

Визуализация многомерных поверхностей и их экстремумов *

А.А. Батюков¹, Е.В. Попов¹, Н. Фогт^{2,3}

¹ Нижегородский государственный архитектурно-строительный университет, 603950, Нижний Новгород, Россия

² Московский государственный университет им. М.В. Ломоносова, 119234, Москва, Россия

³ University of Ulm, Albert Einstein Allee 47, D-89069 Ulm, Germany

Аннотация. В статье описан метод многомерной визуализации гиперповерхности и ее экстремумов по модели Люмиграф. В статье показаны недостатки классического Люмиграфа и сформулированы причины этого. Чтобы преодолеть эти недостатки, авторы вводят концепцию дополнительного экрана. Описаны правила построения изображения на этом экране. Окончательное изображение делает видимой всю гиперповерхность и зоны ее экстремумов. Для возможности визуализации гиперповерхностей размерностью порядка более, чем 5D как развитие модели Люмиграф, разработана модель Exidiagraph и описан способ её построения. Введено понятие псевдо-воксела для отображения гиперповерхности, представленной многомерной таблицей. Разработан программный пакет для визуализации таблично заданных многомерных поверхностей.

Ключевые слова: многомерная визуализация, Люмиграф, Exidiagraph.

1 Введение

Основной особенностью информации является ее многомерность. В связи с этим информация часто теряет свою четкость и возможность представить данные в визуальной форме с помощью стандартных средств визуализации, таких как чертежи, графики, диаграммы и т.д. [1-3]. Эта проблема часто возникает в различных областях человеческой деятельности. Например, в работе [4] сделана попытка систематизировать и кратко описать некоторые типы проблем визуализации, возникающих при обработке результатов обобщенного математического эксперимента в вычислительной газовой динамике. Автор также организовал новые задачи визуального представления многомерных числовых выходных данных. Разработка методик анализа результатов обобщенного вычислительного эксперимента, представленного в виде многомерного объема данных, рассмотрена в [5].

* Работа выполнена в рамках гранта РФФИ № 19-07-01024 А.

Физико-химическое изучение и прогнозирование свойств химических соединений имеют первостепенное значение как с технологической, так и с академической точек зрения. Квантово-химические вычислительные методы становятся полезными инструментами для решения некоторых химических задач, в том числе для изучения механизма реакции перед сложными, дорогостоящими или иногда невозможными экспериментами. Исследование структуры молекул нередко сводится к нахождению множества экстремумов на поверхности потенциальной энергии (PES), то есть к изучению так называемого «конформационного ландшафта» [6], [7]. В общем случае PES имеет многомерный характер. Строго говоря, решение задачи визуализации многомерной поверхности с ее экстремумами весьма актуально для многих прикладных задач в разных отраслях науки и техники. В этой статье рассмотрены общие принципы организации визуализации для многомерной поверхности в зависимости от четырех до шести независимых переменных. Будем называть такую поверхность 5D-7D поверхностью (или гиперповерхностью).

Проблема многомерной визуализации гиперповерхностей очень актуальна и сложна. Существует множество подходов к визуализации и большое количество визуальных таксономий [8]. В качестве подходов к многомерной визуализации можно выделить следующие: 2D и 3D диаграммы рассеяния; Матрица диаграмм рассеяния; Тепловые карты; Карты высот; Линзовые таблицы; Обзорные участки; Иконографические представления; Общие логические диаграммы; Параллельные координаты; Линейный график, Многострочный график; Пиксельные техники; Круговые сегменты; Многомерное масштабирование и графики Саммона; Полярные карты; Анализ главных компонент и главной кривой; Погоня за проекцией; Самоорганизующиеся карты Кохонена и многие другие. Однако анализ в работе [6] показал, что ни один из них не является достаточно подходящим для одновременной визуализации гиперповерхности и ее экстремумов.

2 Визуализация 5D поверхности с помощью Люмиграфа

Существует возможность представить пятимерную гиперповерхность с помощью функции четырех независимых переменных с уровнем энергии, представленным цветом, и применить модель под названием Люмиграф. Обычное параметризованное двухплоскостное световое поле (2PP), расширенное геометрией объекта для процесса реконструкции, было описано Гёртлером [9], который ввел термин Люмиграф. Это обычно применяется, чтобы подчеркнуть тот факт, что прокси-геометрия требуется и используется для реконструкции светового поля. Это подмножество полной пленоптической функции, которая описывает поток света во всех положениях и во всех направлениях. Двухплоскостная установка — это глобальная параметризация, которая описывает луч в пространстве с точками пересечения в двух параллельных плоскостях. Каждая точка пересечения является 2D, и это приводит к четырем параметрам выборки. В нашем случае этот вид луча можно рассматривать как четырехмерную точку. Цвет

точки, рассматриваемый как дополнительный параметр, преобразует эту точку в 5D точку.

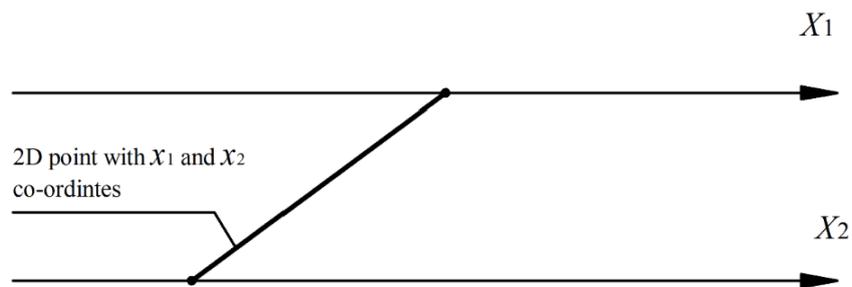


Рис. 1. Представление 2D точки с помощью Люмиграфа

Основная идея может быть объяснена на примере 2D аналога Люмиграфа (см. рис.1). Пусть есть две параллельные оси X_1 и X_2 , связанные с двумя независимыми переменными. Тогда прямая между двумя точками на этих осях адекватна 2D точке с координатами (x_1, x_2) в декартовом пространстве.

Если расширить эту идею до случая четырех независимых переменных, точка может быть представлена линией, показанной на Рис. 2, где две ее конечные точки имеют две пары переменных (D_3, D_1) и (D_6, D_5) , связанных с соответствующей плоскостью.

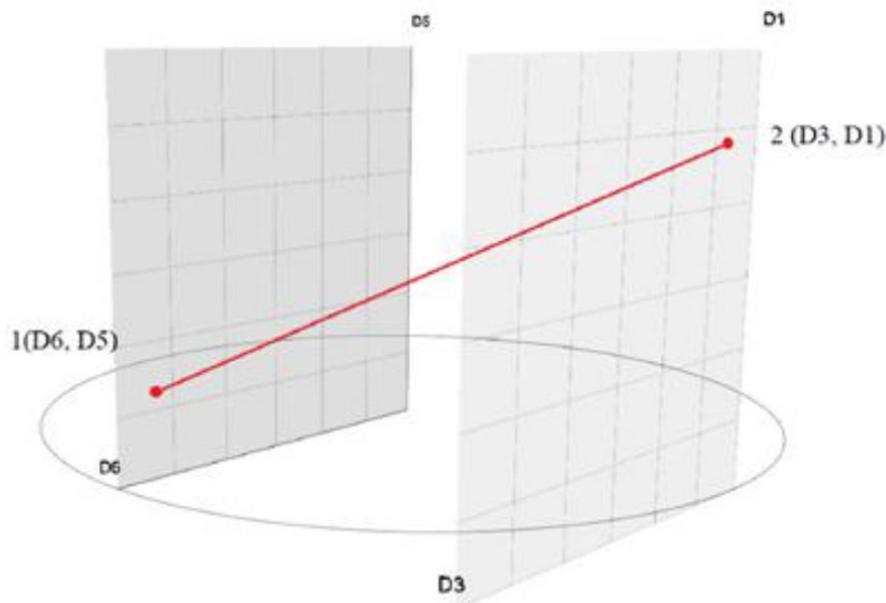


Рис. 2. Точка в зависимости от четырех переменных

В геометрии любая поверхность рассматривается как бесконечное множество точек. Однако в большинстве приложений функция многомерной поверхности представлена в виде конечных таблиц. Это означает, если такие таблицы являются четырехмерными, то их визуализация может быть осуществлена как конечный набор только четырехмерных точек. Люмиграф — это удобная и простая модель для визуализации функции четырех переменных, как показано на рис. 3, где значения функции отображаются в цветном виде. Обычно исходная таблица функций 4D имеет n строк и столбцов в каждом измерении, поэтому количество точек 4D в Люмиграфе равно n^4 . Хотя такое представление очень просто, наглядно и удобно, оно недостаточно ясно для восприятия. Области экстремумов функции при этом практически не видны.

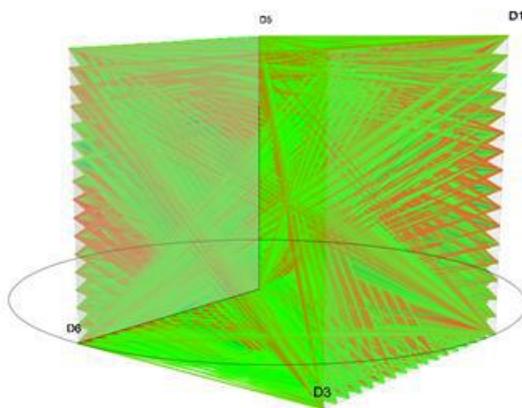


Рис. 3. Люмиграф табличной функции четырех переменных

С целью улучшения визуализации 5D поверхности использован прием отображения Люмиграфа на плоскость (см. работу [10]). Основное требование к такому отображению заключается в том, чтобы вся поверхность и ее экстремумы были видны пользователю. Основная идея здесь — разместить третью плоскость, параллельную двум начальным внутри Люмиграфа для нахождения точки пересечения с ней отрезков. Этот подход назван концепцией дополнительного экрана. Такой способ отображения позволяет осуществить полное 2D-отображение всей 5D-гиперповерхности в виде растрового изображения (см. рис. 4, где условно скрыты 4D-точки).

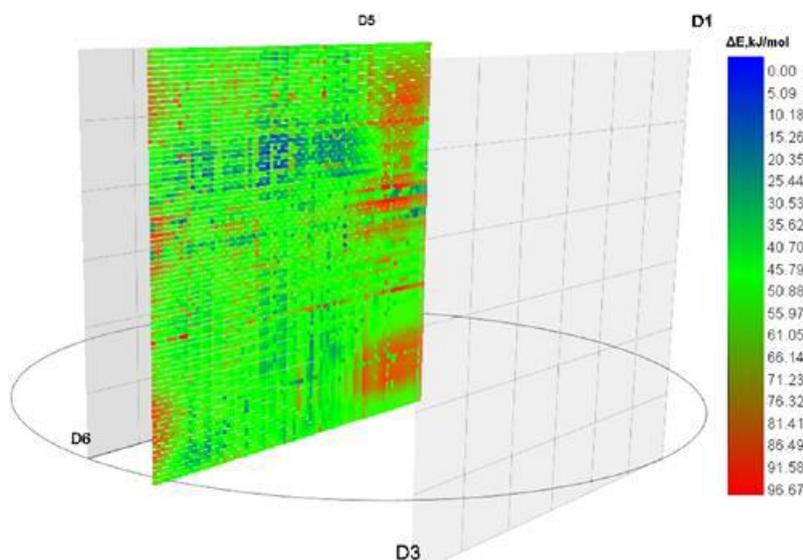


Рис. 4. Отображение табличной функции четырех переменных с помощью растрового изображения

3 Визуализация 6D и 7D поверхностей

Отображение точки в пространстве размерности 6 и выше осуществимо с использованием идеи, реализованной в Люмиграф модели. Как было сказано выше для визуализации точки в пространстве 4D были использованы две параллельные плоскости, в каждой из которых были заданы две двумерные системы координат. В развитие этой идеи была разработана новая модель визуализации точки в пространстве 6D. Реконструкция Люмиграф модели заключалась в добавлении к ней еще одной плоскости с двумерной системой координат, дополняющей модель до пространства размерности 6D. Таким образом была получена система координат, состоящая из трех координатных плоскостей, расположенных так, чтобы все плоскости образовали трехгранник вокруг оси Z (Рис.5). Данная модель получила название Exidiagraph (греч. *éxi diastáseon gráfima* - шестимерный график).

Для отображения в Exidiagraph точек гиперповерхности набор шести координат разбивается на пары. Для случая шести переменных можно получить три пары координат по одной паре для каждой из трех плоскостей. Последовательность выбора координат для разбиения по парам может быть любой. Аналогичный подход можно использовать для визуализации пространств большей размерности, например 8D, 10D и т.д., разбив координаты по парам и добавив координатную плоскость для каждой дополнительной пары координат. С каждой координатной плоскостью можно ассоциировать точку, соответствующую од-

ной пары координат (x, y) . Полученные на координатных плоскостях точки можно считать вершинами треугольника, изображенного на рис.5.

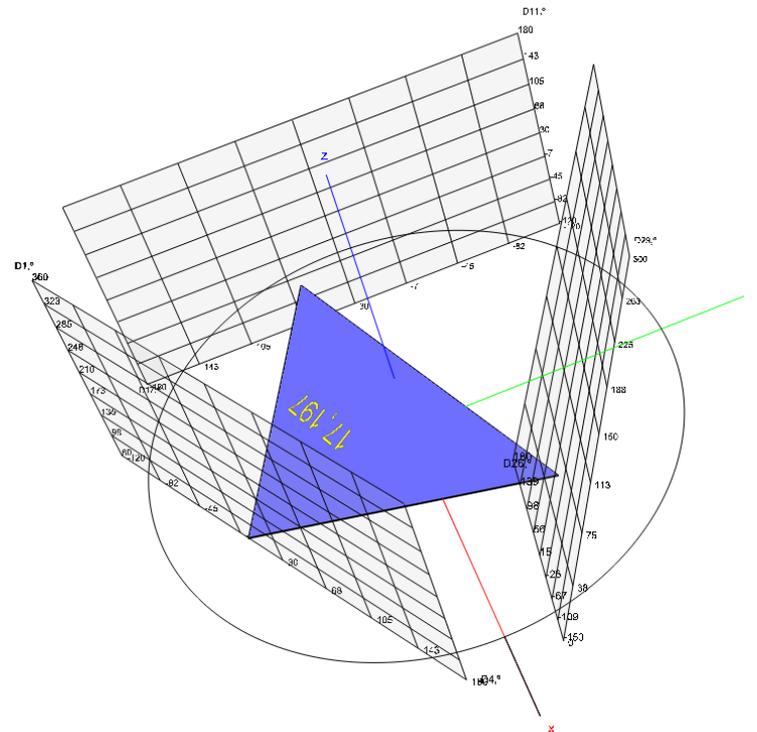


Рис. 5. Отображение точки в виде треугольника в 6D пространстве

Таким образом полученный треугольник однозначным образом соответствует точке, определяемой шестью независимыми переменными. Если данному треугольнику поставить в соответствие определенное значение энергии и отобразить его в виде цвета, ассоциированного с цветовой легендой, то можно получить отображение 7D точки, зависящей от шести независимых переменных в 6D пространстве. Однако, поскольку в геометрическом смысле поверхность любой размерности может рассматриваться как бесконечное множество точек, визуализация всей гиперповерхности с использованием описанной модели представляет собой значительные трудности.

Промежуточный путь преодоление данной трудности может быть найден следующим образом. Точка в 6D пространстве может быть ассоциирована с геометрическим центром треугольника, а соответствие этой точки конкретным координатам может быть визуализировано отрезками прямых линий, связывающих эту точку с точками на соответствующих координатных плоскостях. Эти отрезки назовем линиями связи.

С учетом положительного эффекта от использования вышеописанного подхода к визуализации многомерных точек была разработана псевдо-воксельная

модель визуализации гиперповерхности. В основе модели лежит понятие псевдо-воксел, которое в отличие от термина воксел формируется не в пространстве 3D, а в пространстве 6D. Визуализация строится при этом с учетом того, что гиперповерхность представлена в виде многомерной таблицы. С целью визуализации каждой ячейке многомерной таблицы ставится в соответствие многомерная точка. Затем каждая многомерная точка визуализируется по схеме, описанной в данном разделе выше. В результате визуализации в модели Exidiagraph можно получить картину, представленную на рис.6, где области экстремумов хорошо видны (синим цветом отображены зоны минимумов, красным – зоны максимумов).

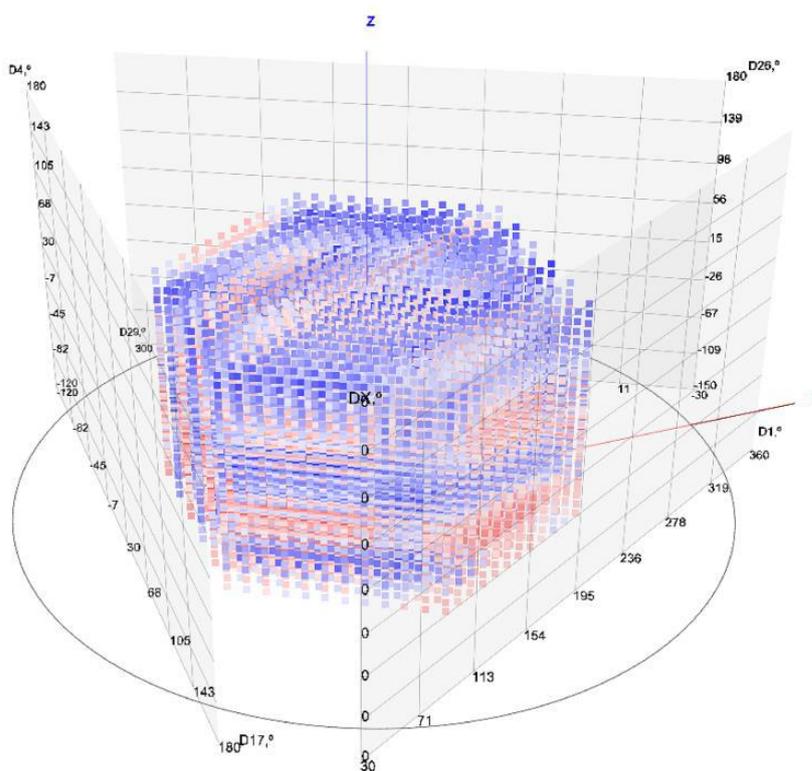


Рис. 6. Псевдо-воксельная модель 7D поверхности в Exidiagraph

Если теперь уровни энергии, соответствующие каждой точке гиперповерхности, нормировать по их истинному значению, то можно получить уточненную псевдо-воксельную визуализацию гиперповерхности, которая является наиболее адекватной визуализацией табличной многомерной гиперповерхности со всеми ее экстремумами (рис. 7).

десятилетия появилось много новых алгоритмов для решения задач глобальной оптимизации. Это привело к успеху в решении широкого круга задач в различных областях, таких как, например, вычислительная химия и биология, информатика, экономика, инженерное проектирование и другие. Но в целом нет эффективных методов, позволяющих минимизировать n -мерные функции. Алгоритмы исходят из исходного предположения с использованием алгоритма поиска, который пытается двигаться в наклонном направлении. Алгоритмы, использующие функцию градиента, минимизируют одномерную линию вдоль этого направления, пока не будет найдена самая низкая точка с подходящим допуском. Затем направление поиска обновляется в соответствии с локальной информацией о функции и ее производных, и весь процесс повторяется до тех пор, пока не будет найден истинный n -мерный минимум.

Алгоритмы, которые не требуют градиента функции, используют другой подход. Алгоритм Симплекса Нелдера-Мида поддерживает $n + 1$ векторов тестовых параметров в качестве вершин n -мерного симплекса. Алгоритм пытается исправить вершину симплекса, имеющую наихудшие параметры, на каждой итерации, используя геометрические преобразования. Итерации продолжаются до тех пор, пока общий размер симплекса не будет достаточно уменьшен. Проблема представляется очень сложной для негладких или дискретных функций. На самом деле, пакеты, например, вычислительной химии производят значения молекулярной энергии в виде большой многомерной таблицы в зависимости от 2, 3, 4 или более переменных, которые соответствуют ряду переменных молекулы. В этом случае табличная гиперповерхность не имеет аналитического выражения, что серьезно затрудняет поиск ее экстремумов. Кроме того, время для расчета энергии критически зависит от ряда переменных, а также от шага расчета для каждой данной переменной. С большим числом переменных и более сжатым шагом время расчета может занять до нескольких месяцев непрерывных вычислений на суперкомпьютерных платформах. Таким образом, это усложняет получение экстремума гиперповерхности с приемлемой надежностью.

Для решения этой проблемы был разработан подход для нахождения локальных минимумов гиперповерхности. Он состоит из двух этапов:

1. Чтобы улучшить поиск минимумов, необходимо заменить исходную многомерную функцию, определенную как регулярная сетка вычисленных точек, на более сжатую регулярную сетку. Значения функций в узлах сжатой сетки рассчитываются путем многомерной сплайн-интерполяции. Существует несколько подходов к многомерной интерполяции. В описанной версии гиперповерхностного процессора авторы применяют процедуру интерполяции многомерного сплайна, разработанную M. Flanagan [11], которая позволяет интерполировать значения функции, заданной таблицей любой размерности. Эта процедура содержит метод для выполнения многомерной кубической сплайн-интерполяции, то есть интерполяции в n -мерных массивах точек данных, $y = f(x_1, x_2, x_3 \dots x_n)$, используя естественные кубические сплайны и где n может принимать любые целочисленные значения. Данные представле-

ны в виде таблицы для массивов $x_1, x_2, x_3 \dots x_n$ и для $y = f(x_1, x_2, x_3 \dots x_n)$. Размерность массивов x_i и этапов и x разбиения могут быть разными. Процедура интерполяции является рекурсивной.

- Следующий этап – последовательное сравнение значений функций в каждом узле сетки со значениями функций в соседних узлах, окружающих его. Если это значение является наименьшим или наибольшим, то это значение можно считать локальным экстремумом. Количество соседних узлов зависит от размерности решаемой задачи. В трехмерном случае таких узлов 8, в четырехмерном случае 26 узлов, в пятимерном случае таких узлов 80 (рис.8). Чтобы сделать экстремумы более точными, сетка должна быть максимально сжатой. Этот подход позволяет находить экстремумы энергетической функции на достаточном уровне надежности, даже если расчет является относительно грубым. В отличие от классического решения по оптимизации, этот подход позволяет очень быстро находить как глобальный экстремум, так и локальные экстремумы. Тем не менее, следует отметить, что конденсация сетки фактически не повышает точность исходного решения. Это может привести к появлению некоторых фиктивных минимумов, которые на самом деле не существуют, в то же время некоторые реальные минимумы могут быть потеряны. Следовательно, весь минимум должен быть подтвержден соответствующим экспериментом.

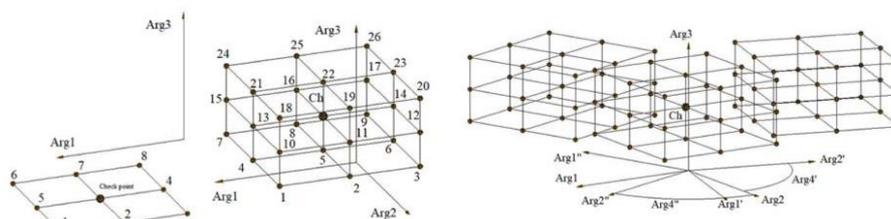


Рис. 8. Схема нахождения минимумов. а) 3D; б) 3D + переменная для цвета = 4D; в) 4D + переменная для цвета = 5D

5 Заключение

В статье описан метод визуализации многомерной гиперповерхности с ее экстремумами на основе модели Люмиграф. Классический Люмиграф недостаточно удобен для восприятия, поскольку он может быть представлен только набором прямых отрезков, как описано в этой статье. Зоны экстремумов функции практически не видны, даже если Люмиграф окрашен в разные цвета. Дополнительный экран, параллельный плоскостям Люмиграфа, содержит растровое изображение всех зон гиперповерхности. Кроме того, такое изображение окрашено в соответствии с его цветовой легендой. Этот механизм отображения многомерной поверхности назван концепцией дополнительного экрана. Благодаря такому механизму мы можем получить более визуально эффективное растровое

изображение функции вместе со всеми ее экстремумами. Этот подход применим для визуализации 5-мерной поверхности, зависящей от четырех переменных. Дальнейшая оценка этого подхода направлена на увеличение размерности отображаемой гиперповерхности. Как развитие идеи, заложенной в модели Люмиграф, была получена система координат, состоящая из трех координатных плоскостей, расположенных так, чтобы все плоскости образовали трехгранник вокруг оси Z. Данная модель получила название Exidiagraph. На базе этой модели разработан программный пакет Exidiagraph для визуализации многомерных поверхностей вместе с их экстремумами на основе языка Java [12], библиотеки Java 3D и библиотеки многомерной сплайн интерполяции Flanagan M. [11]. Программный пакет состоит из трех основных модулей: модуля считывания численных таблиц с представлением их в виде многомерной матрицы, модуля интерполяции матричных значений, модуля отображения многомерной поверхности. Программное обеспечение позволяет визуализировать многомерные поверхности размером до 7D.

Литература

1. Wehrend S. and Lewis C. "A Problem-Oriented Classification of Visualization Techniques", *Proceedings of the 1st IEEE Conference on Visualization '90*, 1990, pp.139-143.
2. Etemadpour R. et al. Choosing Visualization Techniques for Multidimensional Data Projection Tasks: A Guideline with Examples. (2016, 166-186 pp 598. 166-186. 10.1007/978-3-319-29971-6_9.
3. Keim D. A. and Kriegel H.-P. "Visualization Techniques for Mining Large Databases: A Comparison", *IEEE Transactions on Knowledge and Data Engineering*, vol.8, no.6, 1996 pp.923-938.
4. Bondarev A.E. On visualization problems in a generalized computational experiment / *Scientific Visualization*, 2019, volume 11, number 2, pages 156-162, <http://doi.org/10.26583/sv.11.2.12>.
5. Bondarev A.E., Galaktionov V.A., Shapiro L.Z. Processing and visual analysis of multidimensional data / *Scientific Visualization*. V.9, № 5, с.86-104, 2017, DOI: <http://doi.org/10.26583/sv.9.5.08>
6. E. Popov, A. Batiukov, N. Vogt, T. Popova, J. Vogt. Visualization and Analysis of Molecular Potential Energy Surface (Pes) and Its Minima. *IADIS International Conference Interfaces and Human Computer Interaction 2019* (part of MCCSIS 2019), Porto, 2019. pp. 411-415.
7. Popov, E. V., Batiukov, A. A., Vogt, N., Popova, T. P., & Vogt, J. (2020). Visualization and Minima Finding of Multidimensional Hypersurface. In P. Isaias, & K. Blashki (Eds.), *Interactivity and the Future of the Human-Computer Interface* (pp. 282-309). Hershey, PA: IGI Global. doi:10.4018/978-1-7998-2637-8.ch015
8. B. Shneiderman, "The Eyes Have It: A Task by Data Type Taxonomy of Information Visualization," presented at IEEE Symposium on Visual Languages '96, Boulder, CO, 1996.
9. Götler S. et al, The Lumigraph. In *Computer Graphics, Annual Conference Series* (Proc. SIGGRAPH '96): 1996. pp. 43-54.
10. Eugene Popov, Tatyana Popova, Anatoly Batiukov, Natalja Vogt. Visualization of Multidimensional Hypersurface Extrema by Lumigraph (2020). *Scientific Visualization* 12.2: 1 - 8, DOI: 10.26583/sv.12.2.01.

11. Flanagan, M. T. (2014). *JavaScientific Library*
<http://www.ee.ucl.ac.uk/~mflanaga/java/index.html>.
12. Popov E.V., Batiukov A.A. Exi Diastaseon Grafima (ExiDiaGraph). Certificate of a computer program RU 2020612284, 19/02/2020.Application No. 2020611368/69 of 10/02/2020.