

# Визуальная аналитика многомерных динамических данных

© Д.Д. Попов<sup>1</sup> © И.Е. Мильман<sup>1</sup> © В.В. Пилюгин<sup>1</sup> © А.А. Пасько<sup>2</sup>

<sup>1</sup>Национальный исследовательский ядерный университет «МИФИ»,

Москва, Россия

<sup>2</sup>Британский национальный центр компьютерной анимации при университете Борнмута,

Борнмут, Великобритания

drovovmephi@gmail.com igalush@gmail.com VVPilyugin@mephi.ru

apasko@bournemouth.ac.uk

## Аннотация

В статье рассматривается задача анализа изменения некоторых объектов. Каждый объект характеризуется набором из  $n$  численных параметров. Инструментом проведения анализа является метод визуализации. Получение интересных аналитика суждений о рассматриваемых объектах осуществляется за несколько шагов. На первом шаге строится геометрическая интерпретация исходных данных. Затем построенная математическая модель подвергается ряду преобразований. Эти преобразования соответствуют решению первой задачи метода визуализации – непосредственно визуализации исходных данных. Далее человеку предлагается провести анализ полученных изображений и интерпретировать результаты по отношению к рассматриваемым объектам. Иными словами, решение задачи анализа осуществляется с помощью пространственного моделирования исходных данных и последующего исследования аналитиком построенных пространственных объектов.

Предложен алгоритм решения задачи. Описана разработанная интерактивная прикладная программа визуализации, реализующая этот алгоритм. Продемонстрировано, как в процессе работы с программой пользователь может построить суждения об образовании и разрушении кластеров и сгустков из объектов, а также найти инварианты в изменении исходных данных.

## 1 Введение

Среди различных методов анализа данных в настоящее время важное место занимают

перспективные методы визуальной аналитики. Под визуальной аналитикой понимается решение задач анализа данных с использованием способствующего интерактивного визуального интерфейса. Этот термин ввёл Джеймс Томас (James Thomas) в [3], чем обратил внимание научной общественности на роль визуального представления данных в решении исследовательских задач.

Одной из форм визуальной аналитики является решение задач анализа данных методом визуализации. Теоретические обобщения этого метода на примере научных данных рассмотрены в работе [7]. Этот метод заключается в последовательном, в общем случае многократном, решении двух задач:

1. задачи визуализации исходных данных;
2. задачи анализа полученных графических изображений с последующей формулировкой суждений относительно исходных данных.

В данной работе рассмотрен ряд вопросов теории и практики дальнейшего развития этого метода применительно к анализу многомерных динамических данных. Здесь многомерными динамическими данными мы называем числовые данные, представляемые в табличном виде. Они описывают некоторые изменения состояния рассматриваемой совокупности объектов, происходящие с течением времени.

## 2 Подходы к визуализации многомерных данных

За последние 30 лет появилось большое количество научных статей и книг, посвящённых подходам к визуализации. Преимущественно данную тему освещают зарубежные авторы. Найдены публикации европейских и американских исследователей, которые датируются 70-ми, 80-ми годами XX века. Отечественные учёные активно подключились к изучению данного вопроса

---

Труды XVIII Международной конференции DAMDID/RCDL'2016 «Аналитика и управление данными в областях с интенсивным использованием данных», Ершово, 11-14 октября 2016

значительно позже (основной информационный массив российского авторства относится к 2000-2005 гг.).

Среди работ об анализе многомерных динамических данных можно обнаружить большое количество статей посвящённых динамике наноструктур, как, например, [1] и [2]. В [5] проводится визуальный анализ свойств пространственно-временных структур. Данные об объектах, указанных в этих статьях, имеют трёхмерную геометрическую интерпретацию. Также в [5] констатируется существование проблемы визуализации данных, имеющих размерность более 4<sup>1</sup>, адекватной восприятию пользователя. Нас будут интересовать пространства размерности больше 3.

В [8] предложен способ визуализации многомерных данных о состоянии пациента. Производится анализ *одного* объекта, характеризующегося множеством (мощность этого множества > 3) числовых параметров, значения которых изменяются с течением времени. Метод, описываемый в статье, направлен на решение задачи анализа изменений *множества* многомерных объектов. При чём наибольший практический интерес представляют множества большой размерности.

В работе [4] специфика визуализируемых многомерных динамических данных заключается в следующем: 2 или 3 числовые компоненты этих данных имеют пространственную природу. В русскоязычных трудах также рассматриваются способы визуального анализа данных, имеющих такую особенность. Рассматриваемые в них программные комплексы называются геоинформационными системами (ГИС). В ГИС каждому объекту ставится в соответствие некоторый геометрический образ, изображающийся на плоской или объёмной карте местности. В качестве таких геометрических образов могут выступать круги, квадраты, шары, кубы, линии, показывающие траектории движения исследуемых объектов. Числовые компоненты данных, не пространственной природы представляются в отдельных окнах визуального пользовательского интерфейса либо поочерёдно, в соответствии с запросом пользователя, отражаются на карте местности за счёт изменения оптических параметров геометрических объектов (например, цвет, прозрачность) или их размеров, формы. В данной статье числовые характеристики объектов имеют произвольную природу.

Указанные выше программы и методы анализа многомерных динамических данных, предполагающие их использование, различаются не только моделями исходных данных и подходами к их визуализации. Упомянутые различия обусловлены поставленными задачами анализа этих данных. Нас будет интересовать исследование данных с целью

выявления схожих объектов среди рассматриваемой совокупности. Изучается, как изменяются множества подобных объектов, а также, что остаётся неизменным с течением времени.

### 3 Постановка задачи

Пусть имеется множество из  $m$  объектов, каждый из которых характеризуется  $n$  параметрами.

Мы предполагаем, что в рамках рассматриваемой задачи исходно задано  $k$  таблиц вида:

**Таблица 1**  $j$ -ая таблица с данными об объектах

	Параметр 1	Параметр 2	...	Параметр $n$
Объект 1	$x_{11}^j$	$x_{12}^j$	...	$x_{1n}^j$
Объект 2	$x_{21}^j$	$x_{22}^j$	...	$x_{2n}^j$
...	...	...	...	...
Объект $m$	$x_{m1}^j$	$x_{m2}^j$	...	$x_{mn}^j$

Каждая таблица содержит значения параметров объектов в некоторый момент времени. То есть  $j$ -ая таблица содержит данные для момента времени  $t_j$ .

При этом будем считать, что  $t_1 < t_2 < \dots < t_k$ .  $x_{il}^j$  – значение  $l$ -ого параметра  $i$ -ого объекта в соответствующий  $j$ -ой таблице момент времени  $t_j$  ( $l = \overline{1, n}, i = \overline{1, m}, j = \overline{1, k}$ ).

#### 3.1 Геометризация данных

Для решения задачи анализа перейдём к её геометрической интерпретации. Будем считать, что исходные данные являются точками многомерного евклидова пространства  $E_n$  с заданной метрикой  $\rho(x, y)$ . При этом:

- Строка таблицы интерпретируется как точка многомерного евклидова пространства.
- Столбцы таблиц интерпретируются как координаты точек многомерного евклидова пространства.
- Расстояние в евклидовом пространстве трактуется как мера различия объектов.

Используя введённое в пространстве расстояние, можем выделить подмножества точек: сгустки и кластеры. Определения этих подмножеств даны в [6]. Координаты точек известны в моменты времени  $t_1, t_2, \dots, t_k$ .

Введём понятие *геометрического процесса*. *Геометрический процесс* – это множество точек пространства, координаты которых зависят от времени. Таким образом, изначально имеем дискретный геометрический процесс  $P = P(t_j)$ .  $P(t_j)$  задан на множестве  $T = \{t_1, t_2, \dots, t_k\}$ .

$$P(t_j) = \{p_1(t_j), p_2(t_j), \dots, p_m(t_j)\},$$

где  $p_i(t_j)$  -  $n$ -мерная точка.

$$p_i(t_j) = (p_i^1(t_j), p_i^2(t_j), \dots, p_i^n(t_j)),$$

задают координаты точек привычного для человека трёхмерного пространства, и точки с одинаковым значением 4-го параметра выделяют одним цветом.

<sup>1</sup> В случае данных, имеющих размерность 4 в качестве графической интерпретации выступают изоповерхности, то есть 3 числовые характеристики



**Рисунок 1** Метод визуализации

где  $p_i^l(t_j) = x_{il}^j$  –  $l$ -ая координата  $i$ -ой точки.

Для каждой пары точек определено расстояние

$$\rho(p_{i_1}(t_j), p_{i_2}(t_j)) \geq 0, \quad i_1, i_2 = \overline{1, m}, i_1 \neq i_2$$

$$\rho(p_i, p_j) \geq 0$$

Стоит отметить, что расстояние, то есть метрику евклидова пространства, целесообразно выбрать исходя из известных особенностей исходных данных. Существует множество исследовательских работ, описывающих ситуацию, когда в процессе изучения некоторого набора данных эмпирически подбиралась метрика, которая наилучшим образом справляется с поставленной задачей анализа, например с задачей кластеризации. В разработанной программе, подробно о которой будет сказано позже, имеется возможность использовать евклидову метрику, расстояние Минковского, расстояние Чебышева, в частности, и расстояние Махаланобиса.

#### 4 Преобразования геометрических процессов

Для решения поставленной задачи возникает необходимость в преобразованиях исходного дискретного геометрического процесса.

$P(t_j)$  представляет собой упорядоченный по времени набор описаний известных состояний множества исследуемых объектов. Сами же объекты существуют и изменяются непрерывно. Исходя из этого факта, необходимо иметь инструмент преобразования дискретного процесса в непрерывный. Таким инструментом служит интерполяция.

##### 4.1 Интерполяция

Задача интерполяции состоит в поиске такой функции  $F$  из заданного класса функций, что:  $F(t_j) = P(t_j)$ , где  $t_j \in T$ .

Так как

$$P(t_j) = \{p_1(t_j), p_2(t_j), \dots, p_m(t_j)\},$$

$$p_i(t_j) = (p_i^1(t_j), p_i^2(t_j), \dots, p_i^n(t_j)),$$

то поиск  $F$  заключается в поиске таких  $f_i^l$ , принадлежащих заданному классу, что

$$f_i^l(t_j) = p_i^l(t_j).$$

В случае требования непрерывности  $f_i^l$  в результате интерполяции исходного дискретного геометрического процесса получаем непрерывный геометрический процесс  $P(t)$ .

Пусть  $f_i^l$  является алгебраическим двучленом, тогда при любом  $t \in [t_h, t_{h+1}]$ ,  $1 \leq h \leq k-1$ ,  $f_i^l$  будет рассчитываться по формуле:

$$f_i^l(t) = p_i^l(t_h) + \frac{p_i^l(t_{h+1}) - p_i^l(t_h)}{t_{h+1} - t_h} (t - t_h).$$

Интерполяция алгебраическим двучленом или кусочно-линейная интерполяция является самым простым видом интерполяции для которого выполняется свойство непрерывности.

За неимением априорной информации об изменении исходных данных с течением времени выбор способа интерполяции делается в сторону кусочно-линейной, так как она не требует сложных вычислений и при достаточно малых интервалах времени между имеющимися данными хорошо описывает наблюдающуюся зависимость.

##### 4.2 Дискретизация непрерывных геометрических процессов

Пусть  $P(t)$  – непрерывный геометрический процесс, тогда  $P(t_0)$  – его временное сечение.  $t_0$  принадлежит области определения  $P(t)$ .

Из  $P(t)$  можно построить дискретный геометрический процесс, заданный в  $k'$  моментах времени. Для этого выберем моменты времени  $\phi_1, \phi_2, \dots, \phi_{k'}$  и определим  $P'(\phi_h)$  как совокупность временных сечений  $P(t)$ :  $P'(\phi_h) := \{P(\phi_h) | \phi_h \in \{\phi_1, \phi_2, \dots, \phi_{k'}\}\}$ .

#### 5 Метод визуализации в рамках решаемой задачи

Метод визуализации подразумевает последовательное решение двух задач, представленных на рисунке 1.

Для описания первой задачи вводится важнейшее понятие метода – *пространственная сцена*. Пространственная сцена – это совокупность пространственных объектов с заданными геометрическими и оптическими параметрами. Описание геометрических параметров объектов сцены, например, их форма, размеры,

пространственное расположение называется *геометрической моделью сцены*. *Оптической моделью сцены* называется описание оптических параметров, например, цвет, прозрачность объектов.

Решение первой задачи подразумевает построение отображения исходных данных на пространственную сцену. Успех проведения анализа напрямую зависит от того насколько удачным окажется построенное отображение. На этом этапе необходимо выбрать такие модели, которые бы способствовали выявлению закономерностей в данных о рассматриваемых объектах. С одной стороны они должны соответствовать целям визуального анализа, например, анализа формы или взаимного расположения, а с другой обеспечивать возможность последующей интерпретации результатов визуального анализа по отношению к рассматриваемым объектам.

В рамках рассматриваемой задачи задание параметров сцены, атрибутов визуализации и получение изображений или анимационных фильмов происходит до тех пор, пока аналитик не рассмотрит достаточное их количество для формирования некоторого суждения об изменении взаимного расположения заданного множества точек с течением времени. Атрибуты

Рассмотрим подробнее процесс визуализации исходных данных.

### 5.1 Задание динамических исходных данных

Задаётся дискретный геометрический процесс  $P(t_j)$ , для которого будет проводиться визуализация.

### 5.2 Фильтрация

На этом шаге исходные данные проходят предварительную обработку.

Для решения поставленной задачи исходный дискретный процесс  $P(t_j)$  линейно интерполируется, как описано в 4.1. В результате фильтрации получается непрерывный процесс  $P(t)$ .

### 5.3 Мэппинг

На этом этапе исходным данным ставится в соответствие динамическая пространственная сцена  $S(t) = \langle G(t), O(t) \rangle$ , где  $G(t)$  – описание геометрии сцены, а  $O(t)$  – описание оптических параметров сцены.

В каждый момент времени  $t$ ,  $S(t)$  соответствует  $P(t)$ ,  $t \in [t_1, t_k]$ .

$$G(t) = \{Sph_1(t), \dots, Sph_m(t), Cyl_1(t), \dots, Cyl_e(t)\},$$

где  $Sph_i(t)$  – сфера, соответствующая точке  $p_i(t)$ ;  $Cyl_j(t)$  – цилиндр, соединяющий 2 сферы. Цилиндры соединяют только такие сферы  $Sph_{i_1}(t), Sph_{i_2}(t)$ , для соответствующих точек которых выполняется  $c(p_{i_1}(t_j), p_{i_2}(t_j)) \leq d$ ,  $d$  – некоторая константа, выбор которой в процессе решения задачи осуществляется аналитиком.

$$O(t) = \{SphC_1(t), \dots, SphC_m(t), CylC_1(t), \dots, CylC_e(t)\},$$

где  $SphC_1(t) = \dots = SphC_m(t) = SphC$  – цвета сфер,  $SphC$  так же в процессе решения задачи задаётся аналитиком,  $CylC_1(t), \dots, CylC_e(t)$  – цвета цилиндров.

Значение цвета в кодировке RGB рассчитывается по формулам:

$$R(t) = 255 \left( 1 - \frac{c(p_{i_1}(t), p_{i_2}(t))}{d} \right),$$

$$G(t) = 150 \frac{c(p_{i_1}(t), p_{i_2}(t))}{d},$$

$$B(t) = 255 \frac{c(p_{i_1}(t), p_{i_2}(t))}{d}.$$

## 5.4 Рендеринг

Понятие пришло из дисциплины машинная графика. Результатом рендеринга является проекционное изображение  $I$  сцены  $S$ .

Каждому  $t$  соответствует пространственная сцена, а значит и проекционное изображение:

$$I(t) = I(S(t), A(t)),$$

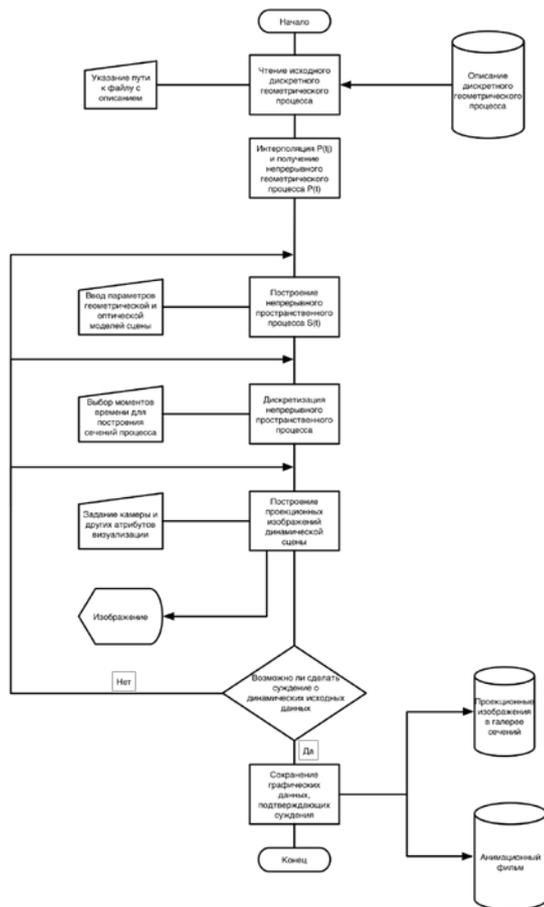
$A(t)$  – атрибуты визуализации: камера, освещение, размер получаемого изображения и другие.

## 6 Описание алгоритма решения задачи

Алгоритм решения задачи состоит из следующих этапов:

1. ввод динамических исходных данных;
2. их интерполяция и получение непрерывного геометрического процесса;
3. задание  $d$  и статических параметров сцены: радиус сфер, радиус цилиндров, подпространство для проекции, цвет сфер;
4. построение непрерывного пространственного процесса, соответствующего непрерывному геометрическому процессу, с использованием заданных статических параметров;
5. построение временных сечений непрерывного пространственного процесса, то есть его дискретизация;
6. задание камеры и других атрибутов визуализации для рендеринга;
7. построение проекционных изображений дискретного пространственного процесса;
8. дальнейшее сохранение проекционных изображений в галерею сечений или создание анимационного фильма;

В случае, если аналитик не может сделать интересующего его суждения относительно исходных данных по полученным проекционным изображениям, алгоритм предусматривает возвраты на этапы 3 и 6. Вышеописанный алгоритм представлен на рисунке 2.



**Рисунок 2** Алгоритм решения задачи

Следует отметить, что важным для практики случаев является анализ статических многомерных данных. Нетрудно видеть, что анализ статических многомерных данных можно рассматривать как частный случай анализа динамических многомерных данных. Он соответствует исходному дискретному геометрическому процессу, заданному в некоторый (единственный) момент времени, то есть  $T = \{t_1\}$ .

Алгоритм предусматривает взаимодействие аналитика и компьютера. Предлагается следующее распределение функций между ними:

- компьютер в процессе решения задачи анализа осуществляет функции расчёта расстояний, построения проекций и графиков зависимости координат от времени;
- аналитик осуществляет зрительное восприятие проекционного изображения на мониторе, анализирует взаимное расположение многомерных точек, выделяет подмножества и задаёт параметры визуализации.

Для реализации данного алгоритма была разработана соответствующая интерактивная программа. Она реализована на базе программного продукта Autodesk 3ds Max, с использованием внутреннего интерпретируемого языка MAXScript. Так же использовалась библиотека, написанная на языке C# в Visual studio 2013.

Эта программа позволяет аналитику совершать следующие действия:

- Ввод исходных данных (описание дискретного геометрического процесса).  
Осуществляется один раз при начале работы с программой.

- Кусочно-линейная интерполяция дискретного геометрического процесса.

- Построение динамической сцены, соответствующей непрерывному геометрическому процессу, и её анимационная визуализация.

Графическое проецирование выполняется в окне 3ds Max с помощью стандартного рендерера.

- Получение информации о точках.

Информацию о точках можно получать в виде таблицы, строки которой соответствуют каждой из отображаемых точек, а столбцы — их координатам. При этом, строки таблицы закрашены тем цветом, которым закрашены точки, т.е. в зависимости от цвета можно определять к какому подмножеству относится данная точка. Эти таблицы можно получать для каждого рассматриваемого момента времени.

- Измерение максимального внутрикластерного расстояния между двумя точками (в исходном n-мерном пространстве).

- Задание параметров сцены.

Параметрами сцены являются: радиус сфер, поставленных в соответствие исходным точкам, их цвет, радиус цилиндров и трёхмерное проекционное пространство. Указанные параметры задаются в начале работы программы, их можно изменять в процессе анализа.

- Использование стандартных средств 3ds Max для работы с пространственной сценой. Так, например, есть возможность использовать следующие средства:

- а) аффинные преобразования сцены;
- б) наложение разнообразных фильтров на изображение;
- в) преобразование оптических характеристик сцены.

- Объединение точек в подмножество.

Каждому подмножеству даётся цвет, который в дальнейшем будет обозначать данное подмножество во время работы программы.

- Построение и сохранение полученных графических проекций многомерных точек.

При анализе удалённых точек важным является то, какие именно координаты вносят больший вклад в расстояние — происходит ли это за счёт всех координат или за счёт

большого отличия только нескольких координат. Для определения этого, предлагается строить графические проекции исходного множества  $(p_i^l, p_i^l)$  и, меняя  $l$ , просматривать все такие проекции.

- Создание анимационного фильма.
- Построение графиков зависимости координат точек от времени.

При проецировании многомерных точек на трёхмерное пространство, особенно в случае большой размерности, проекции точек могут оказаться очень близко друг к другу. Для того, чтобы различить близкие сферы на получаемых изображениях аналитику предлагается воспользоваться такими возможностями изменения пространственной сцены как задание радиусов сфер и цилиндров. Уменьшив их, пользователь программы может более детально изучить точки, проекции которых оказались на малых расстояниях. Кроме того, в случае, если две различные многомерные точки проецируются в одну точку трёхмерного пространства, программа в процессе построения пространственной сцены немного сместит центры соответствующих сфер друг относительно друга и раскрасит их в цвет, отличный от цвета выбранного для сфер сцены по умолчанию.

## 7 Пример использования разработанной программы

С помощью созданной программы был произведён анализ 81 кредитной организации. Данные представляли собой ежемесячные отчёты этих организаций по 9 показателям за 13 месяцев 2013-2014 гг. Данные за первый отчётный месяц соответствуют  $t = 1$ , за последний –  $t = 13$ . Исследовалась принадлежность кредитных организаций в их геометрической интерпретации к сгусткам и кластерам, что позволяет судить об их схожести.

На рисунке 3 приведён ряд изображений, позволяющих сделать вывод, что за промежуток времени  $t \in [7,8]$  удалённая точка, соответствующая выделенной белой окружностью сфере, присоединилась к сгустку.

Как отмечается в [6], рассматриваемая в примере задача, актуальна при проведении аудита кредитных организаций. Априорно аналитик знает «хорошие» организации или же выделяет их в процессе анализа. Например, в первом приближении за множество «хороших» можно принять те, соответствующие точки которых принадлежат одному кластеру. Далее исследуются изменения, происходящие с этим множеством. Например, какая-то точка покинула или присоединилась к выделенному кластеру. Организация, соответствующая этой точке, попадает под особое внимание аналитика, сделанное им

наблюдение может послужить причиной более детального изучения деятельности этой организации. Также для исследователя важны изменения, происходящие внутри одного сгустка или кластера. Если в нём присутствуют точки, для которых с течением времени близость к другим точкам кластера резко изменяется, т.е. они быстро стремятся покинуть кластер или наоборот приближаются к другим точкам, то соответствующие организации также попадают под пристальный контроль. В разработанной программе аналитик фиксирует эти точки, наблюдая за изменением цвета цилиндров, соединяющих сферы.

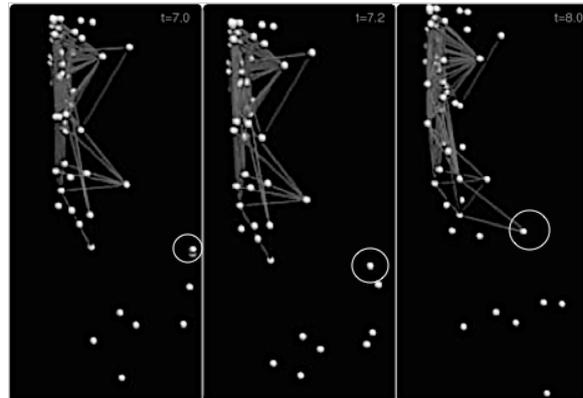


Рисунок 3 Кадры полученного анимационного фильма

Было замечено, что для одной из построенных моделей сцен (рисунок 4) наблюдается следующая закономерность. Большинство сфер, за исключением выделенных белыми окружностями на рисунке, лежат в одной плоскости.

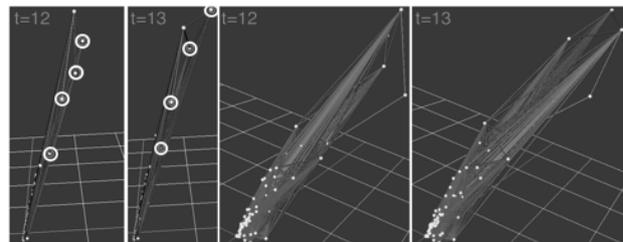


Рисунок 4 Проекционные изображения пространственной сцены

Это позволило выделить зависимость между тремя финансовыми показателями организаций:

$X$  – прощёная задолженность в кредитном портфеле,

$Y$  – выпущенные облигации,

$Z$  – активы нетто.

Она описывается уравнением плоскости:

$$2,231X - 72,672Y + 20,3624Z - 2,58513 = 0$$

Выделенные белыми окружностями сферы и соответствующие им кредитные организации этой зависимости не подчиняются. Зависимость имеет место при  $t \in [1,9] \cup [12,13]$ .

## Заключение

Итак, в рамках данной работы:

- Были созданы математические модели исходных данных, операций над ними.
- Разработан алгоритм решения задачи, на его основе написана программа.
- На примере экономических показателей кредитных организаций продемонстрирован анализ изменения подобия, а так же возможность поиска инвариантов геометрических процессов с помощью разработанной программы.

Инвариантами здесь называются такие свойства исследуемых данных, которые остаются неизменными либо на всём промежутке времени, либо на некоторой его части. В приведённом примере неизменной на выделенных отрезках времени остаётся зависимость трёх координат – показателей финансовых организаций.

В качестве дальнейшего развития функциональных возможностей рассматриваются различные способы интерполяции дискретных геометрических процессов, разработка пользовательского инструментария, позволяющего производить дополнительные построения в пространственной сцене. Например, предоставление возможности построения плоскостей, сфер и других геометрических примитивов.

В настоящей работе не представлены классические методы анализа многомерных данных, предполагающие численное моделирование и автоматический интеллектуальный анализ без участия человека. Тем не менее в настоящее время авторами изучаются эти подходы с целью последующего развития системы до уровня «численно-визуальной» системы анализа данных.

## Литература

- [1] Grottel S. Visual Verification and Analysis of Cluster Detection for Molecular Dynamics / S. Grottel [et al.]. // IEEE Transactions on Visualization and Computer Graphics. – 2007. – Vol. 13. – №6. – P. 1624-1631.
- [2] Sourina O., Korolev N. Visual Mining and Spatio-Temporal Querying in Molecular Dynamics / O. Sourina, N. Korolev // Journal of All rights reserved Computational and Theoretical Nanoscience. – 2005. – Vol. 2. – P. 1-7.
- [3] Thomas J., Cook K. Illuminating the Path: Research and Development Agenda for Visual Analytics. – IEEE-Press, 2005. – 185 p.

- [4] Wallner G. PLATO: A visual analytics system for gameplay data // Computers&Graphics. – 2014. – №38. – P. 341-356.
- [5] Бондарев А.Е. Оптимизационный анализ нестационарных пространственно-временных структур с применением методов визуализации / А.Е. Бондарев // Научная визуализация. – 2011. – том 3. – №2. – С. 1-11.
- [6] Мильман И. Е. Анализ данных о деятельности кредитных организаций с использованием программы интерактивного визуального анализа многомерных данных / И. Е.Мильман [и др.] // Научная визуализация. – 2015. – том 7. – №1. – С. 45-64.
- [7] Пилюгин В. В. Научная визуализация как метод анализа научных данных / В. В. Пилюгин [и др.] // Научная визуализация. – 2012. – том 4. – №4. – С. 56-70.
- [8] Хачумов В. М. Разработка новых методов непрерывной идентификации и прогнозирования состояния динамических объектов на основе интеллектуального анализа данных / В. М. Хачумов, А. Н. Виноградов // Математические методы распознавания образов: 13-я Всероссийская конференция. Ленинградская обл., г. Зеленогорск, 30 сентября - 6 октября 2007г. Сборник докладов. – 2007. – С. 548-550

### Visual analytics of multidimensional dynamic data

Dmitry D. Popov, Igal E. Milman, Victor V. Pilyugin,  
Alexander A. Pasko

This work deals with a problem of analysis of a time variant object. Each object is characterized by a set of numerical parameters. The visualization method is used to conduct the analysis. Insights of interest for the analyst about the considered objects are obtained in several steps. At the first step, a geometric interpretation of the initial data is introduced. Then, the introduced mathematical model undergoes several transformations. These transformations correspond to solving the first problem of the visualization method, in particular obtaining visual representations of data. The next step is for the analyst to analyse the generated visual images and to interpret the results in terms of the considered objects.

We propose an algorithm for the problem solving. Developed interactive visualization software is described, which implements the proposed algorithm. We demonstrate how with this software the user can obtain insights regarding the creation and disappearance of object clusters and bunches and find invariants in the initial data changes.