

Временной анализ 3D-видеопотока

Кулешов С.В.

Аннотация: В статье рассматриваются закономерности внутренней организации видеопотока. Введено понятие темпоральных кривых. Предлагается алгоритм оценки степени межкадровой избыточности видеопотока. Рассмотрен результат анализа реальных видеоданных, обработанных предложенным алгоритмом.

Ключевые слова: видеопоток, темпоральные кривые, компрессия данных.

С момента появления анимации и кинематографа как способа представления зрительной информации «видеопоток» понимался как последовательность неподвижных изображений, сменяющих друг друга через равные промежутки времени. Развитие технологий цифрового представления «движущихся картинок» потребовало расширить определение видеопотока с учетом появившихся механизмов его представления и обработки.

Определим видеопоток как последовательность изображений или их фрагментов, сменяющих или уточняющих друг друга по некоторому закону через заданные промежутки времени, представленная в виде, пригодном для транспортировки, сохранения или просмотра.

Особенностью видеоданных по сравнению с набором не связанных между собой изображений является наличие корреляционной зависимости между соседними кадрами. Такая зависимость является причиной избыточности в представлении видеопотока как последовательности отдельных изображений, причем избыточность между соседними кадрами (inter-frame) обычно многократно превышает избыточность представления внутри кадров (intra-frame).

Для детального анализа особенностей видеопотока и выявления в нем типовых закономерностей воспользуемся следующим представлением:

$$\{V_{i,j,t} \mid i = 1, 2, \dots, w; j = 1, 2, \dots, h; t = 1, 2, \dots, l\}$$

где V – пространственная матрица [1] видеопотока, w – ширина кадра, h – высота кадра, l – количество кадров.

Выполняя сечения ориентации (i) получаем набор двумерных матриц размера $j \times t$, каждой из которых соответствует свое изображение (рисунок 1).

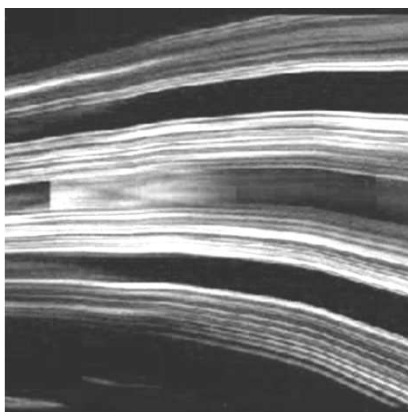


Рисунок 1. К понятию темпоральных кривых

Анализ полученных изображений (временных сечений) позволяет выявить ярко выраженные кривые, образованные неизменными пикселями в последовательных кадрах (темпоральные кривые). Одна из первых работ, в которой для представления видеопоследовательности используется идея «движения как направления» (motion as orientation) [2], была ориентирована на обнаружение движения. Особенность видеопотока, связанная с наличием в нем темпоральных кривых (temporal curves), была использована для создания группы методов компрессии видеоданных [3], но для повышения эффективности подобных методов требуется более детальное изучение поведения темпоральных кривых.

Длина темпоральных кривых зависит от типа контента в видеопотоке – так для систем видеонаблюдения (на которых подавляющая часть кадра остается неподвижной) характерны длины темпоральных кривых до нескольких минут (десятки тысяч кадров). В данном случае обрыв темпоральных кривых обуславливается шумом, изменением освещенности или движением мелких объектов (листва деревьев, трава и т.д.). В более распространенных типах контента (фильмы, телевидение) длина темпоральных кривых в основном определяется сменой сюжета и резким движением объектов.

Оценка средней длины темпоральных кривых L_{TC} на реальных видеоданных произведена в соответствии со следующим выражением:

$$L_{TC} = \sum \begin{cases} 1, & \text{если } |V_{x,y,z} - V_{x',y',z+1}| < \varepsilon, \text{ при } |x - x'| < 1 \wedge |y - y'| < 1 \\ & \text{иначе } 0 \end{cases}$$

В таблице 1 приведены характерные значения L_{TC} для основных типов видеоконтента для $\varepsilon = 1$

Таблица 1

Тип видеоконтента	Средняя длина темпоральных кривых
Музыкальные клипы	10-20
Фильмы «динамичные»	6-8
Фильмы «медленные»	8-14
Системы видеонаблюдения	>10000

Полученные результаты позволяют оптимизировать параметры межкадровой компрессии при сжатии видеопотока. Следует отметить, что некоторые алгоритмы компрессии видеоданных (например MJPEG) не используют возможность устранения межкадровой избыточности и далее не рассматриваются.

Анализируя строки направления (t) в пространственной матрице (рисунки 2-4) для видеопотока, полученного с системы видеонаблюдения с аналоговой черно-белой камерой, можно судить о характеристиках алгоритмов компрессии. Так, при временном анализе видеопотока с предельно низким значением битрейта (рисунок 4) видно нарушение структуры темпоральных кривых по сравнению с исходным видеопотоком.

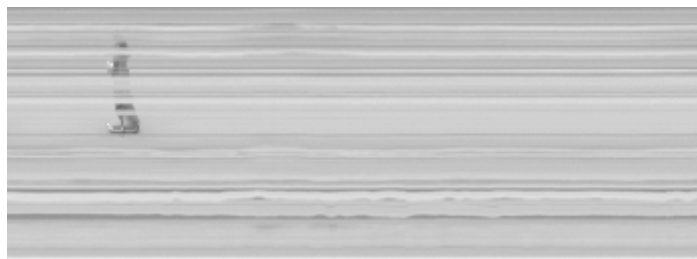


Рисунок 2. Видеопоток без компрессии



Рисунок 3. Видеопоток с компрессией MPEG2, скорость потока 4 Мб/с

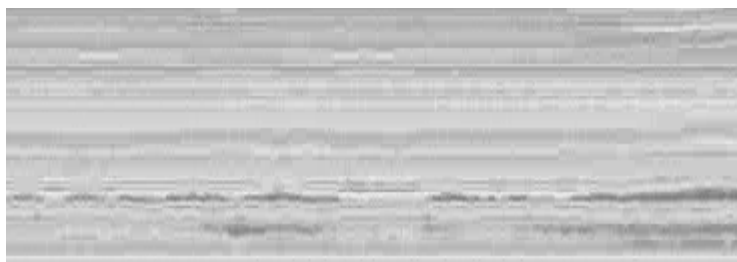


Рисунок 4. Видеопоток с компрессией MPEG2, скорость потока 0,5 Мб/с

Известен ряд методов анализа последовательностей кадров с целью определения направления и величины скорости движения объектов [4] и выделения фрагментов видеоданных [5].

В данной статье для оценки межкадровой (временной) избыточности предполагается использовать количество повторяющихся значений пиксела текущего кадра на каждом последующем кадре (рисунок 5).

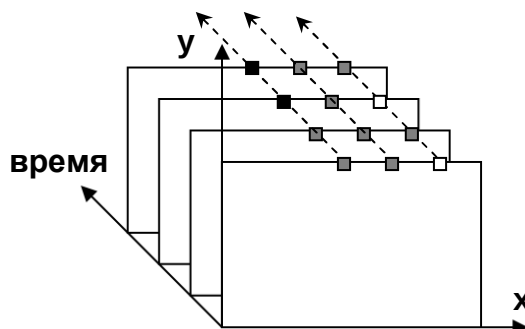


Рисунок 5. Иллюстрация идеи временного анализа последовательности кадров

Сформируем характеристическое изображение по следующему закону

$$H_{x,y} = \sum_{i=1}^{b-1} \begin{cases} 0, & \text{если } I_{k,x,y} \neq I_{k+1,x,y} \\ 1, & \text{если } I_{k,x,y} = I_{k+1,x,y} \end{cases},$$

где b - потенциальный размер блока для GOP-структуры [6] (для результатов, приведенных на рис. 7–9, использован $b=8$), $I_{k,x,y}$ - значение компонента пиксела с координатами x, y в k -ом кадре последовательности.

Примеры характеристических изображений для видеоданных, сохраненных с различными параметрами компрессии кодеками MPEG-2, STVC [3, 7] и h264, приведены на рисунках 7–9.

Для оценки степени структурной выраженности блоков используется алгоритм подсчета количества непересекающихся квадратных областей на характеристическом изображении, сформированном предложенным алгоритмом, содержащем точки с одинаковым значением.



Рисунок 6. Кадр из исходное видеопотка для экспериментов (аналоговая ч/б камера системы видеонаблюдения)

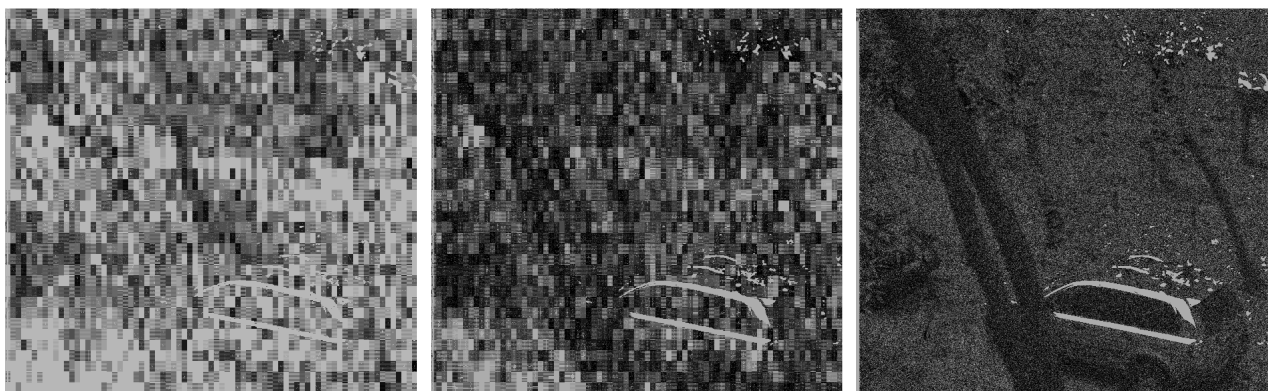


Рисунок 7. Результат работы алгоритма для видеопотока
(а – MPEG-2, поток 0,5 Мб/с; б – MPEG-2, поток 1 Мб/с; в – без компрессии)



Рисунок 8. Результат работы алгоритма для видеопотока
(а – h264, поток 0,15 Мбит/с; б – h264, поток 0,4 Мбит/с)



Рисунок 9. Результат работы алгоритма для видеопотока

(а – STVC, поток 0,14 Мб/с; б – STVC, поток 0,9 Мб/с; в – STVC, поток 0,6 Мб/с)

Полученные результаты позволяют по анализу изображений декомпрессированного видеопотока делать предположения об использованном при сжатии типе кодека и параметрах уровня его компрессии.

В случае применения кодеков группы MPEG характеристические изображения показывают количество и положение статических блоков, которые кодер не включает в поток в В- и Р-кадрах. Кроме того оказывается возможным определить положение ключевого кадра (I-кадра) в потоке.

Указанные параметры способны повысить эффективность работы с видеопотоками, компрессированными кодеками группы MPEG при выполнении операций модификации и повторной упаковки.

Литература

1. Соколов Н. П. Пространственные матрицы и их приложения. — М.: ГИФМЛ, 1960
2. E.H. Adelson, J.R. Bergen, Spatiotemporal energy models for the perception of motion, Journal of the Optical Society of America A 2 (2) (1985) 284-299.
3. Кулешов С.В. Пространственно-временное представление, обработка и компрессия видеопотока. — "Информационно-измерительные и управляющие системы", №4, т.6, 2008. — С. 3
4. Baron, J. L., Fleet, D. J., Beauchemin, S. S., Performances of optical flow techniques. 1994
5. Байгарова Н.С., Бухштаб Ю.А., Евтеева Н.Н., Корягин Д.А. Некоторые подходы к организации содержательного поиска изображений и видеоинформации — Препринт ИПМ им. М.В.Келдыша РАН №78, Москва, 2002.
6. IPTV dictionary —
http://www.iptvdictionary.com/iptv_dictionary_MPEG_GOP_definition.htm

7. Виртуальные и физические каналы цифровой передачи данных. О.В.Цветков. 6, 2007 г., Информационно-измерительные и управляющие системы, Т. 5, стр. 9-14.