Wavelet Transform for NIR Spectral Data Mining: Bacteria Identification

Roumiana Tsenkova*, Hovagim Bakardjian**, Yuji Fujita*, Stefka Atanassova*, Yukihiro Ozaki*** and Kiyohiko Toyoda*

*Department of Environmental Information and Bio-production Engineering,
Faculty of Agriculture, Kobe University, Rokkodai 1-1, Nada-ku, Kobe 657 - 8501, Japan

**National Institute of Bioscience and Human-Technology
Human Informatics Group, 1-1, Higashi, Tsukuba Science City, 305-8566, Japan

***Department of Chemistry, School of Science, Kwansei-Gakuin University, Uegahara, Nishinomiya 662-8501, Japan

INTRODUCTION

Mastitis is a widely spread bacterial infection of the mammary gland in Dairy. Most common pathogens, causing mastitis are *Staphylococcus aureus (Sta aureus)* and *Streptococcus agalactiae (Str agalactiae).* **It causes**

deterioration of milk quality and about 25% reproduction losses.

Near Infrared Spectroscopy has been established as a method for analyzing the ingredients in dairy products and has been demonstrated as a sensitive method for mastitis diagnosis (1,2).

Wavelet transform (WT) gives local frequency information and involves studying certain phenomena at different scales (3). WT has been used (4,5) as a tool for further NIR spectral data mining and early mastitis diagnosis based on raw milk spectra.

The purpose of this investigation was to localize the changes in milk spectra caused by main types of bacteria that cause mastitis, using NIRS and WT.

MATERIAL AND METHODS

Samples: Milk samples with no bacteria and milk samples with different concentration of *Sta aureus* and *Str agalactiae* (concentration from 10^1 to 10^{10} CFU/ml), repectively.

Near Infrared Spectra: NIRSystem 6500 (FOSS NIRSystem, Silver Spring, MD, USA), transmittance, 1mm sample thickness, 1100 to 2500nm.

Spectral Data Treatment:

- Second Derivative, based on Savitzki-Golay polinomial filter (Pirouette 2.0, Infometrics, Inc., Woodinville, WA, USA)
- Wavelet Transform (Mathematica's Wavelet Explorer software package (Wolfram Research Inc., Champaign, Illinois, USA): Exponentially decaying discrete Battle-Lemarie 12-tap wavelet [2]. The wavelet frequency bands (scales) were defined as multiresolution scale from 1 to 6, where:

VLF (very low frequency) scales 1 and 2 were in the range of 128...256 nm and 64...128nm, respectively;

LF (low frequency) scales 3 and 4 were in the range of 32...64 nm and 16...32 nm, respectively;

HF (high frequency) scale 5 and 6 were in the range of 8...16 nm and 4...8 nm, respectively.

RESULTS AND DISCUSSION

The spectra of all samples were transformed as second derivative and examined. Slight differences (order of 10^{-5}) in second derivative spectra of milk with no bacteria and milk with *Sta aureus* and *Str agalactiae* were found at 1158nm, from 1610 to 1750nm, at 1780nm, between 1960-2050nm, at 2174nm and between 2210-2400nm, for *Sta aureus,* and around 1988nm, 2100nm, and 2374nm, for *Str agalactiae*, but they were even smaller compared to differences, observed with *Sta aureus*.

The discrete wavelet multiresolution decomposition that we used allowed even better bacteria identification based on NIR spectra. While the second derivative feature extraction technique could be described as a single-band high-pass filtering operation with output depending upon sampling rate and applied derivation step, our method provides elaborate multi-band spectral information for feature extraction. The efficiency of the multiresolution wavelet representation is due to its pyramidal sampling of time-frequency space, producing approximation and detail output signals at every level (termed 'scale') of the multiresolution pyramid. All the spectral differences found by second derivative transformation were observed by WT, too (Fig. 1,2). New information was obtained for both bacteria as follow: for *Sta aureus* - in the areas of 1100...1200nm and 1800...1900nm (Fig. 3, 4) and for *Str agalactiae* - in the areas of 2200...2300nm and 2400...2450nm (Fig.5, 6).

CONCLUSIONS

Changes in milk spectrum, elevated by 2 main bacteria types causing mastitis: *Staphilococcus aureus* and *Streptococcus agalactiae* were found and localized by NIRS.

It has been found that wavelet transform (WT) could be used as a powerful tool for NIR spectral data mining. In this study WT has been utilized as a "microscope" to provide more detailed information from a single NIR spectrum, when compare to second derivative transformation. It could be used for further investigation not only on mastitis diagnosis, but also for bacteria identification.

REFERENCE

1. Tsenkova, R. N., K.Iordanov and Y.Shinde. Near Infrared Spectroscopy for Evaluating Milk Quality. In: Prospects for Automatic Milking, edited by A.H.Ipema et al., Pudoc Scientific Publishers, Wageningen, Netherlands, p.185-193, 1992.
2. Tsenkova, R. N.: Near Infrared (NIR) Spectroscopy for Mastitis Diagnosis. In Proc. of the 11th National Symposium on Nondestructive Measurement, Tsukuba, p.68-74, Japan, 1995.
3. S.G. Mallat. A Theory for multiresolution signal decomposition: The wavelet representation. IEEE Trans. PAMI, 11, 674-693 (1989).
4. R. N. Tsenkova, Y. Fujita, T. Fukushima, K. Toyoda, S. Atanassova: Wavelet and NIRS for Single Spectrum Mastitis Diagnosis. 9th International Diffuse Reflectance Conference, Chambersburg, Pennsylvania, U.S.A, 1998.
5. R. Tsenkova, S. Atanassova, Y. Fujita, K. Toyoda: Wavelet Transform for NIR Spectral Data Mining, 9th International Conference on Near Infrared Spectroscopy, 12-19 June, Verona, 1999.

ACNOWLEDGMENTS

This work has been supported by the Program for Promotion of Basic Research Activities for Innovative Biosciences, financed by the Japanese Ministry of Agriculture.

The authors also wish to thank Dr. T. Tahara for co-operation in carrying out the experiments and GL Science Ltd., for providing the Pirouette 2.6 software.