In and out (Breas in and breeze out).

Good afternoon, dear colleagues! Today we will consider an important problem for the fertilizer industry – the pattern recognition in the differentiated image for the powder and granulated materials to classify the samples by size.

Slide 2

This slide shows the content of our work and some information about the speaker. Moreover, I am sincerely glad that I can report about problems of our production to you.

Slide 3

So, let us move on to the first point - mineral fertilizers and the role of particle size in this industry.

Slide 4

Mineral fertilizer production is a rapidly growing industry in Russia and the world. At the same time, there is a clear desire to use the basic principles of Industry 4 in it, such as decentralization of production and quality control. It is expressed in active develop of the automated detection and decision-making systems. Including the quality control procedure.

The one of the important parameters of mineral fertilizers is the particle size of granules and powders. In the literature it is often described that the particle size is responsible for the quality of the technological process, the quality of pellets, the agronomic effect and the quality of sample preparation for chemical and physical analyses. (This parameter will be a key parameter in our work. The possibility of its automated detection will be considered – see in room).

Slide 5

(As I mention the purpose of this work is to **develop an automated method for the classification of granular and powder materials by its particle size)** using the optoelectronic image with elimination of the external light efforts (such as brightness, temperature, and gradient). This is necessary because existing approaches have a number of drawbacks. (I would like to note that the described disadvantages do not allow to implement these methods in the automated industrial control of granular materials). It is important to ensure the stability of the method of control to constantly changing external conditions, the key condition is lighting (production facilities are lit up poorly).

Thus, we think that the most perspective and informative method is the optoelectronic control. That's why we chose it and develop the special approach, which allows us to remove the influence of external light on the quality of particle size classification.

For this purpose, about 150 samples of 5 different types of fertilizers were examined. Each sample of fertilizer was prepared as raw granules and grounded powder with two fractions (show on slide). Samples were pressed into tablets in the form of a “sandwich structure” (with boric acid).

Slide 6

Let's consider more details of the proposed algorithm. It will allow us to provide stability of the optoelectronic control and eliminate the effects of external illumination.

Slide 7

To implement the proposed algorithm the pressed samples were placed on a white sheet of paper and were photographed at a fixed distance to the sample surface with natural light illumination.

A simple digital USB video camera was used as an optoelectronic device for obtaining the image. Furthermore, the image was transferred to a computer via USB protocol. (As we can see, there is an irregularity of light for the image with sample).

Slide 8

In the next slide, we can see the main stages of the developed algorithm. I would like to note, that we highlight the area of the sample from photo and converted the image into grayscale. Furthermore, the image was differentiated and smoothed by a median filter. Thus, the surface map for each sample was formed (shown in the figure). (We separate three sample properties: average brightness of the selected image (1) and the average number and average area of founded patterns (2 and 3)). There are also two configurable parameters: the window for median filter and the counter constant for "marching square" algorithm.

Slide 9

With the received information, we built a matrix of «objects-features» and then we carried out the classification of particle size. Within the framework of the work, four types of classifiers are considered: linear without regularization, linear with L1 and L2 regularization, and non-linear “random forest”. The quality of classification was estimated by F-measure (show on slide). All proposed algorithms work in automated mode and were implemented in the Python 3.6 programming language.

Slide 10

Let’s see what we could get in our research.

Slide 11

First of all, we've set up anti-aliasing algorithms and “marching squares” search of patterns. With the help of the grid search approach, parameters were selected from two ranges. The quality metrics to select the values of parameters was the Spirman correlation coefficient between the average number and average size of the calculated patterns.

An interesting effect was found. Thus, the best Spearman correlation achieved with 13 pixels and 1.3 units, respectively. However, with this parameters the values of the number and area of patterns tend to 0. Nevertheless, the increase in the correlation coefficient is caused by the fact that one class falls out completely. Thus, the contour constant should not exceed 1 and the best parameters values are 7 pixels and 0.8 units. With these parameters, we calculated features and construct the “object-features” matrix for further classification.

Slide 12

After the describe stage, we studied the work of four classification algorithms (*list*). (The table shows the values of the F-measure along with the standard deviation of it, as well as the significance of the features for the classification. (*Look at the table*.)) I would like to note that the algorithm parameters were also selected by the grid search method, and the quality of the work was estimated by the cross-validation strategy on 10 stratified fold samples. The **target categorical variable** is the particle size in the pressed sample.

As expected, the random forest algorithm shows the best results (93%). Because the number and area of patterns of the image are related to each other and do not necessarily linear depend on the particle size. However the linear algorithms show the good results too (86%).

As we can see, the **most significant feature** were the average number of defined patterns. Actually, for pressed samples the variability of the patterns area will be small. Consequently, their weight for linear classification will also be small. On the contrary, for the non-linear algorithm the significance of features is slightly equalized.

Separately, we note that the average brightness of pixels plays a smaller role for the classification. This suggests that the differential correction of the external light works well.

Slide 13

In this way, we can draw the following conclusions, which listed on the next slide.

Slide 14

(Pause at 1 minute). I would like to point out that the views expressed in this work are those of the author and do not necessarily represent those of PhosAgro corporation. Also, the data is available here.

Slide 15

Thank you for your attention. I will be happy to answer any questions.

Possible question

For the start, let me apologize for my English. I understand it well, but unfortunately, I have very little conversation practice. So today, I will be using a cheat sheet. I ask you to be kind.

1. What does your company do?

I am the researchers group leader in PhosAgro' corporation and the lecturer in the Cherepovets State University. We are looking for new quality parameters and we are trying to automate the industrial quality control. Recently, I have developed and implemented a device for optical automatic control of granulometric composition.

2. Why not use existing particle size determination methods?

Unfortunately, they are expensive and do not work well in the production environment.

3. What are the reasons for the restrictions in the resolution?

Surface defects must be visible in the image, so that we can calculate patterns for classification.

4. Why did you choose to differentiate?

Our research is the applied work. We have used a set of approaches that give good results and allow us to carry out a qualitative classification. Of course, in the future we have carried out a more complete review of the possible methods.

5. Why did you use Spearman's classification?

It allows us to estimate the monotonous and nonlinear correlation. It is more flexible than the Pearson correlations.

6. Why did you choose F-measure?

This metric is the most informative and harmonic.

7. Your object properties depend on the image.

No, the properties do not depend on the image area and are normalized to this parameter. This is done for greater versatility.

8. Is your approach really new?

Yes, we have not found any literature on the use of these methods to eliminate external lighting conditions. All the described approaches are also used for the first time in the field of industrially produced mineral fertilizers.

9. The fact that the average brightness does not affect the classification does not mean that your approach works well.

With all due respect, I disagree. In a standardized calculation, this parameter reflects the impact of external lighting. Otherwise, the classification works well and this approach is already used in industrial production for optical particle size control.