

ПЕРЕГНАТЬ, НЕ ДОГОНЯЯ

В последнее время на рынке появилось множество мониторов, дающих иллюзию трехмерности изображения. Большинство из них используются в паре с очками той или иной конструкции.

В последнее время на рынке появилось множество мониторов, дающих иллюзию трехмерности изображения. Большинство из них используются в паре с очками той или иной конструкции. Однако еще в 2000 году в России был создан дисплей, не требующий никаких дополнительных средств и дающий неплохое качество объемного изображения. Мы встретились и побеседовали с его создателем, старшим научным сотрудником факультета вычислительной математики и кибернетики Московского Государственного университета Андреем Лукьяницей.



Расскажите, пожалуйста, об устройстве вашего стереодисплея.

- Аппарат состоит из двух жидкокристаллических панелей, между которыми помещена специальная маска. Панели подключены к компьютеру, и на них выводятся изображения, рассчитанные с помощью искусственной нейросети. В результате в заданной области перед дисплеем формируется световое поле, которое соответствует трехмерному объекту, находящемуся за экраном. Идея гениально проста. "Гениальна" в том смысле, что проста. Для того чтобы появилось стереоизображение, необходимо, чтобы правый и левый глаза видели разные картинки. Как этого можно добиться? Обычно используют специальные очки: с красным и синим стеклами, поляризационные, shuttering [Очки, поочередно закрывающие правый и левый глаза. При этом на монитор выводится изображение, соответствующее открытому в данный момент глазу] и т. д. А здесь у вас просто есть два изображения, разнесенных по глубине. Когда вы на них смотрите, каждый глаз видит свою картинку (для левого и правого глаза суммируются разные пиксели, так как лучи идут под разными углами).

Это могли придумать несколько веков назад. На СеВIT 2000 мы ехали еще без дисплея. Уже была готовая численная модель, и мы показывали макет: рассчитали картинку, распечатали их на струйном принтере и закрепили канцелярскими зажимами к разнесенным по глубине стеклам. Если рассматривать такую конструкцию с определенного расстояния на свет, то будет видна объемная картинка. Есть довольно смешные фотографии, как наши высокопоставленные чиновники разглядывали такой самодельный "дисплей".

Вы говорили, что между панелями установлена маска. Зачем она нужна?

- Маска - это диффузор. Она необходима для того, чтобы бороться с муаром. Муар-эффект наблюдается при наложении любых регулярных структур: заборов, сетки рабицы и т. п. Его можно наблюдать и дома: например, если посмотреть на тюль, сложенный вдвое.

На LCD-панелях есть различимая черная решетка: на экран наклеена непрозрачная пленка, в которой вырезаны маленькие отверстия по форме пикселей, а промежутки образуют сетку. И когда мы устанавливаем два экрана друг за другом, эти сетки накладываются и появляется муар. (Кстати, сейчас многие производители LCD-панелей создают трудности для нашей технологии: ради улучшения контрастности они увеличивают ширину черной маски между пикселями.) Поэтому мы и устанавливаем диффузор - это такая тонкая пленка, которая слегка размывает изображение с заднего экрана и тем самым убирает тонкие полоски черной межпиксельной решетки. В результате остается только одна сетка (на переднем экране), а значит, решается проблема муара. Кроме того, диффузор значительно увеличивает угол обзора.

А как вы рассчитываете изображения для экранов?

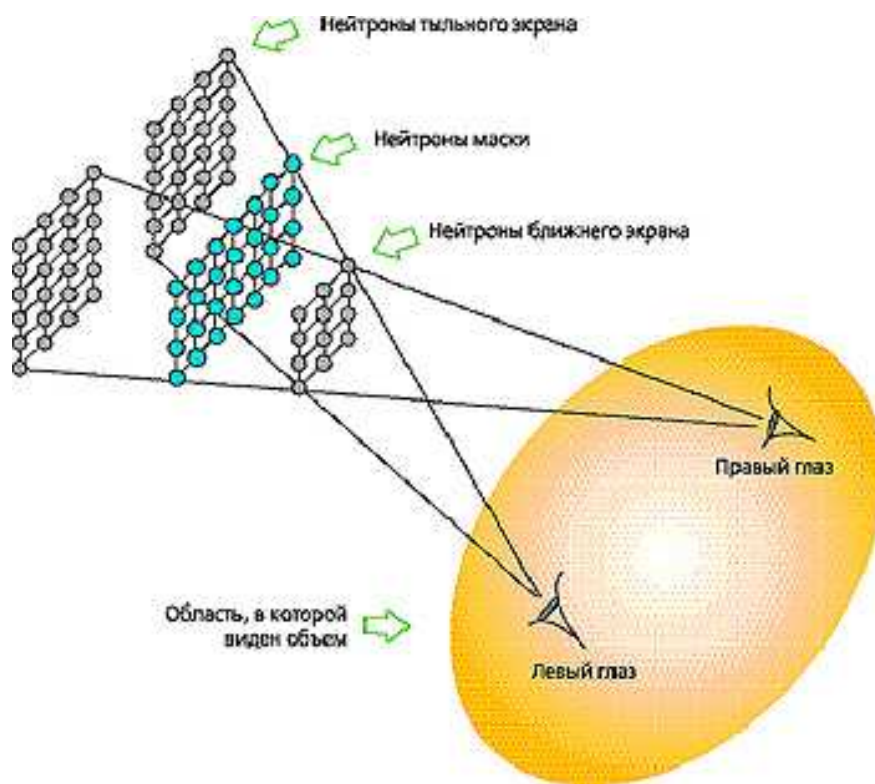
- Для этого решается задача, аналогичная компьютерной томографии. У нас есть стереопара - то есть два изображения, соответствующие тому, что должен видеть каждый глаз. И надо так

подобрать картинки для переднего и заднего экранов, чтобы при проецировании они складывались в заданную стереопару. Решение этой задачи - сложная математическая проблема [Здесь возникает так называемое уравнение Фредгольма первого рода. Эта задача некорректна, а ее решение неустойчиво]. Поэтому было решено использовать нейронные сети. Каждому пикселу ставится в соответствие по нейрону, а связи и весовые коэффициенты вычисляются из соображения геометрии прохождения лучей. Нейроны связаны тогда и только тогда, когда через соответствующие им пикселы проходит луч, попадающий в левый или правый глаз. Коэффициенты же подбираются в соответствии с характеристиками используемых LCD-панелей и диффузора.

Такая система позволяет быстро и эффективно вычислять проекции (фактически на каждый пиксель требуется лишь несколько простейших арифметических операций). Теперь, прогоняя заданные картинки через построенную нейронную сеть и сравнивая получаемые изображения с заданной стереопарой, мы при помощи метода последовательных приближений можем с достаточной точностью построить искомые изображения.

И сколько длится просчет одного кадра?

- Сначала алгоритм был реализован на CPU, и расчет одного кадра занимал не меньше двух минут. На современных процессорах все равно будет порядка одной минуты. В любом случае, для приложений реального времени это слишком много.



Лет шесть назад мой коллега, старший научный сотрудник Института высокопроизводительных вычислительных систем РАН Евгений Епихин заметил, что графическая карта имеет более мощный процессор, чем популярный тогда Pentium. В результате мы стали думать, как использовать GPU для нашей задачи. Оказалось, что при использовании нейросети задача идеально ложится на графический процессор. Довольно быстро мы смогли добиться гораздо более высокой производительности системы (около пятнадцати кадров в секунду).

Подождите, но ведь 2000 год - это максимум второй GeForce. А значит, ни о каких шейдерах и речи не шло?

- Конечно, пиксельных шейдеров тогда еще не было. Но нам удалось обойтись и без них. Мы воспользовались структурой предложенной системы, которая позволила представить трехмерный объект в виде суперпозиции двумерных изображений. А для работы с плоскими

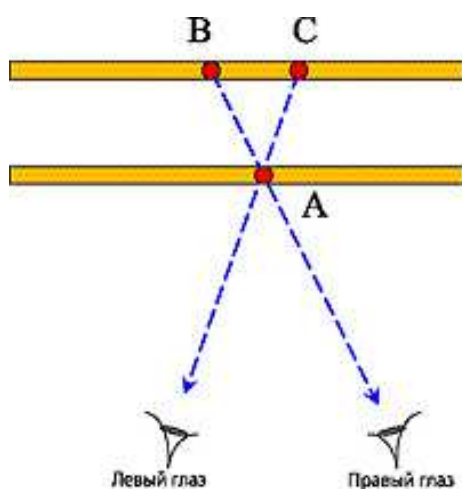
объектами нам хватило стандартных операций с текстурами. Таким образом, мы "подручными средствами" смоделировали действие нашей нейронной сети. При этом получается огромный выигрыш по скорости, поскольку все операции выполняются на GPU, который именно под эти операции и оптимизирован.

Получается, что при таком подходе не учитывается влияние установленной между экранами размывающей маски?

- Да, действительно, до появления пиксельных шейдеров смоделировать действие маски в реальном времени не удавалось. Мы даже пытались заказать специальный DSP-процессор для решения этой проблемы, но по ряду причин и он не подошел. А на современных карточках нам удалось смоделировать более сложную нейронную сеть, учитывающую воздействие маски, и добиться производительности в 30 fps.

А есть реальные устройства на основе вашей технологии? Где они применяются?

- Да, есть. Я тогда работал в фирме NeurOK[www.neurok.ru], и было выпущено несколько готовых дисплеев, которые демонстрировались на крупнейших международных выставках, таких как Comdex и Infocomm. Насколько мне известно, было изготовлено несколько игровых автоматов для американского казино, кажется MGM. Нам удалось скооперироваться с одной фирмой, которая производила специальные установки на основе больших параболических зеркал. Если в такую оптическую систему поместить любой монитор, то изображение трансформируется, в результате чего получается иллюзия объема. Ну а когда туда поставили наши дисплеи, объем увеличился многократно.



А почему не удалось дальше раскрыться?

- Просто рынок пока не готов. Ну, рынка как такового нет. Однажды на Comdex'е к нашему стенду подошел пожилой японец, долго стоял и смотрел, как мы общаемся с посетителями. Потом представился. Оказалось, что он в Sony руководил группой, которая продвигала цветной телевизор. Он сказал, что столкнулся с теми же проблемами, что и мы. Качество первых цветных телевизоров было низким, и все спрашивали: а зачем это нужно? Ведь на черно-белом телевизоре все и так прекрасно видно! Прошло время. И где вы сейчас найдете черно-белый телевизор?

Нам потенциальные производители и консьюмеры говорили, что все хорошо, только глубина изображения мала. То есть хочется, чтобы изображение прямо на тебя высывалось. Но ведь нигде в жизни такого объема нет! Большой объем мы видим только тогда, когда объект приближается к нам вплотную, а при этом мы сразу начинаем нервничать.

Я сотрудничаю с оптиками из МИИГАиК[Московский Государственный университет геодезии и картографии, www.miiigaik.ru], фотограмметрия - это их хлеб, а там стереоизображение используется для восстановления трехмерной структуры. Для обработки снимков специалисты обычно используют стереосистему со специальными очками. Так вот, они мне говорили, что относительно небольшой процент людей может долго работать с такой системой. То есть человек отработал день-второй-третий - и все, жалуется на постоянные головные боли. Мозг просто не может этого выносить, и как говорят, "глаза ломаются". Страшное напряжение, зачем это нужно?

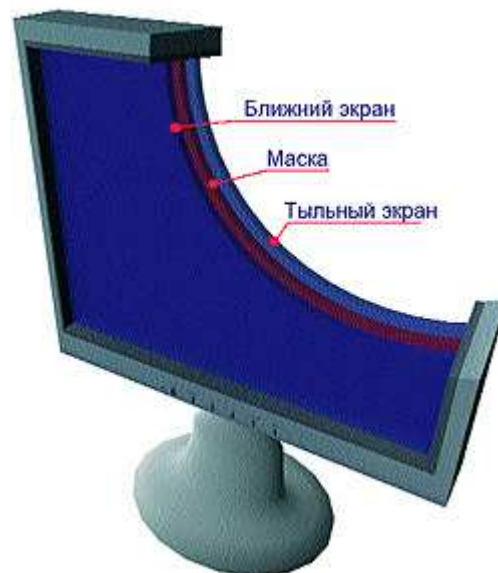
А объясняется все очень просто: в реальной жизни люди привыкли к небольшому параллаксу. Я знакомился с медицинскими исследованиями - для комфортного восприятия угол схождения глаз не должен превышать двух градусов. Если больше, начинает болеть голова.

А насколько реалистичным получается изображение на вашей системе?

- Полный реализм. Глаз четко различает, что это объем. На выставках я обычно сначала делал один экран прозрачным, и видно - картинка плоская. Включаю второй - сразу появляется объемность. Но сказать, что изображение высовывается так, что хочется его потрогать, не могу - этого, естественно, нет. Хотя можно подобрать такие исходные изображения, что глубина будет казаться больше, но в среднем - так, как мы в жизни и видим.

Насколько я знаю, кроме вашей есть еще технологии 3D-дисплеев, не требующие очков.

- Да, сейчас много разных систем выпускается. Например, Sharp производит панели с микролинзовым растром. Но они так устроены, что чуть голову в сторону сдвинешь - нет объема. Еще чуть в сторону - у вас изображение для левого глаза поменялось местами с изображением для правого. Область, в которой видно стерео, очень узкая. Причем настолько, что надо сидеть не шелохнувшись. А некоторые фирмы даже выпускают устройства для слежения за глазами зрителя. Когда наблюдатель меняет положение, то либо экран поворачивается, либо на экран выводится подстроенное изображение.



А у нашего нейродисплея допустимая зона достаточно большая. На выставках по два-три человека одновременно смотрят. Кто по центру, тот вообще классно видит, кто сбоку - чуть похуже (у границы экрана заметно, что есть два отдельных изображения). Отошел подальше - вообще видно прекрасно.

А у вас были еще какие-нибудь проекты с использованием GPU? Нейросетей на GPU?

- Были еще две задачи, связанные с 3D. Во-первых, проекционная система, в которой объем также был виден без очков. В ней использовались четыре проектора, находившиеся на расстоянии нескольких метров от специального экрана. Поскольку проекторы были разнесены по горизонтали на десятки сантиметров, при сведении в одну область экрана возникали трапециевидные искажения. Для их компенсации был разработан специальный алгоритм. На каждый проектор по очереди выводился тестовый прямоугольник, фиксированный камерой. Изображение анализировалось программой, определявшей координаты углов прямоугольника, а затем рассчитывались координаты плоскости, на которую нужно спроецировать изображение для компенсации искажений. После этого на проекторы выводились результаты наложения исходных изображений на соответствующие плоскости, полученные с помощью DirectX.

Вторая задача была связана с преобразованием обычных фильмов в объемные. В мире существует несколько компаний, которые могут это делать, но только с помощью оператора-человека. Нам удалось разработать алгоритмы, позволяющие с неплохим качеством преобразовывать 2D в 3D в реальном времени. Здесь тоже не обошлось без GPU и DirectX.

Сейчас мы разрабатываем систему для распознавания лиц. Как раз при помощи нейронных сетей на графических процессорах. От нее требуется очень высокая производительность. Аналогичные алгоритмы можно будет применять и для распознавания отпечатков пальцев, сетчатки глаза и т. п. Но это пока на стадии проектирования.

В прошлом году у меня на ВМиК была дипломница, Маша Карасева, мы с ней решали задачу восстановления структуры плазмы в установках ТОКАМАК по видеоизображению. Она, правда, использовала OpenGL, и скорость там была, конечно, не такая высокая, как в DirectX, но основные принципы те же самые.

Задача такова: по снимку плазмы, имеющей форму тора, нужно установить ее внутреннюю структуру. Задачи такого типа называются обратными: нам известен результат, и нужно узнать, при каких условиях он получен. Прямая задача проста: каждый слой плазмы светится по-разному, и если нам известно сечение плазмы, можно построить набор вложенных тороидальных поверхностей с заданной прозрачностью и потом "нарисовать" полученный объект, например с помощью OpenGL. Нас же интересовала обратная задача - получение сечения плазмы. Для ее решения сечение в начальный момент формировалось случайным образом, и по нему решалась прямая задача - получение "фотографии" плазмы. Построенное изображение сравнивалось с реальным, вычислялись отличия, сечение корректировалось, и так далее, до достижения необходимой точности.

И что смешно, мы несколько раз докладывали об этой работе, и многие не могли понять, в чем особенность предложенного алгоритма. Понимаете, это ведь очень необычно: получается, что обратная задача томографии решается прямо на видеокarte.

Кстати, я считаю, что вообще многие задачи, родственные томографии, естественным образом ложатся на карту. На кафедре математической физики факультета ВМиК группа профессора А. В. Гончарского занималась проблемами вычислительной диагностики, в том числе и контролем печатных плат. Платы уже давно стали многослойными, и если появлялся брак во внутренних слоях, его невозможно было увидеть. Платы просвечивали под разными углами и решали задачу томографии: восстанавливали структуру каждого слоя. И дальше уже смотрели, какова толщина дорожек, конфигурация и пр. Расчеты проводили еще на БЭСМ-6, и алгоритм работал часами. А на GPU можно делать контроль практически в реальном времени.

Кстати, мы с коллегами разработали венчурный проект, осуществление которого может помочь обогнать зарубежных производителей, не догоняя их. Речь идет об объемном телевизоре. Дело в том, что сейчас практически нет стереоконтента, а уж телевизионное стереовещание - дело будущего. Наш стереотелевизор сможет принимать обычные программы, преобразовывать их в реальном времени в объемные, проводить необходимые расчеты и выводить изображения на стереоэкран. Прimitивный прототип такой системы был создан нами пару лет назад на обычном компьютере, в который вставили карту TV-тюнера и подключили к нейродисплею. Видеокarta вполне справлялась и с преобразованием 2D в 3D, и с расчетами изображений для нейродисплея. Дело осталось за малым - доработать алгоритмы, зашить их в чипы, а главное - найти венчурного капиталиста, кто бы все это финансировал...

Опубликовано в журнале "Компьютерра" №13 от 05 апреля 2006 года