

УДК 519.7:537.8

А.В. Статкус

Национальный технический университет "ХПИ", Харьков

КОДИРОВАНИЕ ИЗОБРАЖЕНИЙ С ПРЕОБРАЗОВАНИЕМ НА ОСНОВЕ МОМЕНТОВ РАКА

Дискретные ортогональные моменты являются мощным средством описания признаков формы изображений в приложениях распознавания образов и анализа изображений. В настоящей статье рассматривается новое множество дискретных ортогональных моментов, сформированных на основе взвешенных дискретных многочленов Рака. Совокупность этих моментов образует линейное унитарное преобразование исходного изображения, называемое в статье преобразованием Рака. В силу унитарности оно удовлетворяет формальным требованиям к преобразованиям, используемым для сжатия изображений. Обсуждаются некоторые свойства многочленов Рака и результаты экспериментальных исследований информационных возможностей преобразования Рака как средства кодирования изображения в сравнении с традиционными преобразованиями с точки зрения степени сжатия, качества восстановления, помехоустойчивости и практической реализуемости.

Ключевые слова: сжатие изображений, кодирование с преобразованием, качество восстановления, помехоустойчивость.

Введение

Актуальность исследования. За последнюю четверть века цифровая обработка видеоданных превратилась в мощную индустрию с годовым оборотом около триллиона долларов, в основе которой лежат международные стандарты сжатия, или кодирования изображения JPEG, MPEG. Эти стандарты в целом обеспечивают широкий спектр противоречивых требований к информационным возможностям видеокодеков, основными среди которых являются степень сжатия, качество восстановления, помехоустойчивость и быстродействие. Однако эти требования неуклонно возрастают, и противоречие между ними обостряется. Во-первых, по мере перехода к форматам видеоданных высокой четкости (HD) и следующего поколения (3D, голографическое телевидение) растет требуемое качество изображений и их разрешающая способность. Во-вторых, возрастает необходимая степень сжатия видеоданных, особенно в мобильном и сетевом сегменте мультимедийных приложений с присущим ему жестким ограничением пропускной способности канала передачи данных. В-третьих, растущая изощренность обработки обычно влечет усложнение кодека и снижение его быстродействия. Повсеместное внедрение более эффективного стандарта H.264 сможет снизить остроту этого противоречия только временно, на ближайшие 5...10 лет. Дальнейший прогресс в этой области прямо связан с разработкой нового поколения методов и алгоритмов кодирования видеоданных.

Обзор существующих методов кодирования. Современный арсенал методов кодирования видео-

данных охватывает широкий круг идей и включает импульсно-кодированную модуляцию (ИКМ), кодирование с предсказанием (дифференциальную импульсно-кодированную модуляцию), статистическое (или энтропийное) кодирование, групповое кодирование (или кодирование с преобразованием) и ряд других методов [1, 2].

Кодирование изображения с помощью ИКМ является первым исторически освоенным методом формирования цифрового изображения и сводится к последовательно выполняемым операциям ограничения спектра аналогового изображения, его дискретизации, квантования дискретных отсчетов (пикселей) изображения и представления их в виде бинарных слов. ИКМ не позволяет существенно сократить объем исходной информации, поскольку не учитывает корреляционные связи между пикселями. Кодирование с предсказанием, в отличие от ИКМ, использует межпиксельные корреляции и заключается в построении прогноза значения каждого текущего пикселя по совокупности соседних и формировании разностного изображения, представляющего собой ошибку прогноза относительно истинного изображения. При этом кодированию и передаче подлежит именно разностный сигнал, а при декодировании исходное изображение восстанавливается из разностного по известному алгоритму прогноза. При точном прогнозе ошибка оказывается в среднем малой и кодируется более коротким кодом, чем исходное изображение при ИКМ, что и обуславливает сжатие данных. Однако и в этом случае возможности сокращения цифрового потока ограничены, главным образом, попиксельным механизмом предсказания. Другой подход к учету статистических

свойств изображения лежит в основе энтропийного кодирования. Поскольку квантованные значения яркости пикселей изображения неравновероятны, сокращение объема сохраняемых данных об изображении возможно путем кодирования уровней яркости словами переменной длины: более вероятные значения яркости кодируются словами с меньшим числом бит, а менее вероятные – с большим. В отличие от этих методов групповое кодирование основано на разбиении изображения на блоки и одновременном кодировании всей группы N_0 пикселей, составляющей блок. При таком методе обычно используют некоторое преобразование блока данных, а повышение эффективности кодирования связано с тремя факторами: значительная часть элементов преобразованного блока (коэффициентов преобразования) N_1 становится настолько малыми по величине, что их можно отбросить без заметной потери качества восстанавливаемых изображений; коэффициенты оказываются значительно декоррелированными в сравнении с элементами блока до преобразования, чем достигается повышение эффективности их дальнейшего энтропийного кодирования и увеличение степени сжатия; нелинейное квантование коэффициентов обеспечивает существенное дополнительное сокращение данных. Наиболее известные из других методов кодирования – гибридные преобразования, кодирование с экстраполяцией и интерполяцией, векторное и фрактальное кодирование. Фрактальное кодирование является одним из самых эффективных по степени сжатия, однако временные затраты при кодировании на несколько порядков больше, чем при декодировании, что затрудняет применение этого метода в приложениях реального времени.

В целом высокая эффективность кодирования с преобразованием привела к тому, что именно этот метод положен в основу всех действующих стандартов сжатия видеоданных. Наиболее распространенным вариантом преобразования в настоящее время является дискретное косинусное преобразование (ДКП). Тем не менее, по указанным выше причинам продолжают исследования, направленные на разработку более эффективных преобразований. В качестве одного из возможных претендентов на эту роль в последнее время активно обсуждается множество дискретных ортогональных моментов изображения, сформированных на основе дискретных многочленов Рака [3] и называемое в настоящей статье преобразованием Рака (ПР).

Цель статьи: Ввести матричное представление преобразования Рака и исследовать его информационные возможности как средства кодирования изображения в сравнении с другими преобразованиями, традиционно используемыми в этой области.

1. Теоретическая часть

1.1. Многочлены Рака и их свойства

Многочлены Рака и родственные им b_j -символы Вигнера возникли в контексте квантовой теории атомных спектров как коэффициенты векторного сложения угловых моментов j_1, j_2, j_3 трех связанных квантовомеханических систем [4, 5]. Они являются ортогональными многочленами дискретной переменной на квадратичной сетке $x(s) = s(s+1)$ [6]. При специальной нормировке они модифицируются во взвешенные многочлены Рака $\hat{u}_n^{(\alpha, \beta)}(s, a, b)$, $n = 0, 1, \dots, M$ и образуют полную ортонормальную систему базисных функций (ПОНС) [3, 6]. Кроме отмеченных выше важнейшими свойствами этих функций являются связь с другими типами специальных функций, включая различные семейства ортогональных многочленов дискретной переменной (Чебышева, Кравчука, Хана и др.), разнообразные виды симметрии, рекуррентные соотношения относительно n и s [3, 5, 6], что значительно упрощает их формирование. В настоящем исследовании используется соотношение между взвешенными многочленами Рака и b_j -символами

Вигнера $W = \begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{23} \end{Bmatrix}$ [5, 6]:

$$\hat{u}_n^{(\alpha, \beta)}(s, a, b) = (-1)^k \sqrt{(2j_{12} + 1)(2j_{23} + 1)} \cdot W, \quad (1)$$

где

$$k = j + j_1 + j_{23},$$

$$j_{12} = n + \frac{\alpha + \beta}{2},$$

$$j_{23} = s,$$

$$j = \frac{a + b + \alpha - \beta - 1}{2},$$

$$j_1 = \frac{\alpha + \beta + b - a - 1}{2},$$

$$j_{12} = \frac{b - a - 1}{2},$$

$$j_{23} = \frac{a + b - 1}{2}.$$

1.2. Совокупность моментов Рака сигнала как линейное унитарное преобразование

Исследование применимости многочленов Рака и некоторых других типов ортогональных многочленов дискретной переменной для анализа изображений предпринято в [3]. Там же содержится библиографический обзор результатов, достигнутых в этой области.

Преобразованием Рака (ПР) будем называть множество моментов Рака, полученных с помощью взвешенных многочленов Рака.

В одномерном случае момент Рака n -го порядка определяется как

$$v_n = \sum_{s=a}^{b-1} \hat{u}_n^{(\alpha,\beta)}(s, a, b) f_{s-a+1}, \quad n = 0, 1, \dots, M, \quad (2)$$

где $f = (f_1, f_2, \dots, f_N)^T$ – одномерный сигнал, вектор длины $N = b - a - 1$. Восстановление сигнала по множеству моментов Рака $v = (v_0, v_1, \dots, v_M)^T$ выполняется как

$$f_{s-a+1} = \sum_{n=0}^M v_n \hat{u}_n^{(\alpha,\beta)}(s, a, b), \quad s = a, a+1, \dots, b-1. \quad (3)$$

Переходя к векторно-матричному представлению и вводя матрицу ПР R и обратную ей R^{-1} :

$$R = \|r_{nm}\|, \quad r_{nm} = \hat{u}_n^{(\alpha,\beta)}(m+a, a, b), \quad (4)$$

где $n = 0, 1, \dots, M$, $m = 0, 1, \dots, b - a - 1$, прямое и обратное ПР имеем в виде

$$v = Rf, \quad f = R^{-1}v. \quad (5)$$

Благодаря тому, что взвешенные многочлены Рака образуют ПОНС, матрица R оказывается унитарной $R^{-1} = R^T$.

Именно это обстоятельство позволяет рассматривать ПР в числе возможных средств кодирования сигнала, поскольку унитарность гарантирует обратимость (то есть возможность восстановления сигнала) и сохранение энергии при преобразовании.

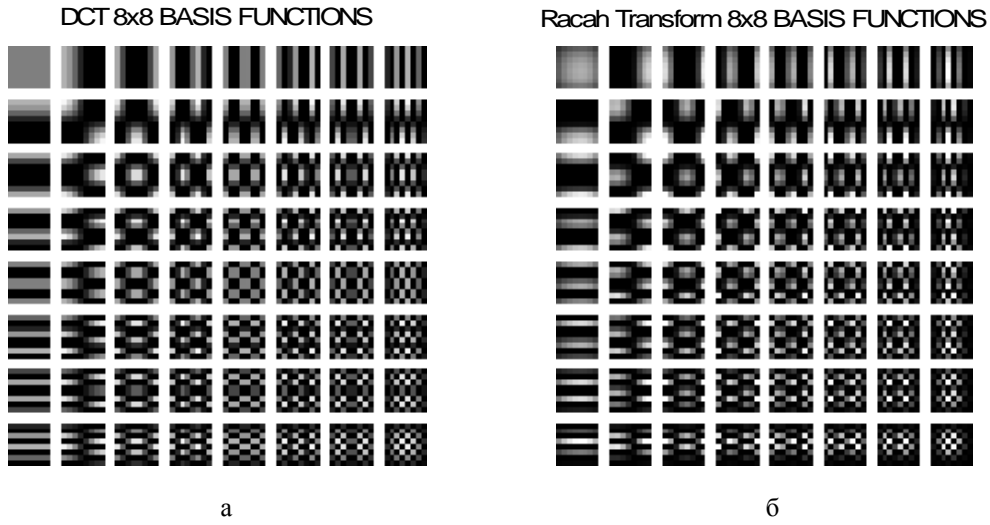


Рис. 1. Двумерные базисные функции для блока 8×8 элементов: а – ДКП, б – ПР при $\alpha = \beta = 1, a = 24$

В двумерном случае смешанный момент Рака $(n + m)$ -го порядка определяется как

$$v_{n,m} = \sum_{s=a}^{b-1} \sum_{t=a}^{b-1} \hat{u}_n^{(\alpha,\beta)}(s, a, b) \hat{u}_m^{(\alpha,\beta)}(t, a, b) f_{s-a+1, t-a+1}. \quad (6)$$

где $F = \|f_{n,n}\|$, $n, m = 0, 1, \dots, M$ – матрица изображения. Восстановленное изображение формируется как обратное двумерное ПР матрицы моментов Рака $V = \|v_{n,m}\|$, $n, m = 0, 1, \dots, M$:

$$f_{s-a+1, t-a+1} = \sum_{n=0}^M \sum_{m=0}^M v_{n,m} \hat{u}_n^{(\alpha,\beta)}(s, a, b) \hat{u}_m^{(\alpha,\beta)}(t, a, b), \quad (7)$$

$$n, m = 0, 1, \dots, M.$$

В векторно-матричном представлении двумерное прямое и обратное ПР имеют вид

$$V = R(RF)^T = RF^T R^T, \quad (8)$$

$$F = R^T \cdot (R^T V)^T = R^T V^T R. \quad (9)$$

2. Сравнительное исследование информационных возможностей

2.1. Информационные возможности кодирования с преобразованием

Под информационными возможностями видеокодеков вообще и методов и алгоритмов кодирования с преобразованием в частности будем понимать: а) степень сжатия, б) качество восстановления, в) помехоустойчивость, г) вычислительную эффективность, количественно оцениваемые соответственно [1, 2, 3]: а) отношением числа коэффициентов преобразования блока изображения размером $N \times N$ элементов до сжатия N_0 и после отбрасывания N_1 несущественных коэффициентов

$$C = N_0 / (N_0 - N_1), \quad N_0 = N^2, \quad (10)$$

б) среднеквадратической ошибкой восстановления (СКО):

$$MSE = \frac{1}{N^2} \sum_{s,t=1}^N (f_{s,t} - \tilde{f}_{s,t})^2, \quad (11)$$

где $\tilde{F} = \|\tilde{f}_{s,t}\|$ – восстановленное изображение размером $N \times N$,

в) зависимостью СКО MSE и от интенсивности шума P при заданном типе шума,

$$MSE = MSE(P); \tag{12}$$

г) числом вычислительных операций.

2.2. Характеристика и результаты исследования

Сравнительная эффективность ПР, ДКП и дискретного преобразования Фурье (ДПФ) исследовалась в смысле (10), (11), (12) при $N = 8$ применительно к тестовым монохромным полутоновым изображениям с длиной кода пиксела 8 бит из библиотеки Матлаба. Дополнительная обработка изображений (таблицы квантования, неравномерное квантование, сглаживание и т.п.) не применялась.

На рис. 1 приведены двумерные базисные функции ДКП и ПР при $N = 8$.

Матрица ПР рассчитывалась в соответствии с (4) при $\alpha = \beta = 1$, $a = 24$. Как видно, базисы этих преобразований весьма близки.

Результаты исследования приведены на рис. 2, 3.

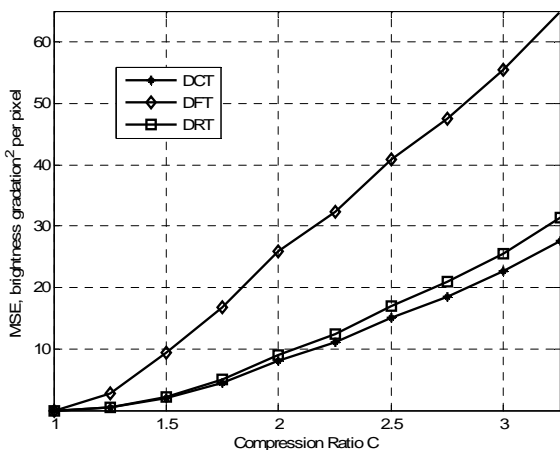


Рис. 2. СКО восстановления для ДКП (DCT), ДПФ (DFT) и ПР (DRT) при $\alpha = \beta = 1$, $a = 24$ в зависимости от коэффициента сжатия MSE(C)

Как следует из анализа зависимостей, при фиксированной степени сжатия качество восстановления ПР значительно (примерно вдвое) превосходит ДПФ, несколько уступая ДКП, а помехоустойчивость практически совпадает у всех трех преобразований. Это противоречит данным [3], где отмечается существенное преимущество ПР перед ДКП по степени сжатия, качеству восстановления и помехоустойчивости.

Поскольку порядок расчета степени сжатия в [3] не указан, расхождение результатов можно было бы объяснить различием в оценке этого показателя

и выбором параметров ПР α , β , a . Но помехоустойчивость ПР при $\alpha = a = 5$ и $\beta = 0$ [3] оказалась равной приведенной на рис. 3, а более чем двукратное сокращение СКО при ПР в сравнении с ДКП при фиксированной степени сжатия по данным [3] вызывает удивление.

В настоящей статье расчеты ПР выполнялись в соответствии с (8), (9). Таким образом, число операций определяется числом умножений трех вещественных матриц размера $N \times N$ и составляет $O(N^3)$

(порядка N^3) операций вещественного умножения и сложения. Это стандартный показатель при отсутствии быстрого варианта преобразования, такого как быстрое преобразование Фурье, который реализуется за $O(N^2 \log_2 N)$ операций.

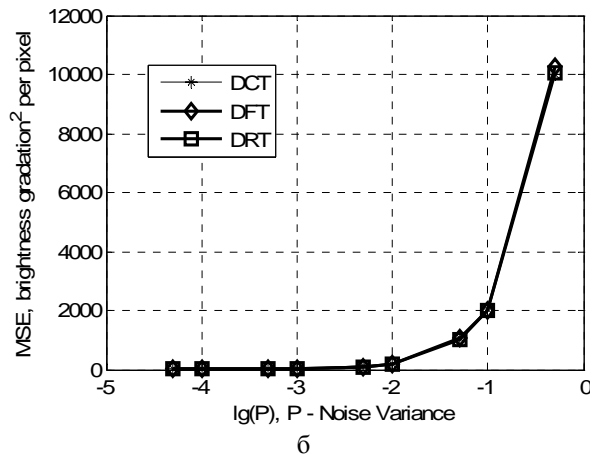
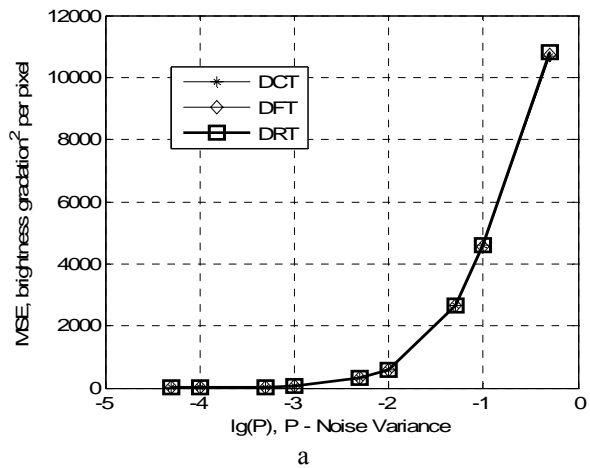


Рис. 3. Помехоустойчивость ДКП (DCT), ДПФ (DFT) и ПР (DRT) при $\alpha = \beta = 1$, $a = 24$. СКО восстановления в зависимости от дисперсии шума MSE(P) : а – гауссов шум; б – шум «соль с перцем»

О разработке быстрого алгоритма ПР не сообщалось, однако можно ожидать, что упомянутые выше свойства симметрии и рекурсии коэффициен-

тов Рака и символов Вигнера допускают его существование.

Заметим, что многочлены $\hat{u}_n^{(\alpha, \beta)}(s, a, b)$ – трехпараметрические функции. Это открывает дополнительные возможности для построения на их основе приложений, адаптирующихся к изменениям условий анализа, что является большим достоинством при унифицированном способе формирования матриц ПР.

Выводы

Выполненное в статье исследование в целом подтверждает наличие у преобразования Рака свойств, ценных для анализа изображений.

Это преобразование нуждается в дальнейшем исследовании, в частности целесообразно рассмотреть обеспечиваемую им степень декорреляции элементов матрицы в сравнении ДКП, ДПФ и другими преобразованиями, а также в сравнении с преобразованием Карунена-Лоэва, обеспечивающим полную парную декорреляцию пикселей и максимальную степень сжатия [1, 2].

Вызывают интерес возможности дополнительного сжатия при неравномерном и неоднородном квантовании коэффициентов преобразования Рака. Актуально изучение возможности построения быстрого алгоритма этого преобразования, а также систематическое исследование характера влияния параметров многочленов Рака на возможности анализа изображения в интересах создания унифицированной многофункциональной адаптивной процедуры.

Список литературы

1. Цифровая обработка телевизионных и компьютерных изображений / Под ред. Ю.Б. Зубарева, В.П. Дворковича. – М.: Международный центр научной и технической информации, 1997.
2. Залмансон Л.А. Преобразования Фурье, Уолша, Хаара и их применение в управлении, связи и других областях / Л.А. Залмансон. – М.: Наука, 1989.
3. Hongqing Zhu. Image analysis by discrete orthogonal Racah moments / Hongqing Zhu, Huazhong Shu, Jun Liang, Lumin Luo, Jean-Louis Coatrieux // Signal Processing. – April, 2007. – V. 87, No. 4. – P. 687-708.
4. Вигнер Е. Теория групп и ее приложения к квантовомеханической теории атомных спектров / Е. Вигнер. – М.: Издательство иностранной литературы, 1961.
5. Вариалович Д.А. Квантовая теория углового момента / Д.А. Вариалович, А.Н. Москалев, В.К. Херсонский. – Л.: Наука, 1975.
6. Никифоров А.Ф. Классические ортогональные полиномы дискретной переменной / А.Ф. Никифоров, С.К. Сулов, В.Б. Уваров. – М.: Наука, 1985.

Поступила в редколлегию 7.07.2011

Рецензент: д-р техн. наук, проф. С.М. Порошин, Национальный технический университет «Харьковский политехнический институт», Харьков.

КОДУВАННЯ ЗОБРАЖЕННЯ З ПЕРЕТВОРЕННЯМ НА ОСНОВІ МОМЕНТІВ РАКА

А.В. Статкус

Дискретні ортогональні моменти є потужним засобом опису ознак форми зображень у додатках розпізнання образів та аналізу зображень. У даній статті розглядається нова множина дискретних ортогональних моментів, що базуються на зважених дискретних поліномах Рака. Сукупність цих моментів породжує лінійне унітарне перетворення зображення, що в статті отримує назву перетворення Рака. Завдяки унітарності воно задовольняє вимогам до перетворень щодо придатності до стиснення зображень. Обговорюються декотрі властивості поліномів Рака та результати експериментальних досліджень інформаційних можливостей перетворення Рака як засобу кодування зображень в порівнянні з традиційними перетвореннями з точки зору ступеня стиснення, якості відновлення, перешкодостійкості та обчислювальної ефективності.

Ключові слова: стиснення зображень, кодування з перетворенням, якість відновлення, перешкодостійкість.

IMAGE CODING WITH TRANSFORM BASED ON RACAH MOMENTS

A.V. Statkus

Discrete orthogonal moments (DOM) are powerful tool for characterization of image shape features in applications of pattern recognition and image analysis. In this paper, a new set of DOM based on the weighted discrete Racah polynomials are considered. This set generate linear unitary transform of initial image called in the paper as Racah Transform (RT). Due to its unitarity, the RT meets formal requirements to transform for use in image compression. Some properties of Racah polynomials and results of comparative experimental research of an information performance of the RT versus the Discrete Fourier and the Discrete Cosinus Transforms are discussed in terms of compression ratio, quality of reconstruction, noise robustness, and computational efficiency.

Keywords: image compression, coding with transform, quality of reconstruction, noise robustness.

