Evaluating Multi-Dimensional Visualizations for Understanding Fuzzy Clusters

Ying Zhao, Feng Luo, Minghui Chen, Yingchao Wang, Jiazhi Xia,
Fangfang Zhou, Yunhai Wang, Yi Chen and Wei Chen

Abstract—Fuzzy clustering assigns a probability of membership for a datum to a cluster, which veritably reflects real-world clustering scenarios but significantly increases the complexity of understanding fuzzy clusters. Many studies have demonstrated that visualization techniques for multi-dimensional data are beneficial to understand fuzzy clusters. However, no empirical evidence exists on the effectiveness and efficiency of these visualization techniques in solving analytical tasks featured by fuzzy clusters. In this paper, we conduct a controlled experiment to evaluate the ability of fuzzy clusters analysis to use four multi-dimensional visualization techniques, namely, parallel coordinate plot, scatterplot matrix, principal component analysis, and Radviz. First, we define the analytical tasks and their representative questions specific to fuzzy clusters analysis. Then, we design objective questionnaires to compare the accuracy, time, and satisfaction in using the four techniques to solve the questions. We also design subjective questionnaires to collect the experience of the volunteers with the four techniques in terms of ease of use, informativeness, and helpfulness. With a complete experiment process and a detailed result analysis, we test against four hypotheses that are formulated on the basis of our experience, and provide instructive guidance for analysts in selecting appropriate and efficient visualization techniques to analyze fuzzy clusters.

Index Terms—Evaluation, multi-dimensional visualization, fuzzy clustering, parallel coordinate plot, scatterplot matrix, principal component analysis, radviz

1 Introduction

Fuzzy clustering [8,22], which is also known as soft clustering, accepts the fact that clusters in data are usually not completely well separated and assigns a membership degree (MD) between 0 and 1 for each datum to every cluster. Traditional hard clustering sorts a datum into a specific cluster without considering uncertainty. Therefore, fuzzy clustering is more practical than traditional hard clustering for many real-world scenarios, such as medical diagnosis [40], weather forecasting [9], and online music services [41].

Fuzzy clusters can be expressed as a MD matrix in which rows and columns describe data items and clusters, respectively. A cell indicates the MD of a datum to a cluster. When the matrix contains numerous data items and a plurality of clusters, it would become complex multi-dimensional data, which makes obtaining insights from the fuzzy clusters challenging for analysts. Many studies have demonstrated that visualization techniques for multi-dimensional data are significantly beneficial to understand fuzzy clusters [1,5,12]. With the assistance of interpretable and interactive methods, analysts are able to explore the answers of a series of complex analysis questions [17,36,50,63]. For example, what is the clustering structure of a fuzzy clustering result? What mutual relationships exist between the clusters of interest? Is a clustering result acceptable?

Empirically, one visualization technique performs well only on a particular analysis task [16,24,34,37]. Therefore, evaluating the ability of multi-dimensional visualization techniques is essential to provide practical guidelines that can help analysts in selecting the appropriate techniques when faced with various analysis tasks. To our knowledge, no previous work has evaluated the ability of the techniques to support the tasks involved in analyzing fuzzy clusters. Most existing approaches are qualitative studies that examine general tasks for multi-dimensional data analysis, such as value retrieval [32,58], class separability [24,31], and correlation judgment [34,53]. Although these general tasks are also required in fuzzy clusters analysis, many special tasks featured by fuzzy clusters are entirely different.

To address this research gap, we conduct a controlled experiment to evaluate multi-dimensional visualization techniques in analyzing fuzzy clusters. First, we carefully select four classic multi-dimensional visualization techniques, namely, parallel coordinate plot (PCP), scatterplot matrix (SPM), principal component analysis (PCA), and radial coordinate visualization (Radviz). Then, we divide the four techniques into two categories: lossy and lossless, and design consistent visual encodings and interactions to evaluate them with uniform comparisons. To guide our experiment design, we carefully define the analytical tasks specific to fuzzy clusters analysis, expand the tasks into representative questions, and formulate four hypotheses on the basis of our experience in practicing the four techniques.

Seven datasets that are frequently used in fuzzy clustering are selected as experiment data. Fifteen graduate students with diverse academic backgrounds are recruited as experiment subjects. We design objective questionnaires to present the tasks that the volunteers have to solve, and observe the accuracy and time of the volunteers using the four techniques to complete the tasks. We also design subjective questionnaires to collect the subjective experience of the volunteers. After preparing the experiment data, volunteers, and equipment, we design a complete experiment process, including volunteer training, formal experiment, and interview.

We record two objective metrics (i. e., accuracy and time) and four subjective metrics (i. e., satisfaction, ease of use, informativeness, and helpfulness) as experiment results, and perform a comprehensive statistical analysis using Shapiro-Wilk, non-parametric Friedman, and Tukey's HSD tests. The analysis results of the two objective metrics and the satisfaction measured with the objective questionnaires suggest that one hypothesis is fully confirmed, but the others are only partially confirmed. The analysis results of the other three subjective metrics measured with the subjective questionnaires indicate which visualization technique has the best overall usability when performing the tasks. After synthesizing all the analysis results, we summarize the ability of

- *Ying Zhao, Feng Luo, Minghui Chen, Yingchao Wang, Jiazhi Xia and Fangfang Zhou are with Central South University. E-mail: {zhaoying,luofeng365,minghui,0903150322,xiajiazhi,zff}@csu.edu.cn.*
- *Yunhai Wang is with Shandong University. E-mail: cloudseawang@gmail.com*
- *Yi Chen is with Beijing Technology and Business University. E-mail: chenyi@th.btbu.edu.cn*
- *Wei Chen is with State Key Lab of CAD & CG, Zhejiang University. E-mail: chenwei@cad.zju.edu.cn.*
- *Jiazhi Xia and Wei Chen are corresponding authors.*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

the four techniques to support our defined tasks and questions. Finally, we discuss the experience we gained from the experiment and present the limitations of this evaluation and further directions.

2 Related Work

2.1 Visualization of Fuzzy Clustering Analysis

Many fuzzy clustering algorithms have been developed ever since the concepts of fuzzy sets and fuzzy partition were first proposed by Zadeh [62] and Ruspini [47]. Fuzzy c-means clustering algorithm (FCM) [6] is the most widely known algorithm because of its simple design and easy conversion into an optimization problem. In this work, we use the FCM algorithm to generate fuzzy clusters for our evaluation. The FCM assigns a probability of membership for each datum to each cluster, which reflects the fuzzy relations among data items and clusters. However, this feature increases the complexity of understanding the clustering result. Klawonn et al. [30] pointed out that multi-dimensional visualization can facilitate the interpretation and understanding of the result. From then on, a variety of techniques have been proposed to analyze fuzzy clusters visually and interactively.

Multi-dimensional visualization techniques, which are based on dimensionality reduction (DR) methods, are frequently used to generate an observable overview of fuzzy clusters. Mao and Jain [38] utilized PCA to map a fuzzy clustering result onto a 2D plane while preserving the distances between pairwise data items and to assist analysts in perceiving the similarities between data items. Abonyi and Babuska [1] proposed a modified fuzzy sammon mapping (FSM) that preserves the distances between data items and cluster centers instead of pairwise data items, which associates data items with clusters presented to analysts. Sharko et al. [50, 51] discovered that Radviz has advantages in visualizing fuzzy clusters because the position of a data item in Radviz can directly indicate its membership distribution to all clusters. Hoppner and Klawonn [25] combined MDS with spheres to help analysts in observing the correlation of cluster memberships. Rueda and Zhang [46] proposed a novel method that maps data items into an irregular hyper-tetrahedron to reflect inter-cluster relationships in a spatial manner.

DR-based visualization inevitably results in the loss of information. To avoid information loss, some researchers proposed lossless visualization techniques to present the complete information of fuzzy clusters. Berthold and Hall [5] visualized a fuzzy clustering result in PCP to enable analysts to investigate the membership distributions of clusters intuitively. Gasch and Eisen [21] employed a heatmap to visualize the sorted membership matrix of a fuzzy clustering result to help analysts discover similarities between clusters as correctly as possible. On the basis of the design of SPM, Cao et al. [7, 35] presented a novel technique called UnTangle Map, which weaves an interactive mesh of triangle-style scatterplots and is suitable for analyzing fuzzy similarities and correlations among clusters.

All the above approaches have their own technical strengths in fuzzy clusters analysis. To our knowledge, no previous work has systematically evaluated their strengths and weaknesses. Therefore, this work selects PCA and Radviz from the above DR-based techniques and PCP and SPM from the obove lossless techniques as representatives to perform an evaluation.

2.2 Evaluation of Multi-dimensional Visualization

A number of visualization techniques are used to visualize multi-dimensional data. Keim and Kriegel [29] classified these techniques into six categories. Most of the techniques that are used for fuzzy clusters analysis belong to the first category, namely, *geometric projection*. The four popular techniques that belong to this category are used in the evaluation and divided into two groups, namely, *lossy* (PCA and Radviz) and *lossless* (PCP and SPM).

Visualizing multi-dimensional data on a 2D/3D plane/space is an inherently ill-posed problem; thus, no single method has no drawbacks. Researchers have gradually realized the importance of comparing and evaluating different techniques. For example, for the lossy approaches, Sedlmair et al. [49] explored the performance of four frequently used DR methods on class separability, and Rubio-Sánchez et al. [45] compared two classic radial projection techniques (i. e. Radviz and star coordinates) on their mapping mechanisms and layout performances. For the lossless techniques, Holten and Wijk [24] evaluated the time and correctness performances of nine PCP variations for cluster identification, and Li et al. [34] compared the effectiveness of PCP and SPM for correlation judgment.

In addition to comparing multi-dimensional visualization techniques from a technical standpoint, evaluations that involve real-world application scenarios are important. These evaluations, along with paying attention to the immediate needs of the actual users and reducing abstract technical comparisons, aim to provide practical guides that help users select the most appropriate technique when faced with an analysis task. A few evaluation studies have been conducted in real-world scenarios. Dimara et al. [14] rigorously evaluated the ability of PCP, SPM, and tabular visualizations to support decision making while selecting holiday travel packages as its application scenario. Marghescu [39] investigated various multi-dimensional visualization techniques with respect to their effectiveness in solving the problem of financial competitor benchmarking. Rzeźniczak [48] examined many techniques to obtain the best technique in facilitating disease recognition based on reference to medical patterns and data of patient's condition. This work focuses on the scenario of fuzzy clusters analysis. This is considered a new scenario different from the above scenario-based evaluations.

As stated in Section 2.1, a number of works have presented their methods of visualizing multi-dimensional fuzzy-clustered data. Some of them provided evaluations in explaining the pros and cons of their proposed methods. Abonyi and Babuska [1] compared the cluster validity performance of FSM and PCA. Cao et al. [7] conducted a brief evaluation to illustrate the advantages of UnTangle Map on PCP, SPM, and PCA. However, such evaluations often tend to only examine several analysis tasks that can highlight the outstanding features of their proposed methods. So far, no study has conducted systematical evaluation followed by carefully summarized analysis tasks controlled under rigorous experimental settings to understand fuzzy clusters.

3 Technique Design

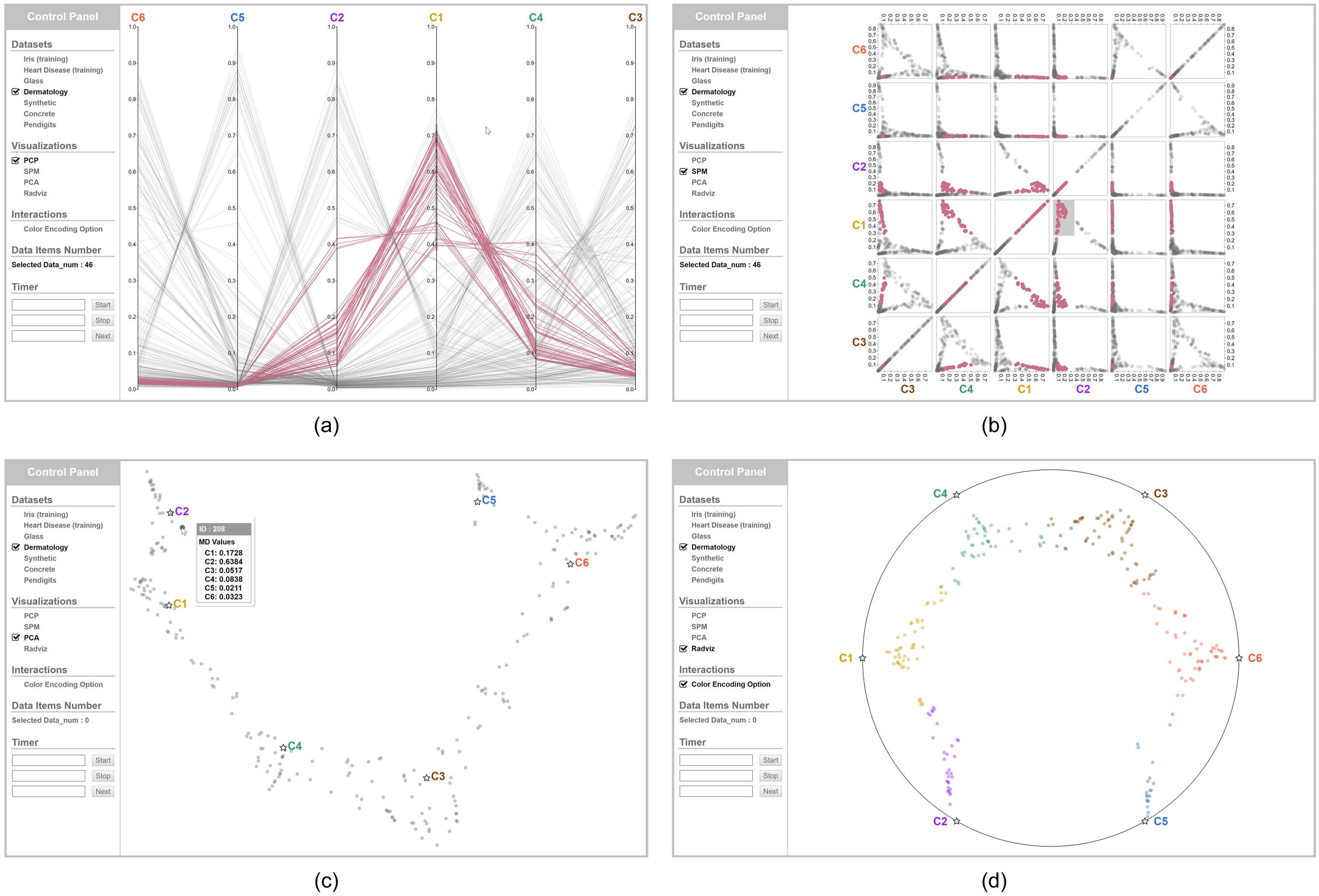
We focus on four visualization techniques, namely, PCP, SPM, PCA, and Radviz. The technique selection is mainly based on three principles: (1) Universality, the candidates must be popular in multi-dimensional data visualization and fuzzy clusters analysis. All the four selected techniques are very common multi-dimensional visualization techniques and have appeared in the literatures related to fuzzy clusters analysis; (2) Diversity, the candidates must cover the lossy and lossless techniques introduced in Section 2.1 to expand the diversity of techniques. Among the four selected techniques, PCA and Radviz are representatives of lossy techniques, while PCP and SPM are representatives of lossless techniques; (3) Comparability, the candidates of the same category should have distinct strategies for embedding data in a 2D plane. In terms of our choice, PCA and Radviz have entirely distinct multi-dimensional data projection strategies, and the visual mappings of PCP and SPM are also quite different.

3.1 Data Model

We use the FCM algorithm proposed by Bezdek et al. [6] to generate fuzzy clusters data (CData) from the original data (OData). OData and CData can be described as matrices. In the OData matrix, each row and column represent a data item and an original dimension (ODim), respectively. In the CData matrix, each row represents a data item, but each column represents a cluster (CDim). The value of each cell indicates the MD to which a specific data item belongs to a certain cluster. The FCM algorithm has two initialization parameters, namely, m for controlling the extent of cluster overlap and c for identifying the number of target clusters. The parameter m is kept consistent for all examined datasets, and c equals the ground truth.

3.2 Visual Encodings

Using various encodings in different visualizations may result in confusion or misinterpretation among volunteers [44]. In this study, we



(a) (b) (c) (d)

Figure 1. Four evaluated visualization techniques and three provided interactions: (a) PCP with range selecting interaction; (b) SPM with range selecting interaction; (c) PCA with hovering interaction; (d) Radviz with color encoding option.

attempt to design consistent visual encodings for all the four techniques as much as possible.

PCP: We use the conventional visual encodings of PCP [26], as shown in Figure 1(a). Each vertical axis represents a CDim. All axes are arranged in parallel. Each polyline represents a data item. The position of a polyline on each axis is proportional to its MD for that CDim.

SPM: We use the generalized SPM developed by Cleveland [10], as shown in Figure 1(b). It assigns CDims to vertical axes and displays all possible pairwise CDims with multiple scatterplots. Data items are drawn as dots in the Cartesian spaces defined by CDims.

PCA: We use the PCA proposed by Jollife [28] to project the multi-dimensional CData into a 2D plane where a dot represents a data item, as shown in Figure 1(c). Conventional PCA projection fails to provide any CDim-related visual information. A pentagram is added to the center of each cluster to be consistent with the other three visualizations. The position of the cluster center is obtained in the original data space and then projected to the 2D plane.

Radviz: We use the original Radviz proposed by Hoffman et al. [23], as shown in Figure 1(d). Each anchor, represented by a pentagram on the circumference, represents a CDim. Each dot in the circle represents a data item, and all the data items are mapped into the circle according to the spring forces from the anchors.

In addition to the above basic visual encodings, other consistent visual designs are found in all the visualizations. First, the default color of the dots in SPM, PCA, and Radviz is gray, with a default opacity of 0.4 and a diameter of 8 pixels. The default color and opacity of polylines in PC are the same as those of the other visualizations, except for a width of 2 pixels. Second, CDims share the same colors and labels, such as C1, C2, and C3, in all the visualizations. We obtain the order of CDims based on the similarities among them [3, 42, 61] to reduce visual clutter and keep the order consistent in PCP, SPM, and Radviz. Lastly, all the visualizations have the same visible area (2400×1800) because a fair comparison requires that all techniques occupy the same space.

3.3 Interactions

Interactions play important roles in visual analysis. In the evaluation, volunteers will struggle to complete some analytical tasks without any interaction. We thus discreetly provide three basic interactions and ensure that the interactions can reveal uniform information in all the four techniques.

Range selection: This interaction helps volunteers focus on a subset of selected data items. In PCP, it is performed by brushing one or more axes. In SPM, Radviz, and PCA, volunteers can execute and adjust a rectangle brush. In SPM, brushing can be carried out in any scatterplot, and the effects will appear simultaneously on all scatterplots. The opacity of the selected data items will change from the default 0.4 to 1, the color will change from gray to red, and the number will be displayed on the control panel side, as shown in Figure 1(a) and Figure 1(b).

Hovering: Hovering over a data item (polyline in PCP or dot in SPM, PCA, and Radviz) will trigger a tooltip, which presents its detailed information, such as ID and all MDs, as shown in Figure 1(c). Meanwhile, the data item will be highlighted with opacity changing from 0.4 to 1 and a doubled size (width of polyline or diameter of dot).

Table 1. Analytical tasks and representative questions.

Tasks		Questions
Data-oriented Tasks	T1. Information recognition of a data item	Q1. Membership information of a data item: What are the maximum and minimum MDs of a given data item?
		Q2. Stability of a data item: Is a given data item stable? (PS: If a data item belongs to one cluster with a dominant high MD, then it is stable; otherwise, it is unstable.)
	T2. Information recognition of a data group	Q3. Stability of a data group: Are more stable data items than unstable data items presented in a given data group?
Cluster-oriented Tasks	T3. Information recognition of a single cluster	Q4. Membership information of a single cluster: What are the maximum and minimum MDs of a given cluster?
	T4. Information recognition of a cluster group	Q5. Dominant cluster: Does a dominant cluster exist in a fuzzy clustering result? (PS: According to the maximum MD principle, if the number of data items partitioned into a cluster is significantly greater than that of other clusters, then the cluster is a dominant cluster.)
		Q6. Similarities between clusters: Given multiple clusters, do they have similar membership distributions?
		Q7. Correlations between clusters: Given multiple clusters, do they have positive correlations? (PS: If the membership degrees of most data items belonging to two clusters increase or decrease simultaneously, then the two clusters have a positive correlation.)

Hovering over a dot in one scatterplot of SPM will highlight all dots that represent the same data item in the other scatterplots.

Color encoding option: This interaction is only permitted when volunteers are completing the specific question T4-Q5 which is defined in Section 4.1. If this option is triggered, then the color of all the data items will turn into the colors of corresponding CDims according to the maximum MD principle, as shown in Figure 1(d). The maximum MD principle is a strategy that is commonly used to convert soft clustering result into hard clustering result [33]. More specifically, when dividing a data item that belongs to multiple clusters into one cluster, this principle states that the data item should be divided into the cluster that corresponds to the maximum MD of the data item.

4 Tasks and Hypotheses

4.1 Task Definition

Various application scenarios have the requirement of understanding fuzzy clusters. For example, online music providers would like to know the music preferences of their customers so as to provide greater personalization services. Generally, people have different music preferences. Some like jazz only, and others may like jazz and pop with different degrees. This can be considered as a kind of fuzzy clusters where different music types are clusters and the customers are data items. The providers are supposed to utilize visualizations to facilitate the analysis of such fuzzy clusters. They would like to find the answers of some specific queries, such as which types of music are more popular than others, whether the preferences of a particular customer or a group of customers are clear, and which groups of customers have similar preferences.

We have not found a well-defined list of analytical tasks specific to fuzzy clusters visualization. To determine the analytical tasks, we first surveyed the user needs of interpreting and understanding fuzzy clusters, as well as the task taxonomies constructed for multi-dimensional visualization and visual analytics via literature research [2, 15, 54, 57, 59, 60]. Then, we conducted pilot experiments to obtain a tentative list of the analytical tasks. By working closely with a data mining expert, we refined and confirmed the tentative task list repeatedly and set up the completed tasks shown in Table 1.

The tasks are divided into two categories, namely, cluster- and data-oriented tasks. Each category involves two specific tasks for identifying information from two perspectives, namely, a single object (data item or cluster) and a group of objects (data items or clusters). To facilitate our questionnaire design, we further extend the four tasks into some representative questions. For a single data item (T1), we mainly examine whether volunteers can identify its maximum and minimum MDs and its stability when using a visualization technique. For the task of a group of data items (T2), we focus on comparing the number of stable and unstable data items. For a single cluster (T3), we only identify its maximum and minimum MDs. As for a group of clusters (T4), volunteers are asked to answer whether a dominant cluster exists in a fuzzy clustering result and whether multiple given clusters have similar membership distributions and positive correlations.

4.2 Hypotheses Definition

We expect to find the right techniques from the four visualization techniques for the above tasks/questions. To guide the design and analysis of our experiments, we formulated four hypotheses based on our experience in practicing the four techniques.

H1: Lossless techniques will perform better in terms of accuracy than lossy techniques within the group of questions (Q1, Q4, Q6, Q7). We formulate this hypothesis because both lossless visualizations (PCP and SPM) provide graphical axes. We believe that graphical axes can facilitate the identification of the maximum and minimum MDs of a given data item (Q1) or cluster (Q4) and depict the MD distribution of all the data items of each cluster (Q6). Furthermore, the comparative views defined by pairwise axes in PCP and SPM make identifying cluster correlations (Q7) convenient for volunteers.

H2: Radviz will outperform the other three techniques within the group of questions (Q2 and Q3) in terms of accuracy, time, and satisfaction. The radial spring-based projection model [23] of Radviz can gather stable data items near relevant dimension anchors. We thus believe that Radviz facilitates the stability determination of data items.

H3: Lossy techniques will perform better in terms of time and satisfaction than lossless techniques within Q5. Data items are colored according to the maximum MD principle to answer Q5. Thus, the performance comparison is translated to whether the visualizations can clearly present colored data items. Empirically, lossless approaches tend to produce higher visual clutter than lossy ones, thereby causing volunteers to suffer from more cognitive load.

H4: The volunteers who specialize in visualization will perform the best in terms of accuracy and time in all tasks. The basis for this hypothesis is the idea that the volunteers with a visualization background are more familiar with the four visualization techniques and more skilled in manipulating the provided interactions than the volunteers with other backgrounds.

5 Experiment Design

The evaluation of this study aims to explore the performance of the four visualization techniques (PCP, SPM, PCA, and Radviz) in analyzing fuzzy clusters. Thus, a within-subject design is selected with independent variables of the four visualization techniques, five datasets, and seven questions. Each volunteer was assigned all the four visualization techniques and any two datasets, so they had to undergo the experiment under 56 ($4 \times 2 \times 7$) conditions. The order of datasets and visualization techniques was randomized to mitigate learning and fatigue effects. The measured dependent variables included two objective metrics, namely, accuracy (percentage) and time (seconds), and four subjective metrics (seven-point Likert scale), namely, satisfaction, ease of use, informativeness, and helpfulness.

5.1 Data and Questionnaire

We initially gathered 20 frequently used datasets in fuzzy clustering. Each of them has been used in a case study in at least one relevant literature. Then, we selected seven datasets from the 20 candidates for evaluation, as shown in Table 2. Our data selection is mainly based on two principles. The first is to select data with moderate sizes because a large data size will cause visual scalability issues and a small data size will be straightforward. Hence, we selected the datasets with 3 to 10 CDims and 100 to 3,000 data items. The second principle is to construct the chosen datasets with dissimilar fuzzy clustering structures. For example, we reserved the Glass dataset but disregarded the Ecoli dataset because their clustering results present similar polygon-style clustering structures.

For each selected dataset, we provided two types of questionnaires. The first type is an objective questionnaire, which includes seven questions described in Section 4.1. All questions were presented in the order of simple to complex to better prepare volunteers for the difficult tasks at the end. We provided four objective questionnaires, which correspond to the four visualization techniques for each dataset.

The details of objective questions were preset with the representative items or clusters selected from the data to ensure that volunteers can observe the typical clustering characteristics of the data. Take the objective questionnaire of Radviz and the Iris dataset as an example. We specified two data items with completely different stabilities for Q1 and Q2: unstable data item A and stable data item B, as shown in Figure 2, and required volunteers to identify the maximum and minimum MDs and stabilities of the two data items. As the three clusters of the Iris dataset have two types of MD distributions, namely, decline and concave, we assigned a decline cluster C1 and a concave cluster C3 for Q4, and asked volunteers to identify the maximum and minimum MDs of the two clusters. We specified the consistent data items and clusters for the four objective questionnaires of the Iris dataset.

The second type of questionnaire is a subjective questionnaire with three statements for each technique: (1) “I think this visualization is easy to use;” (2) “I think this visualization is informative;” (3) “I think this visualization is helpful.” These three statements correspond to three technical usability metrics: ease of use, informativeness and helpfulness, which are commonly and widely used in the evaluation community [14, 52, 55]. The above two types of questionnaires are provided in the appendix with the Iris dataset as the example.

5.2 Participants and Apparatus

We recruited 15 volunteers (eight males and seven females) as the experiment subjects. All volunteers are not color blind and have normal or corrected-to-normal vision. Their age ranges from 20 to 27 (with an average of 25). In addition, they are all graduate students affiliated with the school of information science and engineering in a university and have diverse academic backgrounds. Five volunteers have a data visualization background, five volunteers have a data mining background, and the other five volunteers come from other majors.

All experiments were conducted on a 27.2-inch Dell OptiPlex 7040 desktop with a 3.4 GHz Intel i7 processor, 8 GB of RAM, and a screen resolution of 3840×2160. A standard wired mouse and keyboard were connected to the desktop to enable the volunteers to interact with experimental software and input their answers.

5.3 Experiment Procedure

The purpose and procedure of the study were explained at the beginning of the experiment. All the volunteers were asked to fill in their basic personal information (name, gender, age, and major). Subsequently, the test supervisor explained the visualization techniques, datasets, and questionnaires to the volunteers and demonstrated all the interactions. After ensuring that the volunteers understood all the visualization techniques and questionnaires correctly, we took two datasets (the fuzzy clustering results of the Iris and Heart Disease datasets) that are not related to the formal experiment as training datasets. We used the two datasets to help the volunteers become familiar with the whole procedure, including the use of the four visualization techniques and answering questions. The volunteers were told to use a keyboard or mouse to input or select their answers to each question.

The formal experiment was conducted after the volunteers had mastered the visualization techniques and familiarized themselves with the questionnaires. Each volunteer needed to use two datasets and four visualization techniques to finish two rounds of experiments, and only one dataset was used in one round. During each round of experiment, we first assigned a dataset and a visualization technique to a volunteer randomly and took out the objective questionnaire that corresponded

Table 2. Datasets used for evaluation.

Datasets	Data Items	Dimensions	Clusters
Iris [4], (for training)	150	4	3
Glass [4]	214	9	6
Dermatology [4]	259	34	6
Heart Disease [4], (for training)	303	14	5
Synthetic [18]	750	12	4
Concrete [4]	1030	9	4
Pendigits [4]	2498	63	10

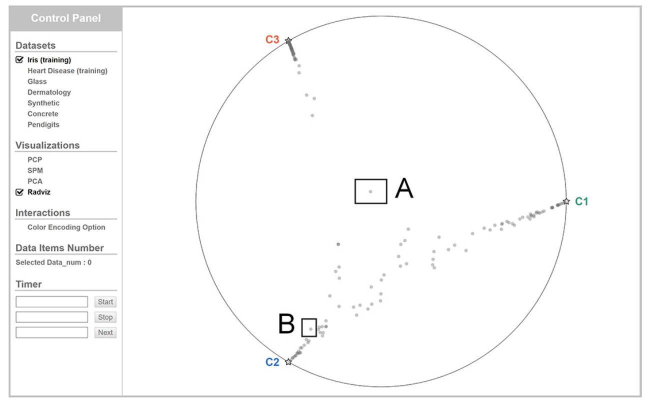


Figure 2. Illustration of selecting two representative data items of the Iris dataset to design the details of objective questions.

to the dataset and the visualization technique. Prior to completing an objective question, the volunteers were asked to read the question carefully. Then, the volunteers pressed the "Start" button to turn on a timer and started solving a question. After the volunteers obtained their answers, they pressed the "Stop" button to shut down the timer and gave the answers. After answering the question, the volunteers must rate their satisfactions with the use of this technique in solving this question on a seven-point Likert scale ranging from 7 (strongly satisfied) to 1 (strongly dissatisfied). Afterwards, the volunteers pressed the "Next" button to start a new question. No strict time limits were imposed during the experiment. After using a visualization technique to complete the objective questionnaire, the volunteers were allowed to rest to avoid fatigue. Later, we assigned another technique and relevant objective questionnaire randomly. Immediately after the volunteers examined all the four techniques, they were provided with a subjective questionnaire and were asked to rate how they agreed with the three statements in the subjective questionnaire based on a seven-point Likert scale ranging from 7 (strongly agree) to 1 (strongly disagree). Then, all the volunteers were required to rest for at least two hours before they started the other round of experiment. Once the volunteers had completed both rounds of experiments, they were encouraged to state their problems and feelings as much as possible. In the formal experiment, we randomized the order of datasets and visualization techniques. We also ensured that each dataset was used six times and each visualization technique was used 30 times.

5.4 Analysis Approach

We fully recorded two objective metrics and four subjective metrics as experiment results. Accuracy and time were measured for each question in the objective questionnaires. We used these two objective metrics to observe the efficiency of each visualization technique in completing each question. Satisfaction was measured to disclose the volunteers' preferences of each visualization when completing each question. We thus asked the volunteers to rate their satisfactions when they completed each question in the objective questionnaires to accurately and timely record their emotions. Ease of use, informativeness and helpfulness were measured with the subjective questionnaires to reveal the overall usability of techniques. We utilized a box plot to mark the outliers when the value of each metric was beyond two standard deviations (upper bound and lower bound). In accordance with standard procedure, the outliers were replaced with medians.

When analyzing experiment results, we first used the Shapiro-Wilk test to examine the normality and found that all results did not follow the normal distributions ($p < 0.05$). Thus, we used a non-parametric Friedman test instead of ANOVA test to examine whether the four visualization techniques have significant differences in the six metrics [13, 43]. If significant differences were found, then we conducted Tukey's HSD test for pairwise comparison of any two techniques. All the tests were performed under the standard significance level $p = 0.05$ to determine the statistical significance of the experiment results. In the pairwise comparison, we applied Bonferroni correction to adjust the significance level. On the basis of the pairwise comparison results, we could rank the four techniques in terms of the six metrics and further investigate why significant differences between them exist.

6 Experiment Results

This section describes the analysis of the experiment results measured with the objective and subjective questionnaires.

6.1 Hypothesis and Objective Questionnaire Analysis

The results measured with the objective questionnaires are shown in Figure 3, Figure 4, and Figure 5. By using the results, we first test against the four hypotheses and then summarize our findings from the perspectives of analytical tasks and techniques.

H1: Influence of graphic axes and interactions. We assumed that lossless techniques will outperform lossy techniques within questions Q1, Q4, Q6, and Q7 in terms of accuracy because PCP and SPM provide graphical axes. This hypothesis can be confirmed partially.

Considering Q1 first, the results are totally different from the hypothesis. The accuracy of the four visualization techniques is greater than 96%, as shown in Figure 3(a), no significant differences among them ($\chi^2(3) = 2.609, p = 0.46 > 0.05$) are found. The hovering interaction is assumed to play a crucial role. The volunteers can easily obtain accurate MDs of a data item from the pop-up tooltip triggered by hovering over a dot or line. The close-to-perfect accuracy of all the techniques in Q1 indicates that the interaction has a significant impact on the performance of the visualization techniques.

The mean accuracy of PCP for Q4 is higher than that of SPM, which partially supports H1. PCP performs the best due to its vertical axes that represent clusters. The volunteers can thus read the MDs of a cluster easily. Both lossy techniques rank second, but no significant pairwise differences between them and PCP are found. This result shows that the volunteers made good use of the hovering interaction, thus addressing the lack of graphic axes of the two lossy techniques. SPM is the last one with significant pairwise differences with any of the other three ($p < 0.05$) because of the much lower resolution of the SPM axes than that of the PCP axes. Therefore, even if SPM has graphic axes, the sight of the volunteers would be blurred when they were identifying information of cluster MDs.

The results of Q6 are completely consistent with H1. The two lossless techniques are significantly better than the two lossy techniques in terms of accuracy, time, and satisfaction. As observed in the experiments, the volunteers were able to fully utilize the graphical axes of PCP and SPM to comprehend the distributions of all the data items on the axes and compare the similarities of the membership distributions of various clusters.

The results of Q7 partially support H1. The mean accuracy of the two lossy techniques (PCA and Radviz) lies between that of SPM and PCP. More precisely, SPM performs well, whereas PCP performs poorly. Significant pairwise differences exist among them ($p < 0.05$, SPM: $\mu = 0.729, \sigma = 0.248$, PCP: $\mu = 0.344, \sigma = 0.262$, PCA: $\mu = 0.521, \sigma = 0.102$, Radviz: $\mu = 0.541, \sigma = 0.310$). SPM and PCP have pairwise axes, which make them more effective in assisting correlation analysis than the two lossy techniques. However, PCP performs worse than expected mainly because of the disabled interactions that allow the volunteers to modify the arrangement of the axes. This result of PCP once again demonstrates the importance of interactions.

H2: Influence of geometric projection mechanisms. We assumed that Radviz's projection mechanism is effective in helping the volunteers determine the stability of data items (Q2 and Q3) in terms of accuracy, time, and satisfaction. This hypothesis is fully confirmed. Radviz has significantly better performances (accuracy, time, and satisfaction) than any other technique. For instance, as shown in Figure 3(b), the time result of Q2 is $Radviz < SPM < PCP < PCA$ ($p < 0.05$, Radviz: $\mu = 19.0, \sigma = 11.361$, SPM: $\mu = 28.4, \sigma = 16.589$, PCP: $\mu = 28.5, \sigma = 17.753$, PCA: $\mu = 28.8, \sigma = 12.338$).

H3: Influence of visual clutter in the visualization results. We assumed that the two lossy techniques are better suited (in terms of time and satisfaction) to identifying the dominant cluster due to the visual clutter issue (Q5). Again, this hypothesis is confirmed partially. Figure 3(b,c) show that the two lossy techniques cost little time and have high satisfaction scores when solving Q5. As expected, SPM significantly costs more time ($\mu = 80.7$ and $\sigma = 43.310$) and has a lower satisfaction score ($\mu = 4.8$ and $\sigma = 1.085$) than the two lossy techniques. Each scatterplot in SPM occupies only a tiny screen space. Thus, the volunteers that suffered from severe visual clutter had to spend more time selecting the data items of interest. Surprisingly, PCP performs almost as well as the two lossy techniques (Time: $\mu = 61.1, \sigma = 19.048$, Satisfaction: $\mu = 5.6, \sigma = 0.669$) because the visual clutter in PCP is mainly located at the bottom of the axes (low MD area). The top areas of the axes are critical in analyzing the dominant cluster.

H4: Influence of volunteers' academic background. We assumed that the volunteers with a visualization background will complete all the questions with the highest accuracy and the least time. This hypothesis can be confirmed partially. When looking into the completion time, the results are consistent with this hypothesis. As shown in Figure 4,

	PCP	SPM	PCA	Radviz
Q1	97.3%	97.9%	96.3%	98.3%
Q2	68.3%	67.5%	66.7%	89.2%
Q3	65.6%	66.7%	63.3%	86.7%
Q4	83.3%	64.2%	78.3%	82.5%
Q5	76.7%	75.0%	80.0%	81.7%
Q6	70.5%	69.5%	50.0%	54.6%
Q7	34.4%	72.9%	52.1%	54.1%

Accuracy (percentage)

(a)

	PCP	SPM	PCA	Radviz
Q1	22.6	20.6	15.2	14.8
Q2	28.5	28.4	28.8	19.0
Q3	97.8	96.9	95.4	66.5
Q4	32.5	30.6	47.7	39.6
Q5	61.1	80.7	48.0	47.1
Q6	40.1	54.1	66.4	65.6
Q7	130.3	60.3	98.1	95.5

Time (seconds)

(b)

	PCP	SPM	PCA	Radviz
Q1	6.4	6.5	6.7	6.8
Q2	5.9	5.8	5.9	6.5
Q3	5.1	5.0	5.2	6.2
Q4	6.1	6.0	5.4	5.4
Q5	5.6	4.8	5.5	5.6
Q6	5.4	5.3	4.2	4.5
Q7	4.9	5.4	3.7	3.7

Satisfaction (seven-point Likert scale)

(c)

Figure 3. Results of mean accuracy (a), mean time (b), and mean satisfaction score (c) in solving each question with the four visualization techniques. Colors indicate groups of no significant pairwise differences, with the winners shown in dark blue, losers in gray, and the ones in between in light blue. Taking Q7 in (a) as an example, the four fields have various colors that reflect the significant differences; the dark blue SPM is the winner, and the gray PCP is the loser, whereas the fields of PCA and Radviz are light blue, which indicates that they are in between and that no significant difference exists between them.

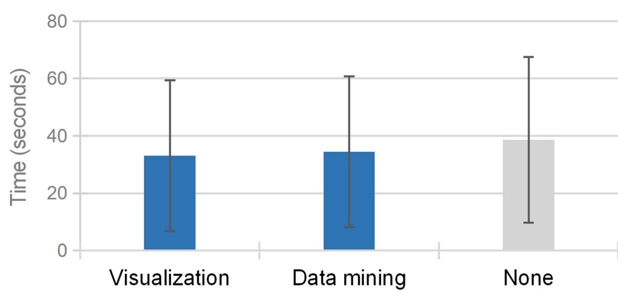


Figure 4. Results of the mean time of the volunteers with different academic backgrounds. Different colors indicate that there are significant differences among corresponding groups. The volunteers with a visualization background or a data mining background have a significant advantage over the volunteers with no background, while no significant difference exists between the former two groups of volunteers.

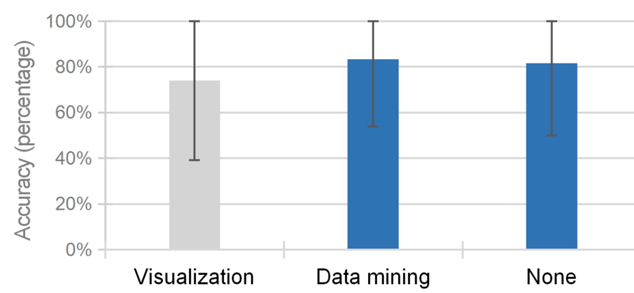


Figure 5. Results of the mean accuracy of the volunteers with different academic backgrounds. Different colors indicate that there are significant differences among corresponding groups. The volunteers with a data mining background or no background have a significant advantage over the volunteers with a visualization background, while no significant difference exists between the former two groups of volunteers.

the volunteers with a visualization background spend the shortest time ($\mu = 33.0, \sigma = 26.290$), followed by the volunteers with a data mining background ($\mu = 34.5, \sigma = 28.902$), and those without any relevant background spend the longest time ($\mu = 38.7, \sigma = 32.252$). As shown in Figure 5, contrary to time, the volunteers with a visualization background gain the lowest accuracy, the volunteers with no background are in between, and the volunteers with a data mining background rank the first. We observed in the experiments that the volunteers with a visualization background were skilled in using interactions and they felt confident in their answers so had not repeatedly verified the answers, whereas the volunteers with a data mining background were interested in exploring the data. One of the volunteers with a visualization background said, "*I'm familiar with this and I don't need to look at the screen over and over again.*" while one volunteer with a data mining background stated, "*It is a little hard to finish all the tasks in a short time, but it's great to explore the data interactively through the system.*"

On the basis of the overall objective questionnaire results, Radviz performs the best among the four techniques. In the three aspects of accuracy, time, and satisfaction, Radviz appears in the top-ranked group 11 out of 21 times, followed by PCP and SPM, both of which appear in the top-ranked group 7 out of 21 times. PCA obtains the worst performance, appearing in the top-ranked group only 5 out of 21 times. For Q1, the four techniques perform almost as well in terms of accuracy with the assistance of the interactions. However, the two lossy techniques require a shorter time and obtain a higher satisfaction score than the two lossless techniques. As for Q2 and Q3, Radviz remarkably outperforms the other three techniques under various accuracy, time, and satisfaction conditions. For Q4, PCA, Radviz, and PCP are similar in accuracy, but PCA and Radviz are slightly worse among the three techniques in terms of time and satisfaction. When it comes to Q5, no significant differences in terms of accuracy are found among the four techniques. SPM requires the longest time and has the lowest satisfaction score among all the techniques. For Q6, the two lossless techniques outperform the two lossy techniques significantly in terms of accuracy, time, and satisfaction. As for Q7, SPM obtains the highest accuracy with shorter time and better satisfaction than the other techniques.

6.2 Subjective Questionnaire Analysis

Figure 6 shows the ratings of the four visualization techniques with regard to the ease of use, informativeness, and helpfulness. Tukey's HSD test reveals the four visualization techniques with different ranks on the three metrics. PCP ($\mu = 5.2, \sigma = 1.440$) and Radviz ($\mu = 4.7, \sigma = 1.557$) are better than SPM ($\mu = 3.8, \sigma = 1.633$) and PCA ($\mu = 3.7, \sigma = 1.688$) for the ease of use ratings. The two lossless techniques (PCP: $\mu = 5.5, \sigma = 1.196$, SPM: $\mu = 5.2, \sigma = 1.289$) significantly outperform the two lossy techniques (Radviz: $\mu = 4.6, \sigma =$

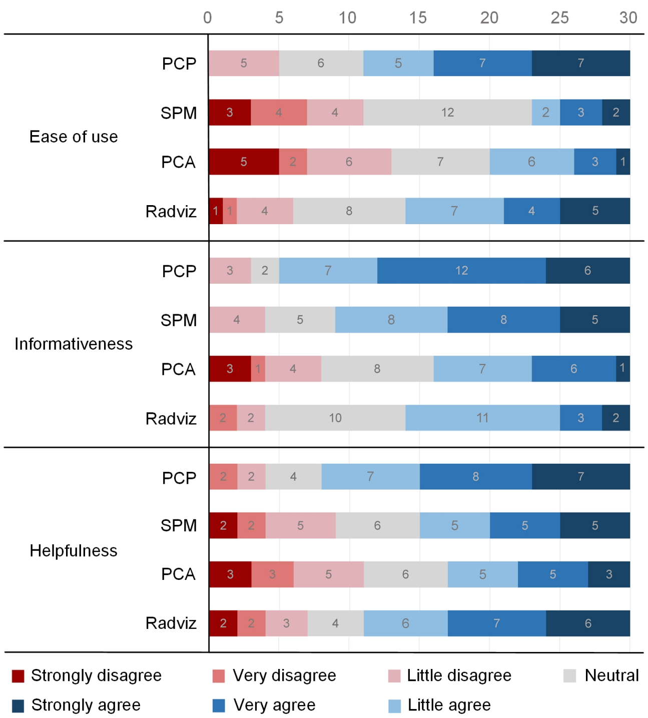


Figure 6. Stacked bar chart of subjective ratings with respect to ease of use, informativeness, and helpfulness of the four evaluated visualization techniques. Each volunteer answered a seven-point Likert scale with the subjective questionnaire after a round of experiment. Thirty ratings for each metric and technique were collected from 15 volunteers with two rounds of experiments.

1.194, PCA:$\mu = 4.2, \sigma = 1.591$) for informativeness. As for helpfulness, PCP receives the highest ratings ($\mu = 5.3, \sigma = 1.484$), Radviz is in between ($\mu = 4.8, \sigma = 1.821$), and SPM ($\mu = 4.5, \sigma = 1.796$) and PCA ($\mu = 4.1, \sigma = 1.814$) are the lowest in rank.

All in all, PCP is the most useful technique. The volunteers deemed that the important advantage of PCP is its graphic axes, which makes most of the cluster-related questions easy to handle. The volunteers commented, "*PCP's axes help me estimate the distribution of all data items on a cluster*" and "*I can immediately read the corresponding value on the axis when identifying the maximum MD of a cluster*". The second most useful technique is Radviz. Almost all volunteers reported that the projection mechanism of Radviz is helpful in identifying the stability of data items and the dominant cluster. As stated by most volunteers, "*The positions of stable data items in Radviz are easy to locate, so we do not need to examine many data items*".

SPM and PCA appear to be the least useful, with both having low ratings in terms of the three metrics. When asked to explain why they assigned low ratings to SPM, the volunteers explained that they found the visual clutter in SPM annoying, "*So many data points are crowded together that I don't know where to set about analyzing*", "*For me, SPM is useful only to compare the correlations between clusters*". PCA received low ratings likely due to its insufficient presentation of the cluster-oriented information, as some volunteers remarked that they can hardly find the information about clusters in PCA.

We briefly analyzed the correlation between informativeness and helpfulness. The experiment results show that the informativeness of PCP, PCA and Radviz has positive correlation with their helpfulness, which is consistent with the common sense that more informative visualizations would be more helpful. However, this positive correlation fails to appear on SPM. Most volunteers admitted that SPM is informative, as they commented, "*SPM shows almost all information of a dataset.*" But they also concerned that the information is repeated many times in SPM, as some volunteers commented, "*There is too much information in SPM you need to keep track of.*" And one volunteer emphasized, "*The similar information is a bit distracting when I'm using SPM.*"

6.3 Summary

The experiment results of objective questionnaires confirmed that no single visualization technique had remarkable ability to well support all the tasks. Considering the data-oriented tasks, Radviz obtained the best overall performance, which mainly benefits from its radial spring-based projection mechanism. As for the cluster-oriented tasks, PCP outperformed the other three techniques due to its vertical axes that represent clusters. SPM performed well when judging the correlations and similarities between clusters because it provides comparative views defined by pairwise axes, but the severe visual clutter in SPM made it perform poorly in the other tasks. PCA performed the worst in most of the tasks mainly due to its insufficient presentation of cluster information. Analyzing the ratings of the four visualization techniques in subjective metrics, we found that the more useful visualization techniques were PCP and Radviz. With regard to the influence of volunteers' academic background, we found that the volunteers with higher accuracy were those who were interested in fuzzy clusters analysis but not familiar with visualization techniques.

7 Discussion

In this section, we discuss the limitations of this evaluation and suggest some interesting aspects for further work.

We evaluated only four visualization techniques, although many other existing techniques represent fuzzy clusters. For example, heatmap is used to visualize the sorted membership matrix to discover similarity patterns [21]. However, we did not evaluate the heatmap due to the difficulty in designing consistent visual encodings and its interactions with the four techniques. We did not include any dataset with two clusters in selecting experiment datasets. Although the application scenarios related to two fuzzy clusters are common, such datasets were not used in the evaluation because they are not typical multi-dimensional data. Nevertheless, the use of data of two clusters in evaluation in the future is worthwhile.

The experiment datasets have diverse clustering structures and data sizes. However, we did not discuss the specific effects of the datasets on the visualization techniques in the experiment result analysis; instead, we examined these effects. The Friedman test showed that no significant differences were found for the datasets in terms of accuracy and satisfaction, because each dataset underwent the experiment with the same number of times. With respect to completion time, a large data size corresponded to a great time cost of the techniques. The above results are consistent with our conventional understanding.

We only provided a fixed dimension reordering strategy for PCP, SPM, and Radviz. Our aim was to simplify interactive operation and experiment design, because it would be a time-consuming and tedious process for the volunteers to manually search a satisfactory dimension ordering, and providing multiple preset reordering strategies would introduce a new experiment condition. Certainly, different dimension reordering strategies can reveal distinct aspects of a dataset [19, 27, 56], which may affect our experiment results [20]. Hence, this is an open question worth exploring in the future.

We preset the questions of the objective questionnaires with the representative data items and clusters, which is beneficial to state clear experimental goals to the volunteers and facilitate quantitative analysis. Nonetheless, the lack of open-ended questions means that the volunteers have no opportunities for free exploration.

The volunteers involved in this evaluation were not extensive. The main reason is that each volunteer needed to spend a long time in the experiment. Although we controlled the data size and the preset questions, the experiment was still time-consuming. Our training session lasted on average three hours. Apart from the training session and rest time, each volunteer took at least four hours to complete all the questionnaires. Hence, it is not clear what the results would be if a large number of volunteers participate in the experiment. Although our analysis results with regard to prior knowledge is similar to that of Dasgupta et al. [11], it is still not convincing. We hope to have an opportunity to conduct experiments involving a wider range of volunteers in the future.

Our experiment result analysis was performed on a range of metrics, which may be related. For example, the completion time should be relevant to the accuracy of the answers. However, the result analysis was based on a single metric analysis without revealing such relationship. The result analysis related to the interactions indicated that the interactions played an important role in reducing the performance differences among the techniques in solving the questions. However, we did not further analyze how performance differences would change with regard to whether the interactions were used or not. Discussing more interaction techniques and evaluating their impact are interesting topics that should be further explored. The four tasks we considered in this study do not cover all user queries in fuzzy clusters analysis. For example, the entire evaluation did not involve a comparison of two fuzzy clustering results. We plan to explore more tasks to better address the majority of user goals.

8 Conclusion

In this work, we conducted a controlled experiment to evaluate multi-dimensional visualization techniques in analyzing fuzzy clusters. Four multi-dimensional visualization techniques were evaluated, namely, PCP, SPM, PCA and Radviz, which are commonly used in both general multi-dimensional data analysis and fuzzy clusters analysis. We first defined analytical tasks and representative questions specific to fuzzy clusters analysis. We then designed objective questionnaires to compare the accuracy, time, and satisfaction in using the four techniques to solve the questions. We also designed subjective questionnaires to collect the experience of the volunteers with the four techniques in terms of ease of use, informativeness, and helpfulness.With a complete experiment process and a detailed result analysis, we provide instructive guidance for analysts in selecting appropriate and efficient visualization techniques to analyze fuzzy clusters. This work also suggests some directions for further research on the evaluation of multi-dimensional visualizations for understanding fuzzy clusters. Moreover, we believe this work will motivate the visualization and visual analytics community to pay urgent attention to evaluation studies involving various real-world application scenarios.

Acknowledgments

We wish to thank the reviewers for their thoughtful comments, the volunteers for their active involvement, and Wei Huang and Yang Shi for the fruitful discussions. This work is supported by the National Science Foundation of China (No. 61672538, 61772315 and 61772456), the Open Project Program of the State Key Lab of CAD&CG, Zhejiang University (No. A1812), the Open Research Fund of Beijing Key Laboratory of Big Data Technology for Food Safety, Beijing Technology and Business University (No. BKBD-2018KF08).

References

[1] J. Abonyi and R. Babuska. Fuzzsam-visualization of fuzzy clustering results by modified sammon mapping. In *Proceedings of the IEEE International Conference on Fuzzy Systems*, pp. 365–370. IEEE, 2004. doi: 10.1109/fuzzy.2004.1375750

[2] R. Amar, J. Eagan, and J. Stasko. Low-level components of analytic activity in information visualization. In *Proceedings of the IEEE Symposium on Information Visualization*, pp. 111–117. IEEE, 2005. doi: 10.1109/infovis.2005.24

[3] M. Ankerst, S. Berchtold, and D. Keim. Similarity clustering of dimensions for an enhanced visualization of multidimensional data. In *Proceedings of the IEEE Symposium on Information Visualization*, pp. 52–60. IEEE, 1998. doi: 10.1109/infvis.1998.729559

[4] A. Asuncion and D. Newman. Uci machine learning repository. `http://archive.ics.uci.edu/ml/datasets.html`, 2007.

[5] M. R. Berthold and L. O. Hall. Visualizing fuzzy points in parallel coordinates. *IEEE Transactions on Fuzzy Systems*, 11(3):369–374, 2003. doi: 10.1109/tfuzz.2003.812696

[6] J. C. Bezdek, R. Ehrlich, and W. Full. Fcm: the fuzzy c-means clustering algorithm. *Computers and Geosciences*, 10(2-3):191–203, 1984. doi: 10.1016/0098-3004(84)90020-7

[7] N. Cao, Y.-R. Lin, and D. Gotz. Untangle map: visual analysis of probabilistic multi-label data. *IEEE Transactions on Visualization and Computer Graphics*, 22(2):1149–1163, 2016. doi: 10.1109/tvcg.2015.2424878

[8] M.-S. Chen and S.-W. Wang. Fuzzy clustering analysis for optimizing fuzzy membership functions. *Fuzzy Sets and Systems*, 103(2):239–254, 1999. doi: 10.1016/s0165-0114(98)00224-3

[9] S.-M. Chen and Y.-C. Chang. Multi-variable fuzzy forecasting based on fuzzy clustering and fuzzy rule interpolation techniques. *Information Sciences*, 180(24):4772–4783, 2010. doi: 10.1016/j.ins.2010.08.026

[10] W. S. Cleveland. *Visualizing data*. Hobart Press, 1993.

[11] A. Dasgupta, S. Burrows, K. Han, and P. J. Rasch. Empirical analysis of the subjective impressions and objective measures of domain scientists' visual analytic judgments. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 1193–1204. ACM, 2017. doi: 10.1145/3025453.3025882

[12] J. V. De Oliveira and W. Pedrycz. *Advances in fuzzy clustering and its applications*. John Wiley and Sons, 2007.

[13] J. Demšar. Statistical comparisons of classifiers over multiple data sets. *Journal of Machine learning research*, 7(1):1–30, 2006.

[14] E. Dimara, A. Bezerianos, and P. Dragicevic. Conceptual and methodological issues in evaluating multidimensional visualizations for decision support. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):749–759, 2018. doi: 10.1109/tvcg.2017.2745138

[15] R. Etemadpour, L. Linsen, C. Crick, and A. Forbes. A user-centric taxonomy for multidimensional data projection tasks. In *Proceedings of the International Conference on Information Visualization Theory and Applications*, pp. 51–62. SCITEPRESS, 2015. doi: 10.5220/0005313400510062

[16] R. Etemadpour, R. Motta, J. G. de Souza Paiva, R. Minghim, M. C. F. de Oliveira, and L. Linsen. Perception-based evaluation of projection methods for multidimensional data visualization. *IEEE Transactions on Visualization and Computer Graphics*, 21(1):81–94, 2015. doi: 10.1109/tvcg.2014.2330617

[17] B. Feil, B. Balasko, and J. Abonyi. Visualization of fuzzy clusters by fuzzy sammon mapping projection: application to the analysis of phase space trajectories. *Soft Computing*, 11(5):479–488, 2007. doi: 10.1007/s00500-006-0111-5

[18] B. J. Ferdosi, H. Buddelmeijer, S. Trager, M. H. F. Wilkinson, and J. B. T. M. Roerdink. Finding and visualizing relevant subspaces for clustering high-dimensional astronomical data using connected morphological operators. In *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology*, pp. 35–42. IEEE, 2010. doi: 10.1109/vast.2010.5652450

[19] B. J. Ferdosi and J. B. Roerdink. Visualizing high-dimensional structures by dimension ordering and filtering using subspace analysis. In *Proceedings of Computer Graphics Forum*, vol. 30, pp. 1121–1130. Wiley, 2011. doi: 10.1111/j.1467-8659.2011.01961.x

[20] M. Friendly and E. Kwan. Effect ordering for data displays. *Computational statistics & data analysis*, 43(4):509–539, 2003. doi: 10.1016/s0167-9473(02)00290-6

[21] A. P. Gasch and M. B. Eisen. Exploring the conditional coregulation of yeast gene expression through fuzzy k-means clustering. *Genome Biology*, 3(11):1–22, 2002. doi: 10.1186/gb-2002-3-11-research0059

[22] R. Hammah and J. Curran. Fuzzy cluster algorithm for the automatic identification of joint sets. *International Journal of Rock Mechanics and Mining Sciences*, 35(7):889–905, 1998. doi: 10.1016/s0148-9062(98)00011-4

[23] P. Hoffman, G. Grinstein, K. Marx, I. Grosse, and E. Stanley. Dna visual and analytic data mining. In *Proceedings of the IEEE Visualization*, pp. 437–441. IEEE, 1997. doi: 10.1109/visual.1997.663916

[24] D. Holten and J. J. Van Wijk. Evaluation of cluster identification performance for different pcp variants. *Computer Graphics Forum*, 29(3):793–802, 2010. doi: 10.1111/j.1467-8659.2009.01666.x

[25] F. Hoppner and F. Klawonn. Visualising clusters in high-dimensional data sets by intersecting spheres. In *Proceedings of the International Symposium on Evolving Fuzzy Systems*, pp. 106–111. IEEE, 2006. doi: 10.1109/isefs.2006.251180

[26] A. Inselberg. Parallel coordinates: visualization, exploration and classification of high-dimensional data. In *Proceedings of Handbook of Data Visualization*, pp. 643–680. Springer, 2008. doi: 10.1007/978-3-540-33037-0_25

[27] S. Johansson and J. Johansson. Interactive dimensionality reduction

through user-defined combinations of quality metrics. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):993–1000, 2009. doi: 10.1109/tvcg.2009.153

[28] I. T. Jolliffe. Principal component analysis and factor analysis. In *Proceedings of Principal component analysis*, pp. 115–128. Springer, 1986. doi: 10.1007/978-1-4757-1904-8_7

[29] D. A. Keim and H.-P. Kriegel. Visualization techniques for mining large databases: a comparison. *IEEE Transactions on Knowledge and Data Engineering*, 8(6):923–938, 1996. doi: 10.1109/69.553159

[30] F. Klawonn, V. Chekhtman, and E. Janz. Visual inspection of fuzzy clustering results. In *Proceedings of Advances in Soft Computing*, pp. 65–76. CRC, 2003. doi: 10.1007/978-1-4471-3744-3_7

[31] A. Konig, O. Bulmahn, and M. Glesner. Systematic methods for multivariate data visualization and numerical assessment of class separability and overlap in automated visual industrial quality control. In *Proceedings of the British Machine Vision Conference 1994*, pp. 1–10. British Machine Vision Association, 1994. doi: 10.5244/c.8.19

[32] X. Kuang, H. Zhang, S. Zhao, and M. J. McGuffin. Tracing tuples across dimensions: a comparison of scatterplots and parallel coordinate plots. *Computer Graphics Forum*, 31(3pt4):1365–1374, 2012. doi: 10.1111/j.1467-8659.2012.03129.x

[33] L. C. Leung and D. Cao. On consistency and ranking of alternatives in fuzzy ahp. *European Journal of Operational Research*, 124(1):102–113, 2000. doi: 10.1016/s0377-2217(99)00118-6

[34] J. Li, J.-B. Martens, and J. J. Van Wijk. Judging correlation from scatterplots and parallel coordinate plots. *Information Visualization*, 9(1):13–30, 2010. doi: 10.1057/palgrave.ivs.9500179

[35] Y. R. Lin, N. Cao, D. Gotz, and L. Lu. Untangle: visual mining for data with uncertain multi-labels via triangle map. In *Proceedings of the IEEE International Conference on Data Mining*, pp. 340–349. IEEE, 2015. doi: 10.1109/icdm.2014.24

[36] M. Liu, J. Shi, Z. Li, C. Li, J. Zhu, and S. Liu. Towards better analysis of deep convolutional neural networks. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):91–100, 2016. doi: 10.1109/TVCG.2016.2598831

[37] S. Liu, W. Cui, Y. Wu, and M. Liu. A survey on information visualization: recent advances and challenges. *The Visual Computer*, 30(12):1373–1393, 2014. doi: 10.1007/s00371-013-0892-3

[38] J. Mao and A. K. Jain. Artificial neural networks for feature extraction and multivariate data projection. *IEEE Transactions on Neural Networks*, 6(2):296–317, 1995. doi: 10.1109/72.363467

[39] D. Marghescu. Multi-dimensional data visualization techniques for exploring financial performance data. In *Proceddings of the 13th Americas Conference on Information Systems*, p. 509. AMCIS, 2007.

[40] F. Masulli and A. Schenone. A fuzzy clustering based segmentation system as support to diagnosis in medical imaging. *Artificial Intelligence In Medicine*, 16(2):129–147, 1999. doi: 10.1016/s0933-3657(98)00069-4

[41] M. Ozer. User segmentation of online music services using fuzzy clustering. *Omega*, 29(2):193–206, 2001. doi: 10.1016/s0305-0483(00)00042-6

[42] W. Peng, M. O. Ward, and E. A. Rundensteiner. Clutter reduction in multi-dimensional data visualization using dimension reordering. In *Proceedings of the IEEE Symposium on Information Visualization*, pp. 89–96. IEEE, 2004. doi: 10.1109/infvis.2004.15

[43] J. Pizarro, E. Guerrero, and P. L. Galindo. Multiple comparison procedures applied to model selection. *Neurocomputing*, 48(1-4):155–173, 2002. doi: 10.1016/s0925-2312(01)00653-1

[44] Z. Qu and J. Hullman. Keeping multiple views consistent: constraints, validations, and exceptions in visualization authoring. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):468–477, 2018. doi: 10.1109/tvcg.2017.2744198

[45] M. Rubio-Sánchez, L. Raya, F. Diaz, and A. Sanchez. A comparative study between radviz and star coordinates. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):619–628, 2016. doi: 10.1109/tvcg.2015.2467324

[46] L. Rueda and Y. Zhang. Geometric visualization of clusters obtained from fuzzy clustering algorithms. *Pattern Recognition*, 39(8):1415–1429, 2006. doi: 10.1016/j.patcog.2006.02.006

[47] E. H. Ruspini. A new approach to clustering. *Information and Control*, 15(1):22–32, 1969. doi: 10.1016/S0019-9958(69)90591-9

[48] T. Rzeźniczak. Visual analysis techniques for medical diagnosis support. *Computer Science and Mathematical Modelling*, 2(2):33–41, 2015. doi: 10.5604/15084183.1197450

[49] M. Sedlmair, T. Munzner, and M. Tory. Empirical guidance on scatterplot and dimension reduction technique choices. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2634–2643, 2013. doi: 10.1109/tvcg.2013.153

[50] J. Sharko and G. Grinstein. Visualizing fuzzy clusters using radviz. In *Proceedings of the 13th International Conference on Information Visualisation*, pp. 307–316. IEEE, 2009. doi: 10.1109/iv.2009.74

[51] J. Sharko, G. Grinstein, and K. A. Marx. Vectorized radviz and its application to multiple cluster datasets. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1444–1427, 2008. doi: 10.1109/tvcg.2008.173

[52] Y. Shi, C. Bryan, S. Bhamidipati, Y. Zhao, Y. Zhang, and K.-L. Ma. Meetingvis: Visual narratives to assist in recalling meeting context and content. *IEEE Transactions on Visualization and Computer Graphics*, 24(6):1918–1929, 2018. doi: 10.1109/tvcg.2018.2816203

[53] J. Soo Yi, R. Melton, J. Stasko, and J. A. Jacko. Dust & magnet: multivariate information visualization using a magnet metaphor. *Information visualization*, 4(4):239–256, 2005. doi: 10.1057/palgrave.ivs.9500099

[54] E. R. A. Valiati, M. S. Pimenta, and C. M. D. S. Freitas. A taxonomy of tasks for guiding the evaluation of multidimensional visualizations. In *Proceedings of the 2006 AVI Workshop on Beyond Time and Errors: Novel Evaluation Methods for Information Visualization*, pp. 1–6. ACM, 2006. doi: 10.1145/1168149.1168169

[55] Y. Wang, L. Gou, A. Xu, M. X. Zhou, H. Yang, and H. Badenes. Veilme: An interactive visualization tool for privacy configuration of using personality traits. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 817–826. ACM, 2015. doi: 10.1145/2702123.2702293

[56] E. J. Wegman. Hyperdimensional data analysis using parallel coordinates. *Journal of the American Statistical Association*, 85(411):664–675, 1990. doi: 10.2307/2290001

[57] S. Wehrend and C. Lewis. A problem-oriented classification of visualization techniques. In *Proceedings of the First IEEE Conference on Visualization*, pp. 139–143. IEEE, 1990. doi: 10.1109/visual.1990.146375

[58] C. Williamson and B. Shneiderman. The dynamic homefinder: Evaluating dynamic queries in a real-estate information exploration system. In *Proceedings of the 15th Annual International ACM SIGIR Conference on Research and Development*, pp. 338–346. ACM, 1992. doi: 10.1145/133160.133216

[59] J. Xia, F. Ye, W. Chen, Y. Wang, W. Chen, Y. Ma, and A. K. Tung. LDSScanner: Exploratory analysis of low-dimensional structures in high-dimensional datasets. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):236–245, 2018. doi: 10.1109/tvcg.2017.2744098

[60] C. Xie, W. Zhong, and K. Mueller. A visual analytics approach for categorical joint distribution reconstruction from marginal projections. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):51–60, 2017. doi: 10.1109/tvcg.2016.2598479

[61] J. Yang, W. Peng, M. Ward, and E. Rundensteiner. Interactive hierarchical dimension ordering, spacing and filtering for exploration of high dimensional datasets. In *Proceedings of IEEE Symposium on Information Visualization*, pp. 105–112. IEEE, 2003. doi: 10.1109/infvis.2003.1249015

[62] L. A. Zadeh. Fuzzy sets. In *Proceedings of the Fuzzy Sets, Fuzzy Logic and Fuzzy Systems*, pp. 394–432. World Scientific, 1996. doi: 10.1142/9789814261302_0021

[63] F. Zhou, W. Huang, Y. Zhao, Y. Shi, X. Liang, and X. Fan. ENTVis: A visual analytic tool for entropy-based network traffic anomaly detection. *IEEE Computer Graphics and Applications*, 35(6):42–50, 2015. doi: 10.1109/mcg.2015.97