

# Параллельная фильтрация в системе визуализации параллельных вычислений

Д.Ю. Горбашевский, А.Ю. Казанцев, Д.В. Манаков

ИММ УрО РАН, Екатеринбург, Россия

## Аннотация

В статье рассмотрены теоретические проблемы параллельной фильтрации в визуальных супервычислениях и примеры ее использования в специализированной системе визуализации параллельных вычислений.

**Ключевые слова:** параллельная фильтрация, визуальные супервычисления.

## 1. ВВЕДЕНИЕ

В ИММ УрО РАН в течение ряда лет ведется разработка системы интерактивной визуализации параллельных вычислений [1-2]. Одной из задач является задача вывода сеток большого и сверхбольшого (десятки и сотни миллионов точек) объема. Проблема заключается в том, что первоначально данные удалены и распределены. И как следствие, возможен объем данных, превышающий память рабочей станции, необходимы конструктивные решения по выбору, отбору, передаче и слиянию данных. Off-line визуализация упрощает решения, но в ряде случаев требует перераспределения данных, что приводит к потерям вычислительного пространства и времени. Наиболее эффективным способом уменьшения объема передаваемых и обрабатываемых данных является фильтрация (предобработка). Разрабатываемую систему следует рассматривать в более широком контексте - "Визуальных супервычислений". Это понятие вводится в работе [3], где указывается, что "Визуальные супервычисления" касаются инфраструктурных технологий для поддержки визуальных и интерактивных вычислений вообще и визуализации в частности в сложных сетевых вычислительных средах. Данная работа посвящена как рассмотрению технологии параллельной фильтрации данных в области визуальных супервычислений, так и применению фильтрации в конкретных случаях визуализации сеточных данных.

## 2. ПАРАЛЛЕЛЬНАЯ ФИЛЬТРАЦИЯ ДАННЫХ

Можно говорить о появлении нового термина - параллельная фильтрация данных или 'parallel data filtering'. Хотя интуитивно понятно, что эта технология связывает параллельную, интерактивную обработку данных с их on-line визуализацией, строгого определения этого понятия пока не существует. Параллельная фильтрация данных является новой и важной технологией, которая лежит в основе целого ряда разрабатываемых систем интерактивной визуализации в области параллельных вычислений [1],[4]. Хотя предварительный отбор или фильтрация данных является хорошо известной технологией, так, например, существуют работы по информационной визуализации, связанные с фильтрацией для распределенных баз данных, а понятие

фильтр используется в ряде графических библиотек (например, VTK), анализа и описания применения этого подхода в рамках визуальных супервычислений до настоящего времени не встречалось.

Какие знания к настоящему времени накоплены в различных областях относительно фильтрации? Так библиотека VTK позволяет отображать различные наборы данных, а также предоставляет простейшие механизмы для вспомогательного анализа данных. Одним из таких механизмов является механизм фильтров. С его помощью можно строить, например, линии уровня. Классом, отвечающим за построение фильтров в VTK, является vtkCutter.

Наибольшее развитие технология фильтрации получила в базах данных. Фильтры представляют собой технические или программные средства для выделения данных, удовлетворяющих каким-либо условиям.

Рассмотрим проблемы и достижения относительно визуальной фильтрации [5]. Все чаще пользователи должны затратить дополнительное время и когнитивные усилия, чтобы динамически определить отображение (вывод на экран), так чтобы вся информация и только та, которая требуется, была легко визуально доступна. Соответствующие интерактивные технологии основаны на задачах выделения информации: фильтрации и линковки. Фильтрация – это процесс удаления незначимых объектов и атрибутов данных. Линковка включает в себя поиск, для того чтобы связать различные визуальные элементы.

Технология фильтрации ограничивает видимые данные подмножествами, представляющими интерес. Интерактивная фильтрация позволяет пользователю быстро и динамически удалять, добавлять и объединять атрибуты и объекты данных на экран. В то время, как она является намного более гибким подходом, чем стандартное действие "показать/скрыть" во многих системах, в действительности она имеет ограничение. Необходимо отменить запрос, чтобы вернуться к предыдущему виду отображения данных. В некоторых случаях менее вмешивающийся метод был бы предпочтительнее.

Общепринятый визуальный подход закрашивание ("brushing") интерактивно выделяет подмножества данных, чаще всего цветом. Он поддерживает визуальное связывание разнородных объектов и поэтому направлен на проблему визуальной фрагментации. В отличие от фильтрации, которая удаляет данные с дисплея, закрашивание дает добавочную информацию, накладывая уточненное изображение на существующую структуру. Однако закрашивание требует собственной кодирующей величины, то есть графический атрибут, используемый для выделения выступающих точек, должен быть сохранен, будь то цвет, форма или текстура. Эти графические атрибуты уже стандартно применяются в визуализации с многими параметрами (множественной визуализации). В частности, цвет удобен для категоризации.

Но, как известно, глаз не чувствителен к изменению цвета на периферии. Необходимы эффективные технологии фильтрации и закрашивания, в которых отсутствуют вышеприведенные недостатки. Большой потенциал имеет применение движения в этих технологиях.

Проанализировав имеющиеся результаты, предлагается сформулировать основные понятия следующим образом

Фильтр – это любая операция над данными, изменяющая их количество. Из этого определения следует, что объединение объектов или слияние баз данных является фильтром. Текстовое сжатие изменяет не количество данных, а их длину, следовательно, не является фильтром. Аналогично, изменение цвета, вращение, перемещение, увеличение (без изменения уровня детализации) объекта меняют некоторые значения, а не количество, то есть эти операции не являются фильтрами.

Параллельная фильтрация данных – это интерактивный процесс над распределенными данными, где на каждом шаге применяется параллельный фильтр, происходит передача отфильтрованных данных и их представление в текстовом или визуальном виде, имеющем вложенный цикл стандартного взаимодействия.

Целью фильтрации данных является получение необходимой (интересующей) информации за минимальное время. Следовательно, операции, которые являются фильтрами, над данными надо выполнять параллельно, минимизировать количество передаваемых данных, применять когнитивную визуализацию.

Важным вопросом является отличие параллельного фильтра от родительского, последовательного фильтра. Если операция выполняется над распределенными данными, которые не являются тесно связанными, то параллельный фильтр совпадает с последовательным.

Параллельный фильтр тесно связан с видом отображения, следовательно, параллельная фильтрация данных относится к распараллеливанию графических алгоритмов. Особое значение эта технология приобретает для визуализации данных большого объема.

Основной задачей при визуализации данных большого объема является сокращение передаваемых данных для последующей визуализации. Так как закрашивание в общем случае применяется уже к сформированной структуре данных, то оно не решает поставленной задачи. Подходы к решению этой проблемы можно разбить на два основных класса – это применение интерактивности, а также сжатие и фильтрация данных [6]. Вообще, решение задачи визуализации данных большого объема требует серьезной инфраструктурной поддержки [3]:

Обеспечение достаточного объема памяти;

Управление механизмом распределения данных;

Выбор самого эффективного алгоритма;

Предоставления удобного и эффективного графического интерфейса.

Все это должно обеспечить получение требуемого решения за приемлемое время.

Данная работа рассматривает фильтрацию в рамках проблемно-ориентированного подхода. Для визуализации передаются отфильтрованные математические данные (“raw” – сырые, необработанные данные), связанные с видом отображения, а не графические примитивы или растровое изображение. Это позволяет воспользоваться аппаратными возможностями видеокарт и в ряде случаев сократить объем передаваемых данных.

Преимущества применения параллельной фильтрации данных (предобработки) на счетных узлах вычислителя сейчас для многих очевидны [7]: обрабатываемые для визуализации данные уже находятся на узлах вычислителя, что позволяет избежать дополнительных обменов; при визуализации можно учесть характерные особенности задачи; возможна on-line визуализация большого объема данных; высокая скорость визуализации. Кроме существенных трудностей при реализации этого подхода основным недостатком является приостановка счетной задачи во время предобработки.

### 3. ПРИМЕНЕНИЕ ПАРАЛЛЕЛЬНОЙ ФИЛЬТРАЦИИ ДАННЫХ

Можно классифицировать фильтры по функциональности, описывая, например, пространственную фильтрацию, фильтрацию по определенным качественным признакам или по значению. Существуют также фильтры, связанные с видом отображения, например, сечение плоскостью, изоповерхность, линии тока. К фильтрам можно отнести такие операции, как сжатие и децимация. В рамках концепции параллельной фильтрации данных можно рассматривать фильтры как преобразования, обеспечивающие целостное восприятие или детализацию. При этом, правда, возможна значительная потеря качества. В связи с этим целесообразно использовать комплексные (множественные) виды отображения. Наконец, можно ввести фильтры, ориентированные на полигональную или воксельную графику. Целесообразно также применять конвейер фильтров или многоступенчатую фильтрацию на одном шаге взаимодействия.

В контексте специализированной системы визуализации данных был использован конвейер фильтров, включающий пространственную фильтрацию и фильтрацию «по значению». Пространственная фильтрация заключается в выделении необходимого объема сеточных данных на основе их пространственного положения. Очевидно, что одновременный вывод на экран всего слишком большого объема данных влечет за собой снижение уровня детализации. Кроме того, из-за большой размерности визуализируемых данных, более далекие от наблюдателя участки могут оказаться вне поля зрения, скрытыми более близкими участками.

Пространственная фильтрация отчасти решает эту проблему, выделяя область интересующей нас информации из общего ее объема. Функции пространственной фильтрации реализованы с помощью метафоры «альфа-сферы». Суть метафоры состоит в том, что из общего объема данных выделяется сферическая область (Рис. 1). При этом визуальная прозрачность объектов, находящихся внутри области, прямо пропорциональна удаленности объектов от

центра области. Таким образом, информация, представляющая наибольший интерес, видна лучше всего, а остальная, тем меньше, чем более удалена от центра области рассмотрения. В связи с внедрением средств виртуальной реальности возникает потребность в видах отображения "изнутри". Вследствие чего метафору альфа-сферы можно расширить до вида от первого лица. При визуализации сеток различной структуры и объема были получены результаты, свидетельствующие об эффективности этих метафор и механизмов пространственной фильтрации в целом.

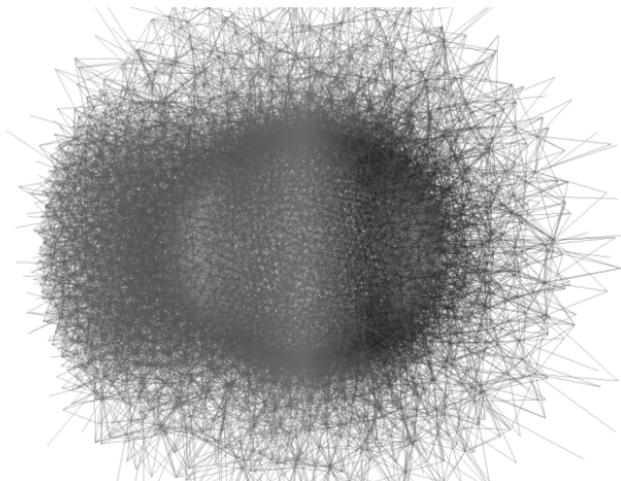


Рис. 1. Пространственная фильтрация данных

Фильтрация «по значению» необходима для отсека ненужных участков визуализируемых данных на основании значения каких-либо характеристик, заданных на этих участках. Для этих целей введен дополнительный компонент пользовательского интерфейса - поле диапазонов. Диапазон представляет собой отрезок  $[a, b]$ , заданный на оси соответствующей характеристики. Диапазоны могут пересекаться и накладываться друг на друга. Результирующая фильтрация проводится по объединению всех диапазонов, расположенных на шкале. Фильтрация данных по значению, как правило, подразумевает вывод численной информации, что является полезным.

Структура сетки задается в виде геометрических примитивов (тетраэдров кубов и др. многогранников) [2]. При пространственной фильтрации выбираются примитивы, попавшие в заданную область фильтрации ("альфа-сфера"). Если одна из вершин примитива попадает в область фильтрации, то считается, что весь примитив принадлежит области фильтрации. К выбранным примитивам применяется фильтрация по значению. Между вершинами примитива производится линейная интерполяция скалярных параметров вершин. При фильтрации, ребро примитива урезается в соответствии с границами диапазона. В результате, получается набор отрезков со значениями на границах. Эти данные и передаются для последующей визуализации. Значению скалярного параметра ставится в соответствие цвет так чтобы обеспечить равномерный переход цвета от синего (наименьшее значение) к красному (наибольшее значение) через зеленый (среднее значение). Для этого используется

пиксельный шейдер DirectX, обеспечивающий преобразование скалярного значения в цвет видеокарты (аппаратная поддержка) (Рис. 2).

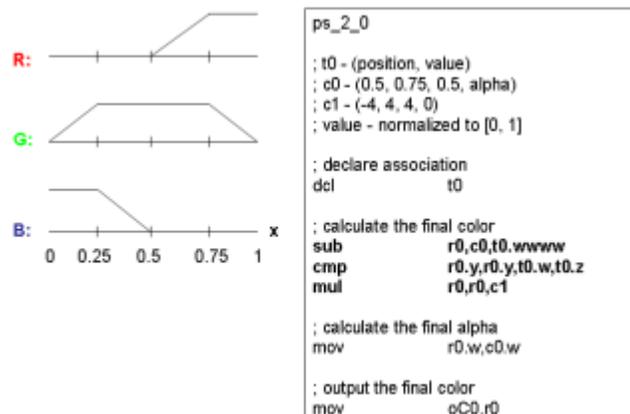


Рис. 2. Преобразование значения скалярного параметра в цвет

#### 4. ЗАКЛЮЧЕНИЕ

Ведутся активные работы по распараллеливанию системы визуализации. В связи с этим, все фильтры, используемые в системе, будут расширены до соответствующих параллельных. Следует заметить, что в текущей концепции распределенной системы визуализации, параллельные фильтры могут быть вынесены в отдельные самостоятельные модули.

В настоящий момент рассматривается off-line визуализация с равномерным перераспределением данных по процессорам, так чтобы количество ребер примитивов на каждом процессоре было примерно одинаково. Для ускорения повторной загрузки данных используется сжатие LZMA (в результате уменьшается объем передаваемых данных по сети). Далее выполняется параллельная фильтрация данных и интерактивная визуализация. В результате должна быть обеспечена возможность отображения данных в реальном времени на сетках порядка сотни миллионов вершин.

Работа выполнена при поддержке гранта РФФИ № 04-07-90120

#### 5. БИБЛИОГРАФИЯ

- [1] Авербух В.Л., Васёв П.А., Горбашевский Д.Ю., Казанцев А.Ю., Манаков Д.В. Система интерактивной визуализации параллельных вычислений // 14-я Международная Конференция по Компьютерной Графике и Зрению ГрафиКон'2004 6-10 Сентября 2004 Москва, Россия. Труды Конференции. МГУ им. М.В. Ломоносова. Стр. 291-294.
- [2] Горбашевский Д.Ю., Казанцев А.Ю. Визуализация сеточных данных большого объема // 15-я Международная Конференция по Компьютерной Графике и Зрению

ГрафиКон'2005 20-24 Июля 2005 Новосибирск, Россия. Стр. 366-367.

[3] K. Brodli, J. Brooke, M. Chen, D. Chisnall, A. Fewings, C. Hughes, N. W. John, M. W. Jones, M. Riding and N. Road, [Visual Supercomputing - Technologies, Applications and Challenges](#), *Eurographics 2004, STAR Reports*, 37-68.

[4] M. Thiebaut, H. Tangmunarunkit, K. Czajkowski, C. Kesselman. Scalable Grid-Based Visualization Framework. Technical report ISI-TR-2004-592, USC/Information Sciences Institute, June 2004.

[5] Lyn Bartram, Colin Ware, Filtering and brushing with motion, *Information Visualization*, v.1 n.1, p.66-79, March 2002

[6] Manakov D., Mukhachev A., Shinkevich A. Visualization of the Distributed Data of Huge Volume. Assembly, Filtration, Sorting. // *Proceedings of the 13-th International Conference on Computer Graphics and Vision Graphicon-2003 Moscow*, September 5-10, 2003. pp. 198-201.

[7] Karpov A.N. Data visualization on parallel computer systems // *Proceedings of the 15-th International Conference on Computer Graphics and Vision Graphicon 2005*, Novosibirsk, June 20-24, 2005, pp. 211-214. (In Russian)

## Об авторах

Горбашевский Дмитрий Юрьевич, аспирант ИММ УрО РАН

Казанцев Алексей Юрьевич, аспирант ИММ УрО РАН

Манаков Дмитрий Валерьянович ИММ УрО РАН, гл. программист, Екатеринбург, ИММ УрО РАН. e-mail: [manakov@imm.uran.ru](mailto:manakov@imm.uran.ru)

# Parallel Filtering in Parallel Computing Visualization System

## Abstract

In the paper the theoretical aspects of parallel filtering in Visual Supercomputing are considered. The example of parallel filtering used in specialized visualization system is described.

**Keywords:** *Parallel Filtering, Visual Supercomputing.*

## About the authors

Dmitriy Gorbachevskiy, Ph.D. student at Institute for Mathematics and Mechanics Urals Branch of Russian Academy of Science

Alexey Kazantsev, Ph.D. student at Institute for Mathematics and Mechanics Urals Branch of Russian Academy of Science

Dmitriy B. Manakov, Institute of Mathematics and Mechanics, Ural Branch of the Russian Academy of Sciences, programmer, Ekaterinburg, RUSSIA. e-mail: [manakov@imm.uran.ru](mailto:manakov@imm.uran.ru)