

### ПРИМЕНЕНИЕ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ ДЛЯ СЖАТИЯ ИЗОБРАЖЕНИЯ

В последнее десятилетие в мире наблюдается значительный интерес к сжатию изображений. Это вызвано стремительным развитием вычислительной техники, графических мониторов, цветных принтеров, а также цифровой техники связи. Изображение представляется в цифровом виде достаточно большим количеством бит. Так, цветная картинка размером 512x512 требует для своего хранения 768 кБайт. Если передавать видеопоследовательность таких картинок со скоростью 25 кадров в секунду, требуемая скорость составит 188.7 Мбит/с.

Различают сжатие изображений без потерь и с потерями. Первое характеризуется незначительными коэффициентами сжатия (от 3 до 5 раз) и находит применение в телевидении, медицине, аэрофотосъемке и других приложениях. При сжатии изображения с допустимыми потерями коэффициент сжатия может достигать сотен раз. Популярность вейвлет-преобразования (ВП) во многом объясняется тем, что оно успешно может использоваться для сжатия изображения как без потерь, так и с потерями. Так, коэффициент сжатия видеосигнала в видеокодеках семейства ADV6xx варьируется от 3 до 350 и больше раз.

Причин успешного применения несколько.

1. Известно, что вейвлет-преобразование хорошо аппроксимирует преобразование Карунена-Лоэва для фрактальных сигналов, к которым относятся и изображения.

2. Дисперсии коэффициентов субполос ортонормального вейвлет-преобразования распределены в широком диапазоне значений. Пусть дисперсии кодируются простым энтропийным кодером. Тогда стоимость кодирования всего изображения есть сумма кодирования субполос. Различные энтропии субполос приведут к стоимости кодирования значительно меньшей, чем при непосредственном кодировании изображения.

3. В результате этого перераспределения дисперсий коэффициенты вейвлет-преобразования имеют существенно негауссовскую статистику и, таким образом, меньшую энтропию, чем гауссовский сигнал той же дисперсии.

4. Наконец, коэффициенты вейвлет-декомпозиции имеют регулярные пространственно-частотные зависимости, которые с успехом используются в ряде алгоритмов кодирования.

Рассмотрим основные проблемы, возникающие при сжатии изображения при помощи вейвлет-преобразования, и возможные пути их решения.

### **10.1. Базовый вейвлет-кодер изображения**

Вейвлет-кодер изображения устроен так же, как и любой другой кодер с преобразованием. Назовем такой кодер базовым. Он состоит из трех основных частей: декоррелирующее преобразование, процедура квантования и энтропийное кодирование. В настоящее время во всем мире проводятся исследования по усовершенствованию всех трех компонент базового кодера. В следующих разделах будут рассмотрены современные, более сложные и эффективные кодеры. Так как в их основе лежит базовый кодер, рассмотрим подробнее его составляющие части.

#### **10.1.1. Выбор вейвлетов для сжатия изображения**

Выбор оптимального базиса вейвлетов для кодирования изображения является трудной и вряд ли решаемой задачей. Известен ряд критериев построения «хороших» вейвлетов, среди которых наиболее важными являются: гладкость, точность аппроксимации, величина области определения, частотная избирательность фильтра. Тем не менее, наилучшая комбинация этих свойств неизвестна.

Простейшим видом вейвлет-базиса для изображений является разделимый базис, получаемый сжатием и растяжением одномерных вейвлетов. Использование разделимого преобразования сводит проблему поиска эффективного базиса к одномерному случаю, и почти все известные на сегодняшний день кодеры используют его. Однако неразделимые базисы могут быть более эффективными, чем разделимые.

Прототипами базисных функций для разделимого преобразования являются функции  $\phi(x)\phi(y)$ ,  $\phi(x)\psi(y)$ ,  $\psi(x)\phi(y)$  и  $\psi(x)\psi(y)$ . На каждом шаге преобразования выполняется два разбиения по частоте, а не одно. Предположим, имеем изображение размером  $N \times N$ . Сначала каждая из  $N$  строк изображения делится на низкочастотную и высокочастотную половины. Получается два изображения размерами  $N \times N/2$ . Далее, каждый столбец делится аналогичным образом. В результате получается четыре изображения размерами

$N/2 \times N/2$ : низкочастотное по горизонтали и вертикали, высокочастотное по горизонтали и вертикали, низкочастотное по горизонтали и высокочастотное по вертикали и высокочастотное по горизонтали и низкочастотное по вертикали. Первое из вышеназванных изображений делится аналогичным образом на следующем шаге преобразования, как показано на рис.10.1.

Известно, что для кодирования изображений хорошо подходят сплайновые вейвлеты. Эксперименты, проведенные рядом исследователей, показывают важность гладкости базисных функций для сжатия. Практически столь же большое значение имеет число нулевых моментов вейвлетов, которое тесно связано с гладкостью. Несмотря на это, некоторые исследователи считают, что важность гладкости для приложений цифровой обработки сигналов остается открытым вопросом. Наиболее широко на практике используют базисы, имеющие от одной до двух непрерывных производных. Увеличение гладкости не приводит к увеличению эффективности кодирования.

|         |         |         |
|---------|---------|---------|
| $HCH_2$ | $VCH_2$ | $VCH_1$ |
| $HCV_2$ | $VCV_2$ |         |
| $HCV_1$ |         | $VCV_1$ |

Рис.10.1. Два уровня вейвлет-преобразования изображения

Д.Вилласенор систематически протестировал все биортогональные блоки фильтров минимального порядка с длиной фильтров  $\leq 36$ . В дополнение к вышеперечисленным критериям учитывалась также чувствительность аппроксимации с низким разрешением к сдвигам функции  $f(x)$ . Наилучшим фильтром, найденным в этих экспериментах, оказался сплайновый фильтр

7/9. Этот фильтр наиболее часто используется в вейвлет-кодерах изображений. В частности, в видеокодеках семейства ADVbxx применяются именно эти фильтры.

Необходимо сделать одно замечание относительно этих результатов. Д.Вилласенор сравнивал пиковое отношение сигнал/шум, получаемое при использовании различных фильтров в простой схеме кодирования. Алгоритм размещения бит, применяемый им, хорошо работает с ортогональными базами. В случае биортогональных фильтров должен применяться другой, более эффективный алгоритм. В силу этой причины некоторые заслуживающие внимания биортогональные фильтры были им обойдены.

Для биортогонального преобразования квадрат ошибки в области преобразования не равен квадрату ошибки в восстановленном изображении. В результате проблема минимизации ошибки становится намного более трудной, чем в ортогональном случае. Можно уменьшить ошибку в области изображения путем применения схемы взвешенного распределения бит, рассматриваемой в пункте 10.1.3. Тогда целый ряд фильтров по своей эффективности становится равным фильтру 7/9. Один из таких базисов – интерполирующий вейвлет Деслаури-Дубук порядка 4, преимуществом которого является то, что коэффициенты фильтра - рациональные числа, кратные степени 2. Оба этих вейвлета имеют 4 нулевых момента и две непрерывные производные.

Семейство многообещающих фильтров было разработано И.Баласингамом и Т.Рамстадом. Процедура разработки заключалась в комбинировании классических методов разработки фильтров с идеями теории вейвлетов. Получившиеся фильтры значительно превосходят популярные фильтры 7/9.

### **10.1.2. Осуществление преобразования на границах изображения**

Для эффективного сжатия необходимо тщательно обрабатывать границы изображения. Методы осуществления вейвлет-преобразования, учитывающие границы, были описаны в главе 3. Альтернативным методом является конструирование граничных фильтров, сохраняющих ортогональность преобразования вблизи границы. Проблеме конструирования граничных фильтров посвящен ряд статей Е.Ковачевич. Как было показано в главе 6, при применении лифтинговой схемы границы учитываются автоматически.

### 10.1.3. Квантование

В большинстве вейвлет-кодеров применяется скалярное квантование. Существуют две основные стратегии выполнения скалярного квантования. Если заранее известно распределение коэффициентов в каждой полосе, оптимальным будет использование квантователей Ллойда-Макса с ограниченной энтропией для каждой субполосы. В общем случае подобным знанием мы не обладаем, но можем передать параметрическое описание коэффициентов путем посылки декодеру дополнительных бит. Априорно известно, что коэффициенты высокочастотных полос имеют обобщенное гауссовское распределение с нулевым матожиданием.

На практике обычно применяется намного более простой равномерный квантователь с «мертвой» зоной. Как показано на рис.10.2, интервалы квантования квантования имеют размер  $\Delta$ , кроме центрального интервала (возле нуля), чей размер обычно выбирается  $2\Delta$ .

Коэффициенту, попавшему в некоторый интервал, ставится в соответствие значение центра этого интервала. В случае асимптотически высоких скоростей кодирования равномерное квантование является оптимальным. Хотя в практических режимах работы квантователи с «мертвой» зоной субоптимальны, они работают почти так же хорошо, как квантователи Ллойда-Макса, будучи намного проще в исполнении. Кроме того, они робастны к изменениям распределения коэффициентов в субполосе. Дополнительным их преимуществом является то, что они могут быть вложены друг в друга для получения вложенного битового потока.

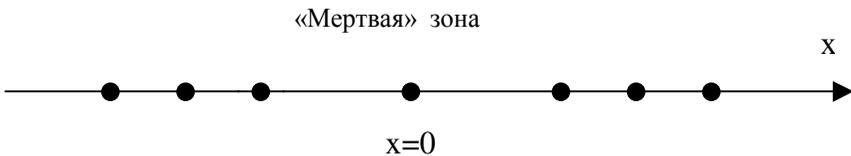


Рис. 10.2. Равномерный квантователь с «мертвой» зоной

### 10.1.4. Энтропийное кодирование

Субоптимальное энтропийное кодирование коэффициентов можно осуществить при помощи алгоритма арифметического кодирования. Кодеру требуется оценить распределение квантованных коэффициентов. Эта оценка

получается путем аппроксимации распределения коэффициентов гауссовской или лапласовской плотностью и вычисления параметров распределения. Оценка параметров может также производиться и в процессе работы, «на ходу». Такой подход имеет то преимущество, что кодер учитывает локальные изменения статистики изображения. Известны эффективные адаптивные процедуры оценивания.

Так как изображение не является случайным гауссовским процессом, коэффициенты преобразования, хотя и некоррелированные, обладают определенной структурой. Энтропийный кодер может использовать эту структуру, осуществляя некоторое предсказание. В ряде работ отмечено, что применение предсказания приводит к незначительному повышению эффективности.

На практике зачастую вместо арифметического кодера используют кодер Хаффмана. Причина этого заключается в меньшем требуемом объеме вычислений, а также в том, что алгоритмы арифметического кодирования запатентованы. Так, только фирма IBM обладает более чем 90 патентами различных вариаций этого кодера. В силу этого в видеокодеках ADVbxx применен кодер Хаффмана.

### 10.1.5. Распределение бит

Итак, последним вопросом, на который необходимо ответить при создании алгоритма сжатия, является следующий: как точно квантовать каждую из субполос? На этот вопрос дает ответ алгоритм распределения бит. Общая идея заключается в определении такого числа бит  $R_j$ , отводимых для кодирования  $j$  субполосы, при котором суммарное искажение  $\sum_j D_j(R_j)$  было бы минимальным с учетом ограничения  $\sum_j R_j \leq R$ . Если известен точный вид функции  $D_j(R)$ , проблема решается с использованием условий Куна-Тукера. Одно из решений заключается в аппроксимации функции  $D_j(R)$  функцией скорость-искажение для гауссовского источника. Однако при низких скоростях кодирования эта аппроксимация будет неточна. Лучшие результаты могут быть получены путем измерения  $D_j(R)$  в диапазоне изменения  $R$  и решения проблемы ограниченной оптимизации с применением метода целочисленного программирования. Данная задача была решена И.Шохамом и А.Гершо.

В случае применения биортогональных вейвлетов возникает дополнительная трудность, заключающаяся в неравенстве среднеквадратической

ошибки (СКО) области изображения и области трансформанты. П.Мулином был сформулирован алгоритм кратномасштабного ослабления, который дает приближенное решение этой проблемы. Данный алгоритм значительно эффективнее обычной минимизации СКО в области трансформанты. В главе 10 будет описан метод распределения бит, примененный в ADV6xx.

Более простым методом является аппроксимация СКО изображения взвешенной суммой СКО субполос. Вес  $\omega_j$  для субполосы  $j$  находится следующим образом: устанавливаем один коэффициент этой полосы в 1, а остальные – в 0. Затем выполняем обратное преобразование. Вес  $\omega_j$  равен сумме квадратов получившихся значений. Распределение бит производится с целью минимизации взвешенной суммы  $\sum_i \omega_j D_j(R_j)$ . Процедура взвешивания дает хорошие результаты, когда используются неортогональные вейвлеты, например вейвлеты Деслари-Дубук, ставшие популярными благодаря лифтинговой схеме (глава 6). Для фильтра 7/9 веса  $\omega_j$  близки к 1, поэтому взвешивание в данном случае нецелесообразно.

#### **10.1.6. Меры искажения, взвешенные с учетом восприятия человеком**

СКО не всегда хорошо согласуется с визуально наблюдаемой ошибкой. Рассмотрим, например, два изображения, которые полностью одинаковы, кроме небольшой области. Хотя визуально разность между этими изображениями хорошо заметна, СКО будет примерно одинаковой. Учет системы человеческого зрения в схеме сжатия является трудной задачей. Было проведено множество исследований, но в силу трудностей с математическим описанием системы зрения человека подходящей меры найдено не было.

Известно, что в человеческом глазу выполняется операция многомасштабного представления изображений. Глаз более чувствителен к искажениям в низкочастотной области. Отсюда существует возможность улучшения визуального качества реконструированного изображения путем взвешивания СКО субполос в соответствии с чувствительностью глаза в различных частотных диапазонах. Веса для наиболее часто используемого фильтра 7/9 были вычислены А.Ватсоном.

## 10.2. Новые идеи в области сжатия изображений, связанные с вейвлет-преобразованием

Базовый вейвлет-кодер, описанный в разделе 10.1, использует общие принципы кодера с преобразованием, то есть основан на эффектах декорреляции и перераспределения энергии. Математическая теория вейвлет-преобразования позволяет создавать совершенно новые и эффективные методы сжатия. Эти методы лежат в основе алгоритмов, описываемых в разделах 10.3 и 10.5. В данном разделе покажем главные идеи этих методов.

Кодирование с преобразованием основано на том, что большая часть энергии сосредоточивается в малом количестве коэффициентов, которые квантуются в соответствии с их значением. Эта парадигма, являясь достаточно мощной, основывается на нескольких предположениях, которые не всегда верны. В частности, предполагается, что изображение порождается гауссовским источником, что не соответствует действительности. С.Маллат и Ф.Фальзон показали, как это несоответствие приводит к неверным результатам при кодировании с низкими скоростями.

Пусть  $Y[n]$  - случайный вектор длиной  $N$ , определенный как

$$Y[n] = \begin{cases} X, & \text{если } n = P, \\ X, & \text{если } n = P + 1 \pmod{N}, \\ 0, & \text{в остальных случаях.} \end{cases} \quad (10.1)$$

Здесь  $P$  есть случайная целая величина, равномерно распределенная от 0 до  $N - 1$ , а  $X$  - случайная величина, с равной вероятностью принимающая значения 1 и -1.  $X$  и  $P$  - независимы. Вектор  $Y$  имеет нулевое среднее и ковариационную матрицу с элементами

$$E\{Y[n]Y[m]\} = \begin{cases} \frac{2}{N} & \text{для } n = m, \\ \frac{1}{N} & \text{для } |n - m| \in \{1, N - 1\}, \\ 0 & \text{в остальных случаях.} \end{cases} \quad (10.2)$$

Ковариационная матрица является циркулянтной, так что преобразованием Карунена-Лоэва для нее является просто преобразование Фурье. Однако преобразование Фурье вектора  $Y$  очень неэффективно с точки зрения коди-

рования. Энергия на частоте  $k$  будет равна  $\left|1 + e^{2\pi \frac{k}{N}}\right|$ . Это означает, что энергия  $Y$  распределена по всей низкочастотной половине базиса Фурье и частично – по высокочастотной половине. Таким образом, преобразование Карунена-Лозва «упаковало» энергию двух ненулевых коэффициентов в примерно  $\frac{N}{2}$  коэффициентов. Конечно, было бы выгоднее кодировать  $Y$  в исходном виде, без всякого преобразования.

Как видно из этого примера, традиционное кодирование с преобразованием может быть улучшено путем введения операторов выбора. Вместо квантования коэффициентов трансформанты в заранее определенном порядке вейвлет-преобразование позволяет выбирать нужные для кодирования элементы. Это становится возможным главным образом благодаря тому, что базис вейвлетов компактен в частотной и пространственной областях. В вышеприведенном примере энергия сигнала была пространственно, но не частотно компактна. Значит, необходимо использовать соответствующий оператор выбора вейвлет-коэффициентов, наиболее эффективно представляющих сигнал. Наиболее значительным результатом этого подхода является создание алгоритма нульдеревья и его разновидностей (раздел 10.3).

Вообще говоря, развитие идей кодирования с преобразованием заключается в снятии ограничения на линейную аппроксимацию изображения, так как оператор выбора является нелинейным. В работах Р.Девора, С.Маллата и Ф.Фальзона показано, что проблема кодирования изображения может быть эффективно решена в рамках теории нелинейной аппроксимации. Отсюда возникает и ряд различий в алгоритмах работы традиционных и вейвлет-кодеров. В случае линейной аппроксимации изображение представляется фиксированным числом базисных векторов Карунена-Лозва. Далее, какое-то число малых коэффициентов трансформанты приравнивается к нулю. Идея нелинейной аппроксимации заключается в аппроксимации изображения путем адаптивного выбора самих базисных функций. Информация о выбранных базисных функциях хранится в бинарной карте значений и передается декодеру, как дополнительная информация. В разделе 10.3 будут описаны нульдеревья, являющиеся исключительно важной структурой данных для кодирования карты значений.

Рассмотренный выше пример показал что изображение неправомерно считать порожденным одиночным гауссовским источником. Для получения большей компактности энергии необходимо адаптировать преобразование к какому-то конкретному, а не к целому классу изображений. В случае если

источник описывается смесью различных распределений, преобразование Карунена-Лоэва не является больше эффективным. В главе 5 были описаны частотно-адаптивные и пространственно-частотно-адаптивные кодеры, в которых происходит разложение изображения в большое количество базисов и выбор из них оптимального по некоторому критерию.

Решетчатое квантование коэффициентов, рассматриваемое в разделе 10.5, гораздо ближе по своей сути к векторному квантованию, чем к кодированию с преобразованием.

Таблица 10.1

**Сравнение кодеров, описываемых в главе по отношению сигнал/шум**

| Тип кодера                              | LENA<br>(бит/пиксел) |      |      | BARBARA<br>(бит/пиксел) |      |      |
|---|----------------------|------|------|-------------------------|------|------|
|   | 1.0                  | 0.5  | 0.25 | 1.0                     | 0.5  | 0.25 |
| JPEG                                    | 37.9                 | 34.9 | 31.6 | 33.1                    | 29.3 | 25.2 |
| Оптимизированный JPEG                   | 39.6                 | 35.9 | 32.3 | 35.9                    | 30.6 | 26.7 |
| Базовый вейвлет-кодер                   | 39.4                 | 36.2 | 33.2 | 34.6                    | 29.5 | 26.6 |
| Нульдереве (Шапиро)                     | 39.6                 | 36.3 | 33.2 | 35.1                    | 30.5 | 26.8 |
| Нульдереве<br>(Саид и Перельман)        | 40.5                 | 37.2 | 34.1 | 36.9                    | 31.7 | 27.8 |
| Нульдереве<br>(R/D оптимизирован)       | 40.5                 | 37.4 | 34.3 | 37.0                    | 31.3 | 27.2 |
| Частотно-адаптивный                     | 39.3                 | 36.4 | 33.4 | 36.4                    | 31.8 | 29.2 |
| Пространственно-<br>частотно-адаптивный | 40.1                 | 36.9 | 33.8 | 37.0                    | 32.3 | 29.7 |
| Частотно-адаптивный +<br>нульдереве     | 40.6                 | 37.4 | 34.4 | 37.7                    | 33.1 | 29.3 |
| Решетчатое<br>квантование               | 41.1                 | 37.7 | 34.3 | -                       | -    | -    |
| Обратное оценивание                     | 40.9                 | 37.7 | 34.6 | -                       | -    | -    |

Итак, развитие идей кодирования с преобразованием заключается в основном во введении некоторого оператора выбора. Информация о выборе должна быть передана декодеру, как дополнительная информация. Она может быть в виде нульдеревьев или в виде обобщенных классов энергии. Метод «обратного оценивания распределения», предложенный К.Рамчандраном, основан на другом подходе. Считается, что дополнительная информация является избыточной и может быть получена декодером непосредственно из данных. Использование данного метода приводит к хорошим показателям кодирования.

В табл.10.1 (см. стр.155) представлены сравнения пикового отношения сигнал/шум для кодеров, которые будут обсуждаться далее. В качестве тестовых использовались полутонные портретные изображения размером 512x512.

Визуальное сравнение восстановленных изображений показывает, что лучшие результаты дают методы, использующие нульдеревья для кодирования коэффициентов. В частности, в этих изображениях лучше выражены контуры и отсутствует размытость мелких деталей.

### **10.3. Кодирование посредством нульдеревья**

Из теории кодирования с погрешностью известно, что оптимальное распределение бит достигается в случае, если сигнал поделен на субполосы, содержащие «белый» шум. Для реальных сигналов это достигается в случае неравномерной ширины субполос: в области НЧ они более узки, чем в области ВЧ. Вот почему вейвлет-преобразование обеспечивает компактность энергии.

Эта компактность энергии ведет к эффективному применению скалярных квантователей. Однако они не учитывают остаточную структуру, сохраняющуюся в вейвлет-коэффициентах, в особенности ВЧ субполос. Современные алгоритмы сжатия все тем или иным образом используют эту структуру для повышения эффективности сжатия.

Одним из наиболее естественных способов является учет взаимосвязей между коэффициентами из различных субполос. В высокочастотных субполосах имеются обычно большие области с нулевой или малой энергией. Области с высокой энергией повторяют от субполосы к субполосе свои очертания и местоположение. И это неудивительно – ведь они появляются вокруг контуров в исходном изображении – там, где вейвлет-преобразование не может адекватно представить сигнал, что приводит к «утечке» части энергии в ВЧ субполосы. Медленно изменяющиеся, гладкие области исходного изображения хорошо описывают НЧ вейвлет-базисы, что приводит к «упаковке» энергии в малом числе коэффициентов НЧ области. Этот процесс примерно повторяется на всех уровнях декомпозиции, что и приводит к визуальной «похожести» различных субполос.

Итак, априорное знание о том, что изображение состоит из гладких областей, текстур и контуров, помогает учитывать эту межполосную структуру. Кодеры, использующие структуру нульдеревя, сочетают учет структуры коэффициентов с совместным кодированием нулей, в результате чего получается очень эффективный алгоритм сжатия.

### 10.3.1. Алгоритм Льюиса и Ноулеса

Впервые идея нульдеревя была предложена А.Льюисом и Г.Ноулесом. В их алгоритме применялась древовидная структура данных для описания вейвлет-коэффициентов (рис.10.3). Такая структура получается в результате применения двухканального разделимого вейвлет-преобразования. Корневой узел дерева представляет коэффициент масштабирующей функции в самой НЧ области и имеет три отпрыска. Узлы дерева соответствуют вейвлет-коэффициентам масштаба, равного их высоте в дереве. Каждый из узлов имеет четыре отпрыска, соответствующих вейвлет-коэффициентам следующего уровня и того же пространственного расположения. Низом дерева являются листовые узлы, не имеющие отпрысков.

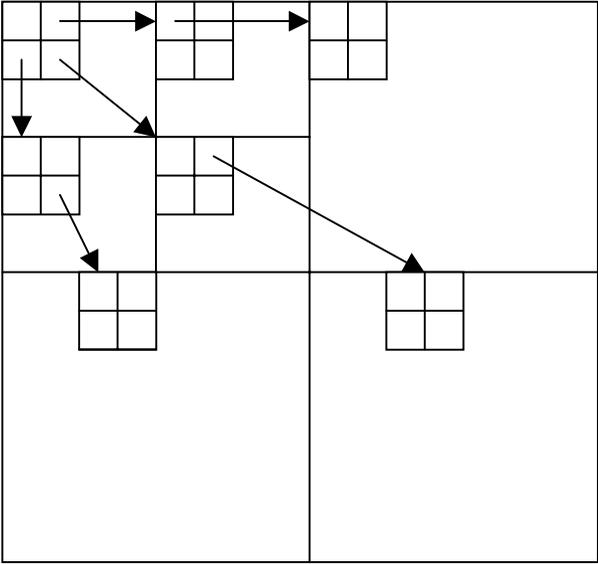


Рис.10.3. Зависимости между коэффициентами вейвлет-преобразования изображения, используемые в алгоритме нульдеревя

Для каждого из коэффициентов самой НЧ области существует три таких дерева, соответствующих трем порядкам фильтрации, как описано в пункте 10.1.1.

Квантование нульдерево основано на наблюдении, что если коэффициент мал, его отпрыски на дереве зачастую тоже малы. Это объясняется тем, что значимые коэффициенты возникают вблизи контуров и текстур, которые локальны. Нетрудно увидеть, что это является разновидностью предсказания. А.Льюис и Г.Ноулес свели это предсказание к минимуму, предположив, что если какой-либо коэффициент незначимый, то все его потомки также будут незначимыми. Дерево или субдерево, которое содержит (по крайней мере, так предполагается) только незначимые коэффициенты, называется нульдерево.

А.Льюис и Г.Ноулес использовали следующий алгоритм квантования вейвлет-коэффициентов. Вначале каждый узел квантуется квантователем, оптимальным для плотности распределения Лапласа. Если значение узла меньше некоторого порога, его потомки игнорируются. Эти потомки будут восстановлены декодером как нули. Иначе осуществляется переход к четырем отпрыскам узла, и процедура повторяется. Если узел не имеет отпрысков (является листом), начинает обрабатываться следующий корневой узел и т.д.

Данный алгоритм является эффективным в силу двух причин. Во-первых, в силу хорошей «упаковки» энергии вейвлет-преобразованием и, во-вторых, за счет совместного кодирования нулей. Для кодирования нулей обычно применяется кодер длин серий. Для повышения эффективности на вход этого кодера коэффициенты должны подаваться в определенном порядке. Например, в JPEG применено зигзагообразное сканирование. Наверное, наиболее важным вкладом этой работы была демонстрация того, что область вейвлет-коэффициентов прекрасно приспособлена для работы кодера длин серий. В самом деле, генерируются большие серии нулей и не надо передавать их длину, так как высота дерева известна. Аналогично JPEG, данный алгоритм является разновидностью скалярного/векторного квантования. Каждый (значимый) коэффициент квантуется отдельно, а символы, соответствующие малым коэффициентам, образуют вектор. Этот вектор состоит из символа нульдерева и последовательности нулей длиной до конца дерева.

Характеристики алгоритма Льюиса и Ноулеса незначительно превосходят JPEG, хотя визуальное качество изображений лучше. Недостатком алгоритма является способ порождения и распознавания нульдереьев. Как было отмечено, если коэффициент мал, то скорее всего его потомки будут малы, но

может быть, и нет. В случае если это не так, обнуляются значимые коэффициенты, и алгоритм Льюиса и Ноулеса ведет к большим искажениям.

Преимуществом этого алгоритма является его простота. Нульдеревья порождаются путем простого сравнения амплитуд коэффициентов, и не требуется дополнительной информации об их местоположении. Однако эта простота дается ценой невысокой эффективности. Детальный анализ этого взаимодействия привел к появлению следующего поколения кодеров с применением нульдеревьев.

### 10.3.2. Алгоритмы Шапиро и Саида-Перельмана

Идеи, лежащие в основе алгоритма Льюиса и Ноулеса, легли в основу многих современных кодеров изображения. Основным недостатком данного алгоритма является возможность ошибочного порождения нульдеревя, так как оно генерируется не из реальных данных, а на основе априорных предположений. Если потомки некоторого узла окажутся больше своего родителя (что нечасто, но все-таки бывает), алгоритм Льюиса и Ноулеса приводит к значительным искажениям.

Методы, обсуждаемые ниже, свободны от этого недостатка.

Шапиро разработал элегантный метод, названный алгоритмом вложенного нульдеревя (Embedded Zerotree Wavelet coder - EZW). Данный кодер основан на передаче и ненулевых данных, и карты значений. Биты, отводящиеся для кодирования карты значений, могут преваляировать в общем потоке бит, особенно на низких скоростях. Однако в карте значений, порождаемой изображением, существует очень большая избыточность, которая и используется для достижения малых скоростей кодирования. Если имеется незначительный родительский узел, то очень вероятно, что потомки также будут незначимы. Так что в большинстве случаев генерируется символ нульдеревя. Если вероятность этого события  $p$  близка к 1, то количество информации  $-p \log p$ , содержащееся в нем, близко к нулю. Значит, среднее число бит, требующихся для кодирования карты значений, мало.

Если один или более потомков незначимого узла является значимым, генерируется символ «изолированного нуля». Вероятность этого события ниже, следовательно, для кодирования требуется большее количество бит. Это плата за то, чтобы не допустить значительного искажения за счет ошибочного порождения нульдеревя.

Алгоритм EZW генерирует вложенный, иерархический код. Подобные кодеры позволяют осуществить прогрессивную передачу изображения с по-

следовательным уточнением его на приеме. При этом изображение вначале аппроксимируется небольшим количеством бит, а потом эта аппроксимация уточняется. Вложенный код имеет то свойство, что при  $R_1 > R_2$  код для  $R_2$  будет префиксом кода для  $R_1$ . Такие коды имеют большой практический интерес по следующим причинам:

- 1) возможность точного регулирования скорости передачи;
- 2) возможность восстановления всего изображения при прекращении приема декодером бит в любой точке. При этом изображение будет максимально хорошего качества для данного числа бит. Это применимо для передачи по каналам с потерями, а также для приложений вещания. В этом случае кодер генерирует высокоскоростной высококачественный поток, который передается по каналам различной пропускной способности декодерам различной вычислительной возможности. Последние выделяют из него нужные им субпотoki;
- 3) возможность быстрого просмотра изображений в удаленной базе данных. Для поиска достаточно и грубой копии, а при нахождении нужного изображения оно декодируется полностью.

Алгоритм Шапиро генерирует вложенный код побитовым способом (рис.10.4). В основе метода EZW лежат следующие основные операции.

Вначале выполняется частичное упорядочивание коэффициентов по амплитуде. Оно реализуется путем сравнения величины каждого вейвлет-коэффициента (ВК) с некоторым порогом  $T$ . Если  $ВК > T$ , то выносится решение о том, что коэффициент значимый, в противном случае – незначимый.

|   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | → | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | → | → | → | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | → | → | → | → | → | → | → | → | 1 | 1 | 1 | 1 | 1 |
| 1 | → | → | → | → | → | → | → | → | → | → | → | → | → |
| 0 | → | → | → | → | → | → | → | → | → | → | → | → | → |

Рис.10.4. Битовый план для сканирования упорядоченных вейвлет-коэффициентов

Сканирование производится от низкочастотных полос к высокочастотным. Для кодирования знака и позиции всех коэффициентов используется двухбитный символ. Этот символ может быть: « $\pm$ » - знак ВК; «0» – показывает, что ВК незначущий; «корень нулядеревя» - показывает, что ВК незначущий вместе со всеми ВК данной пространственной области из более высокочастотных полос. Таким образом, используется межполосная, пространственная корреляция ВК. После вычисления и передачи карты значений для значащих коэффициентов должны быть переданы биты, уточняющие их значение («карта данных»). Далее карта данных и карта значений сжимаются арифметическим кодером. В том случае, если не исчерпан ресурс скорости передачи, порог  $T$  делится на два и процесс повторяется.

Как видно из рис.10.4, верхние ряды бит содержат много нулей, так как многие коэффициенты имеют значение ниже порога. Роль нулядеревя заключается в предотвращении передачи этих нулей. Символ нулядеревя может снова и снова передаваться для данного коэффициента, пока он не станет больше текущего порога. После этого передается его квантованное значение.

А.Саид и В.Перельман улучшили алгоритм EZW. Их версия кодера называется «установка подразделений в иерархических деревьях» (Set Partition In Hierarchical Trees - SPIHT). Имеется общедоступная программная реализация этого кодера, которая очень быстра. Так, сжатие изображения размером  $512 \times 512$  в 100 раз занимает на компьютере P-166 порядка 0.1 секунды. При этом качество восстановленного изображения весьма приемлемо. Вложенные кодеры обладают одной интересной особенностью: чем больше коэффициент сжатия, тем меньше время работы кодера. Это объясняется тем, что требуется осуществление меньшего числа уточнений. SPIHT превосходит EZW примерно на 0.3-0.6 дБ за счет кодирования не одиночных, а параллельных нулядеревьев.

Можно показать, что EZW и SPIHT являются членами большого семейства алгоритмов, в которых карта значений имеет древовидную структуру.

### **10.3.3. Оптимизация нулядеревьев по критерию скорость-искажение**

В рассмотренных кодерах нулядеревья порождались только на основе анализируемых данных. Однако рассмотрим следующий гипотетический пример. Пусть изображение имеет большую равномерную область. Соответствующие ей вейвлет-коэффициенты будут малы, будет генерироваться нулядеревя, и на кодирование тратится малое число бит. Предположим теперь, что среди этой области имеется один резко отличающийся по значе-

нию пиксел. Этот пиксел приведет к появлению большого вейвлет-коэффициента, и нульдереву породиться не будет.

Неточное кодирование одного пиксела не приведет к большому искажению изображения. В нашем примере эффективность кодера может быть существенно повышена путем игнорирования соответствующего коэффициента и построения нульдеревя. Возникает вопрос: каким образом определять, стоит ли отбрасывать коэффициенты, «мешающие» построению нульдеревя.

Введение нульдеревя для группы вейвлет-коэффициентов является, по сути, разновидностью квантования. Значения коэффициентов, которые мы кодируем посредством нульдеревя, не являются в общем случае нулевыми. Значимые коэффициенты также подвергаются квантованию. Если сэкономить часть бит путем порождения больших нульдеревьев, высвободившийся ресурс бит можно направить на более точное квантование значимых коэффициентов. Задачей является оптимальное распределение ограниченного ресурса бит между двумя видами квантователей для достижения меньшего искажения.

Эта задача решена с использованием хорошо известного метода распределения бит. Основным утверждением является то, что для случая оптимального распределения бит наклоны касательных к кривым скорость-искажение для всех квантователей равны. Наклон показывает, насколько искажение увеличивается/уменьшается при обнулении/передаче данного узла. Если один из квантователей имеет меньший наклон, это означает, что при его передаче искажение уменьшится меньше, чем при передаче других узлов. Следовательно, можно передать часть бит от этого квантователя другим. Таким образом, при повторении этой процедуры наклоны всех квантователей будут выровнены.

Ясно, что нульдеревья влияют на уровни квантования ненулевых коэффициентов, так как общий ресурс бит ограничен. Верно и обратное. Поэтому возможен итеративный алгоритм для оптимизации этих двух режимов квантования по критерию скорость-искажение. Вначале фиксируется скалярный квантователь, и ищется оптимальное нульдеревя. Затем оно фиксируется, и ищется оптимальный скалярный квантователь. З.Ксионг было доказано, что эта процедура сходится к локальному оптимуму.

Данный алгоритм незначительно превосходит по эффективности SPIHT, но обладает серьезными недостатками. Во-первых, он намного более сложен. Во-вторых, и, наверное, самое главное, он не порождает иерархический поток бит.

## **10.4. Частотно, пространственно-частотно-адаптивные кодеры**

Данный класс кодеров использует адаптивные ортогональные преобразования, описанные в главе 5. Кодеры на базе вейвлет-пакетов были стандартизированы ФБР США, которое использует их для сжатия изображений отпечатков пальцев. Преимущество адаптивных кодеров проявляется при кодировании сильно нестационарных изображений. Недостатком является относительно высокая вычислительная сложность.

## **10.5. Использование зависимостей между вейвлет-коэффициентами внутри субполос**

Разработка кодера EZW дала импульс к активным исследованиям алгоритмов нульдеревя вейвлетов. Естественная простота структуры нульдеревя, ее вычислительные преимущества, возможность порождать иерархический код – вот наиболее привлекательные черты этих алгоритмов. Модификация алгоритма нульдеревя вейвлетов нашла свое применение и в новом стандарте на сжатие видео: MPEG-4.

Однако не следует думать, что только структура нульдеревя, или использование межполосных зависимостей, может привести к созданию эффективного кодера изображений. Некоторые из очень эффективных вейвлет-кодеров построены на совершенно других принципах. В данном разделе мы рассмотрим два метода, учитывающих внутриволосные зависимости между коэффициентами. Первый метод основан на концепции решетчатого квантования. Второй использует одновременно и внутри- и межполосные взаимосвязи между коэффициентами и рекурсивное оценивание дисперсии вейвлет-коэффициентов. Оба метода являются весьма эффективными.

### **10.5.1. Решетчатое квантование**

Решетчатое квантование (РК) является быстрым и эффективным методом квантования случайных переменных. РК использует корреляционные связи между переменными. При этом ячейки квантователя могут быть непрямоугольной формы, что приводит к недостижимой скалярными квантователями эффективности. Главные идеи РК взяты из работы Г. Унгербоэка по решетчатой модуляции. В данном пункте мы на примерах покажем принципы работы РК, а также усовершенствование его идей для кодирования изображений с низкими скоростями.

Основная идея РК заключается в следующем. Предположим, мы хотим выполнить квантование стационарного равномерного источника без памяти на скорости  $R$  бит/отсчет. Оптимальный скалярный квантователь будет иметь  $2^R$  уровней квантования (символов). В случае РК, вначале источник квантуется более точно,  $2^{R+k}$  символами. Естественно, это приводит к превышению данного ресурса бит, поэтому мы не имеем возможности произвольного выбора символов в каждый момент времени.

Пусть  $k = 1$ . Скалярная кодовая книга из  $2^{R+1}$  символов делится на четыре субкниги из  $2^{R-1}$  символов каждая. Пусть  $R = 2$  (рис.10.5). Субкниги строятся так, чтобы каждая из них представляла уровни реконструкции более грубого квантователя с  $(R - 1)$  бит/отсчет. Четыре субкниги обозначены  $D_0, D_1, D_2$  и  $D_3$ . Также обозначим  $S_0 = D_0 \cup D_2$  и  $S_1 = D_1 \cup D_3$ , где  $S_0$  и  $S_1$  называются суперкнигами.

Как было сказано, ограничение на скорость не позволяет выбирать произвольный символ из  $2^{R+1}$  символов. Однако существует возможность выбирать символы одной из суперкниг. Если известно, какая из суперкниг используется, то для каждого отсчета требуется один бит для обозначения субкниги и  $R - 1$  бит для обозначения кодового слова в ней. Выбор субкниги определяется состоянием автомата с конечным числом состояний, описываемого соответствующей решеткой.

Пример решетки с восемью состояниями приведен на рис.10.6. Субкниги  $\{D_0, D_1, D_2, D_3\}$  используются для пометки ребер решетки, так что бит, обозначающий субкнигу, определяет следующее состояние решетки. Кодирование осуществляется путем посылки одного бита на отсчет для обозначения пути по решетке и  $R - 1$  бит для обозначения кодового слова. Может показаться, что мы вернулись к неоптимальному квантователю со скоростью  $R$ . К чему же все старания? Ответ заключается в том, что мы получили большее количество кодовых слов, так как существует некоторая свобода выбора символов из  $S_0$  или  $S_1$ . Конечно, эта свобода неполная: решение по каждому символу принимается с учетом прошлого и будущего решений, то есть допустимых путей по решетке. Однако и эта гибкость приводит к эффективному кодированию. Доступность в каждый момент времени  $S_0$  и  $S_1$  означает, что уровни квантователя являются «скользящими» и настраиваются на данные, с учетом возможных путей по решетке.

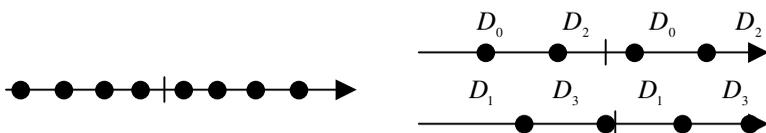


Рис. 10.5. Книги и суперкниги решетчатого квантования

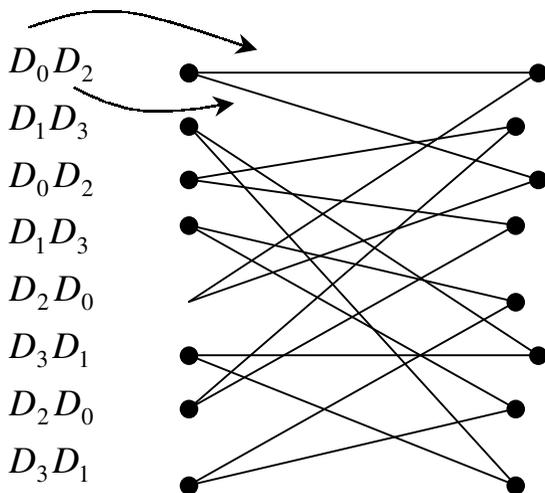


Рис. 10.6. Решетка РК с 8 состояниями

Перед тем как продолжить рассмотрение применения РК к кодированию вейвлет-коэффициентов отметим, что по своей эффективности и вычислительной сложности РК гораздо ближе к векторному, чем к скалярному квантованию. Почему бы не использовать стандартное векторное квантование? Ответ заключается в рекурсивной структуре РК и существовании простого алгоритма динамического программирования, известного как алгоритм Витерби для нахождения кодовых слов РК. Главное отличие заключается в том, что в случае векторного квантования кодовое слово размером  $N-1$  не связано с кодовым словом размером  $N$ . Алгоритм РК автоматически решает проблему увеличения размерности, увеличивая длину решетки.

Обычное РК мало пригодно для кодирования изображений, особенно на низких скоростях. В самом деле, один бит на отсчет используется только для кодирования самой решетки, тогда как интересующий диапазон скоростей находится ниже одного бита. Для кодирования изображения разработан ал-

горитм РК с ограниченной энтропией (ЕСТСQ). ЕСТСQ имеет две особенности: малое число бит для представления решетки (вводится «энтропия состояния») и используется факт, что вероятность появления нулей на выходе кодера очень велика.

### 10.5.2. Субполосные кодеры с РК

В работах Международной конференции по обработке изображений в 1996 году были представлены три похожих, но независимо разработанных алгоритма. Основными их составляющими являются: субполосная декомпозиция, классификация и оптимальное распределение бит по различным подмножествам данных и РК с ограниченной энтропией (ЕСТСQ). Известен кодер, объединяющий эти алгоритмы. Кратко обсудим их главные аспекты.

Рассмотрим субполосную декомпозицию изображения и предположим, что субполосы можно представить как нестационарный случайный процесс  $X$ , чьи отсчеты взяты из распределений с дисперсией  $\sigma_i^2$ . Конечно, в этом случае можно вычислить «среднюю дисперсию» всего случайного процесса и выполнить обычное оптимальное квантование. Но лучших результатов можно достичь путем передачи дополнительной информации о дисперсии каждого отсчета и квантования его в соответствии с его плотностью распределения вероятностей.

Эта идея была впервые предложена В.Ченом и К.Смитом для адаптивного квантования коэффициентов ДКП. Они предложили разделять коэффициенты на четыре группы в соответствии с их «уровнем активности», то есть дисперсией, и кодировать коэффициенты оптимальным квантователем для каждой группы. Вопрос о том, каким образом формировать группы, не рассматривался, и количество коэффициентов в группах было одинаковым.

Тем не менее, можно показать, что равное число коэффициентов в группах не всегда является правильным выбором. Предположим, что имеется  $J$  групп, и отсчеты, соответствующие одному классу  $i \in \{1, \dots, J\}$ , считаются порожденными одним источником  $X_i$ . Пусть источнику  $X_i$  соответствует  $N_i$  отсчетов, а общее число отсчетов во всех группах -  $N$ . Определим вероятность того, что отсчет принадлежит источнику  $X_i$ , как  $p_i = N_i / N$ . Кодирование источника  $X_i$  на скорости  $R_i$  приводит к среднеквадратической ошибке вида

$$D_i(R_i) = \varepsilon_i^2 \sigma_i^2 2^{-2R_i}, \quad (10.3)$$

где  $\varepsilon_i$  - константа, зависящая от вида плотности распределения вероятности. Проблема распределения бит далее решается методом множителей Лагранжа. В результате получаем выражение

$$R_i = \frac{R}{J} + \frac{1}{2} \log_2 \frac{\varepsilon_i^2 \sigma_i^2}{\prod_{j=1}^J (\varepsilon_j^2 \sigma_j^2)^{p_j}}, \quad (10.4)$$

где  $R$  - суммарная скорость, а  $R_i$  - скорость, соответствующая каждой группе. Выигрыш от классификации определяется как отношение ошибки квантования исходного сигнала  $X$  и классифицированного:

$$G_c = \frac{\varepsilon^2 \sigma^2}{\prod_{j=1}^J (\varepsilon_j^2 \sigma_j^2)^{p_j}}. \quad (10.5)$$

Этот выигрыш максимизируется по всем  $\{p_i\}$ . Неудивительно, что процесс оптимизации зачастую приводит к неравномерным  $\{p_i\}$ . Это означает неравномерное распределение коэффициентов по группам. Кроме того, неравномерное распределение требует передачи меньшего количества дополнительной информации: любое распределение  $\{p_i\}$  имеет меньшую энтропию, чем равномерное.

Мы определили выигрыш от классификации для одной субполосы. Обобщение этого результата привело к объединению этого выигрыша с выигрышем от субполосного кодирования. Также там уточнено выражение для количества требующейся для классификации дополнительной информации. Алгоритм кодирования оптимизирует получившееся выражение, и, затем, применяется решетчатое квантование с ограниченной энтропией (ЕСТСҚ).

Практическая реализация данного алгоритма требует учета намного большего количества деталей. Например, классификация отображает определенные уровни энергии сигнала, которые связаны с расположением контуров на изображении и соответствуют различным субполосам. Для уменьшения дополнительной информации могут быть использованы различные методы (например, алгоритм, обсуждаемый в следующем пункте). Другие исследования состоят в поиске альтернативных критериев для классификации, а также применении РК в сочетании с арифметическим кодированием. Эффективность кодирования ЕСТСҚ - одна из самых высоких из ныне известных,

хотя и вычислительная сложность также значительно превышает сложность рассматриваемых в главе кодеров.

### 10.5.3. Моделирование и оценивание смеси распределений

Один из подходов к кодированию вейвлет-коэффициентов изображения заключается в том, что они представляются как случайные переменные, порождаемые смесью распределений. Для каждого отсчета необходимо определить, источнику с каким распределением он соответствует, и квантовать его в соответствии с этим распределением. Так как декодер должен знать, какому распределению принадлежит данный коэффициент, многие кодеры посылают ему эту информацию как дополнительную. Уменьшение количества этой информации становится особенно важным при низких скоростях кодирования, так что управление им является чрезвычайно важным для эффективности кодера.

Все алгоритмы, обсуждаемые здесь, так или иначе используют этот подход. Они отличаются в основном ограничениями, накладываемыми на дополнительную информацию для ее эффективного кодирования. Например, нульдеревья отображают не что иное, как дополнительную информацию. В этом случае предполагается, что данные порождаются смесью источников с очень низкой (нулевой) и высокой энергией и что нули имеют древовидную структуру.

Субполосные кодеры на основе РК, обсуждаемые в предыдущем пункте, также используют этот подход. Информация о разбиении множества коэффициентов на группы с разной энергией передается как дополнительная информация. Очевидно, что существуют методы ее сокращения, также основанные на геометрических ограничениях в расположении коэффициентов той или иной структуры.

Полностью отличный подход к управлению количеством дополнительной информации известен, как «квантование через оценивание». Данный класс алгоритмов называется «кодированием с оцениванием смеси по прошлому» (КОСП).

КОСП использует в качестве модели вейвлет-коэффициентов нестационарный обобщенный гауссовский процесс, чья нестационарность выражается в медленно изменяющейся дисперсии коэффициентов в каждой полосе. Так как их энергия меняется медленно, она может быть предсказана на основе соседей. Следовательно, в случае КОСП, в отличие от предыдущих методов, не передается дополнительная информация, но декодер пытается восстано-

вить ее из принятых данных, отсюда и название – «по прошлому». КОСП основано на предположении, что ближайшие соседи коэффициента, в том числе и его родители по дереву, имеют ту же энергию, что и он сам. Оценка энергии производится по принципу максимального правдоподобия на базе соседних коэффициентов.

Так же, как и в других рекурсивных алгоритмах, где есть квантование, в КОСП должны быть учтены проблемы стабильности и смещения. Так как декодер имеет доступ только к квантованным коэффициентам, кодер тоже обязан использовать только квантованные коэффициенты. Иначе, кодер и декодер будут располагать различным контекстом, и возникнет проблема смещения. Таким образом, появляется добавочная трудность – оценивание дисперсии на основе уже квантованных коэффициентов. КОСП объединяет квантование и максимально правдоподобную оценку дисперсии.

Квантование выполняется равномерным квантователем с «мертвой» зоной (рис. 10.2). Такой квантователь хорошо подходит для квантования с ограниченной энтропией обобщенного гауссовского процесса. Ширину «мертвой» зоны и шага квантователя определяет оптимизационная процедура, основанная на методе множителей Лагранжа. Эта процедура выполняется заранее для различных скоростей кодирования и параметров источника. Результаты хранятся в таблице. Поэтому кодер работает очень быстро.

Еще одна проблема, связанная с алгоритмом оценивания по прошлому, заключается в следующем. Возможен случай, когда все соседи коэффициента будут равны нулю. В этом случае он тоже будет квантован в нуль. При повторении этой процедуры окажется, что все коэффициенты будут квантованы в нуль. Для предотвращения этой опасности в КОСП предусмотрен механизм защиты от распространения ошибки. В случае если все соседи равны нулю, декодеру посылается специальный символ. Это происходит на предварительном этапе кодирования, когда алгоритм проверяет возможность возникновения ошибочных серий нулей для всех коэффициентов. Такие коэффициенты группируются вместе, вычисляется их дисперсия и параметр огибающей обобщенного гауссовского распределения, и передаются декодеру. Далее, каждый раз при возникновении угрозы появления ошибочных нулей кодер и декодер работают с этими параметрами, а не с ближайшими соседями коэффициентов.

КОСП является очень быстрым и эффективным алгоритмом. Поэтому, очевидно, необходимо пересмотреть роль дополнительной информации и механизм ее передачи для вейвлет-кодеров.

## 10.6. Современные направления исследований

Исследования в области сжатия изображений ведутся по разным направлениям. Так, появилась новая интерпретация (описанная в главе 6) вейвлет-преобразования – лифтинговая схема, не основанная на преобразовании Фурье. С использованием этой схемы появилась возможность конструирования новых неразделимых базисов вейвлетов, которые потенциально могут привести к повышению эффективности кодеров. Интересным направлением исследований является изучение нелинейных аналогов вейвлет-преобразования, которые философия лифтинга делает возможным.

Активные исследования проводятся в области кодеров, основанных на классификации и оценивании по прошлому.

Одним из наиболее интересных направлений является разработка кодеров изображения, робастных к ошибкам, возникающим в каналах связи. При этом используется идея совместной оптимизации кодеров источника и канала, а также оптимального сочетания отдельно оптимизированных кодеров.

Особый интерес представляет адаптация вейвлет-кодирования изображения для кодирования видео. Здесь можно сочетать внутрикадровое кодирование с межкадровым предсказанием, как это заложено в стандарте MPEG-4. Можно также рассматривать видеопоследовательность как трехмерный массив и применять трехмерный вейвлет-анализ. Однако этот метод наталкивается на трудности, связанные с фундаментальными особенностями вейвлет-преобразования, как и любого субполосного кодирования. Вейвлет-преобразование не является пространственно-инвариантным в силу присутствия децимации и интерполяции. Эта изменчивость в пространстве мешает компактному представлению видеосигналов. Причина поясняется ниже.

Видеосигналы состоят из кадров. От кадра к кадру информация меняется незначительно. Поэтому существует возможность достичь хороших результатов сжатия, передав одинаковую информацию лишь однажды. Однако вейвлет-преобразование не является инвариантным к сдвигу, следовательно, подобное кодирование невозможно. Аналогичные доводы против трехмерного вейвлет-преобразования приводят и в частотной области.

Итак, в настоящей главе рассмотрено применение вейвлет-преобразования для сжатия изображений. Во всем мире в данном направлении ведутся интенсивные работы. Разработано большое число алгоритмов и кодеров, некоторые из которых стандартизированы.

Современные вейвлет-кодеры основаны на предположении, что изображение порождается источником с флюктуирующей дисперсией. Каждый ко-

дер реализует определенный механизм для отображения локальной дисперсии вейвлет-коэффициентов и квантует их оптимальным или субоптимальным образом в соответствии с дисперсией. Кодеры отличаются друг от друга стратегиями квантования коэффициентов и тем, каким образом происходит оценка и передача значения дисперсии декодеру.

Кодеры, основанные на алгоритме нульдеревя, предполагают у дисперсии наличие двух состояний: нуль или нет. Декодеру передается дополнительная информация о местоположении значимых коэффициентов. Этот процесс приводит к нелинейной аппроксимации изображения. Множества нулевых коэффициентов выражаются в терминах деревьев вейвлетов (Льюис и Ноулес, Шапиро и др.) или комбинаций этих деревьев (Саид и Перельман). Нули передаются декодеру как дополнительная информация, так же как и квантованные данные. Кодеры, основанные на нульдеревьях, учитывают межполосные зависимости вейвлет-коэффициентов.

В частотно-адаптивных кодерах применяются ортогональные адаптивные преобразования – метод вейвлет-пакетов. Локальные флуктуации корреляционных связей используют пространственно-адаптивные кодеры.

Другие вейвлет-кодеры учитывают внутриволновые зависимости между вейвлет-коэффициентами (иногда одновременно и межполосные). Кодеры, основанные на решетчатом квантовании, делят коэффициенты на группы в соответствии с их энергией. Для каждого коэффициента они оценивают и (или) передают информацию о группе и значение квантованного в соответствии с номинальной дисперсией группы коэффициента. Другой новый класс кодеров передает незначительное количество информации о дисперсии. Это показывает, что, возможно, информация о дисперсии имеет большую избыточность, чем считалось раньше.

Многие исследования посвящены разработке алгоритмов кодирования изображения на основе векторного квантования вейвлет-коэффициентов. Эти алгоритмы не рассмотрены в нашей книге. Представление о них вы можете получить, прочтя обзорную статью Р.Грэя (см. Интернет-ссылки).

В заключение можно отметить, что вейвлеты и сопутствующие им идеи внесли неоценимый вклад в теорию и практику кодирования изображений и, по-видимому, будут оставаться основным направлением исследований в этой области в ближайшем будущем.