

МЕТОД ФОРМИРОВАНИЯ БИНАРНОЙ КАДРОВОЙ МАСКИ ДВИЖЕНИЯ

Минаков Е. И., д.т.н., профессор кафедры радиоэлектроники Тульского государственного университета, e-mail: EMinakov@bk.ru;

Калистратов Д. С., аспирант кафедры радиоэлектроники Тульского государственного университета

Ключевые слова: маска движения, кадр, информация, кодирование изображений, видеокодек.

Введение

Задача разделения кадрового изображения на подвижную и неподвижную составляющие играет значимую роль в области видео кодирования, поскольку кодированию подвергается только подвижная часть и от того, насколько качественно она отделена от фона, в серьёзной степени зависит как время обработки кадра, так и качество декодируемого изображения. Визуально легко отделить подвижный объект от фона (рис. 1), однако для технического устройства задача распознавания движения осложняется многими факторами, в числе которых в первую очередь можно выделить неравномерность внешнего освещения и не идеальность применяемых технических устройств. Для обозначения подвижных частей изображения современные видеокодеки [2] используют бинарную кадровую маску движения, которая показывает наличие или отсутствие движения в каждом отдельно взятом пикселе текущего кадра. Неподвижные пикселы обозначаются логическими нулями, а подвижные – логическими единицами. Для вычисления маски движения используется текущий кадр и один или несколько дополнительных кадров, называемых ссылочными. В процессе компенсации движения кадровое изображение разбивается на фрагменты, называемые макроблоками.

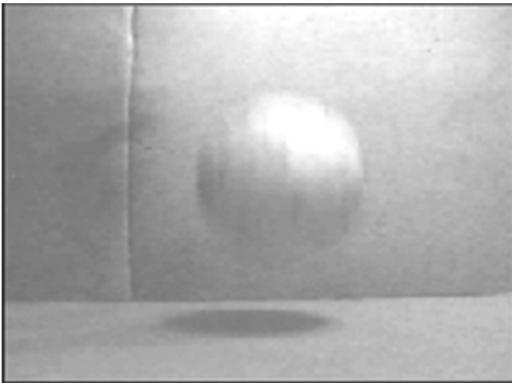


Рис.1. Кадр динамической видеопоследовательности

По маске движения проводится сортировка макроблоков, в результате которой часть блоков признаётся подвижной, а часть неподвижной. Для подвижных блоков кодером проводится поиск их наилучших прогнозов на ссылочных кадрах. Векторы смещений найденных прогнозов кодируются и передаются декодеру, чтобы тот по ним мог восстановить текущий кадр, имея в наличии

Рассматривается метод формирования бинарной кадровой маски движения на основе принципов контекстно-зависимого арифметического кодирования. Предлагается метод формирования бинарной кадровой маски движения на основе динамической регенерации подвижной части маски, проводится структурный синтез и сравнительный анализ моделей видеокодеков, использующих указанные методы.

предыдущий. Поиск хороших блочных прогнозов сопровождается большими временными затратами. В этой связи, в целях повышения временной эффективности видеокодека, актуальной является задача формирования качественной бинарной кадровой маски движения.

Формирование бинарной кадровой маски движения на основе принципов контекстно-зависимого арифметического кодирования

Современные видеокодеки в процессе формирования кадровых масок движения проводят предварительное построение маски по паре последовательных кадров с её последующей коррекцией на основе статистического анализа полученной маски. Руководствуясь принципами контекстно-зависимого арифметического кодирования [2], можно записать следующую формулу для предварительного формирования кадровой маски движения:

$$M_{0,i,j} = \begin{cases} 0, & \text{при } |Y_{t,i,j} - Y_{t-1,i,j}| = 0; \\ 1, & \text{при } |Y_{t,i,j} - Y_{t-1,i,j}| \neq 0, \end{cases} \quad (1)$$

$$(i = 0..(j = 0..cs - 1).ls - 1).$$

где: M – массив масок; Y – массив яркостных кадровых матриц с представлением их элементов относительными единицами яркости из диапазона [0..255]; t – индекс кадра; i, j – индексы пикселов в кадре; ls – количество строк в кадре; cs – количество столбцов в кадре.

Согласно этой формуле, каждый пиксел может быть признан подвижным или неподвижным в зависимости от абсолютного отклонения между значениями яркостей пиксела в текущем и предыдущем кадрах. Коррекция маски движения может быть проведена по формуле:

$$M_{1,i,j} = \begin{cases} 0, & \text{при } 0 \leq P_{1,i,j} \leq p; \\ 1, & \text{при } p < P_{1,i,j} \leq 1, \end{cases} \quad (2)$$

$$(i = 0..(j = 0..cs - 1).ls - 1),$$

где:

$$P_{1,i,j} = \frac{1}{S} \sum_{k=-1}^1 \sum_{l=-1}^1 (M_{0,i+k,j+l})$$

$$с ограничениями \begin{cases} ((i+k) > 0); \\ ((j+l) > 0); \\ ((i+k) < (ls-1)); \\ ((j+l) < (cs-1)). \end{cases}$$

$P_{1i,j}$ – массив вероятностей движения пикселей, значения которого лежат в вероятностном диапазоне [0..1]; p – порог вероятности движения, разделяющий вероятностный диапазон на два поддиапазона с границами [0.. p] и (p ..1]; S – количество соседей текущего обрабатываемого пиксела, включая сам пиксел; k, l – индексы соседних пикселей.

При проведении коррекции маски вычисляются вероятности движения пикселей в контексте движения их пространственных соседей. Конечное заключение относительно движения пиксела делается по попаданиям рассчитанных вероятностей в один из двух вероятностных поддиапазонов. Таким образом, каждый пиксел может поменять свой статус на противоположный. Результат работы алгоритма построения маски движения (рис. 2), использовавшего указанные формулы со значением порога вероятности движения равным $p = 0.5$ и количеством контекстных пикселей $S = 9$, представлен далее. Анализируя полученную маску движения, можно отметить, что эффект влияния помех и слабых отклонений параметров освещения не был устранён. Это означает, что при фрагментации изображения обнаружится много подвижных макроблоков, что, в свою очередь, потребует серьёзных затрат машинного времени на поиск их прогнозов. Уменьшение порога вероятности движения усугубляет действие помех, а при его увеличении вместе с зашумлёнными областями начинают выключаться из обработки действительно подвижные части изображения.

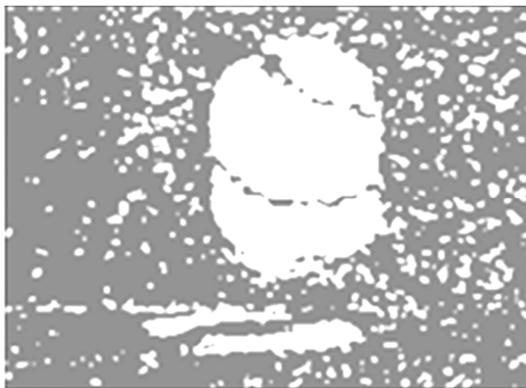


Рис. 2. Бинарная кадровая маска движения, полученная на основе принципов контекстно-зависимого арифметического кодирования

Формирование бинарной кадровой маски движения на основе динамической регенерации подвижной части маски

С точки зрения временной эффективности кодека представляет интерес метод, который без существенных потерь сможет выделить подмножество действительно подвижных пикселей и исключить из обработки подмножество пикселей, движение которых вызвано помехами и

слабыми отклонениями параметров внешнего освещения. Предлагаемый метод также использует начальное формирование маски с её последующей коррекцией. Стартовый вид маски движения определяется формулой:

$$M_{0,i,j} = \begin{cases} 0, & \text{при } |Y_{t,i,j} - Y_{t-1,i,j}| \leq L; \\ 1, & \text{при } |Y_{t,i,j} - Y_{t-1,i,j}| > L, \end{cases} \quad (3)$$

$$(i = 0..(j = 0..cs-1)..ls-1),$$

где L – условный максимум значений помех.

Для исследования эффективности предлагаемого метода использовалось значение этого максимума, равное $L = 25$ отн. ед. яркости. В соответствии с представленной формулой на этапе формирования начального вида маски действительно подвижными признаются только те пиксели, абсолютное отклонение сигналов которых от значений в предыдущем кадре превышает условный максимум помех. При этом (рис. 3) отсекаются как помехи, так и часть действительно подвижных областей кадра. Дальнейшая задача метода состоит в том, чтобы обеспечить динамическую реконструкцию подвижных частей кадра, исключив помехи.



Рис. 3. Начальный вид маски движения с учётом условного максимума значений помех

С этой целью организуется цикл коррекции маски движения, в котором маска пересчитывается заранее фиксированное число раз. Динамическая реконструкция маски ведётся по формуле

$$M_{n,i,j} = \begin{cases} 0, & \text{при } 0 \leq P_{1,i,j} \leq p; \\ 1, & \text{при } p < P_{1,i,j} \leq 1, \end{cases} \quad (4)$$

$$(n = 1..(i = 0..(j = 0..cs-1)..ls-1)..ns),$$

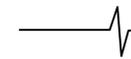
где

$$P_{n,i,j} = \frac{(k1 * P1_{n,i,j} + k2 * P2_{n,i,j})}{2}$$

с условием $k1 + k2 = 2$;

$$P1_{n,i,j} = \begin{cases} \frac{|Y_{t,i,j} - Y_{t-1,i,j}|}{L}, & \text{при } |Y_{t,i,j} - Y_{t-1,i,j}| \leq L; \\ 1, & \text{при } |Y_{t,i,j} - Y_{t-1,i,j}| > L \end{cases}$$

$$P2_{n,i,j} = \frac{1}{S} \sum_{k=-1}^1 \sum_{l=-1}^1 (M_{n-1,i+k,j+l})$$



$$\text{с ограничениями} \begin{cases} ((i+k) > 0); \\ ((j+l) > 0); \\ ((i+k) < (ls-1)); \\ ((j+l) < (cs-1)), \end{cases}$$

n – порядковый номер пересчёта маски; ls – количество пересчётов маски; $P1$ – массив вероятностей движения, определяемых по абсолютным отклонениям сигналов; $P2$ – массив вероятностей движения, определяемых по пикселям предыдущей маски; $k1, k2$ – весовые коэффициенты вероятностных компонентов.

С каждым последующим пересчётом подвижная часть маски начинает расширяться. Это расширение носит не хаотический, а целенаправленный характер. Суммарная вероятность вовлечения очередного пикселя в подвижное подмножество маски зависит как от абсолютного отклонения сигнала этого пикселя, так и от поведения его пространственных соседей. Приоритет между факторами определения движения регулируется весовыми коэффициентами вероятностных компонентов. В роли динамического параметра выступает порядковый номер пересчёта маски. Динамика восстановления кадровой маски движения (рис. 4) со значением порога вероятности движения равным $p = 0.1$, количеством пересчётов маски $ls = 10$ и при значениях весовых коэффициентов $k1 = k2 = 1$ показывает, что происходит существенное ослабление влияния помех на конечный вид кадровой маски движения.

Вместе с тем, наблюдается довольно качественная реконструкция подвижных частей изображения. Конечный вид маски движения после десятого пересчёта (рис. 5) предполагает сравнительно малое количество подвижных макроблоков при фрагментации изображения и, как следствие, существенное сокращение времени обработки кадра в целом.

Структурный синтез и сравнительный анализ моделей видеокодеков, использующих рассмотренные методы формирования бинарной кадровой маски движения

Для исследования эффективности предложенного метода формирования кадровой маски движения, на основе современных стандартов и принципов видеокодирования [1-3], были синтезированы две модели видеокодеков, идентичные по своей структуре (рис. 6). Источник (SRC) и приёмник (RCV) сигналов хранят последова-

$n = 3$



$n = 6$



$n = 9$



Рис. 4. Динамика восстановления кадровой маски движения



Рис. 5. Окончательный вид блочной кадровой маски движения после применения разработанного метода

тельности обрабатываемых кадров. Кодер образован следующей цепочкой элементов: прямой компенсатор движения (MCP), прямой дискретно-косинусный преобразователь (DCT), квантователь (Q), кодер серий переменной длины (RLE) и энтропийный кодер (VLE). В состав декодера входят: энтропийный декодер (VLE^{-1}), декодер серий переменной длины (RLE^{-1}), деквантователь (Q^{-1}), обратный дискретно-косинусный преобразователь (DCT^{-1}), обратный компенсатор движения (MCP^{-1}). Различия моделей заключаются только в способах формирования кадровой маски движения. Первая модель использует бинарную кадрową маску движения, получаемую на основе принципов контекстно-зависимого арифметического кодирования. Вторая модель использует бинарную кадрową маску, получаемую на основе предложенного метода динамической регенерации подвижной части маски.

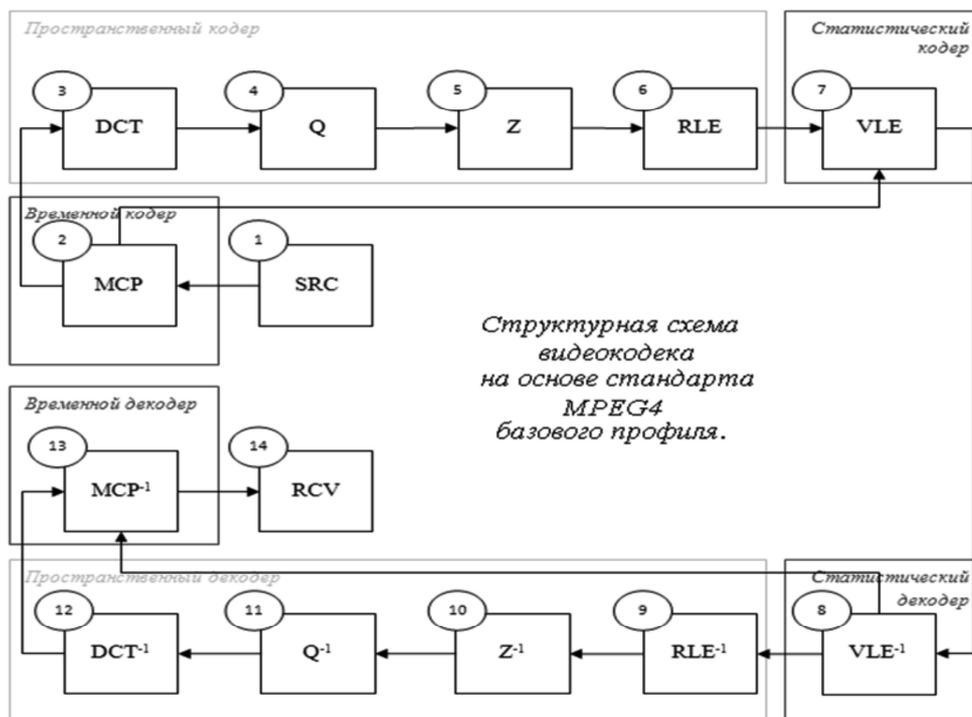


Рис.6. Структурная схема экспериментальных моделей видеокodeков

Для обеих моделей использовались одни и те же значения фундаментальных параметров видеокodeка (табл. 1). Формат кадра, размерность области поиска прогноза, шаг поиска прогноза и размерность макроблока задавались в пикселах. Шаг квантования задавался в относительных единицах яркости.

Таблица.1. Значения ключевых параметров использовавшихся моделей видеокodeков

Параметр	Значение
Режим цветности	чёрно-белый
Формат кадра	256 × 192
Размерность области поиска прогноза	24
Шаг поиска прогноза	2
Размерность макроблока	8
Шаг квантования	4

В ходе сравнительного анализа были получены следующие экспериментальные данные, характеризующие эффективность сравниваемых моделей видеокodeков по выборке из десяти кадров.

Таблица.2. Средние значения параметров качества сравниваемых моделей видеокodeков

Время обработки кадра, T, мс	Объём закодированного кадра, V, бит	Среднее абсолютное отклонение сигнала, E, отн. ед. яркости
Модель 1.	Модель 1.	Модель 1.
162	57138	2.32
Модель 2.	Модель 2.	Модель 2.
141	21342	3.11

По результатам экспериментов (табл. 2) можно сделать вывод, что предложенный метод формирования кадровой маски движения на основе её динамической регенерации обеспечивает существенно лучшие временные и информационные показатели эффективности видеокodeка по сравнению с существующими методами построения маски с применением принципов контекстно-

зависимого арифметического кодирования, но уступает им по качеству декодируемого изображения.

Заключение

В ходе исследования был рассмотрен метод формирования бинарной кадровой маски движения на основе принципов контекстно-зависимого арифметического кодирования, предложен метод формирования бинарной кадровой маски движения на основе динамической регенерации подвижной части маски, а также проведён структурный синтез и сравнительный анализ моделей видеокodeков, использующих указанные методы. Экспериментально было показано, что предложенный метод построения кадровой маски движения способен обеспечить большую временную и информационную эффективность видеокodeка по сравнению с существующими методами при незначительном снижении качества декодируемого изображения.

Литература

- Вернер М. Основы кодирования. – М.: Техносфера, 2004. – 288с.
- Ричардсон Я. Видеокodирование. H.264 и MPEG-4 – стандарты нового поколения. – М.: Техносфера, 2005. – 368с.
- Сэломон Д. Сжатие данных, изображений и звука.– М.:Техносфера, 2004. – 368с.

THE METHOD OF THE ALFA-MASK CONSTRUCTION

Minakov E.I., Kalistratov D.S.

This article presents methods of the alfa-mask construction. Existing and proposed methods are discussed here. It is also contains results of the comparative analysis of the videocodec's models, which use considered methods.