

Модели восприятия визуальной информации

Д.В. Манаков

manakov@imm.uran.ru

Институт математики и механики им. Н.Н. Красовского Уральского отделения РАН, Екатеринбург, Россия

Работа посвящена рассмотрению математических моделей зрительного восприятия. Целью изучения данного вопроса является не только повышение алгоритмического качества систем и устройств визуализации, но и возможность разработки самообучающихся тренажеров, основанных на генерации отрицательной обратной связи, например, с позиций диссипативных или функциональных систем.

Ключевые слова, восприятие, диссипативные системы, методы передачи объема и глубины пространства.

The visual information perception models

D.V. Manakov

manakov@imm.uran.ru

Institute of Mathematics and Mechanics im. N.N. Krasovskogo Ural Branch of the Russian Academy of Sciences, Ekaterinburg, Russia

The paper is devoted to the mathematical models of visual perception. The purpose of this research is both to improve the algorithmic quality of visualization systems and devices, and to develop self-learning simulators based on the generation of negative feedback, for example, from the position of dissipative or functional systems.

Key words: perception, dissipative systems, methods of transferring the volume and the space depth.

1. ВВЕДЕНИЕ

Оценка эффективности визуализации должна рассматривать не только качество визуального дизайна, но и обеспечение соответствующего понимания, что особенно важно в случае когнитивной визуализации. Одним из параметров эффективности теории принятия решения является уровень доверия. Результаты тестирования показывают, что текстовое представление информации вызывает большее доверие, чем визуальное. Это можно объяснить тем, что в тексте меньше многозначности и неопределенности по сравнению с визуализацией. С целью минимизации этих факторов программисты должны максимально доверять математической модели, поэтому верификации программ и визуализации уделяют значительное внимание.

Рассмотрение моделей восприятия-понимания является частным случаем верификации программ визуализации. Разделение человеческого мышления на восприятие и понимание достаточно условно и не всегда оправданно. Принято выделять четыре основных (и связанных с математической формализацией) направления исследования процесса мышления:

1. Кибернетика (например, восприятие как параллельный процесс).
2. Бионика.
3. Нейронные сети.
4. Диссипативные системы (частный случай рассмотрения гомеостаза – функциональные системы).

Последнее направление связано с самоорганизацией, и мы остановимся на нем более подробно.

Целью данной работы является рассмотрение возможных постановок задач основных тематических направлений изучения зрительного восприятия в области компьютерной визуализации, прежде всего с позиций математического моделирования. Несмотря на высокую степень полимодальности человеческого зрения и восприятия в целом тематический круг подобных задач достаточно узок:

1. Сравнение текстовой и визуальной информации.
2. Зрительная фильтрация потока информации.

3. Визуальный анализ и визуализация неопределенности.
4. Оценка качества (эргономичности) дисплеев (реалистичности изображения).
5. Расширенная визуализация (в основном сонификация).
6. Сравнение стереоизображения и других методов передачи объема и глубины пространства (например, параллакс движения) и их интегрирование.
7. Виртуальная и расширенная (дополненная) реальность.
8. Тестирование и тренажеры.

Конечной целью изучения зрительного восприятия для компьютерной визуализации является не только повышение алгоритмического качества систем и устройств визуализации, но и возможность разработки самообучающихся тренажеров, основанных на генерации отрицательной обратной связи, например, с позиций диссипативных или функциональных систем.

2. ДИССИПАТИВНЫЕ СИСТЕМЫ

Диссипативная система - квазистационарная открытая система, характерной особенностью которой является процесс самоорганизации, происходящий в результате действия отрицательного вектора, например, силы трения.

Поясним отличие диссипативной системы от задачи оптимального управления в случае линейных систем:

$$\dot{x} = Ax - Bu,$$

где u – оптимальное управление, которое надо найти. Поскольку диссипативная система является открытой системой управление приходит извне, оно изначально задано. Введение в рассмотрение отрицательного вектора или отрицательной обратной связи сужает круг задач до моделей с насыщением.

Область применения диссипативных систем достаточно широка, начиная от микроуровня - рассмотрения клетки организма, до макроуровня – моделирование устройства вселенной. Теорию хаоса и теорию эволюции можно считать частными случаями диссипативных систем. Близкими по духу являются в области экономики наука синергетика, а в области программирования автономные вычисления. Поскольку данная работа посвящена восприятию визуальной информации, более подробно

остановимся на теории функциональных систем П.К. Анохина.

Человек находится в состоянии гомеостаза (почти статического равновесия). Выделяются два типа функциональных систем. Системы первого типа обеспечивают гомеостаз за счёт внутренних (уже имеющихся) ресурсов организма, не выходя за его пределы – замкнутые системы. Системы второго типа поддерживают гомеостаз за счёт изменения поведения, которое формируется в результате отрицательной обратной связи – открытые (диссипативные) системы.

Функциональные системы второго типа представляют интерес для создания самообучающихся тренажеров, основанных на генерации отрицательной обратной связи, но в этом случае, скорее всего, надо рассматривать задачу оптимального управления. Перспективным направлением является изучение энцефалограммы мозга с целью распознавания обратной связи и валидации моделей восприятия.

Последние исследования в области диссипативных структур позволяют делать вывод о том, что процесс самоорганизации происходит гораздо быстрее при наличии в системе внешних и внутренних шумов. Таким образом, шумовые эффекты приводят к ускорению процесса самоорганизации. Тремор глаз и тремор рук – колебания малой амплитуды, фактически это шум, и их необходимо учитывать, как в процессе моделирования, так и с целью ускорения обучения на тренажерах.

Ключевым понятием диссипативных систем является энтропия. Например, термодинамическая энтропия - мера необратимой диссипации энергии. Поскольку понятие термодинамической энтропии и информационной энтропии (информативности) близки или даже эквивалентны была поставлена цель найти иллюстративный пример динамики диссипативных систем в области информационных технологий. Оказалось проще взять известный пример [5] и рассмотреть его с точки зрения диссипативных систем. Рис. 1. представляет визуальный анализ деятельности кредитных организаций. Он позволяет строить суждения об образовании кластеров (диссипативных структур) и находить объекты, стремящиеся оказаться в кластере (аналогично, модели движения амёб в сторону вещества с наибольшей концентрацией). С точки зрения диссипативных систем на следующем шаге кластер должен разрушиться (обычно по превышению некоторого порогового значения) и на его периферии должен сформироваться новый кластер.

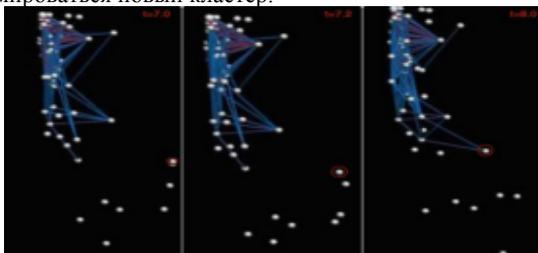


Рис. 1. Кластер и объект, стремящийся оказаться в кластере [5].

Как уже отмечалось, диссипативные системы являются не единственным направлением в области изучения зрительного восприятия и человеческого мышления, но имеют достаточно широкую область применения. Например, они применяются в области лингвистики, а частный случай - функциональные системы, может рассматриваться не только как одна из методик создания тренажеров, но и изучения зрительной фильтрации потока информации – одного из направлений изучения зрительного восприятия.

3. ОСНОВНЫЕ НАПРАВЛЕНИЯ ИЗУЧЕНИЯ ЗРИТЕЛЬНОГО ВОСПРИЯТИЯ

Эти направления были перечислены во введении работы. Конечно, они частично пересекаются между собой, а восприятие пересекается с пониманием, но имеют определенное самостоятельное значение, и поэтому все это можно считать базисом или семантическими единицами ментального пространства визуализации. Для того, чтобы рассматривать компьютерную визуализацию как самостоятельную дисциплину, необходимо сформировать ментальное пространство визуализации со своей семантикой, прагматикой и базисом. Визуальные тексты можно рассматривать в качестве базиса универсального пространства визуализации [4], то есть они имеют основополагающее истинностное значение. Верификация визуализации - один из основных инструментов создания данного пространства, и поэтому тематические направления изучения зрительного восприятия будут представлены не столько в качестве обзора литературы, а и в качестве математических постановок задач, требующих решения.

Кибернетическое направление рассмотрения визуального текста как непрерывного отображения является перспективным, но нерешенных вопросов слишком много. В том числе по модели потока данных, которая имеет частичное пересечение с методикой зрительной фильтрации потока информации [6]. Далее остановимся на этой методике изучения зрительного восприятия и на технологии фильтрации данных, которая имеет важное (близкое к единице) значение в случае обработки данных большого объема, включая результаты параллельных вычислений.

3.1 Зрительная фильтрация потока информации

Данная методика [6] формально связана с функциональными системами второго типа и теорий управления. С целью изучения зрительного восприятия применяется кибернетический подход манипуляции входным потоком информации (изменение двух параметров объектов: цвета и формы). Например, применение светофильтра в рамках натурального испытательного стенда позволят виртуальную блокировку зрительного нерва. Для данной методики напрашивается создание виртуального испытательного стенда: применение очков виртуальной реальности и программное управление потоком информации, включая фильтрацию данных.

Как отмечается [6], у методики хорошая валидность, но на уровне восприятия постоянно возникают труднообъяснимые визуальные артефакты. Например, если в очках виртуальной реальности рассматривать горизонтальную белую линию, то на периферии зрения можно заметить гладкое разложение этой линии на три линии разного цвета. Конечно, причины возникновения подобных визуальных артефактов можно формально объяснить, но не факт, что именно эта модель срабатывает на уровне зрительного восприятия человека. Например, можно рассмотреть взвешенную сумму стереоизображения, основанного на бинокулярной диспаратности, и других видов объемного зрения.

Модели фильтрации данных [3] также кибернетически переносятся в область изучения зрительного восприятия, поскольку стандартный графический конвейер, в том числе модель потока данных, представляет собой последовательность процессов фильтрации, геометрической обработки, рендеринга.

Задача визуализации данных большого объема не имеет решения без постановки задачи фильтрации данных или

семантически близких подходов, например, таких, как агрегация, редукция, кластеризация. Фильтрация данных (фильтр) – любая операция над данными, изменяющая их количество (“не хеширование”). В этом определении “не хеширование” рассматривается, как необходимое условие установление предпорядка. Формально можно говорить о построении идеала.

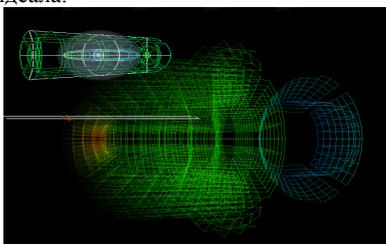


Рис. 2. Пространственная фильтрация с помощью метафоры альфа-сферы.

В качестве иллюстрации можно привести пример пространственной фильтрации данных, реализованный с помощью метафоры альфа-сферы (рис. 2). Если альфа-сферу или фильтр сечения плоскостью рассматривать как эпсилон-покрытия, то возможно применить модели дискретной оптимизации. Построение фильтров для конкретной задачи визуализации или данных можно также рассматривать как задачу оптимального управления [3].

Эпсилон-покрытия и визуальные артефакты можно рассматривать как неопределенности. Существует большое количество англоязычных обзоров по визуальному анализу, визуализации и пространственно неопределенности, часто содержащие противоречивую и недостоверную информацию, и поэтому требующие определенной ревизии. Одной из причин противоречий является то, что неопределенность может быть понятием различных математических дисциплин, например: логики (трехзначная, нечеткая логика), теории принятия решений. Мы считаем, что достаточно рассмотрению неопределенности с позиций математического анализа: предельной неопределенности - неопределенности, которая имеет конечный предел в конкретной метризуемой топологии.

3.2 Визуальный анализ, визуализация и распространение неопределенности

В работе [9] вводится понятие верифицируемой (верифицирующей) визуализации, которая отслеживает, как распространяется погрешность (неопределенность) на всем этапе вычислительного конвейера, включая визуализацию. Подобный подход в общем случае принято называть моделью с неопределенностью, а в частном - визуализацией с неопределенностью. С другой стороны в отображение интервала на интервал (компакта на компакт) нет никакой неопределенности ни с точки зрения восприятия, ни с точки зрения математики. Распространение неопределенности не так интересно, если оно не противоречит оценкам вычислительного метода, необходимо находить визуальные артефакты и объяснять их причины возникновения. Например, визуальные артефакты как следствие: зрительного восприятия, изменения ориентации нормалей (вырожденные ячейки), вычислительной погрешности. Если размер визуального артефакта уменьшился одновременно с уменьшением шага вычислений (аналогия с методом Рунге), то, скорее всего, это следствие вычислительной погрешности.

Неопределенность может быть, если отображение не является взаимно однозначным, тогда надо отображать раскрашенные интервалы. В визуальной семиотике используется понятие визуальной переменной (например, форма объекта) и ставится задача показать с позиций интуитивной логики, что знак этой переменной содержит

неопределенность. Постановка не лишена смысла, но достаточно рассмотрения стандартных математических задач: определение информативных признаков, визуальный анализ чувствительности решения в зависимости от параметров прикладной задачи, например, в области газовой динамики [2]. И ориентироваться надо не на интуитивную логику и семиотику, а на физический или математический смысл, например, в этом случае ламинарное течение, а в этом турбулентное, параметры такие. Разрешение неопределенности и отображение точек бифуркации в виде рогов, воспринимается в рамках трехзначной логики, но математически может быть связано с недифференцируемостью по параметру.

Многопараметрическая модель визуального текста является достаточно общей, и применима для визуального анализа и визуализации неопределенности, поскольку в модели, кроме пространственно-временных характеристик учитываются частные производные визуального текста по его информативным признакам [4]. Двухиндексная нотация неопределенности, в которой нижний индекс обозначает размерность независимых данных, а верхний количество зависимых переменных, пригодна для классификации неопределенности, и заимствована из классификации видов отображения (сущностей) [7], но не учитывает ни происхождения ни способа представления неопределенности.

В работе [10] рассмотрена применимость трех наиболее цитируемых теорий восприятия-понимания для анализа и визуализации неопределенности, основанных на использовании:

1. Визуальных переменных, которые начинают распознаваться на уровне сетчатки глаз.
2. Принципа графического совершенства - это то, что дает зрителю наибольшее количество идей за кратчайшее время с наименьшим количеством чернил в самом маленьком пространстве. (Возможны и другие минимаксные постановки, например, сложность интерпретации, фильтрация данных [3]).
3. Физиологических и когнитивно-психологических (в том числе гештальт-теории) исследованиях, объединяющих два предыдущих направления.

Изучение и моделирование зрительного восприятия невозможно без минимального знания нейрофизиологии особенно для объемного зрения. Структура основного зрительного пути, например, представлена в книге [6].

3.3 Стереозображение и другие методы передачи объема и глубины пространства

Чтобы не путаться будем разделять стереозображение, основанное на диспаратности зрения, как наиболее часто используемое и объемное изображение (зрение) – все остальные методы передачи объема и глубины пространства. Основные механизмы оценки объема и глубины пространства:

Нейрофизиологические: диспаратность зрения, **тремор** (движение глаз с частой 150 Гц и малой амплитудой), аккомодация хрусталика, конвергенция и дивергенция осей глаз, полимодальность зрения и восприятия;

Свойства пространства: параллакс движения, перспектива, тени, цветовые эффекты атмосферы (прозрачность).

Стоит отметить, что компенсаторная функция зрительного восприятия человека очень развита. Она напрямую связана с выживаемостью вида. Стереозрение полностью компенсируется другими способами объемного зрения, среди которых наиболее важную роль играет тремор глаз. Как уже отмечалось, компенсаторная функция может рассматриваться как взвешенная сумма. Кибернетическая модель стереозрения достаточно простая

– разложение в ряд (Фурье, Эйлер, Лапласиан) по двум базисным функциям, в частности по двум (криволинейным) координатам.

Считается доказанным, что во время восприятия цвета работает разностная схема по двум временным точкам, поскольку воспринимается не длина волны, а разность (изменение длины волны). Естественно предположить, что кибернетическая модель тремора глаз также разностная схема. Разностная схема типа крест, с начала выполняется дифференцирование (отображение тернарных треморных лучей в криволинейные координаты колечка тела), а затем обратное интегрирование криволинейных координат в мозолистое тело, соединяющее два полушария головного мозга. Косвенным подтверждением того, что работает разностная схема по трем временным точкам, являются эмпирические данные: частота тремора глаз 150 Гц, а время задержки восприятия одна пятидесятая секунды, следовательно, задержка составляет три такта тремора. Оценка производительности треморной модели зрительного восприятия: количество нейронов глаза 150000000 надо умножить на частоту 150 и результат возвести в квадрат (сложность вычислительного метода N^2), получается примерно 10^{20} . Для сравнения порядок экзафлопных вычислений 10^{18} . Если мы не боги, то производительность параллельных вычислений ограничена сверху зрительным восприятием.

Поскольку производительность зрительного восприятия чрезмерна, имеет право на существование гипотеза о том, что человеческое мышление является артефактом зрительного восприятия с точки зрения искусственных нейронных сетей. Слух и осязание являются врожденными чувствами, а зрение приобретенным. Зрение формируется в процессе обучения нейронной сети, повторяя механизмы этих органов чувств: стереозрению соответствует диспаратность слуха и осязания (левша - правша), тремору зрения - тремор рук. После обучения в нейронной сети осталось свободное место, природа не терпит пустоты и свободное место заняло мышление. Кроме физиологических оснований существуют косвенные психологические подтверждения этой гипотезы, например, решая трудную задачу, люди склонны отключаться от восприятия, тем самым освобождая дополнительное пространство. Данная гипотеза рассматривается всего лишь как модель, в любом случае полимодальность зрения и других органов чувств существенна при изучении зрительного восприятия.

Существуют работы, сопоставляющие стереозрение и свойства пространства: метафора восхода солнца оценивает влияние изменения освещенности Рис. 3. [4], в работе [8] рассматривается модель обработки стереоизображений, которая сначала определяет глубину пространства, полученную из параллакса движения, а затем корректируют бинокулярный дисбаланс.

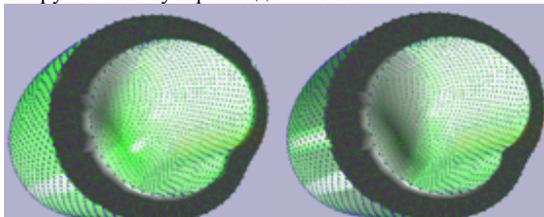


Рис. 3. Вращение источника света вокруг границы сетки.

По ряду причин не все перечисленные направления изучения зрительного восприятия были рассмотрены. В частности, потому что они частично пересекаются между собой, прежде всего на уровне моделей. Например, наивный байесовский подход применяется как в тестирование, так и в разработке тренажеров [1].

4. ЗАКЛЮЧЕНИЕ

Изучение зрительного восприятия является актуальной задачей для компьютерной визуализации. В данной работе предлагается структурирование материалов по тематическим направлениям (узконаправленным видам деятельности) и моделям формализации. Рассмотрение моделей восприятия-понимания является частным случаем верификации (математического доказательства правильности) визуализации. С целью валидации моделей в секторе компьютерной визуализации ИММ УрО РАН существует намерение начать разработку он-лайн сервисов параллельных вычислений. Например, на базе конструктора веб-визуализации реализовать автономную оценку эффективности для семейств алгоритмов с внешней памятью с позиции дискретной оптимизации (транспортная задача).

5. ЛИТЕРАТУРА

- [1] Алешин В.П. Технология виртуальной 3D среды в обратных задачах анализа визуального восприятия и интерпретации изображений // Труды 26-ой Международной конференции по компьютерной графике и зрению 19-23 сентября 2016, Нижний Новгород, с. 9-13
- [2] Бондарев А.Е., Галактионов В.А., Шапиро Л.З. Построение и визуальный анализ параллельных решений для оптимизационных и параметрических исследований в вычислительной газовой динамике // Труды 26-ой Международной конференции по компьютерной графике и зрению 19-23 сентября 2016, Нижний Новгород, с.211-215.
- [3] Манаков Д., Авербух В. Верификация визуализации // Научная визуализация 2016. Кв.1. Том 8. N: 1. Стр. 58 – 94.
- [4] Д. В. Манаков, В. Л. Авербух, П. А. Васёв. Визуальный текст как истинностное подмножество универсального пространства // Научная визуализация 2016. Кв.4. Том 8. N: 1. Стр. 58 - 94.
- [5] Д.Д. Попов, И.Е. Мильман, В.В. Пилюгин, А.А. Пасько. Решение задачи анализа многомерных динамических данных методом визуализации // Научная визуализация 2016. Кв.1. Том 8. N: 1. Стр. 55 – 57.
- [6] Е.А. Трофимов Эргономика зрительного восприятия. М. ООО "АИР" 2013.
- [7] Brodliet, K.; Osorio, R.A.; Lopes, A. A review of uncertainty in data visualization. In *Expanding the Frontiers of Visual Analytics and Visualization*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 81–109.
- [8] Petr Kellnhofer, Piotr Didyk, Tobias Ritschel, Belen Masia, Karol Myszkowski, Hans-Peter Seidel Motion Parallax in Stereo 3D: Model and Applications ACM Transactions on Graphics 35(6) (Proc. SIGGRAPH Asia 2016, The Venetian Macao, Macao).
- [9] Kirby R., Silva C. The need for verifiable visualization // IEEE Computer Graphics and Applications, 28(5) Sep 2008. Pp. 78–83.
- [10] T. Zuk and S. Carpendale. Theoretical analysis of uncertainty visualizations. In *Visualization and Data Analysis*, 2006.

Об авторах

Манаков Дмитрий Валерьянович – ведущий программист ИММ УрО РАН. Его e-mail manakov@imm.uran.ru.