Physicochem. Probl. Miner. Process. 51(1), 2015, 163−172

www.minproc.pwr.wroc.pl/journal/

Physicochemical Problems of Mineral Processing

ISSN 1643-1049 (print)
ISSN 2084-4735 (online)

Received April 20, 2013; reviewed; accepted March 22, 2014

ESTIMATION OF COPPER CONCENTRATE GRADE BASED ON COLOR FEATURES AND LEAST-SQUARES SUPPORT VECTOR REGRESSION

Chuancheng REN[*], Jianguo YANG[**], Chuan LIANG[**]

[*] College of Information Management, Dezhou University, Dezhou, 250323, China, renchuancheng@163.com
[**] National Engineering Research Center of Coal Preparation and Purification, China University of Mining and Technology, Xuzhou, 221116, China

Abstract: In this paper, a new method based on color features of microscopic image and least-squares support vector regression model (LS-SVR) is proposed for indirect measurement of copper concentrate grade. Red, green and blue (RGB), hue and color vector angle were extracted from color microscopic images of a copper concentrate sample and selected for the comparison. Three different estimation models based on LS-SVR were developed using RGB, hue, and color vector angle, respectively. A comparison of three models was carried out through a validation test. The best model was obtained for the hue giving a running time of 30.243 ms, root mean square error of 0.8644 and correlation coefficient value of 0.9997. The results indicated that the copper concentrate grade could be estimated by the LS-SVR model using the hue as input parameter with a satisfactory accuracy.

Keywords: *concentrate grade, copper concentrate, LS-SVR, color features, microscopic image*

Introduction

In copper industry, the content of useful component (grade) is one of primary quality parameters in the final concentrate and a key control parameter in a flotation circuit. The grade determines selling price, so industry increases the grade of final product for the greatest interest (Kaartinen et al., 2006; Malewski and Krzeminska, 2012). The copper grade is the percentage of copper mineral captured into the final flotation concentrate. Usually, the copper grade could be determined by direct and indirect measurements. The direct measurements mainly involves an X-ray fluorescence measurement and chemical assay methods. The X-ray fluorescence measurement is performed in a few copper mineral industries due to high cost (Haavisto et al., 2008). However, the chemical assay method is a time consuming process. Therefore, a faster

http://dx.doi.org/10.5277/ppmp150115

and cheaper indirect measurement is desirable to achieve the estimation of copper grade.

A machine vision system has potential applications for the concentrate grade evaluation, which is highly dependent on the features of froth and concentrate particle appearance. This is due to the fact that flotation operators always roughly evaluate the concentrate grade through the characteristics of froth surface or concentrate particle in order to monitor and improve flotation processes (Tasdemir and Kowalczuk, 2014). Currently, the concentrate grade measurement can be improved by means of machine vision in mineral industry. The final concentrate image features can be used to estimate various parameters, i.e. grade, recovery and other process variables. In particular, if the concentrate grade is measured by an inexpensive machine vision, the flotation performance could be improved through a feedback control. Consequently, a sufficiently accurate image analysis method can interpret the correlation between the characteristic of concentrate particle and grade.

In the past, several methods have already been utilized in order to determine the relationship between the concentrate grade and image analysis. Oestreich et al. (1995) applied a color video camera to characterize the color of dry material, slurry, and flotation froth mixtures of chalcopyrite and molybdenite. In addition, they built a strong correlation between the color vector angle and mineral composition. The color vector angle is used to characterize the color of mineral mixtures. Although the estimated expression was mentioned, there was a little indication of data to validate it directly. A regression method is one of the most prevalent methods for estimating the concentrate grade. Bonifazi et el. (2000) adopted a multiple regression method for evaluating Pb, Zn, Cu and MgO froth grades, which used fractal parameters derived from froth, along with froth color parameters (i.e. RGB, HSI and HSV spectra statistical parameters) as input variables. Additionally, Bonifazi et al. (2002) adopted a simple second-order polynomial approach with the parameters extracted from a 3D model of froth surface and image analysis (color and morphological features) to forecast the mineral grade. When input parameters were mutually independent, Morar et al. (2005) applied a multivariate linear regression model to predict the concentrate grade based on color, velocity and stability information of the froth. In this method, the best prediction model was obtained with a combination of color and stability information. Forbes (2007) gave a description of relationships among froth velocity, bubble size, froth class and concentrate grade by using both linear and non-linear regression model respectively. When the input parameters were highly dependent on each other, Haavisto et al. (2006) applied a partial least squares (PLS) method to suit well the problem. It formed a mapping from input to output variables using a lower dimensional latent basis, and built a set of latent variables that interpreted input and output variables. The PLS method using the measured forth spectrum as input variables, was calculated to estimate the grade of zinc flotation circuit concentrate.

Recently, an artificial neural networks (ANN) method has been increasingly applied in grade estimation. Based on the ANN method, Marais (2010) set up a nonlinear model to estimate the platinum froth grade and recovery. In this nonlinear model, the input variables was extracted from froth images based on the laboratory and industrial data. Based on the ANN and multivariate non-linear regression (MNLR) models, Nakhaei et al. (2012) also developed two simulation models to estimate the grade and recovery of a pilot plant flotation column with the input variables, such as the values of chemical reagents dosage, froth height, air and wash water flow rates, etc. They found that the model based on the ANN approach had a better prediction performance than the MNLR method. Furthermore, Nakhaei et al. (2013) adopted a genetic algorithm to optimize the input variables within the ANN, and applied the GANN approach to predict the metallurgical performance (grade and recovery) of the pilot flotation column with the variables, i.e. chemical reagents dosage, froth height, air, wash water flow rates etc. Yong et al. (2012) proposed a concentrate grade prediction model based on the RBF neural network using the froth image characteristics, in which the input variables were optimized the by immune evolution method to improve predictable accuracy.

All of above mentioned methods are applied to a large amount of samples. In addition, the existing methods also include the preprocessing of input variables and demand a large training set. To offset disadvantage, this paper introduces the LS-SVR method (Suykens and Vandewalle, 1999) into the model of concentrate grade estimation. The LS-SVR method, combining the properties of regression method and the advantages of support vector machine, attracted great interests and become a popular tool for model estimation. It shows a surprising effectiveness in estimating the concentrate grade with a small amount of concentrate samples.

In this study, the concentrate grade of copper flotation process was predicted. Firstly, the colors of microscopic images of copper concentrate were obtained, and the usefulness color features were extracted from these images. Secondly, the estimated model for copper concentrate grade was developed by the least-squares support vector regression model (LS-SVR) method using the training set from the color features and chemical assay.

Materials and methods

Characteristics of copper concentrate

In flotation experiments, copper ore taken from Pulang Copper Mine in China is a typical ore of copper-sulfide minerals. The ore contains mainly chalcopyrite, subordinately bornite, and rarely covellite. The associated minerals are pyrrhotite, molybdenite, magnetite, quartz. The series of copper concentrate were obtained through a laboratory flotation cell. The flotation concentrates were dewatered, dried, and sampled for a chemical assay. The results showed that the sample ranged from 8 to 20% of Cu. The sample was characterized with a complex mineral composition and fine dissemination size, which was a mixture of particle size mainly less than 74 μm. Thus, the macroscopic visualization of sample was extremely difficult to obtain.

The minerals of the copper concentrate sample varied in color. The minerals influencing the overall sample color included: chalcopyrite (brassy yellow), bornite (brown to black with a reddish bronze color on freshly broken surfaces or a typical purplish-bluish tarnish), covellite (a deep metallic indigo blue with iridescent yellow and flashing red), pyrrhotite (bronze), magnetite (black iron), quartz (clear quartz/ white/ pink/brown to black). The above-named minerals displayed a range of colors in the concentrate samples, thus the samples had not a homogeneously distributed colors.

Microscopic image acquisition

In image processing, one of the critical steps is the appropriate acquisition of images. A schematic diagram of acquisition system for color microscopic images of copper concentrate is given in Fig. 1. It consists of a digital monocular microscope, ring of LED light source and computer (PC). The digital monocular microscope with a special design for a CCD camera, was an instrument to help viewing very small particles under a magnification. The camera (3.0 MP camera) had a CCD sensor with a resolution of 2048×1536 pixels. The ring of a LED light source, which mounting around the objective could be adjusted to obtain the best illumination intensity. A concentrate sample image was displayed on a PC screen through a USB cable. The magnification of this image acquisition system ranged from 50 to 300.

Random samplings were performed in each copper concentrate sample to obtain a smaller subsample (around 0.6 g). An 1×3 cm subsample with a flat surface was used for the microscopic image acquisition system. Ten microscopic images were captured from different parts of each subsample (Fig. 2).

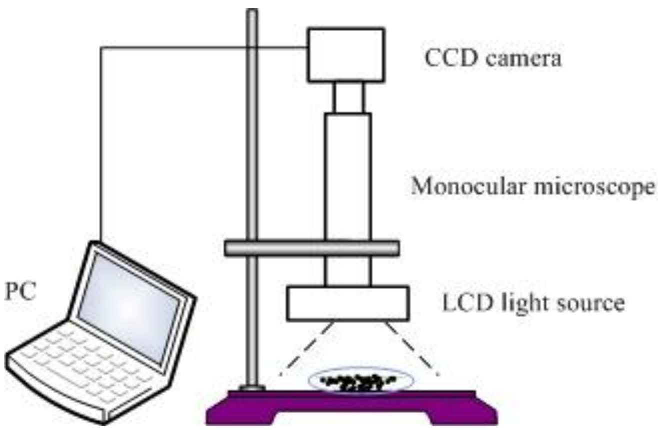
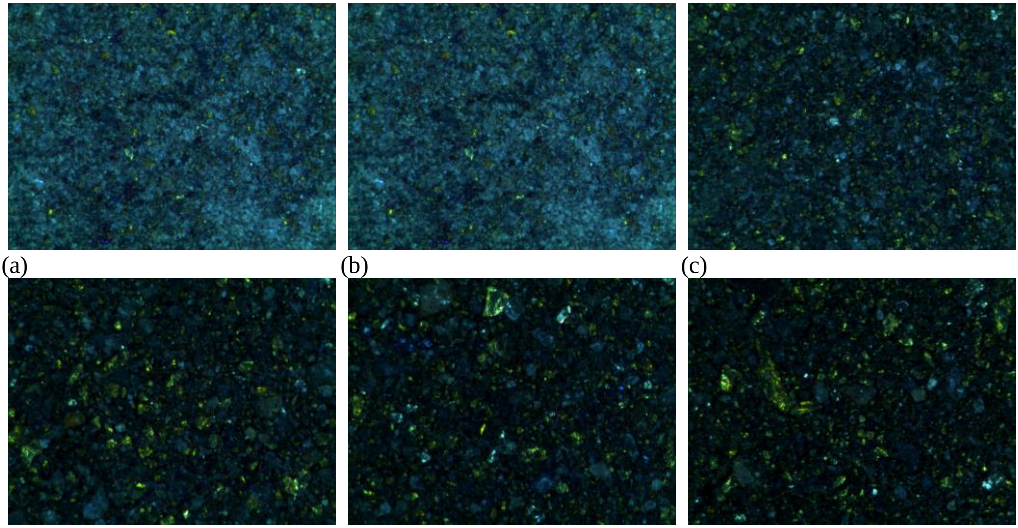


Fig. 1. Schematic diagram of image-acquisition experimental setup

Color features

The image analysis involved extraction of red, green and blue components from microscopic images in the RGB color space, from which the mean values of red, green and blue components were calculated. Other parameters such as the color vector angle and mean value of hue component in the HSV color space were also obtained. These parameters were used to estimate the copper concentrate grade.



(a) (b) (c) (d) (e) (f)

Fig. 2. Example microscopic images from different concentrate grade samples used in study (original magnification × 260): (a) Cu 10.89%, (b) Cu 13.65%, (c) Cu 15.66%, (d) Cu 16.86%, (e) Cu 18.32%, (f) Cu 19.22%

The mean values of red, green and blue components for a concentrate sample are expressed as follows:

$$\begin{cases} R_i = \dfrac{1}{10}\sum_{k=1}^{10}\left(\dfrac{1}{M\times N}\left(\sum_{m=1}^{M}\sum_{n=1}^{N}R_{m,n}\right)\right) \\ G_i = \dfrac{1}{10}\sum_{k=1}^{10}\left(\dfrac{1}{M\times N}\left(\sum_{m=1}^{M}\sum_{n=1}^{N}G_{m,n}\right)\right) \\ B_i = \dfrac{1}{10}\sum_{k=1}^{10}\left(\dfrac{1}{M\times N}\left(\sum_{m=1}^{M}\sum_{n=1}^{N}B_{m,n}\right)\right) \end{cases} \tag{1}$$

where i is an index of all the concentrate samples, and k is a serial number of i-th sample image. $R_{m,n}$, $G_{m,n}$ and $B_{m,n}$ denote a pixel of M×N matrix of red, green and blue components from a color microscopic image, respectively.

The HSV color space is closer to human perception and better color conversion accuracy (Ren and Yang, 2012). The hue component is particularly important, since it refers the spectral composition of color, of which the range is from 0 to 360 degree. The mean value of hue component for i-th sample image can be obtained:

$$H_i = \frac{1}{10}\left(\sum_{k=1}^{10}\frac{1}{M\times N}\left(\sum_{m=1}^{M}\sum_{n=1}^{N}H_{m,n}\right)\right) \tag{2}$$

where $H_{m,n}$ is a value of M×N matrix of hue component in the HSV color space.

The color vector angle (CVA) is insensitive to variations in an illumination component, but sensitive to differences in the hue and saturation, since the luminance (Y) and chromatic components (I, Q) are independently denoted in YIQ color space (Ford and Roberts, 1998). The CVA is defined with (Tkalcic and Tasic, 2003)

$$\begin{cases} I_i = 0.596R_i - 0.274G_i - 0.322B_i \\ Q_i = 0.211R_i - 0.523G_i - 0.312B_i \\ \theta_i = \tan^{-1}(Q_i / I_i) \end{cases} \tag{3}$$

where I_i and Q_i are the mean values of the chromatic components for i-th sample image, respectively. θ_i is the CVA value of i-th sample image.

LS-SVR method

The least-squares support vector regression model (LS-SVR) (Suykens and Vandewalle, 1999; Suykens, 2001) is a nonparametric regression, which is an extension of support vector machines with taking equality instead of inequality constraints and using a squared loss function. The LS-SVR method for function estimation is done in the Kernel-induced feature space, as follows:

$$y = w^T\varphi(x) + b \tag{4}$$

where $x\in R^n$, $y\in R$, w is an unknown coefficient vector, $\varphi(x)$ is a nonlinear map function which maps the input data into a high dimensional feature space, and b is a bias constant. For a given training set $\{(x_i, y_i)\}$, where $i\in N$, and N denotes the number of training data, $x_i\in R^n$ and $y_i\in R$ are i-th input and output data, respectively, the optimization problem with a regularized cost function can be defined as:

$$\begin{cases} \min\limits_{w,e_i} & \frac{1}{2}w^T w + \frac{1}{2}\gamma\sum_{i=1}^{N}e_i^2 \\ s.t. & y_i = w^T\varphi\left(x_i\right) + b + e_i \end{cases} \tag{5}$$

where $\gamma>0$ is a regularization constant which controls the bias-variance trade-off, and e_i is an error term for i-th training data.

The optimization problem (Eq. 5) can be solved using the Lagrangian multipliers, and the resulting LS-SVR method is given by

$$y = \sum_{i=1}^{N} a_i K\left(x, x_i\right) + b \tag{6}$$

where $a_i \in R$ are Lagrange multipliers, x and y are input and output parameters respectively, $K(x, x_i)$ is a certain positive-definite kernel function.

Grade estimation model

The color features of concentrate microscopic images are the key parameters in order to estimate the concentrate grade. Hence, an inherent relationship between the color features and concentrate grade was built by using the LS-SVR method. For grade estimation, 23 groups of sample data were acquired. The 85% groups of dataset were randomly selected as the training and remaining sets. The number of sample such as 3, 6 and 21 were used as the validation tests. The color features and copper concentrate grades measured by the chemical assay are presented in Table 1.

Table 1. Color features and chemical assay for copper concentrate

Samples	The mean values				CVA	Actual Cu%
	Red	Green	Blue	Hue		
1	24.0195	71.8487	90.5993	1.6467	1.102	8.870
2	19.6249	58.8243	69.4159	1.5432	1.1633	10.89
3	15.3936	49.9594	61.9018	1.558	1.1251	11.31
4	15.3363	48.6774	58.834	1.5345	1.1457	11.88
5	15.3134	49.1808	59.1731	1.5376	1.1506	12.31
6	13.6008	45.0216	54.8837	1.5391	1.1412	13.06
7	10.9599	37.8641	44.5042	1.508	1.1764	13.65
8	11.4011	38.7724	44.7765	1.4944	1.1913	14.28
9	15.1702	48.3923	56.635	1.5241	1.175	14.53
10	11.7008	38.825	42.4789	1.4728	1.2388	15.66
11	9.1404	31.6074	33.2607	1.4554	1.2756	16.12
12	9.9737	32.6556	32.0527	1.4259	1.3384	16.68
13	7.4984	27.5487	29.0742	1.4587	1.2738	16.75
14	7.7483	27.446	27.8441	1.4505	1.3081	16.86
15	6.2625	24.1291	25.1321	1.4596	1.286	17.06
16	7.270	26.8296	27.7515	1.4637	1.2919	17.50
17	8.3434	29.6707	31.2064	1.4638	1.2763	17.84
18	9.3754	31.0781	29.7923	1.4202	1.3598	17.97
19	7.0299	25.782	25.4689	1.4522	1.3316	18.32
20	6.7748	25.2065	25.6412	1.467	1.3063	18.41
21	6.5704	23.9987	23.7016	1.4485	1.328	18.62
22	7.1414	25.6158	24.8186	1.4439	1.3488	18.69
23	6.5176	23.8093	22.0975	1.4248	1.3863	19.22

Suppose training set are $\{(R_i, G_i, B_i, y_i)\}$, $\{(H_i, y_i)\}$ and $\{(\theta_i, y_i)\}$, where i is a sample index. With the estimation model from Eq. (6), three models are expressed in the following forms:

$$\begin{cases} y_{RGB} = \sum_{i=1}^{20} \alpha_i K\left((R,G,B)_{new}, ((R_i,G_i,B_i), y_i)\right) + b \\ y_{Hue} = \sum_{i=1}^{20} \alpha_i K\left(H_{new}, (H_i, y_i)\right) + b \\ y_{CVA} = \sum_{i=1}^{20} \alpha_i K\left(\theta_{new}, (\theta_i, y_i)\right) + b \end{cases} \quad (7)$$

where $(R,G,B)_{new}$, H_{new} and θ_{new} are the input parameters, y_{RGB}, y_{Hue} and y_{CVA} are the predicted grades of copper concentrate, respectively.

Results and discussion

In this paper, all proposed estimation grade models were performed using a MATLAB R2010b software. For each prediction model, the radial basis function (Hamers et al., 2002) was selected as the Kernel function, which is defined by $K(x, x_i) = \exp(-\|x - x_i\|_2^2 / \sigma^2)$, where σ means the bandwidth parameter. Parameters σ and γ in Eq. (5) were obtained by an intensive grid search process with the cross-validation approach (An et al., 2007) in order to give high prediction accuracy and stability. The final predicted and actual grade values are presented in Table 2.

In order to evaluate the performances of estimation models based on the LS-SVR, running time, root mean square error (RMSE) and correlation coefficient (R^2) were computed. The statistical results are given in Table 3. The color features for grade estimation are presented in the first column. The running time using only the same validation test is quoted in the next column. The RMSE shows the error between the predicted grades and actual values. The R^2 is reported in the last column.

Table 2. Predicted grade and actual grade measured by chemical assay

Sample	Predicted Cu%			Actual Cu%
	$(R,G,B)_{new}$	H_{new}	θ_{new}	
3	9.68	10.91	10.63	11.31
6	11.90	12.07	11.51	13.06
21	18.11	17.57	17.99	18.62

From Table 3 it can be seen that the running times are almost the same for the hue and color vector angle models and longer for the mean values of red, green and blue model. By comparing the RMSE values from Table 3, it can be found that the copper concentrate grades are predicted well with low error (< 2%) by using three types of color features as the inputs. Furthermore, the estimated model using the hue parameter has more accuracy as its RMSE value is the smallest. It can be also seen that there is a strong correlation between the predicted and actual copper concentrate grades since the values of correlation coefficient R^2 are very high.

Table 3. Statistical analysis of predicted copper concentrate grade based on LS-SVR models

Models	Performances of LS-SVR models		
	Running time (ms)	RMSE, %	R^2
Red, green, blue	56.455	1.1907	0.9977
Hue	30.243	0.8644	0.9997
Color vector angle	30.362	1.0391	0.9926

Conclusions

This paper proposes a new method, based on the least-squares support vector regression model (LS-SVR), to estimate the copper concentrate grade. This method uses color features as inputs with a small amount of samples. The color features, RGB, hue, and color vector angle were extracted from the color microscopic images of copper concentrate sample. Three models of grade estimation based on these color features were created. A comparison of the three models was made by validation test. The results pointed out that there was a strong correlation between the concentrate grade and color feature. It was also found that the model using the hue as the input parameter had more satisfactory performance than the models using the RGB and color vector angle. The model using the hue parameter provided a maximum R^2 value and the least error on the same validation test. In addition, this model required much less running time to estimate the grade than other two models.

References

AN S., LIU W., VENKATESH S., 2007, *Fast cross-validation algorithms for least squares support vector machine and kernel ridge regression*, Pattern Recognition, 40(8), 2154-2162.

BONIFAZI G., MASSACCI P., MELONI A., 2000, *Prediction of complex sulfide flotation performances by a combined 3D fractal and colour analysis of the froths*, Minerals Engineering, 13(7), 737-746.

BONIFAZI, G., MASSACCI P., MELONI A., 2002, *A 3D froth surface rendering and analysis technique to characterize flotation processes*, International Journal of Mineral Processing, 64(2), 153-61.

FORBES, G., 2007, *Texture and bubble size measurements for modelling concentrate grade in flotation froth systems*, PhD Thesis, University of Cape Town.

FORD A., ROBERTS A, 1998, *Colour space conversions*, http://www. poynton. com/PDFs/coloureq. pdf.

HAAVISTO O., KAARTINEN J., HYOTYNIEMI H., 2006, *Optical spectrum based estimation of grades in mineral flotation*, ICIT 2006, Industrial Technology, 2529-2534.

HAAVISTO O., KAARTINEN J., HYOTYNIEMI H., 2008, *Optical spectrum based measurement of flotation slurry contents*, International Journal of Mineral Processing, 88(3), 80-88.

HAMERS B., SUYKENS J.A.K., de MOOR B., 2002. *Compactly supported RBF kernels for sparsifying the gram matrix in LS-SVM regression models*, Artificial Neural Networks—ICANN 2002, Springer, 720-726.

KAARTINEN J., HÄTÖNEN J., HYOTYNIEMI H., MIETTUNEN J., 2006, *Machine-vision-based control of zinc flotation—a case study*, Control Engineering Practice, 14(12), 1455-66.

MALEWSKI J., KRZEMINSKA M., 2012, *Dependence of mine revenue on the grade of copper concentrate*, Physicochemical Problems of Mineral Processing, 48(2), 545-554.

MARAIS C., 2010, Estimation of concentrate grade in platinum flotation based on froth image analysis, PhD Thesis, University of Stellenbosch.

MORAR S.H., FORBES G., HEINRICH G.S., BRADSHAW D.J., 2005, *The use of a colour parameter in a machine vision system, Smart-Froth, to evaluate copper flotation performance at Rio Tinto's Kennecott Utah Copper Concentrator*, Centenary of Flotation Symposium, 147-151.

NAKHAEI F., MOSAVI M.R., SAM A., VAGHEI Y., 2012, *Recovery and grade accurate prediction of pilot plant flotation column concentrate: Neural network and statistical techniques*, International Journal of Mineral Processing, Vol. 110, 140-154.

NAKHAEI F., MOSAVI M.R., SAM A., 2013, *Recovery and grade prediction of pilot plant flotation column concentrate by a hybrid neural genetic algorithm*, International Journal of Mining Science and Technology, 23, 69-77.

OESTREICH, J. M., TOLLEY W.K., RICE D.A., 1995, *The development of a color sensor system to measure mineral compositions*, Minerals Engineering, 8(1), 31-9.

REN CH., YANG J., 2012, *A novel color microscope image enhancement method based on HSV color space and curvelet transform*, International Journal of Computer Science Issues, 9(6), 272-277.

SUYKENS J.A.K, VANDEWALLE J., 1999, *Least squares support vector machine classifiers*, Neural processing letters, 9(3), 293-300.

SUYKENS J.A.K., 2001, *Nonlinear modelling and support vector machines*, Instrumentation and Measurement Technology Conference, 1, 287-294.

TASDEMIR A., KOWALCZUK P.B., 2014, *Application of statistical process control for proper processing of the Fore-Sudetic Monocline copper ore*. Physicochemical Problems of Mineral Processing, 50(1), 249–264.

TKALCIC M., TASIC J.F., 2003, *Colour spaces: perceptual, historical and applicational background*, IEEE, 1, 304-308.

YONG Z., KEJUN J., YUKUN W., 2012, *Flotation concentrate grade prediction model based on RBF neural network & immune evolution algorithm*, Proceedings of the 31st Chinese Control Conference, 3319-3323.