

## **Обзор методов визуализации данных о вербальном и невербальном поведении человека**

*Князев Б.А., аспирант  
кафедры «Системы обработки информации и управления»,  
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана*

*Рыбаков К.А., студент  
кафедры «Системы обработки информации и управления»,  
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана*

*Научный руководитель: Гапанюк Ю.Е., к.т.н., доцент  
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана  
[bknyazev@bmstu.ru](mailto:bknyazev@bmstu.ru)*

### **Введение**

Существует несколько определений понятия невербального поведения человека. «Невербальное поведение» – внешняя форма существования и проявления психического мира личности [1]; индивидуальная, конкретно-чувственная форма действий и поступков, передаваемых индивидом при помощи невербальных средств общения. [2]. При этом существуют следующие формы проявления невербальных средств общения [3]:

- двигательные или кинесика: жесты, мимика, позы, походка, взгляд, манера действий);
- акустические характеристики голоса человека: частота, интенсивность, ритм и др.;
- тактильно-кинестезические;
- ольфакторные;
- пространственно-временные характеристики;
- и другие.

«Вербальное поведение» включает вербальные реакции, такие как говорение, реагирование на слова, запоминание вербального материала и т.д.

Каждая из перечисленных форм проявления невербального и вербального поведений может быть представлена в виде процесса, непрерывно изменяющегося во времени. Временные и частотные характеристики данных процессов отличаются друг от

друга. Существуют различные способы регистрации и измерения этих характеристик, однако, крайние значения при этом следующие:

- частота движения пальцев рук  $\leq 12$  Гц [4,5];
- частота жестов рук и частота движения в целом  $\leq 10$  Гц [5-8];
- частота изменения мимики лица  $\leq 4$  Гц [9];
- скорость движения глаз  $600$  градусов/сек (длительность фиксации в одном положении  $\geq 250$  мс) [10];
- около 90% энергетической составляющей речевого сигнала лежит в диапазоне 100-5000 Гц [11].

Таким образом, из теоремы Котельникова следует, что для исключения значительных потерь исходного сигнала частота дискретизации исследуемых в данной работе невербальных и вербальных сигналов должна быть  $\geq 2.5$  кадров/сек и  $\geq 10$  КГц соответственно. Однако для более детального исследования и исключения потерь необходимо повышение данных частот до 100 кадров/сек (система PRIMAS) и 20 КГц.

Визуализация параметров данных сигналов позволяет повысить эффективность интерпретации и объективной оценки данных. Объективная оценка параметров невербального и вербального сигналов необходима для решения задач [12]:

- безопасности (определение лжи, возбужденности, намерений и др.);
- распознавания языка жестов [6, 13];
- медицинской и психологической диагностики (болезнь Паркинсона [4], шизофрения, алкоголизм [7] и др.);
- повседневного общения людей с компьютером или другим оборудованием;
- робототехники и т.д.

Объективная оценка может осуществляться с помощью эксперта в соответствующей области или может быть автоматизирована с помощью экспертных и продукционных правил или других способов автоматизации принятия решения.

Длительность исследований, записанных на видео и/или аудио материал, может достигать нескольких часов. В общем случае количество отсчетов данных для визуализации  $N$ :

$$N = 3600 \cdot F \cdot L \text{ точек,} \quad (1)$$

где  $L$  – длительность исследования в часах,  $F$  – частота отсчетов в секунду.

Так, например, при стандартной частоте видеок кадров  $F=25$  кадров/сек и длительности видеоматериала  $L=5$  часов количество точек для отображения:

$$N = 3600 \cdot 25 \cdot 5 = 4.5 \cdot 10^5 \text{ точек.}$$

При стандартной частоте дискретизации аудио сигнала  $F_{\text{Дискр}}=11025$  Гц и длительности аудиоматериала  $L=5$  часов:

$$N = 3600 \cdot 11025 \cdot 5 = 2 \cdot 10^8 \text{ точек.}$$

При этом объем данных для визуализации:

$$V = N \cdot B/8 \text{ байт,} \quad (2)$$

где  $B$  – количество бит на один отсчет (англ. *bits per sample*).

В данной работе предлагается обзор методов и алгоритмов визуализации данных большого размера, позволяющих эффективно отображать отсчеты вербального и невербального поведений как процессов, изменяющихся во времени; представляются результаты разработки и исследования характеристик некоторых из этих методов.

### **Подходы к визуализации данных большого размера**

В качестве данных о невербальном или вербальном поведении могут быть отсчеты некоторого процесса  $S$  во времени, график которого необходимо построить (рис. 1, рис. 3, а). Визуализация данного процесса может осуществляться с уменьшением и без уменьшения размерности данных (табл. 1). Снижение размерности, в свою очередь, может осуществляться на уровне визуализации и/или на уровне обработки данных. Проблема отображения данных без их предварительной обработки и упрощения кривой заключается в ограниченности ресурсов электронной вычислительной машины (ЭВМ). Так, при  $B=16$  в соответствии с (2) необходимо отобразить  $V = 2 \cdot 10^8 \cdot 16/8 = 4 \cdot 10^8$  байт  $\approx 0.37$  Гб информации.

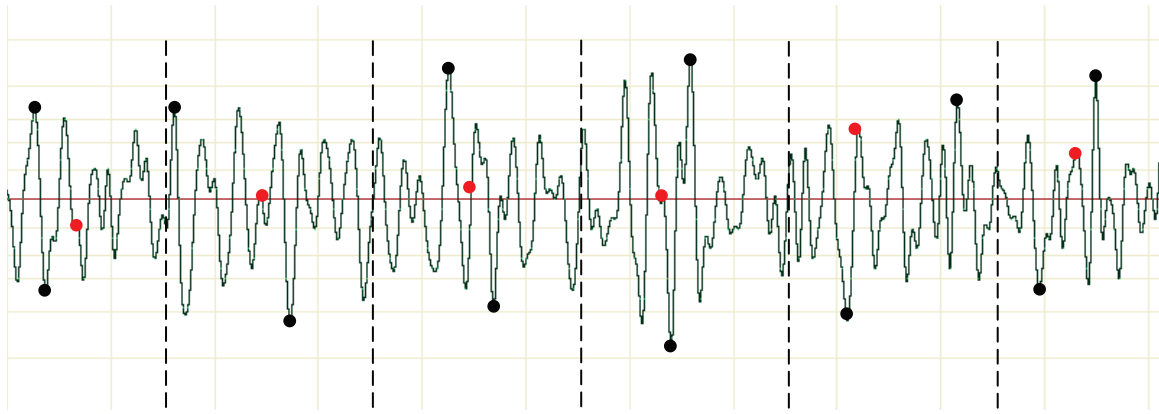
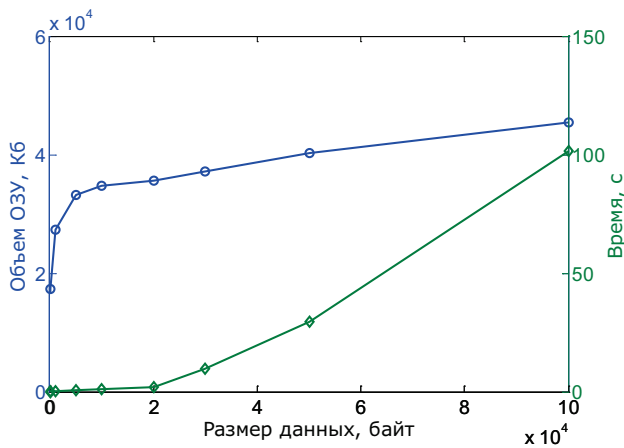
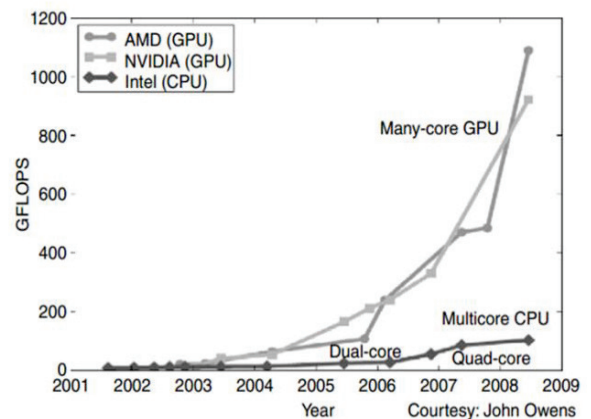


Рис. 1. График исходных данных  $S$ ; черными точками показаны требуемые отсчеты; красными – неточно выбранные отсчеты.

При этом оценка использования ресурсов, используемых ЭВМ для визуализации, показала, что при размере данных всего 100 Кб их отображение занимает около 100 секунд, размер выделенной памяти ОЗУ около 40 Мб (рис. 2, а). Тем не менее, визуализация данных может осуществляться эффективно с помощью встроенных в программную среду, такую как Microsoft Visual Studio, средств. Например, комбинация библиотек C++ и .NET Framework позволяет записывать отсчеты данных с помощью прямого доступа к памяти объекта типа *MapViewOfFile* и последующего доступа к объекту изображения *InteropBitmap* (рис. 3, б).



а)



б)

Рис. 2. а) Оценка производительности построения данных с помощью CPU средствами .NET Framework; б) Сравнение производительности процессоров GPU и CPU [14].

Также существует параллельная реализация данного метода, позволяющая строить и масштабировать до  $10^8$  точек на ноутбуке Dell Latitude D820 [15]. Недостаток данного метода заключается в том, что размер визуализируемых данных ограничен размером

оперативной памяти системы и при построении данных под процесс будет производиться попытка выделения такого объема памяти, какой занимают данные, что вызовет большую загрузку общей ОЗУ и увеличит время отображения данных, то есть сделает работу программы неэффективной. Также необходимы алгоритмы построения кривых между точками, так как метод позволяет отрисовывать только точки. Имеющиеся в открытом доступе библиотеки .NET для визуализации данных также позволяют производить эффективное построение графиков. Так, с помощью библиотеки *DynamicDataDisplay* было построено около  $6 \cdot 10^5$  точек. Другие библиотеки демонстрируют аналогичные результаты.

Таблица 1

Сравнение методов визуализации данных

Методы	Пример методов	Достоинства	Недостатки
Прямой доступ к памяти	MapViewOfFile+InteroperBitmap	Позволяют эффективно строить большие массивы точек	Необходимы алгоритмы построения кривых между точками
Имеющиеся средства	библиотеки .NET DynamicDataDisplay, Oxyplot, ZedGraph и др.	Быстрое построение, удобное использование и разработка	Необходимы алгоритмы уменьшения размерности данных
Аппроксимация	Алгоритм Ramer–Douglas–Peucker	Значительное снижение размерности	Потеря данных
Децимация (downsampling)	Частотные методы	Незначительное снижение размерности	Потеря данных
	Методы агрегации	Значительное снижение размерности	Возможная потеря данных

Среди методов аппроксимации кривой алгоритм *Ramer–Douglas–Peucker* и его разновидности демонстрируют лучшие показатели точности и приближенности к оригинальной кривой [16]. Он основан на рекурсивном вычислении перпендикулярного расстояния между исходной и аппроксимированной кривой, построенной с меньшим количеством точек и не далее, чем на величину  $\epsilon$  от исходной кривой, где  $\epsilon$  – величина ошибки, с помощью которой можно варьировать размер выходных данных. Сложность алгоритма *Ramer–Douglas–Peucker* равна  $O(n \cdot \log n)$  и  $O(n^2)$  в лучшем и худшем случаях соответственно. Также существует эффективная параллельная реализация данного метода

[17]. Несмотря на это, недостаток данного алгоритма заключается в сложности нахождения оптимальной величины  $\epsilon$  для требуемой точности отображения и степени сжатия данных. В результате итоговый график может выглядеть неточно либо коэффициент сжатия будет недостаточен для быстрого построения и навигации по графику (рис. 3, в-д). Более того, данные недостатки увеличиваются с размерностью данных  $S$ , так как время поиска оптимальной кривой является как минимум квадратичной зависимостью от размера исходной кривой [16].

Методы децимации заключаются в выборе каждого  $i$ -го отсчета среди исходных данных по заданному правилу  $R$ , где  $i$  зависит от точности, которую необходимо получить. Недостаток данных методов в том, что, как и у алгоритмов аппроксимации, важные экстремумы могут быть пропущены при выборе неподходящего правила  $R$ . Например, при  $R = i \Rightarrow \text{IndexOf}(i) \text{ Mod } 10 == 0$  будет выбран каждый 10-ый отсчет и важные отсчеты могут быть не учтены (рис. 1).

Частотные методы включают низкочастотные и избирательные фильтры, спроектированные под конкретный тип сигнала. В то же время исследование характеристик вербального и невербального поведений требуется как на низких, так и на высоких частотах.

Под агрегированием в общем случае понимается объединение нескольких элементов в единое целое. Одним из правил агрегирования может быть выбор экстремумов на локальных областях графика. При таком подходе расчеты для каждого блока данных, соответствующих некоторой области графика, могут производиться независимо от других блоков. Тогда для решения задачи визуализации данных с помощью их предварительной агрегации возможно применение параллельных методов, так как они [14]:

- предназначены для вычисления независимых блоков данных;
- имеют необходимую для данной задачи пропускную способность памяти (до 63.4 Гб/сек для графического процессора G80);
- при корректном использовании ресурсов GPU значительно (до 100 раз) превосходят результаты центрального процессора.

Так, пиковая пропускная способность процессора G80 достигает 346 GFLOPS ( $346 \cdot 10^9$  операций с плавающей запятой в секунду), в то время как для процессора CPU Intel E7500 она достигает 23 GFLOPS (рис. 2, б).



а)  $N = 2048$



б) MapViewOfFile+InteropBitmap



в)  $N = 1432, \epsilon = 1$



г)  $N = 843, \epsilon = 3$



д)  $N = 220, \epsilon = 10$

Рис. 3. Реализация алгоритмов средствами .NET:

а) исходный несжатый график; б) график, отображенный методом прямого доступа к объекту изображения; в)-д) графики, аппроксимированные алгоритмом *Ramer-Douglas-Peucker* с параметрами  $N$  и  $\epsilon$

## Заключение

В данной работе приведен обзор и краткий анализ методов и алгоритмов визуализации данных большого размера, таких как данных о вербальном и невербальном поведении. Обоснована актуальность исследования данных методов; проанализированы встроенные в программную среду средства, методы прямого доступа к памяти и объектам изображения, методы децимации и агрегации данных, а также методы аппроксимации кривых, которые эти данные представляют; разработана программная реализация алгоритма *Ramer-Douglas-Peucker* и метода прямого доступа к объекту изображения. Исследуемые методы не продемонстрировали требуемой для данной задачи скорости и точности построения, поэтому в качестве наиболее перспективного метода предложен

метод параллельной агрегации данных. В последующих работах предлагается использование CUDA API для реализации параллельного метода *reduction tree* с целью разработки эффективного метода визуализации данных.

### Список литературы

1. Лабунская В. А. Невербальное поведение (социально-перцептивный подход). Ростов-на-Дону: «Феникс», 1988. 246 с.
2. Горелов И. Н. Невербальные компоненты коммуникации. М.: Наука, 1980. 238 с.
3. Бодалев А.А. Восприятие и понимание человека человеком. М.: Изд-во Моск. ун-та, 1982. 200 с.
4. Analysis of finger-tapping movement// Jobbagy A. [et al.] Neurosci Methods. 2005. 141(1). p. 29-39
5. Ильин Е. П. Психомоторная организация человека: Учебник для вузов. СПб.: Питер, 1-е издание, 2003. 384с.
6. Wilbur R.B., Martinez A.M. Physical Correlates of Prosodic Structure in American Sign Language// Chicago Linguistic Society, 2002
7. York JL, Biederman I. Hand movement speed and accuracy in detoxified alcoholics// Alcohol Clin Exp Res. 1991. 15. p. 982-990
8. Are listeners paying attention to the hand gestures of an anthropomorphic agent: An evaluation using a gaze tracking method// S. Nobe [et al.] Lecture Notes in Computer Science. Springer Verlag Heidelberg. 1998. Vol. 1371. p. 49-59
9. V. Fasel, J. Luetttin. Automatic facial expression analysis: a survey// Pattern Recognition. 2003. Vol. 36, Iss. 1. p. 259-275
10. Quantitative measurement of saccade amplitude, duration, and velocity// Baloh R.W. [et al.] Neurology. 1975. 25. p. 1065-1070
11. Сапожков М.А. Электроакустика. Учебник для вузов. М.: Связь, 1978. 272 с.
12. Алфимцев А.Н. Разработка и исследование методов захвата, отслеживания и распознавания динамических жестов: Дис. канд.техн.наук. Москва, 2008, 167 с.
13. Девятков В.В., Алфимцев А.Н. Распознавание манипулятивных жестов// Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение. 2007. Т. 68, № 3. С. 56-75.
14. David B. Kirk, Wen-mei W. Hwu. Programming Massively Parallel Processors: A Hands-on Approach. San Francisco: Morgan Kaufmann Publishers Inc., CA, USA, 2010, 280 p.



15. How to: High performance graphics in WPF//Tamir Khason – Just code.2008.URL.<http://khason.net/blog/how-to-high-performance-graphics-in-wpf/> (дата обращения: 03.03.2013)
16. P. Heckbert, M. Garland. Survey of Polygonal Surface Simplification Algorithms. New York: ACM Press, Siggraph 97 Course Notes, No. 25, 1997.
17. A real-time parallel implementation of Douglas-Peucker polyline simplification algorithm on shared memory multi-core processor computers. // J. Ma [et al.] 2010 International Conference on Computer Application and System Modeling (ICCASM), 2010. Vol. 4, p. 647-652.