal (side enamel) S. From each part, 3 sampling points were taken: D1, D2, D3; A1, A2, A3; and S1, S2, S3.

Figure 5 a, analysis using pre-processing linear baseline correction. LSW and SSW data classified well on dentin and buccal parts. On the other hand, FW data classified less. Apical part did not well classified with all data neither FW, LSW, nor SSW data.

Figure 5 b, analysis using no pre-processing or original data. It could classify all data based on tooth parts with either FW, LSW, or SSW data. But the results, could not be distinguished between FW, LSW, or SSW data. Thus, this kind of analysis is not good.

Figure 5 c, analysis using pre-processing area normalization. It could classify dentin and buccal either with FW, LSW or SSW. The result of LSW had similar result to LSW and SSW data on pre-processing of linear baseline correction.

Based on PCA score data in figure 5 a-c, dentin part was well classified by all data with all kind of pre-processing. Therefore, dentin has high homogeneity in elements quantities. On the other hands, apical part has its elements separated inhomogenously. To validate which processing is better for analysing LIBS spectra data directly, it is necessary to fit with digital data. Digital data were collected from area under curve from each element signals of wavelength 380 – 660 nm (LSW) and were found 42 areas data. This data were processed using PCA algorithm is revealed in figure 6.

---

**Figure 5 b.** PCA score of FW, LSW, and SSW data taken with no pre-processing

**Figure 5 c.** PCA score of FW, LSW, and SSW data taken with pre-processing area normalization
Based on figure 6, PCA score result was similar to pre-processing area normalization of SSW data. Thus, it was advised to use SSW spectra data and area normalization pre-processing. This procedure is the best for classifying big LIBS data.

4. Conclusion
A short special spectral window SSW (550 – 660 nm) with pre-processing of area normalization is the best procedure to identify homogeneity and tooth surface characteristics using clustering technique with PCA algorithm. Apical, buccal, and dentin surfaces have very low, low and very high homogeneity, respectively.

5. References
[1] Cremers D A and Radziemski L J 2006 Laser-Induced Breakdown Spectroscopy (LIBS) Fundamentals and Applications (UK: Cambridge University Press, Cambridge).
[2] Harmon R S, Russo R E and Hark R R 2013 Applications of laser-induced breakdown spectroscopy for geochemical and environmental analysis, A comprehensive review Spectrochimica Acta Part B: Atomic Spectroscopy, vol. 87, pp.11–26
[3] Noll R 2012 Laser-induced breakdown spectroscopy fundamentals and applications, Heidelberg New York: Springer-Verlag Berlin Heidelberg.
[4] Cisewskia, Snyderb E, Hannigaj J and Oudejansb L 2013 Support vector machine classification of suspect powders using laser-induced breakdown spectroscopy (LIBS) spectral data, journal of Chemometrics, vol. 26, pp.143-149
[5] Kaiser J, Novotný K, Martin M Z, Hrdlička A, Malina R, Hartl M and Kizek R 2012 Trace elemental analysis by laser-induced breakdown spectroscopy, Biological applications Surface Science Reports, vol. 67 (11–12), pp. 233–243.
[6] Rehse S J, Salimnia H and Miziolek A W 2012 Laser-induced breakdown spectroscopy (LIBS): an overview of recent progress and future potential for biomedical applications, Journal of Medical Engineering & Technology, vol. 36(2), pp. 77–89.
[7] Santos D J, Nunes L C, de Carvalho G G A, Gomes M d S, de Souza P F, Leme F d O and Krug F J 2012 Laser-induced breakdown spectroscopy for analysis of plant materials, A review,” Spectrochimica Acta Part B: Atomic Spectroscopy, vol. 71, pp. 3–13.
[8] Pořízka P, Procházková P, Procházk D, Sládková L, Novotný J, Petrílák M and Kaiser J 2014 Algal Biomass Analysis by Laser-Based Analytical Techniques, A Review,” Sensors, vol. 14(12), pp. 17725–52.
[9] Harhira A, Bouchard P, Rifai K, EL Haddad J, Sabsabi M, Blouin A, and Lafflamme M 2017 Advanced Laser-Induced Breakdown Spectroscopy (LIBS), sensor for Gold Mining
Proceedings, The 56th conference of Metallurgists COM 2017-08-30. Proc ISBN: 978-1-926872-36-0

[10] Madhavi Z, Martin, Nicole L, Timothy G, Rials and Stan W D 2005 Analysis of preservative-treated wood by multivariate analysis of laser-induced breakdown spectroscopy, *Spectrochimica Acta Part B*, vol. 60 (7), pp. 1179-85.

[11] Prochazka D, Mazura M, Samek O, Rebrosova K, Parizka P, Klus J, Prochazkova P, Klus, Prochazkova P, Novo J, Kovytny K and Kaisar J 2018 Combination of laser-induced breakdown spectroscopy and Raman spectroscopy for multivariate classification of bacteria, *Spectrochimica Acta Part B* vol.139, pp.6-12

[12] Porizka P, Klus J, Kepes E, Prochazka D, David W. H, and Jozef K 2018 On the Utilization of principal component analysis in laser-induced breakdown spectroscopy data analysis, *Spectrochimica Acta Part B* vol. 48, pp. 65-82, doi 10.1016/j.sab.2018.05.030

[13] James N M and Jane C M 2005 *Statistics and Chemometrics for Analytical Chemistry*. 5th ed. Printed in Great Britain by Ashford colour Press, Gosport, Hants ISBN 0 131 291920. pp.215-219

[14] Suyanto H, Trisnawati N L P, Putra I K and Suprihatin I E 2018 Characterization Of Human Teeth By Laser-Induced Breakdown Spectroscopy *Journal of Physics: Conference Series*, 1120 012018

[15] Morgan M 2009 *AddLIBSTM Laser-Induced Breakdown Spectroscopy Software*- User Manual version 2.2, (USA : Photon Machines, Inc.)

[16] Morgan M 2011 *InsightTM Laser-Induced Breakdown Spectroscopy System*- User Manual version 1.3, (USA : Photon Machines, Inc.)

[17] Liu K, Tian D, Deng X, Wang H and Yang G 2019 Rapid classification of plastic bottles by Laser-induced breakdown spectroscopy (LIBS) coupled with Partial Least Squares discriminant analysis based on spectral windows (SW-PLS-DA)” *J. Anal. At. Spectrom.*, vol.34, pp.1665-71

**Acknowledgements**

We would like to thank The Ministry of Research, Technology, and Higher Education, through Institute of Research and Community Service (LPPM-UNUD) Udayana University, that has funded this research through contract no. 492.6/UN14.4.A/LT/2019 as well as MIPA-Unud Integrated Research Laboratory for providing LIBS equipment.