

К.т.н. Руцков М.В.  
Директор MegaPixel Ltd.  
megapixel@tochka.ru  
095 4129422

## **ВИДЕОДЕТЕКТОРЫ – ВЗГЛЯД ИЗНУТРИ**

От обилия цифровых видеорегистраторов, высыпающих на рынок систем безопасности как грибы после дождя, аж рябит в глазах! У потребителя голова идёт кругом. Причём все системы похожи друг на друга как близнецы-братья. Наиболее ярко это явление наблюдалось на MIPS 2002 в “юго-восточном квартале”. И в этом нет ничего удивительного, поскольку практически все видеорегистраторы собраны из одних и тех же аппаратно-программных компьютерных кубиков. Как следствие – совпадение технических характеристик, если они, конечно, не поданы с некоторой долей лукавства. И на этом равномерном фоне: числа каналов, скоростей записи и отображения, длительности регистрации и т.д., особняком стоит некая “вещь в себе” - Детектор Движения, наличие которого сопровождается в таблице характеристик лишь знаком “плюс”. С таким же успехом можно назвать и телегу, и болид “Формулы-1” – транспортным средством!

Целью настоящей статьи является попытка ответить на волнующие многих вопросы: Что такое Детектор Движения? Так ли он необходим в видеосистемах (подсистемах) безопасности? Каковы перспективы развития данного направления? Сразу хочу отметить, что всё нижеизложенное является только моим субъективным мнением, это просто мысли вслух или взгляд изнутри проблемы.

### **ДЕТЕКТОР ДВИЖЕНИЯ – ЧТО ЭТО?**

Чего только не пишут фирмы в рекламных релизах о своих Детекторах Движения, снабжая их усилительными эпитетами типа: истинный, интеллектуальный, эвристический, макротопологический и т.д. Мне как-то попался даже и – абсолютный! Однако от этого ясности не прибавляется.

Если воспринимать термин Детектор Движения буквально, с точки зрения законов физики, то получается, что это устройство для обнаружения движения некоего объекта. Что тогда понимать под объектом? Можно ли считать объектом нечто волнообразное с плавным изменением яркости? Стоп – пора ставить ограничители. Будем отталкиваться от природы, которая за миллионы лет эволюции создала широчайшую гамму зрительных систем живых организмов. Основное назначение – выживание, т.е. обнаружение опасных и съедобных объектов. Именно это мы и хотим позаимствовать у живой природы для наших систем безопасности. Есть правда ещё и функция размножения, но это уже отдельная тема. Многочисленные нейрофизиологические исследования таких зрительных систем (и в первую очередь человека) показали, что они выделяют именно границы объектов (более популярно - контуры), где резко меняются свойства отражающих или излучающих поверхностей. Вот если бы мы жили в мире Фурье ..., но мы живём в мире с контурами. Хорошему художнику достаточно сделать несколько штрихов – и мы узнаём образ.

Итак, объект обнаружения в общем виде определён – что-то с контурами. Следовательно, на изменения фона, например при заходе солнца за облако, реагировать не надо. Однако многие системы реагируют, что считается ложным срабатыванием. Отсюда и вся неразбериха в терминологии. Их, почему-то называют Детекторами Активности. Причём достаточно сильно укоренилось мнение о том, что Детектор Движения это хорошо, а Детектор Активности – так себе. Однако название Детектор Движения тоже ни о чём не говорит, так как присваивается не группой независимых экспертов, а самими

разработчиками, причём не без участия менеджеров по рекламе. Например, одна известная фирма выпустила семейство WaveLet-кодеков и в порыве рекламной страсти заявила, что, кроме самой компрессии, попутно (без всяких затрат) детектируется и движение. Однако основной задачей любого motion-кодека является кодирование всех кадров с минимальными потерями, т.е. тот поток различий, который и предлагается использовать в качестве Детектора Движения, содержит в интегральном виде всё и вся, и претендует на роль самого всеобъемлющего Детектора Активности.

На самом же деле Детектор Движения для систем безопасности должен, в зависимости от задачи, игнорировать одни виды движения, например: снег, дождь, листья (ну чем не реальное движение), и обнаруживать другие – люди, машины и т.д. Той же системе может быть поставлена и другая задача: обнаруживать только машины (у офиса) или только людей (на скоростной автостраде). И таких вариаций может быть сколь угодно много. В идеале мы стремимся к человекоподобной системе, которой надо только объяснить, что делать. Поэтому простора для творчества здесь предостаточно. Однако и само понятие Детектор Движения становится слишком узким - есть задачи, в которых требуется определять не движение объекта, а его остановку, когда необходимо реагировать на засветку камеры лазером или закрывание объектива рукой и т.д. Считаю, что более уместным был бы термин ВидеоДетектор, с соответствующим описанием его функций.

## **ВИДЕОДЕТЕКТОР – БЫТЬ ИЛИ НЕ БЫТЬ?**

На этот счёт есть самые разнообразные мнения: одни категорически против, другие хотели бы иметь для “плюсика”, ну а третьи – только за. Обычно противники выдвигают следующий аргумент – ВидеоДетекторы не нужны, поскольку они плохо работают и только нагружают процессор. Надо писать всё подряд по принципу: “а как бы чего не вышло”.

Конечно, компьютеры наращивают мощности в геометрической прогрессии, причем заметьте, независимо от пожеланий разработчиков систем безопасности, последние могут только пользоваться благами прогресса. И сразу возникает соблазн реализовать в одном видеорегистраторе побольше каналов, причём с максимальным разрешением, да ещё и работающих в реальном времени. Как следствие, начинается “гонка вооружения” с сопутствующими “рекламными битвами”, которые порождают в воспалённых головах потребителей сверх завышенные ожидания. Примерно в таком виде: 16 каналов, каждый по 25 fps на отображение и запись (всего по 400 fps) в формате 768x576 в цвете (PAL), со временем архивирования около недели! Вроде бы каждый из этих параметров по отдельности достигим, да вот вместе они не живут. Попробую аргументировать.

Одним из главных завоеваний цифровой видеозаписи, в сравнении с аналоговой, является более высокое качество изображения. Действительно, смотреть DVD-диск (500 ТВ-линий) намного приятней, чем VHS-кассету (250 ТВ-линий). Изначально на DVD-диск емкостью в 4.7 Гбайт записывали 133 минуты видео со звуком в формате MPEG-2, т.е. порядка 2 Гбайт/час. Несложный расчёт показывает, что 16 каналов заполнят жесткий диск объёмом в 120 Гбайт (наиболее ходовой на данный момент из соотношения ёмкость/цена) всего за 4 часа! Для увеличения времени регистрации используется хитрый приём - урезание кадра до размера 384x288 (а иногда и того меньше) и компрессия с гораздо более высокой степенью сжатия, чем в DVD. Однако от этого страдает и разрешение, падая даже ниже уровня в 250 ТВ-линий - заколдованный круг. Кроме того, необходимо отметить, что и DVD, а вернее MPEG-2, искажает мелкие детали изображения. Но в художественных фильмах это не страшно, поскольку зрительная система человека эти детали в движении не различает, да и все объекты обычно сняты крупным планом при хорошем освещении. А вот в охранном видео всё наоборот – надо на стоп-кадре рассматривать всякую мелочь (лица, предметы и т.д.) и делать это без помощи специальных алгоритмов (увеличения размеров и повышения контраста мелких деталей) затруднительно. Но они плохо работают на

компрессированных изображениях. Практической пользы от полнокадровой (оба полукадра) записи - тоже никакой, скорее вред. Поскольку полукадры разнесены во времени на 20 мс, то за это время, например при скорости в 6 км/час (2 м/с), объект смещается на 4 см, и начинают на стоп-кадре появляться два носа, четыре глаза и т.д.

Так какой же выход – как и качество сохранить и время регистрации увеличить? Ответ однозначный – писать только по ВидеоДетектору! Однако это не самый главный аргумент за использование ВидеоДетекторов. Основная проблема возникает при работе с архивом. Сколько времени надо для просмотра 3-х часовой кассеты? Да все три часа, если не больше. С появлением цифровых многоканальных видеорегистраторов надо уже смотреть в полиэкран, причём по очереди на каждый канал, чтобы не пропустить какое-либо событие. Физически одному человеку такое не под силу, а с появлением мощных распределённых систем – это уже полный технологический тупик.

Единственное спасение – использовать запись по ВидеоДетектору, который позволяет фиксировать только интересующие фрагменты. Более того, идеология видеодетектирования даёт возможность дифференцированно описывать события. А именно: создается набор виртуальных ВидеоДетекторов, каждый из которых содержит информацию о номере канала, номере зоны в канале, типе детектора и его параметров в данной зоне. Далее с помощью булевой алгебры задаётся событие со своим идентификатором, как функция от выбранных виртуальных ВидеоДетекторов и сигналов от внешних датчиков. Таким образом, в архиве по идентификаторам можно будет получить быстрый доступ к интересующей информации.

Ну и, наконец, основной аргумент – видеонаблюдение должно стать средством активного обнаружения, а не оставаться на роли пассивного регистрирования. Причём речь идёт именно о наружном наблюдении. На периметре, видеодетектирование может, если не заменить, то существенно сократить использование дорогих традиционных средств. А на больших открытых пространствах – альтернативы просто нет.

Однако в многочисленных публикациях по периметральным системам, видеонаблюдение фигурирует только как вспомогательное средство, без упоминания о функции обнаружения, что и соответствует действительности. Основная причина – низкое качество ВидеоДетекторов. В чём же дело?

## **КТО ВИНОВАТ?**

Традиционный вопрос. Рассмотрим процесс создания видеорегистратора собственной разработки. Наиболее часто такое желание приходит в голову торговым фирмам или фирмам-инсталляторам, вынужденных продавать или ставить на объекты системы сторонних производителей. Сроки разработки изначально жёсткие – не более полугода, о каких-либо исследованиях и речи быть не может, конкуренты поджимают. С чего начать?

Первым делом надо определиться с аппаратной платформой, и выбор здесь однозначный – что-то компьютерное. Просто не имеет смысла тратить деньги и время на разработку своего специализированного, невозможно конкурировать с компьютерной индустрией с её много миллиардными вложениями. Далее надо найти хорошего системного программиста с командой. И здесь особых проблем быть не должно, поскольку огромный сектор серверных приложений порождает таких специалистов в достатке.

Теперь немного о специфических особенностях - тех, что и превращают сервер в видеосервер. Всё начинается с оцифровки видеосигнала. Лет десять назад на рынке было пусто, однако после выпуска интегрированных чипов видеodeкодеров, да ещё и со встроенным РСІ-интерфейсом, ситуация резко изменилась. Сначала хватало одной карты с внешним мультиплексором, далее с ростом мощности компьютеров стали появляться многоплатные вариации, ну а сейчас всё чаще используются многоствольные (более одного АЦП) оцифровщики. Причём некоторые такие устройства, доступны в свободной продаже.

Ещё одна специфика видеосерверов – компрессия видеопотоков и отдельных кадров. Ну здесь-то точно мультимедийная индустрия постаралась. Есть и специалисты и готовые к продаже продукты, а также бесплатные кодеры-декодеры с исходными данными, прямо в Интернете. Например, фирма Intel выкладывает оптимизированные под технологии MMX и SSE библиотеки компрессии, только покупайте их процессоры. То же самое можно сказать и о технологиях отображения видеоданных – и AGP-шина и графические акселераторы и т.д.

Складывается радужная картина – даёшь собственный видеорегистратор за пол года. Конечно, не всё так просто, это серьёзная работа, причём успешная её реализация более чем на половину зависит от координатора (руководителя) проекта. Но и такого найти можно, а во всём остальном требуется только одно – разобраться с тем, что уже было создано, и аккуратно собрать в единое целое. И никакого в этом сверх достижения нет - просто время пришло. А то, иногда слышишь буквально следующее: “Сенсация! Впервые в мире реализовано управление мощным видеокомплексом с помощью инфракрасной клавиатуры, или - через Интернет по мобильному телефону, или - по беспроводной связи с карманного компьютера!” Аж дух захватывает.

Ну а теперь – момент истины, что там с ВидеоДетекторами? Увы – дела плохи, дальше некуда. Готовые решения не продаются, специалистов в данной области нет, поскольку их никто и не готовит. Те научные школы, которые были лет 15 назад, практически перестали существовать, одни разъехались, другие переквалифицировались, ну а третьи просто состарились. Да и задачи, решаемые на том этапе, были далеки от сегодняшних запросов. Сейчас уже не достаточно, просто физико-математического образования, необходимо погрузиться в проблему по всей программе научно-исследовательских работ. Однако заказчики не хотят рисковать - тратить деньги и время на изыскания, им нужен продукт, да побыстрей, который можно продавать. Более того, многие наивно полагают, что писать алгоритмы это то же самое, что и писать программы. Вот программисты и пишут!

## ЧТО ДЕЛАТЬ?

А вот это уже не вопрос, а скорее образ мышления в эпоху быстротечных конкурентных схваток. Надо что-то делать! Первым делом – в Интернет, а там ничего и нет. Последние книги по видеоанализу были переведены на русский язык, лет так 15 назад. К сожалению, суровая действительность коммерческого подхода к издательской деятельности, не предусматривает узкоспециализированных изданий – объёмы требуются. Не стоит обольщаться и многочисленными бесплатными приложениями, которые появляются на сайтах зарубежных университетов. Это либо компания по привлечению инвестиций, либо целенаправленная дезинформация. Никто и никогда не выложит “за просто так” свои алгоритмы, имеющие коммерческую ценность, всё это за семью печатями с патентами. Надо начинать думать своей головой. И вот что из этого получается.

Один из распространённых приёмов: на контролируемом изображении расставляют маркеры и смотрят на изменения яркости в них. Если эти изменения больше заданной величины, то объявляется тревога. Ну, это как дырок в заборе насверлить и, вглядываясь метров с трёх, пытаться понять – а что ж там происходит.

С увеличением мощности компьютеров появилась возможность обрабатывать всё поле изображения. Для этого оно разбивается на микроблоки: 2x2, 4x4, NxM пикселей – у кого как получается. Далее вычисляется средняя яркость и всё остальное, как с маркерами. А можно из микроблоков выстраивать макроблоки (зоны) и уже на втором этапе подсчитывать число сработавших микроблоков, словом колдовать и вдоль, и поперёк. Одно только ясно - все изменения фона будут учитываться, как в Детекторе Активности. Кроме того, есть взаимоисключающие моменты: уменьшение размеров микроблоков приводит к усилению влияния шумов, а увеличение – к снижению пространственной чувствительности.

Можно конечно много рассказывать о, такого рода, алгоритмах. Мне и самому посчастливилось в течение нескольких лет ходить кругами по этому минному полю, прежде чем в голове не созрело твёрдое убеждение – всё это не работает! Выход был один – обратиться к матушке-природе, у которой с детекцией движения полный порядок. Пришлось ещё несколько лет изучать многочисленные материалы нейрофизиологических исследований и экспериментировать, прежде чем хоть что-то заработало.

Основная идея заключается в следующем: если контуры объектов не меняют своего положения, то движения нет! Зрительная система человека как раз и выделяет контуры. На этапе предобработки это делается в самой сетчатке глаза. Исходное изображение, созданное светочувствительными элементами (палочками и колбочками), подвергается процедуре свёртки с ядром, имеющим мудрёное название – лапласиан от гауссиана. Трёхмерный график такого ядра напоминает мексиканскую шляпу сомбреро, центральная часть которой имеет положительные коэффициенты, а поля этой шляпы – отрицательные. Строго говоря - это полосовой пространственный фильтр. И таких фильтров в зрительной системе предположительно пять, каждый для своей полосы. Далее по зрительным нервам информация поступает в первичную зрительную кору, где и выделяются контуры. Причём делается это сугубо нелинейными процедурами.

Однако, правильно работающий алгоритм - это только пол дела, надо чтобы он был ещё и быстрым, т.е. аппаратно-программно ориентированным. Например, свертка с ядром размером 13x13 пиксел требует выполнения по 169 операций умножения и сложения, и так для каждой точки изображения. Если все эти вычисления выполнять, как говорится “в лоб”, то время детекции будет исчисляться уже секундами, а не миллисекундами. Поэтому математик, разрабатывающий алгоритм, должен хорошо разбираться как в возможностях современных программируемых вентильных матриц (аппаратная поддержка алгоритма), так и в микроархитектуре процессоров (программная обработка). А то иногда приходится слышать буквально следующее: “Он опередил своё время, вычислительные мощности ещё не доросли до реализации его алгоритмов” или более конкретно: “Свершилось, впервые реально освоена ММХ-технология!”, и это на шестом году от рождества её.

В результате исследований и экспериментов были созданы новые - высокоскоростные нейро-подобные алгоритмы, обладающие высокой пространственной и контрастной чувствительностью, что означает возможность детектирования малых по размеру и слабоконтрастных объектов без ложных срабатываний при флуктуациях освещения. Причём термин “нейро-подобный” отражает тот факт, что осуществляется не точное копирование первичных нейроструктур зрительной системы человека, а используется только сама информационная суть процессов.

Однако, не смотря на кардинальное улучшение качества детектирования, новые алгоритмы не справлялись с задачей охраны периметра. Основная проблема возникала с рядом естественных и искусственных помех, таких как: птицы, насекомые, снег, дождь, блики, свет фар и т.д. Попытка подбора полосовых фильтров и селекции объектов по размеру не приводила к желаемому результату. Решение, однако, “плавало на поверхности”. Анализ показал, что всё эти помехи были высокоскоростными, т.е. имели большие угловые скорости перемещения. Для реализации данной идеи был создан новый тип ВидеоДетектора – SlowSpeedFilter или Фильтр Медленных Скоростей, который успешно прошёл испытания и уже работает на ряде объектов, как средство периметральной защиты.

Ну и, наконец, была создана уникальная технология - SlowDownDetection или Детекция Замедления. Она также известна под названием Детектор Оставленных/Унесённых Предметов. Такой ВидеоДетектор позволяет обнаруживать появление и остановку объектов и/или выявлять их пропадание через заданное время. Алгоритм работает на фоне движения других объектов, причём даже в условиях временного перекрытия последними поля зрения. Первым делом, данная технология ассоциируется как антитеррорное и противокражное средство, однако область приложения её намного шире. А именно: безопасность на скоростных автострадах (особенно в тоннелях) и взлётных полосах – поиск опасных

предметов; обнаружение пожаров; борьба с вандализмом; охрана оборудования на больших открытых пространствах; тот же периметр; и многое другое. Следует отметить, что технология оказалась востребованной потребителями, а некоторые компании даже приступили к созданию аналогичных разработок, что, несомненно, радует.

### ЧТО В ОСТАТКЕ?

Так какие же выводы можно сделать. Самый главный – ВидеоДетекторы не просто желательны в системах видеонаблюдения, а крайне необходимы, иначе технологический тупик. Второй вывод – ВидеоДетекторы могут, и будут работать на периметре, в качестве средства обнаружения. Ну а третий вывод, наверное, шокирует многих: самый главный враг детектирования – предварительная компрессия изображения. Забудьте о ВидеоДетекции, если в системе используются платы с аппаратным сжатием на борту! Однако хочу ещё раз подчеркнуть – всё вышеизложенное является только моим субъективным мнением. А вот объективным является тот факт, что - стереотипы трещат, лёд тронулся, процесс пошёл!

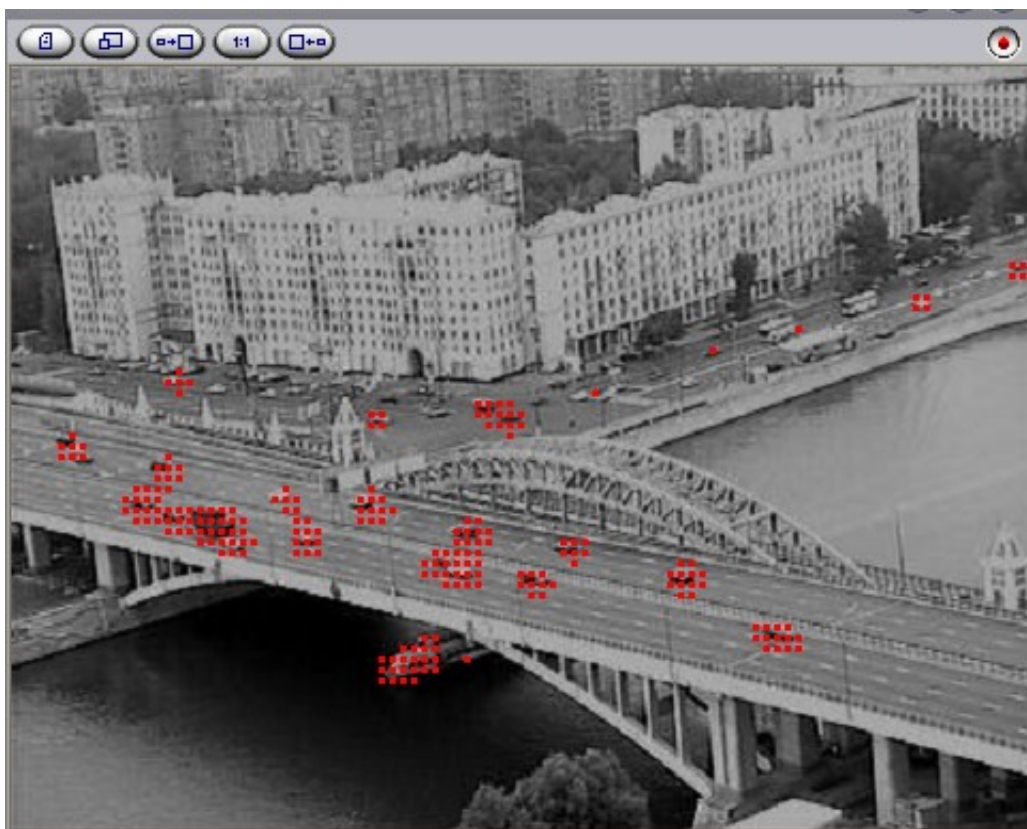
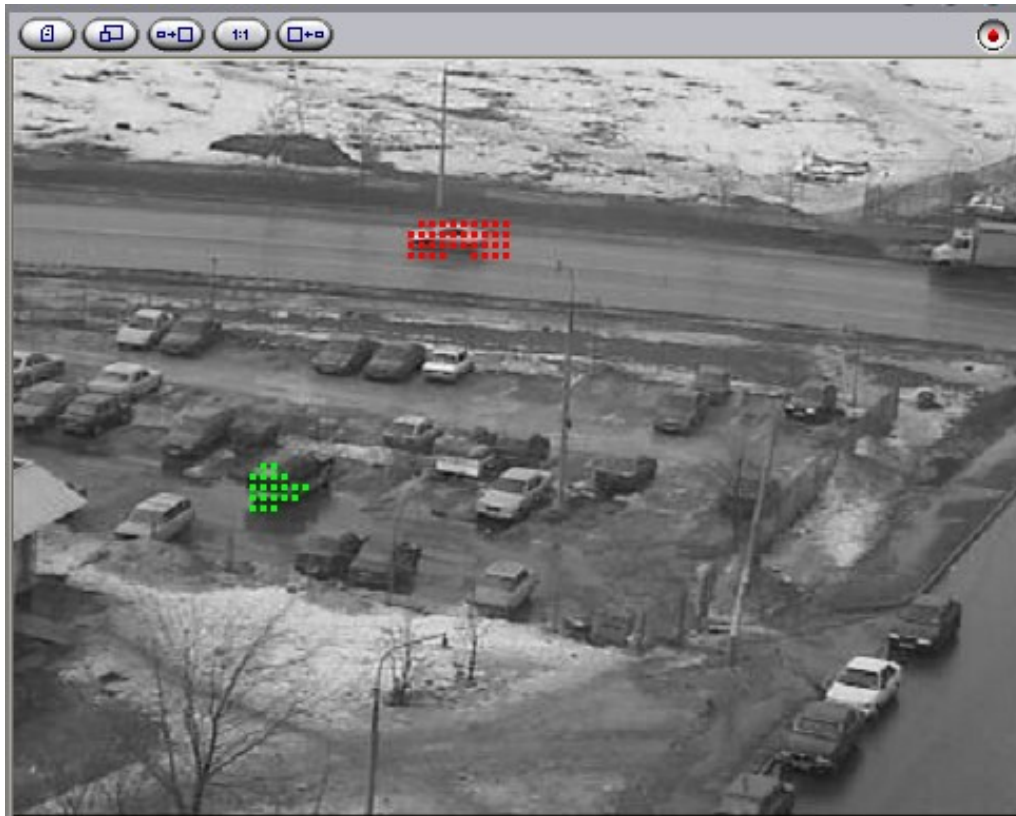


Рис. 1. Автомагистраль: работа в плотном потоке



**Рис. 2.** Автостоянка: детекция движения (красный маркер) и остановки (зеленый маркер)