Journal of Korea Multimedia Society Vol. 12, No. 12, DECEMBER 2009(pp. 1755-1760)

Analysis of Hyperspectral Dentin Data Using Independent Component Analysis

Sung-Hwan Jung†

ABSTRACT

In this research, for the first time, we tried to analyse Raman hyperspectral dentin data using Independent Component Analysis (ICA) to see its possibility of adoption for the dental analysis software. We captured hyperspectral dentin data on 569 spots on a molar with dental lesion by HR800 Micro Raman Spectrometer at UMKC-CRISP (University of Missouri at Kansas City-Center for Research on Interfacial Structure and Properties). Each spot has 1,005 hyperspectral data. We applied ICA to the captured hyperspectral data of dentin for evaluating ICA approach, and compared it with the well known multivariate analysis method, PCA. As a result of the experiment, ICA approach shows better local characteristic of dentin than the result of PCA. We confirmed that ICA also could be a good method along with PCA in the dental analysis software.

Key words: ICA, Hyperspectral data analysis, Raman spectrometer

1. INTRODUCTION

Nowadays, owing to the development of dentistry and related equipments, the prevention and treatment of teeth related disease has become easier than before. It has been partly contributing to human wellbeing. However, as the life span of human being becomes longer, the advanced research related with teeth disease and treatment techniques are still needed to be developed.

According to the statistics, about 34,000 American got some cancer related with teeth disease every year and about 8,000 persons lost their life. It is roughly corresponding to a person to death per an hour [1]. About 42% of 6 to 19 years old students have at least one dental caries or had a treatment on their permanent teeth. About 69% of aged 35 to 44 adults lost at least one permanent tooth because of accidents or gum diseases or dental caries. As a person approaches to the age of about 74, it is published that about 26% of old persons might lost all permanent teeth [2].

To prevent and treat these kinds of dental problems, more research and treatment along with advanced equipments are urgently required. Raman spectrometer has been one of important equipments for dental research and treatment. It is able to get hyperspectral data from teeth or bones.

Hyperspectral data means the huge data which can be captured through wide electromagnetic spectrum domain [3]. As a simple example for easy understanding, a human can see only visual light band. However, a mantis shrimp can see not only this band, but also the broad bands from ultraviolet ray to infrared ray band. Therefore, this mantis shrimp can be better than a human in distinguishing a prey or an object in the water. The hyperspectral data can be different from a multispectral data which has tens to hundreds bands, because the hyperspectral data contains from hundreds to

※ Corresponding Author : Sung-Hwan Jung, Address : (641-773) 9 Sarim_dong Changwon, Gyeongnam Korea, TEL : +82-55-213-3815, FAX : +82-55-286-7924, E-mail : sjung@changwon.ac.kr

Receipt date : Nov. 30, 2009, Revision date : Dec. 18, 2009
Approval date : Dec. 19, 2009

† Dept. of computer Engineering, Changwon National University, Korea

※ This research is financially supported by Changwon National University in 2008.

thousands bands.

Typical research areas of hyperspectral data processing are the remote sensing, agriculture, military, and electronic micro scope area, etc. It has received a big attention from many researchers since the mid of 1980s. Figure 1 shows an example of hyperspectral profile from a hyperspectal dentin data set in our research.

Hyperspectal data has n-dimensional spectrum data on a spot, which represents the characteristic of the spot in the spectrum. In Figure 1, this spectral profile is different from other profiles which come from other areas of a tooth. With these spectral profiles, an interpretation of internal structure of the dentin can be partly possible in terms of bio-chemical realm by analyzing the distributed shapes.

Regarding the hyperspectral data related with dental interface, it is found that multivariate approach using PCA (Principal Component Analysis) is more effective and informative than univariate analysis approach [4].

In fact, current some dental analysis software such as Hyperview 3.0 by PerkinElemr Inc., Isys 50 by Malvern Instrument, and SPF (Structure Property Function) 1.0 by UMKC-CRISP (University of Missouri at Kansas City-Center for Research on Interfacial Structure and Properties) provide analysis functions based on PCA.

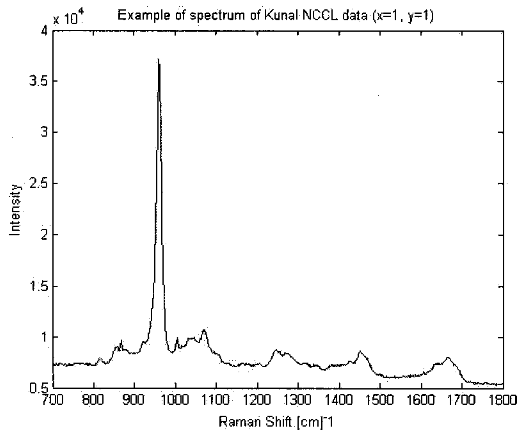


Fig. 1. Example of hyperspectral data extracted from dentin area

As above mention, PCA is one of effective multivariate analyzing methods, but it has some restricted characteristics, compared with ICA (Independent Component Analysis). That is, PCA has the characteristic of orthogonality to the basic direction. It also decides the direction of basic axis by using the 2nd order statistics. On the other hand, ICA has more flexibility about direction than PCA, because it does not need to keep orthogonality. And it also can decide the basic axis by using not only the 2nd order statistic but also the higher-order statistic [5].

ICA has not only application areas such as digital communication area [6], remote sensing area [7] but also medical application areas such as electro-encephalography (EEG), magneto-encephalography (MEG), electrocardiography (ECG), functional magnetic resonance imaging (fMRI) [8-10]. However, most of medical object data are the sampled data in time domain rather than the hyper-spectral data in frequency domain.

In this research, in order to fortify the function of the dental analysis software, SPF 1.0 at UMKC-CRSIP, for the first time, we try to apply ICA to Raman hyperspectral data analysis and want to know the effectiveness of ICA in this kind of application.

2. BSS MODEL AND ICA

2.1 Cocktail-party Problem

BSS (Blind Source Separation) problem can be easily explained by using cocktail-party problem [11]. Suppose that you are in the cocktail-party place now, and then you can hear the mixture sound of background music, people's chat, etc. How can we separate one kind of sound from this mixture sound? To simplify this problem, suppose that two persons (s_1, s_2) are talking now and we are recording their talks by two microphones (x_1, x_2).

The recorded sound in each microphone is mixed with the two persons' voice, according to the each person's relative location. It can be expressed like Equation (1).

$$x_1(t) = a_{11}s_1 + a_{12}s_2$$
$$x_2(t) = a_{21}s_1 + a_{22}s_2 \quad (1)$$

Here, we only know that two persons' voices are different and independent. Based on this assumption and with only given x_1, x_2, to find the independent component s_1, s_2 by estimating coefficients a_{ij} is the object of BSS problem. The method to solve this BSS problem is ICA.

2.2 Blending and Separating Model

The blending and separating BSS general model [12] is expressed in Equation (2). Here, s is the source, non-observable independent component. A is a blending matrix, n is noise which may be in the process of observing, and x is the observed input. B is a separating matrix and y is the estimated source of s. In the Equation, let us ignore the noise n for the simplicity of the model.

$$x = As + n$$
$$\hat{s} = y = Bx \quad (2)$$

In this Equation, the algorithm to find the separating matrix B is ICA algorithm. With this approach, we can solve the BSS problem.

3. ICA ALGORITHM

There are some algorithms of ICA, including Information Maximization (INFOMAX), MISEP, and Joint Approximate Diagonalization of Eigenmatrix (JADE), and others. Here, we will briefly summarize some typical algorithms.

INFOMAX was developed by Bell and Sejnowski in 1995 [13]. But it has some limitations such that it must select the nonlinear function and the function should be fixed. To solve this limitation, in 1998, Almedia proposed MISEP. It adopted the neural network approach to make it adaptive. This can be a generalized INFOMAX [14] which has adaptive characteristics. However, it requires the training of multilayer neural perceptron (MLP). Sometimes its converging is not easy because of the characteristics of input data. JADE was originally developed for digital communication by Cardoso and Souloumiac in 1933 [6]. But it has been used to other application areas including medical area such as EEG and ECG. JADE has an advantage which can separate specific number of independent components from the input data. It can be a good feature to reduce the computation load in case of huge dataset like hyperspectral medical data. However, it also has a limitation having certain of number of extraction components depending on the computation platform.

In this research, we select JADE algorithm because we need to extract a specific number of independent components and to reduce the burden of computation on hyperspectral Raman dentin data. The following is the description of ICA- JADE algorithm [15].

X = the observed input data,
$\hat{R}_x$ = covariance of input data,
$\hat{W}$ = the whitening matrix,
$\hat{Q}_z$= the 4-th order cumulant,
$\hat{B}$ = the separating matrix,
$\hat{s}$ = the desired separating sources

1. form the covariance of input data, $\hat{R}_x$ and compute the whitening matrix, $\hat{W}$.
2. form the 4th order cumulant $\hat{Q}_z$ of the whitened process: $\hat{Z} = \hat{W}\ X$ and compute the n most significant eign pairs $\{\hat{\lambda}_r, \hat{M}_r : 1 \le r \le n\}$.
3. jointly diagonalize the n most significant eigen pairs by a unitary matrix $\hat{U}$.
4. estimate of the separating matrix, $\hat{B} = \hat{W}^{-1}\ \hat{U}$
5. compute the separating source, $\hat{s} = \hat{B}X$.

4. EXPERIMENTAL RESULT AND DISCUSSION

4.1 Experimental environment

In this research, we used as a hyperspectral dentin data, non-caries cervical lesion (NCCL) data which was acquired from a premolar by HORIBA JOBIN YVON company's HR800 Micro Raman Spectrometer at UMKC-CRISP. This data has 569 local observation spots. Each spot provides 1,005 hyperspectral data.

The computing platform used in this research is Win XP, Pentium 4 - 2.8GHz, 1GB RAM PC. We used MATLAB - v7.2 as a software tool.

4.2 ICA and PCA processing results

We extracted 4 components from the hyperspectal data of dentin by ICA-JADE algorithm. At the same time, we extracted 4 components of PCA from the same hyperspectral data for the comparison. The results are shown in the Figure 2 and 3, respectively.

In Figure 2, the right figures show the component profiles of ICA. The left score maps express score vales in pseudo color which represent the amount of the corresponding component at the specific location.

In this figure, the score maps of components represent internal lesion partly. Especially, the score map of component 4 show the internal structure of dentin very well. As a matter of fact, the experimental dental specimen located its tubules in about left-45 degree diagonal direction. So we can recognize its direction by using the scope map.

In Figure 3, the score map of component 4 of

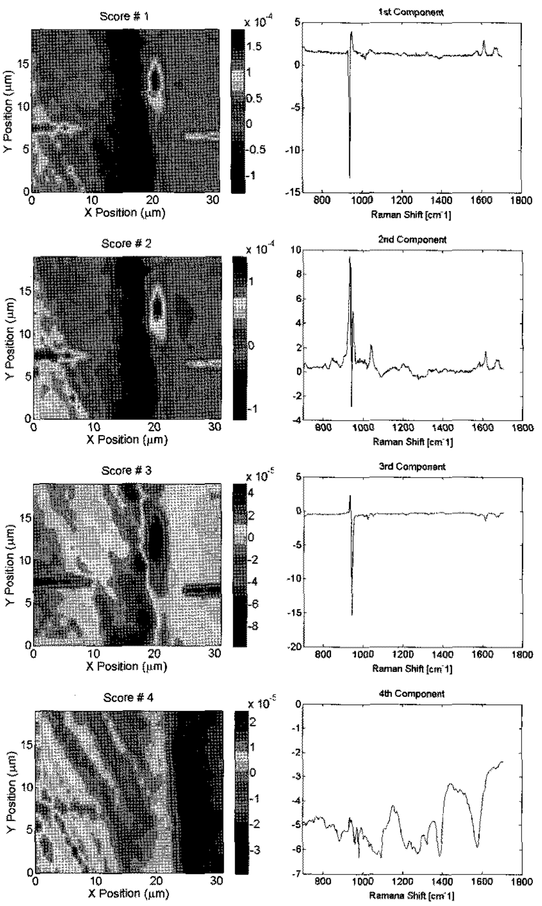


Fig. 2. 2D score maps and profiles of 4 components extracted by ICA

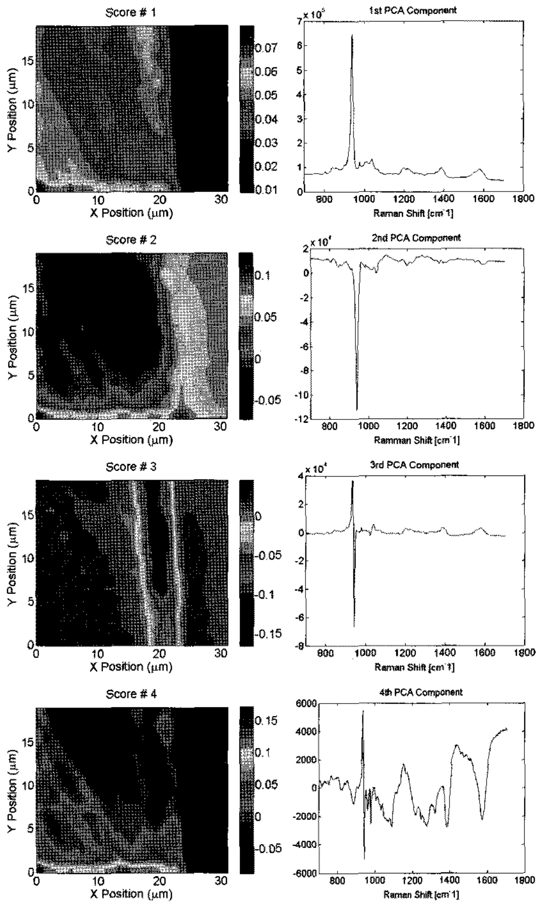


Fig. 3. 2D score maps and profiles of 4 components extracted by PCA

PCA represents a little internal lesion and it shows the internal dentin structure partly.

For the detailed internal structure of dentin, we display 3D score maps of component of ICA and PCA in Figure 4 and 5, respectively. We display the score map of 19×19 ㎛ because of the graphical function's limitation of keeping the same horizontal and vertical length in the MATLAB.

In Figure 4 and 5, the 3D score maps of ICA represent the internal structure of dentin better than those of PCA. Especially, the score map of component 4 of ICA presents the internal structure of dentin in the best.

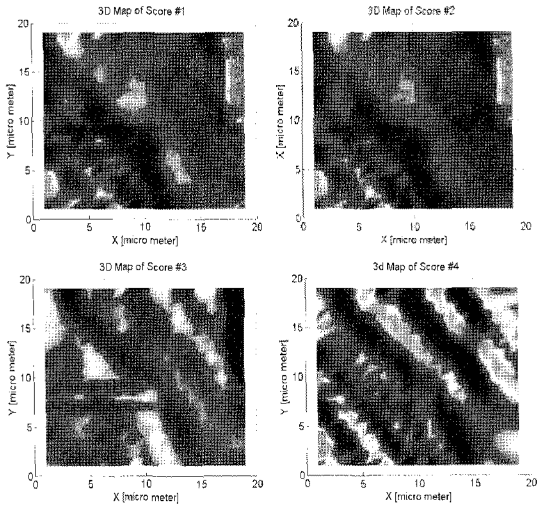


Fig. 4. 3D score maps of extracted independent components by ICA

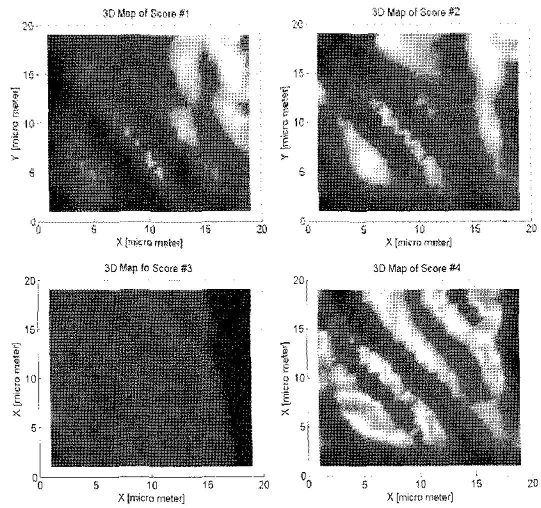


Fig. 5. 3D score maps of extracted independent components by PCA

Through this preliminary experiment of ICA application to the hyperspectral data of dentin, we can confirm that ICA also can be useful like PCA as an effective analysis tool in the multivariate analysis. Further more, we also observe that ICA can represent more detailed internal structure of dentin than PCA.

5. CONCLUSION

This is the first research to try for applying ICA to the hyperspectral data of dentin. We used the hyperspectral data which has 569 x 1,005 sized data from a NCCL molar's dentin. It is captured by HR800 Micro Raman Spectrometer at UMKC-CRISP. We, for the first time, investigate the usefulness of ICA in the area of dental data analysis.

As the results of experiment, ICA presents better information of specific area than the existing multivariate analysis tool, PCA. Especially the internal structural characteristic of dentin is well reflected in the extracted data. Conclusively, through this preliminary research of ICA in the area of dental analysis, we can confirm that it is possible to adopt ICA analysis function in the SPF 1.0, a dental analysis software tool at UMKC-CRISP.

REFERENCES

[1] http://www.oralcancerfoundation.org/facts/

[2] http://www.aaoms.org/dental_implants.php

[3] http://en.wikipedia.org/wiki/Hyperspectral

[4] R.P.G. Thiagarajan, X. Yao, Y-P Wang, P. Spencer and Y. Wang "Application of multivariate spectral analyses in micro-Raman imaging to unveil structural/chemical features of the adhesive/detin interface," Journal of Biomedical Optics, Vol.13, No.1, pp. 014021: 1-9, 2008.

[5] C. Bugli and P. Lambert, "Comparison between Principal Component Analysis and Independent Component Analysis in Electrocencepphalograms Modeling," Vol.48, pp. 1-16, 2006.

[6] J-F Cardoso and A. Souloumiac, "Blind Beamforming for Non Gaussian Signals," IEE-Proceeding-F, Vol.140, No.6, pp. 362-370, 1993.

[7] S. Moussaoui et al., "On the decomposition of Mars hyperspectral data by ICA and Bayesian positive source separation," Neurocomputing, Vol.71, pp. 2194-2208, 2008.

[8] A. Delmore and S. Makeig, "EEGLAB: an open sourxce toolbox for analysis of single-trial EEG dynamics including independent component analysis," Journal of Neuroscience Methods, Vol.134, pp. 9-21, 2004.

[9] J. Xiang, Y. Wang and J.Z. Simon, "MEG Responses to Speech and Stimuli with Speechlike Modulations," Proc. of the 2nd International IEEE EMBS conference on Neural Engineering, Arlington, Virginia, March 16-19, pp. 5-8, 2005.

[10] M. Moosmann et al., "Jointly independent component analysis for simultaneous EEG-fMRI: principle and simulation," International Journal of Psychology, Vol.67, pp. 212-221, 2008.

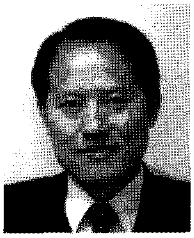
[11] A. Hyvarinen and E. Oja, : Independent Component Analysis Algorithms and Applications," Neural Networks, Vol.13, pp. 411-430, 2000.

[12] J-F Cardoso, C.N.R.S. and E. N. S. T, "Blind signal separation: statistical principles," Proc. of the IEEE, Vol.9, No.10, pp. 2009-2025, 1998.

[13] J. Bell and T. J. Sejnowski, "An Information-Maximization Approach to Blind Separation and Blind Deconvolution," Neural Computation, Vol.7, pp. 1129-1159, 1995.

[14] L. B. Almeida, "Linear and nonlinear ICA base on mutual information - the MISEP method," Signal Processing, Vol.84, pp. 231-245, 2004.

[15] J-F Cardoso, "High-Order Contrasts for Independent Component Analysis," Neural Computation, Vol.11, pp. 157-192, 1999.



Sung-Hwan Jung

He received the B.S., M.S., and Ph. D. degrees from in Electronic Engineering (information and communication major) from Kyungpook National University, Korea in 1979, 1983, and 1988, respectively. He has worked for the Electronic and Telecommunication Research Institute in Korea as a research staff, where he had experienced some national research projects including developing a portable computer. In 1988, he joined the faculty of Department of Computer Engineering at Changwon National University in Korea, where he is currently working as a full professor. From 1992 to 1994, he was a postdoctoral research staff of the Department of Electrical and Computer Engineering at the University of California at Santa Barbara (UCSB). From 1999 to 2000, he also worked for the Colorado School of Mine (CSM) in Golden, Colorado as an exchange professor. From 2008 to 2009, he has experience on the medical information processing at the Dental School of the University of Missouri at Kansas City (UMKC), as a visiting professor. He is an Information System Auditor and P.E. in the area of Information Processing and Electronic Computer. His research interests include content-based image retrieval, steganography, watermarking, internet-based remote monitoring, medical image processing, computer vision and pattern recognition, etc. He is a co-author of seven image processing related books including "Visual C++ Digital Image Processing using Open Source CxImage" and "Practical Computer Vision Programming using VC++ and OpenCV".