

УДК 621.391:519.22

Мороз В. В., канд. техн. наук

## МОДЕЛИРОВАНИЕ СИСТЕМ ПЕРЕДАЧИ ВИДЕО С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ ПРОСТРАНСТВЕННОЙ И ВРЕМЕННОЙ ИНТЕРПОЛЯЦИИ

Одесский национальный университет им. И.И. Мечникова

*Описана модель системы передачи видеопоследовательностей с целью уменьшения объема передаваемой информации. Для этого задействованы механизмы временной децимации, понижения разрешающей способности кадров и статистического сжатия. Восстановление информации на стороне приемника выполняется методами пространственной и временной интерполяции. Предложена модель системы позволяет получать качественное визуальное воспроизведение видео на стороне приемника при минимальной загрузке канала передачи*

### **Введение**

При передаче цифрового видео по каналам связи с невысокой скоростью возникает проблема повышения их пропускной способности. Особенно остро данная проблема стоит перед разработчиками мобильных терминалов, которые готовы предоставлять различные сервисы, включая и потоковое видео, но скорость передачи каналов обычной сотовой связи не позволяет это делать качественно. Возможны два пути решения данной проблемы.

Первое направление связано с развитием высокоскоростных технологий передачи, но оно требует много времени и материальных затрат. К тому же, данное направление не зависит от провайдеров цифрового контента и производителей терминалов.

Второе направление – сжатие данных при их передаче по низкоскоростным каналам. Именно компрессия позволяет значительно увеличить пропускную способность линий при относительно небольших затратах на разработку специальных чипов для мобильных терминалов с соответствующим программным обеспечением.

При такой бизнес-модели, т.е. модели, в которой задействованы только производители терминалов и провайдеры цифрового контента, возможно решить поставленную задачу в кратчайшие сроки. С этой целью исходная видеопоследовательность на стороне передатчика (на стороне провайдера) перед передачей

подвергается сжатию, затем передается по каналу связи приемнику (на терминал связи потребителю видео контента) и после восстановления воспроизводится. Но, как известно, качество воспроизведения зависит от метода сжатия видео. Если использовать для сжатия видео методы, основанные на сжатии опорных кадров и межкадровой разницы без потерь информации, то степень сжатия будет небольшой. При использовании сжатия с частичной потерей информации, даже на небольших фреймрейтах (например, от 12 до 15 кадров/с.), степень сжатия увеличивается, но все же полученные объемы информации в сотни мегабайт не позволяют воспроизводить видео в режиме реального времени с приемлемым качеством. Т.о. даже самые эффективные современные методы сжатия видео не решают проблему.

Поэтому необходимо использовать дополнительные подходы к кодированию видео перед передачей. С этой целью перед передачей предлагается использовать методы временной децимации видеопоследовательностей с последующим понижением разрешения кадров. Затем, полученная промежуточная последовательность подвергается сжатию и передается по каналу связи. Это позволит увеличить скорость передачи в несколько раз. На стороне приемника промежуточная видеопоследовательность подвергается декодированию и последующей пространственно-временной интерполяции. В этой



[1].

*rithm)*

*(Block-Matching Algo-*

( — ),

Так как в пределах соседних кадров происходит незначительное движение объектов, то поиск подобных блоков выполняется в некоторой окрестности исходного блока, который должен быть найден.

Размеры области поиска определяют максимальную величину вектора движения, которая может быть измерена, поскольку этот параметр ограничивает максимальную скорость перемещения объекта, которую возможно зафиксировать. Так в [2] для интерполяции на невысоких битрейтах используется сужение области поиска на основе корреляции последовательных кадров, выбора размера области поиска в зависимости от смещения разницы блоков и информации о классификации блоков по предыдущим кадрам.

Размер области поиска и размер блока оказывают влияние, как на качество интерполяции, так и на вычислительную сложность таких методов. В работах [3] и [4] предлагаются алгоритмы, позволяющие снизить вычислительную сложность при приемлемом качестве с помощью различных подходов. При оценке движения блока большое влияние на поиск скоростей и качество прогнозирования кадров оказывают форма блока, его размеры и распределения векторов движения с несимметричными характеристиками. Для быстрой оценки движения блока предлагается новый алгоритм [3], который в качестве начального шага использует перекрестный поиск шаблона блока и большие/малые ромбовидные шаблоны на последующих шагах. В [4] используется адаптивный выбор порогового значения для согласования блоков в зависимости от общего числа макроблоков. Однако снижение вычислительной сложности таким образом приводит все же к нежелательным визуальным артефактам, особенно при сложном движении.

Одним из вариантов усовершенствования метода на основе ВМА является иерархическая пространственная корреляция. Поиск выполняется последовательно на нескольких уровнях с разными размерами блоков, а значения векторов движения, найденные на некотором уровне, используются в качестве исходных

для получения более точного результата на следующем уровне. Согласование блоков большого размера, с которых начинается поиск, выполняется за сравнительно небольшое количество вычислительных операций. Однако разрешение первого уровня невелико. Найденный вектор используется в качестве предсказания при поиске на следующем уровне с меньшим размером блоков, полученных делением блока первого уровня. Процесс повторяется несколько раз. На каждом последующем уровне увеличивается разрешение. Такой подход сочетает низкую частоту ошибок, характерную для блоков большого размера, и высокое разрешение, которое появляется при использовании малых блоков. Он позволяет находить большое количество векторов с высокой точностью и разрешением, причем для сложного движения. На первом этапе изображение может подвергаться низкочастотной фильтрации передискретизации с целью уменьшения числа отсчетов и сокращения объема вычислений. Низкочастотная фильтрация и передискретизация может быть заменена путем перехода к анализу информации в частотных доменах вейвлетного преобразования [5]. Первый из представленных в этой работе алгоритмов использует поиск ошибки согласования блоков в низкочастотном вейвлетном домене в меньшем масштабе, что позволяет повысить вычислительную эффективность алгоритма, а второй использует неоднородный двойной перекрестный поиск с использованием кратномасштабного анализа пространственно-временного контекста. Главный недостаток такого подхода заключается в том, что маленькие и быстро движущиеся объекты могут быть потеряны на первом этапе, когда используются блоки большого размера в сочетании с фильтрацией.

#### **Временная интерполяция и вейвлетный анализ**

Предложен метод повышения частоты кадров позволяет предсказывать недостающие кадры на основании двух (или в отдельных случаях более двух) ссылочных кадров. Основным моментом в предложенном методе является алгоритм пра-

вильного определения векторов движения блоков. Размеры блока поиска прямо влияют на разрешение и вероятность ложного согласования. Когда блок поиска представляет собой квадрат  $16 \times 16$  пикселей, он содержит довольно значительную часть изображения. Если блоки согласуются, то есть, похожи друг на друга в принятой метрике, то это с высокой долей вероятности говорит о том, что результат оценки правильно отражает смещение объекта. Возможность ложного согласования мала. Но для каждого блока из 256 пикселей находится всего один вектор, то есть разрешение невелико. Для малых блоков с размерами  $2 \times 2$  пикселей достигается высокое разрешение, но частота ошибок становится недопустимо большой, поскольку для блоков такого размера велика вероятность ложных согласований. В задачах сжатия оптимальным блоком, при котором достигается компромисс между разрешением и частотой ошибок, является блок размером  $8 \times 8$  пикселей. Но, для задачи временной интерполяции, ни разрешение, ни частота ошибок, соответствующие такому блоку, не являются приемлемыми.

Используя такие свойства вейвлетного анализа, как локализация частот и масштабируемость, можно варьировать размер блока в зависимости от масштаба изображения. Это, в свою очередь, позволяет на основе информации, находящейся в вейвлетных коэффициентах последовательных кадров, возмещать увеличение количества векторов движения.

Предлагается использовать два уровня вейвлетного преобразования последовательных кадров, т.е. семь частотных доменов. Из них  $LL_2$  – низкочастотный и по три высокочастотных домена на втором и первом уровнях –  $LH_2, HL_2, HH_2$  и  $LH_1, HL_1, HH_1$ .

Для начальной оценки движения используется низкочастотный домен второго уровня  $LL_2$ . Далее, векторы движения на нижних уровнях уточняются с использованием информации о движении, полученной на высоких уровнях разложения.

При этом возможно ложное предсказание из-за свойства сдвига вейвлетно-

го преобразования, которое является результатом процесса децимации [6].

Чтобы устранить присущие данному подходу недостатки, предлагается новый метод, основанный на алгоритме оценки движения, который использует особенности вейвлетных коэффициентов в каждом частотном поддиапазоне.

### Оценка движения на основе вейвлетных коэффициентов

Для двух ссылочных кадров выполняется двухуровневое прямое дискретное вейвлетное преобразование. Низкочастотные поддиапазоны  $LL_2$  разбиваются на блоки размером  $\left(\frac{N}{2^l}\right) \times \left(\frac{N}{2^l}\right)$  пикселей, где  $N$  – размер блока в кадре, а  $l$  – число уровней вейвлетного разложения. После этого производится классификация блоков на основе одной из выше упомянутых метрик. С целью уменьшения вычислительной сложности использовалась средняя абсолютная разность:

$$MAD = \frac{1}{\left(\frac{N}{2^l}\right)} \sum_{i=1}^{\left(\frac{N}{2^l}\right)} \sum_{j=1}^{\left(\frac{N}{2^l}\right)} |I_{LL}^{(1)}(i, j) - I_{LL}^{(2)}(i, j)|,$$

где  $I_{LL}^{(1)}(i, j)$  и  $I_{LL}^{(2)}(i, j)$  – значение вейвлетного коэффициента в положении  $(i, j)$  низкочастотного поддиапазона для первого и второго ссылочных кадров.

В зависимости от значения  $MAD$  для блоков, делается вывод о целесообразности уточнения их векторов движения. Т.е., если  $MAD$  меньше некоторого порогового значения, то данный блок имеет нулевой вектор движения и на нижних уровнях вейвлетного разложения может быть исключен из процедуры уточнения.

В виду того, что ошибки при определении векторов движения на высоких уровнях разложения могут быть распространены на остальные частотные поддиапазоны и, следовательно, будут накапливаться, оценка движения на высоких уровнях должна быть более точной. Поэтому для поддиапазона  $LL_2$  производится полный перебор блоков размера в восьмисвязной зоне поиска  $[-4; 4]$ . Впоследст-

вии, оцененный вектор движения используется как начальный для высокочастотных поддиапазонов этого же уровня в зоне поиска  $[-2; 2]$ .

Как отмечалось выше, блоки классифицируются в зависимости от того какой вектор движения они имеют: нулевой или ненулевой. Блоки с нулевым вектором движения могут не уточняться на последующих уровнях, а соответствующие им вейвлетные коэффициенты, заменяются соответствующими коэффициентами первого ссылочного кадра.

### **Методы пространственной интерполяции**

Пространственная интерполяция имеет довольно развитый математический аппарат, но основной проблемой ее использования в системах передачи видео остается проблема качества на этапе зрительного восприятия. Некачественная пространственная интерполяция до этапа временной интерполяции приводит к ошибкам в определении векторов движения, а после – усиливает видимость ошибочного предсказания векторов движения. Это проявляется в «зашумленности» кадров и дрожании объектов в кадре.

Наиболее популярным методом пространственной интерполяции считается бикубическая интерполяция, которая включена в большинство прикладных программ. Но прямая пространственная интерполяция, как и временная, не является приемлемым инструментом в данном случае. Это подтверждается невозможностью прямых пространственных методов эффективно восстанавливать высокочастотную часть спектра изображения. Поэтому большая часть малых деталей и границ объектов изображения не может быть восстановлена. Восстановленные изображения, как правило, не являются четкими и часто бывают размытыми. В этом состоит основной недостаток пространственной интерполяции.

Для решения качественной проблемы пространственного разрешения кадров предлагается использовать метод, основанный на итеративной экстраполяции ортогонального вейвлетного спектра изображения [7]. Этот метод также успешно

вписывается в модель временной интерполяции и окончательного сжатия межкадровой разницы, которые также основаны на алгоритме дискретного вейвлетного разложения.

### **Выводы**

Предлагаемая модель системы, в основу которой положены предложенные идеи и методы интерполяции, реализована и протестирована на нескольких видеопоследовательностях. Результаты сравнивались с предыдущей реализацией модели, в которой на этапе интерполяции использовались алгоритмы совпадения блоков и бикубической интерполяции. Количественная оценка предсказания кадров производилась на основе метрики *Y-PSNR* интерполированных кадров. Результаты показали, что предложенная модель превосходит и количественно (0.2 – 0.7 децибел) и качественно предыдущую модель. Это объясняется алгоритмами более качественной пространственной и временной интерполяции, позволившими устранить многие визуальные артефакты. Временная интерполяция на основе анализа частотных доменов вейвлетного преобразования дала более качественное предсказание кадров, а повышение разрешения кадров на основе итеративной экстраполяции ортогонального вейвлетного спектра после временной интерполяции привело к более качественной визуальной аппроксимации исходной последовательности. На этапе формирования битового потока был использован алгоритм сжатия, описанный в [8]. Использование данной модели позволило увеличить скорость передачи тестовых последовательностей видео от восьми до 12 раз при визуальном качестве сопоставимом со стандартом *H.264*.

### **Список литературы**

1. Гласман К. Преобразователи стандартов // «625». – М.: «Издательство 625», 2005, № 6. – С. 5 – 11.
2. Hwang-Seok Oh, Heung-Kyu Lee. "Block-matching algorithm based on an adaptive reduction of the search area for motion estimation", *Real-Time Imaging*, vol. 6, Issue 5, October 2000. – P. 407 – 414.

3. *Chun-Ho Cheung, Lai-Man Po.* "A Novel Cross-Diamond Search Algorithm for Fast Block Motion Estimation", IEEE Trans. Circuits And Systems For Video Technology, vol 12, no. 12, December 2002. – P. 1168 – 1177.

4. *Shou-Yi Tseng.* "Motion estimation using a frame-based adaptive thresholding approach", Real-Time Imaging, vol. 10, Issue 1, February 2004. – P. 1 – 7.

5. *Yu Liu, King Ngi Ngan.* "Fast multi-resolution motion estimation algorithms for wavelet-based scalable video coding", Signal Processing: Image Communication, Volume 22, Issue 5, June 2007. – P. 448 – 46.

6. *H.-W. Park and H.-S. Kim.* "Motion estimation using low-band-shift method for wavelet based moving-picture coding," IEEE Trans. Image Process., vol.9, no.4, April 2000. – P. 577 – 587.

7. *В.В. Мороз.* Комбинированный метод сжатия изображений // Сб. научн. тр. Одесск. национ. политехн. ун-та. – 2004. – Вып. 1(21). – С. 168 – 173.

8. *В.В. Мороз.* Эффективное кодирование вейвлетных коэффициентов с использованием адаптивного квантования // Сб. научн. тр. Одесск. национ. политехн. ун-та. – 2006. – Вып. 1(25). – С. 121 – 127.