Arch. Min. Sci., Vol. 59 (2014), No 2, p. 413–425

Electronic version (in color) of this paper is available: http://mining.archives.pl

DOI 10.2478/amsc-2014-0029

DARIUSZ JAMRÓZ*

APPLICATION OF MULTIDIMENSIONAL SCALING TO CLASSIFICATION OF VARIOUS TYPES OF COAL

ZASTOSOWANIE SKALOWANIA WIELOWYMIAROWEGO DO KLASYFIKACJI RÓŻNYCH TYPÓW WĘGLI

Visualization of multidimensional data is a new way of statistical analysis of so-called statistical graphical methods. These methods allow to classify some analyzed objects, including their various features. Facing grained materials problems, like coal or ores many characteristics have an influence on the quality of product. In case of coal, many features must be taken into consideration to determine quality of the material. Apart from most obvious characteristics like particle size, particle density or ash contents there are many others which cause significant differences between considered types of material. In the paper the application of Multidimensional Scaling Method is presented which is one of the multidimensional data visualization techniques. To this purpose, sampling of three types of coal was performed, which were 31, 34.2 and 35 (according to Polish classification of coal types). First, the material was screened on sieves and then divided into density fractions. Next step was to analyze chemically the obtained particle and size fractions of researched coal. Then, the Multidimensional Scaling Method was applied to visualize the investigated set of data. It was proved that the applied methodology allows to identify certain coal types efficiently and can be used as a qualitative criterion for grained materials. However, it was impossible to achieve such identification comparing all three types of coal together. The Multidimensional Scaling Method is new technique of data analysis concerning widely understood mineral processing.

Keywords: multidimensional scaling, MDS, multidimensional data visualization, coal, identification of data, statistical graphics methods, pattern recognition.

Surowce mineralne, które podlegają wzbogacaniu w celu ich lepszego wykorzystania mogą być charakteryzowane wieloma wskaźnikami opisującymi ich, interesujące przeróbkarza, cechy. Podstawowymi cechami są wielkość ziaren oraz ich gęstość, które decydują o przebiegu rozdziału zbiorów ziaren (nadaw) i efektach takiego rozdziału. Rozdział prowadzi się z reguły, w celu uzyskania produktów o zróżnicowanych wartościach średnich wybranej cechy, która zwykle charakteryzowana jest zawartością określonego składnika surowca wyznaczoną na drodze analiz chemicznych. Takie podejście do surowca mineralnego

* AGH UNIVERSITY OF SCIENCE AND TECHNOLOGY, FACULTY OF ELECTRICAL ENGINEERING, AUTOMATICS, COMPUTER SCIENCE AND BIOMEDICAL ENGINEERING, DEPARTMENT OF APPLIED COMPUTER SCIENCE, AL. A. MICKIEWICZA 30, 30-059 KRAKÓW, POLAND. E-mail: jamroz@agh.edu.pl

prowadzi do potraktowania go jako wielowymiarowego wektora $X = [X_1, \ldots, X_n]$. Zasadniczym problemem jest także wybór jednostki populacji generalnej (ziarno, jednostka objętości lub masy), co może decydować o określeniu charakteru wielowymiarowych powiązań cech wektora X. Takimi kierunkami charakteryzowania mogą być wielowymiarowe rozkłady wektora losowego X wraz ze wszystkimi konsekwencjami metody (Lyman, 1993; Niedoba, 2009; 2011; Olejnik et al., 2010; Niedoba i Surowiak, 2012), wielowymiarowe równania regresji wraz z analizą macierzy współczynników korelacji liniowej oraz korelacji cząstkowej (Niedoba, 2013c), analiza czynnikowa (Tumidajski i Saramak, 2009), czy metody wielowymiarowej wizualizacji danych, będące tematem niniejszego artykułu.

Biorąc pod uwagę analizę korelacji pomiędzy badanymi cechami materiałów uziarnionych (węgli) można zidentyfikować jakie jego cechy są ze sobą istotnie powiązane. Jest to swoiste preludium do wytypowania, które cechy węgla powodują istotne różnice pomiędzy jego typami. W artykule poddano badaniu trzy typy węgla, według polskiej klasyfikacji – węgle 31, 34.2 oraz 35, pochodzące z trzech różnych kopalni Górnośląskiego Okręgu Przemysłowego. Można powiedzieć, że z punktu widzenia ich jakości były to węgle energetyczne, semi-koksujące oraz koksujące. Każdy z tych węgli został poddany podziałowi na klasy ziarnowe, przy zastosowaniu odpowiedniego zestawu sit. Następnie każdą z otrzymanych klas ziarnowych rozdzielono w cieczach ciężkich na frakcje densymetryczne. Tak otrzymane klaso-frakcje zostały dodatkowo poddane analizie chemicznej ze względu na szereg cech, tj. ciepło spalania, zawartość siarki, zawartość substancji lotnych, zawartość popiołu, miąższość. Wyniki analiz dla wybranej klasy ziarnowej przedstawiono w tabeli 1. Tym samym otrzymano siedmiowymiarowy zestaw danych, który postanowiono poddać wielowymiarowej wizualizacji za pomocą metody skalowania wielowymiarowego.

Metoda skalowania wielowymiarowego (*multidimensional scaling*, MDS) jest jedną z nowoczesnych metod wizualizacji danych. Tego typu metody są wskazane zwłaszcza w sytuacji gdy ma się do czynienia z zestawem skomplikowanych i złożonych danych. Skalowanie wielowymiarowe jest odwzorowaniem przestrzeni n-wymiarowej w przestrzeń m-wymiarową. Oparte jest na obliczaniu odległości pomiędzy każdą parą n-wymiarowych punktów. Na podstawie tych odległości rozważana metoda ustala wzajemne położenie obrazów tych punktów w docelowej przestrzeni m-wymiarowej. Niech d_{ij} oznacza odległość pomiędzy n-wymiarowymi punktami nr i oraz j. Skalowanie wielowymiarowe polega na takim rozmieszczeniu punktów w przestrzeni m-wymiarowej, by odległość D_{ij} liczona w tej przestrzeni pomiędzy odwzorowanymi punktami nr i oraz j była jak najbardziej zbliżona do d_{ij}. Działanie algorytmu MDS może polegać na iteracyjnej zmianie położenia losowo (początkowo) rozmieszczonych punktów w przestrzeni m-wymiarowej w ten sposób, by funkcja: $S = \sqrt{\sum_{i>j} \left(D_{ij} - d_{ij} \right)^2}$ przyjęła jak najmniejszą wartość. Dla $m = 2$ metoda ta pozwala oglądać wielowymiarowe dane bezpośrednio na dwuwymiarowym ekranie komputera. Metodę tą w sposób szczegółowy oraz zastosowany algorytm przedstawiono w podrozdziale 3.

Rozdział 4 zawiera wyniki eksperymentów. Na rysunkach 1-4 widać, w jaki sposób wzrasta grupowanie punktów reprezentujących trzy różne klasy węgla (31, 34.2 oraz 35) wraz ze wzrostem parametru ITER. Widać, że punkty będące obrazami danych reprezentujących te same klasy węgla zaczynają zajmować osobne podobszary oraz zaczynają się grupować. Czytelność podziału przestrzeni rośnie wraz ze zwiększeniem parametru ITER, więc wraz z dokładniejszym dopasowaniem odległości obrazów punktów D_{ij} w przestrzeni 2-wymiarowej do oryginalnych odległości d_{ij} pomiędzy punktami w przestrzeni n-wymiarowej. Na rysunku 4 pokazano najbardziej czytelny wynik, jaki udało się uzyskać dla danych zawierających trzy typy węgla 31, 34.2 oraz 35. Nastąpiło to przy parametrze ITER = 793. Widać wyraźnie, że obrazy punktów danych reprezentujących próbki węgla danego typu gromadzą się w skupiskach. Można zaobserwować, że na prawie całym obszarze rysunku, skupiska te można od siebie odseparować. Jednak w niektórych częściach przestrzeni obrazy punktów reprezentujących różne klasy węgla zachodzą na siebie. Przez to nie możemy na podstawie tego rysunku stwierdzić, że analizowane dane pozwalają na prawidłową klasyfikację typów węgla. Postanowiono więc przeanalizować dane reprezentujące różne typy węgla parami. Na rysunkach 5-7 przedstawiono parami węgle typu, odpowiednio, 34.2 i 35 (Rys. 5), 31 i 34.2 (Rys. 6) oraz 31 i 35 (Rys. 7). Na każdym z tych rysunków widać czytelnie, że obrazy punktów reprezentujących próbki różnych typów węgla gromadzą się w skupiskach, które łatwo można od siebie odseparować. Przeprowadzona wizualizacja wielowymiarowa przy użyciu skalowania wielowymiarowego pozwala więc stwierdzić, że informacje zawarte w analizowanych siedmiowymiarowych danych są wystarczające do prawidłowej klasyfikacji typów węgla 31, 34.2 oraz 35.

Słowa kluczowe: skalowanie wielowymiarowe, MDS, wizualizacja danych wielowymiarowych, węgiel, identyfikacja danych, statystyczne metody graficzne, rozpoznawanie obrazów.

1. Introduction

Multidimensional statistical methods are the methods of so-called graphical statistical analysis. In case of complicated sets of data containing many sorts of information, the typical statistical approach is not always sufficient to obtain satisfying conclusions. This creates the necessity of searching for new solutions in this aspect. In mineral processing this kind of analysis is necessary mainly in case of materials featuring by many various properties. Often, the searched goal is to produce model equation which describes certain quality parameter. In this case the methods related to multidimensional approximation methods are in use. Such methods were described in many papers (Niedoba, 2009, 2011, 2013a, 2013b; Niedoba & Surowiak, 2012), including typical applications of neural networks or Markov chains (Gawenda et al., 2005; Saramak, 2011, 2013; Snopkowski & Napieraj, 2012; Szostek, 2003; Szostek & Suraj, 2002; Tumidajski, 1997). Apart from traditional modeling also multidimensional visualization methods became the new way of data investigation. These methods allow to recognize and identify similarities and potential differences between analyzed objects, which can be different sorts of grained materials, for example (Brożek & Surowiak, 2005, 2007, 2010). In the paper the analyzed material was coal, which three types were sampled and compared. The quality comparison analysis of researched coals was the goal of the investigation.

The qualitative analysis of multidimensional data (properties of material) obtained from the results of empirical experiments can be carried out by applying the multidimensional visualization method. The results of these analyses can be helpful thanks to materials characteristics as well as the construction of mineral processing models based on this data.

Attempts to depict multidimensional data have been undertaken on many occasions. Among many methods, the following ones can be selected: grand-tour method (Asimov, 1985, Cook et al., 1995), the method of principal component analysis (Li et al., 2000), use of neural networks for data visualization (Jamróz, 2014; Aldrich, 1998; Jain & Mao, 1992; Kraaijveld et al., 1995), parallel coordinates method (Chatterjee et al., 1993; Chou et al., 1999; Gennings et al., 1999; Inselberg, 1985), star graph method (Sobol & Klein, 1989), scatter-plot matrices method (Cleveland, 1984), relevance maps method (Assa et al., 1999). Visualization of multidimensional solids is also possible (Jamróz, 2009). The observational tunnels method (Jamróz, 2001, 2013) makes it possible to achieve an external view of the observed multidimensional sets of points using tunnel radius, introduced by the author (Jamróz & Niedoba, 2013, 2014; Niedoba, 2013b).

Methods of multidimensional data visualization by transformation of multidimensional space into two-dimensional space allow to show multidimensional data on computer screen. This allows to conduct qualitative analysis of these data in the most natural way for human by sense of sight. One of such methods is Multidimensional Scaling Method. This method was used in current paper to present and analyze set of seven-dimensional data describing samples of three various coal types 31, 34.2 and 35 (according to Polish classification of coals). It was decided to check whether this method allows to state that amount of information covered in seven coal features is sufficient to proper classification of coal types or not. The application of various methods to analyze possibilities of recognition of various coal properties becomes more and more interesting issue. Earlier, other visualization methods were applied, including observational tunnels method (Jamróz & Niedoba, 2013, 2014). Application of Multidimensional Scaling Method to this purpose is new way of approaching to this topic.

2. Material characteristics

Three types of coal, types 31 (energetic coal), 34.2 (semi-coking coal) and 35 (coking coal) in the Polish classification were first sampled and then used in the investigation. They originated from three various Polish coal mines and all of them were initially screened on a set of sieves of the following sizes: –1.00, –3.15, –6.30, –8.00, –10.00, –12,50, –14.00, –16.00 and –20.00 mm. Then, the size fractions were additionally separated into density fractions by separation in dense media using zinc chloride aqueous solution of various densities (1.3, 1.4, 1.5, 1.6, 1.7, 1.8 and 1.9 g/cm^3). The fractions were used as a basis for further consideration and additional coal features were determined by means of chemical analysis. For each density-size fraction such parameters as combustion heat, ash contents, sulfur contents, volatile parts contents and analytical moisture were determined, making up, together with the mass of these fractions, seven various features for each coal (in general, 205 samples). The examples of such data were presented in table 1 showing the data for size fractions 8.00-6.30 mm for each type of coal.

TABLE 1

Data for size fraction 8.00-6.30 mm for all three types of coal

Density [Mg/m^3]	Mass [g]	Combustion heat [cal]	Ash contents [%]	Sulfur contents [%]	Volatile parts contents V^a	Analytical moisture W_a
Coal, type 31						
<1.3	933	6857	10.70	0.56	30.05	3.27
1.3-1.4	531.6	6626	12.09	0.67	29.14	3.44
1.4-1.5	79.9	6106	16.20	1.11	28.60	3.64
1.5-1.6	28.9	5272	25.05	1.64	27.79	2.96
1.6-1.7	26.4	4380	34.50	1.50	26.35	2.76
1.7-1.8	28.7	3732	46.60	1.49	24.75	2.80
1.8-1.9	44	3135	50.86	2.14	24.15	2.11
>1.9	760.8	1672	72.24	0.76	12.57	1.25
Coal, type 34.2						
<1.3	601.5	8323	1.42	0.31	30.58	1.15
1.3-1.4	174.6	7958	4.73	0.45	26.37	1.1
1.4-1.5	37.2	6808	15.03	0.53	26.32	1.09
1.5-1.6	21	5838	21.63	0.7	27.21	1.1
1.6-1.7	12.9	5008	31.06	0.9	25.96	1.01
1.7-1.8	12.8	4398	37.59	0.82	24.95	0.95
1.8-1.9	0.8	No data	No data	0.71	No data	No data
>1.9	110	591	81.34	0.19	10.5	0.8
Coal, type 35						
<1.3	389.9	8146	3.88	0.39	20.53	1.32
1.3-1.4	131.2	7813	7.98	0.51	19.69	1.10
1.4-1.5	68.3	6682	19.72	0.56	18.70	1.10
1.5-1.6	43.6	5595	29.19	0.68	18.20	1.14
1.6-1.7	30.4	4526	38.91	0.82	18.24	1.09
1.7-1.8	27.1	3980	44.58	0.81	17.21	1.32
1.8-1.9	25.9	3425	49.87	0.56	16.10	1.47
>1.9	528.1	903	78.62	0.33	10.60	1.35

3. Multidimensional scaling

3.1. Method description

Multidimensional scaling (MDS) is the method based on mapping of n-dimensional space into m-dimensional space. It is based on calculation of a distance between each pair of n-dimensional points. On the basis of these distances the considered method determines mutual location of these points images in destined m-dimensional space. Let d_{ij} mean distance between n-dimensional points of no. i and j. Multidimensional scaling is based on such location of points in m-dimensional space that distance D_{ij} calculated in this space between mapped points of no. i and j is possibly closest to d_{ij}. The operation of algorithm MDS can be based on iterative change of location of randomly (initially) located points in m-dimensional space in the way assuring the function $S = \sqrt{\sum_{i>j}\left(D_{ij} - d_{ij}\right)^2}$ achieving possibly smallest value. For $m = 2$ this method allows to watch multidimensional data directly on two-dimensional computer screen (Kim et al., 2000).

3.2. Algorithm

Initial data set consists of parts described by n features. It can be treated then as set of n-dimensional vectors. Let mark i^{th} initial data vector as $x_i = (x_{i,1}, x_{i,2}, \dots x_{i,n})$. The algorithm serving to realization of visualization by means of multidimensional scaling compounds of several steps:

1. Scaling of initial data. Individual features represented by individual data dimensions must scaled in the way assuring their fitness in the same desired range. It was decided to scale individual coordinates (features) of data set vectors to range (0, 1).
2. Randomization of initial location of points images in 2-dimensional space. Let assume that image of i^{th} point x_i, so the point related to it in 2-dimensional space will be marked as $p_i = (u_i, v_i)$. Because in previous point each coordinate of data vectors was scaled to range (0, 1) the data set fits to n-dimensional box of side length equal to 1. It occurs that the biggest distance between two points of this set can amount the same as length of n-dimensional box diagonal, so is equal to $\sqrt{n}$. That is why the initial coordinates u_i and v_i are being randomized from range $(0, \sqrt{n})$ by means of plate probability distribution.

Next points 3-5 are realized for each possible pair i, j of initial data vectors ITER times (where ITER means parameter accepted in certain moment):

3. For next pair of points x_i, x_j their distance is calculated by means of Euclidean metrics:

$$d_{ij} = \sqrt{\sum_{k=1}^{n}\left(x_{i,k} - x_{j,k}\right)^2} \tag{1}$$

4. For next pair of points images $p_i = (u_i, v_i)$, $p_j = (u_j, v_j)$ their distance is calculated by means of Euclidean metrics:

$$D_{ij} = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2} \tag{2}$$

5. The location of image p_j is changed in the way assuring distance D_{ij} to be possibly closest to d_{ij}:

$$\tilde{u}_j = u_j + 0.01 \frac{\left(u_j - u_i\right)\left(d_{ij} - D_{ij}\right)}{D_{ij}} \tag{3}$$

$$\tilde{v}_j = v_j + 0.01 \frac{\left(v_j - v_i\right)\left(d_{ij} - D_{ij}\right)}{D_{ij}} \tag{4}$$

where $\widetilde{u}$ and $\widetilde{v}$ are values after change. Equations (3) and (4) cause that point p_j is closer to the point p_i when $D_{ij} > d_{ij}$ and farther from p_i when $D_{ij} < d_{ij}$. The allocation of point p_j occurs on straight line passing through points p_i and p_j. The constant 0.01 means the speed of points allocation in proper directions.

In this way the images of multidimensional initial data points in two-dimensional space were obtained. It is sufficient now to present image of each vector on computer screen. It is realized by drawing symbol representing fraction to which related data vector x_i belongs in location of coordinates (u_i, v_i). In this way the image of multidimensional points is created on computer screen for which the mutual distances are possibly preserved.

4. Experimental results

In purpose of visualization of seven-dimensional data describing various coal types the computer system was created based on assumptions presented in the previous chapter. The obtained results were shown on Figures 1-7. These views show the way how seven-dimensional data are transformed by means of multidimensional scaling to two dimensions. The algorithm of multidimensional scaling works in the way assuring preservation of mutual distanced between any two points despite significant reduction of number of dimensions. In this way it is possible to see significant features of seven-dimensional data on computer screen.

On Figures 1-4 it is visible how the grouping of points representing three types of coals (31, 34.2 and 35) grows with increase of parameter ITER. It can be observed that points being images of data representing the same coal fractions start to gather in separated subareas and start to group. The clearness of the space division grows with increase of ITER what means more precise fitting of distance of points D_{ij} images in two-dimensional space to original distances d_{ij} between points in n-dimensional space. Figure 1 presents initial position in which each image of point accept random position on screen. Figure 2 shows that even by small value of ITER = 5 images of points representing the same coal types start to group. Figure 3 presents view by value of ITER = 10. It is visible that further progress in grouping of points images occurs. On Figure 4 the most clear result is shown which was obtained for data containing three types of coal: 31, 34.2 and 35. It occurred by ITER = 793. It is clearly visible that images of data points representing samples of coal of certain point gather. It is also visible that almost on whole area of the Figure these gatherings can be separated. However, in some parts of the space the images of points representing various coal types overlap. That is why it is not possible to state that analyzed data allow to classify coal types properly on the basis of this Figure.

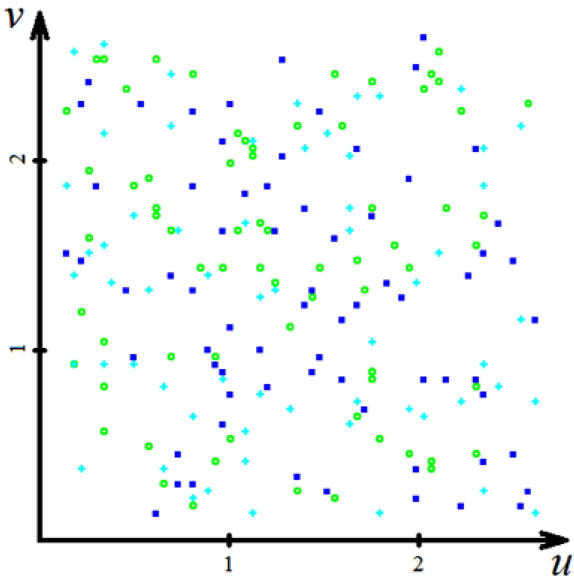

Fig. 1. Initial arrangement in which each image of point representing seven-dimensional data akcept random location. By symbol (▪) the images of points representing samples of coal, type 31 are marked, (+) – samples of coal, type 34.2, (○) – samples of coal, type 35

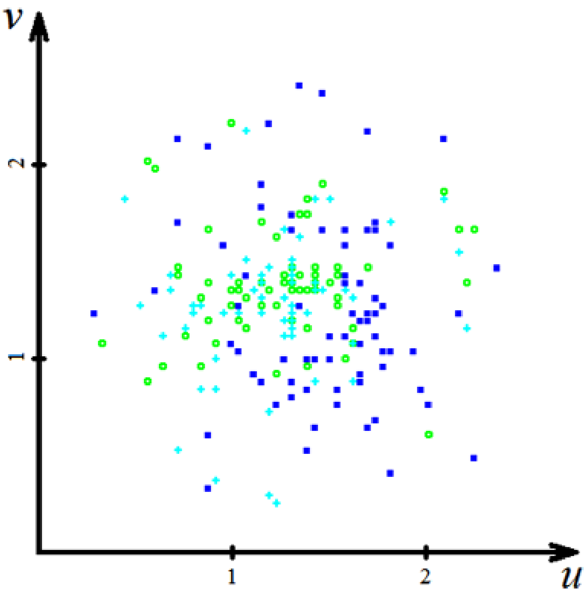

Fig. 2. View of seven-dimensional data representing three various type of coal by parameter ITER = 5. The initial stage of grouping is visible. By symbol (▪) the images of points representing samples of coal, ttype 31 are marked, (+) – samples of coal, type 34.2, (○) – samples of coal, type 35

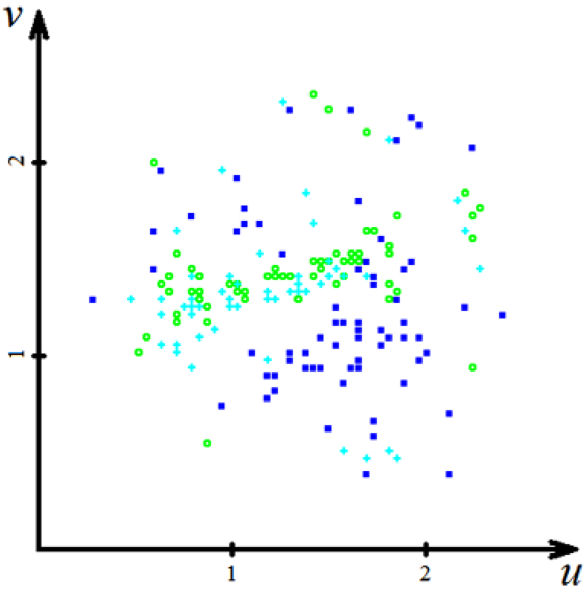

Fig. 3. View of seven-dimensional data representing three various type of coal by parameter ITER = 10. Farther grouping occurs. By symbol (▪) the images of points representing samples of coal, type 31 are marked, (+) – samples of coal, type 34.2, (○) – samples of coal, type 35

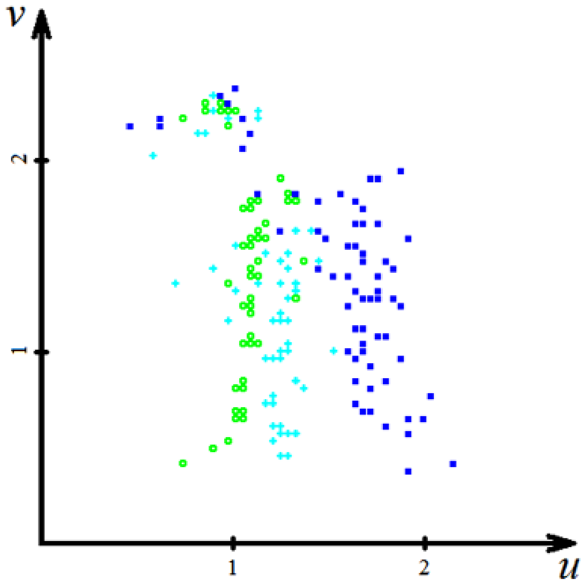

Fig. 4. The most clear view of seven-dimensional data representing three various type of coal obtained by parameter ITER = 793. By symbol (▪) the images of points representing samples of coal, type 31 are marked, (+) – samples of coal, type 34.2, (○) – samples of coal, type 35

In purpose of obtaining more clear results it was decided to present these data by means of multidimensional scaling in some other way. It was decided to analyze the data representing various coal types in pairs. Figure 5 presents the view obtained for data representing coal types 34.2 and 35. It is clearly visible that images of points representing samples of coal, type 34.2 gather and can be easily separated from gatherings of images of points representing coal, type 35.

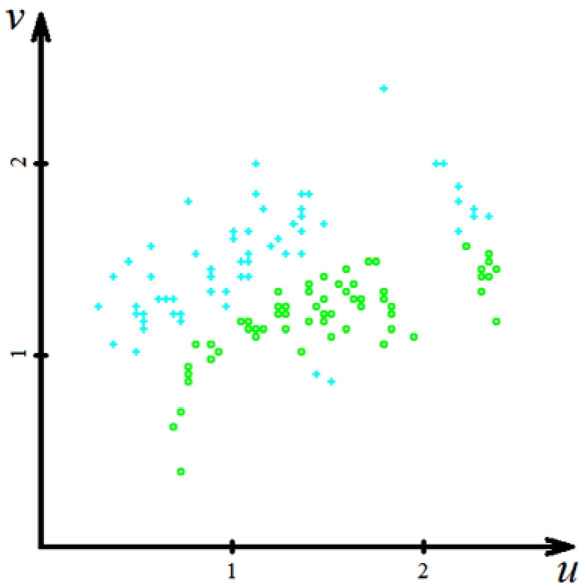

Fig. 5. View of seven-dimensional data representing two types of coal by parameter ITER = 119. By symbol (+) samples of coal, type 34.2 are marked, (o) – samples of coal, type 35

On Figure 6 the view obtained for data representing coal types 31 and 34.2 is presented. Also here it is visible very clearly that images of points representing samples of coal, type 31 gather and can be easily separated from gatherings of images of points representing coal, type 34.2. Figure 7 shows the view obtained for data representing coal types 31 and 35. It is clearly visible that images of points representing samples of coal, type 31 gather and can be easily separated from gatherings of images of points representing samples of coal, type 35.

If it is possible to separate samples of coal, type 34.2 from samples of coal, type 35 (Fig. 5) and it is possible to separate samples of coal, type 31 from samples of coal, type 34.2 (Fig. 6) and also it possible to separate samples of coal, type 31 from samples of coal, type 35 (Fig. 7) then it can be stated that samples of each three types of coal can be separated from each other. Applying multidimensional data visualization by means of multidimensional scaling it is possible to prove that information contained in seven-dimensional data describing samples of three types of coal are sufficient for their proper classification.

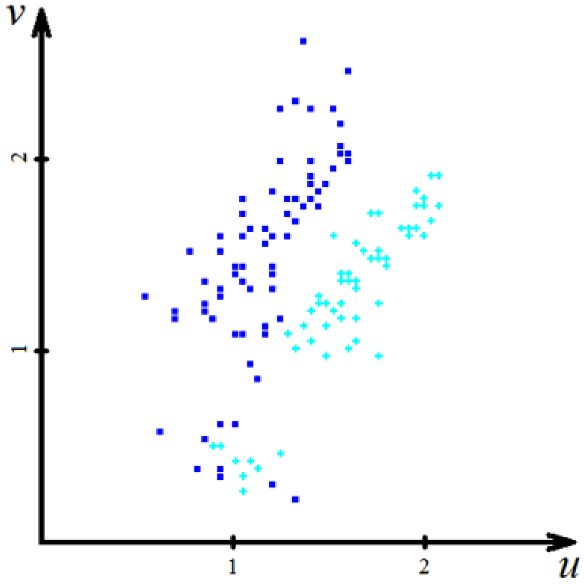

Fig. 6. View of seven-dimensional data representing two types of coal by parameter ITER = 4258. By symbol (▪) samples of coal, type 31 are marked, (+) – samples of coal, type 34.2

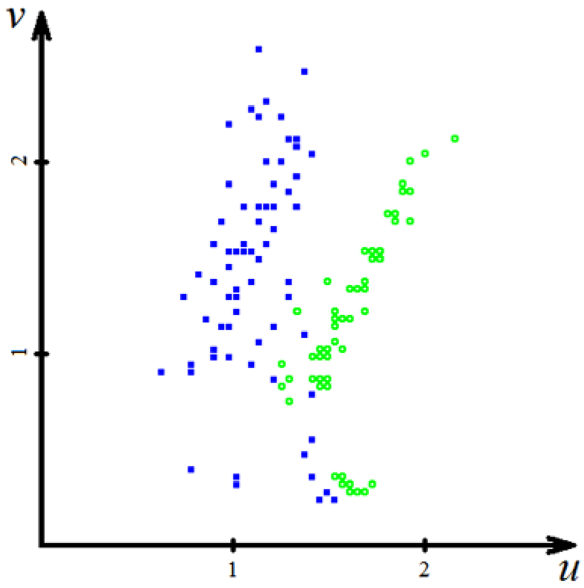

Fig. 7. View of seven-dimensional data representing two types of coal by parameter ITER = 750. By symbol (▪) samples of coal, type 31 are marked, (○) – samples of coal, type 35

It supposes to be noticed that algorithm of multidimensional scaling does not use information about affiliation of points representing data to certain fractions. In this situation this how images of points representing certain fraction will be grouped depends only on some data properties noticed by the algorithm.

5. Conclusions

The conducted experiments based on seven-dimensional visualization by means of multi-dimensional scaling allowed to get following conclusions:

1. Multidimensional visualization by means of multidimensional scaling allows to state that information contained in analyzed seven-dimensional data are sufficient to proper classification of coal types 31, 34.2 and 35.
2. Images of three types of coals on one Figure allowed to state that images of data points representing samples of coal of certain type gather in cluster and can be separated on almost whole area of the Figure. However, in some parts of the space the images of points representing various coal fractions overlapped. That is why it was impossible to state if analyzed data allowed to classify properly coal types on the basis of such view.
3. Only presentation of data representing three various types of coal in pairs allowed to obtain clear results. It allowed to state that images of points representing samples of coal of certain type gather and these gatherings can be separated. That means that data contain information sufficient for proper coal classification.
4. Clearness of results grows with more precise fitting of distance D_{ij} of images of points in 2-dimensional space to original distance d_{ij} between points in n-dimensional space (together with growth of parameter ITER).
5. Clearness of obtained results depends highly on accepted parameters.
6. One of problems occurring by such visualization is necessity of selection of parameters in purpose of view which clearly presents searched information. It is worthy to mention that during conducted experiments the views were obtained by values of ITER from 1 to 10000. These experiments were conducted multiply by renewed generation of random initial values. Sometimes it leads to obtaining more clear results. Also various speeds of learning were accepted before obtaining this one presented in equations 3 and 4. The results presented in the paper are the most clear from the obtained ones during course of experiment.

Acknowledgement

The paper is the result of scientific project no. *N N524 339040*, agreement no. *3390/B/T02/2011/40*.

References

Aldrich C., 1998. *Visualization of transformed multivariate data sets with autoassociative neural networks*. Pattern Recognition Letters, Vol. 19, Iss. 8, 1998, p. 749-764.

Asimov D.,1985. *The Grand Tour: A Tool for Viewing Multidimensional Data*. SIAM Journal of Scientific and Statistical Computing, Vol. 6, No, p. 128-143.

Assa J., Cohen-Or D., Milo T. 1999. *RMAP: a system for visualizing data in multidimensional relevance space*, Visual Computer, Vol. 15, no. 5, p. 217-34. Publisher: Springer-Verlag, Germany.

Brożek M., Surowiak A., 2010. *Argument of Separation at Upgrading in the Jig*. Arch. Min. Sci., Vol. 55, No 1, p. 21-40.

Brożek M., Surowiak A., 2007. *Effect of Particle Shape on Jig Separation Efficiency*. Physicochemical Problems of Mineral Processing, Vol. 41, p. 397-413.

Brożek M., Surowiak A., 2005. *The Dependence of Distribution of Settling Velocity of Spherical Particles on the Distribution of Particle Sizes and Densities*. Physicochemical Problems of Mineral Processing, Vol. 39, p. 199-210.

Chatterjee A., Das P.P., Bhattacharya S., 1993. *Visualization in linear programming using parallel coordinates*. Pattern Recognition, Vol. 26(11), p.1725-1736.

Chou S.-Y., Lin S.-W., Yeh C.-S., 1999. *Cluster identification with parallel coordinates*, Pattern Recognition Letters, Vol. 20, p. 565-572.

Cleveland W.S., McGill R., 1984. *The many faces of a scatterplot*. Journal of the American Statistical Association, Vol.79, p. 807-822.

Cook D., Buja A., Cabrera J., Hurley C., 1995. *Grand Tour and Projection Pursuit*. Journal of Computational and Graphical Statistics, Vol. 4, no. 3, p. 155-172.

Gawenda T., Saramak D., Tumidajski T., 2005. *Regression models of rock materials crushing in jaw crushers*. Scientific Issues of Civil Engineering and Environmental Engineering Faculty of Koszalin University of Science and Technology, series: Environmental Engineering, no 22, p. 659-670 [in Polish].

Gennings C., Dawson K.S., Carter W.H., Jr. Myers R.H., 1990. *Interpreting plots of a multidimensional dose-response surface in a parallel coordinate system*. Biometrics, Vol. 46, p. 719-735.

Inselberg A., 1985. *The plane with parallel coordinates*. Visual Computer, Vol. 1, p. 69-91.

Jain A.K., Mao J.: Artificial neural network for non-linear projection of multivariate data. In: *Proc. IEEE Internat. Joint Conf. On Neural Networks*. Baltimore, MD, Vol. 3, p. 335-340, 1992.

Jamróz D., 2001. *Visualization of objects in multidimensional spaces*. Doctoral Thesis, AGH, Kraków. [in Polish]

Jamróz D., 2009. *Multidimensional labyrinth – multidimensional virtual reality*. [In:] Cyran K., Kozielski S., Peters J., Stanczyk U., Wakulicz-Deja A. (eds.), *Man-Machine, Interactions*, AISC, Heidelberg, Springer-Verlag, p. 445-450.

Jamróz D., 2013. *Application of Multidimensional Data Visualization in Creation of Pattern Recognition Systems*. [In:] Cyran K., Kozielski S., Peters J., Stanczyk U., Wakulicz-Deja A. (eds.), *Man-Machine, Interactions*, AISC, Heidelberg, Springer-Verlag, 2013. (in printing).

Jamróz D., 2014. *Application of multi-parameter data visualization by means of autoassociative neural networks to evaluate classification possibilities of various coal types*. Physicochemical Problems of Mineral Processing, Vol. 50(2), p. 719-734.

Jamróz D., Niedoba T., 2014. *Application of Observational Tunnels Method to Select Set of Features Sufficient to Identify a Type of Coal*. Physicochemical Problems of Mineral Processing, Vol. 50(1), p. 185-202.

Kim S.-S., Kwon S., Cook D., 2000. *Interactive visualization of hierarchical clusters using MDS and MST*. Metrika, Vol. 51, Springer-Verlag, p. 39-51.

Kraaijveld M., Mao J., Jain A.K., 1995. *A nonlinear projection method based on Kohonen's topology preserving maps*. IEEE Trans. Neural Networks, Vol. 6(3), p. 548-559.

Li W., Yue H.H., Valle-Cervantes S., Qin S.J., 2000. *Recursive PCA for adaptive process monitoring*. Journal of Process Control, Vol. 10, iss. 5, p. 471-486.

Lyman G.J., 1993. *Application of Line-Length Related Interpolation Methods to Problems in Coal Preparation – III: Two dimensional Washability Data Interpolation*. Coal Preparation, vol. 13, p. 179-195.

Niedoba T., 2009. *Multidimensional distributions of grained materials characteristics by means of non-parametric approximation of marginal statistical density function*. AGH Journal of Mining and Geoengineering, Iss. 4, p. 235-244 [in Polish].

Niedoba T., 2011 *Three-dimensional distribution of grained materials characteristics*. [In:] Proceedings of the XIV Balkan Mineral Processing Congress, Tuzla, Bosnia and Herzegovina, Vol. 1, p. 57-59.

Niedoba T., 2013b. *Statistical analysis of the relationship between particle size and particle density of raw coal*. Physicochemical Problems of Mineral Processing, Vol. 49, iss. 1, p. 175-188.

Niedoba T., Jamróz D., 2013. *Visualization of multidimensional data in purpose of qualitative classification of various types of coal*. Arch. Min. Sci., Vol. 58, No 4, p. 1317-1331.

Niedoba T., Surowiak A., 2012. *Type of coal and multidimensional description of its composition with density and ash contents taken into consideration*. [In:] Proceedings of the XXVI International Mineral Processing Congress, Vol. 1, p. 3844-3854.

Olejnik T., Surowiak A., Gawenda T., Niedoba T., Tumidajski T., 2010. *Multidimensional coal characteristics as basis for evaluation and adjustment of its beneficiation technology*. AGH Journal of Mining and Geoengineering, Vol. 34, iss. 4/1, p. 207-216 [in Polish].

Saramak D., 2013. *Mathematical models of particle size distribution in simulation analysis of High-pressure grinding rolls operation*. Physicochemical Problems of Mineral Processing, Vol. 49(1), p. 495-512.

Saramak D., 2011. *Technological Issues Of High-Pressure Grinding Rolls Operation In Ore Comminution Processes*. Arch. Min. Sci., Vol. 56, No 3, p. 517-526.

Snopkowski R., Napieraj A., 2012. *Method Of The Production Cycle Duration Time Modeling Within Hard Coal Longwall Faces*. Arch. Min. Sci., Vol. 57, No 1, p. 121-138.

Sobol M.G., Klein G., 1989. *New graphics as computerized displays for human information processing*. IEEE Trans. Systems Man Cybernet., Vol. 19 (4), p. 893-898.

Szostek R., Suraj Z., 2002. *The use covert of Markov chains to analyze the temporal information systems*. Folia Scientiarum Universitatis Technicae Resoviensis, Mathematics, Rzeszow, Vol. 26, p. 263-273 [in Polish].

Szostek R., 2003. *Application of neural networks to approximate probability distributions using Cox distributions*. AGH University of Science and Technology Journals, Electrical Engineering and Electronics, Krakow, Vol. 22(2), p. 114-122 [in Polish].

Tumidajski T., Saramak D., 2009. *Methods and models of mathematical statistics in mineral processing*. Wydawnictwo AGH, Kraków [in Polish].

Tumidajski T., 1997. *Stochastic analysis of grained materials properties and their separation processes*, Wydawnictwo AGH, Kraków [in Polish].

Received: 21 October 2013