

СОЗДАНИЕ МЕТОДА СИНТАКСИЧЕСКОГО ПРЕДСТАВЛЕНИЯ СЕГМЕНТОВ ВИДЕОКАДРОВ С УЧЕТОМ НАЛИЧИЯ КЛЮЧЕВОЙ ИНФОРМАЦИИ

*Медведев Д.О., Баранник В.В, Сорокун А.Д., Стеценко О.Н.,
Додух А.Н.*

Введение

Рассмотрим разработку подхода относительно обработки сегментов ключевой информации. В данном случае предлагается использовать этап дифференциального представления [1].

Дифференциальное представление предлагается проводить для трансформант одномерного ортогонального преобразования.

С одной стороны такой подход позволит для высокоинформативных сегментов $S(Y)_\tau^{(k;l)}$ учитывать укрупненные корреляционные зависимости между несколькими элементами.

С другой стороны, техническая реализация такого подхода будет характеризоваться сокращением количества операций округления вещественных чисел, и иметь меньшую сложность по сравнению с реализацией двумерных ортогональных преобразований.

Это создает условия для использования в дальнейшей обработке более вычислительно сложных алгоритмов кодирования.

Построение метода синтаксического представления сегментов видеокадров

На этапе осуществляется аппроксимация столбцов массивов исходных видеоданных на основе базисов ортогональных преобразований.

В качестве базисов аппроксимации видеоданных используется дискретное косинусное преобразование (ДКП).

Это объясняется следующими достоинствами ДКП-преобразования: является наиболее близким к преобразованию Корнунэна-Лоэва (ПКЛ) в плане декорреляции (для большинства сигналов); обеспечивает лучшие характеристики при обработке оптических изображений, текстурных участков изображений, плавных переходов, сигналов гауссовской природы;

Выполнение одномерного ДКП-преобразования столбцов массива видеоданных реализуется на основе следующего выражения:

$$H(k, \ell) = F(k) X(i, j)_{k, \ell}, \quad (1)$$

где $X(i, j)_{k, \ell}$ – массив видеоданных, образованный на базе одной из плоскостей цветоразностной модели изображения; $H(k, \ell)$ – массив компонент одномерного преобразования, размерностью $q_\ell \times q_c$;

k, ℓ – соответственно индекс строки и столбца элемента массива $X(i, j)_{k, \ell}$, $k = \overline{1, q_\ell}$; $\ell = \overline{1, q_c}$; $F(k)$ – вектор дискретных значений базисных функций ДКП:

$$F(k) = \begin{cases} \frac{1}{\sqrt{q_\ell}} & \rightarrow k=1; \\ \sqrt{\frac{2}{q_\ell}} \cos \frac{(2q_\ell + 1)k\pi}{2q_\ell}, & \rightarrow k = \overline{2, q_\ell}. \end{cases} \quad (2)$$

С учетом системы выражений (2) для вектора базисных функций $F(k)$ соотношение (1) примет вид

$$H(1, \ell) = \frac{1}{\sqrt{q_\ell}} \sum_{k=1}^{q_\ell} X_{k\ell}^{(i,j)}; \\ H(k, \ell) = \sqrt{\frac{2}{q_\ell}} \sum_{k=1}^{q_\ell} X_{k\ell}^{(i,j)} \cos \frac{(2q_\ell + 1)k\pi}{2q_\ell}, \quad k = \overline{2, q_\ell} \quad (3)$$

где $X_{k\ell}^{(i,j)}$ – элемент массива $X(i, j)$, расположенный на пересечении k -й строки и ℓ -го столбца массива видеоданных.

После получения столбцов компонент одномерного преобразования проводится построение дифференциального представления на основе формирования величин разностей между соответствующим компонентами в строках.

Здесь учитывается наличие корреляционных зависимостей между компонентами строк одномерных трансформант дискретного косинусного преобразования (поскольку степень декорреляции сокращается только между элементами столбцов).

Выражение для получения дифференциального описания одномерных трансформант имеет вид

$$y'_{k1} = y_{k1} - y_{k+1,1}, \quad k = \overline{1, q_1}; \quad l = \overline{1, q_2}, \quad (4)$$

где y_{k1} , $y_{k+1,1}$ – компоненты трансформант на позициях k , $k+1$.

Формируется структурный массив Y' элементов дифференциального представления – структура одномерных трансформант дискретного косинусного преобразования (рис. 1).

На рисунке 1 y'_{k1} – компонента расположенная в массиве дифференциального представления (МДП) на пересечении k -й строки и ℓ -го столбца;

$(Q_1 Q_2 / q_1 q_2)$ – количество трансформант размером $q_1 \times q_2$, которое формируется для видеокadra, содержащего $Q_1 \times Q_2$ элементов.

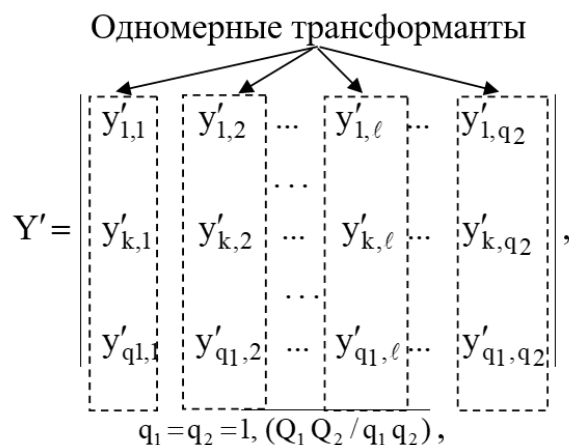


Рис. 1. Структура одномерных трансформант дискретного косинусного преобразования

Второй этап обработки заключается в формировании синтаксического представления структуры одномерных трансформант (СОТ). Синтаксическое представление дифференциального описания компонент одномерных трансформант связано с построением методов кодирования, обеспечивающих сокращение избыточности в СОТ.

Рассмотрим формирование структурной модели описания дифференциального представления одномерных трансформант.

Дифференциальное представление на основе использования одномерного дискретного косинусного преобразования в случае их оценивания в единой двумерной структуре имеет следующие особенности [1]:

1) значения крайних элементов СОТ, расположенных в левой части двумерной структуры зависят от средней яркости исходного сегмента видеокадра;

2) элементы левой области СОТ характеризуют степень насыщенности сегмента видеокадра низкочастотными характеристиками и уровнем семантической сложности. Сюда относятся плавные и ступенчатые изменения уровня яркости. Поскольку данные особенности являются преобладающими, то основная энергия сконцентрирована именно в этой области СОТ. Следовательно, элементам данной области двумерной структуры СОТ соответствуют наибольшие значения динамического диапазона;

3) значения элементов СОТ в правой области зависят от степени насыщенности блока изображения мелкими деталям и всплесками. Для средненасыщенных по уровню информативности сегментов видеокадров высокочастотные перепады являются не характерными. Это приводит к тому, что в данной области СОТ содержится меньшее количество энергии исходного сигнала, а соответственно элементы имеют наименьшие значения динамического диапазона;

4) дискретное косинусное преобразование использует укрупненную по всему одномерному блоку статистику данных, а значит и интегрированную взаимную корреляцию. Коэффициент взаимной корреляции будет близок к 0,9. Это позволяет для широкого класса сегментов видеокадров получать

трансформанты одномерного ДКП со следующими свойствами: компоненты трансформанты с большими значениями сконцентрированы в относительно малой длине столбца, наоборот компоненты с минимальными значениями занимают большую часть столбца. Это приводит к тому, что для значительного класса видеокадров количество элементов СОТ, имеющих низкие значения, будет превышать количество элементов СОТ, имеющих относительно высокие значения;

5) статистические характеристики видеоданных отличаются от гауссовских процессов. Тогда между компонентами трансформант одномерного dct-преобразования существует статистическая зависимость. Поскольку в процессе дифференциального описания учитывается только два соседних элемента, то статистическая зависимость будет существовать и между элементами СОТ.

Поэтому важными характеристиками элементов двумерной структуры СОТ являются:

- величина d_k текущей мощности алфавита компонент строки СОТ, равная разности между максимальным $y'_{k,max}$ и $y'_{k,min}$ минимальным значениями компонент в k -й строке, т.е. $(y'_{k,max} - y'_{k,min}) + 1 = d_k$;

- величина d_ℓ текущей мощности алфавита компонент ℓ -го столбца СОТ, равная разности между максимальным $y'_{\ell,max}$ и $y'_{\ell,min}$ минимальным значениями компонент ℓ -го столбца, а именно $(y'_{\ell,max} - y'_{\ell,min}) + 1 = d_\ell$.

Поскольку неравномерность текущей мощности алфавита характерна как для компонент строк, так и для компонент столбцов СОТ, то в общем случае выполняется неравенство $d_k \neq d_\ell$. Поэтому для снижения значения текущей мощности алфавита конкретной компоненты $y'_{k\ell}$ структуры СОТ необходимо учитывать дифференцированность значений мощности алфавита d_k по строке и d_ℓ столбцу. Другими словами требуется учитывать дифференцированность значений мощности алфавита конкретной компоненты в направлении строки и столбца. В этом случае текущая мощность алфавита компоненты $y'_{k\ell}$ будет равен величине $d_{k\ell}$:

$$d_{k\ell} = \min (d_k ; d_\ell). \quad (5)$$

Понятно, что в этом случае будут выполняться неравенства:

$$y'_{k\ell} < d_{k\ell}; \quad d_{k\ell} < d_k; \quad d_{k\ell} < d_\ell. \quad (6)$$

Отсюда можно допустить, что значения мощности алфавита компонент двумерной структуры СОТ с учетом их дифференциального синтаксического представления будут иметь дифференцированные и ограниченные значения в зависимости от координатных позиций в одномерных трансформантах.

Отсюда столбцы двумерной структуры, составленной из одномерных трансформант в дифференциальном описании могут рассматриваться как комбинаторный объект со следующей трактовкой.

Столбцы структуры СОТ представляют собой перестановку с повторениями элементы y'_{k1} , которой характеризуются ограниченным значением мощности алфавита с учетом двух направлений структуры СОТ.

Для варианта, когда минимальным уровнем является нулевой, абсолютные значения компонент l -го столбца будут принимать допустимые значения в следующем диапазоне:

$$0 \leq y'_{k\ell} \leq d_{k\ell} - 1; \quad k=\overline{1, q_1}. \quad (7)$$

Тогда количество допустимых столбцов структуры СОТ, составленных из q_1 -го элементов y'_{k1} , удовлетворяющих соотношению (7), будет находиться по такой формуле:

$$V_{q_1}^{(1)} = \prod_{k=1}^{q_1} d_{k,1}, \quad (8)$$

где $V_{q_1}^{(1)}$, - количество допустимых столбцов структуры СОТ, элементы которых удовлетворяют ограничению (7).

Тогда в соответствии с комбинаторной интерпретации столбцов двумерной структуры СОТ и выражению (8) количество информации, в среднем содержащееся в одной компоненте $y'_{k\ell}$, будет равно:

$$\bar{N}_1 = (\log_2 V_{q_1}^{(1)})/q_1 = (\log_2 \prod_{k=1}^{q_1} d_{k,1})/q_1 = (\sum_{k=1}^{q_1} \log_2 d_{k,1})/q_1, \quad (9)$$

где \bar{N}_1 - количество информации, приходящееся в среднем на одну компоненту столбца структуры СОТ в случае его комбинаторной трактовки в условиях наличия ограничений (7).

Согласно полученного выражения количество комбинаторной избыточности, которое минимально можно устранить в одном столбце структуры СОТ будет определяться разницей между количеством информации, приходящимся в среднем на одну компоненту y'_{k1} , до и после учета ограничений на величину текущей мощности алфавита.

Природа такой избыточности следует из того, что разница между количеством информации зависит от количества перестановок с повторениями, которое можно составить для различных ограничений на динамический диапазон.

Количество информации будет уменьшаться при понижении значения динамического диапазона. Поэтому раздельное представление компонент структуры СОТ в виде двух массивов: массива Y' абсолютных целочисленных значений компонент одномерных трансформант и двоичного массива G' их знаков (где $g_{k,l}$ (k,l)-й элемент массивы знаков)

$$G = \begin{vmatrix} g_{1,1} & \dots & g_{1,\ell} & \dots & g_{1,n} \\ & & & \dots & \\ g_{k,1} & \dots & g_{k,\ell} & \dots & g_{i,n} \\ & & & \dots & \\ g_{n,1} & \dots & g_{n,\ell} & \dots & g_{n,n} \end{vmatrix} \quad (10)$$

обеспечивает условия для повышения количества устраняемой избыточности.

Различное количество перестановок, а значит и наличие комбинаторной избыточности в структурах СОТ для разных вариантов ограничений на мощность алфавитов их компонент объясняется:

- неравномерными значениями элементов в столбце дифференциального представления;
- остаточной взаимной корреляцией между компонентами соседних столбцов структуры СОТ;
- взаимозависимостью компонент структуры СОТ (поскольку за счет одномерного преобразования и последующего дифференциального синтаксического описания достигается частичное разрушение только линейных зависимостей, а нелинейные зависимости не учитываются);
- неравномерностью распределения значений компонент столбцов двумерной структуры СОТ.

Разработка описания дифференциального представления изображений на основе неравновесных позиционных чисел

В общем случае возможны следующие подходы по распределению количества разрядов для кодового представления компонент столбца структуры СОТ, исключая предварительную оценку вероятностных характеристик и не использующих статистическое перераспределение количества двоичных разрядов [2].

Первый подход базируется на использовании информации о мощности d_{sut} алфавита всей двумерной структуры СОТ (structure of unidimensional transforms). Такой подход соответствует равномерной схеме распределения количества двоичных разрядов, т.е. проводится равномерная пакетирование интенсивности битового потока. Составляющая информационной интенсивности под каждую компоненту в этом случае будет определяться как $(\lceil \log_2 d_{sut} \rceil + 1)$. Соответственно суммарное количество $h_{q_1}^{(1)}$ двоичных бит на представление одного столбца Y_1' структуры СОТ будет оцениваться из условия:

$$h_{q_1}^{(1)} = q_1 (\lceil \log_2 d_{sut} \rceil + 1). \quad (11)$$

Основным недостатком такого подхода является то, что ограничивается уровень снижения информационной интенсивности вследствие высоких значений мощности алфавита, оцениваемой без учета локальности свойств

структуры СОТ. Здесь выделяемое усредненное количество $\bar{h}_{q_1}^{(1)}$ двоичных разрядов по одну компоненту столбца СОТ будет больше, чем величина \bar{N}_1 , рассчитанная по формуле (9), где $\bar{h}_{q_1}^{(1)} = h_{q_1}^{(1)} / q_1$. Понятно, что для данного варианта количество комбинаторной избыточности не будет устраняться в том количестве, которое предусматривается моделью описания столбцов двумерной структуры СОТ.

Для выхода из такой ситуации предлагается учитывать локальность свойств структуры СОТ, что определяет влияние на величину текущей мощности алфавита. Вариантом организации такого подхода является использование позиционных кодовых конструкций.

Здесь требуется учитывать, что структурные свойства двумерной структуры СОТ определяются режимом коррекции частотных составляющих под модель психовизуального восприятия зрительной системы. Параметром, регулирующим режим такой коррекции является величина R - фактор качества, влияющий на значения матрицы квантования (матрицы регулирования частотных составляющих трансформированных сегментов видеокадра).

В данном случае столбец Y'_ℓ его можно рассматривать как позиционное число $Y'_\ell = \{y'_{k\ell}, \dots, y'_{k\ell}, \dots, y'_{q_1, \ell}\}$. При этом система $\Psi'(R)_\ell = \{\psi'(R)_{1, \ell}, \dots, \psi'(R)_{k, \ell}, \dots, \psi'(R)_{q_1, \ell}\}$ оснований зависит от параметра R модели учета психовизуальных свойств зрительной системы, и будет выбираться с учетом значений мощности алфавита по каждому столбцу Y'_ℓ в рамках двумерной структуры СОТ. Тогда значения оснований будут находиться с использованием следующего соотношения:

$$\psi'(R)_{k, \ell} = d_{k, \ell}, \quad k = \overline{1, q_1}. \quad (12)$$

В этом случае кодовое значение $E'(R)_\ell$ позиционного числа Y'_ℓ определяется как:

$$E'(R)_\ell = y'_{1, \ell} v(R)_1^{(1)} + \dots + y'_{k, \ell} v(R)_k^{(1)} + \dots + y'_{q_1-1, \ell} v(R)_{q_1-1}^{(1)} + y'_{q_1, \ell}, \quad (13)$$

где $y'_{k\ell}$ - компонента расположенная на k -й позиции в ℓ -м столбце структуры СОТ; $v(R)_k^{(1)}$ - весовой позиционный коэффициент компоненты $y'_{k\ell}$; q_1 - количество элементов в столбце структуры СОТ.

Весовой коэффициент определяется как произведение оснований младших компонент столбца Y'_ℓ , т.е.:

$$v(R)_k^{(1)} = \prod_{\xi=k+1}^{q_1} \psi'(R)_{\xi, \ell}, \quad (14)$$

или через величину $V_{q_1}^{(1)}$, а именно:

$$v(\mathbf{R})_k^{(1)} = \frac{V_{q_1}^{(1)}}{\prod_{\xi=1}^k \psi'(\mathbf{R})_{\xi, \ell}} = \frac{\prod_{k=1}^{q_1} d_{k, \ell}}{\prod_{\xi=1}^k \psi'(\mathbf{R})_{\xi, \ell}}. \quad (15)$$

В этом случае количество двоичных разрядов $\bar{h}(\mathbf{R})_{q_1}^{(2)}$, которое требуется выделить в среднем под одну компоненту синтаксического представления кодового значения $E'(\mathbf{R})_\ell$ не будет превышать величины \bar{N}_1 , т.е. [3]:

$$\bar{h}(\mathbf{R})_{q_1}^{(2)} = h(\mathbf{R})_{q_1}^{(2)}/q_1 = ([\log_2 E'(\mathbf{R})_\ell] + 1)/q_1 < \bar{N}_1. \quad (16)$$

Это создает условия для сокращения количества комбинаторной избыточности, которое содержится в столбцах двумерной структуры СОТ.

Суммарное количество $h(\mathbf{R})_{q_1}^{(2)}$ двоичных разрядов на представление кодового значения $E'(\mathbf{R})_\ell$ одного столбца структуры СОТ определяется из условия:

$$h(\mathbf{R})_{q_1}^{(2)} = q_1 ([\log_2 E'(\mathbf{R})_\ell] + 1). \quad (17)$$

При этом величина $h(\mathbf{R})_{q_1}^{(2)}$, определяющая длину синтаксического представления одного столбца структуры СОТ будет зависеть от параметра R . В условиях, когда структура СОТ формируется для сегментов ключевой информации (по условию технологии обработки), то требуется обеспечить снижение информационной интенсивности битового потока с сохранением заданного уровня достоверности информации. Поэтому величина параметра R будет иметь ограниченные значения. В этом случае значения мощности алфавита компонент столбцов структуры СОТ будет увеличиваться, и как результат повысится значение величины $E'(\mathbf{R})_\ell$. Отсюда для обеспечения требований:

- с одной стороны относительно исключения вариантов переполнения технологической длины h_{it} кодовых слов необходимо ограничивать количество элементов в позиционном числе;

- с другой стороны относительно исключения маркерных разделителей между кодограммами -

предлагается формировать позиционные числа в пределах отдельных столбцов двумерной структуры СОТ с использованием системы оснований для выделения количества $h'(\mathbf{R})_{q_1}^{(2)}$ разрядов под каждую их кодограмму.

Другими словами, для обеспечения следующих условий:

$$E'(\mathbf{R})_\ell : \bar{h}(\mathbf{R})_{q_1}^{(2)} = ([\log_2 E'(\mathbf{R})_\ell] + 1)/q_1 \leq \bar{h}'(\mathbf{R})_{q_1}^{(2)} < \bar{N}_1;$$

$$\text{при } R \leq R_{\text{nes}};$$

требуется:

1) формировать позиционное число, длина q_{pn} которого выбирается в рамках количества q_1 компонент одного столбца структуры СОТ, т.е.:

$$q_{pn} = q_1;$$

2) определять количество разрядов на представление кодограмм на основе такой формулы:

$$h'(R)_{q_1}^{(2)} = ([\sum_{k=1}^{q_1} \log_2 \psi'(R)_{\xi, \ell}] + 1). \quad (18)$$

В этом случае будет выполняться неравенство

$$\bar{h}(R)_{q_1}^{(2)} \leq \bar{h}'(R)_{q_1}^{(2)} = ([\sum_{k=1}^{q_1} \log_2 \psi'(R)_{\xi, \ell}] + 1) / q_1 \leq \bar{H}_1,$$

где $\bar{h}'(R)_{q_1}^{(2)}$ - количество двоичных разрядов, приходящееся в среднем на одну компоненту столбца структуры СОТ в случае его синтаксического представления как позиционного числа в условиях использования системы оснований для позиционирования их кодограмм.

Согласно принятым условиям построение эффективного синтаксического представления для сегментов ключевой информации определяется следующими выражениями:

$$E'(R)_\ell = \sum_{k=1}^{q_1} y'_{k, \ell} v(R)_k^{(1)} = \sum_{k=1}^{q_1} y'_{k, \ell} \prod_{\xi=k+1}^{q_1} \psi'(R)_{\xi, \ell}; \quad (19)$$

$$h(R)_{q_1}^{(2)} = ([\log_2 E'(R)_\ell] + 1) \leq h'(R)_{q_1}^{(2)} = ([\sum_{k=1}^{q_1} \log_2 \psi'(R)_{\xi, \ell}] + 1); \quad (20)$$

$$\ell = \overline{1, q_2}.$$

Эффективное синтаксическое представление всего сегмента видеокadra ключевой информации в виде кодограмм кодовых значений $E'(R)_\ell$ отдельных столбцов Y'_l двумерной структуры СОТ, рассматриваемых как позиционные взвешенные числа с учетом коррекции частотных составляющих, задается следующим образом:

$$H'(R) = \{h'(R)_1, \dots, h'(R)_\ell, \dots, h'(R)_{q_1}\}; \quad (21)$$

$$E'(R) = \{E'(R)_\ell\}. \quad (22)$$

Здесь $E'(R)$ - вектор кодовых значений для двумерной структуры СОТ, в случае по-столбцовой обработки; $H'(R)$ - последовательность длин кодограмм кодовых значений $E'(R)_\ell$ в условиях использования для их кодообразования системы оснований. Постолбцовая схема обработки и построения синтаксического описания двумерной структуры СОТ обеспечивает снижение интенсивности битового потока с учетом особенностей ключевых сегментов относительно сохранения достоверности информации в условиях высоких значений мощности алфавита. Значит, соотношения (1) – (5) и (12) – (22) задают метод формирования эффективного

синтаксического кодирования сегментов видеокадров, содержащих ключевую информацию на основе формирования кодовых значений столбцов двумерной структуры СОТ, рассматриваемых как взвешенные позиционные числа с учетом коррекции частотных составляющих и последующего кодообразования с использованием системы оснований.

Выводы

Разработан метод формирования эффективного синтаксического кодирования сегментов видеокадров, содержащих ключевую информацию на основе формирования кодовых значений столбцов двумерной структуры СОТ, рассматриваемых как взвешенные позиционные числа с учетом коррекции частотных составляющих и последующего кодообразования с использованием системы оснований.

Постолбцовая схема обработки и построения синтаксического описания двумерной структуры СОТ обеспечивает снижение интенсивности битового потока с учетом особенностей ключевых сегментов относительно сохранения достоверности информации в условиях высоких значений мощности алфавита.

Литература

1. Баранник В. В., Тарасенко Д. А., Баранник Д. В., Медведев Д. О. Технология балансированной обработки динамического видеоресурса для снижения информационной интенсивности в инфокоммуникационных системах. *Безпека інформації*. 2017. №3. С. 163–170.
2. Баранник В. В., Тарасенко Д. А., Медведев Д. О., Хіменко В. В. Технологія обробки передбачених кадрів відеопотоку для бортових інформаційних технологій. *Наукоємні технології*. 2017. №4(36). С. 276–282.
3. V.V Barannik., Yu.N. Ryabukha, V.V. Tverdokhle, D.V. Barannik, “Methodological basis for constructing a method for compressing of transformants bit representation, based on non-equilibrium positional encoding”. 2nd IEEE International Conference on Advanced Information and Communication Technologies, AICT 2017, Proceedings, Lviv, 2017, pp. 188. doi: 10.1109/AIACT.2017.8020096