ПРИМЕНЕНИЕ ОРТОГОНАЛЬНОГО ПРЕОБРАЗОВАНИЯ ВЕЙЛЯ-ГЕЙЗЕНБЕРГА ДЛЯ СЖАТИЯ ИЗОБРАЖЕНИЙ

Асирян Валерий Мишевич
студент группы БЗС1401 МТУСИ
dmc5mod@yandex.ru
Волчков Валерий Павлович
д.т.н., профессор МТУСИ
volchkovvalery@mail.ru

Ключевые слова: *сжатие изображений, ортогональный базис, базис Вейля-Гейзенберга, хорошая локализация, оптимальная функция, преобразование Фурье.*

В работе рассматриваются вопросы сжатия дискретных двумерных сигналов (изображений) с использованием ортогональных преобразований, в том числе ортогонального преобразования Вейля-Гейзенберга, обладающего свойством наилучшей локализации одновременно в частотной и временной областях. Приводится сравнение ортогонального преобразования Вейля-Гейзенберга с другими дискретными методами сжатия сигналов. В работе представлены результаты моделирования оптимального базиса Вейля-Гейзенберга, а также результаты применения данного метода сжатия непосредственно к изображениям.

Введение

В настоящее время важнейшую роль в области цифровой обработки сигналов играют дискретные ортогональные преобразования, которые активно используются для анализа и обработки сигналов. Между тем, аппарат дискретных ортогональных преобразований находит свое место и в области сжатия данных для последующего экономичного хранения или передачи информации.

Одним из известных ортогональных преобразований является дискретное преобразование Фурье, получившее широкое распространение в цифровой обработке сигналов. Однако, на практике большинство сигналов (особенно изображения) являются нестационарными, и возникает необходимость получать информацию о спектре, локализованную в отдельных его фрагментах. Синтез универсального базиса, который позволяет функционально разделять сигнал на определенные фрагменты, а затем внутри них анализировать спектральные особенности сигнала, представляет сложную задачу. Однако, именно с помощью таких базисов можно учесть нестационарные особенности сигнала и получить большую эффективность сжатия. Наиболее важным примером таких базисов являются базисы Вейля-Гейзенберга, получаемые сдвигами по времени и частоте одной функции или целого семейства функций.

1. Синтез оптимального базиса Вейля-Гейзенберга

Первым этапом формирования базиса Вейля-Гейзенберга является выбор формирующей функции. Хорошо известно, что идеально локализованной в частотно-временной области является функция Гаусса. Однако, базис Вейля-Гейзенберга, построенный на ее основе (базис Габора), не является ортогональным. Тем не менее, использование стандартной процедуры ортогонализации, например, процесса Грама-Шмидта, приведет к значительному ухудшению частотно-временной локализации, и, как следствие, такой ортогональный базис Вейля-Гейзенберга не будет оптимальным.

В работах [1, 2] предлагается алгоритм построения оптимального базиса Вейля-Гейзенберга с использованием спектрального разложения. Ниже представлен полный алгоритм синтеза базисов Габора и оптимального базиса Вейля-Гейзенберга.

1.1. Формирование базиса Габора

1) В качестве формирующей функции выбирается дискретизированный усеченный гауссиан вида:

$$g_0(t) = (2\sigma)^{1/4} \exp(-\pi\sigma t^2) \tag{1}$$

На данном этапе важно выбрать такой интервал дискретизации, при котором гауссиан будет обусловлен наилучшим образом. В противном случае, процедура ортогонализации базиса пройдет некорректно.

2) Формируется матрица базиса Габора $\mathrm{G}=[\mathrm{G}_R, \mathrm{G}_I]$, элементы которой определяются выражениями:

$$\begin{gathered}
\mathrm{G}_R(n, lM+k) = g_0[(n-lM)_{\mathrm{mod}\,N}]\exp(2\pi j\frac{k}{M}(n-a/2)) \\
\mathrm{G}_I(n, lM+k) = jg_0[(n+M/2-lM)_{\mathrm{mod}\,N}]\exp(2\pi j\frac{k}{M}(n-a/2)) \\
n = 0,\dots,\mathrm{N}-1, k = 0,\dots,M-1, l = 0,\dots,L-1, \mathrm{N} = LM
\end{gathered} \tag{2}$$

где M – количество сдвигов по времени, L – количество сдвигов по частоте, a – некий параметр. Причем матрица G является прямоугольной, размерности (N, 2N).

1.2. Формирование оптимального базиса Вейля-Гейзенберга

3) Формируется расширенная вещественная матрица Габора $\mathrm{G}_B = [\mathrm{Re}\,\mathrm{G}; \mathrm{Im}\,\mathrm{G}]$.

4) Ищется спектральное разложение квадратной матрицы $\mathrm{G}_B\mathrm{G}_B^* = \mathrm{S}\Lambda\mathrm{S}^*$, где «*» – эрмитово-сопряжение.

5) Вычисляются матрицы $\Sigma = \Lambda^{1/2}, \mathrm{W} = \mathrm{G}_B\mathrm{S}\Sigma^{-1}$

6) Вычисляется оптимальная вещественная ортогональная матрица $\mathrm{V}_{опт} = \mathrm{SW}^*$

7) Из блочного разбиения квадратной матрицы $\mathrm{V}_{опт} = [\mathrm{V}_{1опт}; \mathrm{V}_{2опт}]$ находятся матрицы $\mathrm{V}_{1опт}$ и $\mathrm{V}_{2опт}$ размерности (N, 2N).

8) Строится матрица оптимального базиса Вейля-Гейзенберга $\mathrm{U}_{опт} = \mathrm{V}_{1опт} + \mathrm{j}\mathrm{V}_{2опт}$ размерности (N, 2N).

1.3. Результаты моделирования базисов

Ниже представлены результаты моделирования базисов Габора и Вейля-Гейзенберга при следующих параметрах: M = 8, L = 8, a = 3, σ = 0.1. На рисунке 1 изображены базисные функции во временной области, соответствующие нулевым сдвигам по частоте и по времени. На рисунке 2 представлены те же базисные функции в частотной области, для удобства сдвинутые в середину графика. Из графиков видно, что процедура ортогонализации с использованием спектрального разложения привела к незначительному ухудшению временной и частотной локализации.

На рисунке 3 показано, что матрица оптимального базиса Вейля-Гейзенберга действительно состоит из базисной функции, сдвинутой по времени и частоте.

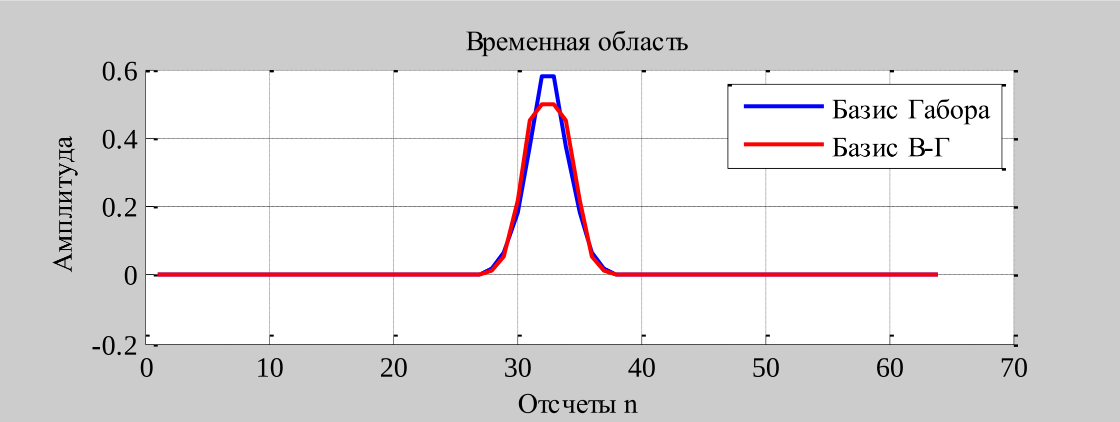

Рис 1. Базисы Габора и Вейля-Гейзенберга во временной области

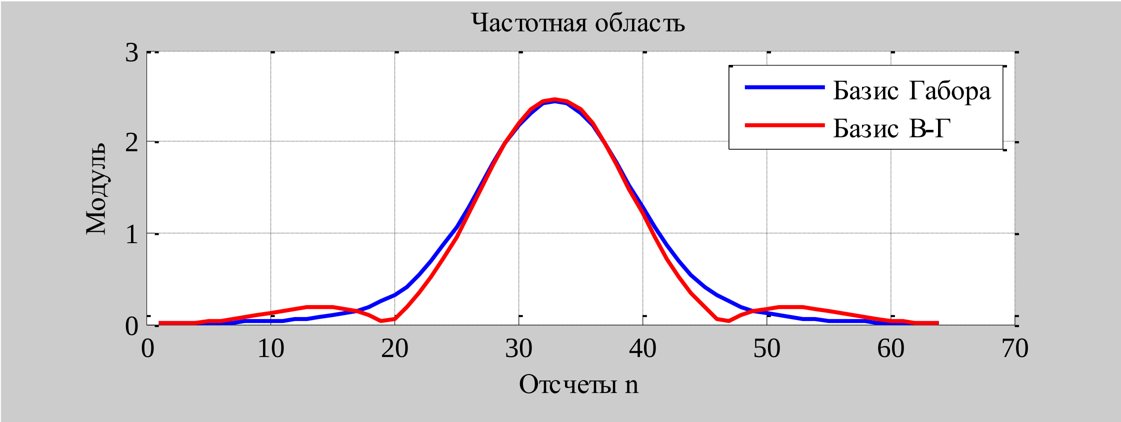

Рис 2. Базисы Габора и Вейля-Гейзенберга в частотной области

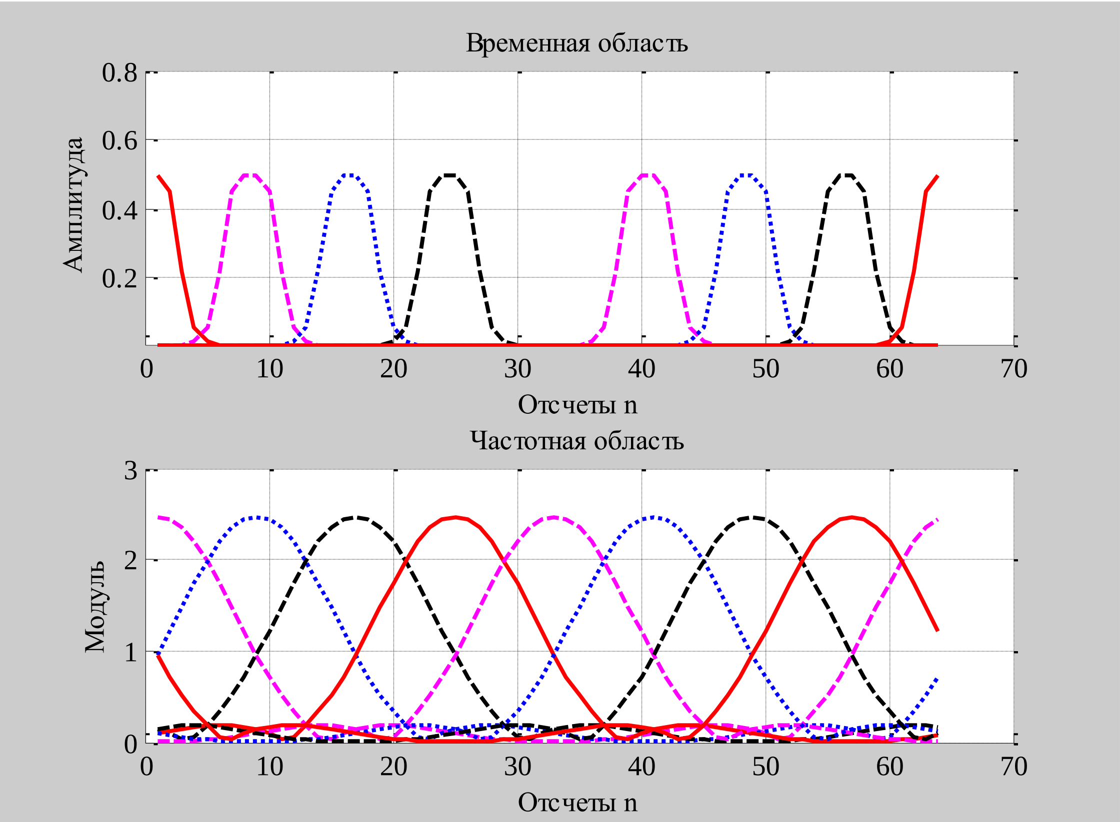

Рис 3. Базисные сдвиги во временной и частотной областях

2. Ортогональное преобразование Вейля-Гейзенберга

Основная идея дискретных ортогональных преобразований заключается в изменении сигнала с целью придания ему другой формы, в которой он имеет, возможно, непривычный вид, но обладает полезными свойствами.

Важнейшим свойством ортогональных преобразований является их обратимость. Это значит, что преобразованный сигнал, изменивший свою форму и вид, можно легко вернуть в первоначальное состояние.

Прямоугольная ортогональная матрица оптимального базиса Вейля-Гейзенберга $\mathrm{U}_{опт}$ является «особенной», так как она ортогональна только в смысле вещественного скалярного произведения. Условие ортогональности в данном случае можно представить в матричном виде:

$$\mathrm{Re}(\mathrm{U}^{*}_{опт}\mathrm{U}_{опт}) = \mathrm{I}_{2\mathrm{N}} \tag{3}$$

где $\mathrm{I}_{2\mathrm{N}}$ – прямоугольная единичная матрица. Поэтому прямое и обратное одномерные ортогональные преобразования Вейля-Гейзенберга будут осуществляться следующим образом:

$$b = \mathrm{Re}(\mathrm{U}^{*}_{опт}a) \tag{4}$$

$$a' = \mathrm{U}_{опт}b \tag{5}$$

Заметим, что для матрицы оптимального базиса Вейля-Гейзенберга не выполняется свойство унитарности, поскольку:

$$\mathrm{U}_{опт}\mathrm{U}^{*}_{опт} = 2\mathrm{I}_{2N} \tag{6}$$

И, поэтому, прямое и обратное двумерные дискретные ортогональные преобразования Вейля-Гейзенберга будут определяться по формулам:

$$\mathrm{B} = \mathrm{Re}(\mathrm{U}^{*}_{опт}\mathrm{A}\mathrm{U}_{опт}) \tag{7}$$

$$\mathrm{A}' = \mathrm{U}_{опт}\mathrm{B}\mathrm{U}^{*}_{опт} / 2 \tag{8}$$

Важно также отметить, что при прямом преобразовании происходит увеличение размера сигнального вектора или матрицы в 2 раза. Это еще одна особенность данного ортогонального базиса.

На рисунке 4 представлены исходное монохромное квадратное изображение *«lena256.jpg»* размеров 256×256 пикселей и его прямое и обратное ортогональные преобразования оптимальным базисом Вейля-Гейзенберга (*DWHT)* при значениях параметров: $M = 8$, $L = 32$, $a = 3$, $\sigma = 0.025$.

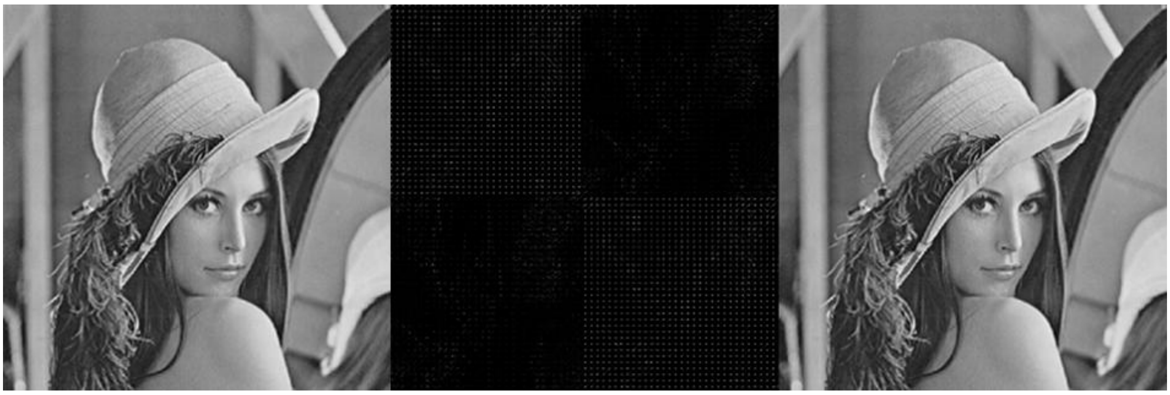

Рис 4. Исходное изображение и его прямое и обратное *DWH*-преобразования

Несложно заметить, что исходное изображение и восстановленное практически не отличаются. В этом можно также убедиться, вычислив норму Фробениуса разности матриц исходного и восстановленного изображений:

$$A_\Delta = A - A' \qquad (9)$$

$$E = \| A_\Delta \|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} h_{ij}^2} \qquad (10)$$

В дальнейшем формула (10) будет служить критерием качественных потерь, то есть главным показателем отличия восстановленного изображения от исходного.

В таблице 1 приведены результаты восстановления исходного изображения *«lena256.jpg»* с использованием преобразования Вейля-Гейзенберга (*DWHT*), преобразования Фурье (*DFT*) и косинусного преобразования (*DCT*). Очень малые ошибки восстановления в данном случае обусловлены только вычислительной погрешностью.

Таблица 1. Результаты восстановления изображения

Базис	*DWHT*	*DFT*	*DCT*
Потери качества: E	2.0479e-10	2.3763e-11	4.9210e-10

3. Сжатие данных по пороговому значению

Сжатие данных по пороговому значению представляет собой процедуру обнуления всех тех значений преобразованного изображения, модуль которых меньше определенного значения T – порога. Данный процесс представляет собой сжатие с потерями. Ниже представлен алгоритм сжатия данных вещественной матрицы B по пороговому значению T.

$$B_{ij} = \begin{cases} 0, \text{ если } |B_{ij}| < T, \\ B_{ij}, \text{ иначе.} \end{cases} \qquad (11)$$

Если матрица B – комплексная, то с порогом сравнивается модуль каждого комплексного элемента $B_{ij} = \mathrm{Re}(B_{ij}) + \mathrm{Im}(B_{ij})$. В этом случае $|B_{ij}| = \sqrt{\mathrm{Re}(B_{ij}) + \mathrm{Im}(B_{ij})}$, а формула (11) формально будет справедлива и для комплексного и для вещественного случая.

Следует заметить, что при оптимально выбранном способе сжатия и пороге T потери будут незначительными, что позволит устранить лишнюю информацию, сохранив при этом целостность и качество восстанавливаемого изображения.

Степень сжатия информации будем определять по формуле:

$$K = N_c / N_o \qquad (12)$$

где N_c – количество сжатых элементов, а N_o – общее количество анализируемых элементов.

Для базисов Фурье и косинусного преобразования общее количество анализируемых элементов матрицы преобразования будет находится по формуле:

$$N_o^{DFT} = N_o^{DCT} = N^2 \qquad (13)$$

где N – высота матрицы исходного изображения.

Так как при прямом ортогональном преобразовании Вейля-Гейзенберга матрицы изображения A размерности (N, N) мы получаем вещественную матрицу преобразования B размерности (2N, 2N), то общее количество анализируемых элементов будет определяться по формуле:

$$N_o^{DWHT} = 4N^2 \qquad (14)$$

На рисунке 5 представлены результаты сжатия изображения *«lena256.jpg»* с использованием оптимального базиса Вейля-Гейзенберга (*DWHT*) при различных пороговых значениях. В таблице 2 приводятся подробные характеристики результатов сжатия.

Рис 5. Восстановленное изображение при значениях порога (T = 10, 25, 50)

Таблица 2. Результаты сжатия изображения при различных пороговых значениях

Базис	*DWHT*		
Пороговое значение: T	10	25	50
Общее кол-во элементов: N_o	262144	262144	262144
Кол-во сжатых эл-в: N_c	220873	246595	255694
Потери качества: E	145.2892	327.8115	563.0511
Степень сжатия: K	0.8426	0.9407	0.9754

4. Сравнение ортогональных преобразований

Чтобы сравнить ортогональные преобразования *DWHT*, *DFT* и *DCT* по критерию качества восстановления, подберем такие пороговые значения T для базисов, при которых степень сжатия исходного изображения будет примерно одинаковой:

$$K = K_{DWHT} \approx K_{DFT} \approx K_{DCT} \tag{15}$$

Для сравнения базисов по степени сжатия, подберем такие пороговые значения, при которых показатели потери качества будут примерно одинаковыми, то есть:

$$E = E_{DWHT} \approx E_{DFT} \approx E_{DCT} \tag{16}$$

В таблицах 3, 4 приведены результаты сжатия исходного изображений с использованием *DWHT*, *DFT* и *DCT* при выбранных согласно условиям 15 и 16 пороговых значениях.

Таблица 3. Результаты сжатия при K = 0.83

Базис	*DWHT*	*DFT*	*DCT*
Пороговое значение: T	9.7	22	20
Общее кол-во элементов: N_o	262144	65536	65536
Кол-во сжатых эл-в: N_c	219699	54780	54488
Потери качества: E	141.2321	395.6927	306.3160
Степень сжатия: K	0.8381	0.8359	0.8314

Таблица 4. Результаты сжатия при E = 270

Базис	*DWHT*	*DFT*	*DCT*
Пороговое значение: T	20	14	17.3
Общее кол-во элементов: N_o	262144	65536	65536
Кол-во сжатых эл-в: N_c	242036	45394	51985
Потери качества: E	270.9123	270.9195	270.5099
Степень сжатия: K	0.9233	0.6927	0.7932

5. Выводы

По результатам исследования представленных выше алгоритмов сжатия можно заключить, что по сравнению с известными ортогональными базисами Фурье и косинусного преобразования оптимальный базис Вейля-Гейзенберга при фиксированной степени сжатия демонстрирует наименьшие потери качества при восстановлении изображения, а при требуемом показателе качества восстановления изображения дает наилучшую степень сжатия данных.

Таким образом, использование оптимального дискретного базиса Вейля-Гейзенберга оказывается весьма эффективным инструментом в задаче сжатия изображений. Это объясняется тем, что изображение представляет собой нестационарный двухмерный случайный процесс, а преобразование Вейля-Гейзенберга позволяет более корректно учитывать эти нестационарные особенности, чем другие известные ортогональные преобразования.

Литература

1. *В.П. Волчков*, «Сигнальные базисы с хорошей частотно-временной локализацией», М.: МТУСИ, 2009;

2. *Д.А. Петров*, «Синтез базиса Вейля-Гейзенберга», М.: МГУ им. М.В. Ломоносова, 2010.

3. *В.П. Волчков, Д.А. Петров*, «Оптимизация ортогонального базиса Вейля-Гейзенберга», Научные ведомости БелГУ, 2009;

4. *Р. Хорн, Ч. Джонсон,* «Матричный анализ», М.: Мир, 1989;

5. *Н. Ахмед, К.Р. Рао*, «Ортогональные преобразования при обработке цифровых сигналов», перевод с англ. Т.Э. Кренкеля, М.: Связь, 1980.