

УДК 621.397.2

СРАВНИТЕЛЬНЫЙ АНАЛИЗ АЛГОРИТМОВ СЖАТИЯ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ДИСКРЕТНОГО КОСИНУСНОГО (DCT) И ЧЕБЫШЕВСКОГО (GDCT) ПРЕОБРАЗОВАНИЙ

Радченко Ю.С., Радченко Т.А., Булыгин А.В.

Введение

В настоящее время стремительно возрастает потребность в услугах по передаче мультимедийной информации по каналам связи (видеоконференция, видеотелефония, беспроводный доступ в телекоммуникационные сети). При этом необходимо обеспечить передачу видеинформации приемлемого качества в реальном масштабе времени. Для решения этой задачи используются алгоритмы сжатия с потерями. В большинстве современных стандартов, например JPEG, MPEG1-4, H.261-264, используется дискретное косинусное преобразование (DCT) для перехода из пространства изображения в более компактное спектральное пространство. Несмотря на высокую популярность этого дискретного преобразования, активно продолжается поиск альтернативных методов сжатия.

К числу последних относится способ разложения сигналов побазису классических ортогональных полиномов Чебышева (алгоритм GDCT) [1,2]. Как показывают эксперименты, он позволяет добиться значительного сжатия, обладает рядом новых сервисных функций, допускает реализацию в виде "быстрых" алгоритмов преобразования и совместим с существующими стандартами на основе DCT. Для выявления преимуществ этого алгоритма и определения области его возможного применения возникает необходимость объективного сравнения различных дискретных преобразований. В данной работе приведены результаты исследования алгоритмов сжатия на основе чебышевского (GDCT) и дискретного косинусного (DCT) преобразований с помощью нескольких критериев качества восстановленного изображения.

Алгоритмы преобразования GDCT и DCT

Пусть в подобласти $\{x,y\} \in \Omega$ наблюдается поле $s(x,y)$, представляющее собой фрагмент изображения в блоке. Базисные функции, используемые для дискретного представления сигнала, имеют вид $\varphi_{mk}(x,y) = \varphi_m(x) \cdot \varphi_k(y)$, где $\varphi_m(x)$, $\varphi_k(y)$ – одномерные функции, основанные на ортогональных полиномах. Тогда для полезного сигнала $s(x,y)$ имеет место пара биортогональных преобразований

$$\begin{aligned} s(x,y) &= \sum_{m,k} C_{mk} p_m(x/a_x) p_k(y/a_y), \\ C_{mk} &= \alpha \iint_{\Omega^*} s(a_x z_1, a_y z_2) \rho(z_1) p_m(z_1) p_k(z_2) dz_1 dz_2 = \\ &= \alpha \int \rho(z_1) p_m(z_1) dz_1 \int s(a_x z_1, a_y z_2) \rho(z_2) p_k(z_2) dz_2. \end{aligned} \quad (1)$$

Проводится исследование алгоритма сжатия изображения с помощью полиномов Чебышева (алгоритм GDCT). Получены количественные характеристики потерь качества восстановленного изображения для различных коэффициентов сжатия. Выполнено сравнение потерь по нескольким критериям в двух вариантах алгоритма GDCT и алгоритма DCT.

В (1) $\alpha = 1/(d_m d_k)$, d_m – норма ортогонального с весом $\rho(z)$ полинома $p_m(z)$, a_x, a_y – характерные размеры подобласти Ω , $z_1 = x/a_x$, $z_2 = y/a_y$. Для последовательного вычисления интегралов в (1) целесообразно использовать квадратурную формулу гауссовского типа

$$\int s(z) \rho(z) dz = \sum_{i=1}^N \lambda_i s(z_i).$$

Здесь z_i – нули полинома $p_N(z)$, ортогонального с весом $\rho(z)$; λ_i – числа Кристоффеля. Узлы и веса $\{z_i\}$, $\{\lambda_i\}$ однозначно определяются видом полинома $p_N(z)$. Соотношения (1) являются общими для разложения по произвольной системе ортогональных полиномов. Для полиномов Чебышева прямое и обратное преобразования (одномерный вариант для нормированного интервала $z \in [-1,1]$) имеют вид

$$\begin{aligned} C_m &= g_m \sqrt{\frac{2}{N}} \sum_{i=0}^{N-1} s(z_i) \cos\left(\pi m \frac{i+0.5}{N}\right), \\ R_M(z) &= g_m \sqrt{\frac{2}{N}} \sum_{m=0}^M C_m \cos(m \cdot \arccos(z)), \end{aligned} \quad (2)$$

где $g_m = 1$ при $m > 0$ и $g_m = 0.5$ при $m = 0$. Согласно (2), точки отсчета $z_i = \cos(\pi(i+0.5)/N)$ изображения $s(z)$ берутся неравномерно. Синтез (восстановление) изображения $R_M(z)$ по M спектральным компонентам выполняется в произвольной точке $z \in [-1,1]$, а не в дискретном наборе точек отсчета, как в DCT. При восстановлении может использоваться сетка отсчетов $z_n = 2n/(L-1) - 1 + \delta$, где δ – субпиксельный сдвиг, $n=0, \dots, L-1$. Если $L \neq N$, то восстановленное изображение подвергается геометрическому масштабированию, как в сторону уменьшения, так и в сторону увеличения размера.

При двумерном GDCT преобразовании в пределах блока из $N \times N$ точек берутся $N \times N$ отсчетов изображения по закону

$$\begin{aligned} x_i &= \text{ROUND}(0.5 \cdot (N-1) \cdot (1 + \cos(\pi(i+0.5)/N))), \\ y_j &= \text{ROUND}(0.5 \cdot (N-1) \cdot (1 + \cos(\pi(j+0.5)/N))), \\ i, j &= 1..N. \end{aligned} \quad (3)$$

Отобранные отсчеты изображения образуют матрицу $S = \|s_{ij}\| = \|s(x_i, y_j)\|$. Эта матрица преобразуется в матрицу спектральных коэффициентов C размером $M \times M$ с помощью прямоугольной матрицы прямого преобразования размером $M \times N$. При обратном преобразовании может использоваться прямоугольная матрица размером $L \times M$. То есть восстановленный блок $R = \|R_{n,r}\|$ имеет размеры $L \times L$. Прямое и обратное преобразование Чебышева (GDCT) и DCT в матричном виде определяются операциями

$$C = \Phi S \Phi^T, \quad R = \Psi^T C \Psi. \quad (4)$$

Матрицы прямого Φ и обратного Ψ преобразований GDCT имеют вид

$$\Phi = \|\varphi_m(i)\|_{NM} = \sqrt{\frac{2}{N}} \begin{bmatrix} \sqrt{0.5} \\ \cos(\pi m \frac{i+0.5}{N}) \end{bmatrix}, \quad m=0..M-1, i=0..N-1, \quad (5)$$

$$\Psi = \sqrt{\frac{2}{N}} \begin{bmatrix} \sqrt{0.5} \\ \cos(m \cdot \arccos(\frac{2n}{L-1} - 1 + \delta)) \end{bmatrix}, \quad m=0..M-1, n=0..L-1. \quad (6)$$

Как видно из (5) и (6), матрицы Φ и Ψ преобразования GDCT, в общем случае, прямоугольные. Матрица прямого преобразования DCT совпадает с (5), где $M=N$. Обратное преобразование осуществляется с помощью матрицы $\Psi = \Phi^T$. Заметим, что в DCT матрицы преобразований квадратные размером ($N \times N = M \times M$).

В отличие от DCT, где отсчеты изображения для преобразования в спектральные компоненты берутся равномерно, алгоритм GDCT предполагает отбор пикселей по закону (3). В общем случае, из-за дробного вида нулей полиномов Чебышева Z_i , $i=1..N$, элементы матрицы S – сэмплы не совпадают со значениями изображений в исходных пикселях. Для уточнения расчетов спектральных коэффициентов предложено [2,3] производить интерполяцию изображения $s(x,y)$ в точках расположения нулей полиномов Чебышева по отсчетам, находящимся вблизи этих точек. Для этого используется четырехточечная интерполяционная формула Бесселя:

$$\begin{aligned} s(x,y) = & 0.25(s_{00} + s_{01} + s_{10} + s_{11}) + \\ & 0.5(u-0.5)(s_{10} - s_{00} + s_{11} - s_{01}) + 0.5(v-0.5)(s_{01} - s_{00} + s_{11} - s_{10}) + \\ & +(u-0.5)(v-0.5)(s_{11} - s_{10} - s_{01} + s_{00}), \end{aligned} \quad (7)$$

где $u=x-x_0$, $v=y-y_0$, $s_{k,l}=s(x_k, y_l)$ ($k, l=0, 1$) – значения в ближайших четырех пикселях, окружающих точку с координатами (x, y) .

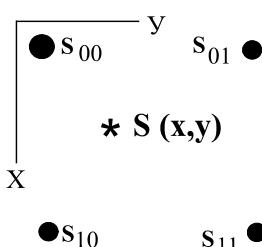


Рис. 1

Вычисление сэмплов с помощью (7) производится в прецизионном режиме работы GDCT, который обоз-

начен в дальнейшем как GDCTi. На рис. 1 представлена конфигурация пикселей и сэмпла используемая в формуле (7). Здесь сэмпл обозначен звездочкой. Применение (7) целесообразно в случае, когда сэмпл находится вблизи центра межпиксельной области.

Общая схема сжатия цветных цифровых изображений на основе преобразования GDCT содержит следующие этапы:

1. Цветовая перекодировка из {RGB} в {YUV}.
2. Субдискретизация матриц Y, U, V.
3. Сжатие первой ступени путем организации сэмплов согласно формуле Гаусса-Чебышева.
4. GDCT- преобразование.
5. Сжатие второй ступени с матрицами квантования

$$Q_{m,k}^Y(q), \quad Q_{m,k}^{U,V}(q),$$

где q -коэффициент, регулирующий сжатие.

6. Сжатие третьей ступени – энтропийное кодирование.

Восстановление производится в обратном порядке с возможностью формирования блоков $R(L \times L)$, $L/N \neq 1$.

Критерии качества восстановленных изображений

Для сравнения алгоритмов сжатия изображения GDCT и DCT было исследовано качество восстановленного изображения в зависимости от степени сжатия. В настоящее время существует множество критериев сравнения двух изображений, отражающих те или иные стороны восприятия изображения [4,5,6]. Из них были отобраны следующие, основанные на разнице значений пикселей: а) PSNR-(пиковое отношение сигнал/шум), б) комплексный критерий MSSIM (mean structural similarity index – усредненный показатель структурного подобия) [5]. PSNR определяется соотношением вида

$$PSNR = 10 \lg \left(\frac{255^2 \cdot v \cdot Nl^2}{\sum_{\eta=1}^v \sum_{n,r=0}^{Nl-1} (s_{nr}^{(\eta)} - R_{n,r}^{(\eta)})^2} \right). \quad (8)$$

В (8) $S_\eta = \|s_{nr}^{(\eta)}\|$, $R_\eta = \|R_{nr}^{(\eta)}\|$ блоки с номером η в исходном и восстановленном изображениях, v -число блоков изображения.

Критерий MSSIM, характеризующий близость изображений S и R по яркости, контрасту и структуре, имеет следующий вид:

$$MSSIM = \frac{1}{v} \cdot \sum_{\eta=1}^v SSIM(S_\eta, R_\eta). \quad (9)$$

Индекс SSIM: structural similarity index – индекс структурного подобия. Индекс $SSIM(S_\eta, R_\eta)$ в (9) выражается следующим образом:

$$SSIM(S_\eta, R_\eta) = \frac{(2 \cdot \mu_S \cdot \mu_R + C_1) \cdot (2 \cdot \sigma_{SR} + C_2)}{(\mu_S^2 + \mu_R^2 + C_1) \cdot (\sigma_S^2 + \sigma_R^2 + C_2)}. \quad (10)$$

Здесь μ_S , μ_R , σ_S^2 , σ_R^2 – выборочные средние и дисперсии в блоках изображений S_η и R_η соответственно, σ_{SR} – корреляционный момент между областями изображений S_η и R_η . C_1 , C_2 – малые константы,



а)



б)

Рис. 2

предотвращающие некорректное поведение критерия при обнулении моментов [5]. Критерий (9) принимает значения от -1 до 1. Причем значение 1 получается только в том случае, если сравнивается одно и тоже изображение.

Исследование влияния степени сжатия на качество изображений

Для выбора параметров алгоритма GDCT и сравнения его с алгоритмом сжатия на основе DCT, были построены два кодека изображения, реализующих названные алгоритмы. Степень сжатия первой ступени алгоритма GDCT определяется отношением $(N_1/N)^2$. Исследовались форматы $N_1/N=12/8$ и $N_1/N=8/6$ в алгоритме GDCT. Спектральные преобразования выполнялись при значениях $M=N$. Размеры исходных блоков $N_1 \times N_1$ в алгоритмах DCT и GDCT при сравнении результатов были одинаковы, то есть 8×8 и 12×12 .

Элементы матриц квантования для спектральных коэффициентов компонент Y,U,V формировались соответственно по законам

$$Q_{m,k}^Y = 1 + (m+k)q, \quad Q_{m,k}^{U,V} = 1 + (m \cdot k)q. \quad (11)$$

Здесь q – регулируемый параметр второй ступени сжатия. Такой выбор матриц квантования обеспечивает более сильное сжатие цветовых компонент. В

качестве тестовых были выбраны изображения: «Lena» и «пейзаж», представленные на рис. 2а и 2б, а также другие такого же класса.

В таблице 1 приведены значения энтропии в «битах на пиксель» $H_{DCT}(q)$ и $H_{GDCT}(q)$ изображения «Lena» для двух вариантов блочного разбиения: 8×8 и 12×12 . Методика определения энтропии приведена в [6]. Как следует из таблицы 1, при одинаковых значениях параметра сжатия q алгоритм GDCT обеспечивает меньшее значение энтропии «бит на пиксель», чем DCT за счет прореживания на первой ступени сжатия.

На рис. 3 приведен относительный проигрыш в энтропии «бит на пиксель» алгоритма DCT по сравнению с алгоритмом GDCT в зависимости от параметра сжатия q , рассчитанный по формуле:

$$\delta H(q) = \frac{H_{DCT}(q) - H_{GDCT}(q)}{H_{GDCT}(q)} \cdot 100\% \quad (12)$$

Как видно из таблицы 1 и рис. 3, с увеличением размеров блоков прореживание на первой ступени сжатия, осуществляемое в GDCT, сильнее сказывается на величине $\delta H(q)$.

Поскольку параметр q не определяет полностью сжатие изображения в алгоритме GDCT, имеет смысл сравнивать характеристики качества восстановления обоих алгоритмов при одинаковых значениях H «бит на пиксель».

Таблица 1

	q	0	10	20	30	40	50	60	70	80	90	100
8×8	H_{DCT}	8.6	1.29	0.977	0.847	0.772	0.723	0.69	0.663	0.644	0.627	0.615
	H_{GDCT}	5.21	1.07	0.826	0.725	0.667	0.631	0.604	0.584	0.569	0.557	0.547
12×12	H_{DCT}	8.45	0.868	0.631	0.53	0.471	0.431	0.404	0.383	0.366	0.353	0.342
	H_{GDCT}	4.12	0.662	0.491	0.417	0.376	0.349	0.33	0.315	0.303	0.295	0.287

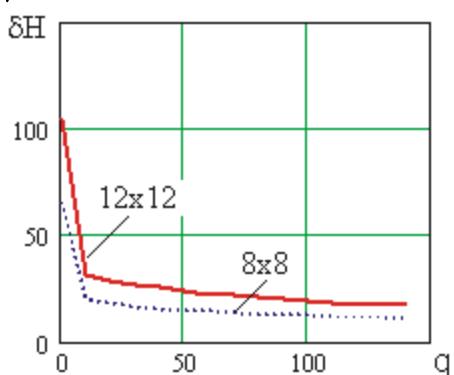


Рис. 3

На рис.4 и рис.5 представлены зависимости характеристик алгоритмов DCT и GDCT от энтропии «бит на пиксель» при разбиении на блоки размером 8×8 . Характеристики MSSIM и PSNR определялись для яркостной компоненты Y. Вычисление сэмплов в алгоритме GDCT производилось согласно (3) без использования межпиксельной интерполяции. Как видно из данных рисунков, алгоритм GDCT имеет некоторое преимущество по сравнению с DCT при большом коэффициенте сжатия, то есть при малых значениях энтропии «бит на пиксель». При малых коэффициентах сжатия изображения характеристики DCT несколько лучше, чем у GDCT. Точка пересечения характеристик зависит от размеров блока, типа изображения.

Качество восстановленного изображения можно улучшить, применяя интерполяцию (7). Результаты количественного сравнения качества восстановлено-

го изображения при разбиении на блоки 12×12 для алгоритмов DCT, GDCT и GDCTi представлены на рис. 6 и рис. 7.

Исследования показали, что межпиксельная интерполяция при нахождении сэмплов целесообразна при небольших коэффициентах сжатия и блоках, содержащих мелкомасштабные фрагменты изображения. Эксперименты с изображением типа «пейзаж» дали результаты, аналогичные представленным на рис. 3-7.

Заключение

На основе проведенных исследований можно сделать ряд практически важных выводов:

1. Алгоритм GDCT предпочтителен при высокой степени сжатия. В этом случае он обеспечивает по сравнению с DCT несколько большее значение критериев MSSIM и PSNR, т.е. лучшее качество восстановленного изображения.

2. Качество восстановленного изображения повышается при использовании в алгоритме GDCT межпиксельной интерполяции для вычисления матрицы сигнальных отсчетов-сэмплов. Интерполяцию следует применять при малых степенях сжатия и для уменьшения искажений на изображениях, блоки которых содержат мелкомасштабные структуры.

3. Алгоритм GDCT (GDCTi), обеспечивая несколько лучшие характеристики при малых значениях энтропии «бит на пиксель», может быть полезен для адаптивной регулировки битового потока при передаче мультимедийной информации по узкополосным и зашумленным каналам связи.

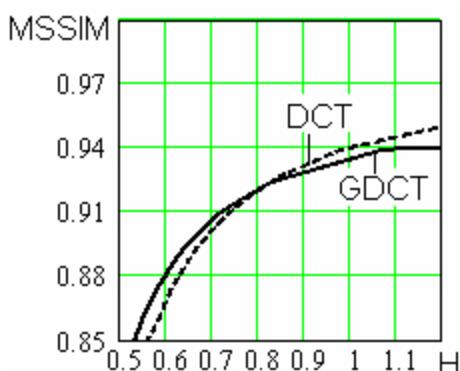


Рис.4

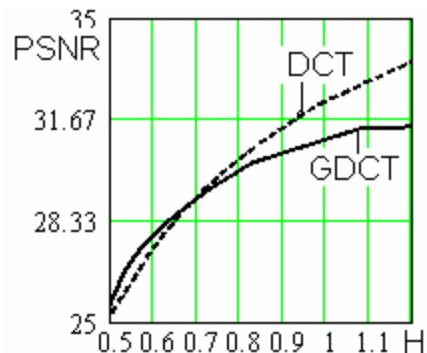


Рис.5

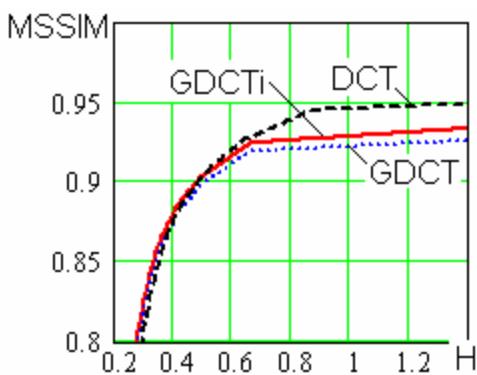


Рис.6

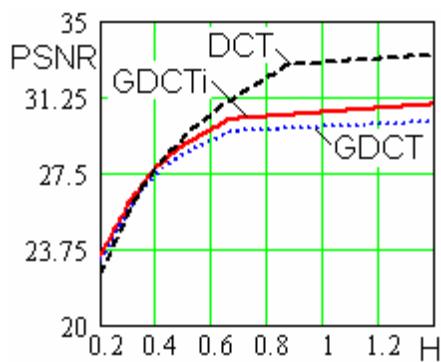


Рис.7

Литература

1. Радченко Ю.С. Алгоритм сжатия изображений на основе полиномиальных преобразований / Ю.С. Радченко // Цифровая обработка сигналов. - 2002. - № 1. - С. 2-6.
2. Радченко Ю.С. Метод сжатия и восстановления изображений на основе быстрых чебышевских преобразований / Ю.С. Радченко // Автометрия. - 2002. - № 4. - С. 32-40.
3. Yuri Radchenko. RESEARCH OF SIGNAL RECOVERY, SUPPRESSION AND PROCESSING ALGORITHMS BASED ON POLYNOMIAL TRANSFORMATIONS. The 6th World Multiconference of Systemics, Cybernetics and Informatics. July 14-18 2002, Orlando, Florida, USA. Proceedings, v. VIX, Image, Acoustic, Speech and Signal Processing III, p.262-266
4. M. Miyahara, K. Kotani, V.R. Algazi Objective Picture Quality Scale (PQS) for Image codong. IEEE Trans. On Comm. 1998, v.46, № 9.
5. Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity", IEEE Transactions on Image Processing, vol. 13, No. 4, pp.600-612, Apr. 2004.
6. Прэтт У. Цифровая обработка изображений: в 2-х т. / У. Прэтт; – перевод с англ. – М.: Мир, 1982. – Т. 1. – 204 с

НОВЫЕ КНИГИ

В конце прошлого 2006 г. в издательствах «Радио и связь» и «Горячая линия-Телеком» завершён совместный проект по выпуску редкой для нашей страны монографии по теории помехоустойчивого кодирования «Теория и алгоритмы многопорогового декодирования» отечественного автора д.т.н. В.В. Золотарёва. Научный редактор книги – член-корреспондент РАН Ю.Б. Зубарев. Положительную рецензию на эту весьма необычную для теории кодирования монографию дал академик РАН В.К. Левин.

Автор представляет по сути совершенно новую теорию, которая уже стала хорошей основой для разработки очень простых в реализации высокоеффективных и чрезвычайно быстрых алгоритмов многопорогового декодирования (МПД). Эта задача исключительно важна для спутниковых, космических и многих других очень дорогих каналов связи с большим уровнем шума. Использование помехоустойчивого кодирования для таких каналов существенно (иногда многократно) повышает их к.п.д., что по существу и определяет ту грандиозную экономическую эффективность применения кодирования и собственно разработок декодеров для подобных каналов.

Монография вышла в свет в год 40-летнего юбилея русского перевода классической книги Месси «Пороговое декодирование». Именно её идеи весьма неожиданным образом автору удалось развить до уровня, при котором сложность нового МПД алгоритма осталась по существу близкой по порядку величины к сложности прототипа, но характеристики даже при весьма высоком уровне шума оказываются практически такими же, как у переборных оптимальных алгоритмов.

МПД по числу операций промежуточных с другими конкурирующими с ними алгоритмами с близкими уровнями эффективности примерно на два порядка. Эти методы разработаны для целого ряда типичных моделей каналов и демонстрируют очень высокий уровень быстродействия как в аппаратных вариантах, так и при программной реализации. Судя по представленным в книге данным, очень высокие характеристики недвоичных алгоритмов класса МПД свидетельствуют о начале совершенно нового периода, когда на базе этих алгоритмов можно строить особо надёжные, с высочайшим уровнем достоверности, системы передачи информации и хранения сверхбольших баз данных. Простота реализации недвоичных МПД и уровень обеспечиваемой ими достоверности на несколько порядков выше того, что можно получить при использовании кодов Рида-Соломона, которые только и могут использовать сегодняшние телевизионные и медиатехнологии. Наверное, многие проблемы в этой области были бы решены быстрее и лучше, если бы применение недвоичных символьных МПД началось 20 лет назад, когда появились первые уже достаточно содержательные результаты по этим символьным декодерам. Невозможно даже перечислить все те новые результаты по простой реализации методов МПД для разных систем и каналов: каскадирование параллельное и с кодами контроля чётности, каналы с неравномерной энергетикой, коды с выделенными ветвями и, конечно, все основные результаты, относящиеся к сходимости решений МПД к решению оптимального декодера. Действительно, в мировой литературе нет данных о том, чтобы аналогичными свойствами обладали бы какие-нибудь другие методы коррекции ошибок.

Заметим, что самые сложные из лучших опубликованных на текущий момент алгоритмов имеют несколько более высокие по энергетическому выигрышу кодирования результаты, чем данные монографии по возможностям МПД. Но эта разница уже сейчас невелика и, судя по динамике развития теории и конкретных разработок по МПД алгоритмам, очень правдоподобно, что эта разница в самом недалёком времени будет минимизирована или даже сведена к нулю. Но это вопросы ближайшего будущего. В то же время, исходя из приведённого в книге описания принципов работы МПД на ПЛИС Xilinx, уже сейчас этот метод допускает реализацию сверхбыстрых МПД со скоростями декодирования, на несколько порядков превышающими производительность других методов.

В целом можно отметить высокую востребованность монографии с точки зрения как чисто теоретических, так и важных прикладных результатов в чрезвычайно конкурентной области теории и техники связи. Они могут существенно поднять уровень отечественной аппаратуры связи, если реализовать в ней хотя бы основные технические решения по созданию систем кодирования, изложенные в этой исключительно полезной книге.

Заслуженный работник
высшего образования РФ
Почетный член РНТО РЭС им. А.С. Попова,
зав. кафедрой радиоуправления и связи
д.т.н., профессор

/Кириллов С.Н./