

Предложения по адаптивному ортогональному преобразованию видеоданных

Proposals for the adaptive orthogonal transformation of video data

Устинов / Ustinov A.

Андрей Александрович

(ust_m_a@mail.ru)

доктор технических наук, профессор.
ФГКВООУ ВО «Военная академия связи имени
Маршала Советского Союза С. М. Буденного»
(ВАС им. С. М. Буденного) МО РФ,
старший научный сотрудник.
г. Санкт-Петербург

Дворников / Dvornikov S.

Сергей Викторович

(practicdsv@yandex.ru)

доктор технических наук, профессор.
ВАС им. С. М. Буденного,
профессор кафедры радиосвязи
г. Санкт-Петербург

Агеева / Ageeva N.

Нина Сергеевна

(ust_m_a@mail.ru)

ВАС им. С. М. Буденного,
младший научный сотрудник.
г. Санкт-Петербург

Цветков / Tsvetkov V.

Василий Валерьевич

(3603326@mail.ru)

ЗАО «НПФ «ТИРС»,
начальник отдела.
г. Санкт-Петербург

Ключевые слова: адаптивное ортогональное преобразование видеоданных – adaptive orthogonal transformation of video data; сжатие изображений – image compression; статистическое кодирование без потерь – statistical coding without loss.

Предлагается метод адаптивного ортогонального преобразования видеоданных. Показано его место в общей структуре классификации методов сжатия изображений. Разработан научно-методический аппарат реализации метода. Представлены результаты оценки его эффективности в сравнении с существующими подходами.

A method of adaptive orthogonal transformation of video data is proposed. Its place in the generic structure of image compression methods classification is shown. Research and methodology tools for the method implementation is developed. The article represents results of evaluation of its efficiency compared to the existing approaches.

Введение

В настоящее время разработано значительное количество различных методов сжатия видеоданных (ВД)

[1–8]. Проведенный анализ показал, что практически все они предусматривают выполнение типовых этапов (технологических операций) преобразования сообщений источников ВД, представленных на рис. 1.

К таковым операциям, согласно рис. 1, следует относить:

- преобразование (как правило, ортогональное) дискретных отсчётов исходных ВД;
- кодирования дискретных отсчётов ВД или коэффициентов их преобразования;
- статистическое кодирование результирующих двоичных последовательностей.

Рассмотрим особенности каждой из указанных операций.

Так, преобразование дискретных отсчётов ВД позволяет минимизировать число битов, используемых при кодировании результатов преобразования.

Операции кодирования преобразованных отсчётов ВД и статистического кодирования результирующих двоичных последовательностей обеспечивают приведение коэффициентов преобразования к двоичному виду для последующей их передачи по цифровым каналам связи.

Поскольку вся совокупность рассмотренных технологических операций составляет сущность каждого из



Рис. 1. Типовые этапы преобразования сообщений источников видеоданных

существующих методов сжатия ВД, то уникальность и возможности последних определяются не указанными операциями, а соответствием проводимых преобразований статистическим свойствам кодируемых изображений. В связи с этим предлагается новый метод сжатия ВД на основе их адаптивного ортогонального преобразования (АОП), обеспечивающий наилучшее согласование со статистическими свойствами ВД по отношению к получившим широкое применение на практике методов дискретно-косинусного преобразования (ДКП) [9] и вейвлет-преобразования (ВП) [6].

Классификация методов сжатия

Алгоритмы, основанные на непосредственном преобразовании дискретных отсчетов ВД в двоичные последовательности, образуют группу методов пространственного кодирования [10]. К данной группе, в частности, относятся методы импульсно-кодовой модуляции и их разновидности (ИКМ, ДИКМ, АДИКМ и др. [1]). Практическое применение указанных методов ограничено вследствие низких коэффициентов сжатия, которые они обеспечивают.

Между тем, в зависимости от вида преобразования дискретных отсчетов ВД, группу методов кодирования с преобразованием можно разделить на методы с ортогональным и неортогональным преобразованием.

В основе всех известных методов ортогонального преобразования лежит преобразование коррелированных дискретных отсчетов ВД в их некоррелированные параметры – коэффициенты преобразования. При этом энергия коэффициентов преобразования перераспределяется так, что при кодировании существенной их части используют незначительное число битов либо вообще отказываются от их кодирования, приравнивая эти значения к нулю. На этом явлении основываются практически все методы сжатия с преобразованием.

В соответствии с проведенным анализом известных методов с ортогональным преобразованием предлагается следующая их классификация, представленная на рис. 2.

Общая характеристика методов сжатия

Проанализируем различные типы ортогонального преобразования. Математическое описание ортогонального преобразования дискретных отсчетов ВД в виде матрицы $\mathbf{A}$ имеет вид

$$\mathbf{D} = \mathbf{U} \mathbf{A} \mathbf{U}^T, \quad (1)$$

где $\mathbf{D}$ – матрица коэффициентов ортогонального преобразования; $\mathbf{U}$ и $\mathbf{U}^T$ – матрицы прямого и обратного ортогонального преобразования соответственно.

Как правило, структура матрицы коэффициентов преобразования $\mathbf{D}$ такова, что позволяет обнулить часть её элементов, т. е. заменить матрицей $\hat{\mathbf{D}}$, имеющей вид $\hat{\mathbf{D}} = \begin{bmatrix} q & 0 \\ 0 & 0 \end{bmatrix}$, а восстановленные ВД получают в виде

$$\hat{\mathbf{A}} = \mathbf{U}^T \hat{\mathbf{D}} \mathbf{U} \quad (2)$$

Матрицу преобразования $\mathbf{U}$ часто называют системой базисных векторов преобразования или базисом [11]. Тип процедуры преобразования определяется видом системы базисных векторов. На рис. 3 приведены примеры двумерных матриц коэффициентов для различных типов ортогонального преобразования. (Здесь и далее все матрицы преобразования представлены с учетом цветовой инверсии их коэффициентов для удобства отображения).

Анализ абсолютных величин матриц коэффициентов преобразования показывает, что для восстановления ВД без существенных искажений возможно обнуление, в среднем, 3/4 коэффициентов от их общего количества.

Используемые в различных методах сжатия процедуры статистического кодирования увеличивают коэффициент сжатия дополнительно в 2...4 раза в зависимости от статистики кодируемых сообщений.

Таким образом, суммарный средний коэффициент сжатия типовых изображений при использовании

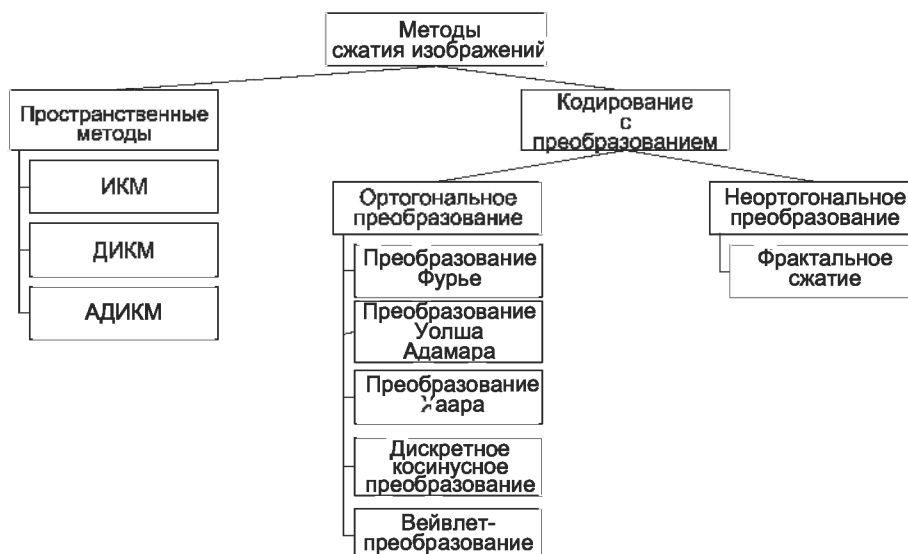


Рис. 2. Классификация методов сжатия изображений

методов кодирования с преобразованием достигает в среднем 25...50 раз.

Наряду с методами на основе ортогональных преобразований можно выделить группу алгоритмов сжатия с использованием неортогональных (аффинных) преобразований. Данные методы известны в литературе как фрактальные методы сжатия [11].

Однако в силу экспоненциальной сложности и низкой устойчивости к канальным ошибкам фрактальные методы сжатия используют обычно для компактного хранения ВД.

Сущность адаптивного ортогонального преобразования

Анализ методов статистического кодирования без потерь показывает, что реализуемые ими коэффициенты сжатия в основном достигли своего предела, определяемого избыточностью источника сообщений. В то же время, использование оптимального ортогонального преобразования для заданного изображения может привести либо к увеличению коэффициента сжатия, при сохранении заданного качества восстановленных ВД, либо к увеличению качества восстановленных ВД при сохранении коэффициента сжатия.

Из теории факторного анализа [12] известно, что какова бы ни была прямоугольная матрица $\mathbf{A}$ размером $P \times N$, всегда существует разложение

$$\mathbf{A} = \mathbf{U} \mathbf{\Lambda} \mathbf{\Lambda}^T, \quad (10)$$

где $\mathbf{U}$ и $\mathbf{\Lambda}^T$ – унитарные матрицы, $\mathbf{\Lambda}$ – прямоугольная диагональная матрица $P \times N$ с невозрастающими неотрицательными диагональными элементами.

Если задано сингулярное разложение (3) матрицы $\mathbf{A}$, то согласно [11]:

- диагональные элементы матрицы $\mathbf{A}$ являются сингулярными числами матрицы $\mathbf{A}$;
- столбцы матрицы $\mathbf{U}$ образуют ортонормированный базис из собственных векторов матрицы $\mathbf{\Lambda} \mathbf{\Lambda}^T$;
- столбцы матрицы $\mathbf{\Lambda}^T$ образуют ортонормированный базис из собственных векторов матрицы $\mathbf{A}^T \mathbf{A}$;
- столбцы матриц $\mathbf{U}$ и $\mathbf{\Lambda}^T$ образуют в совокупности сингулярные базисы матрицы $\mathbf{A}$.

Из выражения (3) следует, что

$$\mathbf{\Lambda} = \mathbf{U}^T \mathbf{A} \mathbf{V} \quad (4)$$

Матрицу $\mathbf{\Lambda}$ можно рассматривать как матрицу коэффициентов преобразования при использовании сингулярных базисов матрицы $\mathbf{A}$. Пример выполнения преобразования (4) показан на рис. 4.

Поскольку матрица $\mathbf{\Lambda}$ является диагональной, то при известных на приеме матрицах $\mathbf{U}$ и $\mathbf{V}$ для восстановления ВД необходимо располагать данными о диагональных элементах матрицы $\mathbf{\Lambda}$. Однако условие того, что матрицы $\mathbf{U}$ и $\mathbf{V}$ должны быть известны в точке восстановления ВД, не позволяет использовать преобразование (4) при практической реализации алгоритмов сжатия. Действительно, если ВД, представленные матрицей $\mathbf{A}$, имеют размер $P \times N$ пикселей, то матрицы $\mathbf{U}$ и $\mathbf{V}$ будут иметь размеры $P \times P$ и $N \times N$ соответственно. Поэтому передача матриц $\mathbf{U}$ и $\mathbf{V}$ на приемную сторону потребует расхода канального ресурса больше, чем непосредственная передача матрицы $\mathbf{A}$.

В связи с этим определим следующую матрицу коэффициентов преобразования исходных ВД

$$\tilde{\mathbf{A}} = \tilde{\mathbf{U}}^T \mathbf{A} \tilde{\mathbf{V}}, \quad (5)$$

где $\tilde{\mathbf{U}}^T$, $\tilde{\mathbf{V}}$ квадратные матрицы размером $P \times P$ и $N \times N$ соответственно, с ограничениями вида:

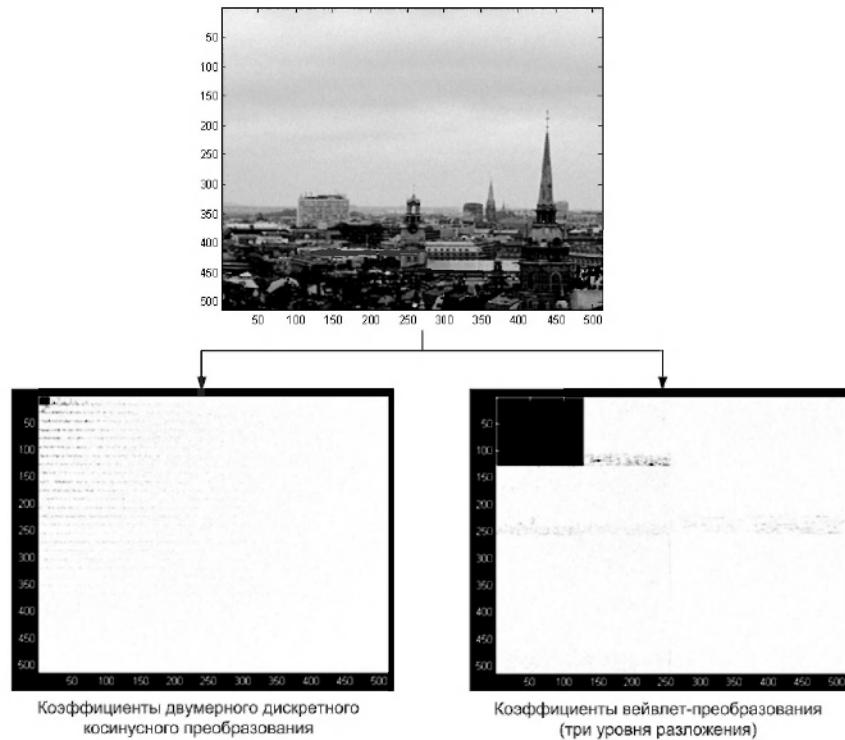


Рис. 3. Визуализация коэффициентов преобразования

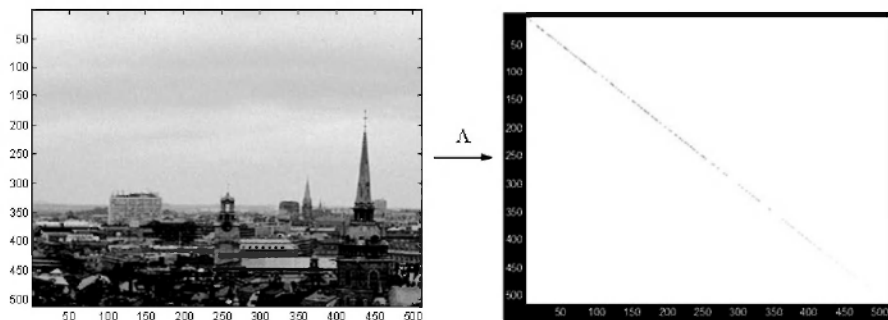


Рис. 4. Пример исходного изображения и матрицы коэффициентов сингулярного преобразования

$$\begin{aligned} \tilde{u}(i, j) = 0, \quad \forall i, j = \overline{1, P}; \quad i \neq 1 + (k-1)m_1, \\ 2 + (k-1)m_1, \dots, km_1; \quad k = 1, 2, \dots, P/m_1; \end{aligned} \quad (6)$$

$$\begin{aligned} \tilde{v}(i, j) = 0, \quad \forall i, j = \overline{1, N}; \quad j \neq 1 + (k-1)m_2, \\ 2 + (k-1)m_2, \dots, km_2; \quad k = 1, 2, \dots, N/m_2; \end{aligned} \quad (7)$$

$$\tilde{\mathbf{U}}^T \tilde{\mathbf{U}} = \mathbf{I}; \quad \tilde{\mathbf{V}}^T \tilde{\mathbf{V}} = \mathbf{I}, \quad (8)$$

Здесь $\tilde{\Lambda}$ – прямоугольная матрица размером $P \times N$ элементов с минимальным числом ненулевых элементов.

Ограничения (6) и (7) означают, что в каждой строке матрицы $\tilde{\mathbf{U}}^T$ и в каждом столбце матрицы $\tilde{\mathbf{V}}$ число ненулевых элементов составляет ровно m_1 и m_2 (величины m_1 и m_2 должны быть кратны величинам P и N , соответственно). Пример структуры матриц $\tilde{\mathbf{U}}^T, \tilde{\mathbf{V}}$ показан на рис. 5.

Если m_1 и m_2 удовлетворяют неравенствам $m_1 \ll P$ и $m_2 \ll N$, то введение ограничений (6) и (7) позволит существенно снизить требования к ресурсу цифрового канала, затрачиваемого на передачу матриц $\tilde{\mathbf{U}}^T$ и $\tilde{\mathbf{V}}$.

Ограничение (8) обеспечивает ортогональность разложения в соответствии с выражением (5).

В ходе теоретического исследования было доказано, что матрицы $\tilde{\mathbf{U}}^T$ и $\tilde{\mathbf{V}}$ полностью определяются собственными векторами корреляционных матриц

$$\left[\sum_{j=1}^N \tilde{a}_{n,j} \tilde{a}_{n,j}^T \right], \quad n=1, 2, \dots, \frac{P}{m_1}$$

и

$$\left[\sum_{j=1}^P \tilde{b}_{j,n} \tilde{b}_{j,n}^T \right], \quad n=1, 2, \dots, \frac{N}{m_2},$$

соответственно.

Здесь $\tilde{a}_{n,j}$ – n -ый фрагмент из m_1 ненулевых компонент j -го столбца матрицы $\mathbf{A}$; $\tilde{b}_{j,n}$ – n -ый фрагмент из m_2 ненулевых компонент j -ой строки матрицы $\mathbf{A}$. Поскольку матрицы найденные преобразования $\tilde{\mathbf{U}}^T, \tilde{\mathbf{V}}$ полностью определяются матрицей исходного изображения, то преобразование (5) будем называть адаптивным ортогональным преобразованием.

В интересах получения количественной оценки было проведено имитационное моделирование по преобразованию фотографии городского пейзажа представленного на рис. 3 и 4. Сжатие данных осуществлялось методом ДКП, вейвлет-преобразования и разработанным методом АОП с различными параметрами. Результаты представлены в табл. 1.

В ходе моделирования осуществлялось квантование полученных коэффициентов преобразования с различным шагом: $kv=4, kv=8, kv=16, kv=32$. Квантование матрицы коэффициентов преобразования $\tilde{\mathbf{A}}$ выполнялось в соответствии с выражением

$$\tilde{\mathbf{A}}_{kv} = \text{round} \left(\frac{\tilde{\mathbf{A}}}{kv} \right), \quad (9)$$

где $\tilde{\mathbf{A}}_{kv}$ – матрица квантованных коэффициентов преобразования; kv – шаг квантования; $\text{round}(\ast)$ – операция арифметического округления.

После вычисления квантованных коэффициентов оценивались два параметра: пиковое отношение сигнал/шум ($PSNR$) и число ненулевых коэффициентов преобразования (NZ). Величина $PSNR$ использовалась для объективной оценки качества восстановленного изображения. Параметр NZ косвенно характеризует реализуемый коэффициент сжатия изображения. Известно [2], что чем меньше число ненулевых квантованных коэффициентов преобразования, тем больше может быть обеспечен коэффициент сжатия.

Имитационное моделирование показывает преимущество АОП по сравнению с ДКП и ВП. Так, при $m_1=32$ и $m_2=32$ число ненулевых коэффициентов при АОП на 30% меньше, чем при ДКП, и более чем на 100% меньше, чем при ВП. Кроме того, восстановленное на основе АОП изображение не имеет артефактов в виде блочности, характерных для ДКП, и артефактов в виде выделения границ фоновых участков, характерных для ВП.

Заключение

Таким образом, разработанный метод АОП имеет явные преимущества по сравнению с известными неадаптивными преобразованиями, как по числу ненулевых коэффициентов преобразования при заданном шаге квантования, так и по устойчивости к канальным ошибкам. Однако более детального рассмотрения требуют вопросы разработки алгоритмов быстрого вычисления собственных векторов матриц малого размера (32×32 элемента) и передачи данных матриц по каналу связи [13], [14]. Данные вопросы являются направлением дальнейших исследований и будут рассмотрены в последующих статьях.

Литература

1. Гонсалес, Р. Цифровая обработка изображений : [пер. с англ.] / Р. Гонсалес, Р. Вудс. – М.: Техносфера, 2006. – 1072 с.

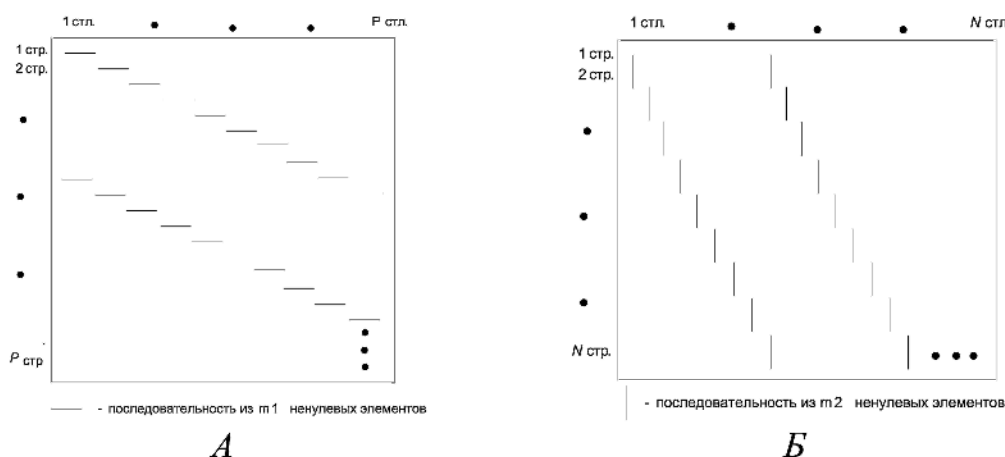


Рис. 5. Визуализация структуры матриц $\tilde{\mathbf{U}}^T = \mathbf{A}$ и $\tilde{\mathbf{V}} = \mathbf{B}$

Таблица 1

Результаты имитационного моделирования

Вид преобразования	Оцениваемый параметр							
	PSNR [дБ]				NZ			
	$kv=4$	$kv=8$	$kv=16$	$kv=32$	$kv=4$	$kv=8$	$kv=16$	$kv=32$
Адаптивное преобразование $m_1=8, m_2=8$	46,88	41,10	36,84	34,27	161688	92088	34838	15820
Адаптивное преобразование $m_1=16, m_2=16$	46,89	41,12	36,96	34,44	159176	88561	32005	12989
Адаптивное преобразование $m_1=32, m_2=32$	46,89	41,17	37,05	34,54	156814	86154	30966	12003
ДКП	46,88	41,07	36,84	34,23	162828	92770	35274	16331
Вейвлет-преобразование	46,89	40,96	36,62	33,72	171135	103176	45249	26745

2. Ричардсон, Я. Видеокодирование H.264 и MPEG-4 – стандарты нового поколения / Я. Ричардсон. – М.: Техносфера, 2005. – 368 с.

3. Зубакин, И. А. Обзор методов кодирования и декодирования изображений / И.А. Зубакин, Ш.С. Фахми // Вопросы радиоэлектроники. Сер. Техника телевидения. – 2008. – № 1. – С. 12–41.

4. Пат. 2419246 Российская Федерация, МПК H04N 7/30. Способ сжатия и восстановления неподвижных полутоновых видеоизображений / С.В. Балакирев [и др.]; патентообладатель ГОУ ВПО «Военная академия связи имени С. М. Буденного» МО РФ. – № 20101022379/09; заявл. 25.01.2010; опубл. 20.05.2011, Бюл. № 14. – 21 с.

5. Пат. 2434358 Российская Федерация, МПК H04N 7/30, G06T 5/10. Способ сжатия и восстановления подвижных цветных видеоизображений / С.В. Балакирев [и др.]; патентообладатель ГОУ ВПО «Военная академия связи имени С. М. Буденного» МО РФ. – № 2010128187/09; заявл. 07.07.2010; опубл. 20.11.2011, Бюл. № 32. – 17 с.

6. Способ сжатия графических файлов методами вейвлет-преобразований / Умбиталиев А. А. [и др.] // Вопросы радиоэлектроники. Сер. Техника телевидения. – 2015. – № 3 (21). – С. 100–106.

7. Дворников, С. В. Устранение межкадровой избыточности подвижных изображений на основе поиска оптимального базиса преобразования во временной области / С.В. Дворников, А.А. Устинов, В.В. Цветков // Вопросы радиоэлектроники. Сер. Техника телевидения. – 2012. – № 2. – С. 45–54.

8. Дворников, С. В. Компенсация движения в видеокодеках, использующих трёхмерные ортогональные преобразования, на основе оптимальных разбиений кодируемых блоков во временной области / С.В. Дворников, А.А. Устинов,

В.В. Цветков // Вопросы радиоэлектроники. Сер. Техника телевидения. – 2013. – № 2. – С. 98–111.

9. Фахми, Ш. С. Адаптивный алгоритм кодирования видеоинформации на основе трехмерного дискретного косинусного преобразования / Ш.С. Фахми, И.А. Зубакин // Изв. вузов России. Радиоэлектроника. – 2010. – Вып. 1. – С. 49–54.

10. Мамаев Н. С. Цифровое телевидение / Н.С. Мамаев, Ю.Н. Мамаев, Б.Г. Теряев. – М.: Горячая линия–Телеком, 2001. – 180 с.

11. Воеводин, В. В. Матрицы и вычисления / В.В. Воеводин, Ю.А. Кузнецов. – М.: Наука. Главная редакция физико-математической литературы, 1984. – 320 с.

12. Лоули, Д. Факторный анализ как статистический метод / Д. Лоули, А. Максвелл. – М.: Мир, 1967. – 144 с.

13. Дворников, С. В. Метод формирования признаков сигналов для систем контроля и диагностики многоканальных устройств радиосвязи / С.В. Дворников // Контроль. Диагностика. – 2009. – № 5. – С. 54–60.

14. Дворников, С. В. Алгоритм распознавания сигналов радиосвязи на основе симметрических матриц / С.В. Дворников, С.С. Дворников, М.А. Коноплев // Информационные технологии. – 2010. – № 9. – С. 75–77.