БИОРТОГОНАЛЬНЫЕ БАЗИСЫ ВСПЛЕСКОВ С РАЦИОНАЛЬНЫМИ МАСКАМИ И ИХ ПРИЛОЖЕНИЯ[1]

А.П.Петухов

Рассмотрены общие принципы построения биортогональных базисов всплесков, ассоциированных с рекурсивными фильтрами. Представлены быстрые алгоритмы разложения функций по данным базисам, в том числе, алгоритмы, опирающиеся на факторизации полифазных матриц. Приведены результаты численного моделирования алгоритмов сжатия изображений, основанные на разложении изображений по предлагаемому базису, которые позволяют сделать вывод о высокой эффективности предлагаемых алгоритмов, как с точки зрения вычислительных затрат, так и с точки зрения величины коэффициентов сжатия.

§1. Введение.

За последние 10 лет базисы всплесков стали мощным инструментом как для многих теоретических задач анализа так и для прикладных проблем, связанных с обработкой сигналов. Наибольшее развитие получила та часть теории базисов всплесков, которая связана с базисами, чьи функции имеют компактные носители. Алгоритмы разложения функций по таким базисам реализуются в виде набора дискретных сверток с числовыми последовательностями, имеющими лишь конечное число ненулевых коэффициентов. В радиотехнике действующие на бесконечных последовательностях перестановочные со сдвигом линейные операторы (т.е. операторы свертки) называют *(цифровыми) фильтрами*. Последовательность чисел, формирующая ядро свертки называют *импульсной характеристикой* фильтра. Таким образом, разложение по базисам всплесков с компактными носителями реализуется при помощи фильтров, имеющих конечную импульсную характеристику (КИХ-фильтров). Преобразование Фурье импульсной характеристики фильтра называют его *частотной характеристикой*. Следовательно, частотная характеристика фильтра с конечной импульсной характеристикой является тригонометрическим полиномом. Во многих разделах математики, например в теории аппроксимации, рациональные функции оказываются более гибким и эффективным инструментом, чем полиномы. Мы рассмотрим здесь базисы всплесков, разложение по которым реализовано на фильтрах с рациональной частотной характеристикой. Разумеется, появление новых видов базисов всплесков дает новые возможности применительно к прикладным задачам обработки сигналов и изображений.

[1] Работа выполнена при поддержке Российского фонда фундаментальных исследований (грант # 97-01-00443).

Несмотря на тот факт, что этот вид фильтров имеет бесконечную импульсную характеристику (БИХ-фильтры), существует их эффективная численная реализация в виде композиции хорошо известных в радиотехнике, так называемых, рекурсивных фильтров. Причем вычислительные затраты на реализацию фильтров с рациональной частотной характеристикой оказываются пропорциональными сумме степеней числителя и знаменателя, что сравнимо с затратами на реализацию разложений по базисам с компактными носителями, которые пропорциональны степеням соответствующих полиномов.

Впервые подобные базисы были подвергнуты систематическому исследованию в работе К.Херли и М.Веттерли [1], где был рассмотрен случай, когда соответствующие базисы ортогональны и образуют кратномасштабный анализ (КМА) пространства $\mathbb{L}^2(\mathbb{R})$. Как известно (см. [4]), за исключением хааровских базисов, не существует ортогональных базисов всплесков с компактными носителями, функции которых с точностью до сдвига аргумента являются четными или нечетными. Однако для базисов, порожденных рекурсивными БИХ-фильтрами, это не так. В [1] были найдены четные и нечетные базисы, ассоциированные с рекурсивными фильтрами. Ортогональность базисов позволяет получать результаты, гарантирующие сходимость и устойчивость алгоритмов разложения, значительно меньшими усилиями, чем в случаях, когда ортогональность отсутствует. Однако для того, чтобы использовать ортогональность нужно, по меньшей мере, иметь в исследуемом пространстве скалярное произведение. Но даже если скалярное произведение определено, для эффективного использования ортогональных базисов в прикладных задачах требуется, чтобы евклидово расстояние было естественным для данной задачи. К сожалению, это не всегда так. В частности, в обработке изображений наиболее эффективными оказались не ортогональные базисы всплесков, а биортогональные пары базисов. Например, для сжатия изображений наиболее часто используются так называемые 9/7-всплески с компактными носителями, полученные в работе [2]. Что касается базисов, построенных в [1], сочетающих ортогональность и симметрию, то ввиду жесткости требований их адаптация к конкретной задаче весьма проблематична. Результаты численного моделирования алгоритмов сжатия изображений с использованием таких базисов показали, что их эффективность достаточно мала по крайней мере для малых порядков рекурсивных фильтров. Напротив, инструмент биортогональных пар базисов, порожденных рекурсивными фильтрами, по гибкости заметно превосходит инструмент биортогональных пар базисов с компактными носителями.

О возможности построения биортогональных пар фильтров с рациональной частотной характеристикой говорится, например, в книге Г.Стренга и Т.Нгуена [7]. Там же имеется несколько примеров таких базисов, получающихся как побочный продукт при построении "хороших" базисов с компактными носителями. Мы будем здесь строить именно биортогональную теорию базисов, ассоциированных с рекурсивными фильтрами. Как мы увидим, использование такого подхода ведет в некоторых задачах к существенной экономии вычислительных затрат.

Изначально теория базисов всплесков развивалась независимо от возникшей несколько ранее в радиотехнике теории субполосного кодирования. Однако скоро стала ясна их почти полная идентичность. Во всяком случае, это можно утверждать в большинстве практически важных случаев. В §2 мы остановимся на их связи. К сожалению, в общем случае вопрос о том, когда пара фильтров, определяющая алгоритм субполосного кодирования, порождает пару биортогональных базисов, остается до сих пор открытым. К настоящему времени для биортогонального случая наиболее полно этот вопрос был исследован в работах [2], [5], но только для случая, когда ассоциированные с субполосным кодированием базисы принадлежат $\mathbb{L}^2(\mathbb{R})$ и имеют компактный носитель. Поэтому в §3 мы временно оставим вопрос о соответствии с базисами всплесков и, оставаясь на платформе субполосного кодирования, займемся построением биортогональных пар фильтров с рациональной частотной характеристикой.

В §4 будет для БИХ-фильтров построен алгоритм реализации субполосного кодирования и восстановления сигналов, основанный на, так называемой, лифтинг-схеме (см. [3], [6]), позволившей получить асимптотически двукратный выигрыш в вычислительных затратах по сравнению с прямой реализацией алгоритмов сверток при реализации субполосного кодирования на КИХ-фильтрах.

В §5 произведен подсчет вычислительной сложности предлагаемых алгоритмов и приведено сравнение с вычислительными затратами при прямой реализации алгоритмов разложения по всплескам.

В §6 мы приведем результаты численного моделирования алгоритмов сжатия изображений, позволяющие сравнить эффективность новых алгоритмов по сравнению с известными как по коэффициенту сжатия, так и по вычислительным затратам. Кроме того, будут приведены соображения, показывающие, что использованные в расчетах фильтры ассоциируются с некоторыми парами биортогональных базисов всплесков пространства $\mathbb{L}^2(\mathbb{R})$.

§2. Кратномасштабный анализ квазибанаховых пространств распределений.

Сначала мы перенесем определение Кратномасштабного анализа на квазибанаховы пространства, отличные от $\mathbb{L}^2(\mathbb{R})$. Напомним, что квазибанаховы пространства порождаются квазинормой, т.е. функционалом, чье отличие от нормы заключается в том, что неравенство треугольника для нее справедливо лишь в виде $\|f+g\| \le C(\|f\|+\|g\|)$, где $C > 1$. Такое ослабление свойств нормы позволяет включить в общее рассмотрение много полезных пространств типа пространств Харди. Полное рассмотрение теории Кратномасштабного анализа пространств периодических распределений можно найти в нашей статье [8] (см. также [9], [11]), случай периодических всплесков на дискретной сетке рассмотрен в [10]. Приводимые ниже определения всплесков в квазибанаховых пространствах непериодических распределений ни в коей мере не являются завешенной теорией. Это лишь набросок возможной схемы построения теории всплесков, который здесь нам необходим только для того, чтобы показать место тех частных базисов, построением которых мы будем здесь заниматься.

Обозначим через S' и S соответственно пространство распределений умеренного роста и пространство основных функций

Определение 1. Набор замкнутых линейных подпространств $\{V^j\}_{j\in\mathbb{Z}}$ квазибанахова пространства X, $S \subset X \subset S'$ называется *кратномасштабным анализом* КМА пространства X, если он удовлетворяет следующим условиям:

a) $\dots \subset V^{-2} \subset V^{-1} \subset V^0 \subset V^1 \subset \dots$;

b) $\overline{\cup V^j} = X$;

c) $f(x) \in V^j \Leftrightarrow f(2x) \in V^{j+1}$;

d) найдется такое распределение $\varphi \in V^0$, называемое *масштабирующей функцией*, что линейная оболочка системы распределений $\varphi(\bullet - k)$, $k \in \mathbb{Z}$ плотна в V^0.

Согласно пункту b) Определения 1 функция $\varphi(x)$ может быть приближена линейными комбинациями функции $\{\varphi(2x-k)\}$. Нас будет интересовать лишь случай, когда КМА образуется вещественными пространствами и имеет место *масштабирующее уравнение*

$$\varphi(x) = \sqrt{2}\sum_{n\in\mathbb{Z}} h_n \varphi(2x-n), \tag{2.1}$$

где h_n — коэффициенты Фурье рациональной функции

$$\sqrt{2}m_0(\omega) = \frac{\mathcal{P}(e^{i\omega})}{\mathcal{Q}(e^{i\omega})} = \sum_{n\in\mathbb{Z}} h_n e^{-in\omega}. \tag{2.2}$$

Ясно, что при этом h_n — вещественная последовательность. В (2.2) функции $\mathcal{P}(z)$ и $\mathcal{Q}(z)$ являются Лорановскими полиномами. Заметим, что мы используем здесь традиционный множитель $\sqrt{2}$, поскольку, если функции $\{\varphi(x-n)\}$ образуют ортонормированный базис пространства $V^0 \in \mathbb{L}^2(\mathbb{R})$, то функции $\{\sqrt{2}\varphi(2x-n)\}$ образуют ортонормированный базис в V^1. Соотношение (2.1) после применения к нему преобразования Фурье может быть переписано в виде

$$\hat{\varphi}(\omega) = m_0\left(\frac{\omega}{2}\right)\hat{\varphi}\left(\frac{\omega}{2}\right). \tag{2.3}$$

Если предположить, что $\hat{\varphi}$ — непрерывная функция и $\hat{\varphi}(0) \neq 0$, то ясно, что $m_0(0) = 1$. Таким образом, воспользовавшись (2.3), предполагая, что $\hat{\varphi}(0) = (2\pi)^{-1/2}$, можно доказать, что

$$\hat{\varphi}(\omega) = (2\pi)^{-1/2}\prod_{k=1}^{\infty} m_0\left(\frac{\omega}{2^k}\right),$$

где произведение в правой части сходится равномерно на компактных множествах.

Пусть $\{V^j\}$ КМА квазибанахова пространства X. Обозначим через Z° пополнение пространства S относительно квазинормы пространства Z. Будем в дальнейшем считать $X = X^\circ$. Для того чтобы ввести понятие всплеск-разложения пространства X нам понадобится КМА $\{\tilde{V}^j\}$ двойственного пространства $Y = X^{*\circ}$. При этом пространства

$$W^j = \{f \in V^{j+1} | \langle f, g\rangle = 0, \text{ для любых } g \in \tilde{V}^j\}$$

будем называть пространствами всплесков. Аналогичным образом определяются двойстенные пространства всплесков

$$\tilde{W}^j = \{g \in \tilde{V}^{j+1} | \langle f, g\rangle = 0, \text{ для любых } f \in V^j\}.$$

Мы не будем здесь обсуждать дополнительные условия для правильного определения пространств всплесков, однако отметим, что КМА $\{\tilde{V}^j\}$ сопряженного пространства должно быть выбранно так, чтобы $V^j \cap W^j = 0$, тогда

$$V^{j+1} = W^j \oplus V^j = \left(\bigoplus_{i=0}^{k} W^{j-i}\right) \bigoplus V^{j-k} = \left(\bigoplus_{i=0}^{+\infty} W^{j-i}\right) \bigoplus \left(\bigcap_{i\in\mathbb{Z}} V^i\right),$$

где значок $\oplus$ означает прямую сумму пространств. Добавка пересечения пространств V^i в последней формуле не могла появиться в случае КМА пространств $\mathbb{L}^2(\mathbb{R})$, поскольку для $\mathbb{L}^2(\mathbb{R})$ такое пересечение всегда пусто. Совсем по другому обстоит дело в общем случае. Нетрудно придумать примеры КМА, имеющие в пересечении полиномы фиксированной степени, δ-функцию Дирака или ее обобщенные производные.

Прокомментируем проведенное построение. С одной стороны, ясно, что классическое определение КМА пространства $\mathbb{L}^2(\mathbb{R})$ нуждается в обобщении на другие функциональные пространства, в которых не определено скалярное произведение. Однако, на первый взгляд, может показаться, что Определение 1 охватывает слишком обширный класс пространств и для приложений можно ограничиться классическими функциональными пространствами, не переходя к обобщенным функциям. С этим можно было бы согласиться, если бы пространства V^j использовались лишь как аппроксимации пространства X. Однако в определении пространств всплесков участвуют два КМА взаимно-сопряженных пространств, причем один из них служит для приближения, а второй — для определения проекции. А поскольку, чем у́же исходное пространство, тем шире пространство сопряженное, то для определения проекции на пространство достаточно гладких функций может оказаться недостаточным наличия функционалов, определяемых обычными функциями. Например, при определении интерполяционных всплесков, примеры которых будут приведены ниже, проекторы задаются сдвигами дельта-функции Дирака — самым простым примером обобщенной функции. Этот пример является вполне естественным следствием того, что пространством, сопряженным к пространству непрерывных функций, является пространство конечных борелевских мер. Заметим, что пространство конечных борелевских мер не является сепарабельным. Поэтому, если $X = \mathbb{C}(\mathbb{R})$ — пространство непрерывных функций, стремящихся к нулю на бесконечности, то пространство конечных борелевских мер $\mathbb{C}^*(\mathbb{R})$ не совпадает с $\mathbb{C}^{*\circ}(\mathbb{R})$. Однако если в качестве X взять чуть более узкое пространство (например, определяемое каким-нибудь модулем непрерывности), то пространство борелевских мер войдет в сепарабельное подпространство пространства X^*.

Напомним, что две системы функций $\{f_n\} \subset X$ и $\{g_n\} \subset X^*$ называются биортогональными, если справедливы соотношения

$$\langle f_k, g_m \rangle = \begin{cases} 1, & k = m, \\ 0, & k \neq m. \end{cases}$$

Предположим, что в пространстве $\tilde{V}^0$ имеется биортогональный (базису $\{\varphi(x-n)\}$)

базис $\{\tilde{\varphi}(x-n)\}$, порожденный рациональной функцией

$$\tilde{m}_0(\omega) = \frac{1}{\sqrt{2}} \frac{\tilde{\mathcal{P}}(e^{i\omega})}{\tilde{\mathcal{Q}}(e^{i\omega})} = \frac{1}{\sqrt{2}} \sum_{n\in\mathbb{Z}} \tilde{h}_n e^{-in\omega}.$$

Ввиду условия биортогональности нетрудно получить, что m_0 и $\tilde{m}_0$ должны удовлетворять условию

$$m_0(\omega)\overline{\tilde{m}_0(\omega)} + m_0(\omega+\pi)\overline{\tilde{m}_0(\omega+\pi)} = 1.$$

Заметим, что поскольку сходимость бесконечных произведений, определяющих функции $\hat{\varphi}$ и $\hat{\tilde{\varphi}}$, влечет равенства $m(0) = \tilde{m}(0) = 1$, получим, что из условие биортогональности имеем $m(\pi)\overline{\tilde{m}(\pi)} = 0$. Это означает, что хотя бы один из сомножителей $m(\pi)$ и $\tilde{m}(\pi)$ равен нулю.

Из определения W^j получим, что преобразование Фурье функции ψ, порождающей базис $\{\psi(x-n)\}$ пространства W^0, может быть вычислено по формуле

$$\hat{\psi}(\omega) = m_1(\frac{\omega}{2})\hat{\varphi}(\frac{\omega}{2}),$$

где m_1 удовлетворяет уравнению

$$m_1(\omega)\overline{\tilde{m}_0(\omega)} + m_1(\omega+\pi)\overline{\tilde{m}_0(\omega+\pi)} = 0.$$

Отсюда

$$m_1(\omega) = \alpha(\omega)e^{-i\omega}\overline{\tilde{m}_0(\omega+\pi)},$$

где α — некоторое π-периодическое распределение. В дальнейшем, скорее следуя традициям в обозначениях, а не сути дела, мы будем для определенности считать $\alpha(\omega) \equiv 1$. Аналогичным образом введем в рассмотрение функцию

$$\tilde{m}_1(\omega) = e^{-i\omega}\overline{m_0(\omega+\pi)},$$

которая порождает базис в двойственном пространстве всплесков $\tilde{W}^0$. Таким образом, имеет место матричное соотношение

$$\begin{pmatrix} m_0(\omega) & m_0(\omega+\pi) \\ m_1(\omega) & m_1(\omega+\pi) \end{pmatrix} \cdot \begin{pmatrix} \tilde{m}_0(\omega) & \tilde{m}_0(\omega+\pi) \\ \tilde{m}_1(\omega) & \tilde{m}_1(\omega+\pi) \end{pmatrix}^* = I, \tag{2.4}$$

где $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Таким образом, уравнение (2.4) дает необходимое условие для того, чтобы функции m_0 и $\tilde{m}_0$ порождали непрерывные масштабирующие функции φ, $\tilde{\varphi}$, которые определяют пару двойственных КМА. Что касается достаточных условий, то в настоящее время эта проблема далека от полного разрешения, хотя в большинстве практически важных случаев не составляет труда показать, что конкретное решение уравнения

(2.4), действительно, генерирует биортогональную пару базисов некоторых двойственных КМА.

В случае домножения функции $m_1(\omega)$ на π-периодическую функцию $\alpha(\omega)$ равенство (2.4) нарушится. Для его восстановления необходимо функцию $\tilde{m}_1(\omega)$ домножить на $1/\overline{\alpha(\omega)}$, получив при этом

$$\begin{pmatrix} m_0(\omega) & m_0(\omega+\pi) \\ m_1(\omega)\alpha(\omega) & m_1(\omega+\pi)\alpha(\omega) \end{pmatrix} \cdot \begin{pmatrix} \tilde{m}_0(\omega) & \tilde{m}_0(\omega+\pi) \\ \tilde{m}_1(\omega)/\overline{\alpha(\omega)} & \tilde{m}_1(\omega+\pi)/\overline{\alpha(\omega)} \end{pmatrix}^* = I. \qquad (2.4')$$

Разумеется, выбор $\alpha(\omega)$ никак не влияет ни на пару двойственных КМА, которые определяются функциями m_0, $\tilde{m}_0$, ни на выбор базисов в соответствующих пространствах V^j и $\tilde{V}^j$, ни на выбор пространств всплесков W^j и $\tilde{W}^j$. Однако варьирование функции α дает дополнительную степень свободы при выборе базисов всплесков в наибольшей степени соответствующих конкретной прикладной задаче. В дальнейшем мы увидим, что базис, наиболее подходящий для сжатия изображений, использует эту степень свободы. Причина невостребованности этого приема в текущей литературе носит объективный характер. Дело в том, что рассматривая биортогональные базисы всплесков, ассоциированные с КИХ-фильтрами мы вынуждены искать полиномиальные решения уравнения (2.4). Если m_0 и $\tilde{m}_0$ — полиномы, то традиционный выбор функций m_1 и $\tilde{m}_1$ также является полиномиальным. Если же мы домножим функцию m_1 на π-периодический полином α, то мы обязаны разделить $\tilde{m}_1$ на $\overline{\alpha}$. Предположим, что найдется такой π-периодический полином α, что вторая матрица в (2.4′) является полиномиальной. В вырожденном случае, когда α — одночлен, т.е. с точностью до умножения на постоянную имеем $\alpha(\omega) = e^{i2k\omega}$, $k \in \mathbb{Z}$, такой выбор приводит к сдвигу элементов базисов в W^{-1} и $\tilde{W}^{-1}$ на 2, т.е. фактически лишь к перенумерации элементов базиса. Если же α — невырожденный полином, то для того, чтобы функция $\tilde{m}_1/\overline{\alpha}$ была полиномом, требуется совпадение нулей $\tilde{m}_1$ с нулями полинома $\overline{\alpha}$, но тогда ввиду π-периодичности полинома α полином $\tilde{m}_1(\omega+\pi)$ также имеет нули в тех точках, где $\alpha(\omega) = 0$. Таким образом, найдется точка ω_0, в которой вектор $(\tilde{m}_1(\omega_0), \tilde{m}_1(\omega_0+\pi))$ становится нулевым, а соответствующая матрица в (2.4), становится вырожденной, что приводит к нарушению равенства (2.4) не только в самой этой точке, но, ввиду непрерывности компонент, и в некоторой ее окрестности. В случае же, когда допускаются рациональные решения уравнения (2.4), нулям определителя одной из матриц в (2.4′) можно поставить в соответствие полюсы другого определителя. Таким образом, традиционный выбор функций m_1 и $\tilde{m}_1$ хотя и вполне оправдан, но является единственно возможным только для полиномиальных решений.

Двойственные КМА $\{V^j\}$ и $\{\tilde{V}^j\}$ в определенном смысле равноправны и могут рассматриваться как в качестве вложенной последовательности пространств в пространстве сигналов, так и для определения операторов проектирования, которые в свою очередь позволяют определить пространства всплесков. Мы в дальнейшем условимся считать, что сигналы проектируются на пространства $\tilde{V}^j$ при помощи пространств V^j. Такая договоренность принята в подавляющем большинстве публикаций.

Введем в рассмотрение функции

$$h(z) = \sum_{n\in\mathbb{Z}} h_n e^{-in\omega} = \sqrt{2} m_0(\omega), \quad \tilde{h}(z) = \sum_{n\in\mathbb{Z}} \tilde{h}_n e^{-in\omega} = \sqrt{2}\tilde{m}_0(\omega), \qquad (2.5)$$

где $z = e^{i\omega}$, и

$$g(z) = z^{-1}\tilde{h}(-z^{-1}), \quad \tilde{g}(z) = z^{-1}h(-z^{-1}). \tag{2.6}$$

Функции h и $\tilde{h}$ будем называть *биортогональной парой фильтров*. Теперь равенство (2.4) в случае вещественных $\{h_n\}$ и $\{\tilde{h}_n\}$ может быть переписано в виде

$$M(z)\tilde{M}^t(z^{-1}) = 2I, \tag{2.7}$$

где символ t означает транспонирование матрицы, а

$$M(z) = \begin{pmatrix} h(z) & h(-z) \\ g(z) & g(-z) \end{pmatrix}, \quad \tilde{M}(z) = \begin{pmatrix} \tilde{h}(z) & \tilde{h}(-z) \\ \tilde{g}(z) & \tilde{g}(-z) \end{pmatrix} \text{ —}$$

так называемые, *модуляционные матрицы*, чьи компоненты в нашем случае являются рациональными функциями. Из (2.6) и (2.7) нетрудно получить, что $\det|M(z)| = \det|\tilde{M}(z)| = -2z^{-1}$.

Поясним смысл и назначение модуляционных матриц M и $\tilde{M}$. Пусть $\tilde{v}^0(x) = \sum_n c_n^0\tilde{\varphi}(x-n) \in \tilde{V}^0$ и требуется найти проекции $\tilde{v}^{-1}(x) = \sum_n c_n^{-1}\tilde{\varphi}(x/2-n)$ и $\tilde{w}^{-1}(x) = \sum_n d_n^{-1}\tilde{\psi}(x/2-n)$ распределения $\tilde{v}^0$ на пространства $\tilde{V}^{-1}$ и $\tilde{W}^{-1}$. Воспользовавшись масштабирующим уравнением (2.1), получим

$$c_m^{-1} = \langle\tilde{v}^0(x), \varphi(x/2-m)\rangle = \sqrt{2}\langle\sum_{n\in\mathbb{Z}} c_n^0\tilde{\varphi}(x-n), \sum_{k\in\mathbb{Z}} h_k\varphi(x-k-2m)\rangle = \sqrt{2}\sum_{n\in\mathbb{Z}} c_n^0 h_{n-2m},$$

Последнее равенство в терминах z-преобразований последовательностей можно записать в виде $c^{-1}(z^{-2}) = \frac{1}{\sqrt{2}}(h(z)c^0(z^{-1}) + h(-z)c^0(-z^{-1}))$, где $c^0(z) = \sum_n c_n^0 z^{-n}$, $c^{-1}(z) = \sum_n c_n^{-1}z^{-n}$. Заметим, что сходимость рядов здесь и в дальнейшем обеспечивается экспоненциальным убыванием последовательностей $\{h_n\}$ и $\{\tilde{h}_n\}$. Аналогичным образом получается равенство $d_m^{-1} = \sqrt{2}\sum_{n\in\mathbb{Z}} c_n^0 g_{n-2m}$, а значит, $d^{-1}(z^{-2}) = \frac{1}{\sqrt{2}}(g(z)c^0(z^{-1}) + g(-z)c^0(-z^{-1}))$.

Таким образом, после перехода к z-преобразованиям матрица M становится матрицей перехода от пространства $\tilde{V}^0$ к прямой сумме пространств $\tilde{V}^{-1}$ и $\tilde{W}^{-1}$, определяемого по формуле

$$\begin{pmatrix} c^{-1}(z^{-2}) \\ d^{-1}(z^{-2}) \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} h(z) & h(-z) \\ g(z) & g(-z) \end{pmatrix}\begin{pmatrix} c^0(z^{-1}) \\ c^0(-z^{-1}) \end{pmatrix}. \tag{2.8}$$

Для восстановления коэффициентов разложения функции в базисе пространства $\tilde{V}^0$ по ее проекциям на пространства $\tilde{V}^{-1}$ и $\tilde{W}^{-1}$ имеет место формула

$$\begin{pmatrix} c^0(z^{-1}) \\ c^0(-z^{-1}) \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} \tilde{h}(z^{-1}) & \tilde{g}(z^{-1}) \\ \tilde{h}(-z^{-1}) & \tilde{g}(-z^{-1}) \end{pmatrix}\begin{pmatrix} c^{-1}(z^{-2}) \\ d^{-1}(z^{-2}) \end{pmatrix}.$$

Разумеется, последние формулы удобны для аналитического представления операции разложения по всплескам и восстановления по проекциям, но содержат много повторяющихся операций и потому не могут быть рекомендованы для непосредственного вычисления коэффициентов.

В каком-то смысле альтернативным методом представления операции разложения пространства $\tilde V^0$ в прямую сумму пространств $\tilde V^{-1}$ и $\tilde W^{-1}$ является его реализация при помощи, так называемой, полифазной матрицы. Для произвольного формального ряда Лорана $f(z)=\sum_n f_n z^{-n}$ введем обозначения $f_e(z)=\sum_n f_{2n}z^{-n}$ и $f_o(z)=\sum_n f_{2n+1}z^{-n}$. Тогда справедливо представление $f(z)=f_e(z^2)+z^{-1}f_o(z^2)$. Если теперь воспользоваться этим представлением, заменив им все компоненты формулы (2.8), то после несложных преобразований правой части получим

$$\begin{pmatrix} c^{-1}(z^{-1}) \\ d^{-1}(z^{-1}) \end{pmatrix} = \begin{pmatrix} h_e(z) & h_o(z) \\ g_e(z) & g_o(z) \end{pmatrix} \begin{pmatrix} c_e^0(z^{-1}) \\ c_o^0(z^{-1}) \end{pmatrix}. \tag{2.9}$$

Именно эта формула является отправной точкой для построения быстрых алгоритмов разложения и восстановления сигналов, объединенных общим названием лифтинг-схема. Матрицу в выражении (2.9) называют *полифазной*. Ее мы будем обозначать через $P(z)$, а полифазную матрицу двойственного КМА — через $\tilde P(z)$. Нетрудно проверить, что

$$P(z)\tilde P^t(z^{-1}) = I. \tag{2.10}$$

Кроме того, поскольку для полифазных матриц P и $\tilde P$ справедливы представления

$$P(z^2)=\frac{1}{2}M(z)\begin{pmatrix} 1 & z \\ 1 & -z \end{pmatrix}, \quad \tilde P(z^2)=\frac{1}{2}\tilde M(z)\begin{pmatrix} 1 & z \\ 1 & -z \end{pmatrix},$$

из равенств $\det M=\det\tilde M=-2z^{-1}$ имеем $\det P=\det\tilde P=1$.

§3. Примеры.

Уравнение (2.7) имеет много рациональных решений. Хотя из дальнейших рассуждений легко видеть, как можно получить все возможные решения такого типа, нас будут здесь интересовать лишь некоторые из них, которые имеют минимальный порядок числителя и знаменателя и соответствуют базисам всплесков, чьи масштабирующие функции являются симметричными относительно целых или полуцелых точек. При этом мы покажем, что после подходящего сдвига аргумента соответствующие базисы всплесков являются в первом случае нечетными, а во втором случае — четными функциями. Поэтому для краткости будем говорить, что имеет место соответственно нечетный или четный случай.

Рассмотрим сначала четный случай. Заметим, что домножение функции $h(z)$ на z^k означает переход от масштабирующей функции $\varphi(z)$ к масштабирующей функции $\varphi(z+k)$, т.е. переход к одной из функций базисной системы сдвигов. Нетрудно видеть, что домножив h, если нужно, на подходящую целую степень переменной z, мы получим, что в (2.5) $h_n=h_{-n}$ для всех $n=1,2,\ldots$. Функция $h(z)$ вместе с каждым полюсом $z=z_0$ обладает и полюсом $z=z_0^{-1}$. Таким образом, рациональная функция h представима в виде отношения двух полиномов $S(z)=\sum_{n=-l}^{l}s_nz^{-n}$ и $Q(z)=\sum_{n=-m}^{m}q_nz^{-n}$ с симметричными последовательностями коэффициентов $s_{-n}=s_n$ и $q_{-n}=q_n$, а значит, инвариантных относительно замены переменной $z\to z^{-1}$.

В нечетном случае, домножив h, если нужно, на подходящую целую степень переменной z, получим $h_n = h_{1-n}$ для всех $n = 1, 2, \ldots$. Ясно, что функция $f(z) = (1+z)h(z)$ имеет симметричную последовательность лорановских коэффициентов. Последовательность лорановских коэффициентов функции $f(z)/((1+z)(1+z^{-1}))$ симметрична, а сама эта функция является рациональной и представима в виде $S(z)/Q(z)$, где S и Q — лорановские полиномы с симметричными последовательностями коэффициентов, инвариантные относительно замены $z \mapsto z^{-1}$. Таким образом,

$$h(z) = \frac{(1+z^{-1})S(z)}{Q(z)}.$$

Покажем, что фильтры h и $\tilde{h}$ могут иметь разноименную четность, т.е., когда одна функция соответствует четному случаю, а другая — нечетному, лишь для пары фильтров $h(z) = (1+z^{-1})/\sqrt{2}$ и $\tilde{h}(z) = \sqrt{2}$. Несмотря на то, что данная пара фильтров приводит к двум КМА, порожденным соответственно базисом Хаара и базисом, состоящим из сдвигов δ-функции, ввиду разрывности хааровских функций эти базисы не образуют биортогональной пары.

Из (2.7) для произвольной пары h и $\tilde{h}$ получим, что

$$h(z)\tilde{h}(z^{-1}) + h(-z)\tilde{h}(-z^{-1}) = 2$$

или, подставляя представление $h(z)$ и $\tilde{h}(z)$ в виде рациональных функций и сокращая при необходимости общие множители числителя и знаменателя, получим выражение вида

$$\frac{A(z)}{B(z)} + \frac{A(-z)}{B(-z)} = 2, \tag{3.1}$$

где A, B — лорановские полиномы от z. Слагаемые в (3.1) должны иметь общие с учетом кратности полюсы. Отсюда, в частности, следует, что полиномы $B(z)$ и $B(-z)$ имеют одни и те же нули. Таким образом, получим, что в полиноме B, по крайней мере после домножения в случае необходимости на подходящий множитель z^k, присутствуют лишь четные степени, т.е. $B(z) = \mathcal{B}(z^2)$, где $\mathcal{B}(z)$ — некоторый лорановский полином. При этом уравнение (3.1) можно переписать в виде

$$A(z) + A(-z) = 2\mathcal{B}(z^2). \tag{3.2}$$

Последнее уравнение является простым источником всех возможных КМА. В самом деле, из уравнения (3.2) видно, что лорановский полином A может быть выбран в определенной степени произвольно. В качестве $A_e(z) = \mathcal{B}(z)$ можно взять произвольный лорановский полином, не имеющий корней на единичной окружности $|z| = 1$. При этом, поскольку $A(-1) = m_0(\pi)\overline{\tilde{m}_0(\pi)} = 0$, полином A_o должен удовлетворять равенству $A_o(1) = A_e(1)$, т.е. должны совпадать суммы коэффициентов полиномов A_o и A_e. Все возможные биортогональные фильтры h и $\tilde{h}$ получаются факторизацией $h(z)\tilde{h}(z^{-1}) = A(z)/B(z)$, где $h(1) = \tilde{h}(1) = \sqrt{2}$. Разумеется, не все факторизации приводят к биортогональным парам базисов всплесков, однако все такие пары базисов порождаются этими факторизациями.

Предположим теперь, что одна из функций h и $\tilde{h}$ порождает четный КМА, а другая — нечетный. Тогда справедливо представление $A(z) = (1+z^{-1})\mathcal{A}(z)$, где $\mathcal{A}(z)$ — лорановский полином с симметричной последовательностью коэффициентов. Подставляя это представление в (3.2), получим

$$(1+z^{-1})\mathcal{A}(z) + (1-z^{-1})\mathcal{A}(-z) = 2\mathcal{B}(z^2), \tag{3.3}$$

где без ограничения общности можно считать, что полиномы $\mathcal{A}$ и $\mathcal{B}$ не имеют общих нулей. Правая часть равенства (3.3) инвариантна относительно замены переменной $z \to z^{-1}$, поскольку она является произведением полиномов, обладающих этим свойством. Значит, и левая его часть инвариантна относительно такой замены. Следовательно,

$$(1+z^{-1})\mathcal{A}(z) + (1-z^{-1})\mathcal{A}(-z) = (1+z)\mathcal{A}(z^{-1}) + (1-z)\mathcal{A}(-z^{-1}).$$

А поскольку $\mathcal{A}(z) = \mathcal{A}(z^{-1})$, то последнее равенство преобразуется к виду

$$(z^{-1} - z)(\mathcal{A}(z) - \mathcal{A}(-z)) = 0,$$

что означает отсутствие у многочлена $\mathcal{A}$ нечетных степеней. Но тогда, вернувшись к (3.3), увидим, что $\mathcal{A}(z) = \mathcal{B}(z^2)$. А значит, ввиду отсутствия у этих многочленов общих делителей имеем $\mathcal{A} \equiv 1$. Данному случаю соответствует биортогональная пара фильтров $h(z) = (1+z^{-1})/\sqrt{2}$ и $\tilde{h}(z) = \sqrt{2}$. Таким образом, h и $\tilde{h}$ могут обладать разной четностью лишь для приведенного выше тривиального случая, который не порождает биортогональные базисы всплесков.

Итак, рассмотрим четные и нечетные биортогональные пары с малыми порядками числителя и знаменателя функций h и $\tilde{h}$. Начнем с нечетного случая. Самая простая правая часть в (3.2) может быть представлена в виде $\mathcal{B} \equiv 1$. На первый взгляд, этот случай может показаться вырожденным, соответствующим биортогональным базисам с компактным носителем. Тем не менее, могло оказаться, что число, бывшее корнем одной из функций $h(z)$ и $\tilde{h}(z^{-1})$ оказалось полюсом другой функции. Тогда соответствующие множители в произведении $h(z)\tilde{h}(z^{-1})$ сократятся и потому никак не будут учтены в (3.1). Специфика КМА вынуждает нас рассматривать не все решения уравнения (3.2). Главное из ограничений заключается в том, что поскольку для нечетного случая $h(-1) = \tilde{h}(-1) = 0$, то $A(z) = (1+z)(1+z^{-1})\mathcal{A}_1(z)$. Таким образом, многочлен $\mathcal{A}_1(z)$ минимально возможной степени равен константе, $\mathcal{A}_1(z) \equiv 1/2$. А учитывая, что рациональные функции $h(z)/(1+z^{-1})$ и $\tilde{h}(z)/(1+z^{-1})$ имеют симметричные последовательности лорановских коэффициентов, получим, что биортогональная рациональная пара фильтров минимально возможного порядка имеет вид

$$h(z) = \alpha\frac{1+z^{-1}}{\sqrt{2}(z^{-1}+a+z)}, \qquad \tilde{h}(z) = \frac{1}{\sqrt{2}\alpha}(1+z^{-1})(z^{-1}+a+z),$$

где $\alpha = 2 + a$ — множитель, который обеспечивает нормировку $h(1) = \tilde{h}(1) = \sqrt{2}$.

Для теории всплесков интерес представляет лишь случай $|a| > 2$. В самом деле, полином $1 + az + z^2$ имеет корни $z_{1,2} = -\frac{a}{2} \pm \sqrt{\frac{a^2}{4} - 1}$. При $|a| \leq 2$ оба корня равны по модулю единице, т.е. $z_{1,2} = e^{\pm i\omega_0}$. Следовательно, в этом случае функция $m_0(\omega)$ имеет особенности в точках $\omega = \pm\omega_0$ и не только несуммируема, но и не является даже распределением.

Рассмотренный пример показывает, как можно варьировать имеющиеся КМА (в нашем случае КМА, порожденные базисами Хаара) путем введения дополнительного множителя $z^{-1} + a + z$ в знаменатель одного из фильтров h, $\tilde{h}$ и того же множителя — в числитель другого. С качественной точки зрения этот прием ведет к "перекачке гладкости" из одного базиса биортогональной пары в двойственный ему базис при сохранении суммарной гладкости базисов. Такой прием может быть использован для варьирования любой биортогональной пары, и это следует иметь в виду, хотя мы о нем далее упоминать не будем.

Теперь рассмотрим менее тривиальный случай, когда полином $\mathcal{B}$ невырожден. В простейшем случае $\mathcal{B}(z) = z^{-1} + a + z$. По приведенным выше причинам здесь также имеем $|a| > 2$. При этом получаем, что удовлетворяющий (3.2) полином $\mathcal{A}_1$ минимально возможной степени имеет вид $\mathcal{A}_1(z) = z^{-1} + (\frac{a}{2} - 1) + z$. Таким образом, существует две биортогональные пары, соответствующие данному выбору полинома $\mathcal{B}$:

$$h(z) = \sqrt{2}\frac{(1 + z^{-1})(z^{-1} + (\frac{a}{2} - 1) + z)}{z^{-2} + a + z^2}, \qquad \tilde{h}(z) = \frac{1 + z^{-1}}{\sqrt{2}}$$

и

$$h(z) = \alpha\frac{(1 + z^{-1})}{z^{-2} + a + z^2}, \qquad \tilde{h}(z) = \frac{(1 + z^{-1})(z^{-1} + (\frac{a}{2} - 1) + z)}{\alpha}, \text{ где } \alpha = \frac{a + 2}{\sqrt{2}}$$

Рассмотрим теперь простейшие решения в четном случае. Поскольку в (3.1) $A(-1) = 0$, то один из полиномов h или $\tilde{h}$ имеет нуль в точке $z = -1$. А так как последовательность его коэффициентов симметрична, то кратность корня четна. Значит, полином $A(z)$ включает в себя множитель $z^{-1} + 2 + z$. Таким образом, в простейшем случае $A(z) = \frac{1}{2}(z^{-1} + 2 + z)$, $\mathcal{B}(z) \equiv 1$,

$$h(z) = \sqrt{2}, \qquad \tilde{h}(z) = \frac{z^{-1} + 2 + z}{2\sqrt{2}}.$$

Данная пара КИХ-фильтров приводит к паре биортогональных КМА, первый из которых порождается сдвигами δ-функции, а другой - сдвигами линейного B-сплайна. На первый взгляд, может показаться, что этот пример является вырожденным и не имеет перспективы найти применение в приложениях. В самом деле, "дефекты" этой пары состоят в том, что одно из пространств двойственной пары состоит из обобщенных функций, кроме того, функции базисов кусочно-линейных всплесков не имеют нулевого среднего. Второе свойство с общепринятой точки зрения кажется особенно неприятным, поскольку считается, что хорошие для аппроксимации (в частности для сжатия информации) базисы должны иметь много нулевых моментов или, что то же самое, полином $h(z)$ должен иметь высокую кратность нуля в точке $z=-1$. Тем не менее, оказывается, что в большинстве случаев использование этой пары базисов для сжатия изображений приводит к бо́льшим коэффициентам сжатия, чем использование вполне благополучных хааровских базисов. Учитывая высокую вычислительную эффективность этих базисов, можно предположить, что они вполне могут быть востребованы в приложениях.

Рассмотрим теперь случай, когда $\mathcal{B}(z)=z^{-1}+a+z$. Для выбранного ранее полинома $\mathcal{A}_1$ возможны четыре варианта четной факторизации $h(z)\tilde{h}(z^{-1})=(z^{-1}+2+z)\mathcal{A}_1(z)$:

$$h(z)=\frac{(z^{-1}+2+z)(z^{-1}+(\frac{a}{2}-1)+z)}{\sqrt{2}(z^{-2}+a+z^2)},\qquad \tilde{h}(z)=\sqrt{2};$$

$$h(z)=\frac{(a+2)(z^{-1}+2+z)}{2\sqrt{2}(z^{-2}+a+z^2)},\qquad \tilde{h}(z)=2\sqrt{2}\frac{z^{-1}+(\frac{a}{2}-1)+z}{a+2};$$

$$h(z)=\frac{z^{-1}+2+z}{2\sqrt{2}},\qquad \tilde{h}(z)=2\sqrt{2}\frac{z^{-1}+(\frac{a}{2}-1)+z}{z^{-2}+a+z^2};$$

$$h(z)=\frac{(z^{-1}+(\frac{a}{2}-1)+z)(z^{-1}+2+z)}{\sqrt{2}(a+2)},\qquad \tilde{h}(z)=\frac{\sqrt{2}(a+2)}{z^{-2}+a+z^2}.$$

Заметим, во всех четырех случаях $\tilde{h}(-1)\neq 0$.

Ясно, что для выполнения соотношения $h(-1)=\tilde{h}(-1)=0$ в четном случае приходится потребовать, чтобы полином $A(z)$ в (3.2) имел в точке -1 нуль по крайней мере четвертого порядка. Таким образом, справедливо представление $A(z)=(1+z)^2(1+z^{-1})^2\mathcal{A}_2(z)$. Пусть снова $\mathcal{B}(z)=z^{-1}+a+z$. Тогда, решая уравнение (3.2) для $\mathcal{A}_2(z)$ вида $bz^{-1}+c+bz$, получим, что

$$\mathcal{A}_2(z)=\frac{6-a}{16}z^{-1}+\frac{a-2}{4}+\frac{6-a}{16}z,$$

где, как и ранее, $|a|>2$. Таким образом, полагая $b=\frac{6-a}{16}$, $c=\frac{a-2}{4}$, получим две различные биортогональные пары

$$h(z)=\frac{2(z^{-1}+2+z)(bz^{-1}+c+bz)}{\sqrt{2}(z^{-2}+a+z^2)},\qquad \tilde{h}(z)=\frac{z^{-1}+2+z}{2\sqrt{2}} \tag{3.4}$$

и

$$h(z)=\frac{4(z^{-1}+2+z)(bz^{-1}+c+bz)}{\sqrt{2}(a+2)},\quad \tilde{h}(z)=\frac{\sqrt{2}(a+2)(z^{-1}+2+z)}{4(z^{-2}+a+z^2)}. \tag{3.5}$$

Формулы (3.4) и (3.5) имеют два важных частных случая, когда алгоритмы разложения по всплескам значительно упрощаются. Первый из этих случаев $a=6$ приводит нас к упрощенной биортогональной паре

$$h(z)=\frac{z^{-1}+2+z}{2\sqrt{2}}.\qquad \tilde{h}(z)=\frac{2\sqrt{2}(z^{-1}+2+z)}{z^{-2}+6+z^2},$$

Второй случай при предельном переходе $a \to \infty$ приводит к известной (см. [2], [4, с. 277]) биортогональной паре базисов с компактным носителем, определяемой функциями

$$h(z) = \frac{(z^{-1}+2+z)(-z^{-1}+4-z)}{4\sqrt{2}}, \qquad \tilde{h}(z) = \frac{z^{-1}+2+z}{2\sqrt{2}}.$$

Наконец, еще две важные пары фильтров получаются при $a = 10/3$. При данном значении параметра a получаются фильтры, имеющие максимальную суммарную кратность нуля в точке -1, равную 6. Фильтры (3.4) трансформируется при этом в фильтры

$$h(z) = \frac{\sqrt{2}(z^{-1}+2+z)^2}{3z^{-2}+10+3z^2}, \qquad \tilde{h}(z) = \frac{z^{-1}+2+z}{2\sqrt{2}}$$

а (3.5) — в фильтры

$$h(z) = \frac{(z^{-1}+2+z)^2}{8\sqrt{2}}, \qquad \tilde{h}(z) = \frac{4\sqrt{2}(z^{-1}+2+z)}{3z^{-2}+10a+3z^2}.$$

Заметим, что кроме двух пар (3.4) и (3.5) в данном случае существует еще четыре различных факторизации, у которых один из фильтров h или $\tilde{h}$ не имеет нуля в точке -1.

§4. Метод лифтинга.

Наибольший прогресс в проблеме ускорения работы алгоритмов разложения функций по базисам всплесков и их восстановления по коэффициентам разложения связан с так называемой лифтинг-схемой, принявшей завершенный вид в работе [3]. Близкие идеи были реализованы в [6] и [15]. Именно на подход, рассмотренный в последней работе [6] мы и будем здесь опираться.

Введем необходимые обозначения. Пусть $\Pi(z)$ — лорановский полином $\Pi(z) = \sum_{n=k_1}^{k_2} p_n z^n$, где $k_1, k_2 \in \mathbb{Z}$, тогда $l(\Pi) \stackrel{\text{def}}{=} k_1$, $u(\Pi) \stackrel{\text{def}}{=} k_2$, $L(\Pi) \stackrel{\text{def}}{=} p_{l(\Pi)}$, $U(\Pi) \stackrel{\text{def}}{=} p_{u(\Pi)}$. Величину $d(\Pi) \stackrel{\text{def}}{=} u(\Pi) - l(\Pi)$ будем называть порядком полинома Π.

Сначала мы рассмотрим случай, когда каждая из рациональных функций h и $\tilde{h}$ не имеет нулей, являющихся полюсами другой функции. В этом случае, как было показано в предыдущем, параграфе знаменатели этих функций являются лорановскими полиномами от z^2, что позволяет записать компоненты полифазной матрицы P в виде:

$$h_e(z) = \frac{E_1(z)}{Q_1(z)}, \quad h_o(z) = \frac{O_1(z)}{Q_1(z)}, \quad g_e(z) = \frac{E_2(z)}{Q_2(z)}, \quad g_o(z) = \frac{O_2(z)}{Q_2(z)}. \tag{4.1}$$

Равенство единице определителя полифазной матрицы влечет выполнение соотношения

$$\det \begin{pmatrix} E_1(z) & E_2(z) \\ O_1(z) & O_2(z) \end{pmatrix} = Q_1(z) Q_2(z). \tag{4.2}$$

Отсюда ясно, что справедливы два неравенства

$$\mathcal{L}(P) \overset{\text{def}}{=} l(Q_1Q_2) - \min\{l(E_1O_2), l(O_1E_2)\} \geq 0;$$

$$\mathcal{U}(P) \overset{\text{def}}{=} \max\{u(E_1O_2), u(O_1E_2)\} - u(Q_1Q_2) \geq 0.$$

Число $\mathcal{D}(P) \overset{\text{def}}{=} \mathcal{L}(P) + \mathcal{U}(P)$ назовем *дефектом* матрицы P.

Матрицы вида

$$T^+(k,a) = \begin{pmatrix} 1 & az^k \\ 0 & 1 \end{pmatrix}, \qquad T^-(m,b) = \begin{pmatrix} 1 & 0 \\ bz^m & 1 \end{pmatrix},$$

где $a, b \in \mathbb{R}$; $k, m \in \mathbb{Z}$ будем называть *элементарными лифтингами.*

Теорема 1. *Любая полифазная матрица P с рациональными компонентами вида (4.1) может быть представлена в виде*

$$P(z) = P_0(z)T^{\sigma(1)}(k_1, a_1)T^{\sigma(2)}(k_2, a_2)\dots T^{\sigma(\mathcal{D}(P))}(k_{\mathcal{D}(P)}, a_{\mathcal{D}(P)}), \tag{4.3}$$

где $P_0(z)$ — матрица с нулевым дефектом, $k_i \in \mathbb{Z}$, $a_i \in \mathbb{R}$, $\sigma(i)$ — подходящим образом выбранный знак; $i = 1, 2, ..., \mathcal{D}(P)$.

Замечание 1. В формулировке теоремы 1 не утверждается, что все элементарные лифтинги в (4.3) невырождены, некоторые из них могут быть равны единичной матрице.

Замечание 2. Факторизация (4.3) определяется не единственным образом.

Доказательство теоремы 1. Для доказательства теоремы нам достаточно показать, что любая полифазная матрица с ненулевым дефектом P может быть представлена в виде $P(z) = P'(z)T^{\sigma}(k,a)$, где $\mathcal{D}(P^1) < \mathcal{D}(P)$.

Сначала докажем следующее вспомогательное утверждение

Лемма 1. *Если для полифазной матрицы $P(z)$ с рациональными компонентами вида (4.1) справедливо неравенство $\mathcal{L}(P) > 0$ (или $\mathcal{U}(P) > 0$), то $l(E_1)+l(O_2) = l(E_2)+l(O_1)$ (или $u(E_1) + u(O_2) = u(E_2) + u(O_1)$) и определитель матрицы*

$$A = \begin{pmatrix} L(E_1)z^{l(E_1)} & L(E_2)z^{l(E_2)} \\ L(O_1)z^{l(O_1)} & L(O_2)z^{l(O_2)} \end{pmatrix} \quad \left(\text{или } B = \begin{pmatrix} U(E_1)z^{u(E_1)} & U(E_2)z^{u(E_2)} \\ U(O_1)z^{u(O_1)} & U(O_2)z^{u(O_2)} \end{pmatrix} \right)$$

равен нулю.

Доказательство Леммы 1. В случае невыполнения равенства $l(E_1) + l(O_2) = l(E_2) + l(O_1)$ мы вошли бы в противоречие с равенством (4.2), поскольку минимальная степень в левой части равнялась бы $\min\{l(E_1O_2), l(E_2, O_1)\}$, что по условию леммы меньше минимальной степени правой части равенства (4.2).

По той же причине определитель матрицы A равен нулю, поскольку в противном случае минимальная степень полинома в левой части (4.2) была бы меньше минимальной степени правой части.

Утверждение относительно матрицы B доказывается аналогичным образом.

Вернемся к доказательству Теоремы 1.

Пусть для определенности $\mathcal{L}(P) > 0$. Тогда из Леммы 1 получим

$$l(E_1) + l(O_2) = l(E_2) + l(O_1). \tag{4.4}$$

Для определенности также будем предполагать, что

$$d(E_1) \leq d(O_1). \tag{4.5}$$

Рассмотрим два случая: 1) $\mathcal{U}(P) > 0$; 2) $\mathcal{U}(P) = 0$.

В первом случае согласно Лемме 1 имеем $u(E_1) + u(O_2) = u(E_2) + u(O_1)$. Вычитая отсюда (4.4), получим $d(E_1) + d(O_2) = d(E_2) + d(O_1)$, что в совокупности с (4.5) дает $d(E_2) \leq d(O_2)$. Таким образом, выполнив преобразование $E_1'(z) = E_1(z)$, $E_2'(z) = E_2(z)$, $O_1'(z) = O_1(z) - az^k E_1(z)$, $O_2'(z) = O_2(z) - az^k E_2(z)$, где $a = L(O_1)/L(E_1)$, $k = l(O_1) - l(E_1)$, учитывая равенство нулю определителя матрицы A, получим, что $l(O_1') = l(O_1) + 1$, $l(O_2') = l(O_2) + 1$. В то же время, справедливы неравенства $u(O_1') \leq u(O_1)$, $u(O_2') \leq u(O_2)$ Следовательно, дефект матрицы

$$P'(z) \stackrel{\text{def}}{=} P(z) \begin{pmatrix} 1 & -az^k \\ 0 & 1 \end{pmatrix} \tag{4.6}$$

строго меньше дефекта матрицы P. Что и требовалось доказать.

Пусть теперь $\mathcal{U}(P) = 0$. Покажем, что преобразование (4.6) понижает дефект матрицы P по крайней мере на единицу. Если $d(E_2) \leq d(O_2)$, то приведенные выше рассуждения остаются в силе. Рассмотрим случай $d(E_2) > d(O_2)$. Поскольку после преобразование (4.6) имеем $\mathcal{L}(P') \leq \mathcal{L}(P) - 1$, то нам остается доказать неизменность величины $\mathcal{U}(P)$. Нетрудно видеть, что $u(O_2') = u(E_2) + k$ и $u(O_1') \leq u(O_1)$. Таким образом, $u(E_2 O_1') \leq u(E_2 O_1)$ и $u(E_1 O_2') = u(E_1) + u(E_2) + k \leq u(O_1) + u(E_2) = u(E_2 O_1)$. Следовательно,

$$\begin{aligned} \mathcal{U}(P') = \max\{u(E_1 O_2'), u(O_1' E_2)\} - u(Q_1 Q_2) \leq \\ \max\{u(E_1 O_2), u(O_1 E_2)\} - u(Q_1 Q_2) = \mathcal{U}(P) = 0. \end{aligned}$$

Таким образом, $\mathcal{D}(P') \leq \mathcal{D}(P) - 1$.

В заключение доказательства Теоремы 1 отметим, что в случае, когда $d(E_1) > d(O_1)$ следует применить элементарный лифтинг, являющийся верхней треугольной матрицей.

Как будет показано в §5, реализация алгоритмов опирающаяся на факторизацию, гарантированную Теоремой 1, имеет двукратный асимптотический выигрыш в вычислительных затратах при фиксированных степенях знаменателей $h(z)$ и $\tilde{h}(z)$ и степени хотя бы одного из числителей, стремящейся к бесконечности, при условии, что обе функции $h(z)$ и $\tilde{h}(z)$ имеют общий вид, т.е. среди коэффициентов их числителей между максимальной и минимальной степенью нет нулей или их относительно мало.

Однако на практике чаще всего приходится иметь дело со всплесками, которые с точностью до сдвига аргумента являются четными или нечетными функциями. Именно такие примеры были рассмотрены в §2.

К сожалению, для последовательностей h_n и $\tilde{h}_n$, симметричных относительно полуцелых точек, нам не удалось найти лифтинг-реализацию, учитывающую симметрию. Важность этой задачи очевидна, поскольку такие последовательности определяют базисы нечетных всплесков (простейшим примером таких функций являются всплески Хаара).

Рассмотрим случай четных всплесков. О возможности учета симметрии при реализации лифтинг-схемы, когда $h(z)$ и $\tilde{h}(z)$ — лорановские полиномы есть упоминание (без доказательства) в [3]. Там же для всех примеров, соответствующих четному случаю, приведена лифтинг-реализация с учетом четности.

Проблема здесь состоит в том, что когда $h(z) = S(z)/Q(z)$, где S, Q — лорановские полиномы, для которых, $S(1/z) = S(z)$, $Q(1/z) = Q(z)$, при реализации формул (2.8) или (2.9) можно сэкономить на учете симметрии числителя, воспользовавшись формулой

$$c_k^{-1} = h_0 c_{2k}^0 + \sum_{m=1}^{M} h_m (c_{2k+m}^0 + c_{2k-m}^0),$$

где $2M = d(S)$. Асимптотически при больших m при учете симметрии количество операций равно $3/4$ числа операций, необходимых в общем случае. Естественно, хотелось бы при лифтинг-реализации дополнительно к двукратному выигрышу получить еще такой же дополнительный выигрыш в арифметических затратах.

Итак, будем считать, что обе рациональные функции $h(z)$ и $\tilde{h}(z)$ имеют последовательности лорановских коэффициентов, симметричные относительно коэффициента с нулевым индексом. Тогда без ограничения общности то же можно сказать и о полиномах E_1, E_2, O_1, O_2 в обозначениях (4.1). Тогда имеет место

Теорема 2. *Пусть полифазная матрица $P(z)$ имеет рациональные компоненты вида (4.1), порожденные последовательностями $h_n = h_{-n}$, $\tilde{h}_n = \tilde{h}_{-n}$. Тогда матрица $P(z)$ имеет четный дефект и может быть представлена в виде*

$$P(z) = P_0(z) \left(\prod_{i=1}^{\mathcal{D}(P)/2} T^{\sigma(i)}(\sigma(i)k_i, a_i) T^{\sigma(i)}((-1) \cdot \sigma(i)(k_i + 1), a_i) \right),$$

где $P_0(z)$ — матрица с нулевым дефектом, $k_i \in \mathbb{N} \cup 0$, $a_i \in \mathbb{R}$, $\sigma(i)$ — подходящим образом выбранный знак.

Доказательство теоремы 2. Покажем сначала, что дефект матрицы $P(z)$ является четным числом. В самом деле, из симметрии коэффициентов нетрудно получить, что $l(E_1) = -u(E_1)$, $l(O_1) = 1 - u(O_1)$, $l(O_2) = -u(O_2)$, $l(E_2) = -1 - u(E_2)$. Следовательно, $l(E_1 O_2) = -u(E_1 O_2)$ и $l(E_2 O_1) = -u(E_2 O_1)$, кроме того $l(Q_1 Q_2) = -u(Q_1 Q_2)$. Поэтому $\mathcal{L}(P) = \mathcal{U}(P)$, что влечет $\mathcal{D}(P) = 2\mathcal{L}(P) = 2\mathcal{U}(P)$.

Таким образом, нам остается показать, что любую матрицу $P(z)$ с ненулевым дефектом, удовлетворяющую условиям теоремы 2, можно представить в виде

$$P(z) = P_1(z) T^{\sigma}(\sigma k, a) T^{\sigma}((-1) \cdot \sigma(k+1), a),$$

где P_1 — матрица, также удовлетворяющая условиям теоремы 2.

Поскольку $\mathcal{L}(P) = \mathcal{U}(P) > 0$, то согласно Лемме 1 имеем $l(E_1)+l(O_2) = l(E_2)+l(O_1)$ и $u(E_1) + u(O_2) = u(E_2) + u(O_1)$, причем определители матриц

$$A = \begin{pmatrix} L(E_1) & L(E_2) \\ L(O_1) & L(O_2) \end{pmatrix}, \quad B = \begin{pmatrix} U(E_1) & U(E_2) \\ U(O_1) & U(O_2) \end{pmatrix}$$

равны нулю. Ввиду симметрии последовательностей h_n и $\tilde{h}_n$ имеем $A = B$. Ввиду того, что $d(E_1)$ и $d(O_2)$ — четные числа, а $d(O_1)$ и $d(E_2)$ — нечетные, имеет место либо выполнение неравенств: а) $d(E_1) > d(O_1)$, $d(E_2) > d(O_2)$, либо выполнение неравенств б) $d(E_1) < d(O_1)$, $d(E_2) < d(O_2)$.

Будем считать, что справедливы неравенства а). Тогда, выбрав $a = L(E_1)/L(O_1)$ и $m = u(E_1) - u(O_1)$, получим, что для полиномов

$$E_1'(z) = E_1(z) - a(z^m + z^{-m-1})O_1(z),$$

$$E_2'(z) = E_2(z) - a(z^m + z^{-m-1})O_2(z)$$

справедливы соотношения $E_1'(z) = E_1'(z^{-1})$, $E_2'(z) = z^{-1}E_2'(z^{-1})$, $d(E_1') = d(E_1) - 2$, $d(E_2') = d(E_2) - 2$. Таким образом, взяв в качестве $P_1(z)$ матрицу, получающуюся из $P(z)$ заменой $E_1 \to E_1'$, $E_2 \to E_2'$, имеем

$$P_1(z) = P(z) \begin{pmatrix} 1 & 0 \\ -a(z^m + z^{-m-1}) & 1 \end{pmatrix} = P(z)T^-(m, -a)T^-(-(m+1), -a) \qquad (4.7)$$

или $P(z) = P_1(z)T^-(-(m+1), a)T^-(m, a) = P_1(z)T^-(m, a)T^-(-(m+1), a)$.

Случай б) рассматривается аналогичным образом. Действуя таким же образом, как и для случая а), получим, что

$$P(z) = P_1(z)T^+(-m, a)T^+(m+1, a),$$

где $a = L(O_1)/L(E_1)$, $m = l(O_1) - l(E_1)$, а матрица $P_1(z)$ отличается от $P(z)$ заменой

$$O_1(z) \to O_1'(z) = O_1(z) - a(z^{-m} + z^{m+1})E_1(z),$$

$$O_2(z) \to O_2'(z) = O_2(z) - a(z^{-m} + z^{m+1})E_2(z).$$

Таким образом, теорема 2 доказана.

§5. Численная реализация алгоритмов субполосного кодирования.

Рассмотрим теперь реализацию алгоритмов разложения и восстановления сигналов, в основе которой лежит схема лифтинга. Сначала поясним, какое преимущество дает факторизация полифазной матрицы, гарантируемая Теоремой 1. Посчитаем сколько арифметических операций (сложения и умножения) требуется для прямой реализации формулы. Для этого воздействие полифазной матрицы $P(z)$ на сигнал в обозначениях §4 удобно представить в виде

$$P(z)\begin{pmatrix} c_e^0(z^{-1}) \\ c_o^0(z^{-1}) \end{pmatrix} = \begin{pmatrix} \dfrac{1}{Q_1(z)} & 0 \\ 0 & \dfrac{1}{Q_2(z)} \end{pmatrix} \begin{pmatrix} E_1(z) & O_1(z) \\ E_2(z) & O_2(z) \end{pmatrix} \begin{pmatrix} c_e^0(z^{-1}) \\ c_o^0(z^{-1}) \end{pmatrix}. \tag{5.1}$$

Нетрудно видеть, что первое матричное умножение на сигнальный вектор в (5.1) может быть реализовано при помощи в среднем за $d(E_1) + d(O_1) + d(E_2) + d(O_2) + 3$ арифметических операций на одну сигнальную точку. Действительно, для получения одного коэффициента в лорановском разложении первой компоненты необходимо $2d(E_1) + 2d(O_1) + 3$ операций, а для получения одного лорановского коэффициента второй компоненты требуется $2d(E_2) + 2d(O_2) + 3$ операций.

Остановимся подробнее на реализации второго сомножителя при помощи приема, который в радиотехнике называется цифровым рекурсивным фильтром. Поскольку для $|a| < 1$ имеем

$$\sum_{n\in\mathbb{Z}} \beta'_n z^n \stackrel{\text{def}}{=} \frac{1}{1-az} \sum_{n\in\mathbb{Z}} \beta_n z^n = \sum_{n\in\mathbb{Z}} \beta_n z^n \sum_{n=0}^{+\infty} a^n z^n = \sum_{n\in\mathbb{Z}} \left(\sum_{k=0}^{+\infty} \beta_{n-k} a^k \right) z^n, \tag{5.2}$$

то

$$\beta'_n = \sum_{k=0}^{+\infty} \beta_{n-k} a^k = \sum_{k=0}^{+\infty} \beta_{n-1-k} a^{k+1} + \beta_n = (\sum_{k=0}^{+\infty} \beta_{n-1-k} a^k) a + \beta_n = \beta'_{n-1} a + \beta_n.$$

Полученная рекуррентная формула называется *элементарным рекурсивным фильтром.*

Если найдется такое N, что $\beta_n = 0$ при $n < N$, поэтому $\beta'_n = 0$, и рекурсивный фильтр позволяет точную реализацию формулы (5.2). Разумеется, в этом случае внутренняя сумма в (5.2) конечная и потому допускает точную реализацию, однако арифметические затраты на ее реализацию значительно выше затрат, которые требует рекурсивный фильтр.

В случае, когда существуют ненулевые b_n с как угодно большими отрицательными номерами, рекурсивный фильтр может быть реализован, начиная с некоторого номера N, поэтому в его реализации неизбежно присутствует погрешность, которая убывает с каждым шагом со скоростью геометрической прогрессии со знаменателем a.

Аналогичным образом оператор, задаваемый формулой

$$\sum_{n\in\mathbb{Z}} \beta'_n z^n \stackrel{\text{def}}{=} \frac{1}{1-az^{-1}} \sum_{n\in\mathbb{Z}} \beta_n z^n,$$

может быть реализован в виде элементарного рекурсивного фильтра $\beta'_n = \beta'_{n+1} a + \beta_n$.

Лорановские полиномы Q_1 и Q_2 в (5.1) могут быть представлены в виде

$$Q_1(z) = Az^{l_1} \prod_{i=1}^{d(Q_1)} \left(1 - \left(\frac{a_i}{z}\right)^{\operatorname{sign}(1-|a_i|)}\right), \quad Q_2(z) = Bz^{l_2} \prod_{i=1}^{d(Q_2)} \left(1 - \left(\frac{b_i}{z}\right)^{\operatorname{sign}(1-|b_i|)}\right),$$

где $\{a_i\}$, $\{b_i\}$ — корни полиномов Q_1 и Q_2, а множители Az^{l_1} и Bz^{l_2} можно без ограничения общности можно считать равными единице. Заметим, что корни полиномов Q_1 и Q_2 не могут равняться по модулю единице, т.к. в этом случае, с одной стороны, у нас получатся неустойчивые фильтры, а с другой стороны, такие функции h и $\tilde{h}$ заведомо не будут соответствовать никакой биортогональной паре КМА. Следовательно, умножение на диагональную матрицу в формуле (5.1) можно реализовать при помощи $d(Q_1) + d(Q_2)$ элементарных рекурсивных фильтров. Поэтому средние арифметические затраты всех рекурсивных фильтров равны $d(Q_1) + d(Q_2)$. Таким образом, полные вычислительные затраты на прямую реализацию формулы (5.1) составляют $d(E_1)+d(O_1)+d(E_2)+d(O_2)+d(Q_1)+d(Q_2)+3$ операций на одну сигнальную точку.

Вычислительные затраты вычисления на лифтинг-реализацию формулы (5.1) приводятся в следующем утверждении.

Теорема 3. *Вычислительные затраты на лифтинг-реализацию формулы (5.1) не превышают $\mathcal{D}(P) + 3d(Q_1) + 3d(Q_2) + 3$ операций на точку.*

Замечание. Если $h(z)$ и $\tilde{h}(z)$ — лорановские полиномы, то в качестве частного случая Теоремы 3 получаем, что затраты равны $\mathcal{D}(P) + 3$, т.е. результат, полученный в работах [3] и [6].

Доказательство теоремы 3. Согласно теореме 1 воздействие полифазной матрицы в формуле (2.9) на сигнальный вектор может быть представлено в виде произведения $\mathcal{D}(P)$ элементарных лифтингов и матрицы P_0 нулевого дефекта. Воздействие каждого лифтинга $T^\sigma(k, a)$ на сигнал состоит в умножении половины сигнальных отсчетов на число a, сдвига координат этих же отсчетов на k позиций и добавления результата ко второй половине отсчетов. Таким образом, стоимость одного лифтинга — равна одной операции на сигнальный отсчет (сдвиг координат, как обычно, не входит в вычислительные затраты. Таким образом, получим, что $\mathcal{D}(P)$ элементарных лифтингов требуют $\mathcal{D}(P)$ арифметических операций.

Обозначив через E_1', E_2', O_1', O_2' числители соответствующих элементов матрицы P_0, получим, что для реализации умножения на матрицу $P_0(z)$ требуется не более $d(E_1') + d(E_2') + d(O_1') + d(O_2') + d(Q_1) + d(Q_2) + 3$ операций на точку. Но поскольку матрица P_0 имеет нулевой дефект, то $d(E_1') + d(O_2') \leq d(Q_1) + d(Q_2)$ и $d(E_2') + d(O_1') \leq d(Q_1) + d(Q_2)$, поэтому общие затраты оцениваются сверху величиной $3d(Q_1) + 3d(Q_2) + 3$, что и требовалось доказать.

Ясно, что из Теоремы 2 для лифтинт-схемы следует возможность дополнительного сокращения на четверть числа арифметических операций при наличии симметрии у последовательностей h_n и $\tilde{h}_n$. Такая возможность появляется из-за то, что для реализации каждого элементарного лифтинга требуется одна операция на сигнальную точку, а как видно из (4.7), композиция двух элементарных лифтингов $T^+(m,a)$ и $T^+(-(m+1),a)$ может быть реализована за 1.5 операций на точку (3 операции с половиной точек).

§6. Приложения к сжатию изображений

Начнем с мотивации применения рассмотренных типов всплесков для обработки сигналов и изображений. Рассмотрим простейшее цифровое устройство, для реализации которого могут быть применены базисы всплесков — эквалайзер. Его предназначение — частотная коррекция сигнала. Принцип действия состоит в следующем. Исходный сигнал $f(t)$ с финитным спектром (преобразованием Фурье) представляется в виде суммы сигналов f_0, f_1,f_2, … ,f_n, имеющихся непересекающиеся спектры расположенные соответственно на промежутках $[0,\omega_1]$, $[\omega_1,\omega_2[$, $[\omega_2,\omega_3[$, … $[\omega_n,\omega_{n+1}]$, объединение которых покрывает весь спектр сигнала f. Далее полученные сигналы складываются с требуемыми весами α_k, что дает на выходе устройства сигнал $\tilde{f} = \sum \alpha_k f_k$. Эта задача может быть решена при помощи Быстрого Преобразования Фурье. Вычислительная сложность такой процедуры равна $O(N \log N)$, где N — число отсчетов сигнала. Ясно что вычислительная сложность алгоритма может быть сокращена за счет разбиения сигнала на короткие участки, однако в этом случае могут возникнуть проблемы, связанные со склейкой отдельных кусков после обработки.

Обычно предполагается, что точки ω_k образуют логарифмическую шкалу, т. е. $\omega_{k+1}/\omega_k = c$, $k = 1,2,\ldots,n$. В случае, когда $c = 2$, удобно воспользоваться представлением сигнала f в виде суммы его проекций на некоторый набор пространств $\tilde{V}^m$, $\tilde{W}^m$, $\tilde{W}^{m+1}$, … , $\tilde{W}^{m+n}$. Где операторы проектирования на $\tilde{W}^{m+k}$ должны достаточно хорошо аппроксимировать полосовой фильтр, пропускающий без искажения гармоники в интервале $[\omega_k,\omega_{k+1}[$ и подавляющего все остальные гармоники. Таким образом, задача сводится к выбору 2π-периодической функции m_0, приближающей функцию

$$\chi(x) = \begin{cases} 1, & |x| < \pi/2, \\ 0, & \pi/2 \le |x| \le \pi. \end{cases}$$

В работе И.Добеши [12] были найдены в каком-то смысле оптимальные решения этой задачи, когда $m_0 = \tilde{m}_0$ — полином, фиксированной степени. Оптимальность выбора заключалась в том, что из полиномиальных решений было выбрано то, которое имеет максимально плоский график вблизи нуля, что было обеспечено максимально возможной кратностью нуля производной полинома m_0 в точках 0 и $\pm\pi$. Такой выбор обеспечивает большое абсолютное значение производных (так называемую, крутитизну фильтра) в точках $\pm\pi/2$.

Ясно, что рациональные функции могут приближать функцию χ со значительно более высокой скоростью. Максимально плоскую характеристику вблизи нуля среди рациональных функций равной степени числителя дают построенные (см., напр., [1]) на основе фильтров Буттерворта

$$h(z)h(z^{-1}) = \frac{2(1+z^{-1})^N(1+z)^N}{(z^{-1}+2+z)^N + (-z^{-1}+2-z)^N}$$

функции $m_0(\omega) = \tilde{m}_0(\omega) = h(e^{i\omega})/\sqrt{2}$.

Если взять для сравнения базисы Добеши и Буттерворта с функциями m_0 с одинаковой равной N кратностью нуля в точках $\pm\pi$, то соответствующие им алгоритмы разложения и восстановления сигналов будут иметь примерно одинаковую вычислительную сложность. В то же время отношение величины производной функции m_0 для всплесков Буттерворта в точках $\pm\pi$ при $N = 2$ к соответствующей величине производной для всплесков Добеши равно 4/3, а при $N = 3$ равно 8/5. При росте N различие производных становится еще более заметным. Если фильтр Буттерворта при $N = 2$ обеспечивает ту же производную (даже немного большую), что и фильтр Добеши при $N = 3$, то для того, чтобы получить ту же производную, которую дает фильтр Буттерворта при $N = 3$, требуется взять фильтр Добеши порядка 7.

К сожалению, подобного рода простая аргументация не может быть серьезным доводом в пользу применения каких бы то ни было базисов для задач сжатия изображений. До настоящего времени нет сколь-нибудь обоснованных соображений о количественных характеристиках базисов, которые влияют на степень сжатия. Одним из таких свойств, которое приводят в большинстве публикаций, является кратность нулей функций $m(\omega)$ и $\tilde{m}(\omega)$ в точках $\pm\pi$. Принято считать, что увеличение кратности нулей ведет к увеличению возможного коэффициента сжатия изображений. Проведенное нами исследование, результаты которого приводятся ниже, в определенной мере опровергают это утверждение.

Нами было проведено численное моделирование, позволяющее сравнить базисы, описанные в §3, и лучшие биортогональные пары базисов с компактными носителями на предмет их эффективности для сжатия изображений.

Все исследования проводились для 8-битных (256 оттенков серого) изображений. Изображение как функция двух переменных периодизировалось при помощи зеркального продолжения за свои края, а затем представлялось в виде коэффициентов разложения по базисам всплесков, являющихся тензорным произведением (см., напр, [4, Глава 10]) одномерных базисов.

Полученные коэффициенты кодировались алгоритмом SPIHT [13] без дополнительного арифметического кодирования. Для восстановления изображения брался начальный участок кода, требуемой длины. Восстановление проводилось по отрезкам кода, равным от 1/10 до 1/150 длины файла, необходимого для записи изображения в стандартном формате, требующем 1 байт на пиксель.

Критерием качества восстановленного изображения было взято измеряемое в децибеллах пиковое отношение сигнал/шум (PSNR), определяемое формулой

$$PSNR = 10 \lg \frac{255^2}{\sum_{i=1}^{N} \sum_{j=1}^{M} |x_{ij} - \tilde{x}_{ij}|^2},$$

где x_{ij}, $\tilde{x}_{ij}$ — целые числа в пределах от 0 до 255, определяющие величину интенсивности пикселя с координатами (i, j) соответственно для оригинального и для восстановленного изображения.

Среди исследованных нами базисов, ассоциированных с рекурсивными фильтрами лучшие результаты для сжатия дала биортогональная пара

$$h(z) = \frac{1+z}{\sqrt{2}}, \qquad g(z) = \frac{z^{-3} - 3z^{-2} + 3z^{-1} - 1}{4\sqrt{2}},$$

$$\tilde{h}(z) = \frac{(1+\alpha)^2}{4\sqrt{2}} \frac{z^{-1} + 3 + 3z + z^2}{(1+\alpha z^2)(1+\alpha z^{-2})},$$

$$\tilde{g}(z) = \frac{(1+\alpha)^2}{\sqrt{2}} \frac{-z^{-2} + z^{-1}}{(1+\alpha z^2)(1+\alpha z^{-2})},$$

где $\alpha = 3 - 2\sqrt{2}$. Такая пара позволяет достичь примерно тех же коэффициентов сжатия, что и лучшие из известных базисов с конечными масками (например, 9/7-всплески из [2]). Различие величины PSNR для приведенной пары и 9/7-базисов для подавляющего числа изображений находится в пределах ± 0.3 dB. Однако наш алгоритм имеет значительное преимущество в вычислительных затратах.

Стандартный пирамидальный алгоритм разложения одномерного сигнала по 9/7-базисам имеет сложность 14 операций (см. [3]) сложения и умножения на сигнальный отсчет. Столько же операций требуется на восстановление сигнала по известным коэффициентам. Мы покажем, что наш алгоритм может быть реализован за 7 операций на один сигнальный отсчет при разложении одномерного сигнала по базису и за 11 операций на отсчет при восстановлении.

Стандартная реализация разложения изображения по двумерным базисам всплесков, заключается в последовательной реализации одномерных алгоритмов сначала по строкам изображения, а затем — по столбцам. Таким образом, соотношения вычислительных затрат для разных базисов сохраняются при переходе к двумерным алгоритмам.

Полифазные матрицы алгоритмов разложения и восстановления имеют вид

$$P(z) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & z \\ -\frac{3z^{-1}-1}{4} & \frac{z^{-1}+3}{4} \end{pmatrix}, \tilde{P}(z) = \frac{(1+\alpha)^2}{\sqrt{2}(1+\alpha z^{-1})(1+\alpha z)} \begin{pmatrix} \frac{3+z}{4} & \frac{1+3z}{4} \\ -z^{-1} & 1 \end{pmatrix},$$

что не позволяет напрямую воспользоваться результатами §4 для реализации алгоритмов разложения и восстановления. Заметим, что причина этого не в том, что матрицы $P(z)$ и $\tilde{P}(z)$ ввиду переброски полинома из знаменателя второй строчки матрицы P в знаменатель второй строчки матрицы $\tilde{P}$ имеют отличные от единицы определители, а в том, что еще до перемещения полинома исходные матрицы имели нулевые дефекты, что делает бессмысленным применение Теоремы 1. Таким образом, вероятно, при реализации одномерного алгоритма декомпозиции сто́ит использовать прямую форму алгоритма разложения (без перехода к полифазной матрице), которая с учетом симметрии фильтров требует 7 операций на сигнальный отсчет.

Что касается алгоритма восстановления, то несмотря на приведенный выше аргумент, нам удалось найти лифтинг-факторизацию

$$\tilde{P}(z) = \begin{pmatrix} 1 & \frac{1+3z}{4} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \frac{(1+\alpha)^2}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{(1+\alpha z^{-1})(1+\alpha z)} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -z^{-1} & 1 \end{pmatrix}.$$

Отсюда видно, вычислительные затраты такого алгоритма равны 11 операции на сигнальную точку. Заметим, что прямая форма реализации алгоритма восстановления требует 15 операций на точку.

В заключение мы отметим, что рассмотренные в этом параграфе фильтры $h(z)$ и $\tilde{h}(z)$ действительно порождают биортогональную пару базисов пространства $\mathbb{L}^2(\mathbb{R})$. К сожалению, нам неизвестен сколь-нибудь общий результат, из которого непосредственно следовал бы данный факт. Один из первых результатов о достаточном условии существования ортогонального базиса всплесков, соответствующего некоторому фильтру $h(z) = \tilde{h}(z)$, удовлетворяющего уравнению (2.7), принадлежит С.Малла [14].

Хотя мы имеем дело с биортогональными парами h и $\tilde{h}$ и не можем воспользоваться упомянутым достаточным условием, однако ввиду положительности в нашем случае функций m_0 и $\tilde{m}_0$ на интервале $(-\pi, \pi)$ доказательство Малла проходит без существенных изменений и для нашей пары рекурсивных фильтров.

Литература

1. Herley C., Vetterli M., *Wavelets and recursive filter banks IEEE Trans. on Signal Proc.* **41** (1993), no. 8, 2536 – 2556.
2. Cohen A., Daubechies I., and Feauveau J.-C., *Biorthogonal bases of compactly supported wavelets Commun. Pure Appl. Math.* **45** (1992), 485 – 560.
3. Daubechies I., Sweldens W., *Factoring wavelet transforms into lifting steps*, J. Fourier Anal. Appl. **4** (1998), no. 3, 247 – 269.
4. Daubechies I., *Ten Lectures on Wavelets*, SIAM, Philadelphia, 1992.
5. Cohen A., Daubechies I., *Nonseparable biorthogonal bases*, Rev. Mat. Iberoam. **9** (1993), no. 1, 51 – 137.
6. Borac S., Seiler R., *Loop group factorization of biorthogonal wavelet bases* (1997), Report.
7. Strang G., Nguyen T.Q., *Wavelets and Filter Banks*, Wellesley-Cambridge Press, 1996.
8. Петухов А.П., *Периодические всплески*, Математический сборник **188** (1997), no. 10, 69 – 94.
9. Петухов А.П., *Кратномасштабный анализ и всплеск-разложения пространств периодических распределений*, Докл. РАН **356** (1997), no. 2, 303 – 306.
10. Петухов А.П., *Периодические дискретные всплески*, Алгебра и анализ **8** (1996), no. 3, 151 – 183.
11. Skopina M.A., *Multiresolution analysis of periodic functions*, East Journal on Approximation **3** (1997), no. 2, 203 – 224.
12. Daubechies I., *Orthogonal bases of compactly supported wavelets*, Comm. Pure and Appl. Math. **41** (1988), 909 – 996.
13. Pearlman W., Said A., *A new fast and efficient image codec based on set partitioning in hierarchical trees*, IEEE Trans on Circuits and Systems for Video Tech. **6** (1996).
14. Mallat S., *Multiresolution approximations and wavelet orthonormal bases of* $\mathbb{L}^2(\mathbb{R})$, Trans. of the Amer. Math. Soc. **315** (1989), no. 1, 69 – 87.
15. Meyer F.G., Averbuch A., Strömberg J.-O., *Fast adaptive wavelet packet image compression* (1998), Report.