Conference materials
UDC 637.073
DOI: https://doi.org/10.18721/JPM.163.261

Application of optical methods for quality control of dairy products using data mining

M.Yu. Zanevskaya[1]✉, M.S. Mazing[2], A.Yu. Zaitceva[2], V.V. Romanova[2]

[1] St. Petersburg State University of Aerospace Instrumentation, St. Petersburg, Russia;

[2] Institute for Analytical Instrumentation RAS, St. Petersburg, Russia

✉ mnevskaya1@gmail.com

Abstract. A method has been developed for express-assessment of the quality of dairy products according to the indicators of optical sensors in the visible and near-IR wavelength range. The use of modern machine learning methods, in particular the principal component method, made it possible to identify groups of samples similar in their properties and determine whether products belong to an industrial or piece manufacturing method. The technique allows you to designate a group of 'references', deviations from it, and is an inexpensive express method for controlling the quality of food products.

Keywords: IR spectroscopy, spectrum analyzer, dairy products

Citation: Zanevskaya M.Yu., Mazing M.S., Zaitceva A.Yu., Romanova V.V., Application of optical methods for quality control of dairy products using data mining, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 16 (3.2) (2023) 349–353. DOI: https://doi.org/10.18721/JPM.163.261

Материалы конференции
УДК 637.073
DOI: https://doi.org/10.18721/JPM.163.261

Применение оптических методов для контроля качества молочной продукции с использованием интеллектуального анализа данных

М.Ю. Заневская[1]✉, М.С. Мазинг[2], А.Ю. Зайцева[2], В.В. Романова[2]

[1] Санкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, Россия;

[2] Институт аналитического приборостроения РАН, Санкт-Петербург, Россия

✉ mnevskaya1@gmail.com

Аннотация. Разработана методика экспресс-оценки качества молочной продукции по показателям оптических датчиков видимого и ближнего ИК диапазона длин волн. Применение современных методов машинного обучения, в частости метода главных компонент, позволило выделить группы схожих по своим свойствам образцов и определить принадлежность продукции к промышленному или штучному методу изготовления. Методика позволяет обозначить группу «эталонов», отклонения от нее и является недорогим и мгновенным методом контроля качества пищевой продукции.

Ключевые слова: ИК-спектроскопия, анализатор спектров, молочная продукция

Ссылка при цитировании: Заневская М.Ю., Мазинг М.С., Зайцева А.Ю., Романова В.В. Применение оптических методов для контроля качества молочной продукции с использованием интеллектуального анализа данных // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2023. Т. 3.2 № .16. С. 349–353. DOI: https://doi.org/10.18721/JPM.163.261

Introduction

Dairy products are the basis of the food basket of the typical consumer not only in the Russian Federation, but around the world. Modern food production requires accurate and rapid methods of controlling the properties and composition indicators of the products [1]. At present, there is a growing tendency in the world to accurately determine the quality of dairy products, since one of the most common practices is the addition of some substance to increase the volume of raw materials [2]. Due to the intensive development of instrumental analysis methods, including infrared spectroscopy, it is increasingly becoming a question not only of determining the main components, such as lactose, fat, protein, dry matter, but also of simultaneously identifying various adulterants [3, 4]. Due to the large volumes of processed dairy raw materials in production, fast and effective methods of monitoring its composition are required [4, 5]. The following methods are used for qualitative and quantitative evaluation of individual components of the composition: photometric, ultrasonic, conductometric, infrared spectroscopy and others [5]. IR-spectroscopy is widespread due to the fact that for a short period of time allows to register the spectrum of most significant chemical compounds in the composition of the studied sample [6, 7].

Thus, the aim of the work was to create a technique for express quality control of dairy products and detection of adulteration in them on the basis of infrared spectroscopy methods with the possibility of training and recognition of images of the studied raw materials.

Materials and Methods

The study of dairy products samples was carried out using the developed hardware and equipment complex consisting of two modules: an optoelectronic module and a computer-information module. Scheme of work and process of research of dairy products are shown in Fig.1. The opto-electronic module is a modern multichannel analyzer of spectra which has an array of 18 fast photosensitive elements, working at wavelengths from 410 to 940 nm, and 3 sources of radiation. These technical characteristics allow the device to obtain spectral data of a product sample in less than 10 seconds [8]. A 500 mA/h battery is built into the system, allowing 5 hours of operation without recharging. There are operating modes of both pulse illumination of the sample and continuous. The data are transmitted via wireless personal networks to a PC. Computation and information module is represented by a complex of statistical algorithms of multidimensional data processing and analysis with the use of machine learning methods (method of main components). The result of the analysis is the assignment of the sample to one of the established similarity groups. Fig. 1 shows schematic diagram of the hardware-software complex and the process of dairy products research.

Two groups of samples of dairy products were studied. The first group included 10 samples of milk of industrial production, the second group included 8 samples of milk from individual farms. The milk samples were produced and purchased at the same time. The samples were stored under the same conditions. Each of them was poured into a specially made black container. Measurements were taken at room temperature.

Results and Discussion

The experimental data obtained during the study were an array of numerical values of each of the 18 elements of the optoelectronic unit. In order to qualitatively analyze the composition of dairy product samples and determine the presence of impurities and adulterations, the spectra of reflected and backscattered radiation in the visible and near-infrared wavelength range were examined (Fig.2).

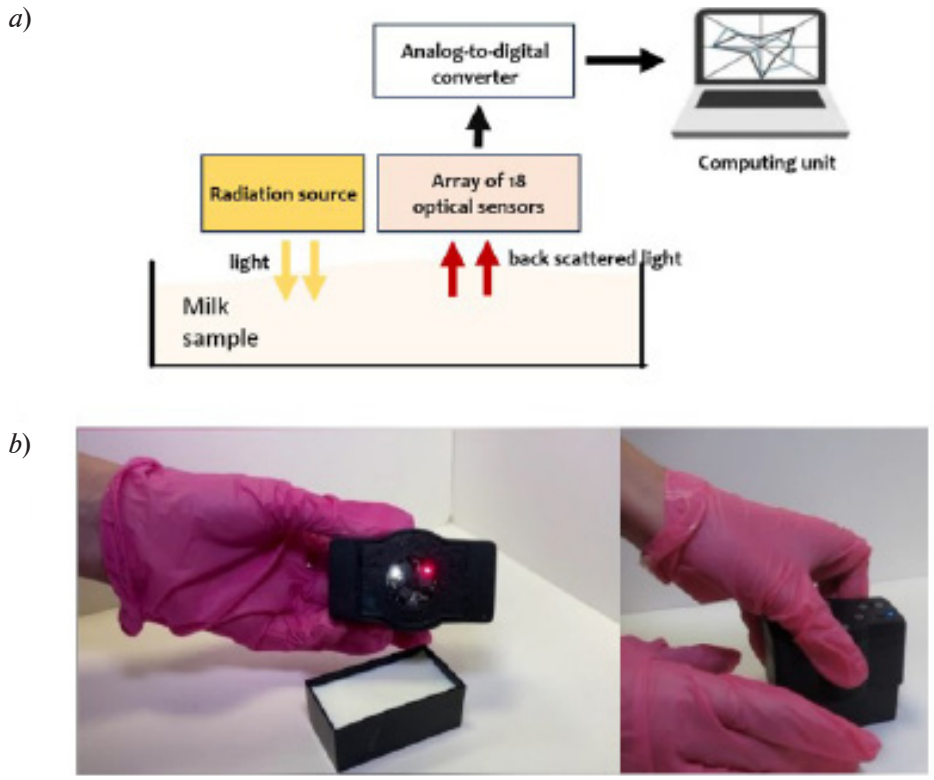

Fig. 1. Schematic diagram of the hardware-software complex (*a*) and the process of dairy products research (*b*)

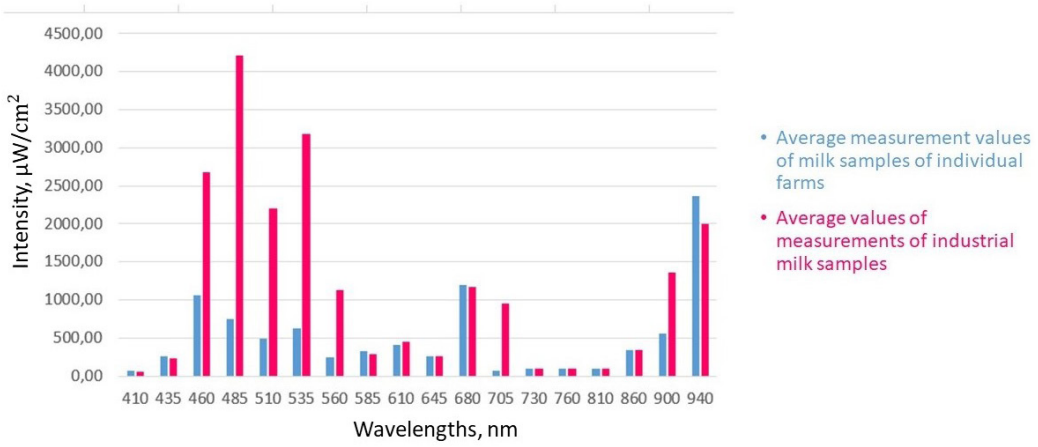

Fig. 2. Spectrum of reflected and backscattered radiation of dairy product samples in the visible and near-infrared wavelength range

During the analysis of the experimental data the following trends are traced: the peaks of the radiation intensity of industrial milk samples are noticeable in the visible range of the spectrum, at wavelengths of 460–535 nm. In contrast, the maximum intensities of milk samples from individual farms are located in the near-infrared wavelength range, at wavelengths of 900–940 nm. This indicates different composition of samples, the presence of higher intensity peaks among industrial dairy products may indicate the presence of impurities in the samples. Intelligent multivariate data processing methods were used to examine the structure of the data and to find relationships between samples and variables, in particular, the principal component method (PCA). PCA is a multivariate statistical analysis aimed at reducing the dimensionality of the data set. It allows us to

identify new variables (principal components) that most accurately reflect the correlation between the original variables and explain the largest proportion of variance in the data [9, 10]. The result of the analysis is a graph, which allows you to represent the relationship between the data. In the Cartesian coordinate system, each point is one of the samples. To determine the optimal number of principal components, a scree plot (Kettel's method) was used, according to which it was necessary and sufficient to leave three components [11]. The following trends are traced: the formation of a cluster of points - samples of industrial milk, which indicates their similarity to each other and a large scatter on the graph of points - samples of farm milk (Fig. 3.).

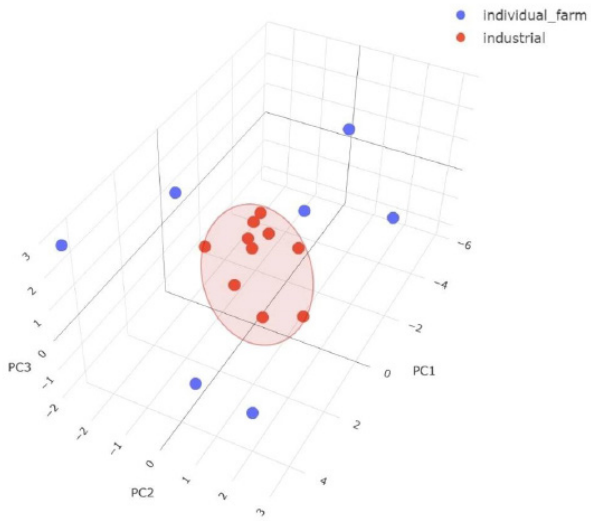

Fig. 3. Result of visualization of multivariate data obtained using optical sensors in the space of the three principal components, where samples marked in blue are dairy products of individual farms; in red - industrial dairy products

The result of intellectual processing of the data obtained with the help of optical sensors were formed groups of product samples. According to the belonging of a particular sample to one of the groups, it is possible to judge about the difference in the composition of the samples, adulteration and the possible presence of impurities. It has been established that the composition and quality of dairy product samples can be assessed using optical spectroscopy methods with subsequent intelligent mathematical processing of the data obtained (including the use of methods of multivariate statistical analysis of data).

Conclusion

Thus, the results obtained testify to the effectiveness of the developed technique based on the optical device - multi-channel spectra analyzer. The presented technique can be used as one of the highly effective methods of controlling the composition and quality of dairy products, which has such advantages as availability for wide application, sufficiently high expressiveness and efficiency, low requirements for the qualification of the operator, etc.

Acknowledgments

The work was supported by the Ministry of Science and Higher Education of the Russian Federation (project 075-15-2021-1057).

REFERENCES

1. **Gastélum-Barrios A., et al.,** Optical methods based on ultraviolet, visible, and nearinfrared spectra to estimate fat and protein in raw milk: A review, Sensors. 20 (12) (2020) 3356.

2. **Taterka H., Zamora A., Castillo M.,** Developing an optical backscatter method for determining casein micelle particle size in heated milk. Food Research International. (161) (2022) 111745.

3. **Rojas Lypez M.,** Analysis of milk adultered with whey powder and liquid by infrared infrared spectroscopy. (2012).

4. **Lypez H., Oropeza I., Betancourt C.,** Determination of calcium, magnesium and potassium concentration in liquid milk of three commercial brands using the atomic spectroscopy technique. Research Journal. (2017).

5. **Calderyn de la Barca A.M., et al.,** Evaluation of human milk consumption by dilution with deuterium and detection by infrared spectroscopy //Perinatol. reprod. hum. (1998).

6. **Holroyd S.E.,** The use of near infrared spectroscopy on milk and milk products. Journal of Near Infrared Spectroscopy. (2013).

7. **Pu Y., et al.,** Recent advances in portable and handheld NIR spectrometers and applications in milk, cheese and dairy powders. Foods. 10 (10) 2021 2377.

8. **Zaitceva A.Y.,** Sensor Intelligent Systems for Monitoring the Oxygen Status of Human Tissues Under Functional Loads //Journal of Pharmaceutical Negative Results. 13 (1) (2022) 6–13.

9. **Pearson K.,** LIII. On lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin philosophical magazine and journal of science. 2 (11) (1901) 559–572.

10. **Hasan B.M.S., Abdulazeez A.M.,** A review of principal component analysis algorithm for dimensionality reduction. Journal of Soft Computing and Data Mining. 2 (1) (2021) 20–30.

11. **Del Giudice M.,** The Square-Root Scree Plot: A Simple Improvement to a Classic Display. (2022).

THE AUTHORS

ZANEVSKAYA Maria Yu.
mnevskaya1@gmail.com

ZAITCEVA Anna Yu.
anna@da-24.ru
ORCID: 0000-0003-0334-2770

MAZING Mariia S.
mazmari@mail.ru

ROMANOVA Veronika V.
venjo@mail.ru
ORCID: 0009-0001-6587-787X

Received 24.07.2023. Approved after reviewing 16.08.2023. Accepted 17.08.2023.