

прикладная видеология

крик касперски, aka мышьх, по-email

широкий ассортимент выбора камер видео-наблюдения превращает свободу выбора в проблему, особенно если ни потребитель, ни продавец не разбираются в технических характеристиках и разницу между avi и mpg чувствуют на смутном интуитивном уровне, не подвластном логическому анализу. давайте попробуем разобраться какие стандарты хранения видеинформации существуют и чем они _реально_ отличаются друг от друга.

введение

Снижение цен на электронику вызвало встречный рост популярности систем видеонаблюдения. И хотя подобная защита во многом носит чисто психологический характер, отпугивающий мелких преступников от кражи и создающий иллюзию безопасности, наличие видеозаписи упрощает "разбор полетов", пускай и реальные случаи задержания носят единичный характер, отчасти вызванный безграмотными настройками охранной системы. Что толку, если силуэт грабителя уверенно отличается от силуэта собаки? Ладно, лицо размазано и неразборчиво, ладно, темно как в танке, так ведь и угол охвата настолько широк, что перспективные искажения не позволяют нам сказать: толи это низенький толстенький мужичок, толи это высокая худощавая блондинка. Конечно, ситуация несколько утрирована, но в целом по стране качество видеозаписи таково, что разобрать кто на ней запечатлен можно только сравнив "портрет" с оригиналом.

Кто виноват? Разрешение современных камер достаточно высоко. Современные алгоритмы сжатия (при правильном их применении) обеспечивают более чем удовлетворительный результат. Наконец, современные накопители данных позволяют записывать тысячи и тысячи часов с подобающим разрешением. Так почему же тогда мы имеем то, что имеем?!

Ответ прост. Обработка видеинформации — сложное дело, требующее предварительной подготовки и управлять охранными видео-комплексами должны специалисты, а не системные администраторы или электрики. Но, увы, таких специалистов в нашем распоряжении нет и быстро они не появятся, поскольку те, кто освоил цифровое видео на уровне чуть выше faq, утекают в более прибыльный бизнес, занимаясь тем, что приносит деньги, в то время как поддержка охранного комплекса — это статья расходов и выделять под нее большой бюджет могут только те организации, которые делают деньги буквально из воздуха (например, банки). А всем остальным — что делать?

Эта статья не претендует на исчерпывающее рецептурное руководство, но, по крайней мере, поможет уберечь вас от популярных ошибок и развеять распространенные заблуждения. Она ориентирована как на тех, кто уже приобрел видео-комплекс, так и на тех, кто только раздумывает — что же ему покупать. Никакой специальной подготовки от читателя не требуется и все термины объясняются по ходу дела.

от камеры до прайса и обратно

Камеры видео-наблюдения можно разделить на две большие категории: первая выдает "сырой" аналоговый видеосигнал (как правило, в системе PAL), вторая — готовое цифровое изображение, сжатое по стандарту MPEG-1/MPEG-2/MPEG-4/x264. Достоинство первой — более низкая цена и (потенциально) более высокое качество, поскольку "добротность" встроенных в камеру энкодеров оставляет желать лучшего, а возможность их ручной настройки сведена к минимуму. Зато камеры со встроенными энкодерами очень часто содержат и встроенный TCP/IP стек с Ethernet-портом, что позволяет их подключать к уже существующей локальной сети, избегая прокладки дополнительных проводов.

Из прочих характеристик в первую очередь хотелось бы выделить прогрессивный и черезстрочечный режимы. Прогрессив обеспечивает более высокое качество на движущихся объектах и позволяет сжимать изображение в MPEG-1 (который черезстрочечной разверти не поддерживает). MPEG-4 — поддерживает, но путем хитрых извращений, которые реализованы не во всех энкодерах. Попытка сжатия черезстрочечного видео в MPEG-1 или MPEG-4 без поддержки прогрессива приводит к ужасным искажениям и существенно ухудшает

сжимаемость. В настоящее время камеры обоих типов практически сравнялись в цене, хотя в целом прогрессив стоит чуть дороже, но он действительно того стоит!

Камеры с аналоговым выходным сигналом меряются количеством телевизионных линий, которое обычно составляет 300, 420, 570 или 600. Естественно, чем больше — тем лучше, но и дороже. Камеры с цифровым выходом указывают разрешение в писелях — 320x240, 352x288, 640x480 или 704x576, которое на самом деле еще ни о чем не говорит, поскольку 352x288 могут соответствовать как 300, так и 420 фактическим линиям. Все зависит от того, в какую строну производитель производит округление — в большую или в меньшую. Впрочем, и "физические" линии аналоговых камер — величина чисто абстрактная, поскольку не учитывает разрешающей способности объектива, качество которого может варьироваться в очень широких пределах и, как показывает практика, практически у всех камер разрешающая способность объектива (штатного) ниже, чем матрицы, причем в некоторых случаях значительно ниже. Камеры со сменными объективами в этом смысле намного более предпочтительны, но и стоят они... скажем так, очень недешево, а хороший объектив стоит вдвое — втрое дороже самой камеры, особенно если речь идет о камерах с высоким разрешением.

Далее — камеры бывают черно-белые и цветные. Цвет — это, конечно, хорошо, вот только... нет, вопрос совсем не в том, что дорого. За цвет приходится платить снижением чувствительности, которая измеряется в люксах (что такое "люкс" рассказано в одноименной врезке) и которая у цветных камер часто падет до 0,1 lux, в то время как большинство черно-белых (даже из недорогой серии) спокойно дают 0,05 lux, что _намного_ больше, поскольку шкала — логарифмическая. Высокая чувствительность позволяет снимать в буквальном смысле слова даже при луне, что особенно актуально для камер наружного наблюдения. К тому же на камерах с инфракрасной подсветкой (а есть и такие) все равно никакого "цвета" не будет. С другой стороны, при низком качестве записи, цвет оказывается чуть ли не единственным источником информации и данные о том, что бандиты приехали на красной машине и были в зеленых куртках крайне важны, так что полностью отказываться от цвета не стоит и тут уже приходится выбирать — то ли покупать цветную камеру с высокой чувствительностью, то ли комбинировать высокочувствительные черно-белые камеры с низко чувствительными цветными.

Чувствительность очень тесно связана как с качеством объектива, так и с его фокусным расстоянием и диафрагмой. Обычно производители приводят чувствительность для минимального фокуса (т. е. широкого угла) с полностью открытой диафрагмой, поэтому при сравнении двух камер помимо чувствительности нужно так же смотреть на фокусное расстояние и диафрагму.

Очень важный (но не всегда упоминаемый в документации) параметр — скорость "затвора" (ES — Electronic Shutter Speed), измеряющая волях секунды и, в зависимости от модели камеры (и уровня освещения), колеблющаяся в интервале от 1/50 (и выше) до 1/100.000 (и ниже). И хотя никакого затвора в традиционном понимании этого слова в видеокамерах нет, его электронный эквивалент остался. Съемка движущихся объектов на 1/50 приводит к размазыванию изображения, которое будет тем больше, чем выше фокусное расстояние объектива. К сожалению, обычно из документации нет никакой возможности узнать реальные характеристики камеры и потому предсказать как она поведет себя при слабом освещении можно только исходя из цены. Дешевая высокочувствительная камера скорее всего просто имеет слишком длинную выдержку, подходящую только для съемок неподвижных объектов. Впрочем, высокая цена еще не гарант обратного.

один, два, четыре, двести шестьдесят четыре

По поводу качества различных алгоритмов сжатия идут непрекращающиеся споры, циркулируют нелепые слухи, мифы и легенды. Одни утверждают, что наименьшие потери дает MPEG1 (не поддерживающий черезстрочечной развертки, кстати), другие же говорят, что, дескать, MPEG1 — дрянь, поскольку, не поддерживаивает разрешений выше 352x240, в то время как MPEG2 легко "тянет" 1280x720.

На самом деле, эти цифры взяты из стандартов на Videodisk и DVD соответственно, где предельное разрешение определяется максимальным битрейтом, то есть количеством информации, которое гарантированно успевает прочитать привод в единицу времени. В видеокамерах таких ограничений нет, и MPEG1 легко "разгоняется" до 4095x4095 при битрейте в 100 Мбайт/сек, что полностью умещается в витую пару (правда, если туда будет "валиться" сигнал с нескольких камер — наступят полные тормоза, но ведь и разрешение 4095x4095 не обеспечивается даже самыми продвинутыми камерами наблюдения!).

Сравнивая различные стандарты сжатия друг с другом, многие обозреватели забывают, что в их основе лежит один и тот же алгоритм, основанный на **дискретном косинусном преобразовании** (*Discrete Cosine Transform* или сокращенно **DCT**), представляющим собой разновидность косинусного ортогонального преобразования для вектора действительных чисел.

В общем случае DTC-преобразование осуществляется умножением вектора на матрицу преобразования:

$$[Y] = [C] \times [X] \times [C^T]$$

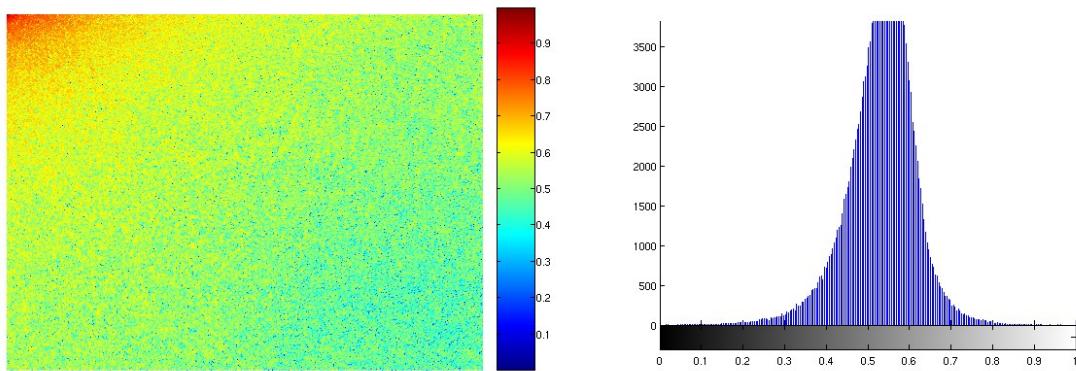
Формула 1 DTC преобразование, записанное в матричной форме, здесь: $[X]$ – начальная матрица, $[C]$ и $[C^T]$ – матрицы с коэффициентами преобразования, где C^T означает транспонированную матрицу

Алгоритмы MPEG1, MPEG2 и MPEG4 разбивают каждый кадр (фрейм) на блоки по 8x8 пикселей, над которыми осуществляется DCT-преобразование — сначала для каждой строки, а затем для каждого столбца матрицы, поэтому такое преобразование часто обозначается как DCT8 (число "восемь" означает 8 векторов).

После выполнения дискретного косинусного преобразования в Y-матрице уже нет пикселей и она превращается в совокупность волн с различными частотами и амплитудами. Низкие частоты, сосредоточенные в левом вернем углу матрицы, соответствуют крупным деталям исходного изображения. Высокие частоты, соответствующие мелким деталям, оккупируют правый нижний угол (см. рис. 1).



DFT



DCT

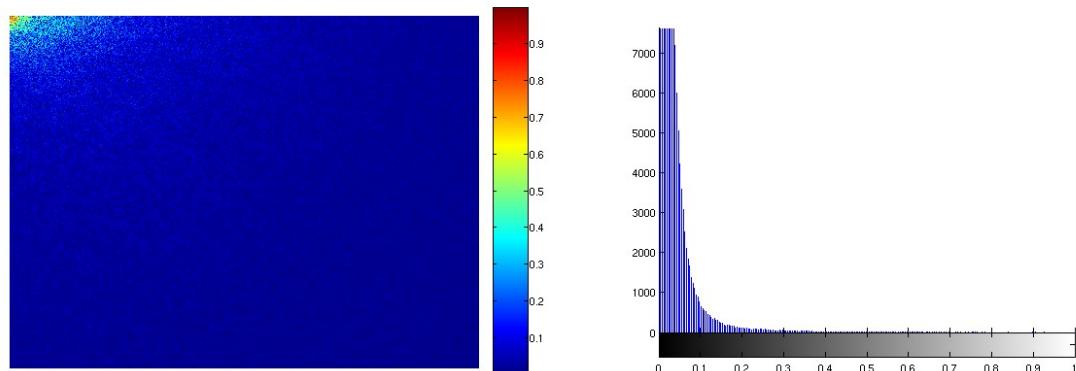


Рисунок 1 наглядная демонстрация DFT- и DCT-преобразований

Суть сжатия заключается в отсечении наименее значимых деталей, практически не воспринимаемых человеческим глазом, но кодируемые наиболее длинными кодами и соответствующие высоким частотам. Отсечение деталей в основном достигается за счет квантования.

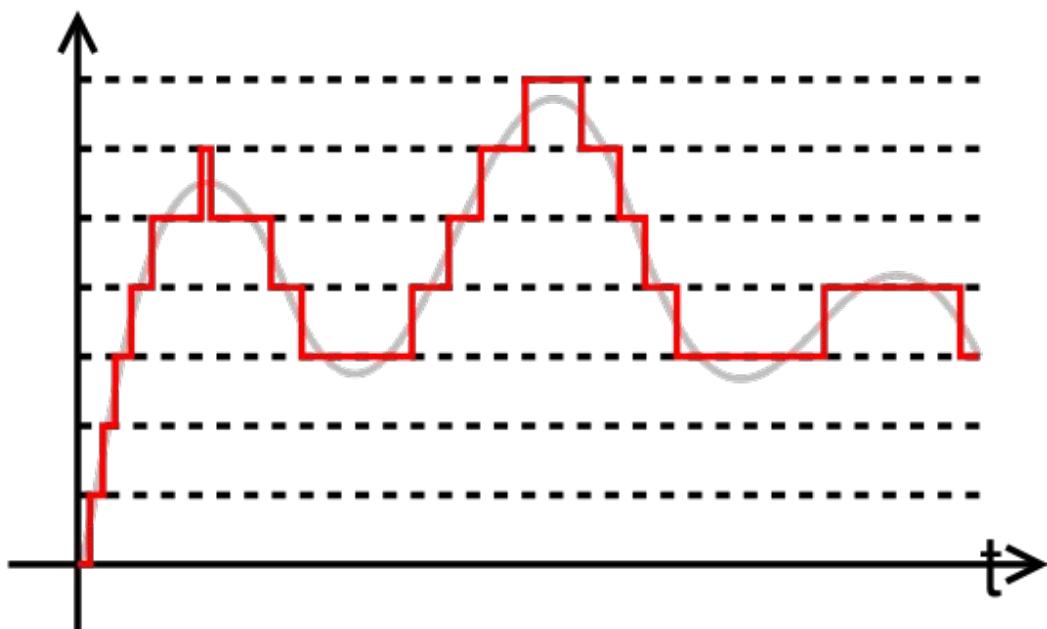


Рисунок 2 сигнал, переведенный в цифровую форму

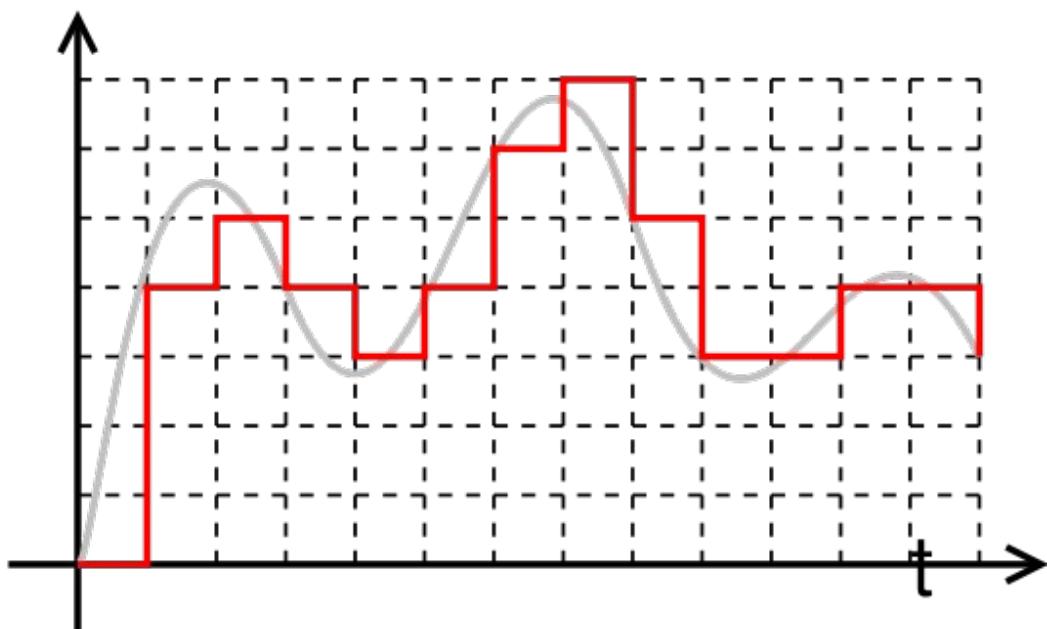


Рисунок 3 оцифрованный сигнал после квантования

Квантованием (применительно к обработке сигналов) называется деление величины сигнала на некоторое целое число, называемое **квантом (quant)**, обычно обозначаемый буквой Q , а обратный ему процесс — **деквантированием**. С математической точки зрения $X/Q = X*Q$, но целочисленная арифметика "огрубляет" сигнал (см рис 2, 3), причем это "огрубление" тем сильнее, чем больше квант. При $Q = 1$ мы не теряем никаких деталей, при Q превышающим половину амплитуды сигнала — теряем все. ("теряем все" — в пределах одного 8×8 блока, который заливается одним цветом, и мы получаем мозаику из 8×8 квадратов, из которой еще кое-что можно разобрать).

Потоковые MPEG-энкодеры, как правило, позволяют менять величину квANTA, которая обычно варьируется от 1 (наилучшее качество) до 31 (наихудшее качество). На самом деле,

31 — это очень большое число и уже при $Q > 6$ на изображение без содрогания смотреть невозможно. В некоторых (достаточно многих) энкодерах качество сжатия определяется в процентах, 100% соответствует Q1, 0% — Q31.

Ниже представлен ряд изображений, сжатых с различным качеством для сравнения:



Рисунок 4 изображение, сжатое с $Q = 1$



Рисунок 5 левая половина — $Q = 1$, правая — $Q = 6$



Рисунок 6 левая половина — $Q = 1$, правая — $Q = 13$



Рисунок 7 левая половина —Q = 1, правая — Q = 31

Теперь поговорим о цветовых пространствах. В большинстве видеокамер используется матрица с красными (Red), зелеными (Green) и синими (Blue) датчиками изображения. RGB-сигнал, снимаемый с матрицы, при преобразовании в PAL, преобразуются в сигнал яркости Y (он же Luminance или Luma) и два цветоразностных сигнала U и V (Chroma), вычисляемых по следующей формуле:

$$Y = 0.299R + 0.587G + 0.114B, \quad U = R - Y, \quad V = B - Y$$

Формула 2 перевод RGB-сигналов в YUV форму

Подбор коэффициентом обусловлен особенностями системы нашего, человеческого, цветоощущения, наиболее всего чувствительной к зеленой (точнее, желто-зеленой) области спектра, которому разработки видеосистем выделили наибольшее количество бит. То есть, даже если не использовать никакого сжатия, изображение все равно будет записано с потерями и качество видеокамеры всегда ниже, чем у фотоаппарата (к черно-белым камерам, работающим только с сигналом яркости, это не относится).

Цветовые пространства алгоритмов видео сжатия привязаны именно к такой схеме. Схеме цветоразностных сигналов. MPEG1 работает с цветовым пространством, именуемым YUV420 (оно же YUV12 и YV12), в которой значение яркости (Y) сохраняется для каждого пикселя, а цветоразностные компоненты (U&V) получаются путем усреднения значений 4'х пикселей из квадрата 2x2. На все компоненты отводится по 8 бит, в результате чего на кодирование одного пикселя расходуется 12 бит (чем и объясняется присутствие данных цифр в названии название). Главным недостатком подобной схемы становится не только низкое цветовое разрешение, но и невозможность работы с черезстрочечным изображением, поскольку объединение пикселей, взятых из соседних вертикальных линий (принадлежащих различным полукадрам) приводит к появлению значительных искажения.

MPEG2 использует цветовую схему YUV422 (она же YUY2), в которой сигнал цветности усредняется у двух соседних пикселей по горизонтали, в результате чего цветовое разрешение возрастает вдвое, а при сжатии черезстрочечного изображения не возникает искажений. К тому же на кодирование каждого сигнала отводится 16 бит.

MPEG4 позволяет использовать как ту, так и другую схему на наш выбор. Точнее, это по стандарту он предлагает на выбор, а в реальной жизни выбор зачастую уже сделан производителем энкодера, реализовавшим поддержку только одной цветовой схемы. Обычно такой схемой становится YV12, которая обладает более лучшим (на %35) цветовым сжатием по сравнению с YUY2. Самое поразительное, что некоторые производители "вставляют" YV12 даже в черезстрочечные камеры, оставляя ее единственной цветовой схемой! Вот еще один повод для отказа от камер со встроенными MPEG-энкодерами — никогда не знаешь какую пакость нам подстроили разработчики.

Считается, что черно-белых камер выбор цветовой схемы некритичен, поскольку, цвета у них нет. Цвета нет, это верно, но сигнал яркости есть, а в YV12 на него отводится на 4 бита меньше — выше сжатие, ниже качество. К тому же, цветоразностный сигнал все равно сохраняется. Некоторые MPEG4 энкодеры имеют специальный черно-белый режим, выбрасывающий цветоразностный сигнал. MPEG1/MPEG2 этого делать не умеют.

Вывод: MPEG4 — кот в мешке, MPEG2 — гарантированное номинальное качество.

>>> врезка сколько в тебе люкс

1 люкс (lux) равен освещённости поверхности площадью 1 м², при световом потоке падающего на неё излучения равном 1 Люмен (lm), который в свою очередь равен световому потоку, испускаемому точечным изотропным источником, с силой света равной одной канделе, в телесный угол, величиной в один стерадиан. А 1 кандела равна силе света, испускаемого в заданном направлении источником монохроматического излучения частотой 540·1012 герц, энергетическая сила света которого в этом направлении составляет (1/683) Вт/ср.

А слабо представить себе 1 люкс в реальности? Вот для этого и пригодится табличка, данная ниже.

освещенность, люксы	что создает
0,00005	свет звезд
0,0001	ночное безлунное небо в облаках
0,001	безоблачное ночное небо
0,01	луна в фазе первой/последней четверти
0,25	полная Луна в безоблачную ночь
10	свет свечи на расстоянии 30 см
50	освещенность жилой комнаты
400	яркий офисный свет
400	восход/закат на чистом небе
32000	солнечный свет в средний день (мин.)
100000	солнечный свет в средний день (макс.)

Таблица 1 освещенность в люксах, создаваемая различными источниками



Рисунок 8 кто следит за тобой?



Рисунок 9 инфракрасная камера слежения видит даже в темноте