



М.В. Руцков

Эксперт, к.т.н.

А теперь рассмотрим еще одно мультимедийное "чудо природы" – компрессию. Суть в том, что объемы видеоданных оцифрованного телесигнала настолько велики, что не лезут никуда – ни в каналы связи, ни на какие-либо сменные цифровые носители. Ну сами посудите, при формате 720x576 пиксел (в цвете) и скорости 25 кадр/с, получается поток порядка 20 Мбайт/с! Сначала был MPEG-1 – этакое баловство, уступающее по качеству легендарным VHS-кассетам. А потом "родили" формат MPEG-2, который практически стал стандартом записи на DVD-диск. Хотя там и есть некоторые рекомендации – суть самого кодека особо не меняется. Это дискретно-косинусное преобразование, которое выполняется над блоками 16x16 пиксел, используя межкадровую разность. Естественно формат MPEG-2 начали успешно использовать и в нашей области, хотя насчет качества можно еще и поспорить. Все-таки одна из основных функций охранного видеонаблюдения – просмотр стоп-кадров из архива.

Однако аппетит приходит во время еды! Захотелось закатывать "голливудское добро" и на обыкновенные CD-болванки, а еще гонять эти произведения кинематографа через Интернет (не безвозмездно, естественно). Сказано – сделано! Появился MPEG-4.

Так что ж это такое – стандарт или формат? На бытовом уровне – второе, а вот если использовать научный подход – первое. Само название MPEG происходит от Motion Picture Experts Group. Вот тут-то – в четвертой версии, в отличие от MPEG-1 и MPEG-2, "товарищи ученые" развернулись по полной программе.

Они решили написать рекомендации лет этак на 50 вперед – каким образом собирать в одну кучу разнородную мультимедийную информацию: видео, аудио, текст, графику и т.д. Причем начали фантазировать на темы того, чего еще и в помине нет. Ну например – выделение на статическом фоне неких объектов типа: человек, автомобиль, лошадь, грабли и т.д. Фактически прозвучало – мол, когда изобретете эти технологии, то вот вам наши рекомендации: как все это в единый мультимедийный поток запикивать! Отсюда самое глупое заблуждение по поводу "фантастиче-

Дурная наследственность-II

Статья печатается в авторской редакции

В предыдущем номере журнала "Системы безопасности" была опубликована статья М.В. Руцкова "Дурная наследственность". В своей статье, публикуемой ниже, он продолжает тему, затронутую ранее

ских" возможностей MPEG-4. Народ, прочитав эти рекомендации, причем не напрямую из самого документа, а из популярных статей некоторых горе-писателей - делает вывод, что в MPEG-4 уже реализованы не просто интеллектуальные видеодетекторы, а еще и распознавание объектов, в придачу!!! Если б это было так, то мы – видеодетекторщики уж давно бы отдохали, а юго-восточные системы гордо маршировали во все стороны!

На самом деле речь идет не о MPEG-4, а о видеокodeках, которые в рамках этого стандарта применяются. Кстати все MPEG-и -1,2,4 используют общие принципы кодирования. Например, метод компенсации движения и другие, но об-



этом чуть позже. Так вот, в первых двух используется дискретно-косинусное преобразование – выполняется аппроксимация внутри блока новыми функциями: 8x8 для MPEG-1 и 16x16 для MPEG-2. Отличия обусловлены разным уровнем вычислительной мощности для своего времени. А вот codeков в рамках MPEG-4 уже столько, что пальцев на руках и ногах не хватит сосчитать. Не буду вдаваться в тонкости (у всех свои секреты), а попробую выделить несколько общих штрихов.

Итак – компенсация движения! В чем смысл? Обыкновенная межкадровая разность уж не позволяет жать шибко сильно – поэтому и придумали сие творение! Сначала можно подумать – как круто: выделяется объект и транс-

сируется вдоль траектории. Да нет же конечно – просто берется область (квадратно-гнездовая), а далее начинается поиск "методом елозинея" – ей подобной, по минимуму или максимуму некой компарирующей функции. Думаю, даже корреляция не используется, в силу ее ресурсоемкости. Определив новую позицию такого квадрата, можно теперь лишь задать вектор смещения и разность. Самое интересное – учитываются только лишь линейные сдвиги! Ежели объект вращается или надвигается, а может быть меняет форму, да и яркость колыхнулась, то алгоритм уходит в глубокие раздумья и эффективность его резко падает. Однако в мультимедийных "фильмах" такое происходит редко – там мы следим за крупными объектами. Даже если они и поворачиваются, то "квадратики", их разбивающие, смещаются практически линейно. А теперь возьмем мелочевку – человек целиком вписывается в квадратик, да еще машет там руками и ногами! В результате все ломается и пропускается, особенно для codeков с постоянным потоком. Но никто этого не замечает – сюжетом увлечены, да и мозг человеческий в движении мелких деталей не видит! А вот в архиве можно лишь более-менее рассмотреть – опорные кадры. Причем некоторые "умельцы" сочиняют самодельные алгоритмы, в которых "опорники" расставлены в километре друг от друга. Тогда совсем – караул! Вот теперь и почувствуйте разницу, что такое – тупо тащить технологии из одной области в другую! Небо и земля!!!

Следующий "феномен" – сегментация. Идея заманчивая – зачем передавать по каналам связи пиксели. Давайте выделять на изображении области с примерно одинаковыми свойствами внутри – текстурами, и отсылать лишь их описание. Таким образом, степень компрессии возрастет еще больше! Естественно, получается – усредненная температура больных в больнице. Если более образно – текстура – это, например, что-то типа: "кирпичной кладки" или "травы", хотя до таких высот в codeках еще не дошли. После сего кодирования и соответствующего декодирования "кладка" замурует маленькую бойницу, из которого дуло торчало, а в "траве" потеряется кузнечик, сидевший там. И сожрала его не лягушка-прожорливое брюшко, а текстурная сегментация (зелененький он был, как и трава)!

Но и на этом мультимедийная индустрия останавливаться не желает. Наверное, все видели шпионские фильмы, в которых резидент и

центр общаются друг с другом шифровками. Причем у обеих сторон одинаковые книжки – передаются не слова, а их координаты (на какой странице, строке и внутри нее). Теперь вообразите, что у вас не книжка, а альбом с картинками – слабо!!! Живое видео подменяется мультимедиа в режиме реального времени! Даже трудно себе представить, к каким последствиям приведет перетаскивание такой "чудо-технологии" в охранное видеонаблюдение. Думаете – не додумаются. Ошибаетесь – еще как! Обязательно найдутся "умельцы", которым лень подумать, а может быть и нечем. Главное – чтобы костюмчик сидел! Вспоминаю времена "перестройки" – порошок стиральный куда-то пропал. И отлично на эту тему "Веселые ребята" из одноименной телепередачи приколотись – рекламный клип выдали. На нем вместо порошка выплывала коробка с "геркулесом". А что, все правильно – внешне не отличить!

Ну и, наконец, поговорим о модном ныне явлении – "предугадывающих" алгоритмах. В первую очередь их собираются использовать в "интеллектуальном" деинтерлейсинге для борьбы с "гребенкой". Задача сама по себе абсурдная – из двух полей, между которыми 20 мс, сделать единый стоп-кадр. А в перспективе – "дорисовывать" фрагменты на слабоконтрастных изображениях. Может быть для целей мультимедиа это и хорошо. Ведь все MPEG-и были задуманы для кинозрителей. Любая компрессия с потерями вносит искажения. Но и они бывают разными – в пределах удовольствия зрительного восприятия человеком и – раздражающи-

ми его! Обычно это раздражение возникает от шибко сильной компрессии, когда появляются, например – JPEG-лапти или та самая "гребенка". Ну и другие, скажем, "диагональные усы" от WaveLet-a. Поэтому цель видеокодеков – "успокоить" глаз человеческий, даже путем микрообмана, фильтруя все вдоль и поперек! Однако у нас – в "царской охране" – совсем другая задача – правда и только правда!!! Объясню – в отличие от Голливуда имеем мелкие и слабоконтрастные детали, которые в первую очередь страдают от компрессоров! И никакие "алгоритмические гадалки" не способны выдать истину! Немного нейрофизиологии и психологии зрения. Мозг (не глаз), причем не только у человека, построен на ассоциативном восприятии. Когда не хватает реальной информации, он сам ее достраивает, исходя из своего жизненного опыта. Сами знаете, что может померещиться в потемках или какие замысловатые "белогривые лошади" возникают в облаках. Это не случайно – это помогает выживать! Лучше десять раз в час зайцу шархануться в сторону, чем единственный и последний раз в жизни быть съеденным соевой, например. Это хорошо было бы для видеодетекторов, да и то в меру! Однако такие алгоритмы требуют по сути создания искусственного интеллектуального зрения с огромной базой "личного опыта"! Причем все очень индивидуально – милиционерам мерещится одно, военным – другое, матерям – третье и т.д. Поэтому, при фиксации видеособытий в охранной области, такие вероятностные "дорисовывания" – абсолютно недопусти-

мы!!! Можно и собаку Тузика нарисовать, и дуло "макарыча", или выдать кадр, в котором один человек тыкает другого не ножом, а бутылкой "Клинского"!

Поэтому сейчас можно лишь констатировать – никакие "интеллектуальные" алгоритмы пока не могут сравниться с интеллектом мозга. Дистанция огромна! Никогда не забуду своего общения с врачами-рентгенологами. Они смотрели на свои пленки, как на творения импрессионистов. И приходилось лишь диву даваться – как они умудрялись из этой каши пикселей выделять те или иные патологии!

Итак, сухой остаток:

1. MPEG-4 изначально предназначен для мультимедийных приложений.
2. В мультимедиа все крупно и контрастно! В охранном видеонаблюдении все – мелко и тускло!
3. В мультимедиа зритель смотрит кино, в охранном видеонаблюдении – стоп-кадры!
4. В мультимедиа главная задача – не раздражать глаз зрителя, пусть даже путем "художественного шаманства", в охранном видеонаблюдении – правда и только правда, эксперты смотреть будут!

Таким образом, делаю вывод – использование MPEG-4 в охранном видеонаблюдении чревато! Это все равно, что криминалисту выдать мобильный телефон с камерой вместо профессиональной фотоаппаратуры!!!

Ваше мнение и вопросы по статье направляйте на ss@groteck.ru

CNB
TECHNOLOGY Inc.

ЦВЕТНЫЕ ВИДЕОКАМЕРЫ ВЫСОКОГО РАЗРЕШЕНИЯ



CNB-GN730/GP730

- 1/3" SONY Super HAD CCD
- Разрешение – 550 твл
- Встроенный вариофокальный объектив (f=4~9 мм)



CNB-A1263

Разрешение 480 твл
220-ти кратный ZOOM



CNB-G1310P

Разрешение 550 твл
Чувствительность 0,3 лк



CNB-MP1310VD

Разрешение 550 твл
Встроенный вариофокальный объектив (f=4~9 мм)

ГалЭл
О П Т И М А Л Ь Н Ы Е
СИСТЕМЫ БЕЗОПАСНОСТИ

Москва, ул. Б.Почтовая,
д.34, стр.12, эт.2
тел.: (495) 221-66-22

www.cnb.ru