

## **ВИДЕОДЕТЕКТОРЫ - ВЗГЛЯД ИЗНУТРИ (ЧАСТЬ ТРЕТЬЯ - ГРАНИ ИНТЕЛЛЕКТА)**

**М. Руцков**

О, интеллект - как много в этом звуке! Я верю друзья - нет, не караваны ракет помчат нас вперёд от звезды до звезды ..., а ещё немного и мы будем зачитываться рекламнотехническими опусами, примерно следующего содержания: “Интеллектуальный видеосигнал от интеллектуальной CCTV-камеры, через интеллектуальные провода поступает в интеллектуальный видеобластер ...”. Далее со всеми остановками - вплоть до интеллектуального монитора, которому уже ничего не остаётся делать, как только с умным видом смотреть в самого себя! Невольно в памяти всплывает “робот-зазнайка” (из одноимённого фантастического рассказа Генри Каттнера) - для целей самосозерцания у него из глазниц на рельсиках выкатывались две камерки и разворачивались друг другу навстречу, ну чем не “замкнутое телевидение”. Нет, мне совсем не хочется анализировать полёт рекламной мысли, которая, наверное, скоро докатится и до галактического интеллекта. А хочется чего-то земного, ну немного помечтать на тему: “Интеллектуальное видеонаблюдение от HomoSapiens - попытка взлома (в хорошем смысле этого слова)”. Сразу предупреждаю категорически, никто, кроме меня, ответственности за нижеизложенные мечты - не несёт!

### **ЗАГЛЯНЕМ В СЕБЯ**

Как же мы заблудились в трёх соснах: летаем в космос, ныряем в бездну океанских глубин, и совершенно забыли о том, что мы и есть обладатели самого великого творения Природы - мозга человеческого! Нам бы себя изучить, не отходя от собственного тела, и создать суперсистемы видеонаблюдения. А мы практически засорили лексикон словом “интеллектуальный”, разбрасывая его - и направо, и налево. Невольно возникает вопрос: а что это такое - интеллект? Смотрим энциклопедию: “интеллект (от лат. intellectus - разумение, понимание, постижение) - относительно устойчивая структура умственных способностей индивида”. Туманно как-то. Да простят меня философы - буду сам излагать, но только отталкиваясь от реальных фактов научных исследований в области нейрофизиологии и психологии.

Итак, вот что утверждает лауреат Нобелевской премии, нейрофизиолог Дэвид Хьюбел : “Головной мозг содержит  $10^{12}$  (миллион миллионов) клеток - по любым понятиям число астрономическое; хотя я не знаю, подсчитывал ли кто-нибудь число клеток в печени человека, было бы удивительно, если бы в ней было меньше клеток, чем в нашем мозгу. И все же никто никогда не скажет, что печень так же сложна, как мозг”. Очевидно, что речь идёт о клетках, обрабатывающих информацию - нейронах, а не о клетках-кирпичиках внутренних органов. Что и подтверждается дальнейшим изречением учёного: “Более веский аргумент в пользу сложности мозга можно усмотреть в соединениях между его клетками. Обычная нервная клетка (нейрон) получает эту информацию от сотен или тысяч других клеток и в свою очередь передаёт информацию сотням или тысячам нейронов”. Можно ли найти рукотворный аналог такой сложности? Нет, не создали пока.

Так как же работает вся эта армада нейронов? Опираясь на данные многочисленных исследований, в научных кругах сложилось мнение, что примерно только 5-10% нейронов

головного мозга задействовано на уровне нашего сознания, а вот вся оставшаяся, “подводная часть айсберга” - подсознание. Переводя с философского языка на технический, можно сказать, что сознание - это интерфейс между нашим “Я” и внешним миром. Интерфейс достаточно развитый: это логическое и образное мышление, речь, обработка сенсорной информации от органов чувств, управление всеми частями нашего тела и многое другое. Однако задачи, решаемые на уровне нашего сознания, достаточно просты и фактически не выходят за рамки конструкции - “if-goto”. Действительно, трудно себе представить что, расхаживая с умным видом по комнате (как в кинокомедии “Весна”) и повторяя сакраментальную фразу: “Масса солнца равна двум квинтиллионам тонн”, можно придумать новую теорию или хотя бы какой-нибудь алгоритм видеодетекции. Или весь день просидеть, размышляя на тему: “Если электрон так же неисчерпаем, как и атом, то исчерпаемы ли те штуковины, из которых он сделан?”.

Конечно, путём только логических рассуждений невозможно создавать шедевры в области науки, искусства, экономики, политики и естественно в сфере безопасности. Для этого нужен интеллект, который как раз и сидит в нашем подсознании. В течении всей жизни в него (в подсознание), как в бездонный колодец, валится информация из внешнего мира, раскладываясь “по полочкам” и превращаясь в индивидуальную базу знаний. Мало того, наше подсознание живёт своей жизнью, постоянно перерабатывая все накопленные данные. Как оно взаимодействует с сознанием - одному богу известно, мы можем только догадываться. В очень упрощённом виде это выглядит примерно так. Сознание являет собой маленького и далеко не умного начальника, который пытается все проблемы решать командно-административными мерами. И вот когда он заикливается на решении очередной задачи, выходящей за пределы её понимания, то следует знаменитое - что делать? Этот призыв, в конце концов, доходит до подсознания, а именно до подчинённых - очень умных, всёзнающих, талантливых, но абсолютно неуправляемых ребят, занятых самосовершенствованием. Те как бы нехотя отвечают: “Вот сейчас всё бросим, и пойдём помогать”. Затем они всё-таки берутся за дело, реализуя народную сказку: “Иди спать, утро вечера мудренее, завтра все твои желания исполнятся”. Кстати Менделеев свою таблицу во сне увидел. А как быть, если цейтнот и времени на сон уже нет? Вспоминаю гениального шахматиста, чемпиона мира - Михаила Таля, который поведал удивительную историю. В одной из партий за шахматную корону, он попал в абсолютно безнадёжную ситуацию, время истекало - решения не было. Вдруг неожиданно пришёл бегемот Корнея Ивановича Чуковского и в голове начала крутиться единственная фраза: “Ох, тяжёлая эта работа - из болота тащить бегемота”. Бегемот во весь мозг - полный паралич сознания. А потом бегемот ушёл, и руки сами собой стали передвигать фигуры. Это тот самый случай, когда цена вопроса была столь высока, что подсознание в аварийном режиме, не дожидаясь пока сознание пойдёт спать - просто заблокировало его бегемотом и без помех решило проблему.

А вообще есть много вариантов торможения сознания, с целью выхода из состояния ступора - у кого как. Например, изобретатель упомянутого “робота-зазнайки” для реализации своих идей применял пиво. Полностью поддерживаю, а от себя добавлю - очень помогает ещё и рок-музыка, Alice Cooreg в частности. Мне даже иногда кажется, что он в некотором роде соавтор моих алгоритмов. Всё это замечательно, одно лишь надо чётко понимать - если нет конкретных знаний, и в голове пусто, то не будет: ни озарений, ни прозрений, ни интуиции, да и вообще ничего не будет, сколько пива ни пей и рок-музыки, ни слушай. Поэтому, как завещал один известный мыслитель, надо - учиться, учиться и ещё раз учиться!

А мозг человеческий, как раз этим и занимается - самообучением. Ключевым свойством этого процесса, является ассоциативность. Информационное воздействие извне вызывает в мозге каскад ассоциаций с целью отыскания аналогий, как ответной реакции. Если ответ не стопроцентный, то отличительные признаки изменяют уже имеющуюся в

памяти информацию, либо в сторону её обобщения, либо - наоборот, в сторону разъединения на независимые аналогии, в соответствии с объективной реальностью, которая, кстати, тоже может заблуждаться. Выглядит это примерно так.



Рис.1. Молчит как рыба - об лёд.

На вопрос ведущего, закончить фразу: “Молчит, как ...”, он первым бабахнул по кнопке и отчеканил - “Партизан”, потом, испугавшись чего-то, тут же воскликнул: “Ой нет - рыба”, ну и наконец, сделав совершенно округлые глаза и недоумённо пожав плечами, машинально добавил - “Об лёд” (рис.1)! “Нет уж, об лёд она - бьётся, а не молчит ”, возразил ведущий. “Понял, теперь буду знать ”, согласился подавленный игрок. “А что? Если партизан действительно шарахнет её (рыбу) как следует об лёд, то может быть она и замолчит, т.е. перестанет биться... насмерть... за свободу... слова... - эх, всё правильно, ответ засчитывается”, подвёл итог ведущий. Вот так, они - мысли, и гуляют ассоциативно в наших головах, чаще ночью, создавая новые теории и алгоритмы! А теперь займёмся непосредственно зрением.

## ТАМ КТО-ТО СИДИТ И СМОТРИТ В ТЕЛЕВИЗОР

Думаю, нет смысла рассуждать о значимости нашего главного органа чувств - зрения. Это очевидно, но хотелось бы понять, как всё это работает. Именно такой вопрос был задан обыкновенной домохозяйке, на что и был получен соответствующий ответ: “Ну, наши глаза - это маленькие телекамеры, от которых изображение поступает прямо в голову. Так, хорошо, а там что? А там кто-то сидит и смотрит в телевизор! А кто???”.

Ну, тогда позвольте и мне высказаться. Конечно, никакого телевизора в голове у нас нет, да и сидеть там некому, кроме мозга, который: смотрит на мир (нашими широко раскрытыми глазами), анализирует эту видеoinформацию и принимает соответствующие решения. Очень условно можно выделить три базовые функции работы зрительного аппарата: бессознательная видеодетекция, обнаружение объектов и распознавание ситуаций.

Теперь по-порядку, почему именно бессознательная видеодетекция. Дело в том, что этот уровень обработки (а вернее предобработки) - вне нашего сознания. Всё происходит очень быстро и автоматически, помимо нашей воли. Главная задача видеодетекции - привлечение внимания. Думаю для большей ясности, надо немного рассказать об устройстве глаза, а именно о его видеосенсоре - сетчатке. Опустим физиологические подробности, главное - это геометрия распределения светочувствительных элементов, коими являются палочки (про колбочки, воспринимающие цвет, говорить не будем, поскольку их всего 6 миллионов, что в 20 раз меньше). Распределены они крайне неравномерно, причём наиболее густо представлены, в так называемой “центральной ямке”. Чтобы было понятно - перевод этой зоны высокого разрешения на рассматриваемый объект, равносильен “наезду” трансфокатора с 30-35-кратным увеличением. А так мы практически ничего и не видим, в чем легко убедиться (рис.2).

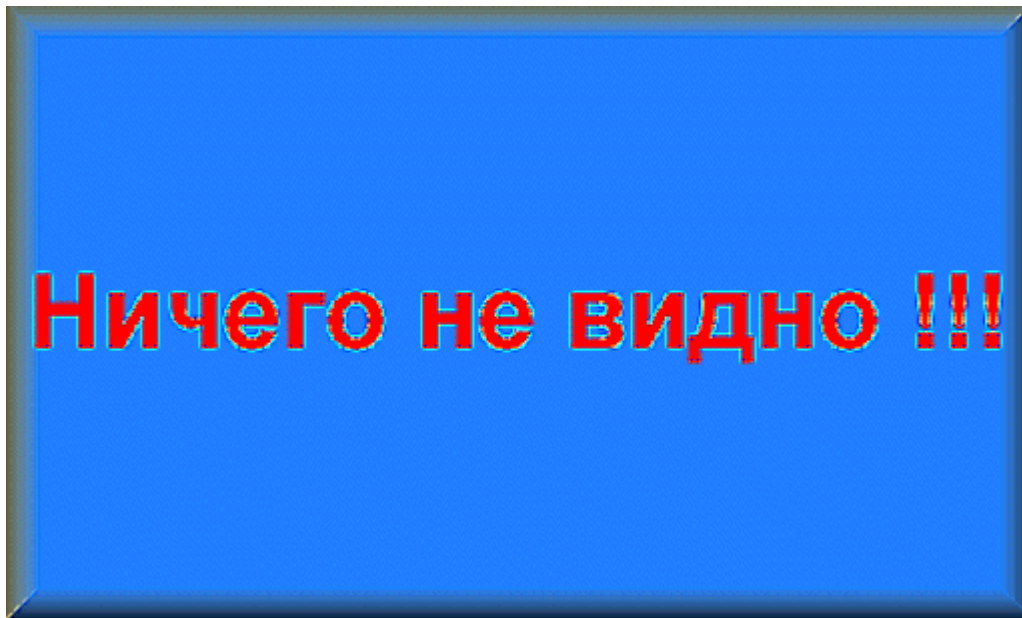


Рис.2. В кабинете окулиста.

Зафиксируйте взгляд на первой букве фразы - мало того, что совершенно невозможно разглядеть буквы в последнем слове, они не видны даже и в конце первого. Почему так Природа придумала - нет бы, сделать максимальное разрешение по всей сетчатке. Ну, тогда наш глаз был бы размером с арбуз, что согласитесь не очень удобно для целей выживания.

Возвращаемся к видеодетекции - а что собственно детектируется. Первым делом - движение, здесь имеем максимальный обзор, практически все 180 градусов. Глаза начинают автоматически поворачиваться в сторону возмущающего воздействия, причём, если этого недостаточно, то докручиваем головой. Далее переходим к исследованию объекта нашего внимания, уже на уровне сознания. Стоп, есть и другие, более быстрые варианты реакции, например такой - он нам достался в наследство от наших далёких предков, а именно: мы уворачиваемся (в целях безопасности) от любых летящих в нас предметов. Видеодетектор движения моментально приводит в действие моторную кору, которая управляет нашей двигательной системой. Знаете, как это часто бывает - сначала мы шарахаемся в сторону, а уже потом, вставая и отряхиваясь, недоумённо спрашиваем: “А что это было?”. Согласитесь - не лучший вариант, например, для хоккейного вратаря. Однако мозг обладает определённой пластичностью, которая позволяет натренировать его (в данном случае) для бессознательной ловли и отбивания шайб. Причём, никакого интеллектуального анализа не требуется - полный автоматизм, нет времени на размышления. Кстати, аналогичные действия могут происходить и на уровне нашего сознания. Например, вот как выглядит классическая оборонительная реакция вида Homo Sapiens - “Сам дурак!”. Не беда, с каждым бывает, можно поправить ситуацию - надо просто мило улыбнуться и добавить: “Но в хорошем же смысле этого слова!”.

В силу ограничений по разрешающей способности, наши глаза всё время находятся в движении, совершая, так называемые, саккады - резкие прыжки, наводящие “центральную ямку” на информативные участки анализируемой сцены (рис.3).

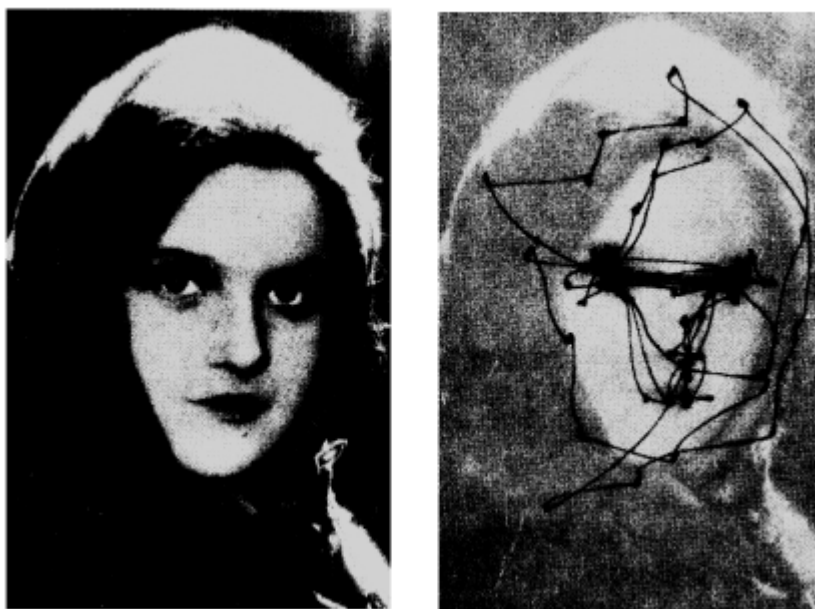


Рис.3. И скачет саккадами глаз.

А “наводчики” - кто? Правильно - видеодетекторы, причём не только движения. Их там целая бригада - что только не детектируется: это и бинокулярность (стереозрение), и контурные перепады (границы объектов), линии, концы линий, причём со всеми вариантами круговой ориентации (точность, лучше 1-го углового градуса), кривизна, цвет объекта и т.д. и т.п. Кроме того, глаза совершают микросаккады (микросканирование), что ещё больше повышает разрешение зрения - этим и объясняется такое обилие нейронов (в первичных зрительных трактах), реагирующих на движение, причём избирательно по каждому

направлению, однако совершенно не относящихся к обнаружению реально движущихся объектов. Хотите проверить? Закройте один глаз, а на другой слегка надавите пальцем через веко, чтоб не дёргался - секунд через 10 наступит полная темнота. Самое интересное - все видеодетекторы работают в ультра параллельном режиме, причём одновременно по каждому участку анализируемого изображения!

Переходим к обнаружению объектов. Под объектом, надо понимать уже нечто осмысленное, что можно описать формально с конкретным названием: человек, машина, дерево, лошадь и т.д. Простейшие объекты представляют собой совокупность пространственных комбинаций признаков, выделенных видеодетекторами. На их обнаружение требуется уже гораздо больше времени, что наглядно показывает следующий эксперимент (рис.4).

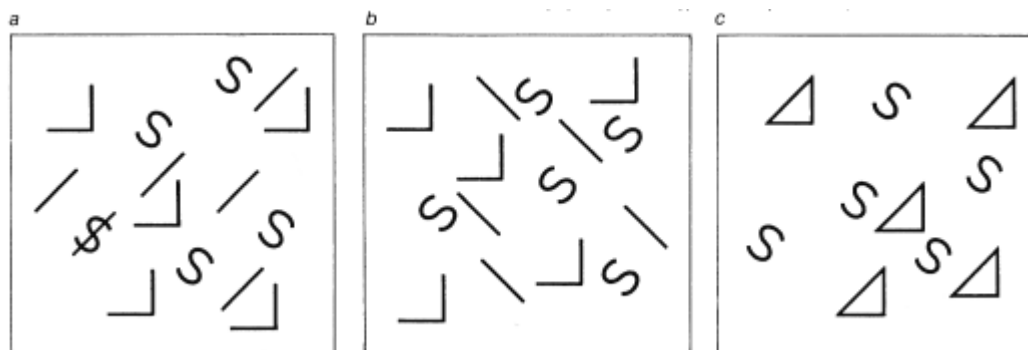


Рис.4. Где тут доллары дают?

Испытуемым сначала предъявляли изображение (а), на котором просили найти знак доллара. А затем показывали, на короткое время (200 мс), два других (b и c) - в результате подопытные очень часто “видели” искомый знак, хотя там его и не было. Это означает, что простейшие признаки успевают отдетектироваться, а вот на окончательную “сборку” объекта - уже не хватает времени.

Ещё дольше детектируются сложные объекты, состоящие из простейших. Примерно так: “Вас будет ждать мужчина в кепке, в голубых джинсах и белых кроссовках, в клетчатом пиджаке, галстук-бабочка, на лице тёмные очки, в правой руке бутылка кваса”. В данном случае требуется уже напряжённая работа нашего сознания, чтобы собрать все эти особенности в единое целое. Однако по мере тренировок, обнаружение часто встречающихся объектов переходит на бессознательный уровень. Взять хотя бы чтение. Сначала мы учимся узнавать буквы по взаимному расположению составляющих элементов, затем читаем слова по буквам (они уже узнаются автоматически), далее начинаем “проглатывать” слова целиком, как иероглифы. Ну и, наконец - машинально читаем текст, думая совершенно о другом, а прочитанное просто напрямую “проваливается” в наше подсознание. Однако самое интересное заключается в том, что всё это рано или поздно неожиданно всплывает из памяти, приводя нас в восторженно-недоуменное состояние - “И почему я такой умный - откуда знаю?” (рис.1).

А теперь высший пилотаж - распознавание ситуаций. Для большей ясности сразу расскажу реальную историю, которая случилась со мной лет так, пять назад. В одной из встреч с силовиками, после демонстрации им видеодетектора движения, состоялся следующий диалог: “Скажите, пожалуйста, а Ваша система может распознать человека с оружием? Это как? Ну, вот идёт человек в пиджаке, а карман оттопырен”. Да уж, выражение

моего лица приняло соответствующий вид (рис.1), и мне как разработчику стало жутко - какая ж гигантская пропасть лежит между нашим желаемым и действительным! Причём, нет ни доли сомнения, что мозг профессионала реально может решать такие задачи. Только вот формального описания данного процесса, мы не дождёмся. Можно лишь эмоционально охарактеризовать некоторые элементы анализируемой ситуации: “неадекватное поведение”, “блуждающий взор”, “вороватое озиране по сторонам”, “подозрительная ухмылка”, “вертлявая походка” и т.д. и т.п. Как запрограммировать это - непонятно, не получается известными методами! Поэтому, писать учебники на данные темы недостаточно - надо всё показывать натурально, в реале или в виде клипов: как обнаружить слежку, как выявить террориста в толпе или как предупредить давку на стадионе - и многое, многое другое. Это - сверхсложные задачи, решение которых возможно только на уровне подсознания. Именно здесь и проявляется наш интеллект, профессионально натренированный на распознавание нетривиальных ситуаций. Поэтому интеллект, сам по себе, крайне специализирован. Та же домохозяйка обладает кулинарным интеллектом: ей достаточно мельком взглянуть на сковородку с котлетами, чтобы скомандовать - "Переворачивать", интеллектуальное крестьянское видение подскажет - "Сеять пора", интеллект оперативного работника примет решение - "Надо брать", а криминальный интеллект, безошибочно углядит очередную жертву в толпе - “Извините, сами мы не местные ...”.

Такие вот чудеса, ну что тут скажешь, остаётся лишь восхищаться нашим мозгом - “Хорошо сидит - глубоко глядит!”. А сами-то мы где сидим, с нашими хвалёными “интеллектуальными” системами видеонаблюдения? Вот этим и займёмся.

### “ХОРОШО СИДИМ”

Хорошо-то, хорошо - да, ничего хорошего! Вот как поднимались к вершинам интеллекта HomoSapiens, так теперь и будем потихоньку спускаться в обратном направлении. Итак - распознавание ситуаций. Тут уж, как говорится, пока делать - нечего. На данном уровне у нас нет достижений, наши алгоритмы не только не могут, скажем, “осознать” динамику поведения человека - они даже не способны обнаружить его как такового, с руками и ногами, совершающими сопутствующие колебательные движения.

Однако периодически, “пытливые умы” вторгаются в высокие научные сферы, сотрясая воздух сенсационными проектами, примерно такого содержания: “Заложены основы технологии выявления террористов в толпе по походке” - дайте денег! Ну, тогда позвольте и мне историю рассказать. В Москве, на станции метро “Павелецкая-кольцевая” в кабину машиниста врывается вооружённый террорист. Далее следует диалог: “Поезд захвачен, гони в Рио-де-Жанейро! Да Вы что, с ума сошли - это ж метро! Ничего не знаю, считаю до числа  $\pi$  - три, два, один, четырнадцать сотых ...! Хорошо, хорошо. Осторожно, двери закрываются, следующая станция - Бруклинский Мост. Как так! А вот так, без дозаправки - не доедем! “.

И ведь действительно - сразу скакнуть в светлое будущее затруднительно. Почему же такое происходит? Ранее этим баловались мужи со степенями и званиями, а теперь, в эпоху демократии - все, кому не лень. Ну что тут скажешь - “целители”, с полным набором самодельных терминов и теорий-учений для избавления цивилизации от любых высокотехнологических болячек. Нет - таких не берут в “космонавты”! Считаю своим долгом, в популярной форме, донести истинное положение дел.

Почему же им - “целителям”, так искренне верят, да и сами они в это уверовали. Всё дело в ассоциативности нашего мышления и страстном желании верить в чудеса. Если снова прибегнуть к собирательному образу “домохозяйки”, то логику её рассуждений можно выразить примерно так: “Я - вижу, камера - видит. Так надо прикрутить к ней лопатку - пусть

сама котлеты и переворачивает”. Хорошо - проехали, спускаемся на ступеньку ниже, в обнаружение объектов.

Но и на данном уровне - аналогичная картина, за небольшим исключением, есть только несколько вариаций на тему: отпечаток пальца, радужка, автомобильные номера и детектор лиц (FaceDetector), о технологии FaceRecognition говорить не буду - язык не поворачивается. Однако, это лишь примитивные “плоские” решения - достаточно погрузиться в реальное 3D-пространство и... Увы, наши алгоритмы бессильны обнаружить, например, такой абстрактный объект, как “кошка”. А ведь она - киска, может в разных ракурсах: бежать, сидеть, лежать, спать, а также совершать другие многочисленные действия. И тут, на арену выходит малыш, ещё говорить-то, как следует не умеющий, и без особых проблем - всех посчитал, всё обнаружил: и киску, и собачку, и шнурки на ботинках, причём даже сообщил, что они - развязались. Так-то вот. Поэтому, в сравнении с его мозгом, наш рукотворный “интеллект” - пока лишь грудка компьютерного металлолома! Но вновь срабатывает ассоциативное мышление: “Ну, если даже ребёнок может, то уж профессиональная система видеонаблюдения и подавно!”.

А “целители” - тут как тут: продули “чакры”, расправили “карму”, заправились “энергетикой” под завязку, подмигнули “третьим глазом” и рванули - прямо в “космос”, при этом их “биополе” засияло всем спектром электромагнитных колебаний, вплоть до жёсткого рентгена! Захотелось даже встать, снять шляпу и запеть: “Я Земля, я своих провожаю питомцев, сыновей, дочерей, долетайте до самого Солнца - и домой!”. Конечно, ничего ужасного в таких полётах нет (сам по молодости летал), наоборот - они даже, в некотором роде, отрезвляют, только вот злоупотреблять этим не надо. Ну что - все “космонавты” вернулись? Отлично - спускаемся ещё на ступеньку вниз, в долину видеодетекторов.

А здесь - жизнь бьёт ключом. Почти в каждой системе есть видеодетекторы: движения, активности (это как бы недоразвитые первые), направления движения и другие экзотические (по названию) особи, кто как придумает. Например, “вентильный” или “скользящий”, ой извините - “скользящий”. Но объединяет их всех, пожалуй, наличие модной ныне приставки - “интеллектуальный”. Ну, это что-то типа титула “Сер”. Вся разница только в том, что последний присваивается английской королевой, а в нашем случае - любым менеджером по рекламе. Иногда попадаются и эмоционально-художественные вариации: “самый-самый”, “вульгарный”, “непотопляемый” и т.д. Но дело не в этом, смею Вас заверить - “интеллектуальностью” тут даже и не пахнет.

## ПАРАД “ТУПЫХ” АЛГОРИТМОВ

Начнём с нашего любимого - видеодетектора движения. Как уже было показано ранее, наши глаза “ничего не видят”, кроме узкой зоны, соответствующей “центральной ямке” сетчатки. А на периферии? Спрашивается, о каком, таком “интеллектуальном” анализе вообще может идти речь - в упор не видно ни зги! Мы даже не понимаем - что движется, зато моментально реагируем.

Теперь наглядный пример из области стереозрения, а именно: видеодетектор глубины или дистанции - как хотите, назовите. Был проведён замечательный эксперимент: взяли, да и насыпали на плоскость случайных точек, продублировали, далее выделили в каждом изображении одинаковые квадратные области и сместили немного в разных направлениях (по горизонтали) - получилась стереопара (Рис.5).



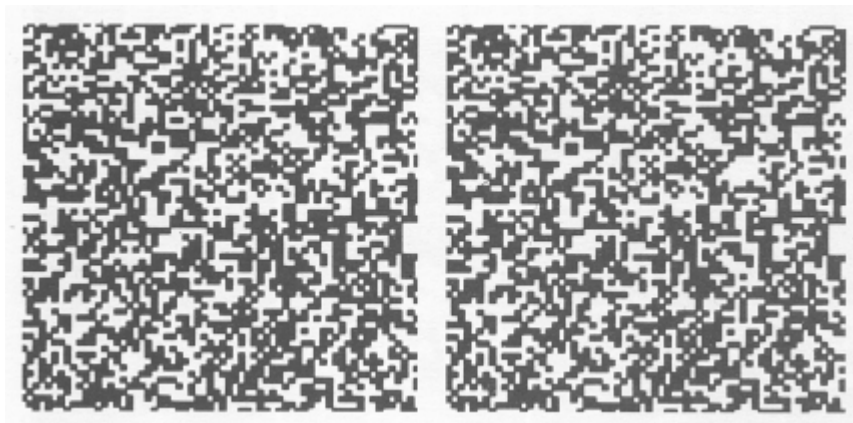


Рис.5. Не верь глазам своим.

Теперь попробуйте совместить фреймы - для этого надо как бы посмотреть вдаль (за объект), а потом навести фокус (слегка прищуриться). И Вы с изумлением увидите, как из этой каши к Вам навстречу выехал квадрат. Самое интересное, распознавать в этих картинках нечего - полная бессмыслица. А теперь вспомним детство и попытаемся найти десять отличий (Рис.6). Вместо того, чтобы анализировать число косичек у внучки и куда повёрнут хвост у мышки, просто посмотрим на это художество, как на стереопару - все отличия и “повыскакивают”.



Рис.6. К вопросу видеонаблюдения в казино.

А вот ещё один удивительный механизм видеодетекции - восстановление формы объекта при движении. Обычно им пользуются как хищники, так и жертвы (каждый в своих интересах), поскольку в неподвижном состоянии их не видно - маскируются. Итак, построили 3D-модель двух цилиндров, накидали на их воображаемые поверхности точки, в случайном порядке (Рис.7) и закрутили - один по часовой стрелке, а другой - против. Если рассматривать каждую проекцию этого действия на сетчатку отдельно (по кадрам), то получаем хаос, но стоит только выдать серию - ух ты, цилиндры в наших мозгах вращаться начали.

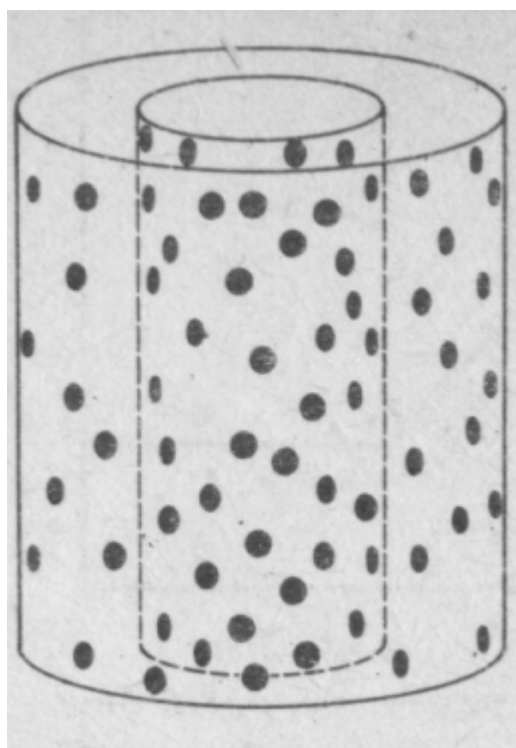


Рис.7. И всё-таки они вертятся.

Думаю - примеров предостаточно. Какие же выводы можно сделать. В зрительной системе HomoSapiens работают “тупые” (в хорошем смысле этого слова) алгоритмы видеодетекции, без обилия всевозможных операторов “if-goto” и рекурсий - изображения буквально продавливаются сквозь нейронные слои, как варёный картофель сквозь сито. Фантастическая мощь - и никакого “интеллекта”.

Конечно, кто-нибудь мне возразит: “А мы вот пойдём - другим путём. Будем использовать: текстурный, морфологический, макротопологический и другие разнообразные анализы, вплоть до сегментации”. Да без проблем, делайте то, что считаете нужным, только вот живую Природу, частью которой Вы и являетесь - не проведёте. Во-первых, такие алгоритмы потребуют более высокого разрешения, что отразится на быстродействии. А во-вторых, они просто не будут работать на приведённых ранее, в качестве примеров, изображениях. Слышу возражения: “Нет таких в реале”. Да, в чистом виде - нет, однако многие будут приближаться к оным, в силу: малой контрастности, наличия разнообразных естественных и искусственных помех, а также маскировки объектов (камуфляж надели).

И не надо морочить людям головы о том, как “космические” алгоритмы бороздят просторы ..., вычисляя у каждой снежинки и дождейки: вектор движения, скорость, ускорение и т.д. Хотите наглядный пример? Берём маленького котёночка, сводим указательные пальцы вместе и совершаем ими кругообразные движения. Любопытная усатая мордочка моментально начинает отслеживать все наши траектории. Затем, резко разводим руки, и видим, как у изумлённой киски, глаза буквально разлетаются в стороны. Вот примерно так, и “разлетаются”, причём в дребезги, “интеллектуальные” алгоритмы слежения за целями, когда тех становится слишком много - аж две! Не усложняйте, именно об этом и писал Генри Каттнер - “робот-зазнайка”, в конце концов, оказался консервным ножом для открывания пивных банок, так уж его изобретатель задумал, но слегка перемудрил. Продолжаем разговор.

## СПАСИ И СОХРАНИ

Итак, подведём предварительный итог. Смею констатировать - наши алгоритмы видеоанализа практически ничего не умеют, кроме примитивной детекции движения. Да и работает всё это не лучшим образом - мы только в самом начале пути, предстоит сделать очень много, в том числе реализовать такие функции, которые Природа и не предусмотрела. Первые шаги уже сделаны.

За последнее столетие человечество кардинально изменило мир, насытив его жуткими кинетическими и потенциальными энергиями. У Эволюции не было времени научить живые существа опасаться остановившихся предметов - угроза всегда исходила от движущихся. Но мы поменялись местами, и теперь сами носимся в самодельных механизмах со страшной скоростью.

Речь пойдёт о новом классе видеодетекторов: SlowDownDetection - SDD (Детекция Замедления). Есть и более популярное название - Детектор оставленных/унесённых предметов. Алгоритм позволяет обнаруживать объекты, которые двигались и остановились на время, более заданного. И/или наоборот, сигнализировать через тот же временной интервал, если некоторые предметы покинули первоначальные позиции. Алгоритм контролирует одновременно все объекты, отдельно по каждой из функций на фоне с движением. Остановилась машина в тоннеле, отвалилась и упала на полосу деталь у взлетавшего самолёта, что-то застряло на железнодорожном переезде - вот лишь некоторые ситуации, приводящие к катастрофе. Однако, область приложений технологии SDD намного шире, о чём и пойдёт речь далее. Но прежде хотелось бы понять - “А зачем всё это видеодетектирование нужно?”.

Мнения явно разделились: одни уверены - будущее за системами с видеоанализом, других вполне устраивает текущее положение дел (в голове большой такой цифровой видеомагнитофон и не более того), третьи - сомневаются, уж больно сильны стереотипы, якобы видеодетектирование вещь ненадёжная и порождает лавину ложных срабатываний. Ну и наконец, есть четвёртая группа “специалистов”, которая в агрессивной манере пытается доказать бесперспективность новых технологий. Причины такого поведения видны как на ладони - конкурентная борьба, а вернее война. Вот так получилось и с технологией SDD: одни порадовались, другие вообще не отреагировали, третьи засомневались ещё больше, а вот четвёртые - встретили в штыки. Первым делом они сознательно сузили область приложения функции, окрестив её антитеррорной (в плане обнаружения оставленных/остановившихся предметов с взрывчаткой). Конечно, террор - явление ужасное, однако в техногенных авариях гибнет на порядки больше людей. А далее последовал “абсолютно бронебойный” аргумент: “Эта технология совершенно бесполезна, поскольку система не сможет отличить ведро с вареньем от ведра с динамитом”. Да уж, смахивает на воинствующий атеизм: “Бога - нет, потому что космонавты на небо летали и никого там не видали”. А если воспроизвести внутренний голос, то выглядит немного иначе: ”Прости нас - господи, космонавтов неразумных, в проблему въехать не можем, мозгов не хватает”.

Хорошо, давайте разбираться вместе. Прежде всего, система видеонаблюдения - не собака-ищейка. И основная её задача - привлечение внимания оператора. Поэтому такую систему уместно сравнить уже с цепной собакой, которая в случае приближения или внедрения на вверенную ей территорию чего-либо, начинает громко лаять. Выходит хозяин и с высоты своего мозга уже сам оценивает ситуацию и принимает решение. А почему собака на цепи? А чтобы кого-нибудь не покусала - мы ведь не можем доверить “тупой” системе какие-либо действия, кроме извещения и отпугивающего лая. Однако собаки (видеодетекторы) разные бывают. Если Ваша зверюга, как бешенная, непрерывно лает: на

Луну и звёзды, на птичек и кошек, на дождь и снег, а также на шорохи листвы, да и вообще беспричинно, то возникает единственное желание - пристрелить! Это к вопросу ложных срабатываний. Вот Вам и ответ на поставленный ранее вопрос - как только видеодетекторы приблизятся по надёжности к традиционным охранным датчикам, то моментально будут востребованы и составят последним достойную компанию.

А технология SDD уже достигла требуемого уровня надёжности и с успехом используется на периметре в качестве средства обнаружения. Это стало возможным за счёт использования интегральных методов - обрабатываются серии изображений, без поиска каких-либо смысловых признаков в отдельных кадрах. Поэтому алгоритм SDD можно смело назвать “тупым”, по аналогии с видеодетекторами в зрительной системе человека. Попробую объяснить в более популярной форме. Лет так 15 назад в “разминке” КВН-овского матча был задан вопрос: “Почему у бегемота красные глаза”, на что и был получен соответствующий ответ - “Чтобы прятаться в помидорах”. Так вот, если в Вашу помидорную оранжерею залезет бегемот в камуфляже и затаится, то система его всё равно обнаружит, хотя на отдельных снимках Вы бегемотика и не увидите.

Теперь рассмотрим некоторые примеры использования данной технологии, в порядке, соответствующем возрастанию времени детекции. Причём, чем больше это значение, тем выше помехоустойчивость. В диапазоне - менее секунды, алгоритм очень хорошо справляется с высокоскоростными помехами: дождь, снег, падающие листья, мошкара перед объективом, летящие птицы и т.д. В интервале 1-5 секунд - система начинает сигнализировать о пробках на дорогах. Значения от 5 до 20 секунд самые ходовые: остановка автомашин на трассах, в тоннелях, перед офисами и других публичных местах. При этом надо четко понимать - основное назначение алгоритма это лишь привлечение внимания службы безопасности. Например, разве не подозрительно, что некий индивид три раза за день присел завязать шнурки перед входом в офис или вертелся перед входной дверью Вашего дома или около машины. Большие времена детекции (1-10 минут) хорошо использовать на открытых безлюдных пространствах, когда нужно предотвратить такие действия как: отпилить, утащить, врезаться, подложить и т.д. А вообще был интересный случай: в расчётном зале одной телефонной компании установили систему, причём время детектирования задали - 5 минут, так она начала как неводом вылавливать оставленные мобильные телефоны и борсетки.

Ну и, наконец, пример сверхбольших интервалов времени. Дело было в одном из городов США. По улице ехала бабушка на самоходной электротабуретке. При этом она, как фея, макала в ведёрко с известью волшебную палочку и прикасалась ею к колёсам стоявших машин. Восхищению моему - не было предела: “Какая прелесть, наверное, это к счастью”. “Это точно” - подтвердил мой местный друг, “через два часа фея вернётся и осчастливит владельцев штрафом за превышение лимита времени парковки!”.

Конечно, можно много рассказывать и фантазировать на тему. Однако то, что реально произошло - просто повергло всех, и меня в том числе, в состояние шока (в хорошем смысле этого слова). Случилось это на стоянке супермаркета в европейской глубинке. Был будний день - народу никого, подъехала молодая мама с грудным ребёнком (он сзади в специальном креслице сидел). Функция SDD отработала через 20 секунд, оператор посмотрел - ничего подозрительного. Далее девушка перешла к ребёнку и стоя на коленках, перевесила его в нагрудный рюкзачок (к себе лицом). Стала выбираться, резко разогнулась, сильно ударилась затылком о ребро дверного проёма и, теряя сознание, мягко упала - ребёнок оказался между ней и землёй. Повторно сработал видеодетектор - остальное было делом техники: подбежали, перевернули, а там дитё с синевой на лице. Слава богу - обошлось. После этого, мне - изобретателю сего, уже не надо мучиться вопросом об актуальности SDD-технологии.

## **”ИНТЕЛЛЕКТ” УМЕР - ДА ЗДРАВСТВУЕТ ИНТЕЛЛЕКТ**

Вот мы и подошли к финалу, по закону жанра следует задать вопрос: “Быть или не быть, интеллектуальным системам видеонаблюдения”. Отвечаю категорически - быть! Но только с одним условием, если кресло председателя займёт “его Величество” - мозг HomoSapiens. Мы имеем уникальную картину. С одной стороны - только человек может, благодаря своему интеллекту, легко обнаруживать объекты и распознавать ситуации. Но он быстро утомляется и не способен, не только одновременно смотреть в десятки окон с изображениями или в квадратор, а даже длительное время контролировать всего лишь одну камеру. С другой стороны, компьютер не устанет - никогда (если только программное обеспечение не кривое) и будет без усталости детектировать события от десятков источников. Одна лишь беда - ума пока не набрался, но ничего - научим.

Итак, вот он, замечательный симбиоз - человек-машина, как органическое дополнение друг друга. Однако видеодетекторам надо подтянуться, чтобы не нервировать биомозги. Слышу традиционный протест сторонников цифровых видеоманитофонов - “А в казино этого, не надо”. Можно подумать, что вся CCTV-индустрия сводится к обслуживанию игорных заведений. Нет, мы должны создавать активные системы, которые не просто регистрируют, а реально предотвращают трагедии и другие негативные проявления. Что толку смотреть в архив: как горят заживо люди в тоннелях и пылающий Конкорд летит - прямо в рай. Но ничего уже изменить нельзя, остаётся лишь, стиснув зубы, мысленно повторять - “Ну как же, так”, и посыпать голову пеплом! Нет у нас такого права, почивать на лаврах достигнутого - работать надо.

Ну вот - все мечты пока вроде и закончились. Вы уж меня извините - кого задел, ненароком. Это от большой любви - в самом хорошем смысле этого слова. Есть такая замечательная народная мудрость: “Бей своих, чтоб чужие боялись”. И правильно будут делать, поскольку российский интеллект - самый мощный в мире, и не иссякнет никогда, хотя бы потому, что любой ребёнок, изучающий с пелёнок великий и могучий русский язык, получает супер гимнастику мозга!

**Об авторе:** Руцков Михаил Вадимович, кандидат технических наук, директор [MegaPixel Ltd.](http://MegaPixel Ltd.), e-mail [mailto: megapixel@tochka.ru](mailto:megapixel@tochka.ru) тел. (095)4129422