

КРИТЕРИИ ОЦЕНКИ ДВИЖЕНИЯ ДЛЯ АЛГОРИТМОВ СЖАТИЯ ИЗОБРАЖЕНИЙ

При кодировании изображений содержащих движущиеся объекты, основной задачей является устранение временной избыточности. Эта задача решается, в том числе, с помощью оценки движения и передачи по каналу связи вместо самих блоков изображения их векторов движения.

Проблема поиска вектора движения $\vec{m} = (m_y, m_x)$ для текущего квадратного блока $\mathbf{V}_t(y, x)$ в момент времени t состоит в том, чтобы ошибка между блоком \mathbf{V}_t и соответствующим блоком \mathbf{C}_τ в момент времени τ была минимальной. Если используется оценка (предсказание) движения “назад” [1], то $\tau < t$. На практике диапазон изменения длины вектора движения \vec{m} ограничен и обычно на много меньше, чем размер его области поиска. Например, в кодеке по стандарту H.263 в штатном режиме работы компоненты вектора по двум координатам могут иметь величину между 16 и 15,5 пикселей. Так как текущий блок имеет размер 16x16 пикселей, это даёт ограничение области поиска в 47,5x47,5 пикселей квадратного блока $\mathbf{S}_\tau(y, x)$ или 48x48 пикселей с целыми значениями (эти блоки показаны на рис. 1).

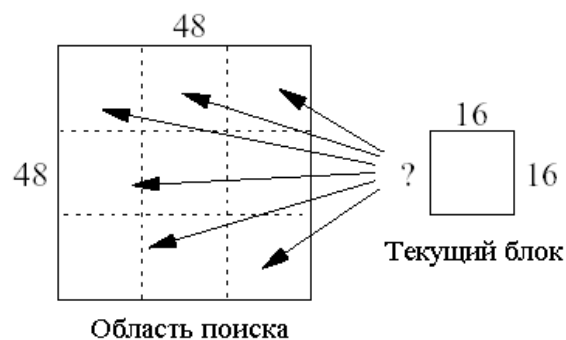


Рис.1

Обычно каждый пиксел в видеопоследовательности имеет три составляющие, которые содержат информацию о цвете и интенсивности. В основном только информация о яркости (интенсивности) используется для оценки движения. Кроме того, вектор движения вначале оценивается только с точностью до пиксела; текущий блок и предполагаемые блоки могут быть позже интерполированы с высоким

разрешением и вектор движения блока с наилучшим соответствием определяется с точностью до половины пиксела, как в стандарте H.263.

Перед использованием алгоритма оценки движения мы должны определить критерий, который измеряет качество оценки движения или определить, как вычислять ошибку оценки. В машинном зрении и визуальных системах наблюдения решается задача извлечения движения трёхмерных объектов, появляющихся в изображении. Это сложно, так как существует много затрудняющих этот процесс факторов. Однако, при сжатии изображений также существует задача что-либо упростить: для уменьшения числа бит необходимых для представления видеопоследовательности, и, в частности, может быть устранена избыточность изображения. В этом случае достаточно определить движение в соответствии с оптическим уравнением и далее эффективно устранить избыточность.

Рассмотрим несколько различных критериев; некоторые из них имеют большую сложность при оценке движения и высокое субъективное качество изображения, другие менее сложны. Простым критерием для оценки ошибки при поиске блоков является корреляция между двумя блоками:

$$\text{COR}(c_y, c_x) = \sum_{y=0}^{B_h-1} \sum_{x=0}^{B_w-1} \mathbf{B}_t(y, x) \mathbf{C}_\tau^{(c_y, c_x)}(y, x) \quad (1)$$

где $(c_y, c_x) = \vec{c}$ - текущий предполагаемый вектор движения, $\mathbf{C}_\tau^{(c_y, c_x)}$ - предполагаемый блок с координатами (c_y, c_x) относительно текущего блока. Для оценки вектора движения этот критерий может быть определён для вектора, имеющего максимальное значение корреляции.

Другой базовый критерий, среднеквадратическая ошибка (MSE), обычно дающий хорошие результаты, может быть вычислен как

$$\text{MSE}(c_y, c_x) = \sum_{y=0}^{B_h-1} \sum_{x=0}^{B_w-1} \left[\mathbf{B}_t(y, x) - \mathbf{C}_\tau^{(c_y, c_x)}(y, x) \right]^2 \quad (2)$$

Для хорошей оценки движения это выражение минимизируется.

Для возведения в квадрат в (2) требуется выполнение операции умножения, что приводит к большим вычислениям. Следовательно, в большинстве практических случаев используется критерий SAD (также часто называемый средней абсолютной разностью).

$$\text{SAD}(c_y, c_x) = \sum_{y=0}^{B_h-1} \sum_{x=0}^{B_w-1} \left| \mathbf{B}_t(y, x) - \mathbf{C}_t^{(c_y, c_x)}(y, x) \right| \quad (3)$$

Существует незначительное отличие в качестве изображения при использовании критериев MSE и SAD: соответствующие эксперименты были опубликованы в [2, 3].

В ряде случаев вычисление операции вычитания становится слишком неэффективным. В некоторых работах [3, 4] предложен критерий соответствия битовых полей (BPM):

$$\text{BPM}(c_y, c_x) = \sum_{y=0}^{B_h-1} \sum_{x=0}^{B_w-1} T[\mathbf{B}_t(y, x)] \oplus T[\mathbf{C}_t^{(c_y, c_x)}(y, x)] \quad (4)$$

где \oplus - оператор исключения T - функция, которая преобразовывает значение яркости пиксела в один бит, равный либо нулю, либо единице. В устройствах оценки движения реализованных в виде специализированных интегральных схем, это выполняется очень быстро, поскольку операция исключения может быть решена эффективно. В методе [3] используется простая пороговая функция

$$T(a) = \begin{cases} 0, & a \geq M \\ 1, & a < M \end{cases} \quad (5)$$

где M – средняя интенсивность изображения. В качестве альтернативного подхода можно преобразовывать изображение в два уровня, как обсуждалось в [4], где копия кадра фильтруется и затем выполняется однобитовое преобразование пиксела, таким образом производится сравнение разложенного на пикселы изображения с оригинальным изображением.

Критерий (4) даёт существенно более низкое качество, чем MSE или SAD. Вместо этого можно использовать критерий, который может быть вычислен как сумма двух различных критериев BPM с M_1 и M_2 , где среднее даётся через две различные области кадра (например, весь кадр \mathbf{F}_t и единичный блок \mathbf{B}_t) [3]:

$$\text{FBPM}(c_y, c_x) = \text{BPM}_1(c_y, c_x) + \text{BPM}_2(c_y, c_x) \quad (6)$$

Это называется соответствием особенных битовых полей (FBPM) и оно лучше, чем проще BPM, но на неподвижном изображении даёт более низкое качество предсказания в системе компенсации движения, чем SAD или MSE.

Согласно [5], критерий MSE (2) является более чувствительным к выбросам и предлагается использовать медианный метод

$$\text{MED}(c_y, c_x) = \text{med} \left\{ \left[\mathbf{B}_t(y, x) - \mathbf{C}_t^{(c_y, c_x)}(y, x) \right]^2 \mid 0 \leq y < B_h, 0 \leq x < B_w \right\} \quad (7)$$

в котором выбирается в качестве критерия медиана квадрата разностей интенсивностей пикселей в области рассмотрения. Также существует более устойчивый критерий [6] - функция Лорентца,

$$\text{LOR}(c_y, c_x) = \sum_{y=0}^{B_h-1} \sum_{x=0}^{B_w-1} \ln \left[1 + \frac{\left(\mathbf{B}_t(y, x) - \mathbf{C}_t^{(c_y, c_x)}(y, x) \right)^2}{2\omega^2} \right] \quad (8)$$

где ω регулирует величину возрастания разностей. Он устраняет эффект выбросов и по нему производится более достоверная оценка.

Другой критерий - соответствия количества пикселей MPC представлен в [7],

$$\text{MPC}(c_y, c_x) = \sum_{y=0}^{B_h-1} \sum_{x=0}^{B_w-1} T \left[\mathbf{B}_t(y, x), \mathbf{C}_t^{(c_y, c_x)}(y, x) \right] \quad (9)$$

где

$$T(a, b) = \begin{cases} 1, & |a - b| \leq \alpha \\ 0, & |a - b| > \alpha \end{cases} \quad (10)$$

α - заранее определённый порог. Критерий (9) считает число пикселей, значения которых находятся вокруг каждого соседнего; для получения оценки движения он максимизируется.

Много критериев ошибки, таких как (2), (3) и (8) могут быть обобщены в соотношение

$$\Phi(c_y, c_x) = \sum_{y=0}^{B_h-1} \sum_{x=0}^{B_w-1} \phi \left(\left| \mathbf{B}_t(y, x) - \mathbf{C}_t^{(c_y, c_x)}(y, x) \right| \right) \quad (11)$$

где $\phi(\varepsilon)$ - критерий оценки функции, когда задано абсолютное значение разности ε .

На рис. 2 изображены оцененные функции для нескольких критериев.

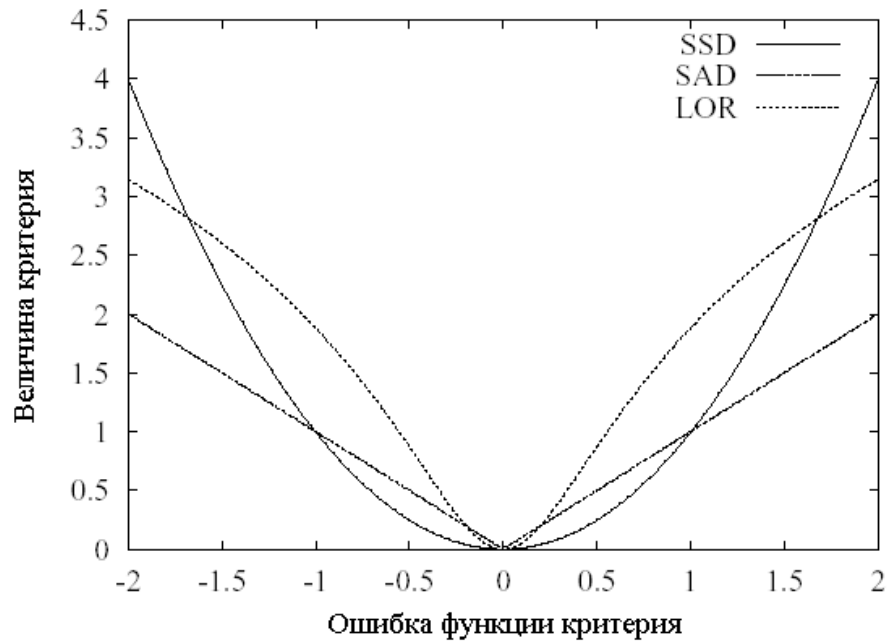


Рис. 2

Также существует относительно новый класс критериев оценки движения, разработанный в [8]. Большинство критериев пытаются минимизировать разность смещённых пикселей; однако, для получения лучшего качества кодированного изображения с заданной скоростью, также используются дополнительные биты для передачи векторов движения, объём которых должен быть минимизирован. Следующие критерии считают не только среднеквадратическую ошибку компенсации движения, но также биты, используемые для передачи:

$$C_{RD}(c_y, c_x) = B(c_y, c_x) + \lambda D(c_y, c_x) \quad (12)$$

где B означает число бит, необходимых для кодирования блока, включая вектор движения, D обозначает MSE остатка изображения с компенсированным движением, а λ - константа, полученная теоретически по зависимости скорость/ошибка.

Критерий MSE может быть использован непосредственно для вычисления $D(c_y, c_x)$; другой критерий может быть использован для его аппроксимации. С практической точки зрения эффективными считаются недавно созданные методы, минимизирующие скорость, что не только быстро, но может также иногда давать даже лучшее качество изображений, чем алгоритмы наилучшего поиска, которые пытаются только максимизировать значение оптического уравнения. Примерами являются “зональный поиск” и “ромбовидный поиск” [9,10], где существует тенденция к минимизации длины вектора движения и тем самым уменьшения количества битов, которые необходимы для кодирования.

До сих пор все представленные критерии использовались для вычисления в пространственной области. Некоторые из них могут быть также вычислены в области преобразования, например, как критерий MSE. Преимущество такого подхода состоит в том, что в общепринятых видеостандартах кодирования изображение будет преобразовано в любом случае (обычно с использованием ДКП). Критерий может быть также модифицирован так, что НЧ составляющие будут учтены с большим весом, что может дать лучшие результаты. С другой стороны, большинство преобразований переводят в пространственную область коэффициенты преобразования, но и имеют другие свойства, такие как изменение фазы на 180° в преобразовании Фурье. Следовательно, необходим критерий, который был бы адаптирован к особенностям преобразования изображения.

Каждый из представленных критериев в пространственной области требует $B_h \times B_w$ итераций, или 256 итераций для всех блоков размером 16x16 пикселей. Для критериев (1)-(4) каждая итерация требует одно сложение и кроме того, корреляция (1) требует умножение, MSE (2) требует сложение и умножение, а SAD (3) требует сложение и операцию получения модуля.

Таким образом, рассмотренные критерии оценки движения при кодировании изображений показывают, что существует некоторая взаимосвязь между скоростью вычисления выбранного критерия, величиной и количеством найденных векторов движения и, в конечном итоге, качеством восстановленного изображения. В ряде случаев достаточно использовать довольно простые критерии оценки движения, такие как COR, MSE или SAD. В других случаях необходимо учитывать содержание изображений и использовать более устойчивые критерии оценок. По-видимому, целесообразны дальнейшие исследования в области построения адаптивных критериев, которые будут учитывать специфику содержания изображения, моменты резкой смены сюжетов и искажения пространственной структуры изображения, возникающие при сжатии.

ЛІТЕРАТУРА

1. ISO/IEC 13818-2. Information Technology - Generic Coding of Moving Pictures and Associated Audio Information: Video. 1995.
2. Wang H., Mersereau R. Fast Algorithms for the Estimation of Motion Vectors. IEEE Transactions on Image Processing 8 (3). 1999.
3. Pourreza H., Rahmati M., & Behazin F. Simple And Efficient Bit-Plane Matching Algorithms For Video Compression. In: Workshop on Real-Time Image Analysis. Oulu, Finland. 2000.
4. Wong P., Au O. Modified One-Bit Transform for Motion Estimation. IEEE Transactions on Circuits and Systems for Video Technology 9 (7). 1999
5. Stiller C., Konrad J. Estimating Motion in Image Sequences - A tutorial on modeling and computation of 2D motion. IEEE Signal Processing Magazine 16 (4). 1999.
6. Konrad J. Motion Detection and Estimation. In: Bovik A. (ed.) Image and Video Processing Handbook. Academic Press. 1999.
7. Cheung C. Fast Motion Estimation Techniques for Video Compression. Ph.D thesis. City University of Hong Kong. 1998.
8. Hoang D., Long P., Vitter J. Efficient Cost Measures for Motion Estimation at Low Bit Rates. IEEE Transactions On Circuits and Systems for Video Technology 8. 1998.
9. Erol B., Kossentini F., Alnuweiri H. Efficient Coding and Mapping Algorithms for Software-Only Real-Time Video Coding at Low Bit Rates. IEEE Transactions on Circuits and Systems for Video Technology 10 (6). 2000.
10. Tourapis A., Au O., Liou M., Shen G., Ahmad I. Optimizing the MPEG-4 Encoder – Advanced Diamond Zonal Search. IEEE International Symposium on Circuits and Systems. Geneva, Switzerland. 2000. Vol. 3