BLIND SOURCE SEPARATION : APPLICATION TO MICROORGANISM RAMAN SPECTRA

Regis Huez, Eric Perrin*, G. D. Sockalingum** and Michel Manfait***
Laboratoire d'Automatique et de Microélectronique*
Unité MéDIAN, CNRS FRE2141, UFR de Pharmacie**
Université de Reims Champagne-Ardenne
Campus du Moulin de la Housse
B.P. 1039, 51687 REIMS Cedex 2, FRANCE
Tel: +33.(0)3.26.91.82.16, Fax : +33.(0)3.26.91.31.06
e-mail: regis.huez, eric.perrin, ganesh.sockalingum, michel.manfait@univ-reims.fr

ABSTRACT

We present in this paper a Blind Source Separation (BSS) application to the separation and identification of microorganism Raman spectra. In the study of microorganisms by Raman spectroscopy, the major difficulty is to dissociate the Raman spectrum (RS) of the bacteria from that of the solid culture medium, on which they grow. This is classically achieved by measuring the pure solid culture medium RS and subtracting its exact value from the bacteria and medium RS. The major advantage of the BSS techniques is the extraction of the bacterial RS without prior measurement from the pure solid culture medium, which permits to be independent of medium variations.

1 INTRODUCTION

The purpose of this paper is to identify the Raman spectral signature of a microorganism growing on a solid culture medium. Raman spectroscopy is a well adapted technique to the study of biological samples *in situ*. Here, live bacteria can be studied in the form of intact microcolonies growing on a solid medium. However, the major difficulty with this approach is the removal of the medium influence (often as a fluorescence background) on the spectrum of the bacteria.

Some methods have been developed to deal with this problem but most of them do not fully succeed in subtracting the exact contribution of the culture medium in the total signal.

The first method consists in measuring the pure solid medium, and subtracting this signal from the bacterial RS. The validation of the obtained solution is made by an estimation of the bacteria RS which is nearly flat except for some representative peaks corresponding to the biological species present in the bacterial cell.

A recently developed method is based on mathematical algebra [1]. Briefly, it consists in performing an orthogonal transformation of the bacteria and medium RS using the only medium RS as one of the second basis. This technique leads to a rather good determination of the two RS (medium and bacteria) but still requires the measurement of medium spectra.

The presented approach is clearly different; it does not consider the orthogonality of the two signals as the principal feature. It mainly involves extracting the most possible statistically independent sources from several medium and bacteria RS. The latter sources being considered as a good estimation of the bacteria and the medium RS.

The main advantage here is that it does not require an *a* priori knowledge of the medium, i.e., there is no need to measure the RS of the pure medium. It therefore increases the possibilities of a precise and non-subjective biochemical interpretation of bacterial spectra.

2 BACTERIAL STUDY BY RAMAN (MICRO) SPECTROMETRY

2.1 Generalities

Routine microbiological phenotypic identification of clinical samples is largely based on nutritional and biochemical characteristics of microorganisms. However, gold standards based on molecular biological techniques are becoming more and more popular for the identification of microorganisms and the detection of specific antibiotic resistance genes. Although very reliable, these techniques are quite expensive, require highly skilled personnel, and are time-consuming. At present, molecular diagnostics are usually second lines of investigation and are seldom the sole basis for microbial identification. A different and novel approach in microbial characterization is the use of vibrational spectroscopic methods.

The use of Fourier Transform InfraRed (FTIR) spectroscopy for microbial identification and characterization has been gaining acceptance since Helm et al [2], Sockalingum et al. [3] have used FTIR spectroscopy to identify bacteria and characterize antimicrobial resistance in such systems, indicating the high molecular information content of this technique.

Raman (micro)spectroscopy, which is complementary to FTIR, is also gaining importance since it offers the advantage of performing measurements directly on-plate without perturbing the colony growth. Furthermore, it has confocal properties and offers high spatial resolution (<2 μm), which makes it suitable for studying young microcolonies. In our study, we have used a Raman microspectrometer equipped with a He:Ne laser excitation at 632.8 nm. The signal was collected on a CCD detector.

The Raman effect corresponds to the inelastic scattering after interaction of a monochromatic light with a

molecular system. The frequency shifts, measured with respect to the incident radiation, observed in the scattering spectrum of the studied system, are related directly to the molecular vibrations, i.e. to the molecular structure, composition, and interactions in the system. The technique being non-destructive, it is therefore well adapted to biological samples. The use of a confocal signal detection scheme enables Raman spectroscopic measurements of very small sample volumes (a few μm^3).

The bacteria are studied here in the form of a microcolony that grows on a solid medium (figure 1). We illustrate here the example of *Staphylococcus aureus* bacteria on Mueller Hinton solid culture medium.

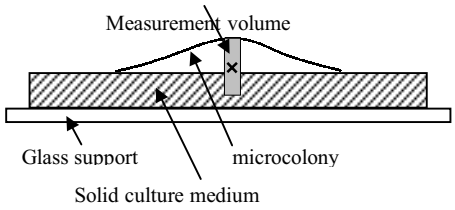

Figure 1 : Measurement specifications

As the measurement volume integrates the microcolony and the medium, the major obstacle of this approach is the presence of signal contributions from the underlying culture medium in the bacterial spectrum. This problem is due to the depth of penetration of the laser light. Because the medium signal contribution is neither negligible nor constant it will interfere with bacterial identification.

2.2 Microcolony study

To determine the bacterial spectrum, we realised a time-dependent set of measurements of the microcolony starting at 6h culture time, the signal being only significant when the colony has reached a certain size.

Depending on the species and the culture times, the microcolony diameter can vary from a few tens of microns (6h) to several hundreds of microns (15h). Its thickness at the centre changes from a few to several tens of microns.

Figure 2 schematically represents a growing colony, showing a population organisation in stratified layers corresponding to old, middle-aged and young bacterial cells. The crosses represent the measurements points. After 6h growth, the microcolony is still quite thin and measurements other than at the centre are much hindered by parasitic fluorescence.

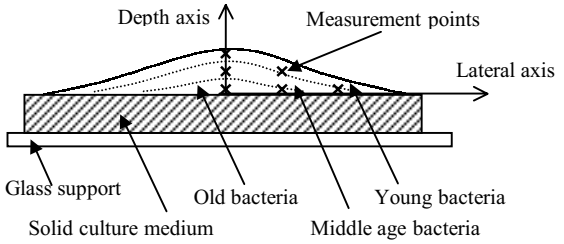

Figure 2 : Microcolony growing and measurement points

To obtain the estimated bacterial spectrum, a possibility is to subtract the medium spectrum (previously measured) from the *in situ* bacterial spectrum. This solution presents three non negligible disadvantages. The subtraction of the exact value is not trivial. Also, the medium can evolve during the measurement. And finally, bacteria-medium interaction can appear and modify the medium spectrum.
Using an estimated spectrum of the medium when the microcolony is present would lead to a better estimation of the RS of the bacteria. In this way the BSS approach can bring a solution.

3 BSS APPLICATION TORAMAN SPECTRUM

3.1 Blind Source Separation

The BSS techniques [4] consist in separating several statistically independent signals named sources from several sensor signals. Three pre-requisites are necessary to apply BSS: need of statistically independent sources, availability of a number of sensors equal or superior to the number of sources, and the existence of a mixing matrix between the sources and the sensors.

The experimental application of BSS techniques is not obvious, it requires an adaptation of the device and a particular definition of the sources, this kind of work has already been done on pyroelectric [5,6] and eddy current sensors [7].

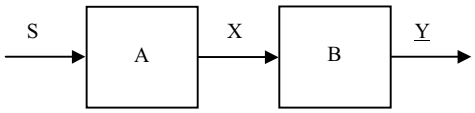

Figure 3 : Principle of BSS

The principle of the BSS is described in figure 3. In the case of two statistically independent sources and two sensors,
the unknown source signals are :

$$S = [s_1\ s_2]^T, \quad (1)$$

and the observed mixed signals perceived by the sensors are :

$$X = [x_1\ x_2]^T. \quad (2)$$

They fit the linear model given below :

$$X = A.S + N, \quad (3)$$

where A is the unknown mixture matrix and N is the additive noise on the sensors. It is always possible to introduce a constant α such as :

$$X = (A/\alpha).(\alpha.S) + N, \quad (4)$$

This formulation points out the uncertainties in the power of the sources noted S. The blind source separation consists of determining a separation matrix $\underline{B}$ (the underlined terms correspond to estimations) such as :

$$\underline{B}.A = I. \quad (5)$$

which leads to:

$$\underline{Y} = \underline{B}.X = \underline{B}.A.S + \underline{B}.N, \quad (6)$$

and finally, we obtained :

$$\underline{Y} = \Lambda.\Pi.S + \underline{B}.N, \qquad (7)$$

The estimated sources named $\underline{Y}$ are equal to the sources S after multiplication with a permutation matrix Π and a coefficient matrix Λ. The permutation and coefficient matrices come from the estimation of $\underline{B}$.

In the case of white noises, the additive sensor noise N is modified by the separation matrix $\underline{B}$ but does not affect the determination of $\underline{B}$. In the following, the additive sensor noise will be neglected in the equations, the estimated sources $\underline{Y}$ will be restored with this additive noise.

Usually, the first step of BSS algorithms is a singular decomposition value. Orthogonal sources are obtained. The solution is non unique. A rotation of the sources which optimises the independence may exist. Matrix $\underline{B}$ is estimated assuming that either the sources are statistically independent and a criterion of independence is maximised [8], or they are also temporally correlated and a criterion of non-correlation [9] is introduced. Knowing that non correlation techniques are used as much as possible due to their robustness, An estimation of the mixing matrix can be obtained :

$$\underline{A} = \underline{B}^{-1}. \qquad (8)$$

3.2 BSS Application

In the microcolony study, the considered sources are the medium and the bacteria RS, they are obviously independent. The sensors are the measured RS.

This type of signal is temporally correlated, the use of second order statistics is sufficient. The algorithm SOBI (Second Order Blind Identification) developed by Belouchrani [9] is used. It is a batch mode (offline) algorithm well adapted for this problematic because it has no time constraint. SOBI can be used directly, it requires the observed mixture signals X, the number of sought sources 'n' and the number 'p' of delayed cross correlation matrices which take into account the temporal correlation. There are two sought sources, therefore 'n=2' and a number of 5 delayed cross correlation matrices is sufficient.

The mixing matrix can be obtained by a judicious exploitation of the measurement set.

- Depth measurements

We have noticed that there exists some spectral differences between spectra taken at different depths within the microcolony. In fact, the contribution of medium RS decreases from the bottom to the top, and that of the bacteria increases, due to its growth. The RS of the bacteria-medium couple at different ratios may be used to obtain a mixing matrix.

- Time measurements

At a fixed location, the 6h RS contains a part of the medium RS and a lesser part of the bacterial RS. Due to the bacterial ageing, the 15h RS contains the same part of the medium RS and a greater part of the bacterial RS. The bacterial RS evolution allows to obtain a mixing matrix between the sources and the sensors.

- Lateral measurements

Along the lateral axis, a similar phenomenon takes place, the older bacterial cells being at the centre of the microcolony and the younger ones at the periphery.

4 RESULTS

The depth and time measurements previously described are presented using the SOBI algorithm.

4.1 Medium RS

A preliminary measurement of the medium has been realised (figure 4). The RS are collected in the spectral interval from 400 to 1650 cm^{-1}. The RS intensities are normalised.

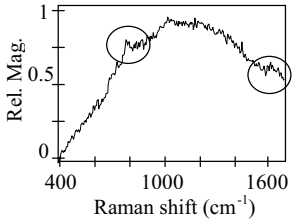

Figure 4 : Medium RS

The medium RS is a broad peak going from 0 to 0.5 magnitude. It is important to notice the two little peaks at 800 and 1600 cm^{-1}. This measurement will allow to confirm whether the estimated medium RS by BSS technique is realistic.

4.2 Depth measurements

Figure 5 shows the obtained results, the upper part concern the measured RS taken 3 and 17 µm above the medium and at 15h, the lower one concerns the estimated RS.

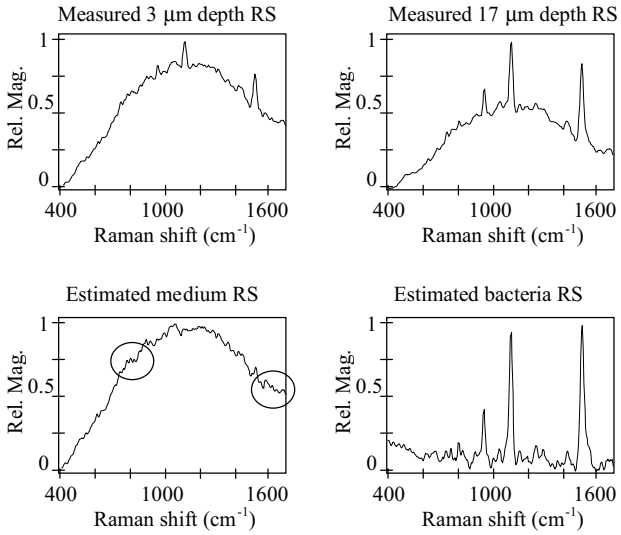

Figure 5 : Depth measurements

The medium is more strongly present on the 3 µm depth measured RS than the 17 µm one, this variation gives us a mixing matrix.

After application of the BSS algorithm, the restored signals, which are the estimated medium and the estimated bacteria RS, are in agreement with predictions. The

estimated medium RS is similar to the preliminary measurement (figure 4) and the estimated bacteria RS is nearly flat except three peaks. The two little peaks located at 800 and 1600 cm^{-1} are not present in the measured RS and therefore are not present in the estimated RS. In this case, the use of the preliminary medium RS would not have been judicious and the BSS application is appropriate.

The determination of the separation matrix $\underline{B}$ allows to find the estimation of the mixing matrix $\underline{A}$.

The mixing is defined by the equation (6) :

$$\underline{Y}=\underline{B}.X \Rightarrow \underline{AY}=\underline{A}.\underline{B}.X$$

by using equation (8) :

$$X = \underline{A}.\underline{Y} \Leftrightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 1.9 \\ 1 & 6.6 \end{bmatrix} . \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

The vectors x_1, x_2 , $\underline{y_1}$ and $\underline{y_2}$ are respectively :

the measured 3 µm and 17 µm depth RS, the estimated medium and the estimated bacteria RS.

The analysis of the estimated mixing matrix coefficients shows that the bacteria influence on the measured RS is three times more important at 17 µm than at 3 µm one. These coefficients agree with the notion of varying ratios previously described.

4.3 Time measurements

Figure 6 shows the results obtained with time measurements. The 6h measured RS is similar to the preliminary medium RS.

After application of the BSS algorithm, the restored signals are in agreement with the predictions.

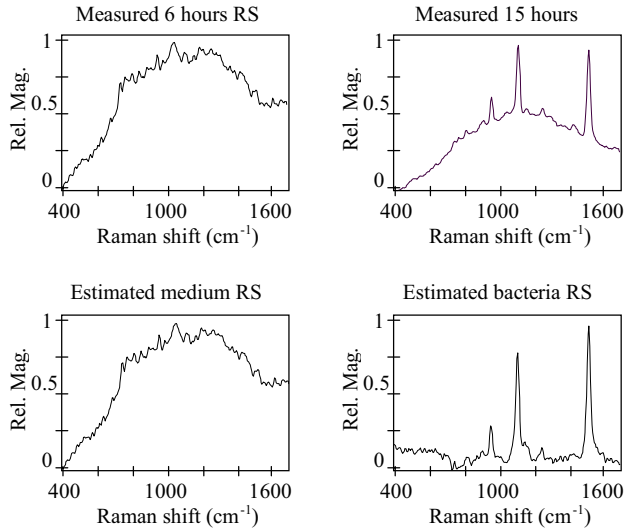

Figure 6 : Time measurements

An analysis of the estimated mixing matrix $\underline{A}$ can give more precisions.

$$X = \underline{A}.\underline{Y} \Leftrightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0.004 \\ 1 & 3 \end{bmatrix} . \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

The estimated medium RS is equal to the preliminary medium RS (x_1=y_1), meaning that 6h growth the bacteria are not visible on the RS. As previously described, the estimated medium RS is different from the preliminary medium RS, principally at 1600 cm^{-1}, where the little peak is absent. Using of this preliminary measurement our approach would not have been exact, thus demonstrating the performances of BSS application.

5 CONCLUSION

We have presented a BBS method applied to RS for the characterization and identification of bacteria and medium spectra. The results clearly demonstrate the possibility to separate the exact spectrum of the bacteria from the medium spectrum. This first result is of a great importance since it permits to be free from the variations of the ratio and the composition of the solid medium. Further work will concern the separation of three sources in view of dissociating the bacteria, the medium, and the bacterial spectral evolution when in interaction with an antibiotic.

REFERENCES

[1] K. Maquelin and al., “Raman Spectroscopic Method for Identification of Clinically Relevant Micro-organisms Growing on Solid Culture Medium“, Analytical Chemistry, Vol. 72, No. 1, pp. 12-19, January 1, 2000.

[2] D. Helm, H. Labichinski, D. Naumann, Microbiol. Methods, Vol. 14, pp. 127-142, 1991.

[3] G. D. Sockalingum and al., M. Biochem. Biophys. Res. Commun., Vol. 232, pp. 240-246, 1997.

[4] J. L. Lacoume, P. O. Amblard, P. Comon, Statistiques d’ordre supérieur pour le traitement du signal, Ed. Masson, Paris, France, 1997.

[5] R. Huez, D. Nuzillard, A. Billat, “Denoising using Blind Source Separation for pyroelectric sensors“, EURASIP JASP, Vol. 1, pp. 53-65, 2001.

[6] R. Huez, D. Nuzillard, A. Billat, De-noising of experimentals signals from pyroelectric sensors by a source separation method, In the Proceedings of the EUSIPCO’98 Conference, Vol. 4, pp. 2073-2076, Rhodes, Greece, 1998.

[7] R. Huez, F. Belloir and A. Billat, “Blind Source Separation Application to Smart Eddy Current Sensor for Metallic Tag Recognition“, EUSIPCO’2000 Conference, Tampere, Finland, Sept. 2000.

[8] J. F. Cardoso, A. Souloumiac, “Blind Beamforming for non Gaussian signals”, IEE Proceedings-F, Vol. 140, no 6, pp. 362-370, 1994.

[9] A. Belouchrani, K. Abed Meraim, J.F. Cardoso, E. Moulines, A blind source separation technique using the second order statistics, IEEE Transaction on Signal Processing, vol. 45, no 2, pp. 434-443, february 1997.