

Популярная Механика

Как работает нейронная сеть: Deep Dream

Деревья, разросшиеся многоцветными пагодами, собаки и овцы в облаках, — а еще глаза, глаза и морды зверей, виданных и невиданных. Новостные ленты и социальные сети переполнены странными картинками, напоминающими творения то ли гения, то ли безумца, — странными фантазиями.

Роман Фишман 28 октября 2015  14099



Вскоре после того, как код Deep Dream оказался в открытом доступе, энтузиасты добавили к нему веб-интерфейс и позволили всем желающим загружать свои картинки для обработки. Уже через пару дней увлечение приобрело характер опасной пандемии.

Как успели выяснить уже все, такие картинки являются побочным продуктом работы искусственной нейронной сети, которую в Google используют для анализа и классификации громадного массива графических файлов, найденных поисковиком в интернете. Это не единственный пример: нейронные сети помогают формировать прогнозы погоды и биржевых котировок, их используют и Facebook, и «Яндекс». Поиск «Яндекса» по картинкам, распознавание и даже перевод текста на них, возможность управлять приложениями с помощью голосовых команд — все это обеспечивают искусственные нейронные сети. Однако даже сами создатели таких сетей не могут во всех деталях объяснить, как они работают: именно в этом вся их суть. Но чтобы понять это, нам понадобится сделать небольшой шаг назад —

К мозгу

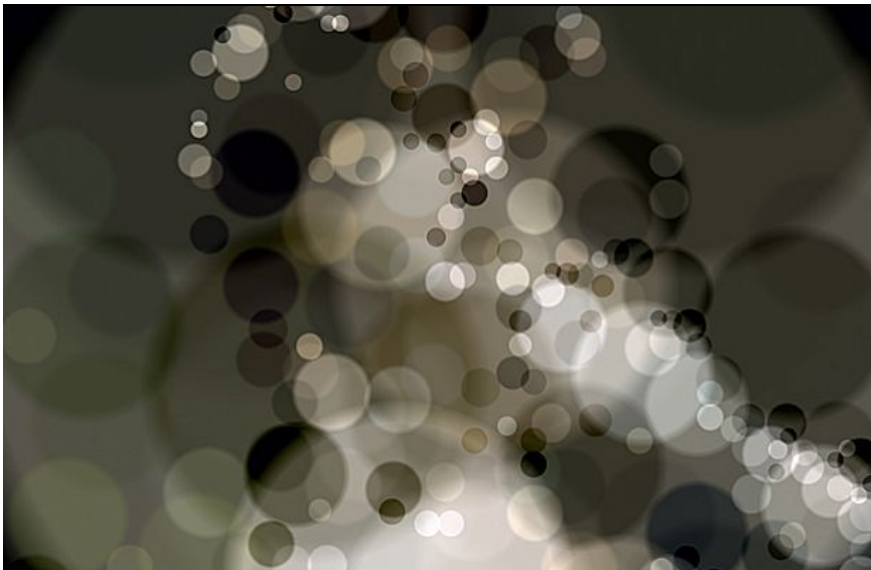
Человеческий глаз может показаться верхом совершенства. Недаром поколения креационистов приводят его в качестве примера такой сложной системы, которая якобы ни за что не могла бы появиться без какого-то изначального плана. Но если бы они увидели ту грубую, необработанную картинку, которую в действительности поставляют глаза, их ждало бы разочарование.

По-настоящему острое цветное зрение охватывает лишь небольшой конус с углом около пятнадцати градусов, окруженный смутными, размытыми, неразличимыми пятнами. К тому же картинка оказывается перевернутой и изрядно подпорченной довольно обширным слепым пятном. Но именно из этого грубого полуфабриката возникает весь объемный, наполненный образами видимый мир, в котором мы живем.

Отвечает за это превращение наш главный орган зрения — головной мозг. Глаза могут видеть что угодно, важно то, что мы способны воссоздать и понять из этой картинки. Выделить на пейзаже пастуха и его собаку, оценить его

возраст и ее размеры на фоне окружающих предметов... Вокруг решения таких сложных задач построены многие функции мозга и даже его анатомия.

Первичная обработка оптических сигналов происходит уже в зрительном нерве, после чего информация попадает в мозг и «аналитические отделы» зрительных центров. Это вовсе не строгое и последовательное движение, при котором шаг за шагом на картинке выделяются все более сложные детали: зрительные центры мозга не имеют четкой линейной иерархии. Они образуют ошеломляюще сложную распределенную сеть, сигналы по которой распространяются параллельно многими путями сразу.



Каждый узел этой сети прямо или косвенно влияет на все остальные, причем их конкретные роли остаются до конца неустановленными. Неизвестно даже точное количество центров, вовлеченных в обработку зрительной информации у человека, хотя обычно говорят о «более чем тридцати», между которыми существует порядка 200 линий прямой и обратной связи.

Разумеется, одни из них можно назвать низшими, ответственными за распознавание простейших особенностей картинки — скажем, форм или движения. Другие, высшие зрительные центры, больше заняты поиском образов и лиц. Однако это разделение довольно условно — достаточно сказать, что информация движется между зрительными центрами не только «снизу вверх», но и «сверху вниз».

Иногда это заставляет нас обманываться зрительными иллюзиями: мозг дополняет «низшие» элементы картинки несуществующими «высшими». В результате мы можем совершенно отчетливо увидеть лицо в смесителе с парой вентилей или обратить внимание на облако в форме вертолета. Похожая картина откроется нам и глубоко внутри программного «мозга» нейронной сети, распознающей изображения. Во всяком случае, у нее те же самые...



Способы мышления

В отличие от мозга, компьютер построен вокруг центрального процессора и математических операций, которые тот выполняет. И как мы неспособны симитировать его бешеную скорость вычислений, так и компьютер плохо приспособлен для выполнения задач, в которых столь замечательно демонстрирует себя мозг. Понимание живой

речи, распознавание образов, обобщения — для этого уже больше полувека и разрабатывается теория и практика нейронных сетей.

Несмотря на свое название, нейронная сеть может пытаться, а может и не пытаться в какой-то степени повторить структуры головного мозга. Однако ее элементарная ячейка, перцептрон, появилась именно как попытка воспроизвести процессы восприятия и обработки мозгом сенсорной информации.

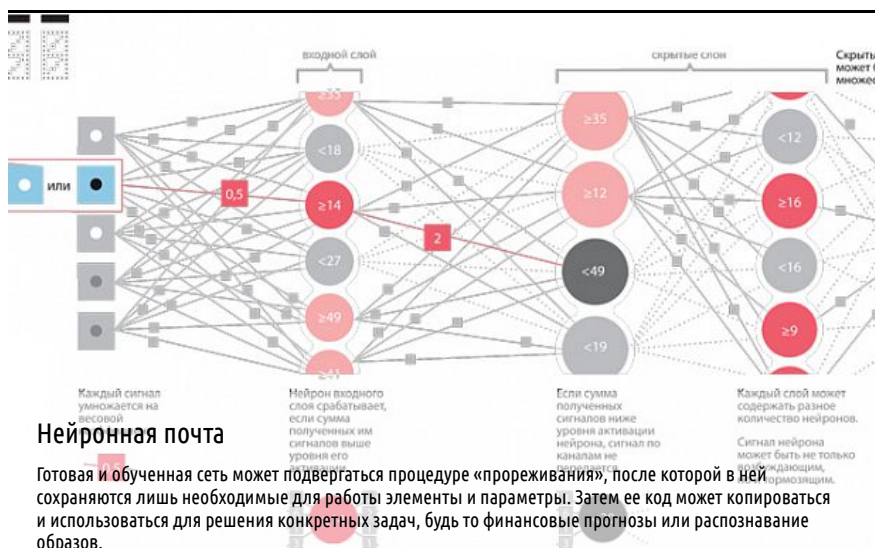
Лишь в редких случаях такую модель воплощают в «железе», в микросхемах, которые имитируют нейроны и связывающие их синапсы. Чаще она создается в виде компьютерных алгоритмов, с помощью которых человеческое мышление пытаются адаптировать под особенности устройства компьютеров. В своем классическом виде перцептрон состоит из входного и выходного слоев нейронов, между которыми находятся один или несколько «скрытых» слоев.



Каждый нейрон получает возбуждающий или тормозящий сигнал от нейронов предыдущего слоя (а первый слой — от входных «рецепторов»). Каждый сигнал умножается на определенный весовой коэффициент, усиливаясь или ослабевая. Нейрон лишь суммирует их, и если в итоге сумма этих взвешенных сигналов превышает определенную пороговую величину, нейрон срабатывает и передает собственный сигнал нейронам следующего слоя.

Такая организация в некоторой степени воспроизводит нейронную структуру «микроколонки», минимальной структурно-функциональной единицы коры головного мозга. Как правило, каждая микроколонка неокортекса содержит шесть слоев и около сотни нейронов.

Однако общее число микроколонок в нашем мозге исчисляется десятками миллионов — запредельное количество для обсчета даже на самых современных компьютерах. Простейшая искусственная нейронная сеть может состоять всего из одного перцептрона. Ведь даже в такой элементарной форме она способна к самосовершенствованию в ходе того, что называется...



Машинное обучение

Никто не рождается суперменом в деле распознавания образов: как и ходьбе, этому надо учиться. Мозг получает зрительный сигнал и добрый совет от родителей: это поле, это пастух, а вот его собака... В следующий раз — «Где здесь собака?» — правильный ответ будет означать похвалу и награду. Так, раз за разом, сотни тысяч раз видя разные предметы с разных сторон, мы совершенствуем этот важный навык, узнаем все новые образы и учимся их узнавать.

Так действует и нейронная сеть, постоянно пересматривая свою работу и адаптируясь для того, чтобы узнавать тот или иной образ. Весовые коэффициенты, которые определяют значимость сигналов, получаемых нейроном от предыдущего слоя, поначалу задаются случайным образом. Однако после того, как система выдала первый — конечно, неверный — результат, она начинает их изменять и уточнять. Обработывая тысячи снимков, нейронная сеть каждый раз вносит изменения в свою работу, пока не начнет стабильно получать правильный результат.

Многое тут зависит от учебника — точнее говоря, от набора обучающих изображений. «На самом деле в интернете картинок не так много, — говорит глава службы функциональности поиска «Яндекса» Андрей Плахов, — особенно в сравнении с тем колоссальным количеством, которые мы видим в течение жизни». Это создает опасность, знакомую разработчикам нейронных сетей: обучение может превратиться в зубрежку.

Если выбрана неполная или однообразная обучающая база, системе будет проще затвердить правильные ответы на все изображения, чем учиться распознавать на них сложные детали. Такую сеть называют «переобученной» (а по-английски «избалованной», *grandmothered*): она великолепно справляется со знакомыми изображениями, но не даст ничего, если ей предложить новую, незнакомую картинку.

Поэтому для обучения нейронных сетей распознаванию картинок, как правило, используют стандартизованные базы. Самой популярной и обширной из них можно назвать ImageNet, поддерживаемую Стэнфордским университетом. База включает больше 14 млн картинок, упорядоченных в 27 категорий верхнего уровня иерархии — от «инструмент» до «гриб» и от «геологическое образование» до «овощ» — и 27 000 категорий всех уровней. На этой библиотеке учится большинство подобных нейронных сетей, и ее случайные особенности могут во многом определить их «способ мышления».

Возможно, именно особенностями обучающей базы объясняется склонность Deep Dream находить повсюду именно животных и дальневосточные пагоды: нейронная сеть обучалась на изображениях, где их было особенно много. Впрочем, так ли это, сказать трудно, ведь даже разработчики толком не понимают...

Что у нее внутри

Современная нейронная сеть может включать десятки, а то и сотни слоев-перцептронов с множеством нейронов в каждом. Как и зрительные центры мозга, они организованы в сложную сеть взаимосвязей — и, как в мозге, эта система пластична. Сконструировав и обучив нейронную сеть, разработчик остается в полном неведении о том, что именно в итоге в ней происходит, как она меняется в ходе обучения, на какие особенности изображений начинает обращать внимание, а какие отбрасывает.

«Биологические аналогии не должны сбивать с толку: нейронные сети — продукт математики и программирования, — говорит Андрей Плахов. — С этой точки зрения нейронную сеть можно представить как невероятно сложную функцию с неизвестной формулой. Запись «формулы» Deep Dream заняла бы, наверное, не один десяток томов. Но в целом мы знаем лишь данные, которые она получает на входе, и результат, который выдает на выходе. Какие именно операции совершаются в процессе, сказать зачастую невозможно».

Видимо, так появился и Deep Dream — из чистого любопытства к тому, как именно работают нейронные сети, ведущие распознавание и классификацию изображений. Самый очевидный способ увидеть результаты работы каждого слоя-перцептрона — заставить его визуализировать эти результаты. Нам, людям, вообще удобнее мыслить картинками, поэтому программисты Google доработали нейронную сеть так, чтобы разные слои, участвующие в распознавании, усиливали на изображении именно те особенности, которые они сами заметили и выделили.

Как и в случае мозга, можно заметить, что одни слои нейронной сети берут на себя задачи ближе к «нижнему» уровню восприятия — такие как выделение границ фигур на картинке, а другие работают с более сложными образами. Это проявляется и в особенностях «художественного стиля», в котором «творят» разные варианты Deep Dream.

Однако и здесь не существует четкого разделения полномочий: все слои нейронной сети взаимосвязаны, как зрительные центры нашего мозга. Возможно, это и делает картинки такими завораживающими, такими похожими на...

Настоящее искусство

Почему картинки, которые создает Deep Dream, воспринимаются как что-то близкое к искусству? Специалисты дали на наш вопрос довольно неожиданный ответ. «Это может быть связано с современной парадигмой зрения, распространившейся с начала XIX века, — предполагает искусствовед Глеб Напреенко. — До того времени наблюдатель уподоблялся бестелесной точке, находящейся как бы за пределами реальности, которую он воспринимает, словно зритель, заглядывающий в камеру-обскуру. Лишь приблизительно двести лет назад субъект получил телесность, его зрение было помещено в глаз, он стал почти соучастником наблюдаемой сцены».

«Этот переход видно и по той роли, которую играет взгляд на живописном полотне: если прежде он скользил по нему, не рискуя натолкнуться на встречный интерес, то уже во французской живописи второй половины XIX века появляются образы, возвращающие взгляд зрителю, смотрящие на него в ответ, — добавляет Глеб Напреенко. — Эта парадигма снова актуализируется в «работах» Deep Dream. Его способность выявлять в любых изображениях рожицы и мордочки соответствует нашему внутреннему стремлению обнаруживать направленные на нас взгляды, таящиеся в вещах».

Сети сходят с ума



Deep Dream уже успел породить «городскую легенду» о том, что обработка алгоритмом нисколько не меняет картины Иеронима Босха с их фантазмагорическими существами. Это не так — можете проверить сами, зато действительно остаются почти теми же фракталоподобные работы английского художника Луиса Уэйна, написанные им в последние годы и после нескольких психозов.

В самом деле, бросается в глаза удивительное сходство изображений, которые порождает Deep Dream, с некоторыми образчиками творчества психически больных или людей, находящихся под действием «психоделических» наркотиков. Это не простое совпадение: среди множества воздействий, которые оказывают эти вещества на мозг, не последнюю роль играет нарушение скоординированной работы зрительных центров.

Статья «Искусство цифровых фантазий: почему нейронным сетям снятся электроовцы» опубликована в журнале «Популярная механика» (№155, сентябрь 2015).

Комментарии

[1 комментарий](#)

Кинотрюки XXI века: как они делаются

С приходом цифровых технологий для кинематографистов, кажется, нет ничего невозможного. Широкий экран легко и убедительно отменяет все физические законы. Но у «оцифровки» есть и обратная сторона. Зритель поначалу удивлялся, а потом и стал требовать все более и более зрелищного действия. Технологии визуализации совершенствуются каждый год, и оказалось, что ничто так быстро не теряет свою реалистичность, как компьютерные эффекты. Другое дело — феерические трюки, исполненные вживую, по-настоящему.



Качественная, полная риска работа каскадера вызывает восхищение независимо от того, сколько лет прошло с выхода картины на экран. Эпизод из фильма «Человек с золотым пистолетом», когда машина после разгона взмывает в воздух, делает оборот и приземляется на другом берегу реки, захватывает не меньше, чем 40 лет назад. И в современном кино лучший результат получается, когда компьютерные эффекты не заменяют, а разумно дополняют настоящие трюки, что с блеском продемонстрировали вышедшие в этом году «Безумный Макс: Дорога ярости» и «Форсаж-7».



Без компьютерной графики «Безумный Макс», конечно же, не обошелся. Она активно использовалась для «строительства» цитадели и оживления пейзажа, ведь съемки проходили на почти плоском, ярко освещенном африканским солнцем безжизненном пространстве.

Небесный форсаж

На разработку самого сумасшедшего эпизода в фильме «Форсаж-7» — сброса автомобилей с самолета — ушло более двух месяцев. Постановщикам потребовались два самолета Hercules C-130 и десяток камер, на которые снимали операторы с вертолетов и скайдайверы. Сброс производился на высоте 5 км. Когда автомобили опустились до 500 м, автоматически сработали парашюты. Разумеется, в автомобилях никто не сидел. Все крупные планы с актерами снимались в павильоне на гидравлических платформах, имитировавших покачивание в воздухе. Съемки приземления потребовали строительства сложной конструкции в горах. Один 300-тонный кран установили вверху выбранного участка дороги, другой внизу и натянули между ними толстый стальной трос, по которому передвигалась тележка с подвешенной на стропах машиной. Скорость снижения машины на парашюте — около 40 км/ч. Как только она разогналась до этой скорости на тросе, крепления строп отстегивались и Dodge, за рулем которого сидел каскадер, плюхался на землю с 2,5 м — при падении с большей высоты повреждений было бы не избежать. Единственное, что пришлось сделать на компьютере, — убрать из кадра тросы и подрисовать парашют.

Автомобиль, на котором будет выполняться трюк, всегда проходит подготовку. Как минимум вваривается каркас безопасности, как максимум дорабатывается двигатель, трансмиссия и подвеска. «Многое зависит от производственного бюджета, — рассказал нам Мартин Иванов, двукратный обладатель премии Taurus, считающейся каскадерским «Оскаром». — Если проект большой и дорогой, то автомобиль готовится за полгода в нескольких экземплярах. Так было на фильмах о Джеймсе Бонде. Для картины «Спектр» (выйдет на экраны в ноябре 2015 года) был построен эксклюзивный Aston Martin в восьми экземплярах. Машина участвует в двух

динамичных погонях — в Австрии и Италии. Здесь мне выпало дублировать главного злодея. Я сидел за рулем 700-сильного Jaguar C-X75, не запущенного в серийное производство. На одном из участков трассы удалось разогнаться до 180 км/ч. На такой скорости снимают с вертолета, но чаще используют полноприводные джипы на базе Porsche Cayenne или Mercedes ML, оборудованные краном Russian Arm».

Трюк выполняется с соблюдением определенного скоростного режима. На киноэкране автомобиль часто движется заметно быстрее за счет увеличения частоты кадров во время монтажа. Например, действие снимают со скоростью 22 кадра в секунду, а воспроизводят с частотой 24. Если внимательно смотреть «Форсажи», можно заметить неестественное ускорение автомобилей.



Dodge Charger, модель 1970 года ручной сборки. Из серийных деталей в нем использовали лишь крышу и элементы приборной панели, да и те пришлось подгонять при установке. Трубная рама этого «крепыша» усилена в лучших традициях внедорожников. 520-сильный мотор LS3 техники переместили в заднюю часть, чтобы приспособить Dodge к прыжкам.

Переворот по-новому

Мартин Иванов должен был работать и над картиной «Безумный Макс: Дорога ярости». Он провел месяц в Австралии, подбирая команду каскадеров-водителей. Но незадолго до старта съемок у продюсеров начались проблемы с финансированием, а когда работу возобновили, Мартин был уже занят в «Форсаже-6» и «Кванте милосердия». «Посмотрев «Безумного Макса», — признается каскадер, — я пожалел, что не смог в нем сняться. В этом фильме отличные трюки».

Для некоторых трюков пришлось придумывать новые методы. Эпизод, когда «Перехватчик» Макса, созданный на базе XB Ford Falcon Coupe 1973 года выпуска, переворачивается несколько раз, режиссер Джорж Миллер попросил сделать максимально эффектным, насколько это вообще возможно. Перевороты и взрывы — самые частые трюки в кино, и техника их выполнения отработана годами. Обычно автомобиль переворачивается при помощи трамплина или пневматической пушки, которая устанавливается вместо пассажирских кресел. По команде второго режиссера или постановщика трюков водитель жмет на кнопку, и пушка выстреливает. Сжатый воздух с силой выталкивает деревянную колодку, и машина из-за резкого толчка опрокидывается. В 1980-х подобные пушки были пиротехническими, поэтому при опрокидывании из-под колес шел дым. Но Гай Норрис, руководивший постановкой трюков (35 лет назад он был простым каскадером на съемках второго «Безумного Макса»), использовать проверенные методы не хотел: рампу трамплина сложно спрятать в пустынном пейзаже, а вылетающая из-под машины деревянная колодка могла попасть в другой автомобиль. Было найдено новое решение. Когда мчащийся на скорости почти в 100 км/ч автомобиль оказывался в нужном месте, водитель резко поворачивал руль и, нажав на кнопку, приводил в действие закрепленное под днищем устройство: плоская стальная пластина с силой ударялась в песок и тут же возвращалась на место, толчок отправлял машину в переворот. На тренировке автомобиль перевернулся восемь раз, но в фильм вошел дубль с пятикратным переворотом.

Лучшие автомобильные трюки

За более чем столетнюю историю кинематографа зрители увидели тысячи эффектных трюков с участием автомобилей. Вот наша пятерка лидеров.



5-е место

Фильм: «Темный рыцарь» Год: 2008 Режиссер: Кристофер Нолан Трюк: сцена, где опрокидывается грузовик, была снята с первого дубля с нескольких камер на реальной улице ночью.

Компьютерная безопасность

Переворачивался на «Перехватчике» сам Норрис. Он же сел за руль автомобиля и при съемках другого сложного и зрелищного эпизода. В одной из сцен 16-тонная колесная фура врезается на полном ходу в другой грузовик. Столкновение происходило на скорости 96 км/ч. Чтобы обеспечить безопасность каскадера, инженеры внесли конструкторские изменения в компоновку кабины. Они разместили ее сбоку автомобиля и установили на рельсах, оборудовав тормозным механизмом. Торможение и нагрузка в момент удара были распределены во времени. В итоге Норрис не получил ни синяка, ни царапины.

Всего для фильма было выполнено 303 трюка, 70 из них повышенной сложности и опасности. В самые загруженные съемочные дни количество каскадеров на площадке переваливало за полторы сотни. Но их лиц зритель не видит. Крупные планы ведущих актеров, яростно «гоняющих» на своих монстрах, снимались на неподвижных грузовиках или же на низкой скорости.

А как же компьютерные технологии? Безусловно, без них не обошлось. Во-первых, с приходом цифровой графики каскадеры получили возможность предварительно моделировать трюк. Трехмерная раскадровка (аниматик) позволяет увидеть желаемый результат, определиться с продолжительностью, ракурсами и крупностью съемки и точно все рассчитать. Во-вторых, «цифра» помогает сделать трюки безопаснее за счет страховочных тросов, которые потом легко стереть. В «Дороге ярости» больше всего компьютерной доработке подверглись сцены песчаной бури — они стали примером виртуозного сочетания реальных съемок и цифровой графики. Режиссер хотел взять за основу проезд настоящих машин, чтобы потом, даже если их все придется заменить цифровыми муляжами, в кадре было точно выдержано направление движения, взаимное расположение автомобилей и перемещения камеры. Машины, которым предстояло вспыхивать и подниматься в воздух, были смоделированы методом фотограмметрии.



Актеры, за редким исключением, не садятся за руль. Их дублирует опытный водитель-каскадер. Один из самых востребованных — Мартин Иванов, снявшийся во множестве известных и зрелищных картин. В «Кванте милосердия» он дублировал Дэниела Крейга.

Компьютерная симуляция огня использовалась во многих эпизодах, но самый зрелищный взрыв — в финальной погоне — настоящий. В момент взрыва на площадке был только бензовоз. Взрывное устройство сработало по радиосигналу, пневмопушки вытолкнули кабину в воздух, клубы огня снимали с разных точек несколько камер,

установленных на автомобилях и вертолете. Вторым «заходом» была снята та же сцена, но с другими машинами, участвующими в эпизоде, и без взрыва. К совмещенной картинке потом на передний план добавили Макса на шесте и чуток изменили место действия — погоня снималась в голой пустыне, а по задумке режиссера кадр должны были окаймлять отроги гор.

По признанию режиссера, у него в запасе есть еще парочка историй о Безумном Максе. А где Макс, там и потрясающие трюки. И, возможно, мы их увидим.

Статья «Настоящий драйв» опубликована в журнале «Популярная механика» ([№154, август 2015](#)).

Комментарии

Написать комментарий