DE GRUYTER OPEN

Arch. Min. Sci., Vol. 60 (2015), No 1, p. 93–106

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

DOI 10.1515/amsc-2015-0007

TOMASZ NIEDOBA*

APPLICATION OF RELEVANCE MAPS IN MULTIDIMENSIONAL CLASSIFICATION OF COAL TYPES

ZASTOSOWANIE MAP ODNIESIENIA W WIELOWYMIAROWEJ KLASYFIKACJI TYPÓW WĘGLA

Multidimensional data visualization methods are a modern tool allowing to classify some analyzed objects. In the case of grained materials e.g. coal, many characteristics have an influence on the material quality. In case of coal, apart from most obvious features like particle size, particle density or ash contents there are many others which cause significant differences between considered types of material. The paper presents the possibility of applying visualization techniques for coal type identification and determination of significant differences between various types of coal. Author decided to apply relevance maps to achieve this purpose. Three types of coal – 31, 34.2 and 35 (according to Polish classification of coal types) were investigated, which were initially screened on sieves and then divided into density fractions. Then, each size-density fraction was chemically analyzed to obtain other characteristics. It was stated 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 presented methodology is new way of analyzing data concerning widely understood mineral processing.

Keywords: relevance maps, multidimensional data visualization, coal, identification of data, pattern recognition

Surowce mineralne, które podlegają wzbogacaniu w celu ich lepszego wykorzystania mogą być (charakteryzują się) 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 prowadzi do potraktowania go jako wielowymiarowego wektora $X = [X_1, ..., X_n]$. Zasadniczym problemem jest także wybór jednostki populacji generalnej (ziarno, jednostka objętości lub masy), co może decydować o kierunkach charakteryzowania 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);

* AGH UNIVERSITY OF SCIENCE AND TECHNOLOGY, FACULTY OF MINING AND GEOENGNEERING, DEPARTMENT OF ENVIRONMENTAL ENGINEERING AND MINERAL PROCESSING, AL. A. MICKIEWICZA 30, 30-059 KRAKOW, POLAND. E-MAIL: tniedoba@agh.edu.pl

– wielowymiarowe równania regresji wraz z analizą macierzy współczynników korelacji liniowej oraz korelacji cząstkowej (Niedoba, 2013b);
– analiza czynnikowa (Tumidajski, 1997; Tumidajski and Saramak, 2009);
– metody wielowymiarowej wizualizacji danych.

W artykule zastosowano nowoczesną metodę wizualizacji wielowymiarowych danych – metodę tzw. map odniesienia (z ang. *relevance maps*). Aby zastosować ww. metodę przeprowadzono doświadczenia na trzech typach węgla, pobranych z trzech kopalni węgla kamiennego, zlokalizowanych w Górnośląskim Okręgu Przemysłowym. Były to węgle typu 31, 34.2 i 35, według polskiej klasyfikacji węgli. Każdą z pobranych prób poddano rozdziałowi na klasy ziarnowe a następnie każdą z klas ziarnowych rozdzielono na frakcje densymetryczne za pomocą rozdziału w roztworze chlorku cynku. Tak otrzymane klaso-frakcje przebadano chemiczne ze względu na wybrane parametry jakościowe węgla. Były to takie cechy jak: ciepło spalania, zawartość popiołu, zawartość siarki, zawartość substancji lotnych oraz miąższość materiału. Otrzymano w ten sposób zestaw siedmiu danych dla każdej klasy ziarnowej i każdego typu węgla. Stanowił on swoisty siedmiowymiarowy zbiór, który postanowiono zobrazować za pomocą techniki wizualizacji bazującej na tzw. mapach odniesienia.

W metodzie map odniesienia na płaszczyźnie służącej do wizualizacji danych zostają rozmieszczone specjalne punkty zwane punktami odniesienia, reprezentujące poszczególne cechy. Do każdej cechy (współrzędnej) zostaje przyporządkowany punkt odniesienia reprezentujący tą cechę. Czyli przy danych 7-wymiarowych umieszczamy na płaszczyźnie 7 takich punktów odniesienia reprezentujących poszczególne współrzędne. Rozkład punktów reprezentujących przedstawiane wielowymiarowe dane odzwierciedla relacje pomiędzy tymi danymi a cechami. Im bardziej *i*-ta cecha występuje w danym obiekcie (czyli *i*-ta współrzędna ma większą wartość), tym bliżej powinien leżeć punkt reprezentujący dany obiekt względem punktu odniesienia reprezentującego *i*-tą cechę (współrzędną). W ten sposób każdy punkt odniesienia reprezentujący daną cechę, dzieli płaszczyznę na obszary bardziej oraz mniej zależne od cechy nr *i* (mniej oraz bardziej odległe od punktu odniesienia reprezentującego *i*-tą cechę). Dokładny opis algorytmu przedstawiono w podrozdziale 3 artykułu.

Za pomocą omawianej metody dokonano wizualizacji danych dotyczących przedstawionych typów węgla. Uzyskane rezultaty przedstawiono na rysunkach 1-9. Widoki te pokazują sposób, w jaki 7-wymiarowe dane zostają przekształcone przy pomocy mapy odniesienia do dwóch wymiarów. Algorytm wizualizacji przy użyciu mapy odniesienia działa tak by pomimo znacznej redukcji liczby wymiarów, w jak największym stopniu odległości pomiędzy punktem reprezentującym konkretny wektor danych a punktami odniesienia zależały od współrzędnych tego wektora danych. W ten sposób na ekranie 2-wymiarowym, możemy zobaczyć istotne cechy danych 7-wymiarowych.

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. Jednak w niektórych częściach przestrzeni obrazy punktów reprezentujących różne klasy węgla zachodzą na siebie. Przez to nie możemy na podstawie tych rysunków stwierdzić, że analizowane dane pozwalają na prawidłową klasyfikację typów węgla.

W celu uzyskania bardziej czytelnych wyników postanowiono przedstawić przy pomocy mapy odniesienia, te same dane w nieco inny sposób. Postanowiono przeanalizować dane reprezentujące różne typy węgla parami. Rysunek 5 przedstawia widok uzyskany dla danych reprezentujących typy węgla 34.2 oraz 35. Widać na nim czytelnie, że obrazy punktów reprezentujących próbki węgla typu 34.2 gromadzą się w skupiskach, które łatwo można odseparować od skupisk obrazów punktów reprezentujących próbki węgla 35. Podobne obserwacje dokonano na podstawie rysunków 6 i 7, gdzie przedstawiono parami, odpowiednio, węgle typu 31 i 34.2 oraz 31 i 35. Przeprowadzona wizualizacja wielowymiarowa przy użyciu map odniesienia pozwala więc stwierdzić, że informacje zawarte w analizowanych siedmiowymiarowych danych są wystarczające do prawidłowej klasyfikacji typów węgla 31, 34.2 oraz 35.

Słowa kluczowe: mapy odniesienia, wizualizacja wielowymiarowych danych, identyfikacja danych, rozpoznawanie kształtów

1. Introduction

In modern statistical researches there is often a need of applying multidimensional statistical methods. Simple regressive analysis is not enough in more complex cases. That is why many new, modern techniques are introduced in scientific works. Of course, there are many methods connected with regressive equations (Niedoba, 2009, 2011, 2013b; Niedoba & Surowiak, 2012), applications of statistical forecasting, Markov chains and neural networks (Gawenda et al., 2005; Saramak, 2011, 2013; Snopkowski & Napieraj, 2012; Szostek, 2010, 2012; Tumidajski, 1997) but also many methods of data-mining are in use. Special type of statistical analysis are multidimensional data visualization methods which main purpose is to recognize differences and similarities between analyzed sets of data. Finding these differences is often a very important issue in mineral processing where processes depend on many material features (Brożek & Surowiak, 2005, 2007, 2010).

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 (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), multidimensional scaling (Kim et al., 2000), scatter-plot matrices method (Cleveland, 1984). 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; Niedoba & Jamróz, 2014).

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 relevance maps method. This method was used in work 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; Niedoba & Jamróz, 2014; Jamróz, 2014a, 2014b, 2014c; Niedoba, 2014). Application of relevance maps 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 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. The examples of such data were presented in table 1 showing the data for size fractions 14.00-12.50 mm for each type of coal.

TABLE 1

Data for size fraction 14.00-12.50 mm for all three types of coal

Coal type 31	**Density [Mg/m^3]**	<1.3	1.3-1.4	1.4-1.5	1.5-1.6	1.6-1.7	1.7-1.8	1.8-1.9	>1.9
	Mass [g]	308.6	292.5	36.1	10.7	25.6	139.0	12.7	601.2
	Combustion heat [kJ/kg]	29508.6	24530.5	12342.7	21423.8	18702.4	16412.2	12887.0	1913.4
	Ash contents [%]	6.41	19.61	16.55	26.10	35.78	37.20	48.20	86.53
	Sulfur contents [%]	0.72	0.70	0.76	1.55	2.28	1.23	1.13	0.40
	Volatile parts contents V^a	34.32	29. 22	28.92	31.08	26.71	29.24	24.05	9.30
	Analytical moisture W_a	3.23	3.36	3.87	3.40	2.40	2.19	2.23	0.91
Coal type 34.2	**Density [Mg/m^3]**	<1.3	1.3-1.4	1.4-1.5	1.5-1.6	1.6-1.7	1.7-1.8	1.8-1.9	>1.9
	Mass [g]	360.5	57.0	25.5	12.2	3.2	15.0	3.6	68.9
	Combustion heat [kJ/kg]	34444.8	32016.5	28893.1	24275.1	20222.2	17383.6	18484.7	2901.5
	Ash contents [%]	2.00	7.67	15.33	33.73	34.30	36.15	27.0	79.33
	Sulfur contents [%]	0.32	0.71	0.83	0.17	0.34	0.34	0.05	0.91
	Volatile parts contents V^a	28.96	24.16	24.58	27.85	No data	27.93	31.75	12.08
	Analytical moisture W_a	1.04	1.87	1.34	0.95	No data	0.37	1.01	0.52
Coal type 35	**Density [Mg/m^3]**	<1.3	1.3-1.4	1.4-1.5	1.5-1.6	1.6-1.7	1.7-1.8	1.8-1.9	>1.9
	Mass [g]	268.7	89.3	39.8	22.0	25.7	29.0	28.1	589.5
	Combustion heat [kJ/kg]	34863.5	31861.5	27494.7	21063.8	20883.8	19213.2	13757.8	2939.1
	Ash contents [%]	2.38	8.97	19.61	35.68	34.62	40.60	52.24	80.57
	Sulfur contents [%]	0.28	0.36	0.56	0.39	1.26	0.38	1.14	0.20
	Volatile parts contents V^a	20.28	20.10	18.83	16.22	19.42	18.86	17.95	10.84
	Analytical moisture W_a	1.45	1.21	1.28	1.32	1.47	1.61	1.51	1.37

3. Relevance maps

3.1. Method description

Relevance maps method on plane serving for data visualization is based on placing special points called relevance points which represent individual features of considered object (Assa, 1999). For each feature (coordinate) the relevance point representing this feature is assigned. That means that by seven-dimensional data set 7 such points are placed on plane which represent individual coordinates. The distribution of the points representing presented multidimensional data shows relations between these data and features. The more i^{th} feature is present in certain object (what means that i^{th} coordinate is higher) the closest point representing certain object according to relevance point representing i^{th} feature (coordinate) should be. In this way each relevance point representing certain feature divides plane on areas more or less dependent on i^{th} feature (more or less distanced from relevance point representing i^{th} feature).

3.2. Algorithm

Set of initial data compounds of elements described by n features. It can be then treated as set of n-dimensional vectors. Let mark i^{th} initial data vector as $x_i = (x_{i,1}, x_{i,2}, \ldots x_{i,n})$. Algorithm serving to realize visualization by means of relevance map consists several steps, which are:

1. Scaling of initial data. Individual features, represented by individual data dimensions are scaled in the way assuring their fitness to the same specific range. It was decided to scale the 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 j^{th} point x_j, what means the point related to it in 2-dimensional space, will be marked as $p_j = (u_j, v_j)$. In previous point each of data vectors coordinates was scaled to range (0, 1). That is why the initial coordinates u_j and v_j were drawn also from range (0,1) by plate probabilistic distribution applied.
3. Randomization of initial location of relevance points in 2-dimensional space. For n-dimensional data n such points are created. These points will be marked as $w_i = (y_i, z_i)$. Initial coordinates y_i and z_i are drawn also from range (0, 1) by means of application of plate probabilistic distribution.

Next steps 4-5 are realized for each initial data vector:

4. For certain j^{th} vector of data x_j the distance of its image p_j from each relevance point w_i is calculated by means of Euclidean metrics:

$$D_{ij} = \sqrt{\left(y_i - u_j\right)^2 + \left(z_i - v_j\right)^2} \tag{1}$$

5. The location of image p_j is changed taking into consideration distance from each relevance point w_i. It is changed in the way assuring the distance D_{ij} from relevance point w_i to be relevant to i^{ith} coordinate of data vector x_j. The goal is to achieve such image p_j which by high value of i^{th} coordinate of point x_j (close to 1) is possibly close to relevance point w_i. For small value of i^{th} coordinate of point x_j its image p_j is far from relevance point w_i:

$$\tilde{u}_j = u_j + 0.01\frac{\left(u_j - y_i\right)\left(1 - x_{j,i} - D_{ij}\right)}{D_{ij}} \tag{2}$$

$$\tilde{v}_j = v_j + 0.01\frac{\left(v_j - z_i\right)\left(1 - x_{j,i} - D_{ij}\right)}{D_{ij}} \tag{3}$$

where $\tilde{u}$ and $\tilde{v}$ are values after change. Equations (2) and (3) cause that point p_j is closer to point w_i when $D_{ij} > 1 - x_{j,i}$ and farther from w_i when $D_{ij} < 1 - x_{j,i}$. The dislocation of point p_j occurs according to straight line crossing through points p_j and w_i. Constant value 0.01 means the speed of points moving in proper directions.

Next steps 6-7 are realized for each relevance point:

6. For given i^{th} relevance point its distance D_{ij} is calculated from each image p_j of data vector by means of equation (1).
7. The location of relevance point w_i is changed taking into consideration distance from each image p_j of data vector x_j. It is changed in the way assuring distance D_{ij} from relevance point w_i to be relevant to i^{th} coordinate of data vector x_j. Similarly like in step 5 the goal is to obtain possibly closest location of image p_j to relevance point w_i by high value of i^{th} coordinate of point x_j (close to 1). For small value of i^{th} coordinate of point x_j its image p_j is far from relevance point w_i:

$$\tilde{y}_i = y_i + 0.01\frac{\left(y_i - u_j\right)\left(1 - x_{j,i} - D_{ij}\right)}{D_{ij}} \tag{4}$$

$$\tilde{z}_i = z_i + 0.01\frac{\left(z_i - v_j\right)\left(1 - x_{j,i} - D_{ij}\right)}{D_{ij}} \tag{5}$$

where $\tilde{y}$ and $\tilde{z}$ are values after change. Equations (4) and (5) cause that relevance point w_i is closer to point p_j when $D_{ij} > 1 - x_{j,i}$ and farther from it when $D_{ij} < 1 - x_{j,i}$. Disslocation of relevance point w_i occur according to straight line crossing through points p_j and w_i. The constant value 0.01 means speed of points moving in proper directions.

The whole procedure covered in steps 4-7 is repeated ITER times (where ITER means the parameter accepted in certain moment).

In this way the image of multidimensional initial data points in two-dimensional space were obtained. It is sufficient then to present image of each vector on computer screen. It is realized by drawing symbol representing fraction to which data vector x_i is related in place of coordinates (u_j, v_j). Additionally, symbols representing relevance points in in screen location (y_i, z_i) can be drawn. In this way the image of multidimensional points is created on computer screen.

4. Experiment results

In purpose of visualization of seven-dimensional data describing various coal types the computer system based on assumptions presented in previous chapter was constructed. The obtained results were presented on Figures 1-9. These views showed the way how seven-dimensional data were transformed by means of relevance maps to only two dimensions. The algorithm of visualization by means of relevance maps works in the way assuring the most possibly the distances between point representing certain data vector and relevance points to depend on coordinates of

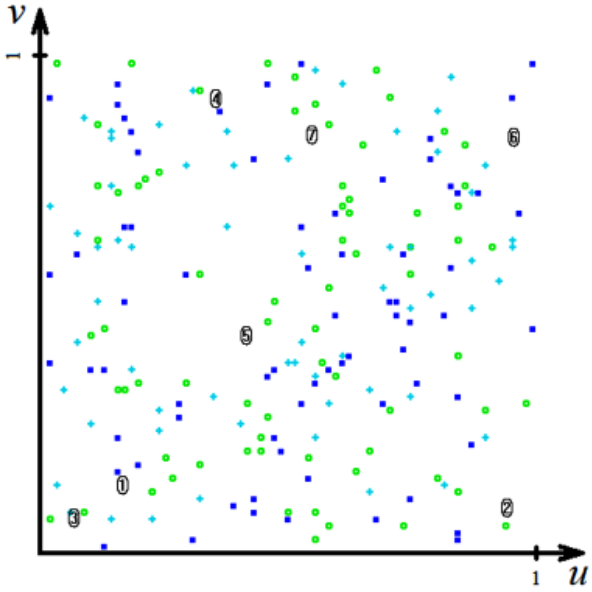

Fig. 1. Initial location in which each image of point representing seven-dimensional data and each relevance point accept random location. By symbol (■) the images of points representing samples of coal, type 31 were marked, by (+) – samples of coal, type 34.2, by (o) – samples of coal, type 35. Digit of value *i* means the relevance point representing i^{th} coordinate

this data vector. This way the projections of important features of seven-dimensional data can be seen on computer screen.

On Figures 1-3 it can be seen how the grouping of points representing Three various coal types (31, 34.2 and 35) grows with increasing of parameter ITER. It can be observed that points being images of data representing the same coal fractions start to gather in clusters in separated subareas. The clearness of separation grows with increasing of parameter ITER what means more precise fitness of distance D_{ij} between images of points and relevance points in two-dimensional space to proper coordinates of seven-dimensional data vector x_j. Figure 1 presents the initial location in which each image of input data vector and each relevance point gather in random position on computer screen. Figure 2 shows that by small value of parameter ITER = 45 the images of points representing the same fractions start to group. Also the relevance points change their locations. On Figures 3 and 4 possibly the most clear results were shown which were obtained for data containing three coal types: 31, 34.2 and 35. Figure 3 was obtained by parameter ITER = 7000. On Figure 4 the result obtained by other setting of random values of initial points images and parameter ITER = 2500 was shown. Both Figures 3 and 4 show clearly that images of data points representing samples of coal of certain type gather in clusters. It can be seen that on almost whole area of the Figure these clusters can be separated. However, in some parts of space images of points representing various coal types overlap. This is the reason of not possible proper classification of coal types on the basis of these Figures.

In purpose of getting more clear results it was decided to present the same data in other way, applying relevance maps to this purpose. It was decided to analyze data representing vari-

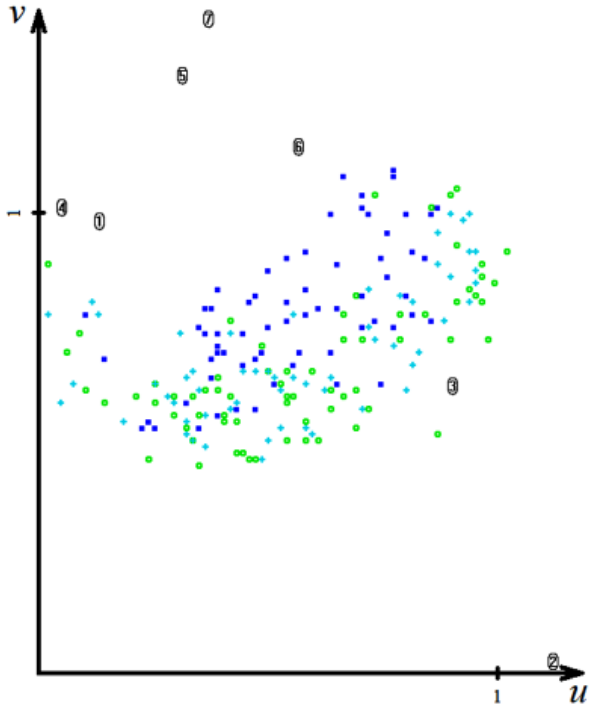

Fig. 2. View of seven-dimensional data representing three various coal type by parameter ITER = 45. The beginning of grouping can be observed. By symbol (■) the images of points representing samples of coal, type 31 were marked, by (+) – samples of coal, type 34.2, by (o) – samples of coal, type 35. Digit of value *i* means the relevance point representing i^{th} coordinate

ous coal types in pairs. Figure 5 shows the view obtained for data representing coal types 34.2 and 35. It can be seen clearly that images of points representing samples of coal, type 34.2 gather in clusters which can be separated easily from clusters of points images representing samples of coal, type 35.

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

If it is possible to confirm the possibility of separation of coal, type 34.2 samples from samples of coal, type 35 (Fig. 5) as well the separation of samples of coal, type 31 from samples of coal, type 34.2 (Fig. 6) and separation of samples of coal, type 31 from samples of coal, type 35 (Fig. 7) then it is possible to state that samples of each type of three types of coal can be efficiently separated. Applying visualization of multidimensional data by means of relevance maps describing samples of three types of coal confirmed that they are sufficient to proper classification of coal types.

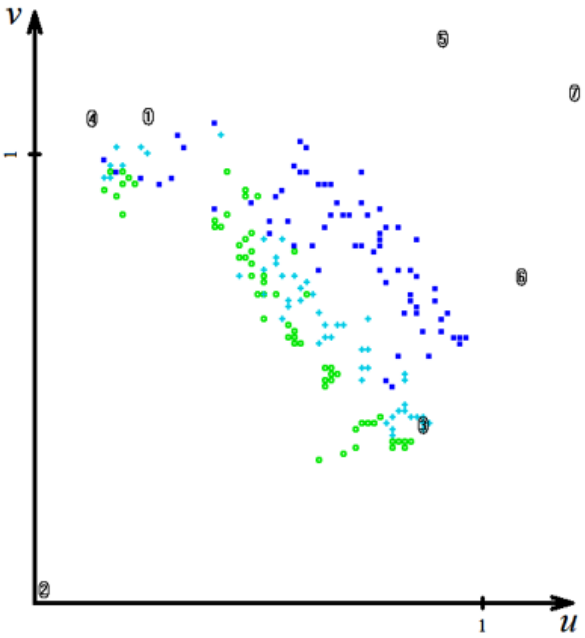

Fig. 3. View of seven-dimensional data representing three various coal type by parameter ITER = 7000. By symbol (■) the images of points representing samples of coal, type 31 were marked, by (+) – samples of coal, type 34.2, by (o) – samples of coal, type 35. Digit of value *i* means the relevance point representing i^{th} coordinate

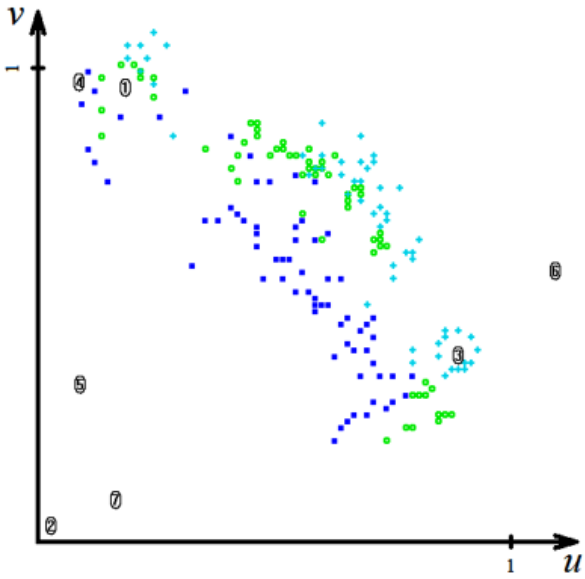

Fig. 4. View of seven-dimensional data representing three various coal type by parameter ITER = 2500. By symbol (■) the images of points representing samples of coal, type 31 were marked, by (+) – samples of coal, type 34.2, by (o) – samples of coal, type 35. Digit of value *i* means the relevance point representing i^{th} coordinate

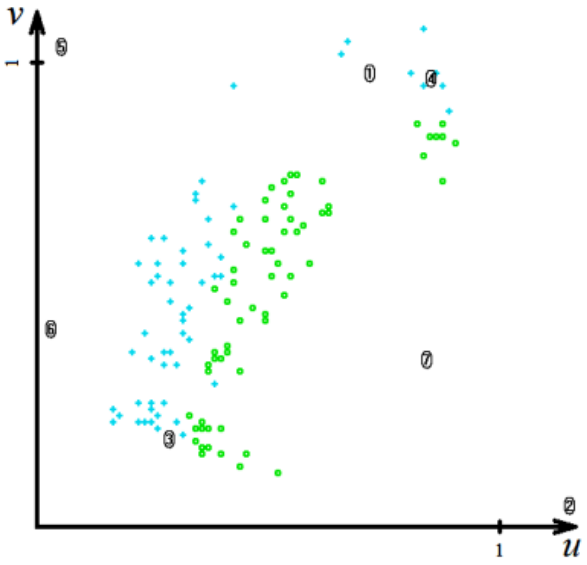

Fig. 5. View of seven-dimensional data representing two various coal type by parameter ITER = 14998. By symbol (+) the images of points representing samples of coal, type 34.2 were marked, by (o) – samples of coal, type 35. Digit of value *i* means the relevance point representing i^{th} coordinate

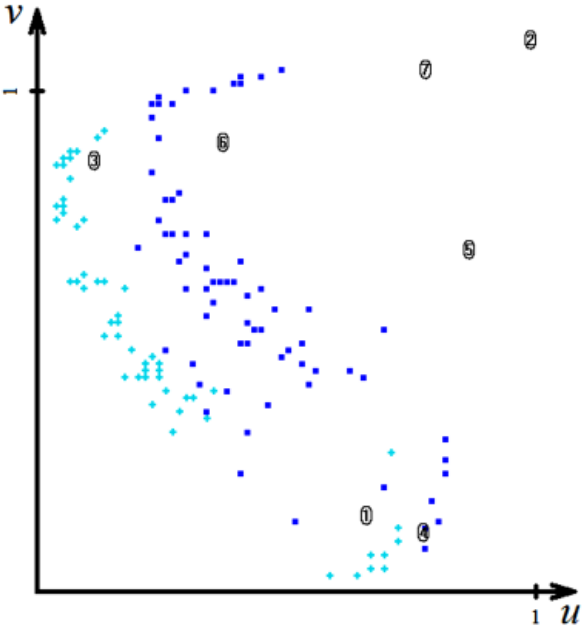

Fig. 6. View of seven-dimensional data representing two various coal type by parameter ITER = 12300. By symbol (■) the images of points representing samples of coal, type 31 were marked, by (+) – samples of coal, type 34.2. Digit of value *i* means the relevance point representing i^{th} coordinate

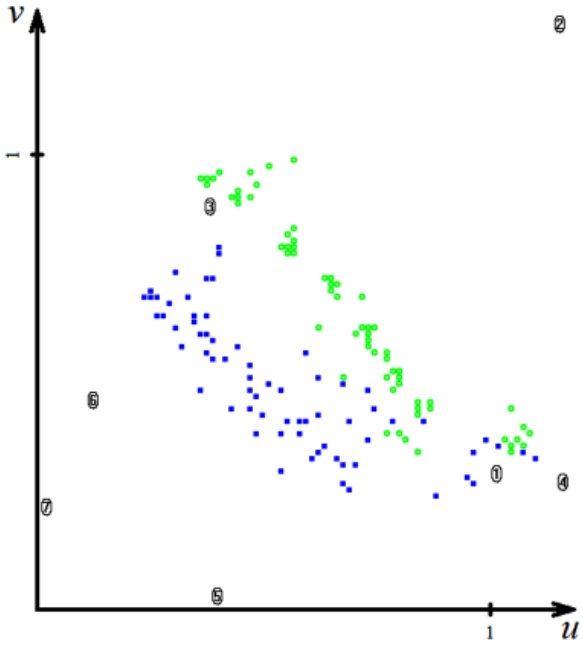

Fig. 7. View of seven-dimensional data representing two various coal type by parameter ITER = 2130. By symbol (■) the images of points representing samples of coal, type 31 were marked, by (o) – samples of coal, type 35. Digit of value *i* means the relevance point representing i^{th} coordinate

It is worthy to notice that algorithm of visualization by means of relevance maps does not use information of 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 properties of these data noticed by the algorithm.

5. Conclusions

The conducted experiments based on visualization of seven-dimensional data by means of relevance maps allowed to get the following conclusions:

1. Multidimensional visualization by means relevance maps allows to state that information contained in analyzed seven-dimensional data is sufficient to proper classification of coal types 31, 34.2 and 35.
2. Imagination of three types of coal on one Figure allowed to state that images of data points representing samples of coal of certain type gather in clusters, which can be separated on almost whole area of the Figure. However, in some parts of space images of points representing various types of coal overlapped. That is why on the basis of such Figure it was impossible to state that analyzed data allowed proper classification of coal types.
3. Only presentation of data representing various coal types in pairs allowed to obtain clear results. They allowed to state that images of points representing samples of coal of certain

type gather in clusters which can be separated. From this it occurred that data contained information sufficient to proper coal classification.

4. Clearness of results grows with more adequate fitness of distance D_{ij} between images of points and relevance points in two-dimensional space to proper coordinates of seven-dimensional data vector x_j (together with growth of parameter ITER).
5. Clearness of obtained results depends highly on accepted parameters.
6. Problem occurring by such visualization is necessity of selection of parameters in purpose of achieving view which clearly presents information searched by the observer. It is worthy to mention that during conducted experiments the views were obtained by the values of parameter ITER from 1 to 20000. Experiments were carried out multiple times by repeated generating of random input values and by other composition of random initial values of points images. Sometimes it lead to obtain more clear results. The observations presented in the paper were the clearest of all achieved results.

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, 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., 1992. *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.

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, Vol. 59, p. 445-450.

Jamróz D., 2014a. *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, Switzerland, Springer International Publishing, Vol. 242, p. 443-450.

Jamróz D., 2014b. *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., 2014c. *Application of multidimensional data visualization by means of multidimensional scaling to qualitative classification of various types of coal*. Arch. Min. Sci., paper in printing, 2014c.

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., 2012. *Multi-parameter data visualization by means of Principal Component Analysis (PCA) in qualitative evaluation of various coal types*. Physicochemical Problems of Mineral Processing, Vol. 50(2), p. 575-589.

Niedoba T., 2013b. *Statistical analysis of the relationship between particle size and particle density of raw coal*. Physicochemical Problems of Mineral Processing, Vol. 49(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., 2011. *Technological Issues Of High-Pressure Grinding Rolls Operation In Ore Comminution Processes*. Arch. Min. Sci., Vol. 56, No 3, p. 517-526.

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.

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., 2010. *Methodology of statistical research*. Folia Scientiarum Universitatis Technicae Resoviensis, Management and Marketing, Vol. 17(4), No. 272, p. 149-157 [in Polish].

Szostek R., 2012. *Modification of Holt's model exemplified by the transport of goods by inland waterways transport*. Folia Scientiarum Universitatis Technicae Resoviensis, Management and Marketing, Rzeszów, Vol. 19(4), No 285, p. 161-168.

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

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

Received: 21 October 2013