The Analysis of the Effectiveness of the Perspective-based Observational Tunnels Method by the Example of the Evaluation of Possibilities to Divide the Multidimensional Space of Coal Samples

Dariusz Jamroz

AGH University of Science and Technology, Department of Applied Computer Science, al. A. Mickiewicza 30, 30-059 Krakow, Poland jamroz@agh.edu.pl

Abstract. Methods of qualitative analysis of multidimensional data using visualization of this data consist in using the transformation of a multidimensional space into a two-dimensional one. In this way, multidimensional complicated data can be presented on a two-dimensional computer screen. This allows to conduct the qualitative analysis of this data in a way which is the most natural for people, through the sense of sight. The application of complex algorithms targeted to search for multidimensional data of specific properties can be replaced with such a qualitative analysis. Some qualitative characteristics are simply visible in the two-dimensional image representing this data. The new perspective-based observational tunnels method is an example of the multidimensional data visualization method. This method was used in this paper to present and analyze the real set of seven-dimensional data describing coal samples obtained from two hard coal mines. This paper presents for the first time the application of perspective-based observational tunnels method for the evaluation of possibilities to divide the multidimensional space of coal samples by their susceptibility to fluidal gasification. This was performed in order to verify whether it will be possible to indicate the possibility of such a division by applying this method. Views presenting the analyzed data, enabling to indicate the possibility to separate areas of the multidimensional space occupied by samples with different applicability for the gasification process, were obtained as a result.

Keywords: Multidimensional data analysis · Data mining · Multidimensional visualization · Observational tunnels method · Multidimensional perspective · Fluidal gasification.

1 Introduction

The qualitative analysis of multidimensional data constitutes a valuable, practical and increasingly used tool for the analysis of real data. It enables to obtain information on characteristics which is directly visible in the two-dimensional

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_64

image representing the multidimensional data, the existence of which we could not even suspect. Various methods of visualization of multidimensional data are used for such an analysis. The perspective-based observational tunnels method used in this paper was presented for the first time in the paper [1]. Its effectiveness for real data placed in the 5-dimensional space of characteristics obtained as a result of the reception of printed text was proven there. It turned out that during the construction of this type of image recognition systems, this method enables to indicate the possibility to separate individual classes in the multidimensional space of characteristics even when other methods fail. During the analysis of 7-dimensional data containing samples belonging to 3 classes of coal, the perspective-based observational tunnels method came first in the ranking of various methods in view of the readability of results [1]. The purpose of this paper is to verify whether this new method is also effective in the case of different real data. This paper presents for the first time the application of perspective-based observational tunnels method for the evaluation of possibilities to divide the multidimensional space of coal samples by their susceptibility to fluidal gasification. This was performed in order to verify whether it will be possible to indicate the possibility of such a division by applying this method. This method has never been applied for such a purpose before. Such an investigation has also a significant practical importance. Indicating the possibilities to divide the space of samples into areas with different applicability to the fluidal gasification allows to conclude that the selected characteristics of coal are sufficient for the correct recognition process of samples of coal more and less susceptible to fluidal gasification. The fluidal gasification itself is in turn significant from the perspective of the coal-based power industry which at the same time emits minimal amounts of pollution.

2 Related Papers

An example of another method of multidimensional data visualization applied during the qualitative analysis is PCA [2, 3] using the orthogonal projection on vectors representing directions of the biggest variation of data. For multidimensional visualization, autoassociative neural networks [4, 5] and Kohonen maps [6, 7] are also used. The method of parallel coordinates [8] consists in placing n axes in parallel representing n dimensions on the plane. In the method of star graph [9], n axes going radially outward from one point are placed on the plane. Multidimensional scaling [10] transforms a multidimensional space into a two-dimensional one in such a way that the mutual distances between each pair of points in the two-dimensional image is as close as possible to their mutual distance in the multidimensional space. In the method of relevance maps [11], n special points $F_1, F_2, ..., F_n$ representing the individual axes are placed on the plane representing the screen. These points and points belonging to the data set are distributed in the two-dimensional image in such a way that the distance from every data point to point F_k is as close as possible to the value of the k-th coordinate of a given data point.

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_64

3 Perspective-based Observational Tunnels Method

The perspective-based observational tunnels method is a new method. It was first presented in the paper [1]. It intuitively consists in the prospective parallel projection with the local orthogonal projection. In order to understand its idea, the following terms must be introduced [1]:

Definition 1. *Observed space X is defined as any vector space, over field F of real numbers, n-dimensional, $n \geq 3$, with a scalar product.*

Definition 2. *Let $p_1, p_2 \in X$ - be linearly independent, $w \in X$. Observational plane $P \subset X$ is defined as:*

$$P = \delta(w, \{p_1, p_2\}) \tag{1}$$

where:

$$\delta(w, \{p_1, p_2\}) \stackrel{def}{=} \{x \in X : \exists \beta_1, \beta_2 \in F, \text{such that } x = w + \beta_1 p_1 + \beta_2 p_2\} \tag{2}$$

The two-dimensional computer screen will be represented by vectors p_1, p_2 in accordance with the above definition.

Definition 3. *The direction of projection r onto the observational plane $P = \delta(w, \{p_1, p_2\})$ is defined as any vector $r \in X$ if vectors $\{p_1, p_2, r\}$ are an orthogonal system.*

Definition 4. *The following set is called hypersurface $S_{(s,d)}$, anchored in $s \in X$ and directed towards $d \in X$:*

$$S_{(s,d)} \stackrel{def}{=} \{x \in X : (x - s, d) = 0\} \tag{3}$$

Definition 5. *A tunnel radius of point $a \in X$ against observational plane $P = \delta(w, \{p_1, p_2\})$ is defined as:*

$$b_a = \psi \xi r + a - w - (1 + \psi)(\beta_1 p_1 + \beta_2 p_2) \tag{4}$$

where:

$$\psi = \frac{(w - a, r)}{\xi(r, r)} \tag{5}$$

$$\beta_1 = \frac{(\psi \xi r + a - w, p_1)}{(1 + \psi)(p_1, p_1)} \tag{6}$$

$$\beta_2 = \frac{(\psi \xi r + a - w, p_2)}{(1 + \psi)(p_2, p_2)} \tag{7}$$

$r \in X$-direction of projection onto observational plane P,
$\xi \in (0, \infty)$ - coefficient of perspective.

The procedure of drawing each point a based on the perspective-based observational tunnels method consists in determining the *tunnel radius* b_a and verifying whether scalar product (b_a, b_a) is smaller than the assumed value of $b_a max$. Additionally, it must be verified whether the *distance of projection* ψ is smaller than the assumed value of ψ_{max}. If the above conditions are met, then the point should be drawn on the screen in position with coordinates β_1, β_2. In the opposite case, the point is not drawn.

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_64

4 Experiment Results

Data to be analyzed was obtained from two hard coal mines, which resulted in obtaining 99 samples in total. Thanks to the upgrading process and chemical analysis, the following 7 values were obtained for each sample: the total sulphur content, hydrogen content, nitrogen content, chlorine content, total coal content, heat of combustion and ash content. In this way, a set of 7-dimensional data was obtained. Based on the Technological applicability card for coals from among the analyzed 99 samples, only 18 samples were marked as those which can effectively be subjected to gasification. In order to conduct the qualitative analysis of the multidimensional data obtained in this way through its visualization, the computer system based on the algorithm presented in the previous point was developed. It was decided to verify whether it will be possible to indicate the possibility to divide the obtained 7-dimensional space of samples into areas with different applicability to the fluidal gasification process using the perspective-based observational tunnels method. The obtained results are presented in Fig. 1-3.

Figure 1 presents the view of the discussed 7-dimensional data with the division into samples of coal more and less susceptible to gasification. It is clearly visible in the figure that images of points representing samples of coal more and less susceptible to gasification occupy separate areas of the figure. It is thus possible to determine the boundary dividing views of points representing different degrees of susceptibility to gasification in the figure. This in turn entails that it is possible to divide areas of the 7-dimensional space occupied by samples with different applicability to the gasification process. Such a possibility results from the existence of the representation used for the visualization, by means of which such a division is possible, as can be seen in the figure. It follows from the above that the perspective-based observational tunnels method enables to indicate the possibility to divide the space of samples into areas with different applicability to the fluidal gasification process. Figure 2 presents the view of the discussed 7-dimensional data with the division into samples of coal more and less susceptible to gasification with the omission of the condition concerning the chlorine content. Also here, despite the omission of the condition concerning the chlorine content, it is clearly visible that images of points representing samples of coal more and less susceptible to gasification occupy separate areas of the figure. Also here, it is possible to determine the boundary dividing views of points representing different degrees of susceptibility to gasification in the figure. This in turn entails that it is possible to divide areas of the 7-dimensional space occupied by samples with different applicability to the gasification process with the omission of the condition concerning the chlorine content. It follows from the above that the perspective-based observational tunnels method enables to indicate the possibility to divide the space of samples into areas with different applicability to the fluidal gasification process even with the change in conditions specifying this applicability. In this particular case, it is especially important, because the chlorine content does not influence the effectiveness of this gasification and it only influences the degree of contamination which appears as a result of

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_64

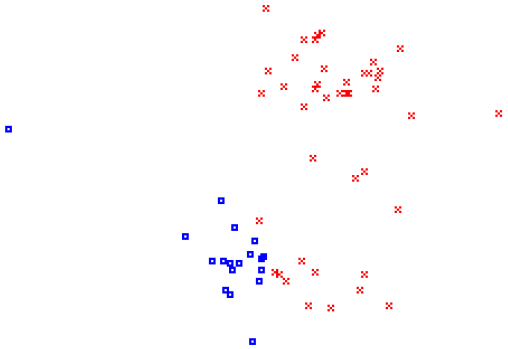

Fig. 1. Images of 7-dimensional points representing samples of coal with lower susceptibility to gasification are marked with symbol x; samples of coal more susceptible to gasification are marked with a circle (o). All points representing samples of coal more susceptible to gasification are visible and may be easily separated from other points.

gasification. However, the allocation of samples changes completely. Comparing Fig. 1 and Fig. 2, it is visible that, in Fig. 1, only 18 samples can be effectively subjected to gasification, while in Fig. 2, with the omission of the condition concerning the chlorine content, from among the same analyzed 99 samples of coal, as many as 78 samples can be effectively subjected to gasification. Figure 3

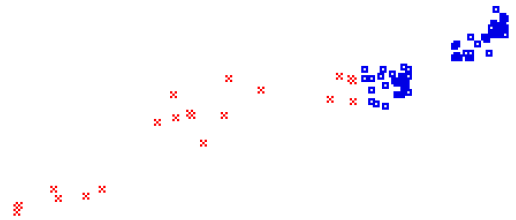

Fig. 2. Images of 7-dimensional points representing samples of coal with lower susceptibility to gasification with the omission of the condition concerning the chlorine content are marked with symbol x; samples of coal more susceptible to gasification are marked with a circle (o).

presents the view of the discussed data with the division according to the place of extraction. It is clearly visible that images of points representing samples of coal with a different place of extraction occupy separate areas of the figure. It is thus possible to determine the boundary dividing views of points representing different places of coal extraction. This in turn entails that it is possible to divide areas of the 7-dimensional space occupied by samples with different places of coal extraction. It follows from the above that the perspective-based observa-

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_64

tional tunnels method enables to indicate the possibility to divide the space of samples into areas with different places of coal extraction. Data analyzed in the

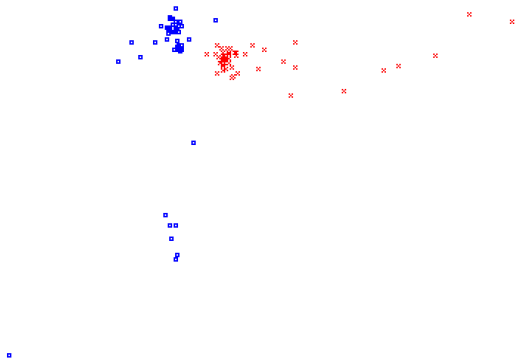

Fig. 3. The view of 7-dimensional data with the division according to the place of coal extraction. Samples coming from different coal mines were marked with different symbols.

paper was 7-dimensional. The perspective-based observational tunnels method can however be used on data with any large number of dimensions. The number of keys necessary to change observational parameters rapidly growing along with the number of dimensions is only a certain limitation [1]. The next papers can focus on verifying the effectiveness of the discussed method on data in the cases in which other methods fail. It can concern both real data and artificially generated multidimensional data.

5 Conclusions

The qualitative analysis of 7-dimensional data using the perspective-based observational tunnels method enabled to draw the following conclusions:

1. It was found that images of points representing samples of coal more and less susceptible to gasification occupy separate areas of the figure. It is thus possible to determine the boundary dividing views of points representing different degrees of susceptibility to gasification in the figure. This in turn entails that it is possible to divide areas of the 7-dimensional space occupied by samples with different applicability to the gasification process. It follows from the above that the perspective-based observational tunnels method enables to indicate the possibility to divide the space of samples into areas with different applicability to the fluidal gasification process.
2. It was found that also with the omission of the condition concerning the chlorine content, images of points representing samples of coal more and less susceptible to gasification occupy separate areas of the figure. Also here, it

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_64

is possible to determine the boundary dividing views of points representing different degrees of susceptibility to gasification in the figure. This in turn entails that it is possible to divide areas of the 7-dimensional space occupied by samples with different applicability to the gasification process with the omission of the condition concerning the chlorine content. It follows from the above that the perspective-based observational tunnels method enables to indicate the possibility to divide the space of samples into areas with different applicability to the fluidal gasification process even with the change in conditions specifying this applicability.

3. It was found that images of points representing samples of coal with a different place of extraction occupy separate areas of the figure. It is thus possible to determine the boundary dividing views of points representing different places of coal extraction. This in turn entails that it is possible to divide areas of the 7-dimensional space occupied by samples with different places of coal extraction. It follows from the above that the perspective-based observational tunnels method enables to indicate the possibility to divide the space of samples into areas with different places of coal extraction.

References

1. Jamroz, D.: The Perspective-based Observational Tunnels Method: a New Method of Multidimensional Data Visualization. Information Visualization 16(4), 346-360 (2017)
2. Jamroz, D., Niedoba, T.: Comparison of selected methods of multi-parameter data visualization used for classification of coals. Physicochemical Problems of Mineral Processing 51(2), 769-784 (2015)
3. Jolliffe, I.T.: Principal Component Analysis, Series: Springer Series in Statistics. 2nd ed. Springer, NY (2002)
4. Aldrich, C.: Visualization of transformed multivariate data sets with autoassociative neural networks. Pattern Recognition Letters 19(8), 749-764 (1998)
5. Jamroz, D.: 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 50(2), 719-734 (2014)
6. Kohonen, T.: Self Organization and Associative Memory. Springer-Verlag (1989)
7. Jamroz, D., Niedoba, T.: Application of multidimensional data visualization by means of self-organizing Kohonen maps to evaluate classification possibilities of various coal types. Archives of Mining Sciences 60(1), 39-51 (2015)
8. Inselberg, A.: Parallel Coordinates: VISUAL Multidimensional Geometry and its Applications. Springer (2009)
9. Akers, S.B., Horel, D., Krisnamurthy, B.: The Star graph: An attractive alternative to the n-cube. In: Proc. International Conference On Parallel Processing, pp. 393-400. Pensylvania State University Press (1987)
10. Kim, S.S., Kwon, S., Cook, D.: Interactive visualization of hierarchical clusters using MDS and MST. Metrika 51, 39-51 (2000)
11. Assa, J., Cohen-Or, D., Milo, T.: RMAP: a system for visualizing data in multidimensional relevance space. Visual Computer 15(5), 217-34 (1999)

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_64