

К УЛУЧШЕНИЮ СЖАТИЯ ДИНАМИЧЕСКИХ ИЗОБРАЖЕНИЙ ПРИ MPEG-2 КОДИРОВАНИИ

Дворкович В.П., Соколов А. Ю.

Научно-Исследовательский Институт Радио, Москва 103064, ул. Казакова 16

Аннотация. Разработан новый алгоритм повышения сжатия при цифровой компрессии динамических изображений на основе пространственной передискретизации кадров, обеспечивающей улучшение компенсации движения. Пространственная передискретизация проводится только для участков изображений, подверженных изменению (движению) по отношению к предыдущим кадрам и сводится к перемещению деталей изображения в пределах долей пиксела, что практически незаметно для зрения. Для рассмотренных тестовых изображений метод позволяет на 4-7 % улучшить коэффициент сжатия.

1. Введение

Стандарт MPEG-2 является в настоящий момент одним из наиболее универсальных принятых международных стандартов для кодирования и передачи цифровых динамических видео изображений [1]. В этом стандарте сжатие исходного сигнала достигается за счет исключения временной избыточности сигнала между соседними кадрами и пространственной избыточности сигнала в каждом кадре. Исключение межкадровой избыточности производится с помощью процедуры компенсации движения. Для компенсации движения используется блочный метод, основанный на минимизации функционала энергии разностного сигнала между текущим и опорным кадрами для макроблока изображения фиксированного размера. Блочный метод анализа движения является вычислительно затратным, и для снижения вычислительной сложности были предложены различные быстрые алгоритмы. Следует упомянуть трехшаговый метод поиска векторов движения [2], иерархический метод [3]), логарифмический метод [4], метод с использованием оптического уравнения [5], метод фазовой корреляции [6, 7], предиктивный метод [8], метод, основанный на модели Маркова векторов движения [9], метод интегральных проекций [10], алгоритм с использованием быстрого преобразования Фурье [11], метод с использованием опорных точек макроблока [12, 13]. Данные алгоритмы характеризуются различной эффективностью компенсации движения и различной вычислительной сложностью. В [8] для поиска векторов движения макроблоков используется функционал энергии, содержащий дополнительный член, пропорциональный коду, необходимому для передачи векторов движения. Такой подход позволяет оптимизировать соотношение размеров кода между передаваемыми векторами движения и сигнала после компенсации движения, что улучшает компрессию. Проблема соотношения размеров кода между передаваемыми векторами движения и сигнала после компенсации движения рассмотрена теоретически в [14].

Помимо блочного алгоритма, в настоящее время разработано множество более эффективных методов компенсации движения. Новые алгоритмы могут использоваться при кодировании динамических изображений при больших степенях сжатия в таких приложениях, как видеотелефония и видео по Интернету, и являются основой для разработке новых международных стандартов компрессии динамических изображений. К данным методам следует отнести модель движения, описываемую проекцией трехмерных поворотов жесткого тела [15]. В [15] разработан объектно-ориентированный подход, позволяющий выделить в изображении объекты, характеризующиеся независимыми полями векторов движения. В [16, 17] рассмотрена модель, в которой гладкое поле векторов движения описывается с точностью до квадратичных членов от координат. При этом поле векторов является общим для достаточно большой области изображения, состоящей из соседних блоков размером 8x8, 16x16 и 32x32 пикселей. В работе [18] предлагается использовать память на несколько кадров для наиболее адекватного выбора опорного кадра для макроблока. В целом, указанные методы, в совокупности с улучшенными методами кодирования сигнала после компенсации движения, повышают сжатие при одинаковом качестве изображения, хотя и выходят за рамки стандарта MPEG-2.

Принципиальной особенностью перечисленных выше методов является то, что улучшение межкадрового сжатия достигается за счет усложнения модели движения. После компенсации движения в рамках стандарта MPEG-2 внутрикадровое сжатие приводит фактически к изменению истинных значений пикселей изображения (обусловленных квантованием коэффициентов спектра) без изменения положения этих пикселей в кадре. С другой стороны, можно незначительно изменить содержание кодируемого кадра таким образом, чтобы данное изменение было незаметно для глаза и улучшало компенсацию движения в рамках той или иной используемой модели. При этом дополнительное сжатие может быть достигнуто за счет небольших (в пределах долей пиксела) пространственных смещений деталей исходного изображения.

Поясним сказанное примером. На Рис.1 изображен одномерный импульсный сигнал, движущийся с постоянной скоростью вдоль оси ординат. Предположим, что данный сигнал должен быть закодирован и передан в рамках стандарта MPEG. При дискретности передаваемого вектора движения после компенсации движения значение разностного сигнала отлично от нуля (кривая 4), и требуется определенное количество бит для передачи разностного сигнала. С другой стороны, сдвинутый сигнал в пределах долей пиксела (кривая 3) визуально мало отличается от исходного, и не требует передачи дополнительной информации после компенсации движения. Поэтому выгоднее передать сдвинутый в пределах долей пиксела сигнал (кривая 3) вместо исходного сигнала (кривая 1).

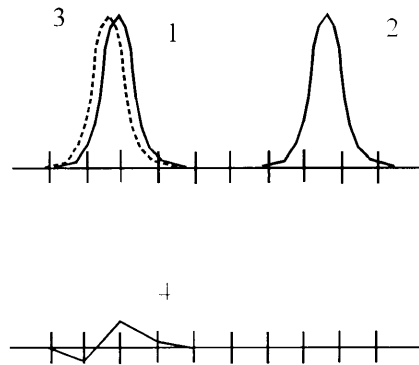


Рис. 1. 1: значение сигнала в текущий момент времени, 2: значение сигнала в предыдущий момент времени, 4: значение сигнала после компенсации движения с точностью до целых значений пикселей. 3: сдвинутое положение сигнала, обеспечивающее нулевой сигнал после компенсации движения.

Можно выделить две принципиально различные возможности пространственной передискретизации элементов изображения в предлагаемом способе:

- 1) произвольно любым способом в пределах чувствительности человеческого зрения к небольшим пространственным перемещениям пикселей;
- 2) определенным способом, который отвечает некоторой физической модели движения.

В качестве первого можно предложить такой алгоритм, в котором каждый пиксел изображения перемещается независимо от других пикселей так, чтобы значение разностного сигнала после компенсации движения в рамках данной модели движения было минимально для данного пикселя. Зрительное восприятие такого сигнала будет характеризоваться небольшим пространственным "дрожанием" пикселей, на которое должно быть наложено ограничение, чтобы дрожание было незаметно для человеческого зрения. Во втором способе, которого мы будем придерживаться в данной статье, пространственные перемещения соответствуют физическому движению объектов. Представим себе, что вместо отображения элементов изображения в строго определенный момент времени (совокупность таких элементов образует кадр изображения в обычном понимании) элементы изображения отображаются на экране декодера в моменты времени, отличающиеся друг от друга на небольшую временную величину, соответствующую разнице между истинным и передаваемым дискретным значением вектора движения (Рис.2). В таком понимании мы приходим к новой модели кадра изображения, элементы которого (в стандарте MPEG такими элементами могут быть макроблоки) имеют временную неопределенность в воспроизведении в пределах долей временного интервала между кадрами. Появление временной неопределенности, с одной стороны, оправдано свойством инерции человеческого глаза, а с другой стороны, расширяет возможности по улучшению компенсации движения в рамках заданного стандарта, и следовательно, увеличения сжатия.

В настоящей работе анализируется новый метод пространственной передискретизации элементов изображения в рамках стандарта MPEG-2, адекватный сформулированной модели. В дальнейшем мы ограничимся случаем, когда передискретизации подвергаются только В-кадры. Изменение пространственного положения объектов в этих кадрах, в отличие от кадров типа I и P, не будет влиять на компенсацию движения и кодирование остальной последовательности кадров. Поэтому предлагаемый способ должен всегда характеризоваться положительным эффектом по увеличению коэффициента сжатия.

Интересно отметить, что в рамках предлагаемой модели стандартный критерий искажений, определяемый по значению среднеквадратичного сигнала шума, становится больше неприменим. Действительно, после пространственной передискретизации значение среднеквадратичной ошибки может быть велико, несмотря на то, что сигнал отличается от исходного лишь незначительным сдвигом по координатам (см. Рис.2). Разработка адекватного критерия визуальной "заметности" вносимых искажений должна включать критерий пространственного отклонения деталей изображения от исходных значений в кадре. Можно предположить, что если, как отмечалось выше, данные изменения связаны с физическими перемещениями объектов, заметность искажений должна значительно уменьшаться вследствие инерции зрения. Разработка адекватного критерия качества является темой отдельного исследования и должна учитывать свойства инерции зрительного восприятия.

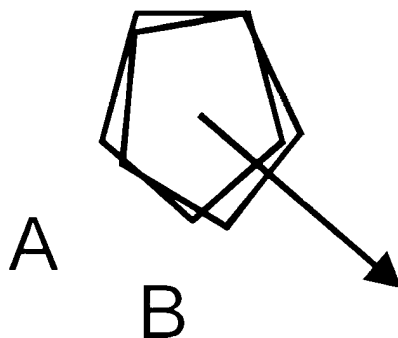


Рис.2 А - реальное положение объекта, В - отображаемое положение объекта.

2. Основные уравнения

Рассмотрим видеопоследовательность кадров, содержащую кадры типа В стандарта [1]. В случае кодирования макроблока В-кадра на кодере вычисляется величина разности

$$C_{x,y} = Y_{x,y}^{(n)} - X_{x,y}, \text{ где}$$

$$C_{x,y} = Y_{x,y}^{(n)} - X_{x,y} \text{ (при предсказании движения в одну сторону)} \quad (2.1)$$

$$X_{x,y} = \frac{1}{2} (Y_{x+V_{1x}, y+V_{1y}}^{(n-\Delta 1)} + Y_{x+V_{2x}, y+V_{2y}}^{(n-\Delta 2)}) \text{ (при предсказании движения в обе стороны)}. \quad (2.2)$$

В (2.1) и (2.2) $Y_{x,y}^{(n)}$ - значение сигнала яркости пиксела в макроблоке в кадре с номером n , имеющего пространственные координаты $\{x, y\}$. $X_{x,y}^{(n)}$ - предсказанное значение пиксела с учетом компенсации движения, Δ - расстояние до опорного кадра, $\{V_x, V_y\}$ - координаты вектора движения. При кодировании с векторами движения с точностью до полпиксела значения $\{V_x, V_y\}$ вычисляются с помощью линейной интерполяции [1].

Мы будем считать, что вектор движения макроблока известен и определен, например, методом полного перебора, и соответственно, значения $X_{x,y}^{(n)}$ вычислены. Сущность предлагаемого метода сводится к замене истинного значения $Y_{x,y}^{(n)}$ в (2.1), (2.2) некоторым другим значением $Y_{x,y}^{new(n)}$, соответствующим небольшому смещению изображения во времени. Целью замены является уменьшение разности $C_{x,y}$. Будем считать, что значения пикселей $Y_{x,y}^{new(n)}$ можно получить исходя из значений $Y_{x,y}^{(n)}$ с использованием интерполяции между пикселями:

$$Y_{x,y}^{new(n)} = Y_{x+\Delta V_x, y+\Delta V_y}^{(n)} \quad (2.3)$$

где $\Delta V_x, \Delta V_y$ - координаты вектора небольшого смещения макроблока.

Величина смещения $\Delta \vec{V} = \{\Delta V_x, \Delta V_y\}$ должна определяться из условия минимизации разностного сигнала:

$$\sum_{x,y} (Y_{x,y}^{new(n)} - X_{x,y})^2 = \min, \quad (2.4)$$

в (2.4) суммирование производится по всем пикселям макроблока. Пусть ориентация вектора $\Delta \vec{V}$ соответствует случаю, изображенному на Рис.3, на котором обозначено $Y_0 = Y_{i,k}$, $Y_1 = Y_{i+1,k}$, $Y_2 = Y_{i,k+1}$, $Y_3 = Y_{i+1,k+1}$. Для вычисления промежуточного значения сигнала яркости в точке $\Delta \vec{V}$ будем использовать интерполяцию по четырем соседним точкам. Предполагая функцию достаточно гладкой, в первом приближении можно считать

$$Y(\Delta \vec{V}) = Y_0 + \Delta \vec{V} \cdot \nabla Y + O((\Delta \vec{V} \cdot \nabla Y)^2),$$

∇Y - градиент функции. В пренебрежении членами второго порядка малости, имеем:

$$Y(\Delta \vec{V}) \approx Y^{(1)} = Y_0 + x_1(\vec{e}_1 \cdot \nabla Y) + x_2(\vec{e}_2 \cdot \nabla Y) = Y_0 + x_1(Y_1 - Y_0) + x_2(Y_2 - Y_0), \quad (2.5)$$

$$Y(\Delta\vec{V}) \approx Y^{(2)} = Y_0 + y_1(\vec{e}_1 \cdot \nabla Y) + y_2(\vec{e}_3 \cdot \nabla Y) = Y_0 + y_1(Y_1 - Y_0) + y_2(Y_3 - Y_0), \quad (2.6)$$

$$Y(\Delta\vec{V}) \approx Y^{(3)} = Y_0 + z_1(\vec{e}_2 \cdot \nabla Y) + z_2(\vec{e}_3 \cdot \nabla Y) = Y_0 + z_1(Y_2 - Y_0) + z_2(Y_3 - Y_0), \quad (2.7)$$

где координаты x, y, z определены как координатами вектора смещения в различных базисах:
 $\Delta\vec{V} = x_1\vec{e}_1 + x_2\vec{e}_2 = y_1\vec{e}_1 + y_2\vec{e}_3 = z_1\vec{e}_2 + z_2\vec{e}_3, \quad y_2 = x_2, \quad y_1 = x_1 - x_2, \quad z_1 = x_2 - x_1, \quad z_2 = x_1.$ (см. Рис. 3).

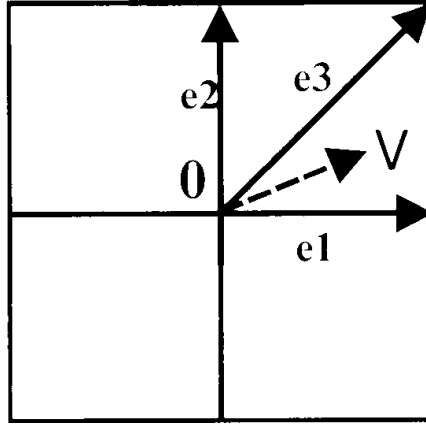


Рис.3. Положение действительной поправки к целочисленному значению вектора движения макроблока.

В (2.5)-(2.7) использовалась приближенная формула для производных: $\vec{e}_k \cdot \nabla Y \approx Y_k - Y_0$.

Для того, чтобы включить в интерполяционную формулу значения функции во всех 4-ех точках, рассмотрим усредненное между (2.5)-(2.7) значение с различным весом для каждого выражения,

$$Y(\Delta\vec{V}) \approx qY^{(1)} + \frac{1-q}{2}Y^{(2)} + \frac{1-q}{2}Y^{(3)} = \quad (2.8)$$

$$Y_0 + (qx_1 + \frac{1-q}{2}y_1)(Y_1 - Y_0) + (qx_2 + \frac{1-q}{2}z_1)(Y_2 - Y_0) + \frac{1-q}{2}(y_2 + z_2)(Y_3 - Y_0)$$

q и $(1-q)/2$ - веса, с которыми выражения (2.5)-(2.7) входят в конечную формулу. В последней формуле выбраны вследствие симметрии (см. Рис.3) одинаковые веса для выражений (2.6), (2.7). Выражая коэффициенты через декартовы координаты вектора x_1, x_2 , имеем

$$Y(\Delta\vec{V}) \approx Y_0 + \alpha_1(Y_1 - Y_0) + \alpha_2(Y_2 - Y_0) + \alpha_3(Y_3 - Y_0) \quad (2.9)$$

где

$$\alpha_1 = \frac{1+q}{2}x_1 - \frac{1-q}{2}x_2, \quad (2.10)$$

$$\alpha_2 = \frac{1+q}{2}x_2 - \frac{1-q}{2}x_1, \quad (2.11)$$

$$\alpha_3 = \frac{1-q}{2}(x_1 + x_2) \quad (2.12)$$

В (2.5)-(2.7) независимыми величинами являются только координаты вектора $\{x_1, x_2\}$, по которым необходимо произвести минимизацию в (2.4). С другой стороны, в (2.8) присутствует новый параметр q , поэтому минимизацию можно проводить по 3-ем независимым параметрам $\{x_1, x_2, q\}$. Нетрудно видеть, что в общем случае якобиан преобразования (2.10)-(2.12) $\partial\{\alpha_1, \alpha_2, \alpha_3\} / \partial\{x_1, x_2, q\} \neq 0$, поэтому величины $\{\alpha_1, \alpha_2, \alpha_3\}$ являются независимыми, и, следовательно, минимизацию можно проводить по величинам $\{\alpha_1, \alpha_2, \alpha_3\}$

После нахождения оптимальных значений $\{\alpha_1, \alpha_2, \alpha_3\}$ необходимо вычислить, какому вектору движения они соответствуют. Формулы обратного преобразования имеют вид:

$$x_1 = \alpha_1 + \alpha_3, \quad (2.13)$$

$$x_2 = \alpha_2 + \alpha_3, \quad (2.14)$$

$$q = \frac{\alpha_1 + \alpha_2}{\alpha_1 + \alpha_2 + 2\alpha_3}. \quad (2.15)$$

Координаты вектора движения $\{x_1, x_2\}$ определяются простыми выражениями (2.13), (2.14). Необходимо наложить условие, чтобы вектор движения находился внутри рассматриваемой области. Величины $\Delta V_x, \Delta V_y$ не должны превышать доли пиксела, $0 \leq x_1, x_2 \leq p$, откуда

$$0 \leq \alpha_1 + \alpha_3 \leq p, \quad (2.16)$$

$$0 \leq \alpha_2 + \alpha_3 \leq p, \quad (2.17)$$

где p - диапазон допустимого размера смещения элементов изображения при передискретизации, $0 \leq p \leq 1/2$. Величины $\{\alpha_1, \alpha_2, \alpha_3\}$ находятся из условия (2.4):

$$H = \sum (X_{0i} - Y_{0i} - \alpha_1(Y_{1i} - Y_{0i}) - \alpha_2(Y_{2i} - Y_{0i}) - \alpha_3(Y_{3i} - Y_{0i}))^2 = \min. \quad (2.18)$$

В (2.18) $X_0 = X_{i,k}$ - значение сигнала яркости в текущем кадре, суммирование производится по всем точкам макроблока (индексу i в (2.18)). Необходимыми условиями минимума являются

$\partial H / \partial \alpha_1 = \partial H / \partial \alpha_2 = \partial H / \partial \alpha_3 = 0$, откуда получаем линейную систему 3-ех уравнений для определения $\{\alpha_1, \alpha_2, \alpha_3\}$:

$$a_{11}\alpha_1 + a_{12}\alpha_2 + a_{13}\alpha_3 = -s_1, \quad (2.19)$$

$$a_{21}\alpha_1 + a_{22}\alpha_2 + a_{23}\alpha_3 = -s_2, \quad (2.20)$$

$$a_{31}\alpha_1 + a_{32}\alpha_2 + a_{33}\alpha_3 = -s_3, \quad (2.21)$$

где $a_{lm} = \sum (X_{0,i} - Y_{l,i})(Y_{0,i} - Y_{m,i})$, $s_l = \sum (X_{0,i} - Y_{0,i})(Y_{0,i} - Y_{l,i})$ ($l, m=1,2,3$).

Система (2.19)-(2.21) решается стандартными методами линейной алгебры. В вырожденном случае (детерминант матрицы равен нулю) передискретизацию проводить не будем (положим $\alpha_1 = \alpha_2 = \alpha_3 = 0$). Поскольку вектор смещения может иметь 4 различных ориентации (Рис.3), необходимо выбрать наилучшие значение среди значений, соответствующих 4-ем квадратам на Рис.3.

После определения вектора смещения для данного макроблока производится замена значений попиксельной яркости согласно (2.9).

При кодировании сигнала цветности необходимо учитывать ту модель цветности, которая используется в данном приложении. Например, при использовании распространенной модели 4:2:0 пространственное разрешение цветоразностных сигналов Сг и Сб в 2 раза меньше пространственного разрешения сигнала яркости. В этом случае передискретизация сигналов цветности производится с половинными значениями векторов движения, найденными выше, например (см. Рис.3):

$$Y_{Cr}(\Delta \vec{V}) \approx Y_{Cr}^{(1)} = Y_{Cr0} + \frac{x_1}{2}(Y_{Cr1} - Y_{Cr0}) + \frac{x_2}{2}(Y_{Cr2} - Y_{Cr0}) \quad (2.22)$$

$$Y_{Cb}(\Delta \vec{V}) \approx Y_{Cb}^{(1)} = Y_{Cb0} + \frac{x_1}{2}(Y_{Cb1} - Y_{Cb0}) + \frac{x_2}{2}(Y_{Cb2} - Y_{Cb0}) \quad (2.23)$$

3. Результаты экспериментов

Тестирование нового алгоритма осуществлялось на модели программного MPEG-2 кодера, полностью соответствующего стандарту [1]. Для сравнения результатов выбирались постоянные значения коэффициентов квантования для всех макроблоков в кадре. Выбор типа макроблока в В-кадре осуществлялся после процедуры пространственной передискретизации. Для этого пространственная передискретизация проводилась независимо для трех значений $X_{x,y}$ в (2.1)-(2.3), соответствующих типу кодирования макроблоков с предсказанием по времени вперед, назад, либо в обе стороны. Решение от выборе типа кодирования осуществлялось по минимальному значению мощности сигнала после компенсации движения:

$$f = \sum_{i,k} C_{i,k}^2. \quad (3.1)$$

Макроблок кодировался без компенсации движения [1], если $f > g$, где

$$g = \sum_{i,k} (Y_{i,k} - \langle Y_{i,k} \rangle)^2, \quad \langle Y_{i,k} \rangle = \frac{1}{256} \sum_{i,k} Y_{i,k} \quad (3.2)$$

При кодировании каждый 15-ый кадр являлся опорным, на каждый Р-кадр приходилось два В-кадра. Результаты вычислений приведены в Табл. 1,2 для различных тестовых последовательностей. Выбирались такие последовательности или их участки, в которых отсутствует смена сюжета.

Табл.1. Последовательность "Футбол", разрешение 352x240 пикселей, формат цветности 4:2:0, 105 кадров. 25 кадров/с.

3-я Международная Конференция DSPA-2000

Коэффициент квантования	Размер файла (байт)			Выигрыш по сжатию [%]		
	$p=0$	$p=0.25$	$p=0.5$	$p=0$	$p=0.25$	$p=0.5$
4	4717155	4539104	4433831	0	3.9	6.4
8	2665627	2575045	2508722	0	3.5	6.2
16	1365542	1337365	1313466	0	2.1	3.9

Табл.2. Последовательность "Теннис", разрешение 352x240 пикселей, формат цветности 4:2:0, 60 кадров, 25 кадров/с.

Коэффициент квантования	Размер файла (байт)			Выигрыш по сжатию [%]		
	$p=0$	$p=0.25$	$p=0.5$	$p=0$	$p=0.25$	$p=0.5$
4	1555862	1499434	1456107	0	3.8	6.8
8	830824	798822	773192	0	4.0	7.5
16	413488	403713	396726	0	2.4	4.2

Для рассмотренных видеопоследовательностей эффект дополнительного сжатия составляет порядка 4-7 % и зависит от максимально допустимого значения p вектора смещения, увеличиваясь с увеличением p . Визуальный анализ закодированных последовательностей показывает, что при $p=0.25$ искажения сигнала практически незаметны для глаза. При $p=0.5$ становятся заметными небольшие искажения, которые проявляются как слабое дрожание элементов изображения с частотой, равной частоте смены кадров.

На Рис.4 показаны восстановленные кадры последовательности "Футбол" при различных значениях p . Из Рис.4 видно, что статические изображения без и с применением нового алгоритма практически не отличаются друг от друга.



$p=0$

$p=0.25$

$p=0.5$

Рис. 4 Восстановленный кадр последовательности "Футбол" при различном значении максимального вектора смещения при передискретизации.

4. Заключение

Как следует из расчетов, метод передискретизации приводит к дополнительному эффекту сжатия порядка 4-7 % и возрастает с увеличением максимального значения допустимого вектора смещения. С другой стороны, при больших значениях вектора смещения ($p=0.5$) становятся заметными мелкомасштабное "дрожание" элементов изображения. Поэтому практически приходится ограничиваться значениями $p=0.25$, при котором эффект дополнительного сжатия незначителен (порядка 4%), хотя и положителен.

Литература:

- [1] ISO/IEC 13818-2. Information Technology - Generic Coding of Moving Pictures and Associated Audio Information. Part 2: Video./ Ed.1 JTS I/SC 29, 1994. Trans. Communi., Vol. COM-29, pp.1799-1808, Dec., 1981.
- [2] T. Koga et al. «Motion compensated interframe coding for video conferencing», National Telecom. Conf., pp. G.5.3.1-G.5.3.5., New Orleans, LA, Nov.-Dec. 1981.
- [3] L.D. Vos, "VLSI - architecture for the hierarchical block matching algorithm for HDTV applications", SPIE/VCIP, Vol.1360, pp.398-409, Cambridge, MA, Nov. 1990.
- [4] J.R.Jain, A.K. Jain, «Displacement measurement and its application in interframe image coding», IEEE Trans. Commun., Vol. COM-29, pp. 1799-1808, Dec., 1981.

- [5] A. M. Tekalp, Digital Video Processing, Prentice Hall PTR, Prentice-Hall, Inc, USA, 1995.
- [6] J. Watkinson. Textbook for Engineers on Motion Compensation. - Snell&Wilcox Ltd., 1994.
- [7] В.Г. Дворкович, В.В. Нечепаяев, А.Ю. Соколов, «Предварительная обработка изображения для анализа движения методом фазовой корреляции», в материалах 1-ой Межд. конф. «Цифровая обработка сигналов и ее применения», 30 июня - 30 июля 1998 г., Россия, Москва, МЦНТИ, т. III, стр.143.
- [8] F. Kossentini, Y.W. Lee, M.J.T. Smith, R/K/ Ward, «Predictive RD Optimized Motion Estimation for Very Low Bit-Rate Video Coding», IEEE Journal On Selected Areas in Communication., Vol.15, Num.9, Dec.1997, p.1752
- [9] S. Kim, C.-C. J. Kuo, "Fast motion vector estimation with a Markov model for MPEG", Proc. of The Int. Society for Optical Engineering, Digital Video Compression: Algorithms and Technologies 1995, 7-10 Feb. 1995, San Jose, California, Vol. 2419, p.210.
- [10] V.-C. Lin, S.-C. Tai, "Fast Full-Search Block-Matching Algorithm for Motion-Compensated Video Compression", IEEE Trans. on Communications, Vol. 45, No.5, May 1997, p. 527.
- [11] В.П. Дворкович, В.В. Нечепаяев, «Компенсация движения с использованием преобразования Фурье», в материалах 1-ой межд. конф. «Цифровая обработка сигналов и ее применения», 30 июня - 30 июля 1998 г., Россия, Москва, МЦНТИ, т. III, стр. 149.
- [12] А.В. Дворкович, В.П. Дворкович, Ю.Б. Зубарев, А.Ю. Соколов, «Способ анализа векторов движения деталей в динамических изображениях», Патент РФ № 2137194, от 15.07.98.
- [13] Ю.Б. Зубарев, В.П. Дворкович, В.В. Нечепаяев, А.Ю. Соколов, "Методы анализа и компенсации движения в динамических изображениях", Электросвязь, II, 1998 г.
- [14] G.M. Schuster, A.K. Katsaggelos. «A Theory of Optimal Bit Allocation Between Displacement Vector Field and Displaced Frame Difference», IEEE Journal On Selected Areas in Communication, Vol.15, Num.9, Dec.1997, p.1739.
- [15] S.-C. Han, J. W. Woods, "Adaptive Coding of Moving Objects for Very Low Bit Rates", IEEE J. on Selected Areas in Communications, Vol.16, No.1, Jan. 1998, p.56.
- [16] "Description of Nokia very low bitrate video coder/ Version 3.0", Proposal LBC97029 to ITU for advanced video coding, Nokia Research Center, Feb. 96.
- [17] «Core Experiment on Motion Compensated Prediction Using Quadtree Segmentation and Polynomial Motion Fields.» Contribution M1189 to ISO/IEC JTC1/SC29/WG11, October 1996.
- [18] T. Wiegand, N. Farber, Bernd Girod, "Error-Resilient Video Transmission Using Long-Term Memory Motion-Compensated Prediction", Proposal Q15H24 to ITU, August 1999.



ON THE IMPROVEMENT OF COMPRESSION OF DYNAMIC IMAGES CODED ACCORDING TO MPEG-2

Dvorkovich V.P., Sokolov A.Yu.

Radio Research & Development Institute, Kazakova str. 16, Moscow 103064

Abstract. A novel method of additional compression of dynamic images within the framework of MPEG-2 standard is proposed and investigated. It is proposed to take into account the non-integral part of motion vectors of the image elements (macroblock) by the replacement of the original frame by the frame in which the elements are shifted on the non-integral part of motion vectors. Taking into account that in MPEG-2 standard the motion vector accuracy is half of pixel element this shift will be less than the half pixel and almost not visible to human eye. Such displacement can be also considered as the displaying of picture elements at the varying times being equal to the constant temporal reference of the frame plus some addition corresponding to the motion with non-integral part of motion vectors. In this sense such displayed frame can be considered as the generalization of usual frame. Because the spatial (temporal) displacements correspond to physical motion of the elements it is believed that due to inertia of human eye it is almost not visible. On the other hand, the appearance of the additional degree of freedom allows to reduce the energy of motion compensated frames in MPEG-2 standard and thus achieve additional compression.

Because in MPEG-2 standard the coding of each P-frame is based on the context of previous P- or I-frame the proposed concept cannot be applied directly to P-frame coding, however, it can be used for coding of B-frames. The context of B-frame doesn't impact the coding of other frames, and thus, the replacement of B-frames by a frame in which macroblocks are shifted by some small values will not disturb the coding of other frames. Therefore, the proposed method will be characterized always with positive effect.

To find the non-integral displacements of macroblocks linear interpolation is used, in which case the problem is reduced to minimization of quadratic functional, corresponding to the energy of motion compensated signal in the macroblock.

The following Tables demonstrate the effect of such coding on MPEG-2 file size. In these Tables parameter p is equal to the maximum amplitude of spatial displacement of pixels in the macroblock.

Table 1. "Football", resolution 352x240 pixels, 105 frames, 25 fps

Quantizer scale	File size (bytes)			Effect [%]		
	$p=0$	$p=0.25$	$p=0.5$	$p=0$	$p=0.25$	$p=0.5$
4	4717155	4539104	4433831	0	3.9	6.4
8	2665627	2575045	2508722	0	3.5	6.2
16	1365542	1337365	1313466	0	2.1	3.9

Table 2. "Tennis", resolution 352x240 pixels, 60 frames, 25 fps

Quantizer scale	File size (bytes)			Effect [%]		
	$p=0$	$p=0.25$	$p=0.5$	$p=0$	$p=0.25$	$p=0.5$
4	1555862	1499434	1456107	0	3.8	6.8
8	830824	798822	773192	0	4.0	7.5
16	413488	403713	396726	0	2.4	4.2

Next Figure shows the restored B-frames for various values of p.



Fig.1 Restored frames from "Football".

In Fig.1 the displacements are not visible on the static image even for the value of $p=0.5$. However, when dynamic images are considered, the tests with various dynamic images show that maximum values of p should not exceed 0.25, otherwise the small tremble of picture elements becomes visible. For $p=0.25$ the proposed method provide only a small effect of additional compression (see Tables 1,2), which is, however, positive.