

## ТЕХНОЛОГИЧЕСКАЯ КОНЦЕПЦИЯ ДИФФЕРЕНЦИРОВАННОЙ ОБРАБОТКИ СЕКМЕНТОВ ВИДЕОКАДРА С УЧЕТОМ КЛЮЧЕВОЙ ИНФОРМАЦИИ

МЕДВЕДЕВ Д.О.

Обосновывается подход для формирования ядра технологической концепции дифференцированной обработки сегментов видеокadra с учетом ключевой информации. Это позволит: с одной стороны, снизить сложность процессов синтаксического представления; с другой – обеспечить заданный уровень достоверности видеoinформации, т.е. осуществлять режим обработки с контролируемой потерей качества реконструируемых видеокadров.

### 1. Введение

Для снижения информационной интенсивности битового потока в условиях сохранения требуемого уровня достоверности информации предлагается осуществлять дифференцированную обработку сегментов видеокadra. В этом случае обработка сегментов, содержащих ключевую информацию, будет проводиться с учетом сохранения заданной достоверности. Наоборот, обработка базовых сегментов, несущих фоновую нагрузку, предлагается с наибольшим снижением информационной интенсивности. Такая концепция позволяет сформировать дифференцированное синтаксическое представление сегментов видеокadra с учетом ключевой информации

### 2. Построение ядра технологической концепции дифференцированной обработки сегментов видеокadra

Ядром реализации предложенной концепции дифференцированной обработки с учетом ключевой информации необходимо выбирать класс методов, обеспечивающих возможность снижения информационной интенсивности с контролируруемыми параметрами. Такими параметрами в данном случае являются количество сокращаемой психовизуальной избыточности и вычислительная сложность обработки [1]. Значит, процессы компрессии и восстановления организуются с контролируемой потерей качества визуальной оценки видеoinформации. Под контролируемой обработкой с учетом модели психовизуального восприятия понимается возможность методов обеспечить требуемую достоверность реконструируемого видеокadra. В случае обработки сегментов ключевой информации требуется обеспечить режим без её потери.

Базовой составляющей методов данного класса является предварительная обработка, основанная

на использовании ортогональных преобразований, wavelet-систем и аффинных преобразований.

Технологии, базирующиеся на таких подходах, реализованы в форматах на JPEG-платформе (используется дискретное косинусное преобразование (ДКП)) и платформе JPEG2000 (используется wavelet-преобразование).

Данные методы позволяют формировать эффективное синтаксическое представление видеокadров как без потери качества визуальной оценки (режим loss-less), так и с потерей качества. Регулирование потерь качества и степени сжатия осуществляется путём: выбора типа и схемы трансформирования сегментов видеокadra; квантизации компонент трансформант; дополнительной регенерирующей фильтрации на приемной стороне.

В то же время для методов такого класса характерны следующие проблемные стороны:

1. В режиме ограниченного сокращения психовизуальной избыточности (режим сохранения требуемого уровня достоверности) проявляется существенная зависимость уровня снижения информационной интенсивности битового потока от степени информативности и семантической сложности сегментов цифровых изображений, в том числе изображений аэрокосмического мониторинга.

При этом степень снижения информационной интенсивности остается ограниченной.

Данные недостатки объясняются низкими значениями степени компрессии насыщенных изображений (в среднем до 2 раз) и дополнительными временными задержками на обработку. Здесь наибольшая сложность процессов обработки соответствует методам арифметического кодирования и методам LZW.

2. В режиме устранения психовизуальной избыточности проявляется зависимость уровня снижения информационной интенсивности от качества визуальной оценки реконструируемых видеокadров.

Наибольший уровень ухудшения визуальной оценки характерен для высокоинформативных фрагментов видеокadров. Здесь наблюдается наиболее быстрое падение качества и достоверности видеoinформации с ростом уровня снижения информационной интенсивности [2].

Это ограничивает возможность эффективного использования таких технологий в энергоэффективных телекоммуникационных системах. Данные проблемные стороны функционирования методов с контролируемыми параметрами обработки обусловлены следующими причинами:

1) ограниченная степень снижения информационной интенсивности для методов на платформах JPEG и JPEG 2000 при режиме  $h \geq 35$  дБ. Наибольшее снижение степени компрессии проявляется при обработке высоко- и средне-насыщенных цифровых изображений;

2) реализация двумерных ДКП и dwt-преобразований для бортовых вычислительных комплексов связана с необходимостью затрачивать от 10 до 70% временных затрат от суммарного времени на обработку. Время на их выполнение в процессе преобразования изображений может достигать десятков секунд. Это обусловлено выполнением большого количества операций умножения в пространстве как целочисленной, так и вещественной арифметики;

3) базовыми методами для устранения статистической избыточности в трансформированных изображениях являются коды Хаффмана и арифметическое кодирование, для которых свойственно:

– в случае обработки насыщенных изображений в режиме контролируемого качества происходит генерирование коротких нулевых цепочек. Отсюда длины таких цепочек будут иметь равномерное распределение и, как следствие, незначительное количество статистической избыточности. Это приводит к резкому снижению коэффициента компрессии видеоданных;

- существует необходимость использовать маркерные кодовые последовательности, разделяющие неравномерные кодовые комбинации статистических кодов. Это приводит к увеличению объема сжатого представления.

Поэтому дополнительное уменьшение информационной интенсивности битового потока обеспечит дополнительные возможности относительно повышения разрешающей способности [3]. При этом требуется учитывать ограничения на временной ресурс по обработке данных.

Проведенные исследования существующих технологий компрессии видеокадров показали, что их совершенствование требуется проводить в направлении:

– с одной стороны, снижения сложности процессов синтаксического представления;

– с другой - обеспечения заданного уровня достоверности видеoinформации, т.е. осуществления режима обработки с контролируемой потерей качества реконструируемых видеокадров.

В качестве ядра концепции дифференцированной обработки видеокадров *предлагается* использовать трансформацию изображений на основе ортогональных преобразований и перевод обрабатываемых данных в дифференциальное пространство. Это позволит использовать механизмы для управления коэффици-

ентом снижения уровня информационной интенсивности и уровнем достоверности информации, а также за счет дифференцирования создать механизм для сокращения вычислительных затрат.

Для реализации выдвинутых направлений возможны следующие варианты дополнительного снижения уровня информационной интенсивности битового потока:

– совершенствование адаптивных вероятностно-статистических моделей трансформант. Но это, с одной стороны, приводит к увеличению времени на обработку и к росту количества разрядов на представление служебной информации. С другой - эффективность статистических методов по степени компрессии практически достигла своего максимального уровня. Избыточность статистических кодов относительно энтропии Марковского источника не превышает в среднем 10 – 20 %;

– дальнейшее развитие путей для дополнительного устранения психовизуальной избыточности при обработке трансформант преобразований. Например, увеличить количество отбрасываемых из дальнейшей обработки компонент трансформант. В то же время в условиях обработки цифровых изображений с высоким уровнем информативности и семантической сложности это влечет за собой разрушение семантической структуры либо отдельных фрагментов, либо всего видеокadra.

Таким образом, для методов, использующих ортогональные преобразования и дифференцирование, нужно применять технологии сокращения избыточности путем выявления новых закономерностей и устранения избыточности не только статистической, но и психовизуальной природы. Для этого *предлагается* дополнительно выявлять в трансформированных изображениях структурные закономерности. Данный подход базируется на учете структурных преимуществ трансформированных сегментов видеокadra.

### **3. Разработка метода синтаксического представления базовых сегментов**

Рассмотрим разработку подхода относительно обработки базовых сегментов видеокадров. В данном случае в процессе формирования синтаксического представления необходимо учитывать возможность коррекции трансформант таких сегментов под особенности зрительного их восприятия. Другими словами, допускается использование коррекции частотных компонент трансформированных базовых сегментов под модель психовизуального восприятия [4]. Предложенная схема создает условия для повышения ко-

личества устраняемой психовизуальной избыточности.

Основными технологическими этапами такого подхода являются следующие.

Выполняется двумерное ДКП-преобразование, которое реализуется на основе выражения:

$$C(k, \ell) = F(k) X(i, j)_{k, \ell} F^T(\ell); \quad (1)$$

где  $X(i, j)_{k, \ell}$  – массив базового сегмента, образованный на основе цветоразностной плоскости изображения;

$k, \ell$  – соответственно индекс строки и столбца элемента массива  $X(i, j)_{k, \ell}$ ,  $k = \overline{1, q_1}$ ;  $\ell = \overline{1, q_2}$ ;

$C(k, \ell)$  – матрица компонент трансформанты ДКП-преобразования сегмента видеокadra;

$F(k), F^T(\ell)$  – соответственно вектор дискретных значений базисных функций ДКП и его транспонированный вид

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

Для трансформант сегментов, содержащих фоновые области видеокadra, характерны следующие особенности [5]:

- 1) значение компоненты в верхнем левом углу трансформанты с координатами (1; 1) пропорционально средней яркости фонового сегмента, т.е. несет информацию о среднем фоне;
- 2) наибольшая энергетическая составляющая фонового сегмента сосредоточена в ограниченном количестве компонент трансформанты;
- 3) преобладающей областью трансформанты является область низких энергетических характеристик той или иной её особенности.

В зависимости от учета в процессе обработки трансформант ее особенности формируется конкретная технология кодирования для исключения избыточности [6]. Процесс реализуется на очередном технологическом этапе обработки трансформированных сегментов видеокadra.

С учетом того, что кодирование проводится для фоновых областей видеокadra, предлагается, с одной стороны, исключить этап дифференциального представления компонент смежных столбцов трансформанты, а с другой - использовать более высокий уровень коррекции частотных компонент под модель психовизуального восприятия изображений для повышения количества сокращения психовизуальной

избыточности. Для формирования кодового значения  $E''(R)$  предлагается учитывать следующую интерпретацию двумерной трансформанты ДКП [7].

Трансформанта  $Y''$  представляет собой перестановку с повторениями, на компоненты  $y''_{k, \ell}$  которой наложены ограничения на мощность алфавита  $d''_{k, \ell}$ :

$$d''_{k, \ell} = \min(d''_k; d''_\ell). \quad (3)$$

Здесь  $d''_k, d''_\ell$  – мощность алфавита соответственно для компонент  $k$ -й строки и  $\ell$ -го столбца трансформанты.

В этом случае количество  $V_{q_1 \times q_2}^{(2)}$  различных двумерных трансформант, составленных из  $q_1 \times q_2$  количества элементов, удовлетворяющих соотношению (6), равно:

$$V_{q_1 \times q_2}^{(2)} = \prod_{k=1}^{q_1} \prod_{\ell=1}^{q_2} d''_{k, \ell}. \quad (4)$$

Согласно комбинаторной интерпретации двумерной трансформанты и соотношению (4) количество  $\bar{N}_2$  информации, в среднем содержащееся в одной компоненте  $y''_{k, \ell}$ , равно:

$$\bar{N}_2 = ([\sum_{k=1}^{q_1} \sum_{\ell=1}^{q_2} \log_2 d''_{k, \ell}] + 1) / q_1 \times q_2, \quad (5)$$

где  $\bar{N}_2$  – количество информации, приходящееся в среднем на одну компоненту двумерной трансформанты в случае его комбинаторной трактовки как цельного объекта для ограничений (2). Для устранения количества комбинаторной избыточности в двумерной трансформанте, определяемой как разница между количеством информации, приходящимся в среднем на один элемент, до и после учета ограничений на мощность алфавита ее компонент, предлагается использовать двумерные позиционные кодовые конструкции. Это объясняется тем, что количество комбинаторной избыточности зависит от количества перестановок с повторениями, которое можно составить для различных ограничений на компоненты двумерной трансформанты с учетом текущей мощности их алфавита. Тогда двумерная трансформанта  $Y''$  представляется как двумерное неравновесное позиционное число, на элементы которого наложены ограничения, определяемые их мощностью алфавита [8].

Такая система будет зависеть от устраняемой психовизуальной избыточности, количество которой

зависит от параметра  $R$ , определяющего уровень коррекций частотных компонент [9]. При этом согласно комбинаторной трактовке двумерной трансформанты  $Y''$  в качестве оснований выбираются значения мощности алфавита компонент, т.е.:

$$\psi''(R)_{k,\ell} = d''_{k,\ell}, \quad q_1 = q_2 = \overline{1}, (Q_1 Q_2 / q_1 q_2). \quad (6)$$

В этом случае кодовое значение  $E''(R)$  двумерного позиционного числа  $Y''$  определяется как:

$$E''(R) = \sum_{k=1}^{q_1} \sum_{\ell=1}^{q_2} y''_{k,\ell} v(R)_{k,\ell}^{(2)}, \quad (7)$$

где  $v(R)_{k,\ell}^{(2)}$  - весовой коэффициент компоненты  $y''_{k,\ell}$ .

В случае обхода элементов в направлении столбцов сверху вниз, а затем по строкам слева – направо значение весового коэффициента  $v(R)_{k,\ell}^{(2)}$  будет равно:

$$v(R)_{k,\ell}^{(2)} = \prod_{\xi=\ell+1}^{q_2} \psi''(R)_{k,\xi} \prod_{\eta=k+1}^{q_1} \prod_{\xi=1}^{q_2} \psi''(R)_{\eta,\xi}. \quad (8)$$

Технологию выбора количества двоичных разрядов на представления кодового значения  $E''(R)$  предлагается строить с учетом системы оснований. Тогда количество разрядов на представление кодовой двумерной трансформанты будет определяться на основе следующего выражения:

$$h''(R)^{(2)} = ([\sum_{k=1}^{q_1} \sum_{\ell=1}^{q_2} \ell \log_2 \psi''(R)_{k,\ell}] + 1). \quad (9)$$

В этом случае будет выполняться неравенство

$$([\log_2 E''(R)] + 1) / q_1 \leq \overline{h''(R)}^{(2)} \leq \overline{H}_1,$$

где  $\overline{h''(R)}^{(2)}$  - количество двоичных разрядов, приходящееся в среднем на одну компоненту двумерной трансформанты в случае его синтаксического представления как двумерного позиционного числа в условиях использования системы  $\Psi''(R)$  оснований для позиционирования их кодовых значений. Здесь количество двоичных разрядов  $\overline{h''(R)}^{(2)}$ , которое требуется выделить в среднем под одну компоненту синтаксического представления кодового значения  $E''(R)$ , не будет превышать величины  $\overline{H}_1$  [10]. Это создает условия для сокращения количества комбинаторной избыточ-

ности, которое содержится в двумерной трансформанте дискретного косинусного преобразования.

Отсюда согласно принятым условиям построение эффективного синтаксического представления для базовых сегментов определяется такими выражениями:

$$E''(R) = \sum_{k=1}^{q_1} \sum_{\ell=1}^{q_2} y''_{k,\ell} v(R)_{k,\ell}^{(2)}; \quad (10)$$

$$h''(R)^{(2)} = ([\sum_{k=1}^{q_1} \sum_{\ell=1}^{q_2} \ell \log_2 \psi''(R)_{k,\ell}] + 1) \leq q_1 \cdot q_2 \cdot \overline{H}. \quad (11)$$

Эффективное синтаксическое представление всей трансформанты базового сегмента видеокадра в виде кодограмм кодовых значений  $E''(R)$  двумерных позиционных взвешенных чисел с учетом коррекции частотных составляющих определяется величиной  $h''(R)^{(2)}$ .

В этом случае переполнение технологической длины  $h_{it}$  кодовых слов исключается путём выбора более высоких значений параметра  $R$ .

Созданный подход относительно синтаксического представления двумерной трансформанты обеспечивает снижение интенсивности битового потока с учетом особенностей базовых сегментов относительно возможности дополнительного повышения количества устраняемой психовизуальной избыточности. Значение  $E''(R)$  будет зависеть от детальности базового сегмента. Значение кода будет тем меньше, чем больше отношение площади, имеющей слабенеющуюся яркость, к площади изображения передаваемого объекта. Следовательно, соотношения (10) – (11) задают метод формирования эффективного синтаксического кодирования базовых сегментов, содержащих фоновую информацию видеокадра на основе формирования кодового значения двумерной трансформанты, рассматриваемой как взвешенное двумерное позиционное число с учетом коррекции частотных составляющих и последующего кодообразования с использованием системы оснований.

**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. Barannik. V.V, Ryabukha Yu.N., Podlesnyi S.A. “Structural slotting with uniform redistribution for enhancing trustworthiness of information streams”.

Telecommunications and Radio Engineering (English translation of *Elektrosvyaz and Radiotekhnika*), 2017. №76 (7), pp.607. doi: 10.1615 / TelecomRadEng.v76.i7.40. 4. *Barannik V.V., Ryabukha Yu.N., Tverdokhle V.V., Barannik D.V.*, “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. P. 188. 5. *Barannik V., Podlesny S., Tarasenko D., Barannik D., Kulitsa O.* “The video stream encoding method in infocommunication systems”. *Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, 2018 14th International, Proceedings of the 14th International Conference on TCSET 2018 Lviv, 2018, pp. 538-541, doi: 10.1109/TCSET.2018.8336259. 6. *Barannik V.V., Shulgin S.S.* “The method of increasing accessibility of the dynamic video information resource”. *Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET)*, Proceedings of the 13th International Conference on TCSET 2016 Lviv, 2016. P.621. 7. *Zhang Y., Negahdaripour S. and Li Q.*, "Error-resilient coding for underwater video transmission," *OCEANS 2016 MTS/IEEE Monterey*, Monterey, CA, 2016. P. 1-7. 8. *Wang S., Zhang X., Liu X., Zhang J., Ma S. and Gao W.*, Utility-Driven Adaptive Preprocessing for Screen Content Video Compression, in *IEEE Transactions on Multimedia*, vol. 19, no. 3. P. 660-667, March 2017. 9. *Stankiewicz O., Wegner K., Karwowski D., Stankowski J., Klimaszewski K. and Grajek T.*, Encoding mode selection in HEVC with the use of noise reduction, 2017 International Conference on Systems, Signals and Image Processing (IWSSIP), Poznan, 2017. P. 1-6. 10. *Barannik V.V.*. Creation of the rule of formation of positional structural and weight numbers in the conditions of codes formation with given length / *Barannik V., Krasnorutskij A.* // *Information security.* – 2016. – T 20. – 6 p.

#### Transliterated bibliography:

1. *Barannik V.V., Tarasenko D.A., Barannik D.V., Medvedev D.O.* Tekhnolohiya balansyrovanoi obrabotky dynamycheskoho vydeoresursa dlia snyzheniya ynformatsyonnoi yntensyvnyosti v ynfokommunikatsionnykh systemakh. *Bezpeka informatsii*. 2017. №3. S. 163–170.  
 2. *Barannik V.V., Tarasenko D.A., Medvedev D.O., Khimenko V.V.* Tekhnolohiia obrobky peredbachenykh kadriv videopotoku dlia bortovykh informatsiinykh tekhnolohii. *Naukoiemni tekhnolohii*. 2017. №4(36). S. 276–282.  
 3. *Barannik V.V., Ryabukha Yu.N., Podlesnyi S.A.* “Structural slotting with uniform redistribution for enhancing trustworthiness of information streams”. *Telecommunications and Radio Engineering (English translation of Elektrosvyaz and Radiotekhnika)*, 2017. №76 (7), pp.607. doi: 10.1615 / TelecomRadEng.v76.i7.40

4. *Barannik V.V., Ryabukha Yu.N., Tverdokhle V.V., Barannik D.V.* 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.  
 5. *Barannik V., Podlesny S., Tarasenko D., Barannik D., Kulitsa O.* The video stream encoding method in infocommunication systems // *Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, 2018 14th International, Proceedings of the 14th International Conference on TCSET 2018 Lviv, 2018, pp. 538-541, doi: 10.1109/TCSET.2018.8336259  
 6. *Barannik V.V., Shulgin S.S.* The method of increasing accessibility of the dynamic video information resource // *Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET)*, Proceedings of the 13th International Conference on TCSET 2016 Lviv, 2016, pp.621.  
 7. *Zhang Y., Negahdaripour S. and Li Q.* Error-resilient coding for underwater video transmission // *OCEANS 2016 MTS/IEEE Monterey*, Monterey, CA, 2016, pp. 1-7.  
 8. *Wang S., Zhang X., Liu X., Zhang J., Ma S. and Gao W.* "Utility-Driven Adaptive Preprocessing for Screen Content Video Compression // *IEEE Transactions on Multimedia*, vol. 19, no. 3, pp. 660-667, March 2017.  
 9. *Stankiewicz O., Wegner K., Karwowski D., Stankowski J., Klimaszewski K. and Grajek T.* Encoding mode selection in HEVC with the use of noise reduction // 2017 International Conference on Systems, Signals and Image Processing (IWSSIP), Poznan, 2017, pp. 1-6.  
 10. *Barannik V.V.*. Creation of the rule of formation of positional structural and weight numbers in the conditions of codes formation with given length / *Barannik V., Krasnorutskij A.* // *Information security.* – 2016. – T 20. – 6 p.

Поступила в редколлегию 10.06.2018

**Рецензент:** д-р техн. наук, проф. Баранник В.В.  
**Медведев Денис Олегович**, аспирант кафедры ИСИ ХНУРЭ. Научные интересы: обработка информации. Адрес: Украина, 61166, Харьков, просп. Науки, 14, e-mail: domedvedo@gmail.com  
**Medvedev Denis Olegovich**, postgraduate student of the Department of ISI of the Kharkov National University of Radio Electronics. Scientific interests: information processing. Address: 61166, Kharkiv, Nauka ave.14, e-mail: domedvedo@gmail.com