A Visualization and Analysis Method by Multi-Dimensional Crossplots from Multi-Well Heterogeneous Data

Maojun Cao, Zhiyong Gao, Ye Yuan, Zhao Yan and Yihong Zhang

https://doi.org/10.3390/en15072575
A Visualization and Analysis Method by Multi-Dimensional Crossplots from Multi-Well Heterogeneous Data

Maojun Cao 1, Zhiyong Gao 1, Ye Yuan 2,*, Zhao Yan 1 and Yihong Zhang 1

1 School of Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China; caomaojun@nepu.edu.cn (M.C.); gzy@stu.nepu.edu.cn (Z.G.); zyan@stu.nepu.edu.cn (Z.Y.); zyh@stu.nepu.edu.cn (Y.Z.)
2 Research Institute of Petroleum Exploration and Development, Petrochina, Beijing 100083, China
* Correspondence: yuanye_riped@petrochina.com.cn

Abstract: Crossplot is an important tool for data visualization analysis in well logging processing and interpretation. Generally, the color of the points in the crossplot is calibrated by the data to improve the plane crossplot from two dimensions (2D) to three dimensions (3D). However, this method is limited to the logging data. In many cases, it is necessary to use crossplots for comprehensive evaluation with the well logging, geology, mud logging, and oil testing data, or to deeply analyze the data from multiple wells and multiple layers. The traditional crossplots cannot meet the demands. To expand the traditional crossplot to multiple dimensions, we proposed an augmented-dimensional visualization and analysis method by the crossplot with multi-well and multi-dimensional heterogeneous data, which is developed based on the traditional crossplots. Firstly, we built the match matrix of the augmented-dimensional heterogeneous data from different depths and set up the crossplots of the heterogeneous data after depth normalization. In this way, the attribute calibrations of the points in the crossplots by logging data were realized, and the auxiliary figures of the augmented-dimensional crossplots were established. In addition, the crossplots were expanded to more dimensions by collaborative analysis of the multiple crossplots and synchronous display in different single-well and multi-well modules with the feature points projection. Secondly, we established the display method of the crossplots from multi-well data by quadtree index. The display and interaction performance of the augmented-dimensional crossplots under the conditions of large amount of data from multiple wells were greatly improved by optimizing the overlapped and covered points in the crossplots. This method can provide more information about the reservoirs and realize the comprehensive well logging comparison, analysis, and visualization from different views. It has been developed and integrated into a CIFLog software platform and has been widely used in many domestic oilfields.

Keywords: CIFLog; crossplot; multi-well; heterogeneous data; augmented-dimensional depth match matrix; quadtree

1. Introduction

Crossplot is an evaluation technique that uses mathematical statistics to express the relationship between formation logging parameters. In logging interpretation and data processing, the commonly used crossplots include crossplots plate, frequency crossplots, Z-value crossplots, and histograms. Crossplot is an interpretation chart used to represent the relationship between two logging parameters of lithology. They are all theoretical charts based on the logging response relationship of rocks, such as lithology-porosity logging crossplots, M-N and MID crossplots used to identify formation lithology, and crossplots used to identify clay minerals and other minerals in the formation. The frequency crossplot is an intuitive digital chart that counts the values of A and B curves of each sampling point on the well section in the x-y plane and falls on the number of sampling points.
(i.e., the frequency number) of each unit grid, referred to as the frequency crossplots. The Z-value crossplot is a data chart made by introducing the third curve Z (called Z curve) on the basis of the frequency crossplots. The number of Z-value crossplots represents the average level of the third Z of the corresponding sampling point in each unit grid on the frequency crossplots of the same well section. Crossplots can check the quality of the logging curve, correct the curve, determine the formation lithology combination, identify the formation mineral composition, analyze the pore fluid properties, determine the formation interpretation model, calculate the logging interpretation parameters, and calculate the geological parameters of the formation. Crossplots play a significant role from the beginning to the end of the logging processing and interpretation process.

In 2004, Liu Xianhong and Huo Hao [1] applied the background deviation distance method in AVO crossplot analysis. In 2006, Wu Lei, Xu Huaimin et al. [2] established a sample set by crossplots based on geochemical and petrological studies of core samples, and used cluster analysis and distance discriminant method to determine the initial weight. In 2010, Liang Limei and Yu Gaoming and others [3], based on BP neural network, designed a neural network simulation crossplot chart for the development of medium and late reservoir fluid identification. In 2011, Yu, Guang, Aguilera, and Roberto [4] proposed a practical method based on Pickett crossplots for preliminary and accurate quantitative evaluation of shale gas formation. In 2012, Xu and Li et al. [5] used the crossplot method to identify lithology and fluid type of M oilfield abroad. In 2016, Ranye, Wang Guwen et al. [6] quantitatively characterized diagentic facies of tight oil reservoirs by logging crossplots. In the same year, Siyamak Moradi, Mohammad Moeini [7] determined shale volume distribution and effective porosity by crossplots. In 2017, Wang Xiaoqiang and Zhang Jun [8] proposed a conventional logging identification method of fracture-cavity type based on the decision tree of crossplots combined with the differences in lithology and physical properties between different types of fracture-cavity. In the same year, Junhwan Choi, Bona Kim, et al. [9], proposed a probabilistic phase analysis method using the 3D crossplots of stochastic forward modeling results and solved the problem when the results of stochastic forward modeling were plotted on the 2D crossplots. Problems where the boundaries of the probability density function can overlap each other are indistinguishable. In 2018, Leti Teklu Wodajo’s [10] research developed a crossplot analysis using time-lapse seismic refraction tomography and resistivity tomography to evaluate earth dams and levees, incorporating their unique seismic resistance and resistivity properties into the crossplot analysis. In January of the same year, Samit Mondala, Ashok Yadav, et al. [11] studied the petrophysical characteristics of different depositional environments, the influence of key geological factors on cross-plot trends, and the established depositional models in the drilling area. Finally, they reduced the uncertainty in characterization of untapped potential reservoirs. In March of the same year, OE Austin, OE Agbasi et al. [12] used the data of Well X-26 in an oilfield in the southwestern coastal sedimentary belt to analyze the crossplots of rock properties in the Niger Delta oilfield to identify fluids and lithology. In June 2019, P. Connolly [13], by using the normalized elastic property reflectance pair as the base vector, determined that the position of the reflectance point in the intercept gradient space can be directly determined from the elastic property comparison, which can more intuitively understand the structure of the intercept gradient crossplot. C. G. Okeugo, K. M. Onuoha [14] et al. used cross-plots to distinguish lithofacies and predict fluids for accurate positioning of new wells. In July, Un Young Lim, Richard L. Gibson, and others [15] proposed a new method of crossplots based on the Zoeppritz equation, and applied this method to the estimation of total organic carbon (TOC) content in shale. In November, Leti T. Wodajo [16], discussed the use of crossplot analysis combined with laser resistivity tomography and seismic refraction tomography to assess the integrity of earth dykes and dykes applied to the Francis Levee site Geophysical surveys conducted. In 2020, Dr Muhsan Ehsan et al. [17] used the Talhar shale as an example to identify lithofacies and clay mineralogy types through crossplots. Ahmed E. Radwan [18] used the crossplot to estimate the type of oil and gas by using the synthetic acoustic curve, the
original density curve, the original neutron curve, their porosity, and variable values such as M, N, DTMAA, and RHOMAA. Finally, take the Gulf of Suez in Egypt as an example for application. In May of the same year, J Choi [19] proposed a cross-plot method to help distinguish potash ore, and the method was applied in the basins of Michigan, Nova Scotia, Saskatchewan, and southeastern New Mexico. In 2021, MT Zakaria and N Mohd Muztaza [20] used the crossplot analysis method, combining seismic refraction and 2D resistivity to study landslide activity in the local area. In the same year, GB Azuoko, A Ekwe et al. [21] differentiated the reservoir contents into fluid and lithology through crossplots, and carried out detailed crossplot analysis of petrophysical properties.

However, from the perspective of data type, display type, and auxiliary interpretation ability, the crossplot needs to solve the following problems. (1) Various logging data, including conventional continuous logging curve data of logging data well logs, discrete data, logging table data, oil test data, mud logging data, and other different logging data. (2) The way of crossplot is complex. Due to the variety of crossplot data, it is necessary to support different types of crossplots, such as: crossplot of continuous well logs with different depth sample levels, crossplots of continuous logging curve and discrete data, crossplots of discrete data and logging table data, and so on. (3) The display and interactive speed of the crossplot is required to improve during logging processing and interpretation. In addition, complex auxiliary interpretation and interactive operations in crossplot are needed. Therefore, in the case of large amount of data, an efficient display and fast interactive technology in crossplot is the key problem of multi-well and multi-dimensional auxiliary interpretation. In this paper, the augmented-dimensional depth match matrix of heterogeneous data is constructed. Through the study of the core components of the augmented-dimensional depth match matrix, the functions of logging data to scale the properties of crossplot points, such as color light and darkness and symbol size, and the augmented-dimensional auxiliary map of crossplot are realized. The depth selection in the augmented-dimensional depth match matrix realizes the cooperative analysis of multi-crossplots. The feature points of the crossplot are projected to other multi-dimensional application modules for synchronous display, and more heterogeneous data are used to assist crossplot analysis. Based on mathematical function, basic operator and logic operator analysis technology, more logging curves are used to scale the crossplot, filter the crossplot data points that meet the conditions, and expand the display dimension of the crossplot. The display method of the crossplots from multi-well data by quadtree index is constructed to optimize the overlapping and occlusion data points in the crossplot and improve the dimensional display and interactive performance of the crossplot under the condition of large data amounts of multi-well.

The second section of this paper expounds the basic principle of crossplot augmented-dimensional display analysis and designs the theoretical model framework. The third section gives the key technologies and implementation details of crossplot augmented-dimensional display analysis method in detail. The fourth section shows the application effect of crossplot augmented-dimensional display. The fifth section and last section conclude the full text and propose further work. At present, this method has been fully integrated into CNPC’s new generation logging software CIFLog [22–28], and has been applied in Daqing and Xinjiang oilfields.

2. Analysis Model of Augmented-Dimensional Display of Crossplots

The multi-well and multi-dimensional heterogeneous crossplot augmented-dimensional analysis method refers to the use of multi-well logging data to increase the extended crossplot display dimension (dimension > 3) by integrating all available data, such as geology, logging, and oil testing. Other application modules are used as the extended dimensions to assist crossplot analysis. At the same time, interpretation charts and augmented dimension auxiliary figures are provided to provide more reservoir reference information for logging interpretation. Among them, heterogeneous data refer to data with different sources and structures.
The bottom layer of CIFLog platform adopts CIFPlus data storage format \[29,30\]. The bottom layer of this format uses table structure and data block to store data, in which the physical structure takes 4096 bytes as its unit, and each 4096 bytes is called a record block. The CIFPlus data storage format can store data such as logging, geology, logging, and oil testing, and supports the expansion of any heterogeneous data. At the same time, this format can sample logging data at equal and unequal intervals. Complex types of data can be stored, which can meet future needs to store unknown storage types. Therefore, on the basis of this storage format, this method automatically identifies heterogeneous data of different categories and loads it into the crossplot for dimensionally increased display.

In order to realize the augmented-dimensional display analysis of multi-well heterogeneous crossplot, this paper designs the model framework of Depth Matcher-Data-Model-View-Controller (DDMVC) crossplot, as shown in Figure 1. After more heterogeneous data of different categories are normalized according to depth, the augmented-dimensional display on the crossplot is realized. The framework encapsulates each function of the crossplot into independent module units, which are interconnected through the defined standard service interface. The internal function expansion of the module unit does not affect other module units within the framework, which realizes the loose coupling and scalability of the model framework.

**Figure 1.** The visualization and analysis model framework of augmented.

Dimension of multi-well heterogeneous crossplot DataSource is a data source module. By default, the CIFPlus storage format reading and writing interface in CIFLog platform is encapsulated, and other data storage formats can also be extended, even data in the database.

DepthMatcher is a dimension-increasing depth alignment matrix module. First, the depth range is displayed according to the needs of the crossplot, the depth range is intercepted, the heterogeneous data of different types in the depth range is uniformly matched according to the depth and interval, and then the data is loaded into this module to form a data grid matrix. The DepthMatcher module saves the unified depth as an index in the matrix. When the data feature points are selected on the crossplot, they can be matched to the corresponding depth dimension. Subsequently, the Controller module is used to send the selected depth to the system through the message queue. Other application modules perform response processing to realize functions such as synchronous display and analysis with other logging application modules. The Controller receives the interactive operation messages of other application modules and displays them in the rendezvous diagram. The Map data mapping module establishes a data index for the deeply aligned data. Map does not need to care about the source and category of heterogeneous data and establishes the mapping relationship between the data and the view rendering module view through Map. View is a crossplot view rendering module, which performs drawing rendering according
to the index and data provided in the Map. For example, the first column of data in the grid corresponds to the X-axis of the crossplot, the second column of data corresponds to the Y-axis, the third column corresponds to the color axis, the fourth column corresponds to the symbol size scale, and the fifth column corresponds to the color intensity scale, etc. Realize the mapping of multi-dimensional logging data to view for display. The Controller module is used to respond to event messages such as user interaction and data refresh, such as selecting abnormal points on the crossplot and projecting them to the logging diagram to view abnormal intervals.

3. The Realization of Augmented-Dimensional Display Technology of Crossplots

3.1. Approach

In order to solve the problem of augmented crossplot analysis between different heterogeneous data, this paper constructs the augmented-dimensional depth match matrix technology under the augmented display analysis model of crossplots and realizes the unified calibration of heterogeneous data based on depth for multi-dimensional display of crossplots. At the same time, this paper constructs the filter technology to provide more curve expression calculation and logical condition logging filter, realizes the calibration and condition logging filter of crossplots display points by using more heterogeneous data, and filters the crossplot points in the abnormal range of crossplots. This paper established the display method of the crossplots from multi-well data by quadtree index and improved the efficient display of multi-well and multi-dimensional heterogeneous crossplot and the ability of fast interactive picking of data points. In addition, the multi-well and multi-dimensional heterogeneous crossplot is analyzed for more flexible and convenient auxiliary users. Finally, through the crossplot communication technology, the crossplot can play the greatest role and realize the use of multiple wells to provide help for the evaluation of complex reservoirs.

3.2. Augmented-Dimensional Depth Match Matrix for Heterogeneous Data

3.2.1. Alignment Grid Matrix

In the conventional crossplot, the attributes such as the symbol size of data points and the brightness and darkness of color are all fixed values. In order to make full use of the two-dimensional plane crossplots to display more logging reference information and use heterogeneous data to scale these attributes, this paper constructs the augmented-dimensional depth match matrix technology. According to different heterogeneous data categories, the data are formed into alignment grid matrix according to the depth, and then the matrix data are dynamically allocated to the crossplot view. The view calibrates these display attributes according to the heterogeneous data value and realizes the augmented-dimensional display of the crossplot. Based on the grid matrix technology, the crossplot can use the depth dimension to realize the collaborative analysis of the multiple crossplots and synchronous display in different single-well and multi-well modules with the feature points projection to improve the display dimension. For example, feature points are selected in the crossplots. The augmented-dimensional depth match matrix is mainly divided into four core components, including depth alignment area, heterogeneous data loading area, data filter and depth, and data matrix cache. According to the feature points, the corresponding depth intervals can be projected in the multi-well cross-section, and the augmented-dimensional depth match matrix framework is shown in Figure 2.
3.2.2. Deep Alignment Area

The depth alignment area is based on different heterogeneous data categories. It instantiates the corresponding depth loader and adds it to the heterogeneous data depth loading queue, including interpretation conclusion depth loader, logging data depth loader, core analysis depth loader, etc. The depth loader is defined as an interface that supports subsequent heterogeneous data category extensions without affecting other rendezvous function module units. In order to improve the depth alignment efficiency, the depth loading queue of heterogeneous data dynamically adjusts the priority order according to the loaded data category and depth range. The priority of heterogeneous data queue is discrete data with the least depth point > tabular data with the least depth point > discrete data > tabular data > continuous curve with the least depth range > continuous curve. Priority queues ensure deep matching from the least heterogeneous data, reduce the number of traversals (in turn), and improve processing speed.

3.2.3. Data Filter Approach

Data filter is based on the aligned depth results, supports multiple logging filter conditions, limits the range of heterogeneous data values, and only retains data points that meet the filter conditions for display analysis, fitting and statistics. For example, the filter condition ‘GR0’ can be set in the crossplots, that is, only data points with natural gamma values less than 60 and natural gamma values greater than 0 are displayed on the crossplots.

Filters provide string expression parsing that supports ‘and’, ‘or’, and ‘not’ multi-condition logging filter. The effective logging filter of the data on the crossplot is realized, and one or more logging curves in the well are used to filter out the crossplot points that are not concerned. The technology mainly realizes the following functions:

1. Unit conversion. When logging data is inconsistent with the unit of crossplot scale, unit conversion of logging curve is needed. For example, the unit of density logging curve (DEN) is kg/cm$^3$, but the current crossplot is calibrated according to g/cm$^3$, so the unit conversion of logging curve is needed, and the curve expression is entered in the crossplot ‘1000 * DEN’.

2. Curve correction. When the curve is corrected, the data is projected to the interpretation chart of the crossplot. When the distribution of the vast majority of data points is
dense but deviates from the lithology line of the interpretation chart, the additional correction of the curve is needed. That is, the additional correction of the x-axis or y-axis curve in the crossplot is added to the $\Delta x$ and $\Delta y$, so that most data points are located near the known lithology line of the interpretation chart.

(3) Multi-curve conditional logging filter and curve effective calibration range. For example, natural gamma logging is a method to measure the natural gamma ray intensity of rock strata. Rocks generally contain different amounts of radioactive elements, and constantly emit rays. The more argillaceous they contain in sedimentary rocks, the stronger their radioactivity is. Interpreters pay more attention to the distribution of sandstone and filter out possible mudstone data points. Therefore, logging filter conditions can be set in the crossplot ‘GR < 60&&GR > 0’. That is, only crossplot points with natural gamma values less than 60 and greater than 0 are displayed.

However, there are the following difficulties in the construction of expression analysis technology:

(1) Expression analysis is complex. In the analysis process, the logging curve name, operator, and mathematical function in the expression need to be identified and classified. When the curve name contains operators, or the curve name is a function keyword, the analysis is more complex.

(2) Expressions should not only support operators, but also support a variety of functions, even logical expressions, and also consider the priority and operation order.

(3) There are many logging data points and a large amount of calculation, so the calculation speed is required.

Based on the above requirements and difficulties of expression analysis and calculation, combined with the characteristics of logging data, this paper studies the expression analysis and calculation technology, provides support for a variety of mathematical functions and operators, and realizes efficient and fast curve expression calculation function.

Expression Parsing

The mathematical expression includes variables, constant brightness, operators, and functions. The purpose of analysis is to extract characters from the expression, determine the meaning of each character, and form a data structure that can be calculated later. Finally, the data is calculated. Expression analysis mainly includes lexical analysis, syntax analysis, and data structure generation.

The main work of the method analysis is to extract keywords and judge whether each keyword is correct, for example, whether the function name is correct, whether the variable name has started in letters, etc. The main task of grammar analysis is to judge the whole mathematical expression, for example, two operators in the expression are adjacent and other errors. The data structure for subsequent calculation is formed in the process of syntax analysis. Expression parsing has two main storage data structures: stack structure and binary tree structure [31–33]. Among them, for some complex expressions, for example, when the function parameters are uncertain, the method can be nested, and the priority of different operators is different, so it is difficult to use the stack top structure. Therefore, this paper uses the binary tree structure to store the expression. The process of expression parsing is the process of constructing the binary tree reservoir structure. The expression parsing flowchart is shown in Figure 3.
Suppose the mathematical expression is $1000 \times \text{DEN} + \text{Max}(60, \text{GR})/2$

Where DEN and GR are logging curves, and Max is the maximum function. The generated expression binary tree is shown in Figure 4.

![Binary tree structure generated by expression.](image)

Figure 3. The flow chart of expression analysis.

When the logging curve contains a keyword or mathematical operator, braces “{” and “}” are added to the keyword. For example, when the natural gamma name is Gr-1, the curve name contains the operator “-”, so it becomes [Gr-1] when writing the expression. At this time, the lexical analyzer will take the operator in braces as the curve name by default, and the lexical check is correct.

Expression Evaluation

After the binary tree is formed, the expression can be calculated. From the binary tree structure chart generated by the expression, it can be seen that the parent node is an operator and the leaf node is an operand. Therefore, the calculation of the expression can be completed by using the middle order traversal of the binary tree, and the operation speed is fast. The parsing process of the expression parsing library is shown in Figure 5.
Taking the Z-value crossplot of well GS1 in Central Sichuan as an example, the display results after setting the filter condition “GR < 60” are shown in Figure 6.

(a)

Figure 6. Cont.
Figure 6. The display result of crossplots data filter. The X-axis represents the compensated neutron log (CNL) and the Y-axis represents the acoustic time (AC). (a) The crossplot display results before no filter conditions are set. (b) The crossplot display results after setting the filter conditions.

3.2.4. Heterogeneous Data Loading Area

As shown in Figure 7 below, the heterogeneous data loading area reads the data according to the depth after alignment and the heterogeneous data storage structure, and finally forms the depth and data matrix cache to provide data for the crossplot data mapping Map module unit. The Map module constructs the data index according to the data column in the depth and data matrix cache and displays the location of the corresponding data through the index query when the crossplot is refreshed or interacted.

Figure 7. Heterogeneous data loading flow chart.

3.2.5. Depth and Data Matrix Cache

The cache area mainly stores all kinds of information related to data communication and interaction in memory [34–36], including data index, adjacent well information, processing parameters, intermediate results and resources, etc. The cache area manages the
multi-well and multi-dimensional heterogeneous data uniformly. It can be seen from the schematic diagram that after receiving the data access request, the cache area first calculates the location and size of the data by verifying the address area, and then searches the data in the search area. When the corresponding data are found, it is hit, and can directly read the cache data. If not hit, you need to use the loading area to load the data into memory. The loading area first verifies whether the cache area is full. If full, first clean up the data by replacement policy, then load the data, and finally read the cache data. The cache structure principle is shown in Figure 8.

Figure 8. Schematic diagram of heterogeneous data processing flow of augmented-dimensional depth match matrix.

It provides temporary cache data support for multi-well multi-dimensional modules. The cache information is shown in Table 1.
### Table 1. Cache Information Table.

<table>
<thead>
<tr>
<th>Cache Information</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Curve list</td>
<td>Curve data used in multi-well multi-dimensional module processing</td>
</tr>
<tr>
<td>Layer list</td>
<td>Multi—well Multi—dimensional Module Processing Deep Layer</td>
</tr>
<tr>
<td>Exegetical mode</td>
<td>Interpretation Model Formula for Multi—well Multi—dimensional Module Processing</td>
</tr>
<tr>
<td>Processing parameters</td>
<td>Processing parameters formed by multi-well multi-dimensional module processing, such as skeleton density</td>
</tr>
<tr>
<td>resource information</td>
<td>Resource information for multi-well and multi-dimensional use, such as drawing templates, color labels, etc.</td>
</tr>
<tr>
<td>Adjacent well date</td>
<td>adjacent well test results, lithology, physical properties, water analysis data and other information</td>
</tr>
</tbody>
</table>

Based on this caching technology, the data cache loading and data processing process refer to Figure 9 and description below.

![Figure 9. Heterogeneous data processing flow of alignment matrix of augmented dimension depth.](image)

The main design objectives of the cache area are as follows:

1. Management of current interpretation and processing data flow, and coordination of multi-well and multi-dimensional module interactions

   In the process of module processing, the buffer area will cache the processing area, data source, curve, and other information, and also temporarily store the intermediate results or processing parameters obtained by calculation. Other modules can obtain the intermediate results and processing parameters in the buffer area, display them in real time, and share data between modules.

2. Improving access efficiency

   In the process of multi-well multi-dimensional interaction and display, multi-well work area data will be used, which usually contain dozens or even hundreds of well data. Especially when the well contains big data such as imaging, it directly accesses the data in the disk file every time, which is slow and seriously affects the efficiency of multi-well
multi-dimensional interaction. Therefore, the cache area opens up an independent memory space for multi-well multi-dimensional interaction of the platform and caches the attribute information and curve information of the work area to memory. When the module accesses the cache area, the cache area retrieves the data in memory. If not hit, the cache area will load the key data into memory for management and scheduling, and then return the data to the application from memory. When the same data is accessed later, it can be read directly from memory, which effectively improves the access efficiency.

(3) Ensuring multi-well and multi-dimensional heterogeneous data consistency

The platform shares the cache data. When a module modifies the data, the cache will force other modules to synchronize the data, which effectively ensures the consistency of the data.

Figure 9 shows the multi-dimensional data loading and processing data flow diagram of the crossplot of a well in Daqing Oilfield. Based on the dimension-increasing depth alignment matrix, the data depth in the well is first intercepted according to the depth range of the interval, which greatly reduces the amount of data loaded into the memory. Read the discrete curve depth, table data depth, GR conventional curve depth, and other conventional curve depths according to the depth of the interval, and regularize them according to a uniform sampling interval to form a data depth matrix, and then filter the data values through the data filter, the filter only selects data with porosity > 0.02 for crossplot display, and finally forms a grid data matrix to provide multi-dimensional logging data display for the crossplot. Under multi-well conditions, each well corresponds to an increased-dimensional depth alignment matrix, and the data in the matrix is partially cached in memory according to the depth index to improve data access efficiency.

3.3. The Display Method of the Crossplots from Multi-Well Data by Quadtree Index

There are two main factors that affect the response speed of crossplot: (1) access to hard disk logging data; (2) Data points of multi-well crossplot are shown. The CIFLog platform provides CIFPlus data caching technology. The augmented-dimensional depth match matrix technology constructed in this paper can effectively reduce the frequency of accessing hard disk data and improve the efficiency of data access. However, the crossplots data points displayed, especially in the case of multiple wells and large amounts of data, need to traverse each data point for rendering. When selecting feature points on the crossplots, it is necessary to determine whether each data point is in the selection area, which seriously reduces the display and interaction performance of the crossplots.

Through the study, it is found that most of the data points are concentrated in a small area, and the data points are occluded and overlapped with each other. The repeated drawing of occluded points and overlapped points seriously reduces the performance of the crossplot. Therefore, this paper proposes the display method of the crossplots from multi-well data by quadtree index. The basic principle is based on quadtree spatial index. The basic idea of quadtree index is to divide the geographic space recursively into different levels of tree structure. It divides the space of the known range into four equal subspaces recursively until the tree level reaches a certain depth or meets a certain requirement. The structure of quadtree is relatively simple, and when the spatial data objects are evenly distributed, it has relatively high efficiency of spatial data insertion and query. Therefore, quadtree is one of the commonly used spatial indexes in GIS. Quadtree has high efficiency for region query.

Therefore, the two-dimensional plane of the crossplots is equally divided into four equal subspaces. According to the distribution of data points, the subspace is recursively divided into multi-well quadtree subspaces at different levels, so that it is recursively continued until the tree level reaches a certain depth or meets some requirements. The crossplot points are stored on the leaf nodes, while the intermediate nodes and the root nodes are used to store the statistical information of the sub-nodes, such as the total number of leaf nodes, the average value of leaf nodes, and the maximum and minimum values.
Figure 10 shows the distribution diagram of only eight data points on the crossplot: the left diagram is the distribution diagram of data points on the plane crossplot, and the right diagram is the tree storage structure of data points. The multi-well quadtree subspace is divided recursively according to the distribution of data until the subspace reaches the crossplots display accuracy, that is, multiple data points are inserted into the subspace and not divided. At this point, when the subspace contains multiple data points, only the last inserted data points are displayed, and the other points are not displayed as cover points, which can filter out the occlusion and cover data points to the maximum extent. Especially in the case of large number of wells, this can greatly improve the crossplot data point display speed and interactive performance.

Based on the display method of the crossplots from multi-well data by quadtree index, the feature points on the crossplot can be selected by obtaining all the grid cells intersecting the selected area and the grid, and then traversing the intersecting grid cells to obtain the feature points. The user selects data points through ellipses in the crossplot, as shown in Figure 11. The crossplot traverses the multi-well quadtree index, and only node a intersects the selection area. Therefore, excluding the other three intermediate nodes, it further traverses all the child nodes under node a, and finally obtains the selected data points a and b. This method can greatly reduce the number of traversal data and the amount of calculation and improve the selection and interaction speed of cross plot feature points.

3.4. The Crossplot Communication Technology

The crossplot communication technology uses other logging figures to increase the display dimension of the crossplot. Especially in the case of complex formation conditions, it can jointly reflect the real information of the formation from more angles, so that the crossplot can use other logging figures to assist in displaying more dimensional information, such as: (1) The Crossplot Analysis, which selects abnormal points in the cross-plot, and use other logging figures to assist in analyzing the unreasonable causes of abnormal points. (2) The Neutron-Density Crossplot Discriminating Lithology Interpretation Plate,
according to the distribution of lithologic lines where data points are located to distinguish different lithologies, which selects specific lithologic data points on the crossplot to check the corresponding depth of the data in the logging figures. Imaging curve responses are all informative, further assisting in fine logging interpretation. (3) The Neutron-Density Crossplot, which mainly is used to determine the formation porosity and discriminate lithology, and the porosity-resistivity crossplot is used to determine the fluid properties, so the relevant lithology data points can be selected in the neutron-density crossplot and determine the fluid properties on the porosity-resistivity crossplot.

The crossplot communication technology utilizes the message event processing methods, which excluded three functional components: event source, event object, and event listener. The event source is the source that triggers the event. When the event source generates an event, the event listener can listen to the event and respond to the event. The event object is used to pass event data. The event listener is responsible for listening to the events emitted by the event source, and when it listens to the event, it will automatically call the corresponding method for processing. The message event processing process is as follows: first, the listener is registered with the event source, when the user triggers the event source event, the event source generates an event message object and sends it to the event listener, and finally the event listener notifies the module processor to process the crossplot communication technology the schematic is shown in Figure 12.

Figure 12. Schematic diagram of the event listener model.

As the event source, the crossplot can be quickly selected based on the display method of the crossplots from multi-well data by quadtree index in this paper, and the corresponding depth index can be queried according to the selected crossplot data points, and the depth index sequence can be established. It is encapsulated into an event object and sent to all figures for corresponding processing. Other figures are used as event listeners to wait to receive event objects, so that the crossplot can use other figures to respond to the selected data in real time.
4. Application Effect

4.1. Method

At present, this method has been applied to the integrated logging processing and interpretation platform CIFLog. This paper gives the following application examples: (1) Carrying out dimension increasing crossplot analysis on multi-well and multi-sections in well log, construct multi-well and multi-section crossplot, and display multiple layers and multi-dimensional data in the same crossplot; (2) The productivity prediction crossplot is established to compare the multi-dimensional attribute characteristics of reservoirs between different wells; (3) Other application modules of the platform are used to assist the cross plot augmented-dimensional display and analysis, so as to accurately reflect the real formation information and provide more reservoir reference information for interpreters.

4.2. Analysis of Dimension Increasing Crossplot of Multi-Well and Multi-Sections

The common crossplots analysis tool is to add color dimension to the two-dimensional plane crossplots and display three-dimensional logging data, such as Z-value crossplot. However, under multi-well and multi-sections, the Z-value crossplot cannot scale the information of multi-well and multi-sections. Therefore, based on the multi-well crossplots augmented-dimensional display method proposed in this paper, the neutron-density-acoustic time difference-natural gamma crossplots of multi-well and multi-sections can be constructed. As shown in Figure 13, color is used to scale multiple wells, symbolic attributes to scale logging intervals, GR to scale color brightness, and AC to scale symbolic dimensions. At the same time, in order to facilitate the analysis of multi-well interval data, multi-well multi-crossplot layered section display is realized. The two-dimensional plane crossplot can scale more multi-dimensional logging information and can be better analyzed on multi-well and multi-section crossplots.

Figure 13. Cont.
Figure 13. Results of augmented dimension crossplot for multiple wells and multiple layers. (a) Multi-well, multi-layers, and multi-dimensional data are displayed in the same crossplot. The X-axis represents the compensated neutron log (CNL) and the Y-axis represents the acoustic time (AC). (b) Independent display of multi-well and multi-section augmented dimension crossplot. The X-axis represents compensated neutron log (CNL) and the Y-axis represents the acoustic time (AC).

4.3. Productivity Prediction Crossplot

Productivity prediction is a basic task of reservoir logging evaluation in oil and gas exploration. Under similar stratigraphic background and engineering technical conditions, the electrical and physical characteristics of the reservoir are the most important factors to control productivity. A conventional crossplot can only add color dimension on the two-dimensional plane for three factor analysis and cannot realize the comparison of multi-dimensional attribute characteristics of reservoirs between different wells. The crossplot augmented-dimensional display analysis technology proposed in this paper can effectively solve this problem and realize the comparison of electrical properties, physical properties, reservoir thickness, and other characteristics between target well reservoir and reference well reservoir, and provide technical support for reservoir productivity prediction of target wells. Figure 14 shows the Carboniferous block of Xinjiang Oilfield. In the multi-well crossplot, all tested Carboniferous oil wells and target well DT1 (the first crossplot on the upper left) are established for multi-well similarity analysis. The target well DT1 is projected into the tested oil wells by using the multi-well multi-dimensional crossplot projection technology, and the gray point in the figure is the projection point. According to the ranking of similarity from high to low, it can be seen from the figure that well DT1 has the highest similarity with well C54, so the reservoir parameters of target well DT1 are predicted.

As shown in Figure 15, the scatter area is shown by the conventional crossplot (the left area shown in Figure 15), and the green point and red point are the porosity-resistivity distribution of the reference well (C54 well) and the target well (DT1 well). The productivity prediction chart (the right area shown in Figure 15) can focus on the comparison of reservoir porosity-resistivity distribution, reservoir lower limit interpretation chart, and reservoir effective thickness. By calculating the reservoir gravity center, the reservoir similarity is compared by using the distance between the reservoir gravity center of the target well and the reference well. The interpretation chart of reservoir lower limit is the chart of porosity and electrical lower limit determined by adjacent well data. The blue area is water layer, the red area is oil layer, and the brown area is dry layer. The effective thickness of the target well is compared with the effective thickness of the reference well to
determine the effectiveness of the target well reservoir. Through the comparative analysis of reservoir productivity with reference wells, the productivity of the target reservoir can be directly predicted.

**Figure 14.** Application results of augmented dimension crossplot for multiple wells and multiple layers in Xinjiang oilfield. The X-axis represents the porosity (POR) and the Y-axis represents the true formation resistivity (RT).

**Figure 15.** Augmented dimension auxiliary function of crossplot of effectiveness analysis and productivity prediction. The X-axis represents porosity (POR) and the Y-axis represents true formation resistivity (RT).

### 4.4. Multi-Module Auxiliary Crossplots Augmented-Dimensional Analysis

Figure 16 shows a well example in Changqing Oilfield. Based on core analysis data, a two-dimensional crossplot of oil, gas, and water is established on the crossplot, and then the
target reservoir is selected to project onto the crossplot to quickly determine the distribution of oil and water layers in the target interval. It also supports the selection of feature points on the crossplot and the projection to the logging drawing to view the morphological characteristics of other logging data, so as to realize the use of other application modules on the platform to assist the augmented-dimensional display analysis of crossplot.

**Figure 16.** Multi-Module Auxiliary Crossplots Augmented-Dimensional Analysis (a) Application modules based on augmented dimension crossplot of depth alignment matrix. The X-axis represents compensated neutron log (CNL) and the Y-axis represents density (DEN). (b) Logging Drawing Display Feature Data Points Synchronously.
Figures 17 and 18 are the crossplots of multiple wells. The feature area is selected to display the target layer synchronously in the well profile module to assist the crossplot analysis.

**Figure 17.** Synchronous display of data in multi-well multi-section multi-dimensional crossplot based on deep alignment matrix. The X-axis represents the porosity (POR) and the Y-axis represents the true formation resistivity (RT).

**Figure 18.** Based on depth alignment matrix technology, select the feature area in the multiple crossplot to assist the crossplot analysis of connecting well profile.

5. Discussion

This section briefly introduces the limitations and advantages of this paper and gives the main research content and future development direction of logging crossplot.
Limitations: In order to meet the needs of field applications, more convenient full interactive functions need to be provided in the future, mainly including more free drag and zoom, automatic scale adjustment, and full interactive operation. In order to adapt to the coming of the intelligent era, it is very necessary to add artificial intelligence algorithms for data dimensionality reduction, which can better improve the operation, display speed, and efficiency. In order to further analyze and verify the data points in the crossplot by using the crossplot in paper materials, such as literature and periodicals, it can also provide rapid data point extraction, fitting, and scheme decision-making.

Advantages: This paper covers all display styles of logging crossplot, mainly including conventional crossplot, Z-value crossplot, frequency crossplot, histogram, multi-well and multi-layer section display crossplot, multi-well and multi-layer section display histogram, multi-curve correlation analysis crossplot, and error analysis diagram. Moreover, this paper realizes many advanced functions such as inputting data expression, multi conditional filtering data points, and multi-well section distribution statistics. What is particularly exciting is the opening of 217 rich secondary development interfaces.

Crossplot technology is an important data visualization analysis tool in logging processing and interpretation. On the basis of conventional crossplot, it further expands and develops, makes full use of heterogeneous data such as logging, geology, logging, and oil testing, expands the crossplot dimension increase to more dimensions, and provides a multi-well and multi-dimensional heterogeneous data crossplot dimension increase display analysis method, so as to provide more reservoir reference information for interpreters and realize multi angle logging comprehensive comparison, analysis, and display.

A visualization and analysis method by multi-dimensional crossplots from multi-well heterogeneous data is currently one of the most important research contents of crossplots. Several key technologies studied in this paper still need to be further studied in the future as follows:

1. In the future, artificial intelligence methods will be used to conduct dimensionality reduction-assisted logging cross-plot evaluation for logging high-end imaging data [43–46].
2. Make full use of offset well information and other results data and interpretation parameters to achieve quantitative interpretation of crossplots [47,48].

6. Conclusions

In this paper, a visualization and analysis method by multi-dimensional crossplots from multi-well heterogeneous data is proposed, and an augmented-dimensional depth match matrix is constructed. The technology can automatically align logging, geology, logging and oil testing, and other heterogeneous data according to depth, and load them to the crossplot display. The augmented dimension of the crossplot is extended to more dimensions. Based on this technology, the expansion support for various heterogeneous data can be quickly realized. The display method of the crossplot from multi-well data by quadtree index is constructed, which can effectively reduce the number of drawing data points of the crossplot and improve the display and interaction ability of the crossplot under the condition of large amount of multi-well data. The display and interactive performance were tested comprehensively by loading 160 wellhead data at the same time, and the display and interaction performance of the augmented-dimensional crossplots under the conditions of large amount of data from multiple wells were greatly improved. The crossplot analysis tool has been applied in logging evaluation of multiple well working areas in China, and the analysis function and interactive performance meet the requirements of field application.

Author Contributions: Conceptualization, M.C. and Z.G.; methodology, Y.Y. and M.C.; investigation, Z.Y. and Y.Z.; data curation, Y.Y.; writing—original draft preparation, Z.G. All authors have read and agreed to the published version of the manuscript.
Funding: This research was funded by HeiLongJiang Natural Science Foundation of China “Quantum Intelligent Optimization and Application in Real-time Adjustment of Logging Interpretation Model While Drilling” (grant No. LH2019004).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References
18. Radwan, A.E. Hydrocarbon type estimation using the synthetic logs: A case study in Baba Member, Gulf of Suez, Egypt. AAPG Datapages Search Discov. Artic. 2020, 20475. [CrossRef]


