Вестник Ивановского государственного университета

Серия «Естественные, общественные науки»

Вып. 2 / 2009 С. 121 – 128

С. И. Хашин

СРАВНИТЕЛЬНАЯ ЭФФЕКТИВНОСТЬ ДИСКРЕТНЫХ ВЕЙВЛЕТОВ МАЛОГО ПОРЯДКА

Предлагается метод сравнения эффективности разных вариантов дискретного вейвлет-преобразования. Сравниваются вейвлеты Хаара, Добеши D4 и D6. С ними же сравнивается эффективность дискретного преобразования Фурье длины от 4 до 16.

The method of comparison of efficiency of different variants of discrete wavelet-transformations is offered. Haar, Dobeshi D4 and D6 wavelets are compared. Efficiency of Discrete Fourier Transformation of length from 4 to 16 is compared also.

Ключевые слова: вейвлет, компьютерная графика, эффективность.

Key words: wavelet, computer graphics, efficiency.

УДК 519.68.

Введение

Каждая компонента RGB-цвета (или YUV-цвета) для изображения размером $mx \times my$ задается матрицей размера $mx * my$. Будем рассматривать эту матрицу как вектор из $\mathbb{R}$-пространства L размерности $mx * my$ с обычной евклидовой метрикой.

Одна из основных идей дискретного вейвлет-анализа [1, 3, 4, 9] заключается в разложении пространства L в прямую сумму двух подпространств $L = L' \oplus L''$ — “низкочастотной” и “высокочастотной” компонент.

Помимо различных “биологических” объяснений пользы от такого разложения [2, 5], существует и более формальное. Если взять “случайное” разложение, то можно ожидать, что длина проекции вектора на каждое из подпространств составит примерно $1/\sqrt{2}$ от длины исходного вектора, т. е. примерно 71 %. Однако для реальных изображений длина высокочастотной компоненты не просто меньше, а меньше на порядок. Это и позволяет подтвердить вывод, что большая часть информации содержится в низкочастотной компоненте.

Рассмотрим для примера разложение пространства матриц размера $mx * my$ при четном mx: $L = L'_0 \oplus L''_0$. В качестве L'_0 (низкочастотная компонента) возьмем матрицы $A = \{a_{ij}\}$, у которых $a_{2k+1,j} = a_{2k,j}$, а в качестве L''_0 (высокочастотная компонента) — матрицы, у которых $a_{2k+1,j} = -a_{2k,j}$. Это простейшее разложение назовем *разложением Хаара.*

Для стандартного тестового файла “lena.bmp” размером $512 * 512$ точек длина проекции на подпространство L''_0 составляет лишь 4.281 % от длины исходного вектора.

Для выбранного набора из 45 высококачественных изображений длины проекций на подпространство L''_0 в процентах от длины исходного вектора оказались следующие:

7.526, 9.910, 16.512, 11.588, 6.736, 3.423, 6.304, 7.580, 5.192, 3.026, 4.312, 9.877, 4.270, 6.251, 3.007, 10.310, 5.279, 6.074, 5.823, 5.647, 3.241, 8.851, 3.379, 8.841, 6.707, 6.497, 8.816, 10.995, 7.502, 3.535, 7.680, 16.024, 2.651, 7.187, 2.773, 8.874, 15.144, 5.865, 2.359, 6.456, 4.855, 4.354, 2.608, 5.300, 11.815.

Полученные числа лежат в довольно широком интервале: от 2.65 до 16.5, среднее значение равно 6.910, а среднеквадратичное отклонение — 3.479. Таким образом, среднее отклонение ненамного меньше среднего значения рассматриваемой величины. Поэтому использовать эту величину для оценки качества разложения в прямую сумму не очень корректно, даже после усреднения по большому количеству изображений.

В настоящей работе предлагается другая, более стабильная численная характеристика разложения пространства в прямую сумму.

Дискретные ортогональные вейвлеты

Пусть L — пространство действительнозначных функций на **Z** с конечным носителем. На пространстве L будем рассматривать обычное скалярное произведение и обычную евклидову норму. Через R_k обозначим оператор сдвига на k:

$$R_k(f)(x) = f(x-k).$$

Определение. Пусть $f(x)$ — функция из L. Обозначим через $\rho_2(f)$ подпространство в L, порожденное всеми сдвигами функции f на четное число:

$$\rho_2(f) = \{R_{2k}f, k \in \mathbb{Z}\}.$$

Определение. Пусть $a(x)$, $b(x)$ — пара функций из L, удовлетворяющая условиям:

1. $(a,a) = (b,b) = 1$;
2. $(a, R_{2k}(a)) = 0, \ \forall k \in \mathbb{Z}, k \neq 0$;
3. $(b, R_{2k}(b)) = 0, \ \forall k \in \mathbb{Z}, k \neq 0$;
4. $(a, R_{2k}(b)) = 0, \ \forall k \in \mathbb{Z}$.

Эти условия означают, что система функций $W = (R_{2k}a, R_{2k}b)$, полученных из пары (a,b) сдвигом на все четные числа, образует ортонормальный базис в пространстве L. В этом случае будем называть базис W *ортогональным вейвлет-базисом*, а саму пару функций (a,b) — ортогональным вейвлетом.

Таким образом, каждый ортогональный вейвлет задает разложение пространства L в прямую сумму двух подпространств: $L = L' \oplus L''$, где L' — подпространство, порожденное векторами $(R_{2k}a)$, а L'' — векторами $(R_{2k}b)$. При этом векторы $(R_{2k}a)$ образуют ортонормальный базис в пространстве L', а векторы $(R_{2k}b)$ — в пространстве L''.

Пример. Функции Хаара. Пусть функции $h_0(x), h_1(x)$ равны 0 при отрицательных x, а для неотрицательных:

$$\begin{aligned} h_0(x) &= \alpha, \quad \alpha,\ 0,\ \ldots, \\ h_1(x) &= \alpha,\ -\alpha,\ 0,\ \ldots, \end{aligned}$$

где $\alpha = 1/\sqrt{2}$. Пространство L' состоит из функций, значения которых в двух соседних точках равны, а L'' — из функций, значения которых в двух соседних точках равны по абсолютной величине и противоположны по знаку.

Других вейвлетов, носители которых сосредоточены в двух соседних точках, не существует.

Ортогональные вейвлеты с носителем длины 4

Предположим, что функции (a, b), составляющие ортогональный вейвлет, имеют не более 4 ненулевых компонент: (a_0, a_1, a_2, a_3) и (b_0, b_1, b_2, b_3) соответственно. Тогда наложенные на них условия будут выглядеть так:

1. $a_0^2 + a_1^2 + a_2^2 + a_3^2 = 1,$
2. $b_0^2 + b_1^2 + b_2^2 + b_3^2 = 1,$
3. $a_0 a_2 + a_1 a_3 = 0,$
4. $b_0 b_2 + b_1 b_3 = 0,$
5. $a_0 b_0 + a_1 b_1 + a_2 b_2 + a_3 b_3 = 0,$
6. $a_0 b_2 + a_1 b_3 = 0,$
7. $a_2 b_0 + a_3 b_1 = 0.$

Из уравнений (1) и (3) получим, что вектор (a_0, a_1, a_2, a_3) должен иметь вид

$$(\sin\alpha\cos\varphi,\ \sin\alpha\sin\varphi,\ -\cos\alpha\sin\varphi,\ \cos\alpha\cos\varphi)$$

для некоторых α, φ, а из уравнений (2), (4) получаем вид для вектора (b_0, b_1, b_2, b_3):

$$(\sin\alpha_1\cos\varphi_1,\ \sin\alpha_1\sin\varphi_1,\ -\cos\alpha_1\sin\varphi_1,\ \cos\alpha_1\cos\varphi_1)$$

для некоторых α_1, φ_1. Уравнения (6) и (7) будут выполняться лишь в случае $\varphi = \varphi_1$, а уравнение (5) будет выполнено лишь в случае $\alpha_1 = \alpha + \pi/2$. Общее решение системы будет выглядеть так:

$$\begin{aligned} (a_i) &= (\sin\alpha\cos\varphi,\ \sin\alpha\sin\varphi,\ -\cos\alpha\sin\varphi,\ \cos\alpha\cos\varphi)\ , \\ (b_i) &= (\cos\alpha\cos\varphi,\ \cos\alpha\sin\varphi,\ \sin\alpha\sin\varphi,\ -\sin\alpha\cos\varphi) \end{aligned}$$

для некоторых α, φ.

Для использования в компьютерной графике (как и для многих других приложений) желательно выбрать вейвлет-базис так, чтобы скалярные произведения исходных данных f_i с векторами из второй половины базиса $R_{2k}b$ имели наименьшее среднеквадратичное значение. Для этого естественно выбрать вектор (b_0, b_1, b_2, b_3) так, чтобы он был ортогонален полиномам степени 0 и 1, т.е. выполнялись условия

$$b_0 + b_1 + b_2 + b_3 = 0$$
$$0 \cdot b_0 + 1 \cdot b_1 + 2 \cdot b_2 + 3 \cdot b_3 = 0.$$

Решением этих уравнений будет вейвлет Добеши $D4$. Он получается, если в предыдущих формулах положить $\alpha = \pi/12$, $\varphi = \pi/3$:

$$(b_0, b_1, b_2, b_3) = \left(\frac{\sqrt{3}+1}{4\sqrt{2}}, \frac{-\sqrt{3}-3}{4\sqrt{2}}, \frac{3-\sqrt{3}}{4\sqrt{2}}, \frac{\sqrt{3}-1}{4\sqrt{2}}\right)$$

или, приближенно:

$$(b_0, b_1, b_2, b_3) \approx (0.4829629131, -0.8365163037,\ 0.2241438682,\ 0.1294095227).$$

Вейвлеты Добеши порядков 6 и 8

Вычисления, проведенные в предыдущем разделе, можно распространить и на вейвлеты с носителями большей длины. Например, для носителя длины 6 коэффициенты вектора $b = (b_0, b_1, b_2, b_3, b_4, b_5)$ должны удовлетворять уравнениям:

$$\begin{aligned}
&b_0^2 + b_1^2 + b_2^2 + b_3^2 + b_4^2 + b_5^2 = 1,\\
&b_0b_2 + b_1b_3 + b_2b_4 + b_3b_5 = 0,\\
&b_0b_4 + b_1b_5 = 0,\\
&b_0 + b_1 + b_2 + b_3 + b_4 + b_5 = 0,\\
&b_1 + 2b_2 + 3b_3 + 4b_4 + 5b_5 = 0,\\
&b_1 + 4b_2 + 9b_3 + 16b_4 + 25b_5 = 0.
\end{aligned}$$

Решение этой системы единственно с точностью до знаков и перестановки чисел в обратном порядке:

$$\begin{aligned}
b_0 &= -0.3326705529500825764,\\
b_1 &= 0.8068915093110925474,\\
b_2 &= -0.4598775021184914324,\\
b_3 &= -0.1350110200102546676,\\
b_4 &= 0.08544127388202667206,\\
b_5 &= 0.0352262918857095335.
\end{aligned}$$

Для аналогичного вейвлета длины 8 получаем:

$$\begin{aligned}
b_0 &= 0.2303778133088965008 6,\\
b_1 &= -0.7148465705529156470 9,\\
b_2 &= 0.6308807679298589078 8,\\
b_3 &= 0.0279837694168598542 11,\\
b_4 &= -0.1870348117190930840 8,\\
b_5 &= -0.0308413818355607636 27,\\
b_6 &= 0.0328830116668851997 35,\\
b_7 &= 0.0105974017850690321 05.
\end{aligned}$$

Разложение Фурье

Помимо вейвлет-разложения существуют и другие способы разбиения пространства в сумму низкочастотной и высокочастотной компонент. Наиболее известное — дискретное преобразование Фурье.

Пусть N — четное натуральное число, $F_0, \ldots, F_{2N-1}$ — ортонормальный базис Фурье длины N. В N-мерном пространстве обозначим через F' подпространство, порожденное первыми $N/2$ базисными векторами, через F'' — остальными $N/2$ векторами. Получаем разложение N-мерного пространства в прямую сумму двух подпространств половинной размерности: $F = F' \oplus F''$.

Пусть a — матрица изображения размером $mx * my$, ширина которого кратна N. Разобьем каждую строчку матрицы на блоки длиной N и каждый блок разложим в прямую сумму компонент из F' и F''. Таким образом, получаем разложение пространства размерности $mx * my$ в ортогональную сумму двух подпространств $L = F'_N \oplus F''_N$.

Замечание. При $N = 2$ разложение Фурье совпадает с разложением Хаара.

Эффективность разложения пространства в прямую сумму

Пусть пространство L разложено в прямую сумму двух подпространств: $L = L' \oplus L''$. Если вектор v разложен в сумму $v = v' + v''$, $v' \in L'$, $v'' \in L''$, то

$$||v|| = \sqrt{||v'||^2 + ||v''||^2}.$$

Конечно, длины проекций вектора подпространства L' и L'' могут иметь любые отношения. Но в компьютерной графике мы рассматриваем не произвольные векторы, а достаточно конкретные: их координатами являются яркости точек в некотором изображении. В этом случае отношение длины проекции к длине исходного вектора может оказаться довольно стабильным. Однако, как было показано выше, эта стабильность совершенно недостаточна для того, чтобы оценивать качество разложения в прямую сумму: этот показатель слишком сильно меняется от одного изображения к другому.

Оказывается, можно найти гораздо более стабильный показатель. Для этого наряду с данным разложением рассмотрим еще и самое простейшее разложение такого вида, разложение Хаара $L = L'_0 \oplus L''_0$. Вектор v, соответствующий некоторому изображению, разложим в прямую сумму двумя способами:

$$v = v'_0 + v''_0 = v' + v'',$$

и в качестве численной характеристики качества разложения $L = L' \oplus L''$ рассмотрим отношение

$$\frac{||v''||}{||v''_0||}.$$

Разумеется, для произвольного вектора это отношение может быть каким угодно, но для векторов, соответствующих реальным изображениям, ситуация сравнительно стабильная. Например, для разложения, соответствующего вейвлету Добеши D4, для того же набора из 45 изображений получаются следующие значения эффективности (в процентах):

81.42, 89.56, 97.77, 95.32, 95.71, 75.78, 81.90, 87.30, 86.35, 70.60, 86.31, 77.25, 86.48, 87.51, 74.62, 86.12, 85.30, 90.54, 90.02, 86.24, 95.22, 92.00, 94.70, 94.66, 96.35, 95.15, 97.05, 95.06, 92.22, 93.99, 90.01, 93.18, 96.48, 95.49, 93.52, 96.56, 93.35, 93.16, 89.46, 96.06, 93.37, 95.53, 94.10, 95.15, 95.69.

Среднее значение равно 90.44 % и среднеквадратичное отклонение — 6.47 %.

В работах [6, 7, 8] рассматриваются другие подходы к оценке качества вейвлетов, однако предлагаемый метод является значительно более точным и надежным.

В работе [10] рассматривается метод оценки качества ортонормальных базисов, основанных на аналогичных идеях.

Сравнение эффективности

Будем сравнивать эффективность всех методов на стандартном изображении "lena.bmp" размера 512*512 точек, преобразованном к серому цвету, и на уже упоминавшемся выше наборе из 45 высококачественных изображений (табл.).

Подпространство	Эффективность на "lena.bmp", %	Эффективность на 45 изображениях, %	Среднеквадратичное отклонение на 45 изображениях, %
Хаар	100.00	100.00	0.00
Dobeshi-4	79.37	90.44	6.47
Dobeshi-6	74.29	87.60	8.20
Dobeshi-8	73.01	86.25	9.01
Fourier-4	85.03	93.36	5.18
Fourier-6	79.54	89.92	7.30
Fourier-8	76.89	88.05	8.43
Fourier-16	72.59	85.42	9.93

Таким образом, можно сделать вывод, что разложение по вейвлетам Добеши примерно на 5 % более эффективно, чем разложение Фурье той же длины.

Оптимальный ортогональный вейвлет длины 4

Каждый ортогональный вейвлет длины 4 определяется 8 числами (a_0, a_1, a_2, a_3) и (b_0, b_1, b_2, b_3), удовлетворяющими условиям, приведенным в начале работы. Если взять только высокочастотную компоненту (b_i), то она должна удовлетворять условиям:

$$\begin{aligned} &1.\ b_0^2 + b_1^2 + b_2^2 + b_3^2 = 1, \\ &2.\ b_0 b_2 + b_1 b_3 = 0. \end{aligned}$$

Четверок чисел, удовлетворяющих этим двум условиям, много. Для выделения наиболее “высокочастотной” компоненты можно применить подход Добеши, заключающийся в том, чтобы наложить на функцию b условие ортогональности многочленам малой степени, в нашем случае это ортогональность многочленам нулевой и первой степени:

$$\begin{aligned} &3.\ b_0 + b_1 + b_2 + b_3 = 0, \\ &4.\ b_1 + 2b_2 + 3b_3 = 0. \end{aligned}$$

При этих условиях решение будет единственно с точностью до знака и перестановки в обратном порядке: это вейвлет Добеши (см. выше).

Необходимость условия (3) для минимизации веса высокочастотной компоненты сомнений не вызывает, это подтверждают и эксперименты. Однако про условие (4) этого сказать уже нельзя. Посмотрим, что получится в результате отказа от этого условия.

Общее решение уравнений (1—3) можно представить в следующем виде. Обозначим через r длину вектора (b_0, b_1): $r = \sqrt{b_0^2 + b_1^2}$, при этом $0 \leqslant r \leqslant 1$ и в содержательных случаях $0 < r < 1$. Нетрудно видеть, что задание параметра r однозначно определяет вектор (b_0, b_1, b_2, b_3), удовлетворяющий условиям (1—3):

1. Находим $\lambda = \sqrt{1/r - 1}$;
2. Находим $\varphi = \operatorname{arctg} \frac{\lambda+1}{\lambda-1}$;
3. Находим b_i по формулам:

$$\begin{aligned} b_0 &= r\cos\varphi, \\ b_1 &= r\sin\varphi, \\ b_2 &= -\lambda b_1, \\ b_3 &= \lambda b_0. \end{aligned}$$

Вейвлет Добеши получается по этим формулам (с точностью до перестановки) при

$$r = \frac{\sqrt{3} - 1}{2\sqrt{2}} \approx 0.2588.$$

Чтобы понять, насколько вейвлет Добеши отличается от оптимального, проверим эффективность вейвлетов при различных r для прежнего отобранного множества из 45 изображений.

Из этих вычислений следует, что оптимальное значение r для различных изображений лежит в пределах от 0.25 до 0.27, причем их эффективность различается лишь в сотых, иногда в десятых процента.

В целом можно сделать вывод, что в пределах погрешности наших измерений вейвлет Добеши D4 является наиболее эффективным среди всех ортогональных вейвлетов с носителем длины 4.

Библиографический список

1. *Ватолин Д. С.* Алгоритмы сжатия изображений. М.: Изд-во МГУ, 1999. 80 с.
2. *Ватолин Д., Ратушняк А., Смирнов М.* Методы сжатия данных. М.: Диалог-МИФИ, 2003. 382 с.
3. *Воробьев В. И., Грибунин В. Г.* Теория и практика вейвлет-преобразования. СПб.: ВУС, 1999. 206 с.
4. *Добеши И.* Десять лекций по вейвлетам. М.: РХД, 2001. 461 с.
5. *Дремин И. М., Иванов О. В., Нечитайло В. А.* Вейвлеты и их использование // УФН. 2001. Т. 171. № 5. С. 465—501.
6. *Злобин А. С.* Эффективность базисных функций вейвлет преобразований // Докл. 48-й конф. МФТИ. Секция радиотехники и защиты информации. М.: МФТИ, 2005.
7. *Умняшкин С. В., Кочетков М. Е.* Анализ эффективности использования дискретных ортогональных преобразований для цифрового кодирования коррелированных данных // Изв. вузов. Электроника. 1998. № 6. С. 79—84.
8. *Умняшкин С. В.* Эффективность применения ортогональных преобразований для кодирования дискретных сигналов с точки зрения корреляционной теории // Интеллектуальные системы в производстве. 2003. № 1. С. 100—123.
9. *Фрейзер М.* Введение в вейвлеты в свете линейной алгебры. М.: Бином, 2008. 488 с.
10. Хашин С. И. Оптимальный ортонормальный базис в компьютерной графике // Вестн. Иван. гос. ун-та. Естественные, общественные науки. 2008. Вып. 2. С. 122—127.