

Unsupervised Random Forest Based Analysis of Biomedical Raman Images

Khalifa Mohammad Helal

Department of Mathematics, Comilla University, Cumilla, Bangladesh,
Email: khalifahelal@gmail.com

Received 3 January 2020; accepted 18 February 2020

Abstract. Raman imaging, a live cell imaging technique, is expected to provide a new means to diagnose the diseases, where the images contain both the spatial and spectral information of molecular vibrations of components of biosystems. Biomedical Raman images are the big data having rich molecular information. Useful data analysis is needed to analyze the complex Raman spectra to make the Raman imaging an efficient diagnostic tool. This study involves the application of unsupervised random decision forest (URF) and rate-distortion theory (RDT) based clustering, to analyze the Raman hyperspectral data in order to gain insights into the underlying spectral distributions and discrimination between spectra of various samples. The data are the Raman images of non-alcoholic fatty liver disease (NAFLD) of liver tissues of three different types of rats that were fed with three different diets. URF provides the proximity matrix (representing the similarity among the super pixel spectra (obtained by the extended simple linear iterative clustering (SLIC)), which is then transformed into Euclidean space, and taken as the input of the RDT clustering to extract the meaningful groups having the spectra with similar characteristics. Hierarchical clustering path plot is produced to show how the set of spectra hierarchically makes a group in full spectral dimension as a function of the number of clusters. We expect our proposed approach could be a nice tool for the quantitative analysis of biomedical Raman image data which could be effective to analyze various diseases.

Keywords: Raman images, unsupervised random forest, cluster analysis, rate-distortion theory

AMS Mathematics Subject Classification (2010): 62H30, 62H35, 62H86

1. Introduction

Conventional method to diagnose different disease mainly relies on histopathology, but it often suffers from inter observer variability [1,2,3,4], and it does not provide the detailed chemical information about the samples. Raman microscopic imaging [5,6,7], a live cell imaging technique, is expected to be a prominent alternative to disease diagnosis. It provides the molecular information at the micrometer scale which allows us to study the

cells and tissues at cellular and sub cellular level to facilitate the histological diagnosis of the diseases [7,8]. Raman images contain both the spatial and spectral information of molecular vibrations of constituents of the bio-systems. The different disease states of cells or tissues can lead to Raman spectral differences which can be compared to differentiate the characteristics of those cells and tissues. But the differences from one sample to another in different pathological conditions are very small and difficult to observe in raw Raman spectra [9]. Moreover, the signals in the raw spectra are very weak and are contaminated with noise, auto-fluorescence signals etc. [9]. Therefore, the systematic analysis of Raman images is crucial important to make the complex Raman spectra more interpretable aiding disease diagnosis, with which the true and useful information can be revealed. In this connection, this work deals with the analysis of biomedical Raman images using unsupervised random forest (URF) [10,11,12] based proximity measure combined with information theoretical rate-distortion theory (RDT) [13,14] based clustering to explore the underlying structure of the data.

URF is the unsupervised setting of Random decision forest (RF), a classification and regression ensemble machine learning algorithm, in which results of multiple uncorrelated and weak learner decision trees are aggregated into one final results by simple majority voting [11,12]. URF produces similarity measure by classifying two-class classification problem using supervised RF [11,12], where a synthetic data set, randomly generated from the original data is labeled as class 1 while the original data set is labeled as class 2 [10,11,15]. This proximity measure can be transformed to Euclidean space, which can be used to apply in unsupervised methods, such as principal coordinates analysis (PCoA) [16] and clustering [17,18], to reveal the hidden structure in a data set. In many biological studies, URF was used as an unsupervised learning to extract the meaningful groups in the data [10]. For instance, in [19] URF was used for tumor class discovery, and it was employed on his tone markers of prostate cancer in [20]. To reveal the meaningful groups (clusters) within the tissue images based on their spectral differences due to different phenotypic changes in pathological states, we employed information theory based unsupervised clustering using rate-distortion theory (RDT) [13,14]. RDT takes the dissimilarity matrix as an input and clusters a set of spectra into a smaller number of groups having similar characteristics.

The data used in this work are the Raman hyperspectral images of liver tissues of non-alcoholic fatty liver disease (NAFLD) [21] from three different types of rats which were fed with three different diets. The data is used with the prior permission from the experimentalists and the corresponding authors of [22]. NAFLD is the common liver disorder that affects men, women and children [21]. There are two types of NAFLD: one is simply 'non-alcoholic fatty liver (NAFL)' and the other is 'non-alcoholic steatohepatitis (NASH)' [23,24]. This NASH may lead to liver cirrhosis and/or liver carcinoma [23] which may be irreversible.

Unsupervised Random Forest Based Analysis of Biomedical Raman Images

In this study, at the beginning, the Raman images of rat liver tissues of NAFLD are segmented using extended simple linear iterative clustering (SLIC) [22,25] to generate set of superpixel spectra (the perceptual grouping of pixels having similar characteristics [25]). Here, we choose approximately 100 pixels inside a superpixel and segment the Raman hyperspectral images to obtain superpixel spectra. URF is employed to the set of superpixel spectra to obtain a proximity matrix which is then transformed into a dissimilarity matrix to use it as an input of PCoA [16] for visualization, and an input of RDT clustering algorithm for detecting underlying patterns of the NAFLD data. Based on three dietary states of NAFLD, we use RDT clustering with 3 classes. Confusion matrix obtained from the clustering outputs suggests how the diet states are classified in terms of Raman characteristics. To know the diversity of the dietary states, we generate hierarchical clustering path plot as a function of the number of clusters, which shows how the data makes a group in terms of number of clusters.

2. Methods

The experiments were performed by the groups of researchers at the Department of Pathology and Cell Regulation, Kyoto Prefectural University of Medicine, Kyoto, Japan and the Department of Gastroenterology and Hepatology, Kyoto Prefectural University of Medicine, Kyoto, Japan. The details of the experimental system, Raman measurement, and evaluations of liver tissues can be found in [22]. The data is used in this study with the prior permission from the corresponding authors of [22].

The dataset

The data are the Raman hyperspectral images of rat liver tissues. A total of 48 rat liver tissues were collected after 2,4, 8, and 16 weeks of feeding period. 16(= 4×4) rats were fed with each diet (standard diet (SD), high-fat diet (HFD), and a high-fat high-cholesterol diet (HFHC)) for Raman analyses.

Data preprocessing

Since pure spectral features are hidden underneath noise and contamination [9], the raw Raman data must be preprocessed before quantitative analysis to reduce the influence of those contaminations in the analysis and to extract the true spectral information as much as possible relevant to the NAFLD states. Following subsection will give a brief discussion of the preprocessing schemes of the Raman image data.

Bias correction

The first step is to correct the Raman hyperspectral images from the bias which is assumed to be noiseless but quite large constant value in intensity. For each hyperspectral image, bias is calculated as the minimum intensity of the entire image, and is subtracted from each spectrum of the corresponding Raman image.

Baseline removal

In our data, auto-fluorescence contributions (background spectra/baseline) were estimated using modified multi-polynomial fitting method, which is a recursive polynomial fitting algorithm [26]. It recursively estimates the unwanted contribution (anomaly) from the background. Each single pixel spectrum S is first fit to a polynomial p , and then a modified spectrum is constructed as the minimum intensities of original spectrum and the fitted polynomial. We compute n_b as the number of spectral points of S below the fitted polynomial p , and continue the procedure until a terminative criterion holds. To make the terminative criterion consistent for all the spectra in our application, we iterate polynomial fitting until $n_b \leq 0.01 n_w$, where n_w is the total number of wavenumbers. 8th order polynomial was chosen for fitting the spectra with the above-mentioned stopping criterion. Finally, the baseline corrected spectrum is obtained by subtracting the baseline from the original spectrum. The baseline correction procedure for a single spectrum is illustrated in Fig. 1.

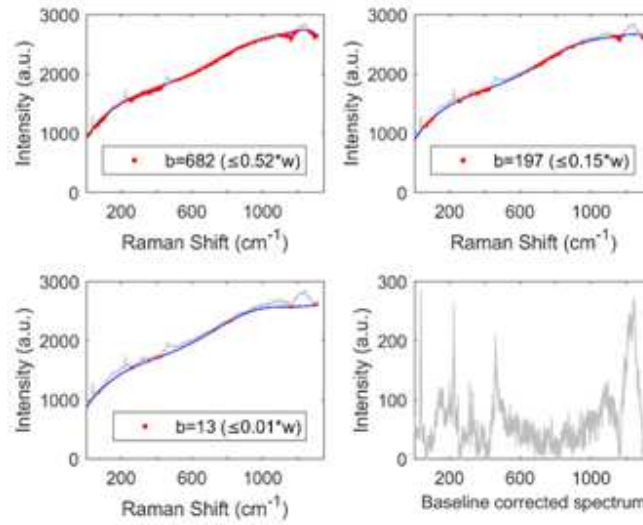


Figure 1: Sample spectrum with baselines at different number of iterations by recursive polynomial fitting.

Segmentation averaging and removal of silent region

After bias and baseline correction, we segment the Raman tissue images into superpixels using extended SLIC [22,25]. The size of each superpixel is chosen so that each superpixel contains approximately 100 pixels. One averaged spectrum is computed from one superpixel, and in this way, we collect a set of superpixel spectra across all the 48 Raman images of liver tissues. All the superpixel spectra are then cropped by removing the silent region ($1801-1799 \text{ cm}^{-1}$), where there are no Raman signals in the tissues [27].

Normalization

Taking into account the signal intensity variation among the spectra due to the concentration difference of the tissue samples, area normalization is used for making the spectra comparable. The area normalized spectrum $s'_w(i)$ of the i -th superpixel spectrum $s_w(i)$ is derived as

$$s'_w(i) = \frac{s_w(i)}{\sum_{w=1}^W s_w(i)}.$$

3. Analysis of the Raman image data

For the quantitative analysis of Raman image data, we performed unsupervised random forest (URF) to derive the dissimilarity matrix among the spectra and rate-distortion theory (RDT) based clustering to detect the underlying structures of the data.

Unsupervised random forest (URF)

URF [10,11,12] is expected to be a promising approach for extracting underlying patterns in the data. URF algorithm is the extension of supervised RF which is developed on the assumption that any structured data should be distinguishable from its randomly generated version [10,11,15]. Appending an original data set with its randomly generated version, synthetic data set, URF forms an artificial two-class classification problem which is then modeled using supervised RF, where a synthetic data set is labeled as class 1 while the original data set is labeled as class 2 [10,15].

URF algorithm is composed of the following steps [10,15]:

- Step 1. A synthetic dataset (same size as original data set) is created by randomly sampling equal number of observations from the marginal distributions of the original variables.
- Step 2. Concatenate the synthetic dataset below the original data, and labeled the original and synthetic data set as “class 1” and “class 2”, respectively, to form a two-class classification problem.
- Step 3. Apply supervised RF to this artificial two-class classification problem.
- Step 4. Derive the proximity matrix using the observations only from the original dataset.

Proximity matrix provides the similarity between the Raman spectra. In RF, proximity between two spectra is defined as the number of times that these two spectra share the same terminal nodes of a RF tree, divided by the total number of trees in RF [10,11,15]. The proximities between spectra construct a proximity matrix $PM = [P_{i,j}]_{n \times n}$, n is the number of spectra, which is symmetric, positive definite [10]. RF dissimilarity between two spectra is defined by $1 - [P_{i,j}]$. Dissimilarity matrix DM is constructed by the all the pairwise dissimilarities between the spectra, which is used as an input of RDT clustering and principal coordinates analysis (PCoA) [16] to detect the structure of the NAFLD data.

Clustering using rate-distortion theory (RDT)

Clustering [17,18] is unsupervised algorithm which groups the elements (spectra) of a data set into subgroups (clusters) having the elements (spectra) with similar Raman information.

We use information theory based clustering by using rate distortion theory (RDT) to Raman hyperspectral image data of NAFLD. RDT [13,14] mainly addresses the phenomenon “compressing the data keeping an acceptable level of distortion”. That is, it finds the number of clusters needed to describe the data within a minimal expected distortion. RDT [13,14] clusters a set of spectra by returning the conditional probabilities of assigning a given spectrum into a given cluster.

Suppose $\mathbf{S} = \{s_1, s_2, \dots, s_N\}$ be the set of N spectra and $\mathbf{C} = \{C_1, C_2, \dots, C_N\}$ is the set of N_c clusters. RDT clusters the sets of spectra \mathbf{S} into set of clusters \mathbf{C} by minimizing the rate $I(\mathbf{C}; \mathbf{S})$ constrained by the mean distortion $\langle D(\mathbf{C}, \mathbf{S}) \rangle$ through the minimization [28,29,30] of the functional [13,14,31,32],

$$\mathcal{F}[p(C_k|s(i))] = I(\mathbf{C}; \mathbf{S}) + \beta \langle D(\mathbf{C}, \mathbf{S}) \rangle, \quad (1)$$

where $p(C_k|s(i))$ is the conditional probability of assigning a given spectrum $s(i)$ into a given cluster C_k ,

$I(\mathbf{C}; \mathbf{S})$ is the degree of compression of the spectra into clusters which is defined as

$$I(\mathbf{C}; \mathbf{S}) = \sum_{k=1}^{N_c} \sum_{i=1}^N p(C_k|s(i)) p(s(i)) \log \frac{p(C_k|s(i))}{p(C_k)}, \quad (2)$$

and $\langle D(\mathbf{C}, \mathbf{S}) \rangle$ is the mean distortion among the spectra defined as the mean of the pair wise distance between all pairs of spectra within the set of clusters \mathbf{C} , averaged over all clusters which is defined as

$$\langle D(\mathbf{C}, \mathbf{S}) \rangle = \sum_{k=1}^{N_c} p(C_k) \left[\sum_{i,j=1}^N p(s(i)|C_k) p(s(j)|C_k) d_{i,j} \right], \quad (3)$$

Here, $p(C_k)$ is the marginal probability of the cluster C_k , β is the Lagrange multiplier (control parameter) which controls the softness of the clustering.

The parameters of RDT algorithm is the number of clusters N_c and β . We choose $N_c = 3$ to cluster the data based on three different dietary states, and β is chosen large enough to make RDT as a hard clustering algorithm.

4. Results and discussions

URF is employed to the set of all superpixel spectra obtained from all the 48 Raman images of liver tissues from SD, HFD, and HFHC groups, where RF comprised of 5000 fully grown trees was run to obtain the proximity matrix. The proximity matrix is then converted into dissimilarity matrix to employ it into PCoA and RDT clustering.

First, we apply PCoA to the dissimilarity matrix. Fig. 2 represents the visualization results of PCoA as a usual scatter plot of spectra in 2-D embedded space, where the spectra are colored according to three different diets. We see that the spectra are clearly separable into three groups. We observe that the SD group of spectra are located far-away from those of the HFHC group, while HFD spectra are located closed to both SD and HFHC spectra, and are largely varied. Hence, it is revealed that Raman spectral features is significantly larger in HFD group than those at SD and HFHC groups, from which we would biologically expect that HFD group has larger diversity of microchemical environment than SD and HFHC group.

This larger dispersion of HFD is harmonized with the results illustrated in Fig. 3 which shows that the mean of the dissimilarities of the spectra from each of SD, HFD and HFHC group are 0.4207, 0.4723 and 0.3770 with the corresponding standard deviations

Unsupervised Random Forest Based Analysis of Biomedical Raman Images

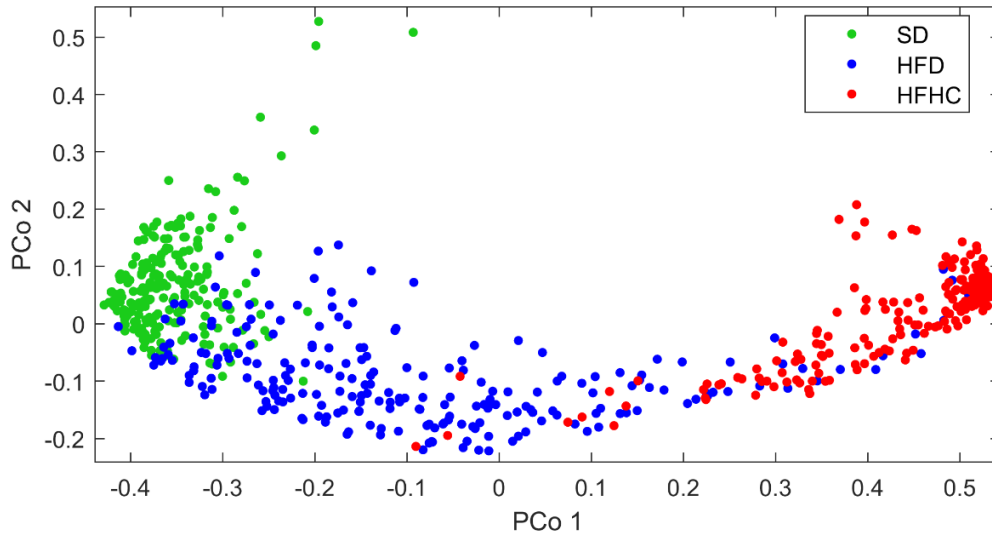


Figure 2: PCoA score plot of the spectra in 2-D embedded space obtained using the URF dissimilarity matrix.

0.0870, 0.1348 and 0.1079, respectively. Here, the variation of the dissimilarities of the HFD spectra is larger than that of SD and HFHC.

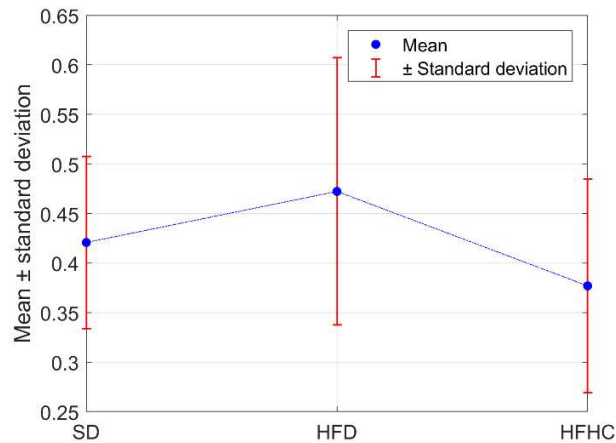


Figure 3: The mean and standard deviation of the distances among the spectra in different diet groups.

We perform RDT classification with 3 classes on the dissimilarity matrix. We have 3 different clusters corresponding to the different parts/regions across all the liver tissue having identical biochemical environment. The distributions of the spectra within the clusters are shown in the projected 2-D PCoA space [Fig.4] (each color corresponds to individual cluster).

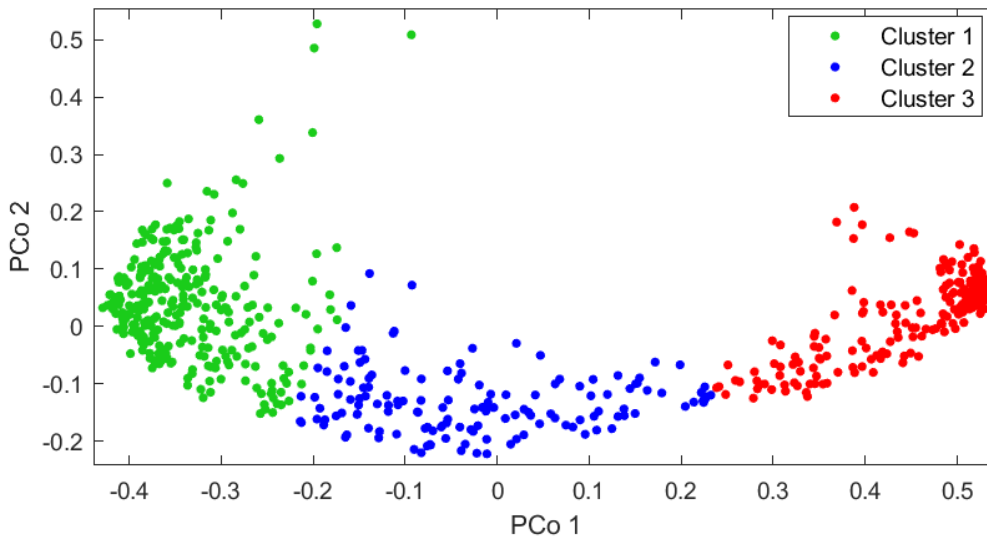


Figure 4: Distributions of the spectra within three clusters in the 2-D PCoA space.

Visual comparison of Fig. 2 and Fig.4 suggests that RDT well classifies the SD and HFHC groups (correspond to green and red clusters), but the classification is not well performed in case of the HFD group, suggesting that the HFD group is largely diverse in terms of Raman information.

To quantify how well RDT classifies the dietary states, we use performance matrix. Assuming three dietary states as true classes (truth labels), we compare those to the RDT clustering outputs (the predicted classes). We find the best-matching pairs between the truth labels and RDT clustering labels such that the number of true positive (TP) is maximized.

Table 1: Confusion matrix for RDT clustering

		Predicted			
		SD	HFD	HFHC	Total
Actual	SD	211	13	0	224
	HFD	17	130	77	224
	HFHC	0	0	224	224
	Total	228	143	301	672

The confusion matrix of RDT clustering to classify SD, HFD and HFHC groups is shown in Table 1. It shows that, RDT classifies the SD, HFD and HFHC groups with the accuracy of $\approx 94\%$, $\approx 58\%$, and $\approx 100\%$, respectively. We see that the prediction of HFD

Unsupervised Random Forest Based Analysis of Biomedical Raman Images

is less accurate, which indicates that HFD is a heterogeneous group having diverse microchemical environment. This result also harmonizes with the results shown in Figs. 2 and 3.

Prespecified three clusters may not be enough to sufficiently quantify the diverse chemical environments of the dietary states. Increasing the number of clusters may provide better insights into the heterogeneity of the chemical environment of different diet states. We calculate the hierarchical clustering path plot of the dietary states as a function of the number of clusters, which displays how the subset of spectra will belong to each cluster in full dimension. Fig. 5 to Fig. 8 display the hierarchical clustering path plots of all the spectra, and the subset of spectra from SD, HFD and HFHC groups. Here, each marker point represents the mean index of a cluster corresponding to the number of clusters. Each line represents the path of the spectra, i.e., how they distributed if we increase the number of clusters. The thickness of the line indicates the probability of “path of those spectra” belonging to clusters.

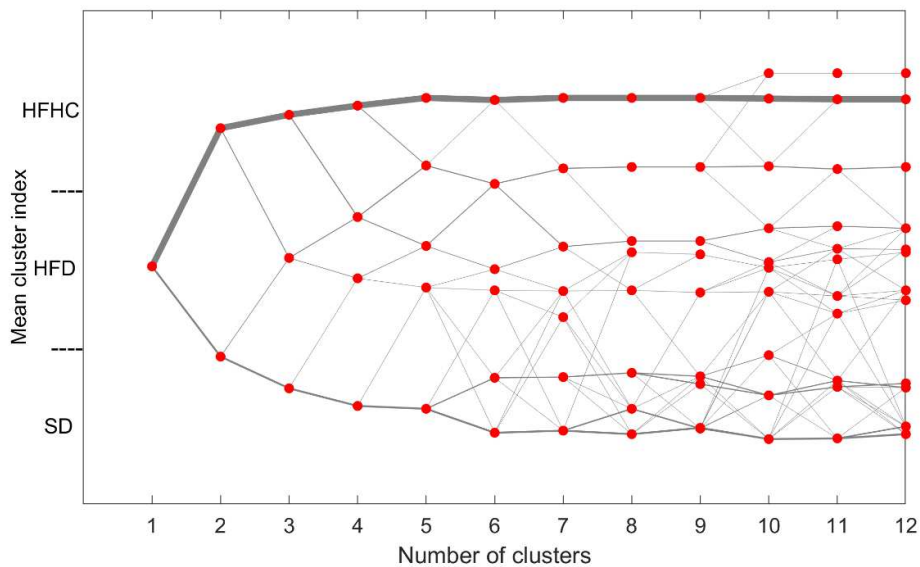


Figure 5: Hierarchical clustering path plot for all the spectra from SD, HFD and HFHC groups

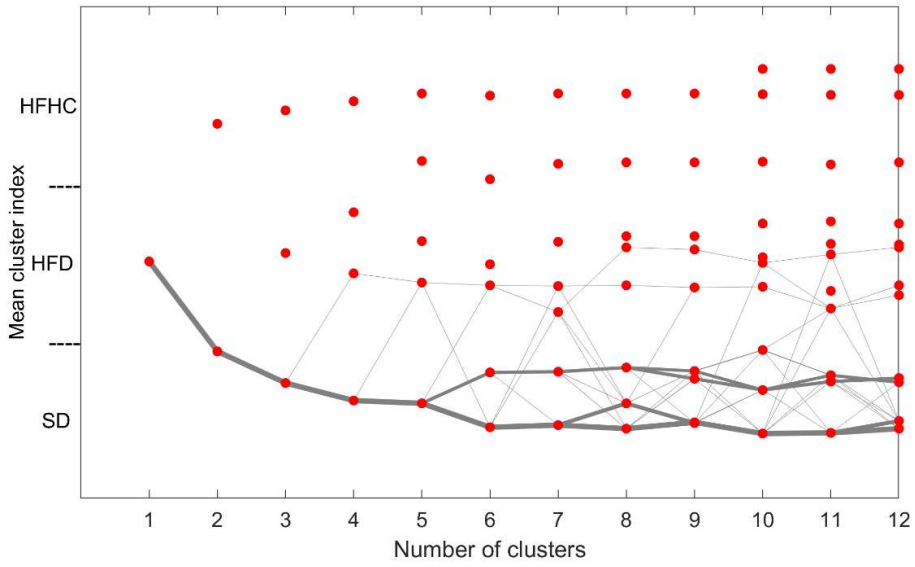


Figure 6: Hierarchical clustering path plot for all the spectra from SD group.

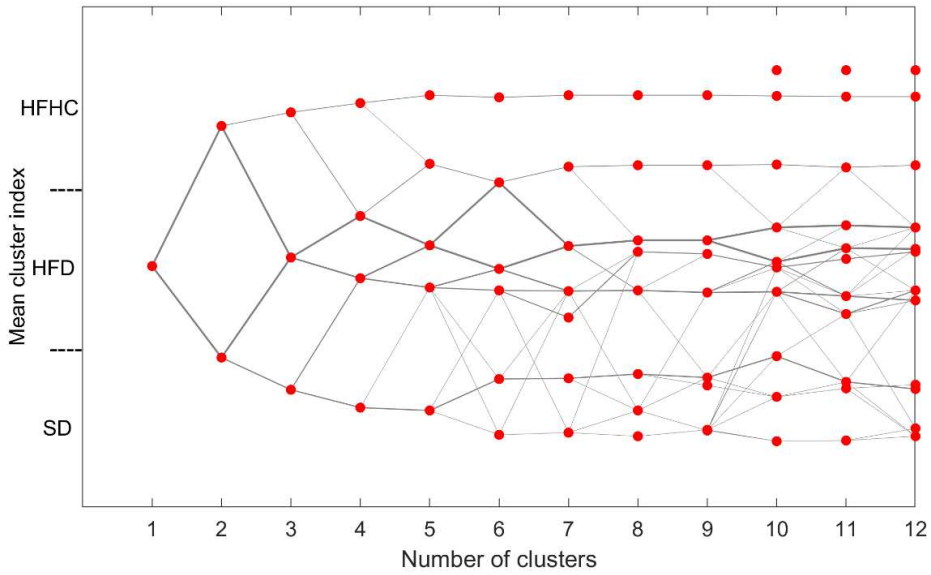


Figure 7: Hierarchical clustering path plot for all the spectra from HFD group

Unsupervised Random Forest Based Analysis of Biomedical Raman Images

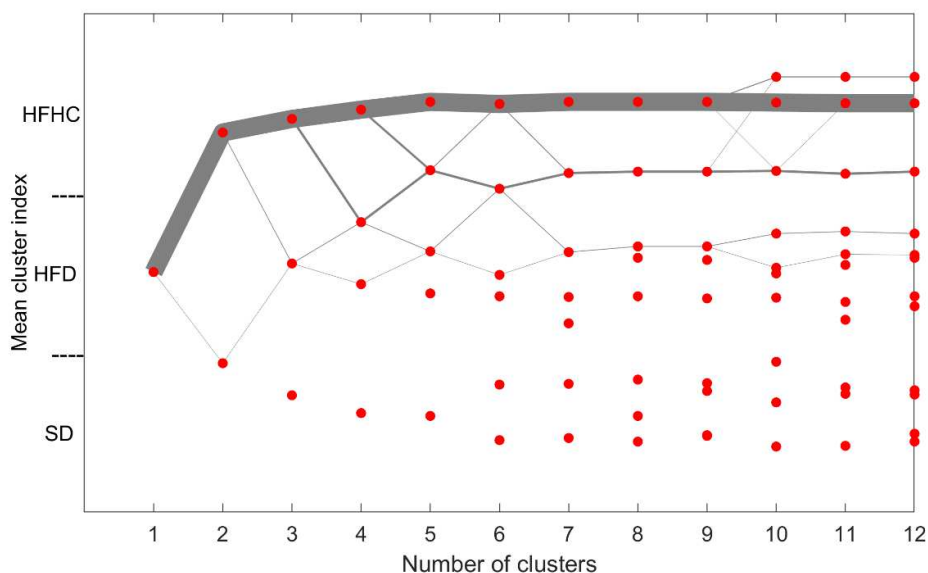


Figure 8: Hierarchical clustering path plot for all the spectra from HFHC group.

Fig. 5 shows how the set of superpixel spectra from all the three states hierarchically makes a group in full dimension. In Fig. 6, we see that SD spectra rarely change microchemical environment. The high probable path represents that most of the spectra in SD group belongs to single cluster indicating a good classification. Fig. 7 shows that HFD spectra are largely diffused indicating a larger diversity of microchemical environment which results the less accuracy of the RDT clustering. Finally, Fig. 8 illustrates that HFHC diet state quickly changes microchemical environment, and most of the spectra are localized in a single cluster indicating a good prediction.

5. Conclusions

This work dealt with the quantitative analysis of biomedical Raman images of non-alcoholic fatty liver disease (NAFLD) using unsupervised random forest (URF)-derived proximity measure combined with rate-distortion theory (RDT) based clustering, to explore the underlying information of the data. URF derived the dissimilarity matrix from the spectra, which was used as the input of unsupervised learning algorithms, namely, principle coordinate analysis (PCoA) and RDT based clustering. RDT revealed the hidden patterns in the data where Raman spectra are visible into separable groups. RDT clustering distinguished SD and HFHC into separate groups with high accuracy, but in case of HFD the accuracy was not high. It was found that HFD spectra were largely diffused compared to SD and HFHC spectra, indicating a larger diversity of microchemical environment in this HFD group. Hierarchical clustering path plot shed light upon how a set of spectra hierarchically makes a group in full spectral dimension.

Khalifa Mohammad Helal

Acknowledgements. I thank Prof. T. Komatsuzaki, Graduate School of Life Science, Research Center of Mathematics for Social Creativity, Institute for Electronic Science, Hokkaido University, Sapporo, Japan and Prof. Y. Harada, Department of Pathology and Cell Regulation, Kyoto Prefectural University of Medicine, Kyoto, Japan for giving the permission to use the Raman image data of NAFLD. I am also thankful to all the authors of [22] and the anonymous referees for their valuable comments for the further improvement of the paper.

REFERENCES

1. J.Yan, Y.Yu, J.W.Kang, et al., Development of a classification model for non-alcoholic steatohepatitis (NASH) using confocal Raman micro-spectroscopy, *J. Biophotonics*, 10 (2017) 1703-1713.
2. S.F.Cramer, L.M.Roth, S.E.Mills, et al. Sources of variability in classifying common ovarian cancers using the World Health Organization classification, Application of the pathtracking method, *Pathology Annual*, 28 (1993) 243-286.
3. S.C.Hollensead, W.B.Lockwood and R.J.Elin, Errors in pathology and laboratory medicine: consequences and prevention, *J. Surg. Oncol.*, 88 (2004) 161-181.
4. S.S.Raab, Variability of practice in anatomic pathology and its effect on patient outcomes, *Seminars in Diagnostic Pathology*, 22 (2005) 177-185.
5. G.J.Puppels, F.F.M.de Mul, C.Otto, et al., Studying single living cells and chromosomes by confocal Raman microspectroscopy, *Nature*, 347 (1990) 301-303.
6. M.Okada, N.I.Smith, A.F.Palonpon, et al., Label-free Raman observation of cytochrome c dynamics during apoptosis, *Proc. Natl. Acad. Sci.*, 109 (2012) 28-32.
7. A.S.Haka, Z.Volynskaya, J.A.Gardecki, et al., In vivo margin assessment during partial mastectomy breast surgery using Raman spectroscopy, *Cancer Res.*, 66 (2006) 3317-3322.
8. L.A.Austin, S.Osseiran and C.L.Evans, Raman technologies in cancer diagnostics, *Analyst*, 141 (2016) 476-503.
9. R.Gautam, S.Vanga, F.Ariese, et al., Review of multidimensional data processing approaches for Raman and infrared spectroscopy, *EPJ Techniques and Instrumentation*, 2 (2015) 8.
10. N.L. Afanador, A.Smolinska, T.N.Tran and L.Blanchet, Unsupervised random forest: a tutorial with case studies, *Journal of Chemometrics*, 30 (2016) 232-241.
11. L.Breiman, Random forests, *Mach Learn*, 45 (2001) 5-32.
12. A.Liaw and M.Wiener, Classification and regression by random forest, *R News*, 2 (2002) 18-22.
13. J.N.Taylor, C.B. Li, D.R. Cooper, et al., Error-based extraction of states and energy landscapes from experimental single molecule time-series, *Sci. Rep.*, 5 (2015) 9174.
14. N.Slonim, G.S.Atwal, G.Tkacik and W.Bialek, Information-based clustering, *Proc. Natl. Acad. Sci.*, 102 (2005) 18297-18302.

Unsupervised Random Forest Based Analysis of Biomedical Raman Images

15. L.Breiman, Manual on setting up and understanding Random forests, (1999) https://www.stat.berkeley.edu/~breiman/Using_random_forests_V3.1.pdf.
16. B.G.M.Vandeginste, L.M.C.Buydens, S. de Jong, et al., *Handbook of Chemometrics and Qualimetrics: Part B*, Vol. 20B, Elsevier: Amsterdam (1997), 146.
17. L.Kaufman and P.J.Rousseeuw, *Finding Groups in Data: an introduction to Cluster Analysis*, John Wiley and Sons (1990).
18. A.K.J.Jain and R.C.Dubes, *Algorithm for Clustering Data*, Prentice Hall (1988).
19. T.Shi, D.Seligson, A.S.Beldegrun, et al., Tumor classification by tissue microarray profiling: random forest clustering applied to renal cell carcinoma, *Modern Pathology*, 18(4) (2005) 547–557.
20. D. B.Seligson, S.Horvath, T.Shi, et al., Global histone modification patterns predict risk of prostate cancer recurrence, *Nature*, 435 (7046) (2005) 1262–1266.
21. K.Kochan, E.Maslak, C.Krafft, et al., Raman spectroscopy analysis of lipid droplets content, distribution and saturation level in non-alcoholic fatty liver disease in mice, *J.Biophotonics*, 8 (2015) 597–609.
22. K.M.Helal, J.N.Taylor, H.Cahyadi, et al., Raman spectroscopic histology using machine learning for non-alcoholic fatty liver disease, *FEBS Letters*, 593 (18) (2019) 2535-2544.
23. N.Malhotra and M.D.Beaton, Management of non-alcoholic fatty liver disease in 2015, *World J. Hepatol.*, 7 (2015) 2962–2967.
24. E.Hashimoto, K.Tokushige, and J.Ludwig, Diagnosis and classification of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis: Current concepts and remaining challenges, *Hepatol. Res.*, 45 (2015) 20-28.
25. R.Achanta, A.Shaji, K.Smith, et al., SLIC superpixels compared to state-of-the-art superpixel methods, *IEEE Trans. Pattern Anal. Mach. Intell.*, 34 (2012) 2274–2282.
26. C.A.Lieber and A.Mahadevan-Jansen, Automated method for subtraction of fluorescence from biological Raman spectra, *Appl. Spectrosc.*, 57 (2003) 1363-1367.
27. A.F.Palonpon, M.Sodeoka and K.Fujita, Molecular imaging of live cells by Raman microscopy, *Curr. Opin. Chem. Biol.*, 17 (2013) 708-715.
28. H.A.Khalifa and N.M.Alharbi, Solving multi-objective optimization problems through unified approach, *Annals of Pure and Applied Mathematics*, 14 (2) (2017) 327-336.
29. D.K.Biswas and S.C.Panja, Advanced optimization technique, *Annals of Pure and Applied Mathematics*, 5 (1) (2013) 82-89.
30. K.Bharathi and C.Vijaylakshmi, Optimization for flexible job shop scheduling by evolutionary representation, *Annals of Pure and Applied Mathematics*, 7 (1) (2014) 29-34.
31. R.E.Blahut, Computation of channel capacity and rate-distortion functions, *IEEE Trans. Inf. Theory*, 18 (4) (1972) 460.

Khalifa Mohammad Helal

32. S.Arimoto, An algorithm for computing the capacity of arbitrary discrete memory less channels, *IEEE Trans. Inf. Theory*, 18(1) (1972) 14.