

Interactive Visual Categorization of Spinel-Group Minerals

M. L. Ganuza^a, G. Ferracutti^b, F. Gargiulo^b, S. M. Castro^a, E. A. Bjerg^b, E. Gröller^{c,d}, K. Matković^e

^aVyGLab Research Laboratory and Departamento de Cs. e Ing. de la Computación, Universidad Nacional del Sur, Bahía Blanca, Argentina

^bINGEOSUR and Departamento de Geología, Universidad Nacional del Sur, Bahía Blanca, Argentina

^cTUWien, Vienna, Austria

^dUniversity of Bergen, Norway

^eVRVis Research Center, Vienna, Austria

Abstract

Spinel-group minerals are excellent indicators of geological environments and are of invaluable help in the search for mineral deposits of economic interest. The geologists analyze them by means of Barnes and Roeder's contours. In this paper, we present a collection of novel, interactive methods, which assist geologists in the categorization of spinel-group minerals. We fully integrate Barnes and Roeder's contours using a polygonal representation. This makes it possible to efficiently superimpose user-provided point data over the contours, and to automatically rank the contours based on the number of enclosed points. We also allow the expert to create contours for the user-provided point data. Once user contours are created, they can be compared with Barnes and Roeder's contours. During the analysis, the user can drill-down by means of brushing. As we deal with specific data, we apply two novel brushing techniques, i.e., the percentile brush and the contour brush. The novel brushing mechanisms along with the interactive comparison speed-up the analysis significantly. We evaluate the newly introduced approach and the resulting novel workflow using real-world data from different locations in Argentina. According to the domain experts, the classification of spinel minerals needs several minutes now, while it took a few days with the current state of the art approach in the domain.

Keywords: Interaction Techniques, Interactive Classification, Spinel Mineral Analysis, Visual Analytics

1. Introduction

A big challenge in geosciences is to characterize a particular geologic region in terms of its tectonic settings. To accomplish this, a great amount of data obtained from rocks and minerals formed in different geological environments must be processed and analyzed. In this sense, the analysis of spinel minerals is one of the most reliable tools.

Spinel-group minerals provide useful information regarding the geological environment in which the host rocks were formed. By studying a particular suite of rocks with spinel-group minerals, it is possible to reconstruct the tectonic setting where the suite has been formed and the history of events it was subjected to through the geological time.

Geologists primarily use triangle plots, scatterplots, and the 3D spinel prism representation to compare populations of analyses [1]. These diagrams usually represent 2D projections on the faces of the spinel prism (see Figure 1).

In 2001, Barnes and Roeder [2] defined a set of contours corresponding to compositional fields for spinel-group minerals. These fields were defined using point-density contour-plots (PDCPs) based on the spinel prism. Figure 2 shows an example from Barnes and Roeder's contour collection, depicted through three projections of the prism.

At present, geologists use these contours as empirical tectonic discriminators that empower them to estimate the tectonic environment where a spinel with a particular composition has been formed. In order to do this, they compare their data, con-

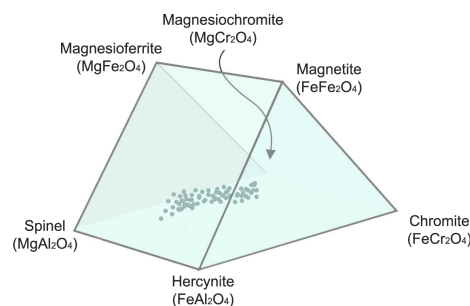


Figure 1: The Magnetite Prism is an example of a commonly used spinel prism. It is a prismatic space whose vertices correspond to the end-members of the compositional space. All points inside the prism represent various compositions of the six end-members.

sisting of several analyses, with the PDCPs from Barnes and Roeder. This is a tedious task. The current state of the art includes an exhaustive manual comparison of new data with the Barnes and Roeder's contours [2]. The actual task is to find contours that contain (or partially contain) points corresponding to the new data. Manipulating and analyzing large datasets of spinel compositions is a highly time-consuming process.

In this paper, we introduce a semi-automatic, interactive detection of tectonic settings for an arbitrary spinel dataset according to the dataset and contours published by Barnes and Roeder. In order to achieve this, several techniques have been developed and incorporated into Spinel Explorer++, an exten-

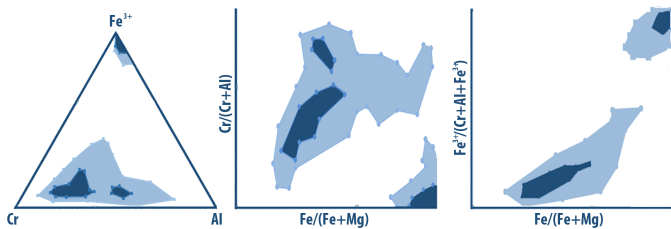


Figure 2: Barnes and Roeder’s contours for Basalts depicted in the three projections of the Magnetite prism. 90th percentile contours (in light blue) and 50th percentile contours (in dark blue) are shown.

sion of the Spinel Explorer [3].

The novel interactive categorization is the result of a thorough study of geologists’ tasks, identified during numerous evaluation sessions with domain experts. All Barnes and Roeder’s contours are now integrated in the system, and can be depicted in three projections of the 3D spinel prism: two rectangular faces—represented by two scatterplots—and one triangular face—represented by a triangle plot. The new methodology is integrated in Spinel Explorer++, which significantly improves the scientists’ workflow. For the first time, we integrate user-based point data and point-based contours with Barnes and Roeder contours for simultaneous comparisons.

The high-level contributions of the paper can be summarized as: (1) an interactive, semi-automatic categorization of spinel data using Barnes and Roeder’s contours, (2) an interactive data-density-based contour creation and contour-based comparison, and (3) an evaluation of the proposed approach in a geoscience context including domain experts’ feedback. In order to accomplish this high-level contributions, various improvements and novel techniques have been necessary. They include: (a) what we call a contour brush, (b) what we call a percentile brush, and (c) an integrated, fully operational tool, used by domain experts. While the approach has been designed for a specific problem it can be generalized and it is applicable to dense scatterplots.

We evaluate the introduced approach and the resulting workflow using real-world data from different locations in Argentina. This paper is written by an interdisciplinary group of visualization and geology experts. If the pronoun *we* is used in the paper it stands for this interdisciplinary group. In cases where it refers only to a sub-group it will be explicitly mentioned.

2. Related Work

If the analysis of large amounts of data becomes too cumbersome in various domains, interactive visual analysis often can provide new solutions. Visual analytics combines the strengths of human perception and cognition with those of a computational analysis [4, 5, 6]. Interactive visual analysis provides an interactive and iterative exploration and analysis framework, where the user guides the analysis, supported by a variety of computational analysis tools. This helps the domain expert to explore and analyze the data and to understand complex and often hidden relationships between certain data aspects. The visual information seeking mantra—overview first, zoom and fil-

ter, then details-on-demand—as identified by Shneiderman [7], summarizes the most typical pattern in interactive visualization.

This paper builds on the Spinel Explorer [3] that introduced the interactive triangle plot and the 3D spinel prism, exploiting the well known coordinated and multiple-views principle. Still, the spinel categorization was done based only on the observations made by geologists. In this paper we present an interactive categorization based on the Barnes and Roeder’s contours. This is a major step forward, as the most cumbersome part of the workflow, i.e., the comparison of user-provided point data with the contours, is significantly improved and semi-automatized.

Bruckner et al. [8] described a visual query system, which is related to our contour categorization. They deal with 3D brain images, whereas we have 2D contours. The queries themselves are also different, we have user-provided point data, i.e., point sets, which should be compared with the contours. Wu et al. [9] deal with boundary data. They analyze boundary changes over time and support finding patterns that dynamically emerge. The work is related to ours, but we compare point sets with boundaries and we investigate domain specific plots. We also introduce different alternatives for the comparison of point sets and 2D contours. Schmidt et al. [10] deal with 3D point clouds and analyze different meshes generated thereof. We deal with 2D contours and an evaluation of different contour-generation techniques is not in our focus.

Visual analytics has been successfully employed in many domains. It has also been applied to scientific data [11]. Coordinated multiple-views [12] are often employed as a proven concept in visual analysis. The main idea is to depict various dimensions using multiple-views and to allow the user to interactively select (i.e., brush) subsets of the data in a view. All the corresponding data items in all linked views are then consistently highlighted. Wang Baldonado et al. [13] describe guidelines for using multiple-views. The positive feedback we obtained concerning the Spinel Explorer [3] motivated us to further exploit coordinated multiple-views.

In geology, it is usual to plot spinel minerals on prismatic spaces [1]. Such plots are difficult to generate by hand, therefore, scientists mostly use binary and ternary plots to evaluate correlations between chemical elements or oxides. In 2012, Ganuza et al. [14] presented the SpinelViz tool. The application consists of an interactive 3D viewer, which enables the experts to view and explore the spinel prism with several datasets at the same time. SpinelViz provides the capability to manipulate, view, plot, and project data in 2D and 3D, which helps the user to gain a better insight into the data distribution. However, to complete another step of the visualization process, it is necessary to compare the user data with the Barnes and Roeder’s contours. According to the best of our knowledge, there is no previous tool, which supports interactive semi-automatic spinel categorization integrated with the plots that the experts are used to employ in the traditional workflow.

3. Domain Background—Spinel-Group Minerals

The main goal of spinel mineral-analysis is to classify spinels in order to show chemical characteristics and to discriminate

processes related to their formation.

In this process, the geologists aim to identify chemical groups and sub-groups of spinel minerals and the tectonic environment to which these groups are related. These chemical groups and sub-groups are determined by the chemical variations in the dataset. In addition, they intent to discriminate whether these groups and sub-groups belong to the same tectonic setting or to different ones.

Essentially, the geologists want to identify the spinel-group minerals in their samples and assign them to a certain geological process in a particular tectonic setting. The process is not straightforward and cannot be fully automatized. There is no sharp boundary between spinel minerals, and domain experts' knowledge and experience is crucial in the process. Interactive support tools are very helpful for the domain experts.

Spinel-group minerals are constituents of mafic, ultramafic, and metamorphic rocks. They make up a solid solution with 22 end-members [15]. There are eight end-members, which are especially useful in the analysis. They contain two partially overlapping groups of six members, which sum to one. These eight members are commonly used for the mineral representation on special diagrams—triangular prisms. The Magnetite Prism and the Ulvöspinel Prism are used to depict the two groups. Each vertex of a prism represents one end-member. Depending on the ratios of the elements in a mineral, the mineral composition is plotted in a specific position inside the prism (see Figure 1).

Since the 3D prisms are difficult to analyze, the geologists mostly use four 2D plots: (1) the $Cr-Al-Fe^{3+}$ triangular plot, representing the projection onto the triangular face of the Magnetite Prism, (2) the $Cr/(Cr+Al)$ vs. $Fe^{2+}/(Mg+Fe^{2+})$ scatterplot, representing the projection onto the base of the Magnetite Prism, (3) the $Fe^{3+}/(Cr+Al+Fe^{3+})$ vs. $Fe^{2+}/(Mg+Fe^{2+})$ scatterplot, which represent the projection onto the lateral face of the Magnetite Prism, and (4) the TiO_2 vs. $Fe^{3+}/(Cr+Al+Fe^{3+})$ scatterplot.

3.1. Barnes and Roeder's Contours

Barnes and Roeder [2] compiled a global database of about 26000 samples of spinels from igneous and metamorphic rocks. From this database, they extracted and defined a set of contours corresponding to compositional fields for spinel-group minerals (see Figure 2). In order to do this, they classify all members of a point set—each corresponding to an analysis—into a two-level hierarchy, grouping the points in eight major categories [2]. These categories are further subdivided into subcategories and for each subcategory they generate the spinel compositional plots (one for each representative face of the prism).

Barnes and Roeder contoured all the generated plots to aggregate the data, such that the n^{th} percentile contour encloses the most densely packed $n\%$ of the data. They defined and tabulated the 50th and 90th percentile contours and these contours are effectively used as a quantitative basis of comparison. This point-density approach is used to derive a meaningful compositional field from a point set and it is also a useful one to compare point sets from different analyses.

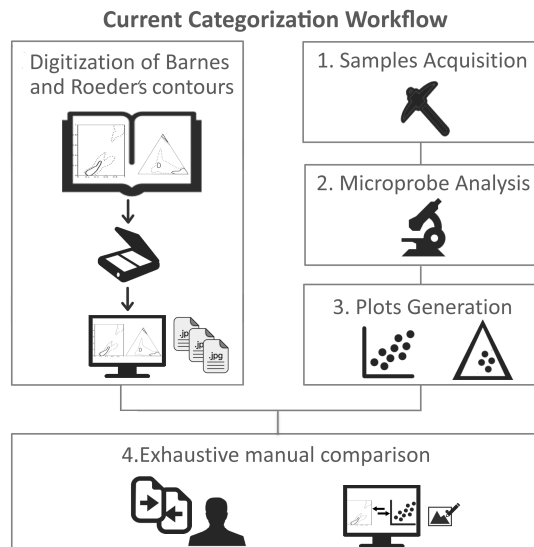


Figure 3: Current workflow for the categorization of spinel-group minerals.

3.2. Current Categorization Workflow

There are mainly two standard tasks in the spinel-categorization workflow. One of them is to confirm the tectonic setting to which the spinel sample under consideration belongs to. The other task to categorize a spinel sample whose tectonic setting is practically unknown. In both cases, a comparison with the Barnes and Roeder's contours must be conducted.

The current workflow for the categorization of spinel-group minerals is illustrated in Figure 3. The first step is the sample acquisition. In this step geologists take in-situ rock samples, which are representative of the problem they want to solve. The second step involves the study of the rock samples with a petrographic microscope and the chemical analyses of the minerals of interest. This step results in the ratios of the 22 end-members for each sample, i.e., a point set. In the third step, with the acquired mineral analyses, the experts construct different scatterplots and triangle plots. With these they study the relations between the several chemical elements that constitute each spinel mineral. In these diagrams, geologists can also discriminate the chemical variations in the point set and recognize chemical groups and sub-groups. From these plots, the experts generate the four diagrams provided by Barnes and Roeder. Interested readers in these first three steps are referred to previous work [3, 16] for a more detailed description. Eventually, after diagrams have been created, they are compared with the Barnes and Roeder's contours. The Barnes and Roeder's contours are available as hard-copy printouts, or as scanned images.

The last step of the workflow consists of an exhaustive manual comparison of all Barnes and Roeder's contour diagrams with the plots generated in the third step. The geologists either place the contours next to the generated plots or they superimpose plots using a drawing program that supports layers. This is considered to be the most advanced method in the field. The generated plots do not contain contours. They depict a set of points, and geologists check, which contours (partially) encompass them. Finally, geologists have to decide (visually), which

of the digitized contours match the best with their own data. The best fit will be the most probable tectonic setting for the spinel-group mineral dataset under consideration. If the dataset matches with more than one tectonic-setting compositional-field, geologists must rely on their background knowledge concerning the geological context of the study area in order to select the most probable tectonic setting.

The manual comparison of the plots with Barnes and Roeder's contours can take several days. Usually, the geologists need to compare the Barnes and Roeder's contours for all categories and subcategories (and for all the projections of the prism) with the plots generated in the third step.

The current categorization workflow is error-prone and tedious. However, according to the best of our knowledge, there was no software tool that helps the experts in the adoption of automated methods in the conventional practice.

We propose a novel categorization approach that improves the current workflow significantly. Instead of several days, the categorization takes only several minutes now.

4. Interactive Spinel Categorization

The current classification workflow is limited due to several reasons. The contours are available as images only and the users can only compare their own point-based diagrams with contour plots. The generation of contours for a sample, i.e., a user-provided point set, could improve the comparison. Together with geologists, we have identified the most important requirements that would improve the comparison significantly. The requirements can be summarized as:

R1. Automatically rank Barnes and Roeder's contours based on the number of members of the point set, which they encompass.

R2. Allow the creation of contours from the user-provided point set for different percentiles.

R3. Provide a superimposition of the contours created from the user-provided point set and the Barnes and Roeder's contours in order to allow the experts a visual comparison.

R4. Provide an efficient way of reducing the user-provided point set (e.g., outliers removal) and allow the experts the comparison of reduced subsets, as well.

In this paper we introduce the Spinel Explorer++, which provides solutions for all the identified requirements. This integrated, fully operational tool results in a novel workflow for the categorization of spinel-group minerals. The novel workflow improves the current procedure significantly.

4.1. Spinel Explorer++ Overview

Spinel Explorer++ is based on the Spinel Explorer system [3]. The main goal of Spinel Explorer is to provide a unified system for the exploration of spinel minerals. The Spinel Explorer includes scatterplots, triangle plots and spinel prisms, which are well known and established visual representations in the geology domain. In addition, it includes parallel coordinates, histograms, and statistics overviews. All views are linked and support interactive selections.

The positive feedback we obtained concerning the Spinel Explorer motivated us to design the Spinel Explorer++ including

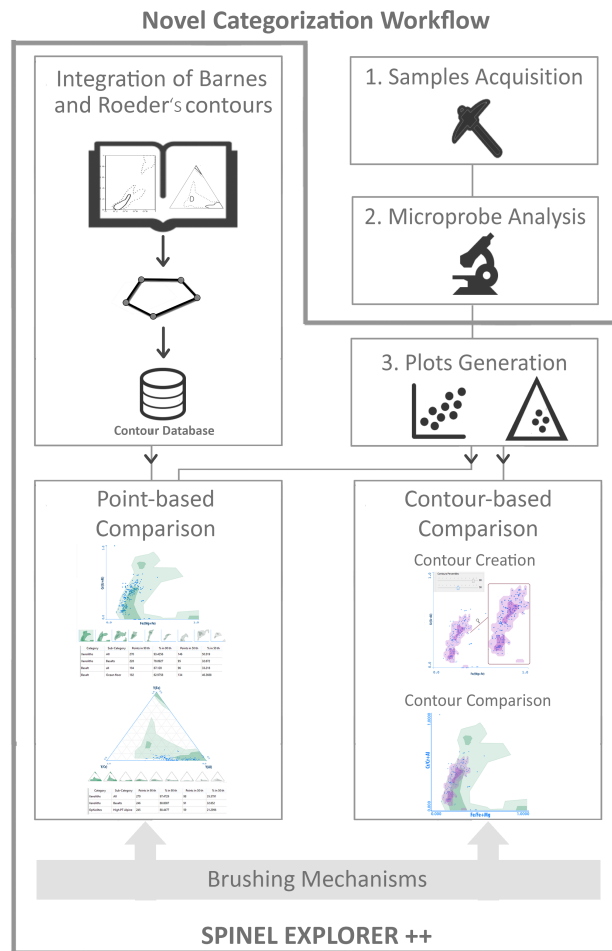


Figure 4: Novel workflow for the categorization of spinel-group minerals.

those standard views. The domain experts already had a clear understanding of how they wanted the data to be displayed, avoiding the need to consider other design alternatives. The existing Barnes and Roeder's contours were a strong guiding example to design our own techniques in a similar fashion. As Barnes and Roeder's contours are defined on the projections of the spinel prism, we digitized those contours and integrated them in the scatterplot and triangle plot views. Finally, two new brushes has been introduced, which follow conventional operations of the geologists.

4.2. Novel Categorization Workflow

We propose a novel workflow for the categorization of spinel-group minerals. The new procedure is illustrated in Figure 4. The first and the second step are carried out in the same way as in the current workflow described in Section 3.2. The integration of the Barnes and Roeder's contours involves their transformation into polygons. The polygons are stored in a contour database. The corresponding contours can be shown in the scatterplot and the triangle plot during the whole categorization process with a very short query time. Instead of a book of contours and a collection of bitmap images, we have the contour database stored in the system, which enables the experts to display the required

plots. There is no need any more to use drawing programs for comparison, with a tedious alignment step.

The analysis begins when the expert loads a sample, i.e., a point set, into the Spinel Explorer++. In order to study chemical variations, chemical groups, and sub-groups, and in order to classify spinel minerals, the Spinel Explorer++ allows the experts the construction of the four diagrams suggested by Barnes and Roeder. In addition, many other diagrams, for example, parallel coordinates or histograms are also supported (Step 3).

At this stage of the workflow, and based on the requirements summarized before, we provide two different techniques, i.e., a *point-based* and a *contour-based* approach.

With the *point-based* approach we compare the point-based diagrams generated in Step 3 with the Barnes and Roeder's contours stored in the contour database. The user can select which contour to show (if the experts conjectures a certain category), or the system can search for the contours that contain most of the members of the point set. Spinel Explorer++ automatically ranks Barnes and Roeder's contours based on the number of points, which they include. For the ranking in the *point-based* approach, we compute the number of points from the sample that belong to each Barnes and Roeder's contour. The contours are defined as 2D polygons, where the checking if a point belongs to a polygon (contour) can be efficiently computed [17]. The user can provide two thresholds, one for the 50th percentile contours and the other one for the 90th percentile contours. All contours, for which the number of points inside is larger than the user-provided threshold, are listed as potential spinel categories. These categorization lists are provided for each of the four 2D projections separately. Additionally there is an aggregated intersection list, which shows only categories that appear in all four categorization lists. All potential contours are displayed in a preview below the corresponding diagram of the point set. The preview is shown as a strip of small contours sorted by the number of points enclosed. In these small contours, transparency is used to reflect the level of overlap between the point set and the represented contour. Now, the user can select a small preview contour to compare the contour with the plotted point set. Figure 5 shows three projections of the prism with the preview contours and categorization lists. The lists are different for the various projections, and the aggregated intersection list of potential spinel categories is shorter. Entries for both, the 50th and the 90th percentile contours, are shown in the lists. Quantitative information on the number of points inside the contours and comparison speed, make such a categorization superior to the current state of the art.

In the *contour-based* approach we create contours from the user-provided point set first. This novel interaction allows the experts the creation of different percentiles contours from the point set. The user has full control of the contour-creation parameters and can freely create contours in the range from the 1st to the 100th percentile. This allows the user the visualization of different percentiles contours of the point set, which helps the experts to gain a deeper insight into their data. The Spinel Explorer++ provides contour creation for all projections of the prism. Once the contours are created, they can be overlaid and compared with Barnes and Roeder's contours. The contour generation

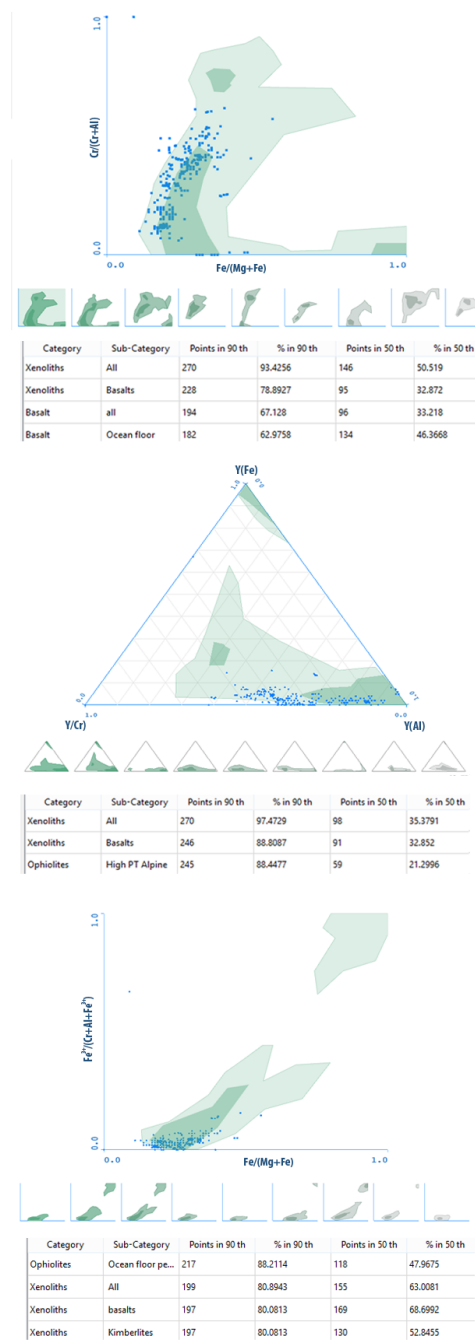


Figure 5: Preview contours and categorization lists. Xenoliths in Basalts depicted in the three projections of the prism. For each projection a preview of the contours and a categorization list is provided. Although the lists are different for the three projections, entries 1 and 2 appear in all of them. Barnes and Roeder's contour for Xenoliths in Basalts is superimposed on the user-provided point set. In the preview contours, transparency is used to reflect the level of overlap between the point set and the represented contour.

is density-based and we follow the algorithm as described by Barnes and Roeder. The following section provides more details on the algorithm itself.

4.3. Point-Density Contour-Plot

We start out with a sample, i.e., a set of points. The area of a particular 2D plot is partitioned into rectangular or triangular

Algorithm 1 ContourCreation(pointSet, percentile)

```
count ← size(pointSet)
listOfCells ← CalculateDensity(pointSet)
index ← 0
summedDensity ← 0
while (summedDensity < (count/100) * percentile) do
  summedDensity+ = listOfCells[index].Density
  insert listOfCells[index] in listOfInsideCells
  index ++
marchingSquares(listOfInsideCells)
```

Algorithm 2 CalculateDensity(pointSet)

```
listOfCells ← create all the cells of the grid
for each point in pointSet do
  indexCell ← index of the grid cell that contains point
  listOfCells[indexCell].Density ++
sort listOfCells based on density
return listOfCells
```

grid cells. The cell size is one-tenth the side length of the corresponding plot. For each grid cell the number of included points is counted. This number is called the density of the grid cell. Grid cells are sorted according to their densities into a list in descending order. For a specific percentile value we first determine those grid cells that belong to the inside of the corresponding percentile contour. We traverse the sorted list of grid cells and sum up their density values until the percentile value is reached. Cells further down the list are outside the contour. After this classification step each grid cell is either inside or outside and the contour is created automatically using a simplification of the marching squares algorithm [18].

Algorithm 1 describes a generalization of the contour creation process for triangular and rectangular plots. The *ContourCreation* algorithm receives a sample, i.e., a *pointSet* with 2D coordinates for each point, upon which the contour will be generated and the *percentile* of the contour to be created. This algorithm calculates which points in the *pointSet* lie inside the new contour of the given *percentile*. The *percentile* must be an integer value in the range of 1-100. The first step is to calculate the density for each grid cell. Algorithm 2 receives the *pointSet* and returns the *ListOfCells* sorted according to density in descending order. The *listOfCells* is traversed and cell densities are summed (*summedDensity*) until the desired *percentile* is achieved. This determines the cells inside the contour (*listOfInsideCells*).

With the *listOfInsideCells*, iso-lines between inside and outside cells are traced applying a simplification of the marching squares algorithm [18]. A contour may consist of several disjoint parts.

Figure 6 shows an example of two density-based contour-creations for the three projections of the prism. In this particular case the 50th and 90th percentile contours have been created. The density-based contour-creation is available at any time of the analysis session. The user can create contours, change percentiles, and hide contours at any time.

There is not a pre-established order between point-based and contour-based comparison steps in the analysis workflow.

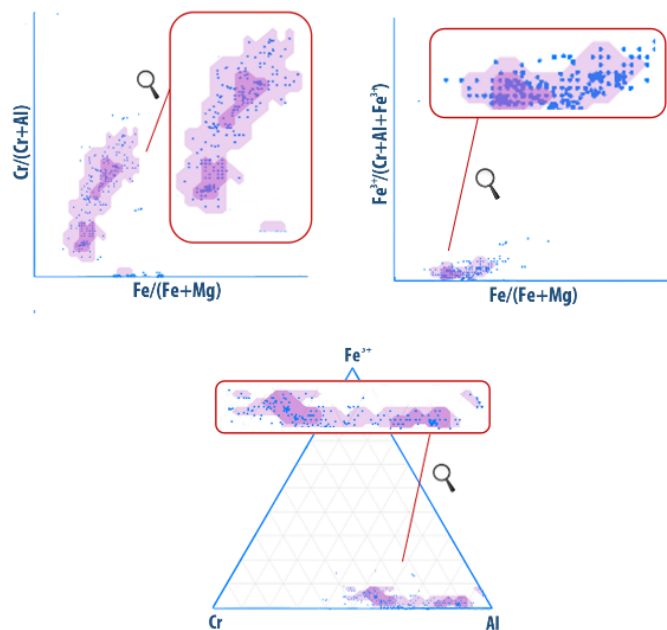


Figure 6: Contour creation for three projections of the prism. The 50th percentile contours (dark purple) and the 90th percentile contours (light purple) for a user-provided point set have been created. The user can interactively select the percentile of the contour to be generated.

Users may perform the semi-automatic categorization first, and after that, for further exploration, create new contours and compare them with the potential contours returned by the Spinel Explorer++ in the point-based comparison step. Otherwise a user could start with the contours-based comparison and proceed with a point-based analysis. The user can interactively loop between all possibilities during the interactive analysis process.

As stated in the requirements list, the experts also want to exclude some members of the point set from the analysis. If there are clear outliers, for example, they might influence the analysis so that it yields wrong results in the end. Brushing—interactive selection of points in the plots—is a well established technique in the visual analysis. Standard brushes include: rectangles for scatterplots, triangles for triangle plots, or interval brushes for parallel coordinates, for example. All of these are supported in Spinel Explorer++. A table view shows data for all brushed items and if we select points outside of a contour they can be individually examined. In order to support exploration, the new workflow involves two novel brushing mechanisms, called *percentile brush* and *contour brush*. The following section describes the proposed brushes in more detail.

4.4. Brushing Mechanisms

In order to support exploration, we developed two novel brushing mechanisms in addition to the conventional rectangular brushing, and the grid-based brushing in the triangle plot [3].

The concept of linking and brushing is key to interactive visual analysis [19, 20]. The term brushing was defined by Becker and Cleveland [21] in 1987 and since then, different brushing mechanisms and brushing shapes were investigated [22, 23, 24]. Radoš et al. [25] propose a collection of interactive techniques

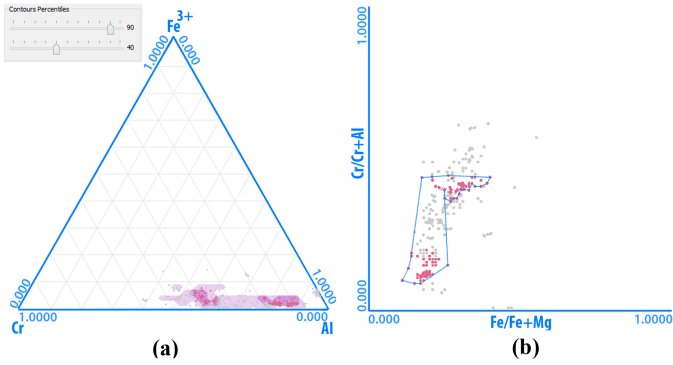


Figure 7: Percentile and contour brush. (a) A percentile brush is active, selecting all the points that fall into the 40th percentile of the density-based contour. (b) The corresponding concave contour (contour brush) is drawn.

targeted specifically towards reproducible and quantitative visual analytics. They introduce the concept of percentile brushing, which constrains the brush extent so that it always contains a predefined number of items, like 10%. This brush can be moved freely, snapped to a conventional grid, or snapped to a percentile grid. When moved, the extent of the brush is continuously adapted so that it always selects a predefined number of items.

Our *percentile brush* has a different meaning, since it selects all points belonging to a given contour by a simple click. Depending on the configuration, only points belonging to the 90th or 50th contour are brushed. In this way experts can explore hidden inter-relations in the dataset. The percentile brush is provided not only for Barnes and Roeder’s contours, but also for the density-based contours created by the user. The percentile brush has different meanings, depending if it is applied to the Barnes and Roeder’s or to the density-based contours. If it is applied to a Barnes and Roeder’s contour, percentile brushing allows the user the identification of those members of the point set that fall inside the contour. If this novel interaction is applied to density-based contours, it allows the user to identify, by a simple click, all points with higher densities than the percentile of the generated contour (see Figure 7).

The *contour brush* is created based on the linked views. The user brushes some data in an arbitrary view and the corresponding points in the scatterplots or triangle plot are highlighted. The highlighted points are used to create a contour. A concave hull is drawn around the points and intersections with all other contours are computed. Now, instead of the number of points in the contour, the contours’ areas can also be compared. The experts choose the best alternative according to their current task and needs. Note that the concave hull is different from a density-based contour. It is always possible to create density-based contours for a subset of points. The contour brush, offers an additional way in the analysis. Figure 7 shows an example of a percentile brush for a density-based contour on the left and a concave hull (contour brush) enclosing selected points on the right. A case study, which illustrates how all the newly proposed techniques can be used, is described in the following section.

5. Evaluation

In this section we describe one scenario, which we have analysed using our novel exploration tool. This scenario was selected and investigated by our cooperating geology experts. A group of three researchers has been involved in the design of our proposed tool as well as in the selection of the case study and in the evaluation of the final visualization. They are professors—one being the head, and the other two members—of a large Geological Research Institute. All of them have more than 20 years of experience in this domain. They have worked on the chemical composition of spinel-group minerals and their mineralogical classification. Further they applied spinel-group minerals as proxies in the prospection of base metals and platinum group minerals and in the identification of geotectonic environments. Besides, they have been Spinel Explorer users for three years.

The evaluation of our novel categorization method was conducted in the following way. First, a thorough explanation of the novel interactive methods included in the Spinel Explorer++ was given. The explanation included the novel brushing mechanisms, the Barnes and Roeder’s contours loading, and the contour-generation features. The explanation was followed by some questions about the tool, in order to ensure that the novel application functionalities have been understood. This first part of the session was successful. Since the experts are experienced Spinel Explorer users, they successfully got acquainted with the new techniques.

The second part of the session had two goals. The first goal was to verify that our proposal is capable of identify the tectonic setting to which the spinel sample under consideration belongs to, but in a faster way. The second goal was to prove that our proposal is capable of helping in the generation of new hypotheses. For this second part a specific case study was set up, and is described in Section 5.1.

The third part of the evaluation session was used to collect informal user feedback from domain experts. The user feedback is summarized in Section 5.2.

5.1. Case Study: Xenoliths in Basalts from Argentina

We evaluated the novel categorization method in a case study. The user-provided point set corresponds to Xenoliths from different locations in Argentina.

In the first stage of the case study we focus on the novel categorization. Then, we explore the potential of Spinel Explorer++ by investigating relationships between the spinels collected in different locations. The overall feedback of the domain experts was very positive.

The *user dataset* consists of 531 data samples. Each data sample corresponds to the chemical composition of an analyzed point of the spinel samples. All data samples have 60 attributes, including oxide values, cation values, and end-member values corresponding to each analyzed point.

Geologists are aware of the geological context and the regional framework where the rocks were collected in the field. Based on this they assume that the dataset corresponds to the Xenoliths in Basalts. The goal is to confirm this hypothesis.

The analyst configures the system to show the database through the triangle plots and the scatterplots corresponding to the projections of the spinel prism as investigated by Barnes and Roeder.

Since the Spinel Explorer++ allows the experts to put any Barnes and Roeder's contour on top of the loaded data, the geologists superimpose the Xenoliths in Basalts contours to check if the point set lies inside the contours. As shown in Figure 8, indeed most of the points lie inside the contours specified by Barnes and Roeder for Xenoliths in Basalts. This gives experts a positive feedback about their hypothesis. However, this is not sufficient. They would like to verify that these contours are the best match for their data. The Spinel Explorer++ makes it possible to search for the most similar contours and to quickly select the best fit among all the tectonic settings. The categorization lists are provided for each projection separately.

Figure 5 shows that at the top of two of the three lists are the Xenoliths, but for one projection (the $Fe^{3+}/(Cr + Al + Fe^{3+})$ vs. $Fe^{2+}/(Mg + Fe^{2+})$) the best fitting is the Ophiolites contour. This is not an expected result.

The compositional ranges for spinels are very similar for some categories. This can result in similar contours for different categories, which may overlap in some of the diagrams and lead to undesirable results. In this context, the Spinel Explorer++ is used to explore the two contours that appear first in the categorization list: the Ophiolites (in first place) and the Xenoliths (in second place). Figure 9(a) shows the Barnes and Roeder's contours for Ophiolites overlaid with the user-provided point data. Applying the percentile brush, it is possible to select those points, which belong to the intersection of both contours. Figure 9(b) shows a composite percentile brush, where all points belonging to both contours are selected. This selection reveals that most members of the user-provided point set are lying in the intersection of both contours. It also shows that the point set fits better into the 50th percentile contour for Xenoliths than in the 50th percentile contour for Ophiolites. In fact, exploring the quantitative information in the categorization lists, reveals that for Ophiolites 47.96% of the point set lies inside the 50th percentile contour. For Xenoliths more than the 63% of the point set lies inside the 50th percentile contour. Therefore if we give the 50th percentile threshold a value greater than 48%, the categorization list will change and the first contour in the ranking will be the Xenoliths, as expected.

A viable way to continue the analysis process is to create the density-based contours for the point set in the three projections of

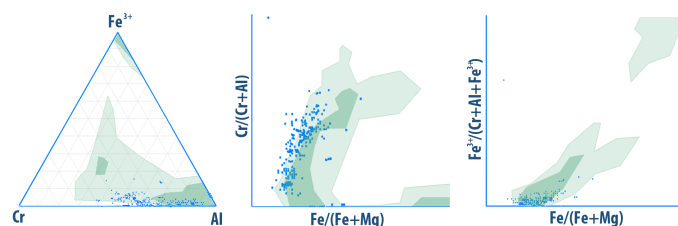


Figure 8: Barnes and Roeder's contours for Xenoliths in Basalts depicted simultaneously with the user provided point set.

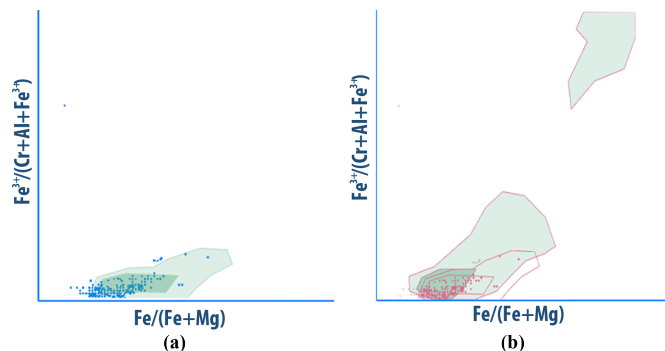


Figure 9: (a) Barnes and Roeder's contours for Ophiolites overlaid with the user-provided point set. (b) Applying a percentile brush, it is possible to select those points which belong to the intersection of both contours.

the prism. Figure 6 shows the density-based contours created for the three projections of the prism. The 50th percentile contours and the 90th percentile contours have been created.

The density-based contours for the point set reveal the existence of two different populations of points (see Figure 10(a)). In the 50th percentile contour on the base projection of the prism ($Cr/(Cr + Al)$ vs. $Fe^{2+}/(Mg + Fe^{2+})$) two compositional groups show up. For some reason these groups are concentrated in two different locations of the plots. Group A is richer in aluminum than group B. Group B, which seems to have a higher density of points, is richer in chrome.

The density-based contour-creation makes it possible to detect this kind of different compositional groups in a very short time and without the need of any additional tool. The reason for the observed cluster formation depends on the original locations of the samples, the types of host rocks, the mineral associations, and the processes that the rocks have been through the geological time. The reasoning about possible causes leads to further explorations.

Another important observation emerges from this visual analysis session. Points in the triangular and lateral projections of the prism agglomerate close to the horizontal axis because of the little amount of Fe^{3+} in the samples' composition. The lack of Fe^{3+} can be seen in the prism view, where all points are positioned close to its base. To study this deviation, the experts overlaid the density-based contours with the Barnes and Roeder's contours for Xenoliths. Now the experts can easily compare visually and quantitatively the density-based contours with the Barnes and Roeder's contours. Figure 10(b) shows as the 50th percentile contour of the user data has the same positive trend that the 50th percentile contours for Xenoliths defined by Barnes and Roeder. However, the density-based contour reveals a displacement towards the edge of the Barnes and Roeder's contour for Xenoliths. This finding points out that the Xenoliths from the user-provided point set are slightly richer in Mg than those used by Barnes and Roeder's to generate the contours.

5.2. User Feedback

After conducting the case study, a positive feedback from the domain experts indicated that our approach is useful, fast, and intuitive. In summary, the experts concluded that the newly

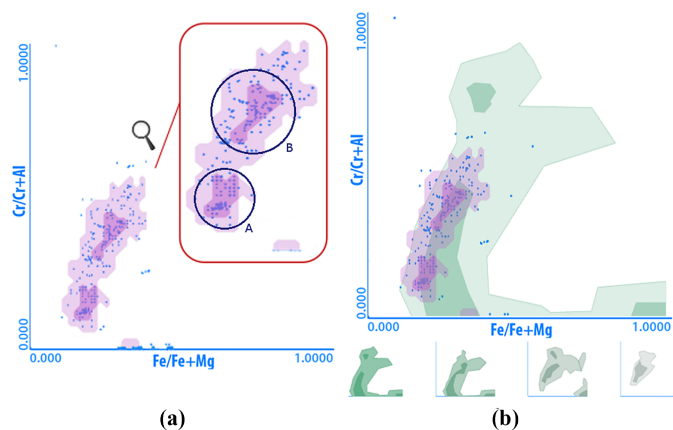


Figure 10: (a) Two compositional groups detected from density-based contour-creation. (b) Density-based contours of the user dataset overlapped with the Barnes and Roeder's contours for Xenoliths. The 50th percentile contour of the user has the same positive trend as the 50th percentile contour of Barnes and Roeder for Xenoliths. However, the density-based contour reveals a displacement towards the edge of the Barnes and Roeder's contour for Xenoliths.

introduced semi-automatic comparison of contours in Spinel Explorer++, makes it possible to confirm the experts' hypothesis very fast, and without the need of any additional tool. The time needed for the confirmation was few minutes, compared to at least an hour using the conventional approach. Moreover, the geologists are much more certain in the results, when using the newly proposed approach. They see immediately if some alternative categorization is possible. The additional drill-down by means of interactive brushing opens new, previously unavailable, possibilities for contour-based spinel-data exploration.

6. Discussion

The main contribution of the presented interactive tool is to help geologists to determine the tectonic setting based on the analysis of their data. This data are the chemical analysis obtained from carefully chosen points of an spinel rock.

A certain amount of error might be introduced in the first two steps of the workflow (samples acquisition and microprobe analysis). With the uncertainty introduced in the data acquisition and analysis processes, the need to understand the effects of those errors in other steps of the novel categorization workflow arises. Barnes and Roeder's contours are the result of a thorough study and compilation of a global dataset. The dataset has been subjected to a verification process and all data resulting from a defective acquisition or an erroneous measurement have been eliminated. Therefore, after this step no uncertainty remains and the generated Barnes and Roeder's contours are considered to be established patterns in the domain. On the other hand, correctness can not be fully assumed for the user dataset. In fact, a contour created from the user-provided point set may change with small variations in the input. Fortunately, the Spinel Explorer++ has proven to be of great help in the detection of erroneous points. Linked multiple views and interactive selections allow the expert to detect the defective points and check their

chemical composition. In these particular cases, the geologists decide whether to eliminate the points from the diagram or not.

With regard to the semi-automatic categorization process, if a hypothesis test is done, two types of errors are possible: type I (false positives) and type II (false negatives). A type I error results in the incorrect rejection of a true hypothesis, while a type II error is incorrectly retaining a false null hypothesis. In our approach, the possibility of a type I error is very small. The Spinel Explorer++ automatically ranks Barnes and Roeder's contours based on the number of points, which they include. So if the loaded dataset corresponds to a Barnes and Roeder's contour, then it will appear in the categorization list. Nevertheless, a type II error may happen if the correct contour is not ranked in the first place of the list, and the Spinel Explorer++ suggests another contour that fits better to the user dataset. This error can occur if two or more Barnes and Roeder's contours overlap and the user dataset is mostly located in the overlapping area. In this case, the Spinel Explorer++ provides the functionality to superimpose the involved Barnes and Roeder's contours. A visual inspection of the overlapping area allows the expert to detect this characteristic and to make the appropriate decision.

Finally, as this paper builds on the Spinel Explorer [3], it is important to discuss the main differences between this system, and the Spinel Explorer++ presented here. The Spinel Explorer introduced the interactive triangle plot and the 3D spinel prism exploiting the well known coordinated and multiple-views principle. Still, the spinel categorization was done based only on the observations made by geologists. The Spinel Explorer++ integrates all Barnes and Roeder's contours and, for the first time, integrates user-based point data and point-based contours with Barnes and Roeder's contours for simultaneous comparisons. In this paper we present an interactive categorization based on the Barnes and Roeder's contours. This is a major step forward, as the most cumbersome part of the workflow, i.e., the comparison of user-provided point data with the contours, is significantly improved and semi-automatized.

7. Conclusions and Future Work

Geologists analyze the spinel-group minerals in order to characterize a particular geologic region taking into account its tectonic settings. Barnes and Roeder defined a collection of contours corresponding to compositional fields for spinel-group minerals. These contours are used as empirical tectonic discriminators, enabling the estimation of the tectonic setting where a particular spinel could have formed. The current state of the art includes a tedious and error-prone manual comparison of Barnes and Roeder's contours with diagrams generated from the acquired data.

In this paper we incorporate several novel techniques into the Spinel Explorer++ in order to assist geologists in the categorization process. We introduce a semi-automatic, interactive detection of tectonic settings for an arbitrary spinel dataset. All Barnes and Roeder's contours are fully integrated, and can be depicted in three projections of the 3D spinel prism, i.e., two scatterplots and a triangle plot. The system supports interactive data categorization based on the Barnes and Roeder's contours

and provides density-based contour creation for different percentiles. This allows the experts the creation, visualization, and comparison of custom contours obtained from the user-provided point set.

The newly introduced approach and the resulting novel workflow has been evaluated using real-world data from different locations in Argentina. We received a positive feedback from the domain experts, i.e., they found the new approach fast and intuitive. They achieved the data categorization in a few minutes, a task that previously required several days.

The comparison is based on the shape of density contours, which enclose subsets of points belonging to different percentiles both in the scatterplots and triangle plots. It is important to point out that the interactive data categorization through density contours can be applied in a wide range of cases where it is required to reveal or to compare intrinsic spatial patterns. The contour approach can be basically applied to all types of scatterplots with a high point density, so that drawing individual dots would lead to heavy overdrawing. Density-based approaches through kernel-density estimation have been employed in many domains [26]. A density-based representation might be interesting to selectively calculate a few iso-contours (the percentile curves like in our case) and compare these with other contours. This would be a very general approach to high-density scatterplots. While the transformation of a point-set of a scatter plot into a density-based representations has already been done, navigating, comparing, and interacting with contours have potential.

A novel interaction, i.e., the percentile brush, was also implemented that allows the experts the selection of all points belonging to a particular Barnes and Roeder's contour by simple picking. Finally, when displaying linked data in the plots, new concave contours can be created around selected points.

At the moment, the contours of the tectonic fields are specified using only the scatterplots and triangle plots. The contours should be redefined in 3D diagrams for a better understanding of the compared datasets. Such a comparison is currently impossible with available tools, but might be useful to detect some deeply hidden information in the data. The interpretations would be difficult at the beginning, because geologists are used to work with spinel compositions only in 2D diagrams. The redefinition of the contours in 3D, together with a more formal evaluation are the main directions for future research.

References

- [1] Haggerty, S.E.. Oxide mineralogy of the upper mantle. *Reviews in Mineralogy and Geochemistry* 1991;25(1):355–416.
- [2] Barnes, S.J., Roeder, P.L.. The range of spinel compositions in terrestrial mafic and ultramafic rocks. *Journal of Petrology* 2001;42(12):2279–2302.
- [3] Ganuza, M.L., Ferracutti, G., Gargiulo, M.F., Castro, S.M., Bjerg, E.A., Gröller, E., Matkovic, K.. The spinel explorer - interactive visual analysis of spinel group minerals. *IEEE Transactions on Visualization and Computer Graphics* 2014;20(12):1913–1922.
- [4] Keim, D., Andrienko, G., Fekete, J.D., Görg, C., Kohlhammer, J., Melançon, G.. Visual analytics: Definition, process, and challenges. In: Kerren, A., Stasko, J.T., Fekete, J.D., North, C., editors. *Information Visualization*. Berlin, Heidelberg: Springer-Verlag. ISBN 978-3-540-70955-8; 2008, p. 154–175.
- [5] Keim, D.A., Kohlhammer, J., Ellis, G., Mansmann, F., editors. *Mastering The Information Age - Solving Problems with Visual Analytics*. Eurographics; 2010.
- [6] Cook, K.A., Thomas, J.J.. *Illuminating the path: The research and development agenda for visual analytics*. Tech. Rep.; Pacific Northwest National Laboratory (PNNL), Richland, WA (US); 2005.
- [7] Shneiderman, B.. The eyes have it: A task by data type taxonomy for information visualizations. In: *Proceedings of the 1996 IEEE Symposium on Visual Languages*. VL '96; Washington, DC, USA: IEEE Computer Society. ISBN 0-8186-7508-X; 1996, p. 336–343.
- [8] Bruckner, S., Soltészová, V., Gröller, E., Hladuvka, J., Bühler, K., Yu, J.Y., Dickson, B.J.. Braingazer - visual queries for neurobiology research. *IEEE Transactions on Visualization and Computer Graphics* 2009;15(6):1497–1504.
- [9] Wu, W., Zheng, Y., Qu, H., Chen, W., Gröller, E., Ni, L.. Boundaryseer: Visual analysis of 2d boundary changes. In: *Visual Analytics Science and Technology (VAST), 2014 IEEE Conference on*. 2014, p. 143–152.
- [10] Schmidt, J., Preiner, R., Auzinger, T., Wimmer, M., Gröller, M.E., Bruckner, S.. YMCA- your mesh comparison application. In: *IEEE VIS 2014*. VIS; IEEE Computer Society; 2014..
- [11] Matković, K., Gracanin, D., Klarin, B., Hauser, H.. Interactive visual analysis of complex scientific data as families of data surfaces. *IEEE Transactions on Visualization and Computer Graphics* 2009;15(6):1351–1358.
- [12] Roberts, J.C.. State of the art: Coordinated & multiple views in exploratory visualization. In: *Proceedings of the Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization*. CMV '07; Washington, DC, USA: IEEE Computer Society. ISBN 0-7695-2903-8; 2007, p. 61–71.
- [13] Wang Baldonado, M.Q., Woodruff, A., Kuchinsky, A.. Guidelines for using multiple views in information visualization. In: *Proceedings of the Working Conference on Advanced Visual Interfaces*. AVI '00; New York, NY, USA: ACM. ISBN 1-58113-252-2; 2000, p. 110–119.
- [14] Ganuza, M.L., Castro, S.M., Ferracutti, G., Bjerg, E.A., Martig, S.R.. SpinelViz: An interactive 3d application for visualizing spinel group minerals. *Computers & Geosciences* 2012;48(0):50 – 56.
- [15] Gaines, R.V., Skinner, H.C.W., Foord, E.E., Mason, B., Rosenzweig, M.. *Dana's New Mineralogy*. John Wiley & Sons Inc; 1997.
- [16] Gargiulo, M., Bjerg, E., Mogessie, A.. Spinel group minerals in metamorphosed ultramafic rocks from río de las tunas belt, central andes, Argentina. *Geologica Acta* 2013;11(2):133–148.
- [17] Haines, E.. Graphics gems iv. chap. Point in Polygon Strategies. San Diego, CA, USA: Academic Press Professional, Inc. ISBN 0-12-336155-9; 1994, p. 24–46.
- [18] Ho, C., Wu, F.C., Chen, B.Y., Chuang, Y.Y., Ouhyoung, M.. Cubical marching squares: Adaptive feature preserving surface extraction from volume data. In: *Computer graphics forum*; vol. 24 (3). Wiley Online Library; 2005, p. 537–545.
- [19] Weber, G.H., Hauser, H.. *Interactive visual exploration and analysis*. In: *Scientific Visualization*. Springer; 2014, p. 161–173.
- [20] Kehrer, J., Hauser, H.. Visualization and visual analysis of multifaceted scientific data: A survey. *IEEE transactions on visualization and computer graphics* 2013;19(3):495–513.
- [21] Becker, R.A., Cleveland, W.S.. Brushing scatterplots. *Technometrics* 1987;29(2):127–142.
- [22] Cleveland, W.C., McGill, M.E.. *Dynamic graphics for statistics*. CRC Press, Inc.; 1988.
- [23] Ward, M.O.. Xmdvtool: Integrating multiple methods for visualizing multivariate data. In: *Proceedings of the Conference on Visualization'94*. IEEE Computer Society Press; 1994, p. 326–333.
- [24] Martin, A.R., Ward, M.O.. High dimensional brushing for interactive exploration of multivariate data. In: *Proceedings of the 6th Conference on Visualization'95*. IEEE Computer Society; 1995, p. 271.
- [25] Radoš, S., Splechtina, R., Matković, K., Duras, M., Gröller, E., Hauser, H.. Towards quantitative visual analytics with structured brushing and linked statistics. In: *Computer Graphics Forum*; vol. 35 (3). Wiley Online Library; 2016, p. 251–260.
- [26] Lampe, O.D., Hauser, H.. Interactive visualization of streaming data with kernel density estimation. In: *Pacific Visualization Symposium (PacificVis), 2011 IEEE*. IEEE; 2011, p. 171–178.