

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

3/2008


НАУЧНО - ТЕХНИЧЕСКИЙ ЖУРНАЛ

- МНОГОСКОРОСТНАЯ
ОБРАБОТКА СИГНАЛОВ

- ЦИФРОВАЯ ФИЛЬТРАЦИЯ

- ВЕЙВЛЕТ-ТЕХНОЛОГИИ

- РАДИОВИДЕНИЕ



“Глубоко уверен в том, что издание журнала является очень своевременным и актуальным, имеющим важное значение для ученых и специалистов в области теории и техники цифровой обработки сигналов...” - академик РАН Котельников В.А. (журнал “Цифровая обработка сигналов”, №1, 1999г.)

100 лет

академику В. А. Котельникову

ISSN 1684-2634



9 771684 263005 >

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

Научно-технический журнал
№ 3/2008

Издается с 1999 года
Выходит четыре раза в год

ГЛАВНЫЙ РЕДАКТОР
Ю.Б. ЗУБАРЕВ

ЗАМЕСТИТЕЛИ ГЛАВНОГО РЕДАКТОРА:
В.В. ВИТЯЗЕВ, В.П. ДВОРКОВИЧ

РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

Б.А. Бабаян, Ю.И. Борисов,
С.А. Грибачев, Г.В. Зайцев, Р.В. Зубарев,
А.П. Кирпичников, М.И. Кривошеев,
Н.А. Кузнецов, М.С. Куприянов, А.А. Ланнэ,
В.Г. Мистюков, С.Л. Мищенко,
А.А.Петровский, Ю.Н. Прохоров,
А.Н. Соловьев, Ю.Г. Сосулин,
В.В. Шахильдян, Ю.С. Шинаков

Адрес редакции:

107031 Москва, Рождественка, 6/9/20, стр. 1
Тел.: (495) 621-71-08, (495) 621-16-39
Факс: (495) 621-16-39
E-mail: nto.popov@mtu-net.ru
tor@rgta.ryazan.ru
http://www.dsps.ru

Для писем:

129090 Москва, а/я 48.

Издатель:

Российское научно-техническое общество радиотехники,
электроники и связи им. А.С. Попова
Ответственный редактор: П.И. Зудков
Компьютерная верстка: В.В. Андряхина
Дизайн: И.Е. Артюхина

Подписной индекс по каталогу
ОАО «Роспечать» – **82185**

Подписано в печать 30.09.08
Формат 60x90/8.

Гарнитура «Arial». Печать ризографическая.
Бумага офсетная. Печ.л. 6,5. Тираж 700 экз.

Заказ № 5385. Отпечатано в
ООО НПЦ «Информационные технологии»
Рязань, ул. Островского, д. 21/1
тел.: (4912) 98-69-84

Издание зарегистрировано в Министерстве
Российской Федерации по делам печати,
телерадиовещания и средств
массовых коммуникаций.
Свидетельство о регистрации ПИ № 77-1488
от 14.01.2000

УЧРЕДИТЕЛИ:

ЗАО «Инструментальные системы»
ФГУП «НИИ радио»

В НОМЕРЕ:

- Витязев В.В., Витязев С.В., Зайцев А.А.
**Многоскоростная обработка сигналов: ретроспектива
и современное состояние (часть 2)** 2
- Никитин Д.А., Ханов В.Х.
**Синтез рекурсивных цифровых фильтров по
импульсной характеристике, определяемой
элементарной математической функцией** 10
- Горшков А.П., Грызлова Т.П.
**Система формирования образа и классификации
временных рядов по характерным
последовательностям** 15
- Орешко Н.И., Князева Т.Н.
**Вейвлет-технология анализа и очистки сигналов от
шума** 21
- Макаров В.А., Павлов А.Н., Тупицын А.Н.
**Сортировка нейронных спайков на основе
параметрического вейвлет-анализа с адаптивной
фильтрацией** 26
- Прохоренков А.М., Качала Н.М.
**Цифровая фильтрация сигналов в промышленных
системах управления** 32
- Клионский Д.М., Орешко Н.И., Геппенер В.В.
**Декомпозиция на эмпирические моды и ее
использование при анализе дробного броуновского
движения** 37
- Вишнеевский К.П., Чижиков В.И., Барышенский Д.С.,
Жучкова В.В.
Новый симметричный алгоритм шифрования 46
- Балусов И.Л., Приоров А.Л., Хрящев В.В.
**Частотные свойства двумерных рекурсивных
цифровых фильтров второго порядка с
симметричными коэффициентами** 49
- Заикин С.А., Мойбенко В.И., Парфиров В.А.
**Алгоритм синтезирования апертуры для получения
радиолокационного изображения при маловысотном
полете** 56
- Григоренко Д.В., Ручкин В.Н.
**Компенсация межсимвольных помех посредством
обратной связи по решению для сигналов с
решетчатой кодированной модуляцией** 60
- Фиксированная или плавающая точка? Теперь это не
вопрос** 63

Подписной индекс по каталогу
ОАО «Роспечать» – **82185**

МНОГОСКОРОСТНАЯ ОБРАБОТКА СИГНАЛОВ: РЕТРОСПЕКТИВА И СОВРЕМЕННОЕ СОСТОЯНИЕ (ЧАСТЬ 2)

Витязев В.В., Витязев С.В., Зайцев А.А.

Введение

Многоскоростная обработка сигналов (МОС), как эффективный инструмент построения цифровых информационно-телекоммуникационных систем, работающих в реальном времени, стала предметом интенсивных исследований с начала 70-х прошлого столетия. В работе [1] представлен тематический обзор эволюции теории и техники МОС с позиции научного вклада, который внесли в эту область исследования российских ученых и специалистов в период с начала 70-х до конца 80-х, т.е. на этапе, когда закладывались теоретические основы многоскоростной фильтрации и построения систем цифровой частотной селекции сигналов на основе эффектов прореживания по времени и по частоте [2]. В предлагаемой вашему вниманию второй части теоретического обзора [1] мы продолжаем систематизированное краткое описание основных результатов исследований в области МОС в период с конца 80-х прошлого столетия до наших дней.

Оптимальное проектирование многоскоростных систем на цифровых сигнальных процессорах

В представленном ранее тематическом обзоре методов и алгоритмов многоскоростной частотной селекции сигналов [1] был сделан вывод о высокой вычислительной эффективности многоступенчатой реализации цифровых полосовых фильтров и их наборов. Вместе с тем стала очевидной необходимость разработки методики оптимального проектирования подобных структур цифровых фильтров. Предстояло ответить на вопрос: сколько ступеней преобразования и с какими коэффициентами децимации (интерполяции) необходимо выбрать, чтобы достигнуть наилучшего результата с позиции вычислительной эффективности? Уже в своих первых работах, посвященных многоступенчатой реализации цифровых узкополосных фильтров с заданными параметрами частотной избирательности, М.Крошье и Л.Рабинер [3,4] в качестве критерия оптимизации использовали приведенное число операций умножения или требуемое число ячеек памяти данных. В то время такой подход к задаче оптимизации был вполне оправдан, так как именно операция умножения считалась определяющей общие вычислительные затраты, а память данных была жестко ограниченной небогатыми внутренними ресурсами микропроцессоров и микроЭВМ.

С появлением цифровых сигнальных процессоров (ЦСП) ситуация с оценкой вычислительных затрат резко изменилась. Операция умножения на ЦСП выполняется за один-два командных такта, так же как и операции

Рассматривается дальнейшая эволюция теории и техники многоскоростной обработки сигналов в период с конца 80-х прошедшего столетия до наших дней. Изложение материала ведется с позиции вклада, который внесли в эту область исследований работы российских ученых и специалистов.

сложения (вычитания) и обращения к памяти данных. Поэтому стало необходимым и одновременно возможным перейти к оценке вычислительной эффективности алгоритма через оценку временных затрат на его реализацию в числе командных тактов, что обеспечивает достаточно полное и точное описание основного критерия оптимизации – минимизации приведенных к периоду дискретизации вычислительных затрат [2,5]. Но требуемые вычислительные затраты на реализацию того или иного алгоритма обработки являются далеко не единственным критерием оптимизации. По-прежнему важно учитывать внутренние ресурсы памяти данных и коэффициентов ЦСП, разрядность представления данных, особенно для 16-разрядных процессоров с фиксированной запятой. Могут выдвигаться и другие ограничения с позиции задачи многокритериальной оптимизации: на длительность переходного процесса, помехоустойчивость алгоритма, уровень собственных шумов. Но, главное, - это достижение максимальной точности воспроизведения желаемых частотных характеристик в условиях ограниченных ресурсов ЦСП. Поэтому многокритериальная оптимизация структуры и параметров цифровых систем и устройств ЦОС, учитывающая все ограничения, накладываемые конкретными условиями реализации, является важным этапом их проектирования [2,6].

В работах [2,5] методика оптимального проектирования на ЦСП рассматривается на примере многоступенчатой реализации узкополосного низкочастотного (НЧ) фильтра в классе КИХ-цепей. На рис. 1 показан общий вид $(m+1)$ -ступенчатой структуры фильтра, включающей m ступеней децимации (интерполяции) на основе фильтров N_i -го порядка с функциями передачи $H_i(j\omega)$, $i=1, m$, обеспечивающих понижение и соответственно повышение частоты дискретизации в $v = \prod_{i=1}^m V_i$ раз, основного фильтра N_0 -го порядка с функцией передачи $H_0(j\omega)$, работающего на предельно низкой частоте дискретизации. Основным фильтром формирует заданную прямоуглольность АЧХ проектируемого узкополосного КИХ-фильтра. Предполагается, что многоступенчатая структура проектируемого фильтра эквивалентна по свойствам частотной избирательности некоторому НЧ фильтру N_f -го порядка с функцией передачи $H(j\omega)$.

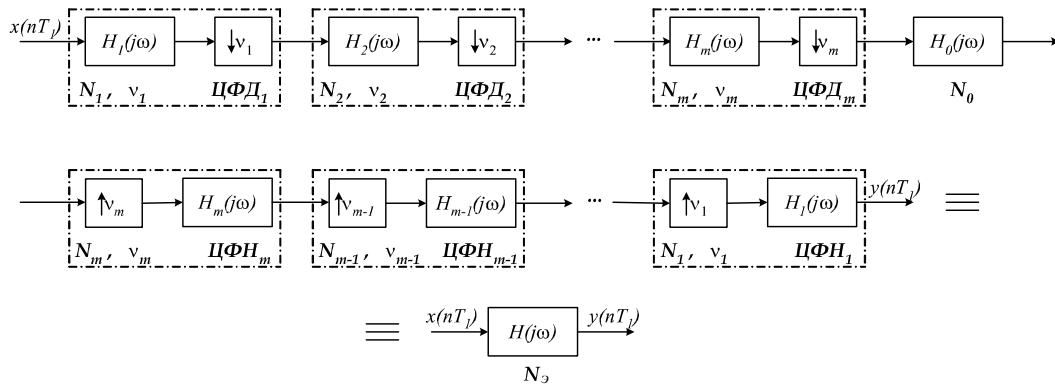


Рисунок 1. Общий вид $(m+1)$ -ступенчатой структуры НЧ фильтра N_y -го порядка

Прямую задачу оптимального проектирования многоступенчатой структуры цифрового фильтра сформулируем следующим образом: на множестве многоступенчатых структур $G_F^{A_l} \in G_F^A$, $l = 0, 1, \dots, m$, класса КИХ-фильтров G_F^D , реализуемых с использованием вторичной дискретизации (децимации), найти подкласс $G_F^{Dk} \in G_F^D$ и составной оператор $F^k \in G_F^{Dk}$ с заданной структурой L^{Dk} , такие, что

$$\begin{cases} \Phi(G_F^{Dk}, F^k) = \max_{\omega=-\pi, \pi} p(\omega) |H_B(j\omega, F^l) - H(j\omega)| \rightarrow \min_{G_F^{Dk}, F^l}; \\ V(G_F^{Dk}, F^k) \leq T_{\text{don}}; & S(G_F^{Dk}, F^k) \leq S_{\text{don}}; \\ Q(G_F^{Dk}, F^k) \leq Q_{\text{don}}; & D(G_F^{Dk}, F^k) \leq D_{\text{don}}, \end{cases} \quad (1)$$

где $\Phi(G_F^{Dk}, F^k)$ — целевая функция, выбираемая из критерия минимизации максимального отклонения строго воспроизводимой в подклассе G_F^{Dk} функции передачи $H_B(j\omega, F^k)$ от желаемой $H(j\omega)$; $p(\omega)$ — весовая функция чебышевского приближения; $V(G_F^{Dk}, F^k)$ — временные (вычислительные) затраты на программную реализацию оператора $F^k \in G_F^{Dk}$, приведенные к периоду дискретизации T_1 входного сигнала; $S(G_F^{Dk}, F^k)$ и $Q(G_F^{Dk}, F^k)$ — емкость внутрикристалльной памяти данных и емкость памяти программ, требуемые на программную реализацию оператора $F^k \in G_F^{Dk}$; $D(G_F^{Dk}, F^k)$ — дисперсия собственного шума на выходе цифрового устройства, реализующего оператор $F^k \in G_F^{Dk}$; $T_{\text{don}}, S_{\text{don}}, Q_{\text{don}}, D_{\text{don}}$ — совокупность ограничивающих факторов, определяемых конкретными условиями программно-аппаратной реализации оператора $F^k \in G_F^{Dk}$. Допустимое время обработки T_{don} равно периоду дискретизации T_1 входного сигнала; ограничения на предельную емкость внутрикристалльной памяти данных S_{don} и внешней (внекристалльной памяти программ) Q_{don} определяются семейством ЦСП и способом организации памяти цифрового устройства, а допустимый уровень собственных шумов D_{don} зависит от требований, накладываемых на точность воспроизведения желаемых характеристик фильтра.

Задачу оптимального проектирования в форме (1), как было показано в [2], удобно свести к последовательному решению двух задач: задачи выбора подкласса $G_F^{Dk} \in G_F^D$ и оператора $F_{\text{opt}}^k \in G_F^{Dk}$, максимизирующего значение порядка N_y , эквивалентного по свойствам частотной избирательности НЧ фильтра с функцией

передачи $H_B(j\omega, F_{\text{opt}}^k)$, строго воспроизводимой в классе КИХ-цепей N_y -го порядка, при заданных ограничениях на программно-аппаратную реализацию цифрового устройства:

$$\begin{cases} \Phi_l(G_F^{Dk}, F_{\text{opt}}^k) = N(G_F^{Dl}, F^l) \rightarrow \max_{G_F^{Dl}, F^l}, & l = 0, 1, 2, \dots, m; \\ V(G_F^{Dk}, F_{\text{opt}}^k) \leq T_{\text{don}}; & S(G_F^{Dk}, F_{\text{opt}}^k) \leq S_{\text{don}}; \\ Q(G_F^{Dk}, F_{\text{opt}}^k) \leq Q_{\text{don}}; & D(G_F^{Dk}, F_{\text{opt}}^k) \leq D_{\text{don}}, \end{cases} \quad (2)$$

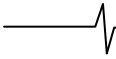
и задачи чебышевского приближения, решаемой при заданной структуре L^{Dk} и оптимальных значениях параметров $N_{i \text{opt}}, v_{i \text{opt}}, i = 0, k, (k+1)$ -ступенчатой структуры оператора $F_{\text{opt}}^k \in G_F^{Dk}$:

$$\Phi_2(\bar{F}_{\text{opt}}^k) = \max_{\omega=-\pi, \pi} p(\omega) |H_B(j\omega, \bar{F}^k) - H(j\omega)| \rightarrow \min_{\bar{F}^k}, \quad (3)$$

где $\bar{F}_{\text{opt}}^k = \{\bar{F}_0, \bar{F}_1, \dots, \bar{F}_k\}$ — представление оператора F^k в подклассе G_F^{Dk} , раскрывающееся на множестве ступеней преобразования при заданных структуре L^{Dk} и значениях параметров N_i, v_i через совокупность представлений F_i его компонент $F_i, i = 0, k$.

Решение задачи оптимального проектирования в форме (2) сводится фактически к структурному синтезу в классе операторов $F^l \in G_F^{Dl}, l = 0, m$, предполагающему определение оптимального числа ступеней k_{opt} и параметрическую оптимизацию $(k_{\text{opt}}+1)$ -ступенчатой структуры с определением оптимальных значений порядков фильтров-дециматоров (интерполяторов) $N_{i \text{opt}}$ и оптимальных значений коэффициентов прореживания $v_{i \text{opt}}$ для каждой i -й ступени преобразования, $i = 0, k_{\text{opt}}$. По окончании структурного синтеза и параметрической оптимизации на каждой i -й ступени решается задача чебышевского приближения в форме (3) с использованием известных методов и средств машинного синтеза.

Обратную задачу оптимального проектирования многоступенчатой структуры цифрового фильтра сформулируем в тех же обозначениях следующим образом: на множестве многоступенчатых структур $G_F^{Dl} \in G_F^D, l = 0, 1, \dots, m$, класса КИХ-фильтров G_F^D , реализуемых с использованием вторичной дискретизации, найти подкласс $G_F^{Dk} \in G_F^D$ и составной оператор $F^k \in G_F^{Dk}$ вида $F^k = L^{Dk}\{F_0, F_1, \dots, F_k\}$ с заданной структурой L^{Dk} , такие, что



$$\left\{ \begin{array}{l} \Phi^*(G_F^{lk}, F^k) = V(G_F^{lk}, F^l) \rightarrow \min_{G_F^{lk}, F^l}; \\ \max_{\omega \in -\pi, \pi} p(\omega) |H_B(j\omega, F^k) - H(j\omega)| \leq \varepsilon_{don}; \\ S(G_F^{lk}, F^k) \leq S_{don}; \quad Q(G_F^{lk}, F^k) \leq Q_{don}; \quad D(G_F^{lk}, F^k) \leq D_{don} \end{array} \right. \quad (4)$$

где $\Phi^*(G_F^{lk}, F^k)$ — целевая функция, выбираемая из критерия минимизации приведенных временных затрат на программно-аппаратную реализацию оператора F^k в подклассе G_F^{lk} .

По аналогии с методикой решения прямой задачи оптимального проектирования многоступенчатой структуры фильтра обратную задачу в форме (4) можно свести к последовательному решению двух задач: обратной аппроксимационной задачи чебышевского приближения в классе КИХ-цепей G_F , устанавливающей значение минимального порядка $N_{\mathcal{Y}}$ эквивалентного НЧ фильтра с функцией передачи $H_B(j\omega)$, строго воспроизводимой в классе КИХ-цепей N -го порядка, и задачи выбора подкласса $G_F^{lk} \in G_F^{lk} \subset G_F$ операторов $F_k \in G_F^{lk}$, минимизирующих объем приведенных временных затрат на программно-аппаратную реализацию многоступенчатой структуры фильтра $N_{\mathcal{Z}}$ -го порядка при известных ограничениях

Постановка задачи оптимального проектирования в форме (1) и (4) носит общий характер и требует раскрытия математического описания целевой функции и области ограничений для каждой $(l+1)$ -ступенчатой структуры фильтра, реализуемого в конкретных условиях (при всех $l = 1, m$). Поэтому решению задачи оптимального проектирования предшествует этап формализации входящих в (1,2) и (4) выражений общего вида. Функциональную зависимость между порядком $N_{\mathcal{Z}}$ эквивалентного НЧ фильтра и параметрами $N_i, v_i, i = 0, l, (l+1)$ -ступенчатой структуры, представленной на рис. 1, запишем в виде

$$N_{\mathcal{Z}} = N_0 \nu = N_0 \prod_{i=1}^l v_i, \quad (5)$$

где N_0 — порядок основного фильтра, формирующего АЧХ проектируемого фильтра с заданным показателем прямоугольности на частоте дискретизации, в ν раз меньшей частоты дискретизации входного сигнала $x(nT_1)$; ν — общий коэффициент прореживания многоступенчатой структуры фильтра-дециматора (интерполятора).

Следовательно, при увеличении коэффициента прореживания ν в допустимых пределах обратно пропорционально уменьшается порядок N_0 основного фильтра. В то же время согласно (5) порядок $N_{\mathcal{Z}}$ эквивалентного НЧ фильтра остается неизменным. Таким образом, потенциальная возможность увеличения порядка $N_{\mathcal{Z}}$ эквивалентного НЧ фильтра непосредственно связана с возможностью увеличения порядка основного фильтра N_0 : чем большими резервами по времени обработки и емкости памяти располагает основной фильтр, тем выше порядок проектируемого фильтра. Предоставление соответствующих резервов под основной фильтр обеспечивается прежде всего эффективной организацией многоступенчатой структуры фильтра-дециматора (интерполятора) в условиях реальных ограничений на программно-аппаратную реализацию проектируемого фильтра. Достижение максимально допустимого значения коэффициента прореживания ν при минимальных вычислительных и аппаратных

затратах предоставляет дополнительные резервы под реализацию основного фильтра и, как следствие, дает возможность увеличения его порядка N_0 .

Пусть даны оценки приведенных временных затрат и затрат памяти на программно-аппаратную реализацию i -й степени фильтра-дециматора в виде функций вида

$$\begin{aligned} V_{T_i}(G_F^{ll}, N_i, v_i); \quad S(G_F^{ll}, N_i, v_i); \\ Q(G_F^{ll}, N_i, v_i), \quad i = \overline{1, l}. \end{aligned} \quad (6)$$

Аналогично оценки приведенных временных затрат и затрат памяти на программно-аппаратную реализацию i -й степени фильтра-интерполятора представим в виде системы функций:

$$\begin{aligned} V_{T_i}(G_F^{ll}, N_i, v_i); \quad S_l(G_F^{ll}, N_i, v_i); \\ Q(G_F^{ll}, N_i, v_i), \quad i = \overline{1, l}, \end{aligned} \quad (7)$$

а оценки затрат на реализацию основного фильтра — системой функций вида

$$V_{T_{l+1}}(G_F^{l\Phi_l}, N_0); \quad S_0(G_F^{l\Phi_l}, N_0); \quad Q_0(G_F^{l\Phi_l}, N_0). \quad (8)$$

Используя введенные ранее описания (6) — (8), объединенные оценки приведенных к периоду T_1 временных затрат и затрат памяти на реализацию $(l+1)$ -ступенчатой структуры фильтра запишем в виде

$$\left\{ \begin{array}{l} V_{T_1}(G_F^{ll}, F_l) = \sum_{i=1}^l \frac{1}{\prod_{j=0}^i v_j} [V_{T_i}(G_F^{ll}, N_i, v_i) + V_{T_i}(G_F^{ll}, N_i, v_i)] + \frac{1}{\nu} V_{T_{l+1}}(G_F^{l\Phi_l}, N_0); \\ S(G_F^{ll}, F_l) = \sum_{i=1}^l [S_i(G_F^{ll}, N_i, v_i) + S_i(G_F^{ll}, N_i, v_i)] + S_0(G_F^{l\Phi_l}, N_0); \\ Q(G_F^{ll}, F_l) = \sum_{i=1}^l [Q(G_F^{ll}, N_i, v_i) + Q(G_F^{ll}, N_i, v_i)] + Q_0(G_F^{l\Phi_l}, N_0). \end{array} \right. \quad (9)$$

Формализованное описание функции $D(G_F^{ll}, F^l)$, определяющей зависимость дисперсии собственного шума $(l+1)$ -ступенчатой структуры фильтра от оптимизируемых параметров $N_i, v_i, i = 0, l$, является самостоятельной задачей, решение которой на множестве подклассов $G_F^{ll}, l = 1, m$, должно предшествовать решению общей задачи оптимального проектирования $(m+1)$ -ступенчатой структуры фильтра. Вместе с тем, как показывает практика построения подобных систем, поиск эффективной многоступенчатой структуры и расчет оптимальных значений ее параметров достаточно провести, по крайней мере на первом этапе проектирования, без учета ограничений на допустимый уровень собственных шумов. Выбор числа ступеней преобразования и расчет оптимальных значений ее параметров по критерию, фактически обеспечивающему минимизацию суммарного порядка всех преобразующих фильтров, одновременно приводит и к уменьшению влияния собственных шумов: чем меньше объем вычислений, тем меньше уровень выходного шума. Поэтому в дальнейшем решение задачи оптимального проектирования в формах (1) и (4) рассматривается без учета ограничений на допустимый уровень собственного шума на выходе фильтра. Оценка влияния собственных шумов производится на втором этапе — этапе статистического анализа оптимизированной $(k_{opt}+1)$ -ступенчатой структуры с использованием общепринятой шумовой модели [7].

Общая методика решения прямой и обратной задач оптимального проектирования состоит в пошаговой параметрической оптимизации двух-, трех- и в общем случае $(l+1)$ -ступенчатой структур до достижения такого оптимального значения числа ступеней k_{opt} , дальнейшее увеличение которого приводит к уменьшению эффективности многоступенчатой структуры фильтра в смысле принятого критерия качества (1) или (4). Поскольку на практике число таких ступеней не превышает трех — четырех ступеней, то достижение оптимальных значений параметров структуры фильтра обеспечивается уже на первых этапах расчета.

В последующих работах [8,9] методика оптимального проектирования на сигнальных процессорах нашла применение при синтезе цифровых полосовых фильтров с высокой прямоуглобностью АЧХ и многоступенчатой реализации узкополосных фильтров на основе эффектов прореживания по частоте с использованием квадратурной модуляции импульсной характеристики гребенчатого и сглаживающего фильтров. Естественным шагом на пути дальнейшей разработки методики оптимального проектирования на ЦСП стали математическая формализация и решение задачи оптимизации многоступенчатых пирамидальных структур набора цифровых полосовых фильтров с использованием эффектов прореживания по времени и по частоте [10]. Эффективность применения методики оптимального проектирования многоступенчатых фильтров-дециматоров на сигнальных процессорах семейства ADSP-218х иллюстрирует работа [11].

МОС в задачах адаптивной фильтрации

Одним из быстро развивающихся направлений в области ЦОС в конце 70-х – начале 80-х стала адаптивная обработка сигналов [6]. Основным свойством адаптивных систем является изменяющееся во времени функционирование с саморегулированием в соответствии с заданным критерием оптимизации. Обычно адаптивные системы (фильтры) разделяют на два класса: системы без обратной связи и системы с обратной связью (с самообучением) [12,13].

Первый класс адаптивных систем реализует процесс саморегулирования путем измерения только характеристик входного сигнала или окружающей среды, а второй – учитывает влияние системы на выходной сигнал, как правило, используя обучающий сигнал.

Примером построения адаптивной системы без обратной связи являются поисковые системы ЦОС, предназначенные для исследования частотно-временной структуры сигнала, поиска скрытых периодичностей, разделе-

ния и распознавания источников излучения в широкой полосе частот панорамного обзора. Характерная область применения - радиотехнические системы, в том числе устройства радиолокации, радионавигации и радиопеленгации. Впервые об эффективной многопроцессорной реализации подобных систем ставится вопрос в работе [14]. Оригинальным решением поставленной задачи является разработка многоступенчатой адаптивной поисковой системы на базе типового однопроцессорного модуля БПФ с элементами самонастройки.

Общая структура приемника панорамного обзора может строиться по схеме M -компонентной системы анализа-синтеза сигналов на базе одноступенчатой реализации набора полосовых фильтров. При этом повышение разрешающей способности панорамного приемника достигается путем увеличения числа частотных каналов M . Однако такой подход приводит к пропорциональному увеличению вычислительных затрат и емкости памяти данных. В то же время в реальных системах «сверхразрешение» по частоте необходимо только в тех участках области частот, где присутствуют близко расположенные узкополосные составляющие спектра входного сигнала с медленно меняющимися огибающими. Кроме того, в других частотных поддиапазонах может наблюдаться более динамичное изменение спектральных составляющих, сохранение характера которых потребует расширения полосы частот анализа. С этой целью в систему вносятся элементы адаптации к частотно-временной структуре сигнала.

Адаптивная система частотной селекции сигналов строится на основе двух принципов синтеза самонастраиваемой структуры набора цифровых фильтров-дециматоров (ФДМ) с растущей памятью: принципа многоступенчатости структуры и принципа избирательности каналов [2]. В соответствии с первым принципом набор ФДМ реализуется по многоступенчатой пирамидальной структуре [1], что дает возможность предварительного анализа частотно-временной структуры сигнала на каждом шаге преобразования. В соответствии со вторым принципом между ступенями преобразования включаются решающие устройства (РУ), выделяющие только информативные частотные каналы, выходные сигналы которых после децимации подвергаются дальнейшей обработке с повышенным разрешением по частоте.

На рис. 2 представлена структурная схема адаптивного приемника панорамного обзора, реализованного с использованием указанных выше принципов построения поисковой системы с частотно-временным разделением каналов [15].

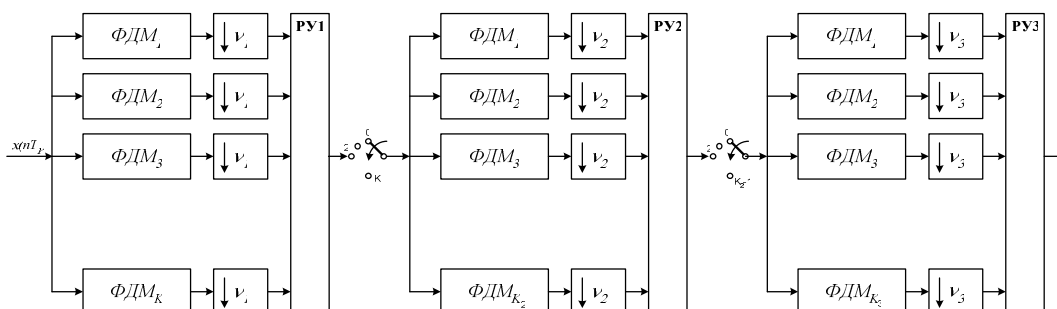


Рис. 2. Адаптивный приемник панорамного обзора

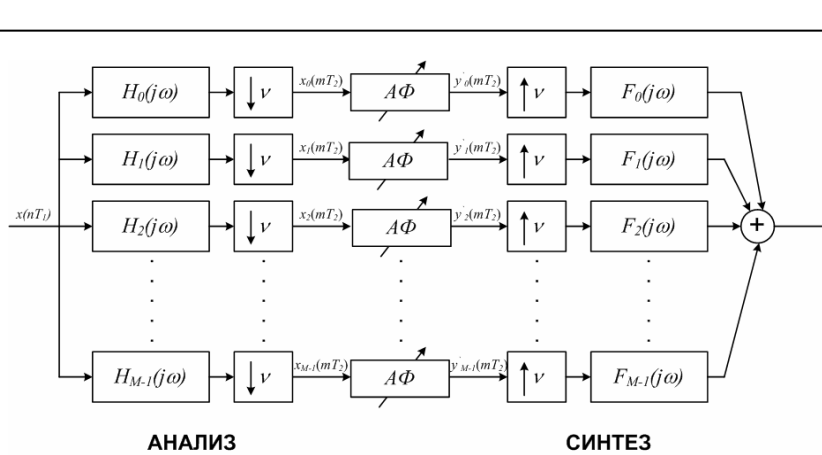


Рис.3. Общая структурная схема адаптивного КИХ-фильтра с использованием системы анализа-синтеза сигналов

На первой ступени преобразования набор из K_1 фильтров-демодуляторов (ФДМ) разбивает весь диапазон рабочих частот на K_1 поддиапазонов с понижением частоты дискретизации выходного сигнала в V_1 раз. По данным на выходе набора ФДМ решающее устройство PY_1 оценивает «информативность» каждого из частотных поддиапазонов и настраивает набор из K_2 ФДМ последующей ступени преобразования на выделение информативных частотных составляющих с одновременным увеличением разрешающей способности в V_1 раз. В зависимости от спектральной структуры входного сигнала все K_2 ФДМ второй ступени могут быть сконцентрированы в $R \leq K_1$ частотных поддиапазонах. На второй ступени преобразования с помощью набора из K_2 ФДМ в каждом поддиапазоне выделяется K_2 субполос с понижением частоты дискретизации в V_2 раз. Решающее устройство PY_2 анализирует по принятому признаку «информативность» каждого из K_2 частотных каналов и настраивает набор из K_3 ФДМ последующей ступени преобразования. При необходимости процесс преобразования по аналогичной структуре может быть продолжен до достижения необходимой частотной избирательности в каждом из поддиапазонов. Отметим, что структура адаптивной системы в результате окончательной настройки приобретает пирамидальный характер с сохранением только тех «ветвей» дерева, которые несут полезную информацию. При изменении частотно-временной структуры принимаемого сигнала автоматически перестраивается и структура адаптивного приемника. Кроме того, исключение из последующего анализа неинформативных поддиапазонов позволяет значительно уменьшить общие вычислительные затраты и емкость памяти данных.

Адаптивные фильтры с предварительным обучением – относительно новый класс систем с саморегулированием, использующих, как правило, градиентные методы поиска оптимума рабочей функции, в качестве которой выступает средний квадрат отклонения выхода фильтра от обучающего сигнала, нашли широкое применение в задачах идентификации динамических систем и их обратного моделирования, эхо-компенсации, линейного предсказания и других [12,13]. В силу того, что на этапе обучения подобные системы используют обратную связь, одной из проблем их реализации стала проблема устойчивости, которая особенно остро проявляется при синтезе в классе БИХ-цепей. Поэтому в подавляющем большинстве случаев на практике при-

меняют адаптивные КИХ-фильтры. Однако для динамических систем и процессов с ярко выраженными частотноизбирательными свойствами (наличие узкополосных составляющих, спектральных переходов, глубоких замираний и т.п.) порядок адаптивного КИХ-фильтра может достигать тысячи и десятки тысяч единиц. Это затрудняет процесс адаптации и приводит к значительному увеличению вычислительных затрат.

Решающим шагом к принципиально новому способу синтеза адаптивного КИХ-фильтра высокого порядка стало эффективное сочетание многоскоростной и адаптивной фильтрации. В частности, еще в работе [2] было отмечено, что система анализа-синтеза сигналов с применением набора ЦФДМ-ЦФМ (демодуляторов-модуляторов) и промежуточным понижением частоты дискретизации является мощным инструментом эффективной реализации систем ЦОС во многих приложениях.

В соответствии с общей идеей построения подобных систем применительно к адаптивной фильтрации структурная схема обработки сигналов принимает вид, показанный на рис.3.

Сигнал $x(nT_1)$ на этапе анализа разбивается на M частотных каналов с частично перекрывающимися (в переходной зоне АЧХ) характеристиками. После выделения комплексной огибающей на выходе каждого частотного канала и децимации ее отсчетов с коэффициентом V , пропорциональным числу каналов M , выполняется адаптивная фильтрация по всем M частотным каналам. Предполагается, что такой же процедуре частотного разделения подвергается и обучающий сигнал. Таким образом, задача синтеза одного адаптивного КИХ-фильтра высокого порядка сводится к задаче синтеза M адаптивных КИХ-фильтров относительно малого порядка (порядок канальных АФ уменьшается пропорционально числу каналов M). Кроме того, каждый адаптивный фильтр $A\Phi_k$, $k = \overline{1, M}$, работает на пониженной в V раз частоте дискретизации, что в V раз уменьшает требование к скорости обработки. На этапе синтеза все частотные составляющие с выходов M адаптивных фильтров используются для восстановления единого выходного сигнала $y(nT_1)$ с помощью набора фильтров-интерполяторов ΦI_k , $k = \overline{1, M}$, повышающих частоту дискретизации в V раз. Предполагается, что банк (набор) фильтров-дециматоров и фильтров-интерполяторов предварительно синтезируется таким образом, чтобы вносить минимальные дополнительные

искажения принятой процедуры обработки. В частности, для этой цели могут использоваться банки фильтров «с полным восстановлением» [16,17].

Одним из первых приложений адаптивной системы анализа-синтеза сигналов стали обработка и сжатие сейсмических сигналов и геофизических данных [18,19,20]. Далее последовала серия работ, связанных с адаптивным подавлением мультigarмонических помех в устройствах ближней радиолокации и радиометрии [21,22]. Но наивысший всплеск работ был посвящен в эти годы решению задач прямого и обратного моделирования динамических систем на основе многоскоростной адаптивной фильтрации применительно к проблемам эхо-компенсации и борьбы с межсимвольной интерференцией, а также субполосному кодированию в системах телекоммуникаций [23-31].

составляющих (рис. 5). Главным инструментом для такого разбиения являются банки фильтров анализа $H_i(z), i = 0, M - 1$ и банки фильтров синтеза $F_i(z), i = 0, M - 1$. Поскольку главная цель кодирования источника сигнала – максимальная степень сжатия информации, то желательно, чтобы банк фильтров анализа обеспечивал максимальную децимацию, а банк фильтров синтеза – полное восстановление.

Применение субполосного кодирования приводит к хорошим результатам, если известны априорные сведения о характере распределения энергии исходного сигнала. При обработке речи, аудиосигналов и изображений такие сведения имеются благодаря большому опыту кодирования подобных сигналов.

К числу относительно новых приложений многоскоростной адаптивной фильтрации следует отнести ра-

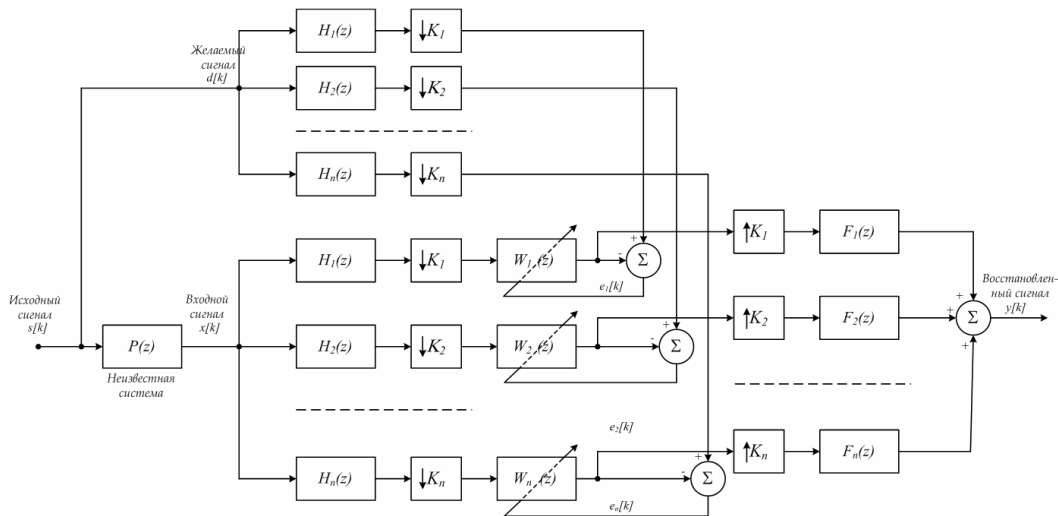


Рис. 4. Функциональная схема устройства обратного адаптивного моделирования с разбиением на субполосы

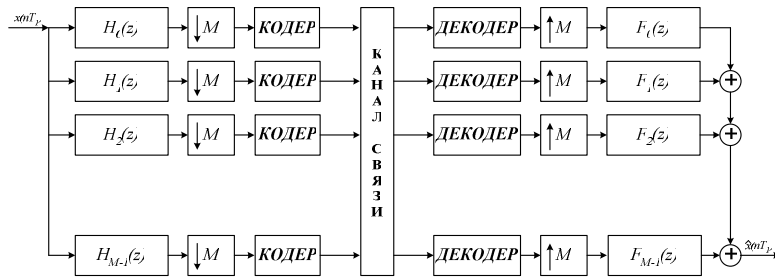
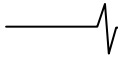


Рис. 5. Схема субполосного кодирования сигналов

На рис. 4 представлен пример реализации субполосного адаптивного фильтра, решающего задачу обратного моделирования. Сигнал $x(nT_1)$, искаженный неизвестной системой (каналом связи), разбивается на M субполос банком фильтров анализа $H_i(z), i = 0, M - 1$. Затем в каждом субполосном канале производится прореживание отсчетов выходного сигнала с коэффициентами децимации $k_i, i = 1, M$. Такой же банк фильтров анализа используется для разделения на субполосы желаемого (обучающего) сигнала $d(nT_1)$. Далее производится независимая настройка каждым адаптивным фильтром $W_i(z), i = 1, M$ и восстановление отфильтрованного сигнала $y(n)$ банком фильтров синтеза $F_i(z), i = 1, M$.

Принцип субполосного кодирования заключается в разделе исходного сигнала $x(nT_1)$ на M частотных

диовидение [32]. Необходимость использования многоскоростной обработки в задачах формирования радиолокационного изображения (РЛИ) и селекции наземных движущихся целей (НДЦ) обусловлена тем, что ширина спектра частот траекторного сигнала при переднем и передне-боковом секторных обзорах земной поверхности может быть значительно уже по отношению к частоте повторения зондирующих импульсов, а, следовательно, целесообразны предварительная узкополосная фильтрация и децимация отсчетов траекторного сигнала [33]. Кроме того, как показано в [33], и в режиме панорамного кадрового обзора в значительной степени эффективна предварительная обработка траекторного сигнала с помощью набора цифровых фильтров-дециматоров и последующим формированием РЛИ на пониженной частоте дискретизации. Поскольку траек-



торный сигнал подвержен нестабильностям и изменению ширины спектра в зависимости от азимутального направления диаграммы направленности антенны (ДНА), система формирования РЛИ и селекции НДЦ должна быть адаптивной, перестраиваемой по центральной частоте и ширине полосы пропускания фильтров-дециматоров. Результаты разработки и исследования эффективности многоскоростной адаптивной фильтрации в задачах радиовидения можно найти в работах [33-38].

Описанные выше системы адаптивной многоскоростной обработки на основе децимации и интерполяции сигналов не лишены одного общего недостатка: роста «шума децимации» с увеличением коэффициента децимации V и с уменьшением требований высокой частотной избирательности набора фильтров-дециматоров за границами зоны их полосы пропускания. Поэтому, как и для обычных систем частотной селекции [2], так и для систем адаптивной многоскоростной обработки существует альтернативный подход, использующий прореживание по частоте с помощью набора адаптивных гребенчатых фильтров (АГФ) [39,40].

В основе предлагаемого альтернативного подхода лежит идея представления импульсной характеристики адаптивного узкополосного фильтра в виде линейной свертки импульсной характеристики цифрового гребенчатого фильтра (ЦГФ), имеющего периодическую частотную характеристику, и цифрового сглаживающего фильтра (ЦСФ) с относительно короткой реакцией, выделяющего из множества спектральных составляющих на выходе ЦГФ полосу частот обучающего сигнала. С целью увеличения соотношения сигнал/шум на входе АГФ целесообразно изменить порядок включения фильтров, поставив первым полосовый сглаживающий фильтр, выделяющий по априорной информации (или полученной по результатам предварительного спектрального анализа) основную полосу частот обучающего сигнала и подавляющего все мешающие частотные составляющие (аналогично фильтру-дециматору на входе последующего адаптивного фильтра).

На рис. 6 представлена схема двухэтапного прямого моделирования динамической системы с использованием децимации импульсной характеристики адаптивного фильтра. Блок оценивания настраивает ЦСФ на заданную полосу частот обучения, а АГФ выполняет основную функцию – адаптацию к обучающему сигналу в выделенной полосе частот.

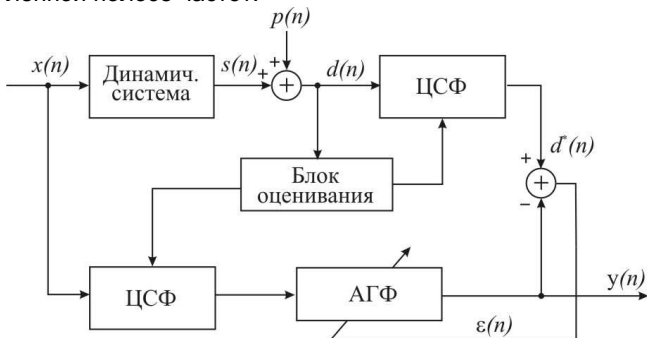


Рис. 6. Схема двухэтапного прямого моделирования динамической системы с использованием адаптивного гребенчатого фильтра.

Преимущества предлагаемого метода очевидны и связаны с уменьшением размерности вектора оптимизируемых весовых коэффициентов в число раз, определяемое степенью узкополосности обучающего сигнала по отношению к полосе рабочих частот.

Заключение

Подводя окончательные итоги представленного тематического обзора, можно утверждать, что многоскоростная обработка сигналов и ее сочетание с адаптивной фильтрацией являются эффективным инструментом решения задач цифровой обработки сигналов в широком спектре приложений, прежде всего в радиотехнических и телекоммуникационных системах. Вместе с тем остаются нерешенными еще многие проблемы, возникающие в связи с новыми условиями технической реализации, новыми требованиями к построению систем ЦОС и новой областью их применения. А значит – впереди новые исследования и новые результаты, подтверждающие тезис [6]: «В основе эффективной реализации систем ЦОС, как с позиции достижимых показателей качества, так и с позиции минимума аппаратно-программных затрат, лежит сочетание концепций многоскоростной, оптимальной и адаптивной обработки цифровых сигналов».

Литература

1. Витязев В.В., Витязев С.В., Зайцев А.А. Многоскоростная обработка сигналов: ретроспектива и современное состояние // Цифровая обработка сигналов, 2008. № 1. -С.12-21.
2. Витязев В.В. Цифровая частотная селекция сигналов. М.: Радио и связь, 1993. -240 с.
3. Rabiner L.R., Crochiere R.E. A novel implementation for narrow-band FIR digital filters // IEEE Trans. Acoust., Speech and Signal Processing. V. ASSP-23. Oct., 1975. -P. 457-464.
4. Crochiere R.E., Rabiner L.R. Optimum FIR digital filter implementations for decimation, interpolation and narrow band filtering // IEEE Trans. Acoust., Speech and Signal Processing. V. ASSP-23. Oct., 1975. -P. 444-456.
5. Витязев В. В. Оптимальное проектирование многоступенчатых структура цифровых фильтров на процессорах обработки сигналов // Электросвязь. 1992. № 4. -С. 23—27.
6. Зубарев Ю.Б., Витязев В.В., Дворкович В.П. Цифровая обработка сигналов – информатика реального времени // Цифровая обработка сигналов. 1999. № 1. -С. 5-17.
7. Рабинер Л., Гоулд Б. Теория и применение цифровой обработки сигналов. М.: Мир, 1978. -848 с.
8. Витязев В.В., Горюнов Ю.Н. Оптимальное проектирование цифровых полосовых фильтров с высокой прямоугольностью // Электросвязь. 1995. № 4. –С. 30-32.
9. Витязев В.В., Морозов Е.А. Оптимальное проектирование цифровых полосовых фильтров на процессорах обработки сигналов // Электросвязь. 1995. № 12. –С. 29-31.
10. Витязев В.В., Бодров К.А., Сорокин Д.В. Оптимальный синтез многоскоростных систем фильтрации на цифровых сигнальных процессорах // Цифровая обработка сигналов и ее применение-DSPA'1999: Труды 1-й Международной науч.-техн. конф. М., 1999. Т. 1. -С. 50-54.
11. Витязев В.В., Зайцев А.А. Оптимальное проектирование многоступенчатых структур фильтров-дециматоров на сигнальных процессорах // Цифровая обработка сигналов.

2001. № 2. -С. 2-9.
12. Уидроу Б., Стирнз С. Адаптивная обработка сигналов. – М.: Радио и связь, 1989. – 440 с.
 13. Адаптивные фильтры \ Под ред. К.Ф.Н. Коузена и П.М. Гранта. – М.: Мир, 1988. – 392 с.
 14. Витязев В.В., Морозов Е.А., Широков В.А. Многопроцессорная поисковая система цифровой обработки сигналов \ Методы и микроселектронные средства цифрового преобразования и обработки сигналов. Труды 1-й Международной науч.-техн. конф., Рига, 1989. –С. 27-29.
 15. Витязев В.В., Бодров К.А., Иванов С.И. Адаптивная многоскоростная фильтрация узкополосных процессов \ Цифровая обработка сигналов и ее применение-DSPA'1999: Тр. 1-й Международной науч.-техн. конф. М., 1999. Т. 1. -С.150-160.
 16. Зайцев А.А. Методы построения банков цифровых фильтров: тематический обзор \ Цифровая обработка сигналов. 2003. № 1. -С. 2-10.
 17. Сорокин Д.В. \М-полосовые фильтры: расчет и применение в системах анализа-синтеза цифровых сигналов с полным восстановлением \ Труды 3-й Международной науч.-техн. конф. «Цифровая обработка сигналов и ее применение-DSPA'2000», Москва, 2000, Т.2, -С. 218-222.
 18. Витязев В.В., Зайцев А.А. Оптимальное проектирование многоступенчатых структур фильтров-дециматоров на сигнальных процессорах \ Цифровая обработка сигналов. №2. 2001. -С.2-9.
 19. Витязев В.В., Зайцев А.А. Применение банков цифровых фильтров с полным восстановлением в целях сжатия сейсмических сигналов \ Труды 4-й Международной науч.-техн. конф. «Перспективные технологии в средствах передачи информации – ПТСПИ'2001, Владимир-Суздаль.
 20. Витязев В.В., Зайцев А.А. Перспективы использования цифровых банков фильтров в целях сжатия сейсмических сигналов \ Труды 4-й Международной науч.-техн. конф. «Цифровая обработка сигналов и ее применение-DSPA'2002», Москва, Т.1, -С.160-162.
 21. Витязев В.В., Иванов С.В. Метод адаптивного подавления аддитивных помех в цифровых системах радиометрии \ Труды 4-й Международной науч.-техн. конф. «Цифровая обработка сигналов и ее применение-DSPA'2002», Москва, Т.1, -С.206-209.
 22. Витязев В.В., Иванов С.В., Езерский В.В. Метод адаптивной компенсации мультигармонической помехи в устройствах радиометрии \ Цифровая обработка сигналов. №2. 2002. -С. 27-31.
 23. Витязев В.В., Линович А.Ю. Субполосные адаптивные системы на основе банков фильтров с немаксимальной децимацией в задаче обратного моделирования \ Труды 6-й Международной науч.-техн. конф. «Цифровая обработка сигналов и ее применение-DSPA'2004», Москва, Т.1, -С. 19-22.
 24. Линович А.Ю., Витязев В.В. Субполосная адаптивная фильтрация в задачах обратного моделирования \ Цифровая обработка сигналов. 2004. № 1. -С. 41-48.
 25. Гусинская Е.И., Зайцев А.А. Оптимизация банка фильтров в задачах субполосного кодирования \ Цифровая обработка сигналов. 2004. № 3. -С. 18-28.
 26. Линович А.Ю. Методы многоскоростной обработки сигналов в задачах обратного моделирования динамических систем \ Цифровая обработка сигналов и ее применение-DSPA'2005: Тр. 7-й Международной науч.-техн. конф. Москва, 2005. Т. 1. -С. 50-54.
 27. Кузнецов Е.П., Витязев В.В. Цифровая обработка сигналов в задачах эхо-компенсации (часть1) \ Цифровая обработка сигналов. 2006. № 3. -С. 8-19.
 28. Кузнецов Е.П., Витязев В.В. Цифровая обработка сигналов в задачах эхо-компенсации (часть2) \ Цифровая обработка сигналов. 2006. № 4. -С. 20-28.
 29. Линович А.Ю. Многоканальный эквалайзер на основе субполосной адаптивной фильтрации \ Труды 9-й Международной науч.-техн. конф. «Цифровая обработка сигналов и ее применение-DSPA'2007», Москва, 2007. Т.1, -С. 124-127.
 30. Кузнецов Е.П. Новый метод эффективной реализации субполосного адаптивного эхо-компенсатора \ Труды 10-й Международной науч.-техн. конф. «Цифровая обработка сигналов и ее применение-DSPA'2008», Москва, 2008. Т.1, -С. 269-272.
 31. Кузнецов Е.П. Методы и алгоритмы адаптивной эхо-компенсации: сравнительный анализ эффективности применения цифровой обработки сигналов \ Цифровая обработка сигналов. 2007. №2.С. 26-34.
 32. Кондратенков Г.С., Фролов А.Ю. Радиовидение. Радиолокационные системы дистанционного зондирования Земли. Учебное пособие \ Под ре. Г.С. Кондратенкова. – М.: Радиотехника, 2005. – 368 с.
 33. Витязев В.В., Колодько Г.Н., Витязев С.В. Способы и алгоритмы формирования радиолокационного изображения в режиме доплеровского обужения луча \ Цифровая обработка сигналов. 2006. № 3. -С. 31-41.
 34. Витязев В.В., Колодько Г.Н., Воронков Д.В. Формирование радиолокационного изображения в режиме фокусируемого синтезирования апертуры ДНА \ Цифровая обработка сигналов. 2006. № 4. -С. 34-40.
 35. Витязев С.В., Колодько Г.Н. Моделирование и исследование эффективности формирования радиолокационного изображения в режиме доплеровского обужения луча \ Труды 9-й Международной науч.-техн. конф. «Цифровая обработка сигналов и ее применение-DSPA'2007», г. Москва, март 2007.Т.1, –С. 254-257.
 36. Витязев В.В., Колодько Г.Н., Витязев С.В. Селекция наземных движущихся целей на основе многоскоростной адаптивной обработки траекторного сигнала \ Цифровая обработка сигналов. 2007. № 1. -С. 41-50.
 37. Витязев С.В., Якунин С.А. Применение многоскоростной обработки сигналов для построения доплеровского фильтра обнаружения движущихся объектов в задачах радиовидения \ Труды 10-й Международной науч.-техн. конф. «Цифровая обработка сигналов и ее применение-DSPA'2008», Москва, 2008, Т.1, -С. 156-158.
 38. Витязев В.В., Андреев Н.А. Оптимальное проектирование на сигнальных процессорах многоступенчатой структуры адаптивного узкополосного фильтра-дециматора предварительной обработки траекторного сигнала \ Труды 10-й Международной науч.-техн. конф. «Цифровая обработка сигналов и ее применение-DSPA'2008», Москва, Т.1, -С.156-158.
 39. Витязев В.В., Витязев С.В. Методы синтеза узкополосного адаптивного КИХ-фильтра на основе многоскоростной обработки \ Цифровая обработка сигналов. №4. 2007. -С.15-20.
 40. Витязев С.В. Адаптивная цифровая фильтрация узкополосных процессов \ Труды 10-й Международной науч.-техн. конф. «Цифровая обработка сигналов и ее применение-DSPA'2008», Москва, 2008, Т.1, -С. 159-161.

СИНТЕЗ РЕКУРСИВНЫХ ЦИФРОВЫХ ФИЛЬТРОВ ПО ИМПУЛЬСНОЙ ХАРАКТЕРИСТИКЕ, ОПРЕДЕЛЯЕМОЙ ЭЛЕМЕНТАРНОЙ МАТЕМАТИЧЕСКОЙ ФУНКЦИЕЙ

Никитин Д.А., Ханов В.Х.

Введение

В современном арсенале методов синтеза рекурсивных цифровых фильтров мало методов на основе расчета и аппроксимации его импульсной характеристики. Практически — только метод Прони, который хоть и достаточно проработан, но, тем не менее, обладает рядом недостатков. Например, для его применения необходимо заранее знать минимальный достаточный порядок фильтра. Отсутствие разнообразия при выборе таких методов обусловлено прежде всего их невысокой востребованностью, так как при синтезе цифрового фильтра обычно задаются другими исходными данными. Тем не менее, в некоторых случаях [1, 2] может понадобиться поиск коэффициентов фильтра именно исходя из некоторого количества начальных отсчетов его импульсной характеристики. В этой статье предлагается такой метод, который к тому же позволяет совместить отыскание коэффициентов цифрового рекурсивного фильтра с определением минимально необходимого их количества.

Алгоритм синтеза цифрового БИХ фильтра

Пусть имеется последовательность $Y = \{y_k\}$ длины L , элементы которой являются значениями неизвестной элементарной математической функции одного переменного $f(x)$ при $x = 0, 1, 2, \dots, L - 1$. То есть $y_0 = f(0)$, $y_1 = f(1)$, $y_{L-1} = f(L-1)$. Необходимо найти цифровой рекурсивный фильтр, у которого первые L отсчетов импульсной характеристики в точности совпадают с элементами Y .

Нахождение цифрового фильтра (ЦФ) означает отыскание всех его коэффициентов a_i и b_i , которые однозначно определяют процесс фильтрации (а значит и импульсную характеристику) в соответствии с рекурсивным выражением:

$$y(n) = \sum_{k=0}^N b_k x(n-k) - \sum_{k=1}^M a_k y(n-k), \quad (1)$$

где $x(n)$ — значение на входе фильтра на n -ом такте, $y(n)$ — значение на выходе фильтра на n -ом такте, a_k и b_k — действительные коэффициенты, M и N определяют количество коэффициентов a_k и b_k соответственно.

Импульсной характеристикой называется реакция дискретной системы (в нашем случае это фильтр) на единичный импульс при нулевых начальных условиях. Используя выражение (1), можно построить систему линейных уравнений, определяющую взаимосвязь коэффициентов фильтра с его импульсной характеристикой.

Предлагается алгоритм синтеза цифрового БИХ-фильтра по его импульсной характеристике, которая определяется элементарной математической функцией или линейной комбинацией таких функций.

$$\left\{ \begin{array}{l} b_0 = y_0 \\ b_1 - y_0 a_1 = y_1 \\ b_2 - y_1 a_1 - y_0 a_2 = y_2 \\ \dots \\ b_N - (a_1 y_{N-1} + a_2 y_{N-2} + \dots + a_{M-1} y_0) = y_N \\ -(a_1 y_{N-2} + a_2 y_{N-3} + \dots + a_{M-1} y_1 + a_M y_0) = y_{N+1} \\ \dots \\ -(a_1 y_{L-2} + a_2 y_{L-3} + \dots + a_M y_{L-N-1}) = y_{L-1} \end{array} \right. \quad (2)$$

Здесь неизвестными являются коэффициенты a_k и b_k , а выход фильтра $Y = \{y_k\}$ является его импульсной характеристикой и в нашей постановке задачи известен заранее. Система уравнений (2) имеет L строк, в правой части каждой из которых стоит отсчет импульсной характеристики. Поскольку в (2) $M+N+1$ неизвестных, то для того, чтобы система могла иметь решение, должно выполняться условие

$$L \geq M + N + 1 \quad (3)$$

Чтобы начать решать систему уравнений (2) не хватает только знания M и N . Пусть $N=M-1$ (таким образом, количество коэффициентов a_k и b_k будет равным), а M выберем любое, удовлетворяющее неравенству (3). Теперь построим систему из первых $M+N+1$ уравнений системы (2), назовём её начальной.

$$\left\{ \begin{array}{l} b_0 = y_0 \\ b_1 - y_0 a_1 = y_1 \\ b_2 - y_1 a_1 - y_0 a_2 = y_2 \\ \dots \\ b_N - (a_1 y_{N-1} + a_2 y_{N-2} + \dots + a_{M-1} y_0) = y_N \\ -(a_1 y_{N-2} + a_2 y_{N-3} + \dots + a_{M-1} y_1 + a_M y_0) = y_{N+1} \\ \dots \\ -(a_1 y_{N+M-1} + a_2 y_{N+M-2} + \dots + a_M y_{M-1}) = y_{N+M} \end{array} \right. \quad (4)$$

Для построения начальной системы необходимо $M+N+1$ членов последовательности Y . Анализ системы в виде (4) затруднителен, поэтому перепишем её в матричном виде.

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -c & 0 & 0 \\ 0 & 0 & 1 & -a-b-c & -c & 0 \\ 0 & 0 & 0 & -4a-2b-c & -a-b-c & -c \\ 0 & 0 & 0 & -9a-3b-c & -4a-2b-c & -a-b-c \\ 0 & 0 & 0 & -16a-4b-c & -9a-3b-c & -4a-2b-c \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -c & 0 & 0 \\ 0 & 0 & 1 & -a-b+c & -c & 0 \\ 0 & 0 & 0 & -2a & -a-b & -c \\ 0 & 0 & 0 & 0 & -2a & -a-b \\ 0 & 0 & 0 & 0 & 0 & -2a \end{pmatrix}$$

Аналогично для расширенной матрицы:

$$\bar{P} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & -c \\ 0 & 1 & 0 & -c & 0 & 0 & -a-b-c \\ 0 & 0 & 1 & -a-b-c & -c & 0 & -4a-2b-c \\ 0 & 0 & 0 & -4a-2b-c & -a-b-c & -c & -9a-3b-c \\ 0 & 0 & 0 & -9a-3b-c & -4a-2b-c & -a-b-c & -16a-4b-c \\ 0 & 0 & 0 & -16a-4b-c & -9a-3b-c & -4a-2b-c & -25a-5b-c \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -c & 0 & 0 & 0 \\ 0 & 0 & 1 & -a-b+c & -c & 0 & 0 \\ 0 & 0 & 0 & -2a & -a-b & -c & 0 \\ 0 & 0 & 0 & 0 & -2a & -a-b & 0 \\ 0 & 0 & 0 & 0 & 0 & -2a & 0 \end{pmatrix}$$

Таким образом, ранги матриц P и \bar{P} равны при любых параметрах a , b и c , $a \neq 0$. Построим теперь те же матрицы, но с L строками, $L > M + N + 1$. После аналогичных эквивалентных преобразований матрицы приводятся к следующему треугольному виду:

$$P \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -c & 0 & 0 \\ 0 & 0 & 1 & -a-b+c & -c & 0 \\ 0 & 0 & 0 & -2a & -a-b & -c \\ 0 & 0 & 0 & 0 & -2a & -a-b \\ 0 & 0 & 0 & 0 & 0 & -2a \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\bar{P} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -c & 0 & 0 & 0 \\ 0 & 0 & 1 & -a-b+c & -c & 0 & 0 \\ 0 & 0 & 0 & -2a & -a-b & -c & 0 \\ 0 & 0 & 0 & 0 & -2a & -a-b & 0 \\ 0 & 0 & 0 & 0 & 0 & -2a & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Значит $\text{rang}(P) = \text{rang}(\bar{P}) = 6$, и добавленные $L - M - N - 1$ уравнений линейно-зависимы от уравнений начальной системы. Поэтому решение, найденное для начальной системы, состоящей из 6-ти уравнений, будет удовлетворять последовательности отсчётов квадратичной функции любой длины большей 5. Это решение представляет собой 6 коэффициентов рекурсивного цифрового фильтра: $b_0, b_1, b_2, a_1, a_2, a_3$.

Таким образом, для любой последовательности Y длины $L \geq 6$, члены которой являются равноотстоящими отсчётами квадратичной функции $y_i = f(i) = a \cdot i^2 + b \cdot i + c$, $a \neq 0$, $i = 0, \dots, L-1$, существует рекурсивный цифровой фильтр третьего порядка, импульсная характеристика которого совпадает с Y .

Доказательство для других видов f проводится аналогично. Главное отличие состоит в том, что различным

функциям может соответствовать цифровой фильтр разного порядка.

Минимально достаточные значения M и N фильтров, соответствующие определённым элементарным функциональным зависимостям f , а также определяемое ими минимально необходимое для расчёта количество начальных отсчётов импульсной характеристики, представлены в таблице 1. Полиномиальные функции для краткости приведены только до 4-й степени.

Приведённые в таблице 1 зависимости являются достаточно простыми, и скорее всего члены исходной последовательности будут подчиняться им в редких случаях. Гораздо более реальным представляется случай, когда значения членов импульсной характеристики y_i подчиняются некоторой конечной сумме элементарных функций, то есть их линейной комбинации. Назовём функции, для которых существуют конечные M и N (и они известны) — *функциями, известными алгоритму A1*. Можно доказать, что, если значения коэффициентов импульсной характеристики ЦФ подчиняются конечной сумме таких функций, то для такого фильтра могут быть найдены M и N , и он может быть рассчитан. Доказательство этого проводится аналогично, приведенному в примере 1. Примеры минимальных значений M и N для фильтров с импульсными характеристиками, которые являются суммами элементарных функций известными алгоритму A1, приведены в таблице 2.

Из таблиц 1 и 2 видно, что чем более сложна функциональная зависимость, тем больше коэффициентов у соответствующего фильтра. Однако рост числа коэффициентов невелик и фактически пропорционален росту параметров функциональной зависимости.

Проиллюстрируем работу алгоритма расчёта ЦФ на конкретном примере.

ПРИМЕР 2. Пусть дана последовательность $Y = (48, 35, 24, 15, 8, 3, 0, -1, 0, 3)$. Значения членов Y определяются выражением $y_i = i^2 - 14i + 48$, $i = 0, 1, 2, \dots, 9$, то есть они являются отсчётами квадратичной функции, взятыми с шагом 1 (но заранее это неизвестно, известна только сама последовательность). Надо найти рекурсивный ЦФ, у которого начало импульсной характеристики совпадает с Y .

Таблица 1.

Минимально достаточные значения M и N для элементарных функций

Зависимость		M	N	L_{min}
Линейная	$y_i = k \cdot i + b$	2	1	4
Квадратичная	$y_i = a \cdot i^2 + b \cdot i + c$	3	2	6
Кубическая	$y_i = a \cdot i^3 + b \cdot i^2 + c \cdot i + d$	4	3	8
Полином 4-й степени	$y_i = a \cdot i^4 + b \cdot i^3 + c \cdot i^2 + d \cdot i + e$	5	4	10
Периодическая, T — количество отсчётов, приходящееся на один период	произвольная	T	$T - 1$	$2T$
Показательная	$y_i = k \cdot a^i + c$	2	1	4
Синусоида	$y_i = a \cdot \sin(b \cdot i + c) + d$	3	2	6
Косинусоида	$y_i = a \cdot \cos(b \cdot i + c) + d$	3	2	6

Таблица 2.

Минимальные M и N для линейных комбинаций функций

Зависимость		M	N	L_{min}
Сумма линейной и периодической с периодом 2	Например, $y_i = a \cdot i + b + c \cdot (-1)^i$	3	2	6
Сумма квадратичной и периодической с периодом 3		5	4	10
Сумма двух синусоид	$y_i = k_1 \cdot \sin(f_1 \cdot i + \varphi_1) + k_2 \cdot \sin(f_2 \cdot i + \varphi_2) + c$	5	4	10
Сумма показательной, квадратичной и синусоидальной функций	$y_i = k_1 \cdot a^{bi} + k_2 \cdot i^2 + k_3 \cdot i + k_4 \cdot \sin(f \cdot i + \varphi) + c$	6	5	12
Сумма трёх синусоид	$y_i = k_1 \cdot \sin(f_1 \cdot i + \varphi_1) + k_2 \cdot \sin(f_2 \cdot i + \varphi_2) + k_3 \cdot \sin(f_3 \cdot i + \varphi_3) + c$	7	6	14

Необходимо выбрать какое-то начальное количество коэффициентов фильтра. Выберем $M = 5$ (N , как указывалось выше, равно $M - 1$). Построив начальную систему уравнений и вычислив ранги её матриц, находим что она совместна, но обладает неполным рангом (8, при порядке системы равно 10). Поэтому, в соответствии с алгоритмом А1, уменьшаем M и N на единицу. При $M = 4$ ранг матрицы системы также меньше порядка системы (7 при порядке системы равно 8). Наконец, СЛАУ (5), построенная для $M=3$, получается совместной и определённой. То есть ранги матрицы системы уравнений и расширенной матрицы максимальны и равны.

$$\begin{cases} b_0 = 48, \\ b_1 - 48a_1 = 35, \\ b_2 - 35a_1 - 48a_2 = 24, \\ -24a_1 - 35a_2 - 48a_3 = 15, \\ -15a_1 - 24a_2 - 35a_3 = 8, \\ -8a_1 - 15a_2 - 24a_3 = 3. \end{cases} \quad (6)$$

После этого систему (6) можно решить при помощи любого стандартного метода. Находим: $b_0 = 48$, $b_1 = -109$, $b_2 = 63$, $a_1 = -3$, $a_2 = 3$, $a_3 = -1$.

Затем дополняем систему до полной (10 уравнений, 6 неизвестных). Теперь в ней учтены все члены последовательности Y . Находим ранги матриц системы и убеждаемся, что они не изменились. Это говорит о том, что все члены Y не противоречат найденному решению, даже члены с 6-го по 9-й, которые не участвовали в его нахождении. Для дополнительной проверки вычислим первые 10 отсчётов импульсной характеристики найденного фильтра и убедимся, что они совпадают с исходной последовательностью.

$$y(n) = \sum_{k=0}^2 b_k x(n-k) - \sum_{k=1}^3 a_k y(n-k)$$

$$y(0) = 1 \cdot b_0 = 48$$

$$y(1) = 1 \cdot b_1 - y(0) \cdot a_1 = 35$$

$$y(2) = 1 \cdot b_2 - y(1) \cdot a_1 - y(0) \cdot a_2 = 24$$

$$y(3) = -y(2) \cdot a_1 - y(1) \cdot a_2 - y(0) \cdot a_3 = 15$$

$$y(4) = -y(3) \cdot a_1 - y(2) \cdot a_2 - y(1) \cdot a_3 = 8$$

$$y(5) = -y(4) \cdot a_1 - y(3) \cdot a_2 - y(2) \cdot a_3 = 3$$

$$y(6) = -y(5) \cdot a_1 - y(4) \cdot a_2 - y(3) \cdot a_3 = 0$$

$$y(7) = -y(6) \cdot a_1 - y(5) \cdot a_2 - y(4) \cdot a_3 = -1$$

$$y(8) = -y(7) \cdot a_1 - y(6) \cdot a_2 - y(5) \cdot a_3 = 0$$

$$y(9) = -y(8) \cdot a_1 - y(7) \cdot a_2 - y(6) \cdot a_3 = 3$$

Заключение

Как видно, описанная методика довольно проста и может быть существенно модифицирована и оптимизирована. Все формы импульсных характеристик, к которым можно применить методику, определяются возможными функциями f , о которых говорилось выше. Это полиномиальные, показательные и периодические функции, либо их линейные комбинации. Так как полиномиальные и показательные функции являются расходящимися, то фильтр с такой импульсной характеристикой будет неустойчивым. Поэтому, с точки зрения расширения возможностей применения предлагаемой методики, для синтеза устойчивых цифровых фильтров важен поиск нерасходящихся функций f , а ещё лучше —

сходящихся к нулю. Полученный таким образом фильтр будет заведомо устойчивым. Тем не менее, применение алгоритма возможно и в случаях, когда устойчивость синтезируемого фильтра не имеет значения [1].

В завершение необходимо сказать, почему N было принято равным $M - 1$. В этом случае количества коэффициентов фильтра a_k и b_k получаются равными. Такое соотношение было определено опытным путём. Если выбрать N меньшим, чем $M - 1$, то решения вообще не будет существовать — система уравнений (5) будет несовместна. Если выбрать N большим, чем $M - 1$, то все коэффициенты b_k найденного фильтра начиная с M -го будут нулевыми. Поэтому дополнительные коэффициенты b_k не будут вносить никаких изменений на процесс вычисления импульсной характеристики в соответствии с (1).

Необходимо также еще раз отметить, что если рассмотренный алгоритм А1 находит решение, то импульсная характеристика рассматриваемого цифрового фильтра соответствует какой-то элементарной математической функции, либо же некоторой конечной сумме таких функций. Поэтому по окончании анализа полученных коэффициентов алгоритм А1 может стать основой методики отыскания отрезков элементарных функций в числовых последовательностях [2].

Таким образом, предложен и обоснован на примерах алгоритм синтеза цифрового БИХ-фильтра по его импульсной характеристике, работоспособный в том случае, если она соответствует элементарной математической функции или линейной комбинации таких функций.

Литература

1. Никитин Д.А. Сжатие временных рядов с использованием блочной интерполяции // Информационные технологии моделирования и управления. Науч.-технич. журнал. — Воронеж: Научная книга, 2007. — Вып. 1 (35). — С. 85–89.
2. Ханов В.Х., Никитин Д.А. Алгоритм анализа числовых последовательностей // Вестник СибГАУ. — Красноярск, 2006. — № 6 (13). — С. 10-14.

Уважаемые коллеги!

Предлагаем вам принять участие в формировании тематических выпусков журнала «Цифровая обработка сигналов» и размещению рекламы продукции (услуг) Вашей фирмы на его страницах. В случае положительного решения просим представить в редакцию журнала Ваши предложения по плановому размещению информационных материалов и макет рекламы продукции (услуг) Вашей фирмы с указанием желаемого её месторасположения: обложка (2-я, 3-я или 4-я стр.), цветная внутренняя полоса (объем полосы).

В 2009 году планируется выпуск 4-х номеров журнала (тираж до 1000 экз.). Журнал распространяется по подписке через агентство «Роспечать» в России, СНГ и странах Балтии (индекс 82185), а также на Выставках: «Цифровая обработка сигналов и ее применение – DSPA'2009», «ExpoElectronica», «Связь Экспокомм», «ЭЛЕКТРОНИКА: компоненты, оборудование, технологии» (г. Москва) и др.

Размещение рекламы Вашей фирмы на страницах журнала «Цифровая обработка сигналов» на плановой основе (не менее 2-х полных или 4-х половинчатых рекламных полос в течение года) предоставит Вам следующие возможности и права:

1. Первоочередное право расположения рекламных материалов на всех обложках (кроме 1-й) и страницах журнала.
2. Публикация представленных Вами рабочих (рекламных) материалов (статей) объемом до 6 полос в каждом очередном номере (в счет оплаченной рекламы).
3. Установка баннера Вашего сайта (или логотипа вашей организации) на 1-й странице сайта журнала «Цифровая обработка сигналов» (www.dsps.ru) в течение всего года, что привлечет внимание к продукции (услугам) Вашей фирмы новых участников на рынке DSP-технологий (ежедневно фиксируется до 100 и более посещений сайта www.dsps.ru).
4. Предоставление до 10 экз. очередного выпуска журнала.

Ориентировочная стоимость рекламных услуг:

4-я (внешняя) страница цветной обложки - 20 тысяч рублей.

2-я и 3-я (внутренние) страницы цветной обложки - 13 тысяч рублей.

1\2 цветной внутренней полосы - 7 тысяч рублей.

1\2 черно-белой внутренней полосы – 1 тысяча рублей.

Ждем Ваших предложений.

С наилучшими пожеланиями, зам. главного редактора

УДК 004.931

СИСТЕМА ФОРМИРОВАНИЯ ОБРАЗА И КЛАССИФИКАЦИИ ВРЕМЕННЫХ РЯДОВ ПО ХАРАКТЕРНЫМ ПОСЛЕДОВАТЕЛЬНОСТЯМ

Горшков А.П., Грызлова Т.П.

Введение

Задача классификации временных рядов является одной из самых актуальных научных задач в области технической диагностики и контроля, автоматического анализа сигналов и изображений, управления и множества других задач автоматизации интеллектуальных функций человека. Ключевой проблемой при решении задачи классификации временных рядов в рамках математического подхода теории распознавания является определение эффективного признакового пространства (пространства образов). Обычно эта задача решается эвристически специалистами предметной области. Известны методы автоматизированного поиска признаковых пространств для классификации временных рядов [1, 2]. Построение подобных систем связано с решением двух задач: определением семейства признаков и разработкой алгоритма поиска эффективного пространства признаков в рамках этого семейства. Для формирования правила классификации при заданном признаковом пространстве может быть использован один из методов математического подхода теории распознавания образов [3, 4]. В настоящей работе показано, что эффективность семейства признаков, отобранных из характерных последовательностей сигналов с помощью алгоритмов и методов системы «Гиперкуб», выше, чем у систем классификации временных рядов, использующих другие типы образов.

Методы классификации временных рядов

Известные методы классификации временных рядов разрабатываются либо на основе представления сигналов как точек в признаковых пространствах, либо на основе структурного описания, либо непосредственно по сигналам, т.е. отличаются по типу образа.

Образы в признаковых пространствах, как правило, вычисляются методами ЦОС, теории динамических систем или прикладной статистики.

Распространённым типом образа является структурное описание, когда сигналы представляются в виде цепочки производных элементов, каждому из которых соответствует сегмент сигнала [2]. В структурном распознавании образ характеризуется множествами производных элементов и отношений между ними [3].

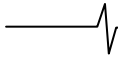
Во многих методах классификации временных рядов используют непосредственно сигналы, без определения признаков или структурных элементов. В частности, к ним относится классификация на основе метода к ближайших соседей и некоторого расстояния на множестве сигналов. Одним из наиболее эффективных методов сравнения сигналов является метод динамической трансформации

Представлено решение широкого класса задач классификации временных рядов. Решение ищется в автоматизированной системе формирования признакового пространства на основе статистик характерных последовательностей временных рядов. Приведены сравнительные оценки вероятности распознавания на известных тестовых задачах «Пистолет – палец», «Переходные процессы», «Полупроводниковая пластина».

времени (Dynamic Time Warping) [6 – 8]. Результаты исследований разработанной нами системы формирования образа и классификации временных рядов по характерным последовательностям сопоставляются с альтернативными подходами: методом динамической трансформации времени, системой Zeus и системой обобщенного формирования признаков для структурного распознавания образов.

Метод динамической трансформации времени – это процедура сравнения временных рядов на основе динамического программирования. При сравнении двух сигналов (вычислении расстояния) производится их нелинейное наложение. Классификация выполняется по правилу ближайшего соседа. Базовая процедура динамической трансформации времени заключается в следующем. Пусть заданы два временных ряда: $\mathbf{x} = x(t) = x_0 x_1 \dots x_{N-1}$ и $\mathbf{y} = y(t) = y_0 y_1 \dots y_{M-1}$. Построим матрицу D размера $M \times N$, в каждой ячейке которой указаны квадратичные отклонения: $D_{i,j} = (x_i - y_j)^2$. Расстояние DWT – это минимальная сумма значений ячеек, соответствующих пути из ячейки $D_{0,0}$ в ячейку $D_{N-1,M-1}$. Другими словами, необходимо найти путь с минимальной суммой. Значение этой суммы и будет выражать расстояние между последовательностями. Существует несколько модификаций метода DWT с различными ограничениями на область допустимых ячеек в матрице. Например, в работе [8] область допустимых значений матрицы определяется на основе обучающей выборки, что позволяет повысить точность классификации [6 – 8].

В системе Zeus признаки для распознавания формируются при помощи генетического программирования. Структура признака описывается контекстной грамматикой. В качестве структурных элементов используются операции над сигналами: автокорреляция; определение окна заданием его центра и ширины; индекс максимального (минимального) значения; индекс точки пересечения сигналом заданного значения и другие. На основе алгоритмов генетического программирования производится поиск систем классификации в целом. То есть, одновременно с формированием признаков выполняют объединение их в признаковое пространство, обучение классификатора (например, линейного дискриминатора Фишера) и оценку эффективности полученной системы на обучающей выборке [1].



В методе обобщенного формирования признаков для структурного распознавания классификация производится на основе структурного представления временных рядов, а именно, разбиения их на части и аппроксимации полученных сегментов функциями: константы, линейной, экспоненциальной, синусоиды, кусочно-линейных в форме треугольника и трапеции. Применяется также совмещенное описание, когда разные сегменты описываются различными функциями. Несмотря на структурное представление сигналов, классификация производится в признаковом пространстве. Вектор признаков описывает ряд параметров разбиения на части и аппроксимации, включая позицию сегмента, условный номер функции аппроксимации, параметры аппроксимации, ошибку аппроксимации сегмента и другие [2].

Семейство признаков на базе характерных последовательностей

Предлагается формировать признаки на основе анализа появления похожих подпоследовательностей сигналов. Характерная последовательность (ХП) cs – это элементарная последовательность (ЭП), которая неоднократно встречается в сигнале (кластер «Одинаковых» ЭП). Для формализации процедуры сопоставления последовательностей вводится расстояние на множестве ЭП одинаковой длины l и максимально допустимый порог расстояния Th_p , при котором ЭП считаются одинаковыми. Распространенными расстояниями для сравнения последовательностей являются метрики Чебышева и СКО:

$$\rho(x, y) = \max_{i=1..l} |x_i - y_i| \quad (1)$$

$$\rho(x, y) = \frac{1}{l} \sum_{i=1..l} (x_i - y_i)^2 \quad (2)$$

Наряду с ними используется расстояние на основе коэффициента корреляции $\rho_{x,y}$:

$$\rho(x, y) = 1 - \rho_{x,y} \quad (3)$$

Основная операция, производимая над сигналом – это сопоставление его ЭП с эталоном. При сопоставлении производится перебор ЭП сигнала с заданным в отсчётах шагом и их сравнение с эталоном по заданному порогу. Семейство признаков для формирования образа включает в себя частотные признаки на основе ХП. Частотный признак f на основе ХП вычисляется как отношение количества ЭП сигнала,

расстояние между которыми и эталоном не превосходит порога, к общему числу перебранных ЭП.

Значение частотного признака можно интерпретировать как оценку вероятности попадания последовательностей сигнала x в область координатного пространства последовательностей $D_{cs} : \{x | \rho(cs, x) < Th_p\}$, где cs – эталонная последовательность, а Th_p – порог. Выбор расстояния определяет форму области D_{cs} . Порог позволяет управлять её размерами. Например, ХП с метрикой (1) в пространстве последовательностей соответствует гиперкуб с центром в эталонной последовательности, параллельными осям ребрами и стороной $2T$.

Алгоритмы формирования признаков

Пусть S – обучающая выборка, а F – искомое признаковое пространство. Поиск признакового пространства F заключается в последовательном выполнении алгоритмов формирования множества возможных признаков F_2 (рис. 1) и алгоритма поиска эффективного признакового пространства $F \subset F_2$ (рис. 2).

Множество возможных признаков строится на основе множества ХП (CS). Множество ХП формируется с помощью кластерного анализа последовательностей сигналов обучающей выборки S . При этом используется простой алгоритм кластерного анализа, изложенный в [3]. Для проведения кластерного анализа задаётся расстояние $\rho(x, y)$, порог Th_p , размер l и шаг перебора последовательностей. Формирование множества ХП производится для нескольких размеров последовательностей, расстояний и иногда для нескольких значений порогов. Таким образом, кластерный анализ последовательностей выполняется неоднократно, а полученные множества ХП объединяются. Длины последовательностей задаются на основе визуальной оценки графиков сигналов.

При настройке порогов для кластерного анализа используется его вероятностное представление. Порогу в $n\%$ соответствует такое расстояние d , для которого вероятность появления в сигналах обучающей выборки двух последовательностей с расстоянием, не превосходящим d , составляет $n\%$. Обычно используется порог в 7% или 10%. Иногда дополнительно рассматривается порог в 25% или 30%. Параметры кластерного анализа для трёх задач классификации временных рядов, рассмотренных ниже, представлены в таблице 1.

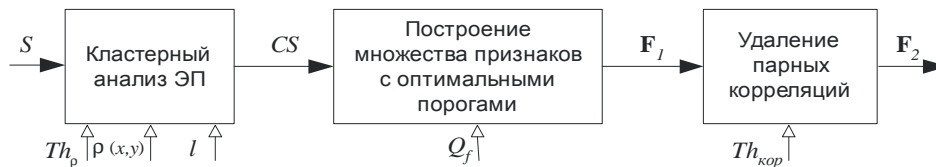


Рис. 1. Структурная схема алгоритма формирования множества возможных признаков для семейства частотных признаков на основе ХП

Таблица 1.

Параметры кластерного анализа последовательностей

Задача	Размеры последовательностей	Порог расстояния
«Переходные процессы»	5, 10, 25, 50, 75	10%, 30%
«Полупроводниковая пластина»	5, 10, 20, 30, 40, 50, 75, 100	10%, 30%
«Пистолет - палец»	10, 20, 30, 40, 50, 75, 100	7%, 30%

На основе множества ХП формируется множество признаков с оптимальными порогами за счёт оптимальной (по критерию максимума заданной меры информативности признаков Q_f) установки порогов для каждой из ХП. После этого из множества полученных признаков удаляется по одному из парно коррелированных по пороговому значению коэффициента корреляции $Th_{кор}$ признаку.

При автоматизированном формировании множество возможных признаков F_2 имеет большие размеры (больше 500), поэтому выполняется отбор заданного числа наиболее информативных признаков. При этом формируется множество признаков $F_3 \subset F_2$. Далее выполняется случайный поиск с адаптацией (метод СПА), для которого необходимо задать меру информативности признаков пространства Q_F [4]. Информативность признакового пространства определяется кластерными свойствами расположения множеств точек, соответствующих разным классам.

При использовании в системе преобразования кластеризации, оно включается в алгоритм Q_F , и информативность оценивается для преобразованного пространства. В отличие от алгоритмов оценки информативности пространств признаков Q_F алгоритмы вычисления информативности признаков Q_f удаётся сделать более быстрыми: с линейной оценкой вычисления относительно размера обучающей выборки, что и определяет необходимость их отдельного использования.

Время синтеза зависит от таких параметров задачи, как количество классов, размер обучающей выборки и длина реализации сигналов. Программный комплекс «Гиперкуб» позволяет решать задачу повышения качества уже разработанной системы диагностики, если она работает по принципу вычисления признаков. Для этого должна быть предоставлена таблица значений признаков существующей системы классификации для сигналов обучающей выборки. Помимо технической диагностики, система «Гиперкуб» позволяет решать и другие задачи классификации временных рядов. Готовятся публикации с успешными результатами решения трех задач классификации временных рядов, представляющих собой модельные данные, сигналы из областей технической диагностики и распознавания изображений.

Сравнительная оценка эффективности систем классификации

Испытания проводились на трёх наборах временных рядов, доступных в англоязычном Интернете: «Переходные процессы» («trace»), «Пистолет - палец» («gun-point») и «Полупроводниковая пластина» («wafer») [5 – 8, 2].

Множество сигналов «Переходные процессы» является подмножеством набора данных для распознавания переходных процессов (the Transient Classification Benchmark), впервые представленного в [9]. Оно сформировано для воспроизведения недостатков оборудования ядерных электростанций. Полный набор данных включает 16 клас-

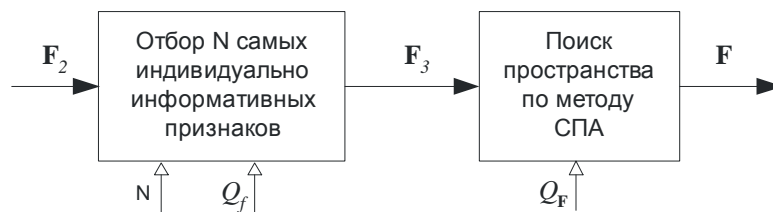


Рис. 2. Структурная схема алгоритма поиска информативного пространства признаков

Таблица 2.

Параметры формирования признаков

Параметр	Значение
Максимально допустимое значение коэффициента корреляции между признаками при формировании множества с оптимальными порогами	0,9
Количество N наиболее индивидуально информативных признаков, передаваемых на СПА	500
Количество эпох в методе СПА	1000
Количество шагов в эпохе	50

В таблице 2 представлены значения управляющих параметров алгоритмов формирования признаков, использованные при решении тестовых задач классификации временных рядов.

Программная система классификации временных рядов «Гиперкуб» создана на платформе C++ Builder 6 и позволяет за одну–две недели получить алгоритм автоматической диагностики, приложение и библиотеку с функциями вычисления признаков и принятия решения. Для синтеза процедуры классификации необходима обучающая выборка – множество классифицированных сиг-

сов по 50 экземпляров каждого. Каждый экземпляр характеризуется 4 сигналами. Как и в ряде работ [5 – 8, 1] используется уменьшенный набор данных, включающий только второй сигнал для классов 2 и 6, а также третий сигнал для классов 3 и 7. Таким образом, получаются 4 класса, каждый из которых характеризуется одним сигналом с длиной реализации 275 отсчётов. Уменьшение количества рассматриваемых процессов, характеризующих один экземпляр, связано с избыточностью такого описания. На рис. 3 представлены примеры процессов набора данных «Переходные процессы».

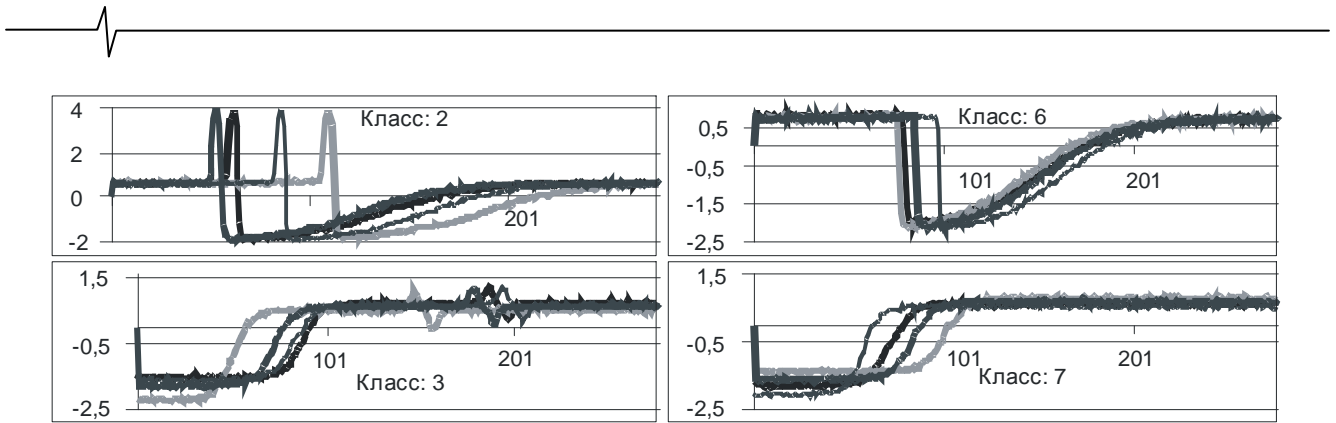


Рис. 3. Примеры процессов набора данных «Переходные процессы»

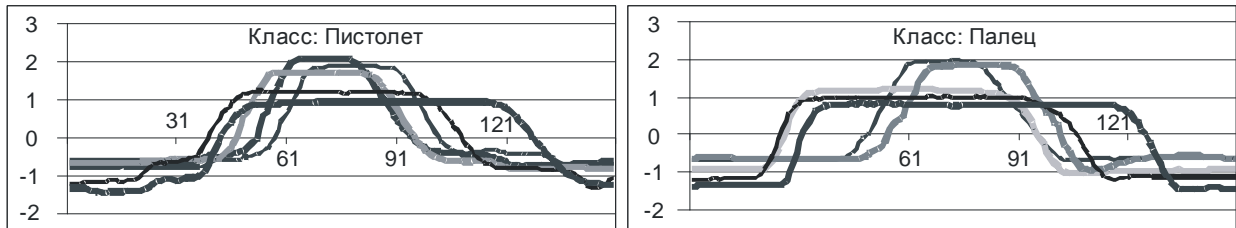


Рис. 4. Примеры процессов набора данных «Пистолет - палец»

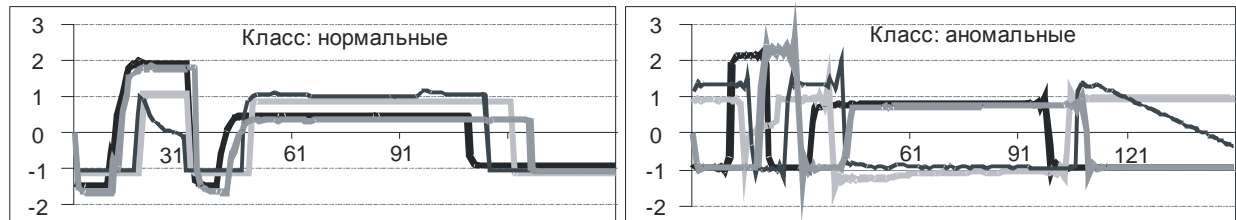


Рис. 5. Примеры процессов набора данных «Полупроводниковая пластина»

Набор данных «Пистолет - палец» относится к области видеонаблюдения. Имеются процессы двух классов, каждый из которых представлен сотней экземпляров. Процессам класса «пистолет» соответствует следующая последовательность действий:

- актёр держит руки по швам;
- достаёт пистолет из кобуры на бедре и, примерно в течение одной секунды, наводит его на цель;
- возвращает пистолет обратно в кобуру.
- Порядок действий при формировании процессов класса «палец»:

- актёр держит руки по швам;
- показывает указательным пальцем на цель в течение приблизительно одной секунды;
- возвращает руку на место.

Сигналы каждого класса построены с участием двух актёров: мужчины и женщины. Процессы отражают изменение координаты x центра тяжести правой руки актёра. Длина реализации составляет 150 отсчётов [6 – 8]. Примеры процессов представлены на рис. 4.

Набор данных «Полупроводниковая пластина» является коллекцией временных рядов, содержащих последовательности показаний интенсивности излучений двух различных длин волн (405 и 520 нанометров), регистрируемых в процессе напыления при производстве полупроводниковых устройств [2]. Используется набор данных, где отсутствует информация о том, какой длине волны соответствует тот или иной сигнал [6 – 8].

Каждый сигнал классифицируется как нормальный или аномальный. Аномальные сигналы представляют ряд проблем, возникающих при производстве полупроводниковых устройств. Всего имеется 7164 сигнала. Среди них 762 – аномальных. Обучающая выборка: 717 сигналов (10%), из них 77 аномальных. Тестовая выборка: 6447 сигналов, из них аномальных – 685. Примеры сигналов показаны на рис. 5.

В работах [4 – 7] представлены результаты экспериментальной проверки метода динамической трансформации времени (Dynamic time warping, DWT) для классификации рассмотренных временных процессов. Для оценки вероятности правильного распознавания (p) в [5] набор данных разбивается на тестовую и обучающую выборки. Оценка p вычисляется как отношение правильно распознанных процессов тестовой выборки к общему её объёму. Такой критерий обозначим как n -ТТ, где n – процентное выражение доли тестовой выборки. В работах [6 – 8] оценка p производится при помощи более оптимистичного критерия. При этом перебираются все экземпляры набора данных. Проверка правильности распознавания каждого из них производится по ближайшему соседу на основе оставшихся. Такой критерий обозначим 1-NN (leave one out).

В [1] получена оценка p для классификации наборов данных «Переходные процессы» и «Пистолет - палец» на основе системы Zeus. Оценка p производилась по методу скользящего контроля по 10 блокам. Обозначим этот критерий: 10-CV.

Таблица 3.

Оценка вероятности правильного распознавания – p

Метод		«Переходные процессы»		«Полупроводниковая пластина»		«Пистолет - палец»	
		Критерий	$p, \%$	Критерий	$p, \%$	Критерий	$p, \%$
На основе ХП с расстоянием	(1)	76-ТТ	100	90-ТТ	96.13	75-ТТ	99.3
	(2)	76-ТТ	100	90-ТТ	95.18	75-ТТ	94.6
	(3)	76-ТТ	100	90-ТТ	99.94	75-ТТ	99.3
	(1), (2) и (3)	76-ТТ	100	90-ТТ	99.98	75-ТТ	100
Метод динамической трансформации времени	[5]	50-ТТ	100	86-ТТ	99.5	75-ТТ	91.3
	[6 - 7]	1-NN	100	-	-	1-NN	99
	[8]	1-NN	100	-	-	1-NN	99.5
Система Zeus [1]		10-CV	100	-	-	10-CV	98.5
Обобщенное формирование структурных признаков [2]		-	-	10-CV	98.6	-	-

При вычислении критерия обучающая выборка случайным образом разбивается на 10 непересекающихся подмножеств одинакового или почти одинакового размера. Обучение и тестирование производится 10 раз. Каждый из блоков по очереди используется в качестве тестовой выборки, а обучение производится по 9-и оставшимся блокам. В качестве оценки p используется среднее из полученных 10-и значений. Таким образом, размер тестовой выборки при каждом тестировании составляет 10% размера выборки.

В [2] представлены результаты классификации процессов набора данных «Полупроводниковая пластина» при помощи метода обобщенного формирования признаков для структурного распознавания. Оценка p производится по процедуре скользящего контроля 10-ТТ. В работе представлены результаты для всех функций аппроксимации. Результаты получены отдельно для сигналов разной длины волны 405 и 520. Кроме того, вместо общей вероятности p даны отдельно вероятности правильного распознавания для сигналов каждого из классов: p_a и p_n . Для сопоставления значения p мы взяли наибольшие показатели по разным функциям аппроксимации. Значение p вычисляется на основе p_a и p_n по следующей формуле: $p = (C_a \cdot p_a + C_n \cdot p_n) / (C_a + C_n)$, где C_a и C_n – количества сигналов каждого из двух классов.

В таблице 3 представлены оценки вероятности правильного распознавания для всех трёх наборов данных в сравнении с другими подходами. На основе ХП системы распознавания строятся для каждого расстояния (1), (2) и (3) в отдельности и на объединенном множестве признаков для всех трёх расстояний. Решение, основанное на характерных последовательностях, оказывается более эффективным. Результаты тестирования на независимой выборке являются в данном случае более пессимистичными, чем оценки 1-NN и 10-CV. Прежде всего, это связано с уменьшением объема обучающей выборки. Как показано в [6, 7], для набора данных «Пистолет - палец», даже незначительное уменьшение объема обучающей выборки приводит к заметному ухудшению результатов.

Заключение

Анализ трёх различных задач классификации временных рядов показал, что формирование образа на основе ХП позволяет получить более эффективные системы распознавания, чем системы на основе динамической транс-

формации времени (DWT) [5–8], а также структурного описания при помощи системы Zeus [1] и обобщенного формирования структурных признаков [2]. Построенные системы легко поддаются интерпретации и имеют достаточно простую структурную схему.

В [10] представлено решение задачи диагностики состояния подшипников трансмиссии ГТД по вибрационным сигналам, основанное на представленном семействе признаков. Успешное применение частотных признаков на основе ХП при работе с различными типами сигналов позволяет сделать вывод о гибкости и, возможно, универсальности предложенной системы классификации временных рядов.

Применение расстояния на основе коэффициента корреляции для рассмотренных задач является более эффективным, чем использование метрик Чебышева и СКО. В работе [10], напротив, показана преимущественная эффективность метрик СКО и Чебышева. Это не позволяет отказаться от использования ни одной из рассматриваемых метрик в общем случае. Кроме того, наилучшие результаты получаются именно при объединении признаков, полученных для всех трёх расстояний.

Предложенная система применима для синтеза процедур классификации временных рядов различных типов. Малые временные затраты на синтез делают применение этой технологии экономически целесообразным.

Литература

- Eads D., Glocer K., Perkins S., Theiler J. Grammar-guided feature extraction for time series classification. Neural Information Processing Systems, 2005.
- Olszewski R. T. Generalized Feature Extraction for Structural Pattern Recognition in Time-Series Data. PhD thesis, Carnegie Mellon University, Pittsburgh, PA, 2001.
- Ту Дж., Гонсалес Р. Принципы распознавания образов. – М: Мир, 1978.
- Загоруйко Н.Г. Методы распознавания и их применение. - М: Советское радио, 1972.
- Keogh E. UCR Time Series Archive: www.cs.ucr.edu/~eamonn/TSDMA/, 2007.
- Xiaopeng Xi, Keogh E., Shelton C., Li Wei & Ratanamahatana C.A. Fast Time Series Classification Using Numerosity Reduction. International Conference on Machine Learning, 2006.
- Ratanamahatana C. A., Keogh E. Everything you know about dynamic time warping is wrong. In 10th ACM SIGKDD Interna-

tional Conference on Knowledge Discovery and Data Mining Workshop on Temporal Data Mining, 2004.

8. Ratanamahatana, C. A. and Keogh. E. Making Time-series Classification More Accurate Using Learned Constraints. SIAM International Conference on Data Mining, April 22-24, 2004.
9. Roverso D. Multivariate temporal classification by windowed wavelet decomposition and recurrent neural networks. In 3rd ANS

International Topical Meeting on Nuclear Plant Instrumentation, Control and Human-Machine Interface, 2000.

10. Горшков А.П., Грызлова Т.П., Комаров Б.И., Шепель В.Т. Диагностика состояния подшипников трансмиссии газотурбинных двигателей в пространствах статистик характерных последовательностей вибраций. // Авиационно-космическая техника и технология 10 / 36. - Харьков, ХАИ, 2006.

ПОЗДРАВЛЕНИЯ ЮБИЛЯРУ



Заместителю Главного редактора нашего журнала, лауреату премии правительства РФ в области науки и техники, доктору технических наук, профессору ДВОРКОВИЧУ Виктору Павловичу – 70 лет!

Свой юбилей Виктор Павлович встречает на новом ответственном государственном посту – заместителя директора ФГУП «Главный радиочастотный центр», проработав до этого 45 лет в ФГУП «НИИ радио». Вся творческая судьба Виктора Павловича была связана с этим ведущим российским институтом, начальником отдела в котором он был последние годы.

Понимая жизненную важность и научно-техническую значимость продвижения в России новых информационных технологий, опирающихся на цифровую обработку сигналов и изображений в реальном времени, Виктор Павлович стал одним из инициаторов организации и проведения в нашей стране международной научно-технической конференции: «Цифровая обработка сигналов и ее применение», которая уже в 10-й раз про-

шла в марте этого года. И с самого первого дня профессор Дворкович В.П. является бесменным руководителем секции: «Обработка изображений».

С созданием научно-технического журнала: «Цифровая обработка сигналов» в 1999 году, став заместителем Главного редактора, Виктор Павлович проводит огромную редакционную работу. Фактически ежегодно под его руководством на страницах нашего журнала формируется тематический выпуск «Цифровая обработка и передача изображений».

Виктор Павлович известен как один из ведущих российских ученых, плодотворно работающих в области современных телекоммуникаций, цифрового телевидения, обработки изображений. Он автор более 200 научных работ (в том числе более 10 книг), имеет более 70 авторских свидетельств и патентов.

Отмечая активную научную и организаторскую деятельность Дворковича В.П., как ученого и руководителя творческого коллектива, члена экспертного Совета ВАК РФ, члена оргкомитета МНТК «Цифровая обработка сигналов и ее применение», заместителя Главного редактора нашего журнала, редакционная коллегия журнала «Цифровая обработка сигналов» поздравляет Виктора Павловича с 70-летием! Крепкого Вам здоровья, новых научных достижений, верных друзей и семейного благополучия.

Главный редактор
Зам. Главного редактора

Ю.Б. Зубарев
В.В. Витязев

УДК 519.218.82

ВЕЙВЛЕТ-ТЕХНОЛОГИЯ АНАЛИЗА И ОЧИСТКИ СИГНАЛОВ ОТ ШУМА*Орешко Н.И., Князева Т.Н.***Введение**

В многочисленных статьях и книгах по теории и практике использования вейвлетов приведен необходимый инструментарий по отдельным аспектам обработки данных с их использованием. Однако для большинства пользователей, не знакомых с тонкостями и проблемами в вейвлет-анализе, представляет значительную трудность выбрать, например при очистке от шума реальных данных, многочисленные настроечные параметры, приводящие к эффективной обработке. К ним, в первую очередь, относятся следующие: какой вид базисной функции (т.е. какой тип фильтра) выбрать, какой должна быть длина фильтра, сколько нужно задать уровней очистки от шума, какое выбрать пороговое правило и тип порога, как скомпенсировать граничные эффекты и т.д. И самое главное как определить эффективность примененной обработки.

Учитывая все расширяющуюся применимость вейвлетов в инженерной практике, эта задача становится все более актуальной. Необходимы некоторые процедуры автоматизации и структуризации при выборе и настройке всех параметров. Все это объединено в понятии технологии использования вейвлетов, которой и посвящается данная статья. Так как рассматриваемый предмет очень широк, то мы ограничим свое внимание одним из важных его аспектов, а именно задачей очистки от шума одномерных данных.

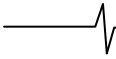
Задача очистки от шума имеет глубокие корни, начиная еще от исследований Гаусса. Многочисленные подходы, предложенные в регрессионном анализе и теории Калмановской фильтрации, преимущественно опираются на полное или частичное знание параметрической модели. В параметрической регрессии неизвестная зависимость предполагается принадлежащей известному семейству функций, зависящих от конечномерного параметра. Как правило, упомянутое семейство функций представляет собой линейную комбинацию базисных функций, заданных априори. Размерность семейства либо берется с запасом, либо определяется по выборке. Если выбор конечномерного семейства определяется знанием закона, объясняющего данную зависимость, а также при недостатке данных или вычислительных средств такой подход оправдан. В других ситуациях попытка втиснуть изучаемую зависимость в узкое, заданное априори, семейство может привести к искажению даже качественной информации об изучаемой зависимости, например, о ее особых точках, изменении характера зависимости на разных участках и др. Методы непараметрической регрессии, разработанные для

Приводится описание основных этапов вейвлет-технологии анализа и очистки от шума сигналов, включающей решение многочисленных задач, таких как выбор порогов и правил усечения, выбор наилучшего базиса, работу с выборками произвольного объема, обработку на границах, уменьшение эффекта Гиббса и др. Рассматривается мультиполосный анализ, позволяющий определить поведение флуктуаций на разных масштабах, выявить причины их происхождения, удалить ненужные циклические компоненты.

анализа таких случаев, не предполагают включения искомой зависимости в заданное априори конечномерное семейство. Предварительное задание параметрической модели может оказаться слишком ограничительным или чересчур малой размерности для аппроксимации непредвиденных характеристик, в то время как непараметрическое сглаживание предоставляет гибкие средства анализа неизвестных регрессионных зависимостей. За прошедшее десятилетие были предложены различные нелинейные (пространственно адаптивные) оценщики в непараметрической регрессии. Наиболее популярными являются ядерные методы с переменной полосой пропускания, классификационные и регрессионные деревья, адаптивные регрессионные сплайны. Несмотря на то, что эти методы достигают асимптотически оптимальных оценок, они имеют большой объем вычислений и обычно предназначены для очистки от шума регулярных функций.

В 90-е годы прошлого столетия в зарубежной литературе по непараметрической регрессии нашли широкое применения нелинейные методы вейвлет-пороговой обработки. Эти оценщики образуют подмножество старого класса непараметрических регрессионных методов оценивания, так называемые методы ортогональных рядов. При этом в отличие от алгоритмов линейного сглаживания данных, вейвлет-пороговая обработка обладает способностью адаптироваться к той или иной регулярности. Кроме того, эти методы оценивания легко применить, используя быстрые алгоритмы, так что они очень полезны с практической точки зрения.

Вейвлеты представляют собой новое семейство базисных функций с компактным носителем, которые используются для эффективного представления реальных сигналов. Они сочетают такие важные свойства, как: различную степень гладкости, локализацию во временной и частотной областях, представление сигнала на различных масштабах, что позволяет анализировать мелкомасштабные флуктуации и особенности высокого порядка, разреженность представления сигнала, быструю численную реализацию. Вейвлет-анализ представляет сигнал через систему сдвинутых и



масштабных версий единственной базисной вейвлет-функции. По вейвлет-коэффициентам можно определить разрывы и резкие изменения в функциях и их производных. Сигналы с резкими локальными изменениями могут хорошо представляться лишь малым числом вейвлет-коэффициентов, что невозможно при использовании стандартных ортонормальных базисов, для которых характерны многочисленные "компенсирующие" коэффициенты для представления разрывов и подавления эффекта Гиббса. Согласно принципу Гейзенберга, невозможно одновременно точное воспроизведение сигнала и во временной, и в частотной областях. Разреженность вейвлет-преобразований является отличительной чертой вейвлет-базисов и обусловлена возможностью регулирования ограничений в принципе Гейзенберга в зависимости от поведения исходного сигнала. Таким образом, вейвлеты автоматически находят компромисс между частотной и временной локализациями. Все эти аргументы являются подтверждением того, что вейвлет-базис - это подходящий инструмент для создания эффективных процедур анализа и обработки данных.

Модель входного сигнала и дискретное вейвлет-преобразование

Поскольку в реальных ситуациях мы имеем дело с сигналами, представляющими собой конечную последовательность отсчетов, то в данной статье рассматривается только дискретный вейвлет-анализ.

Входной сигнал y - N -мерный вектор, представляется в виде выборки неизвестной функции f на равнодискретной временной сетке t_i . В стандартной постановке предполагается также, что ξ_i - независимые случайные переменные, принадлежащие одной и той же функции нормального распределения. Тогда

$$y_i = f(t_i) + \sigma \xi_i, \quad i = 0, \dots, N-1,$$

где σ - среднеквадратическое отклонение шума. Задача заключается в том, чтобы получить оценку для функции f из зашумленных данных $y = (y_0, \dots, y_{N-1})$, при неизвестной параметрической модели f .

В данной статье мы рассматриваем более реалистичную модель, приближенную к ситуации реальных измерений, а именно: наличие одиночных и множественных сбоев в данных, временных разрывов, нестационарного (гетероскедастического) шума, резких изменений в сигнале и его производных. Длина входной выборки предполагается произвольной (т.е., в общем случае, не равная степени двойки).

Существующие на данный момент вейвлет-преобразования можно разделить на две основные группы: неизбыточные (общее число вейвлет-коэффициентов при уровне разложения J соответствует объему выборки N) и избыточные (число вейвлет-коэффициентов на уровне J равно $(J+1)N$).

Неизбыточные преобразования основаны на традиционном дискретном вейвлет-преобразовании с использованием различных способов пороговой обработки, порогов и правил усечения. Методы этой группы работают с выборками, объем которых кратен степени двойки.

В группе избыточных вейвлет-преобразований одним из наиболее эффективных алгоритмов является максимально накладывающееся дискретное вейвлет-

преобразование (МНДВП) [1]. Несмотря на то, что МНДВП является избыточным неортогональным преобразованием, оно обеспечивает более качественную очистку от шума, инвариантно относительно сдвига и способно работать с выборками произвольного объема. Все основные этапы обработки и анализа данных, перечисленные ниже, были адаптированы под данное преобразование.

Отличительной чертой МНДВП является нормирование коэффициентов высокочастотного и низкочастотного фильтров на $\sqrt{2}$ и отсутствие операции децимации. На первом уровне вейвлет-преобразования вейвлет-коэффициенты $W_1(i)$ и аппроксимирующие коэффициенты $V_1(i)$ рассчитываются по формулам:

$$W_1(i) = \sum_{l=0}^{L-1} g_l y((i-l) \bmod N), \quad V_1(i) = \sum_{l=0}^{L-1} h_l y((i-l) \bmod N),$$

$$\text{где } (i-l) \bmod N = \begin{cases} i-l, & \text{если } i-l \geq 0, \\ i-l+N, & \text{в остальных случаях,} \end{cases}$$

g_l, h_l - коэффициенты высокочастотного и низкочастотного фильтров соответственно. Они могут быть получены из обычных коэффициентов фильтров g_l^*, h_l^* (например, Добеши) путем нормирования на $\sqrt{2}$, т.е. $g_l = g_l^* / \sqrt{2}$ и $h_l = h_l^* / \sqrt{2}$, таким образом,

$$\text{что } \sum_{l=0}^{L-1} g_l = 0 \text{ и } \sum_{l=0}^{L-1} h_l = 1, \text{ где } L - \text{длина фильтра.}$$

Вейвлет-коэффициенты и аппроксимирующие коэффициенты j -го уровня вычисляются по формулам

$$W_j(i) = \sum_{l=0}^{L-1} g_l V_{j-1}((i-2^{j-1}l) \bmod N),$$

$$V_j(i) = \sum_{l=0}^{L-1} h_l V_{j-1}((i-2^{j-1}l) \bmod N)$$

при этом $V_0(i) = y(i)$. Эти два соотношения составляют прямой алгоритм МНДВП. Обратное МНДВП вычисляется с помощью обратного алгоритма по следующей формуле:

$$V_{j-1}(i) = \sum_{l=0}^{L-1} g_l W_j((i+2^{j-1}l) \bmod N) + \sum_{l=0}^{L-1} h_l V_j((i+2^{j-1}l) \bmod N).$$

Основные этапы вейвлет-технологии анализа и очистки от шума сигналов

Для эффективной очистки от шума сигналов в условиях реальных измерений предлагается вейвлет-технология, которая включает следующие этапы обработки и анализа данных:

- 1) предварительная обработка и анализ данных;
- 2) анализ, классификация, кластеризация, сегментация и отбор участков для обработки данных;
- 3) разведочный анализ для адаптивного выбора способов и параметров обработки;
- 4) очистка от шума;
- 5) анализ и интерпретация результатов обработки.

Предварительная обработка и анализ данных. Реальные данные нередко содержат большое количество одиночных и множественных сбоев, поэтому требуется проведение предварительной обработки данных с целью их обнаружения, удаления и заполнения образовавшихся разрывов, коррекции обнаруженных сбоев.

Одиночные и множественные сбои возникают в результате грубых ошибок измерений. Эти ошибки лучше всего выявить и исключить из дальнейшего рассмотрения в самом

начале обработки данных.

Предлагается автоматизированный метод обнаружения выбросов [2], основанный на МНДВП. Основная идея метода обнаружения выбросов опирается на три наблюдения. Первое основное наблюдение заключается в том, что выброс соответствует скачку в свойствах непрерывности функции или ее производных, поэтому он характеризуется большими значениями вейвлет-коэффициентов. Второе наблюдение заключается в том, что при корректном выборе длины фильтра (то есть, чтобы число его нулевых моментов соответствовало гладкости исходного сигнала) вся гладкая составляющая не проходит в вейвлет-область, а переходит в область аппроксимирующих коэффициентов. При этом вся информация о тонких особенностях сигнала остается в вейвлет-области. Третье состоит в том, что в вейвлет-области общий вид одиночных выбросов имеет форму, схожую с используемым высокочастотным вейвлет-фильтром. Пример отбраковки аномальных значений с использованием описанного метода представлен на рис.1. Данный метод может использоваться для сигналов, имеющих гладкую основу или особенности, представленные достаточным числом отсчетов.

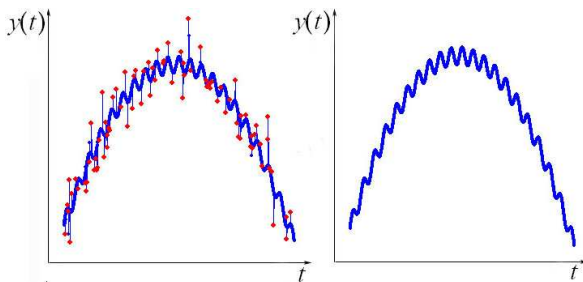


Рис.1 Результаты отбраковки аномальных значений.

После отбраковки аномальных данных возникает задача заполнения пропусков на местах найденных выбросов (если мы остаемся в рамках дискретных вейвлет-преобразований, а не используем лифтинговую схему). Для восстановления пропущенных данных можно использовать методы, основанные на стандартных процедурах интерполяции и аппроксимации (линейной, квадратичной, сплайновой) или на соответствующих методах заполнения пропусков в регрессионном анализе [3]. При дальнейшей очистке от шума (при формировании порогов) восстановленные на основе аппроксимации участки не учитываются, поскольку не содержат шум. Если же используется интерполяционная схема заполнения, то необходимо учитывать возникающие корреляционные связи.

Анализ, классификация, кластеризация и отбор участков для обработки данных. На данном этапе выполняются:

1) отбор участков на основе анализа погрешностей, который заключается в выделении погрешностей на основе вейвлет-преобразования, используя его известную разделяющую способность, и последующей классификации участков измерений на основе интервального значения среднеквадратического отклонения погрешностей, асимметрии распределения шума или эксцесса; проверки на гетероскедастичность и однородность, независимость и коррелированность;

2) отбор участков по характеристикам производных на основе вейвлет-преобразования;

3) классификация участков по режимным изменениям в данных, найденных путем анализа вейвлет-коэффициентов.

Разведочный анализ для адаптивного выбора способов и параметров обработки. Важное значение имеет этап разведочного анализа данных, который позволяет выбрать: наиболее подходящий метод очистки от шума, число уровней, длину фильтра, наилучший базис и подобрать различные параметры к конкретным методам (например, наилучшую длину блока для усечения вейвлет-коэффициентов при блочной обработке).

Разведочный анализ включает две подгруппы алгоритмов.

Первая подгруппа алгоритмов содержит способы адаптивного выбора базиса, зависящего от сигнала по критериям, определяющим стоимости аппроксимации сигнала в заданном базисе. После чего производится минимизация или максимизация (в зависимости от выбранного критерия, то есть выпуклости или вогнутости функции стоимости), тем самым обеспечивая поиск наилучшего базиса. Выбор наилучшего базиса осуществляется по следующим критериям: энтропийному, основанному на коэффициентах Шутца и Джини, мере Эмлена, L^p - энергии, критерию Аткинсона [4].

Вторая подгруппа основана на диагностике остатков, полученных при различных базисах, с помощью критериев независимости (Бокса-Пирса, Ли-Маклеода, Льюнга-Бокса, множителей Лагранжа) и нормальности (Д'Агостино-Пирсона, Жарку-Бэра, Гэри, Шапиро-Уилкса, Шапиро-Франка, Лилли-Форса) [5]. Наилучшему базису соответствует случай с наиболее независимыми остатками или ближайшими к нормальным.

Очистка от шума. В большинстве ситуаций, при работе с реальными данными модель сложного динамического процесса, как правило, неизвестна. Кроме того, многие процессы, с которыми приходится работать, характеризуются переменной степенью гладкости, то есть изменяющимся количеством непрерывных локальных производных вплоть до разрыва. Необходимо также учитывать и разнообразные статистические свойства шума (независимый или коррелированный, гомоскедастический или гетероскедастический). В таких условиях для очистки от шума прекрасно зарекомендовала себя вейвлет-технология, основанная на использовании различных типов базисных функций с компактным носителем и на различных схемах выбора порогов.

Этап очистки от шума содержит большое число процедур, специально ориентированных на очистку от шума: мультиполосный анализ, методы компенсации краевых эффектов, выбор типа и длины фильтра, различные методы вейвлет-преобразований, различные пороги и правила усечения вейвлет-коэффициентов.

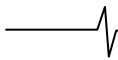
Общая схема очистки от шума включает следующие шаги:

1. Прямое вейвлет-преобразование для исходного сигнала $\{y_t : t = 0, \dots, N - 1\}$

$$W = w y,$$

где W - вейвлет-коэффициенты, w - матрица преобразования, N - длина выборки.

2. Процедура усечения вейвлет-коэффициентов: $W_{thr} = T(W)$, где T - оператор преобразования вейвлет-коэффициентов, зависящая от способа пороговой обработки и правила усечения.



3. Обратное вейвлет-преобразование $f_{оцен.} = w^T W_{thr}$, в результате которого получаем оценку истинной функции f .

Непосредственно перед вейвлет-преобразованием сигнала, для улучшения результатов очистки от шума на границах, рекомендуется использовать процедуру компенсации краевых эффектов. Искажение результатов очистки от шума на границах возникает из-за кругового сдвига в начале сигнала в результате прямого МНДВП, а в конце сигнала в результате обратного МНДВП. Компенсация краевых эффектов достигается с помощью дополнения данных слева и справа необходимым числом отсчетов, которое определяется в зависимости от длины фильтра и числа уровней. На j -ом уровне декомпозиции длина фильтра определяется по формуле $L_j = (2^j - 1)(L - 1) + 1$ [1], где L - исходная длина фильтра. Число отсчетов для дополнения на каждом уровне j должно равняться $L_j - 1$, поскольку именно столько вейвлет-коэффициентов находится под действием кругового сдвига. Дополнение данных выполняется с помощью прогнозирования как традиционным способом (симметрическое или периодическое), так и с помощью экстраполяции по оцененной на границе некоторой параметрической (например, полиномиальной) модели с адаптивно оцениваемой структурой и порядком. Добавленные отсчеты должны быть удалены после завершения процедуры очистки от шума.

Процедура усечения вейвлет-коэффициентов заключается в нахождении оценки вейвлет-коэффициентов, которые соответствуют сигналу с помощью оператора, ослабляющего шум, но сохраняющего сигнальную составляющую.

Можно выделить два основных подхода к вейвлет-очистке сигналов от шума: классический и байесовский. До недавнего времени классический подход включал различные линейные алгоритмы очистки от шума, а нелинейные процедуры были специализированными и сложными. В 90-х годах Donoho и Johnstone предложили использовать очень простой пороговый алгоритм, который в подходящем базисе может давать почти оптимальную нелинейную оценку. В связи с этим, стали появляться различные эффективные алгоритмы нелинейной пороговой обработки.

Линейные вейвлет-оценители относятся к классу проекционных оценителей, где проекционные операторы используют вейвлет-ядра. Эти операторы являются линейными относительно данных. В работах Wahba (1990), Antoniadis (1996) и Amato & Vuza (1997) описывается линейный подход, использующий методы сглаживающего сплайнового оценивания. В любом случае линейный подход используется для оценивания достаточно регулярных функций.

Для функций с низкой степенью регулярности обычно применяется нелинейная пороговая обработка (Donoho, Johnstone (1994)) или нелинейные методы усечения (сжатия). Классические методы нелинейной пороговой обработки эквиваленты оцениванию сигнала путем его усреднения с помощью ядра, которое адаптировано к гладкости сигнала. Пороговая обработка в общем случае заключается в модификации вейвлет-коэффициентов в соответствии с заданным типом порога и правилом усечения. Приравнение вейвлет-коэффициентов нулю эквивалентно локальному усреднению зашумленных данных, что имеет место только, если выделенный истинный сигнал окажется гладким. В настоящее время разработано большое количество различных порогов (универсальный, минимаксный, обобщенный порог перекрестной проверки, порог, рассчитанный на основе про-

верки множественных гипотез, порог на основе несмещенного риска Стейна, экспоненциальный растущий порог и др.) и правил усечения (жесткое, мягкое, гиперболическое, гаррота, полумягкое, сигмоидальное, гибридное).

В последнее время также стала популярной вейвлет-обработка данных с использованием байесовских правил (Vidacovic, Muller (1999)). Байесовские правила подобны обычным пороговым правилам и строятся по реальным статистическим моделям вейвлет-коэффициентов с учетом априорной информации об истинном сигнале. Большинство реальных байесовских правил легко вычислить путем моделирования или представить в аналитической форме.

Немаловажное значение имеет способ пороговой обработки (рис.2).

В случае глобальной пороговой обработки шумоподавление сигнала осуществляется на основании статистических характеристик каждого уровня разложения.

Локальная пороговая обработка обладает большей адаптивностью к исходным данным в сравнении с глобальной пороговой обработкой. При этом в случае общей пороговой обработки общий порог применяется ко всем вейвлет-коэффициентам различных уровней. При многоуровневой обработке для каждого уровня рассчитывается свой порог, а при поэлементной обработке - для каждого вейвлет-коэффициента.

Блочная пороговая обработка сочетает в себе основные свойства глобального и локального способов обработки, поскольку предполагает применение того или иного оператора к блокам соседних коэффициентов детализации.

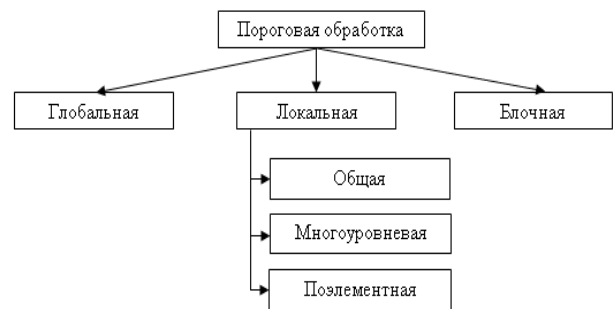


Рис. 2. Способы пороговой обработки.

Существенное влияние на очистку от шума оказывает выбранный базис. В зависимости от вида сигнала в качестве базисной функции для вейвлет-преобразования могут использоваться следующие типы фильтров: Daubechies, Symmlet, Coiflet, Haar, Сплайн, Ojanen, Bathlet, Legendre, Remez, Maxflat и др. Выбор наилучшего базиса для конкретного сигнала осуществляется с помощью разведочного анализа, описанного выше. В случае использования МНДВП выбор вейвлет-фильтра менее критичен.

Мультиполосный анализ. Произвольный входной сигнал можно рассматривать в виде суммы разнотипных составляющих: тренда, различных циклических компонент, флуктуаций – мелкомасштабных колебаний вокруг составляющих, аномалий - резких изменений, локальных особенностей. Инструментом разделения сигналов на такие составляющие является мультиполосный анализ. С его помощью можно определить поведение флуктуаций на разных масштабах, выявить причины их происхождения, удалить ненужные циклические компоненты.

Мультиполосный анализ позволяет посмотреть результаты разделения сигнала на составляющие как во временной, так и в вейвлет-области. Преимуществом представления сигнала по частотным полосам во временной области является то, что мы получаем проекции на детализирующее и аппроксимирующее пространство для различных частотных полос. Их легче анализировать, поскольку в вейвлет-области за счет формы используемого вейвлет-фильтра частотные составляющие, соответствующие различным полосам, сдвигаются, и общая картина становится менее понятной.

Модификации алгоритмов очистки от шума. В реальных ситуациях при обработке данных нередко сталкиваются с изменяющимися в процессе измерений условиями наблюдений. Это приводит к тому, что на различных сегментах наблюдений погрешности измерений принадлежат нормальному распределению, но с различной дисперсией, то есть мы имеем дело с так называемыми гетероскедастическими моделями. В этом случае методы очистки данных от шума, основанные на стандартной вейвлет-технологии, не дают хороших результатов. Для решения этой проблемы (для случая кусочно-постоянного уровня шума) был разработан двухступенчатый способ очистки от шума [6], где на первом этапе на основе кластеризации вейвлет-коэффициентов определяются все параметры модели погрешности с кусочно-постоянной дисперсией, а затем на втором этапе производится расчет модифицированных порогов для вейвлет-коэффициентов по оцененным кластерам. При этом для кластеризации вейвлет-коэффициентов используется алгоритм, основанный на сравнении дисперсий вычисленных погрешностей с применением F-распределения Фишера и дальнейшего уточнения границ. Преимущества этого алгоритма состоят в том, что, во-первых, число кластеров априорно не задается, а автоматически определяется в процессе кластеризации, во-вторых, алгоритм обеспечивает высокую точность разделения выборки на кластеры за счет разработанной процедуры уточнения границ. Пример результатов очистки от шума показан на рис.3.

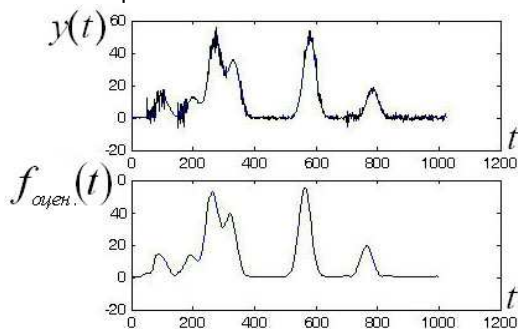


Рис.3. Результаты очистки от шума сигнала при наличии гетероскедастических погрешностей.

Среди реальных сигналов нередко встречаются сигналы с резкими пиками, так называемые мультипиковые процессы. При очистке от шума таких сигналов стандартными методами информация о пиках искажается, и в итоге мы получаем совершенно другой тип сигналов, чем хотелось. Шумовая составляющая сосредоточена в основном на самых тонких уровнях вейвлет-декомпозиции. При увеличении числа уровней скорость спада вейвлет-коэффициентов зависит от гладкости исходного сигнала. Таким образом, если пик в исходном сигнале представляется большим количеством отсчетов, то в вейвлет-области на тонких уровнях соответствующие ему вейвлет-коэффициенты будут на порядок меньше шума,

то есть пик обнаружить нельзя. С увеличением числа уровней картина меняется, шумовые коэффициенты уменьшаются, а резкие изменения сигнала выделяются более отчетливо. Поэтому была разработана модификация, включающая в себя этап обнаружения резких пиков (с помощью МНДВП и анализа максимальных кривизн) и алгоритм очистки от шума, основанный на сегментной очистке с фиксированными узлами. При этом практически полностью гасится эффект Гиббса. Примеры результатов очистки от шума показаны на рис.4.

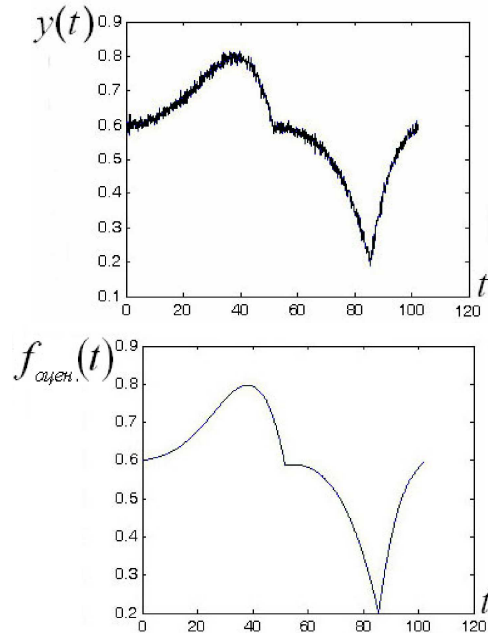


Рис.4. Результаты очистки от шума сигнала с резкими изменениями.

Анализ и интерпретация результатов обработки.

Оценка результатов обработки осуществляется на основе диагностики остатков с использованием критериев независимости и нормальности.

Заключение

На основе описанной вейвлет-технологии был создан программный комплекс в среде MatLab. Разработанные методы прошли апробацию в реальных условиях и показали более высокую эффективность по сравнению с ранее существующими традиционными вейвлет-методами очистки от шума.

Литература

1. Percival D., Walden A. Wavelet methods for time series analysis // London: Cambridge University Press. 2000. P. 594.
2. Орешко Н. И., Князева Т.Н. Автоматизированный поиск выбросов с использованием вейвлет-преобразования // Труды Российского НТОРЭС им. А.С. Попова. Серия "Цифровая обработка сигналов и ее применение", Выпуск: X-1, Москва, 2008, стр. 145-149.
3. Rao C. R., Toutenburg H. Linear models: least squares and alternatives // New York: Springer-Verlag. 1999. P. 426.
4. Маршалл А., Олкин И. Неравенства: теория мажоризации и ее приложения // М.: Мир. 1983. 574 с.
5. Henry C. Thode., Jr. Testing for normality // New York: State University of New York at Stony Brook. 2002, P. 479.
6. Орешко Н.И., Князева Т.Н. "Очистка от шума траекторных данных при наличии гетероскедастических погрешностей" // СПб: Известия СПбГЭТУ "ЛЭТИ". Сер. "Информатика, управление и компьютерные технологии". 2007. Специальный вып. 2. С. 31-41.

СОРТИРОВКА НЕЙРОННЫХ СПАЙКОВ НА ОСНОВЕ ПАРАМЕТРИЧЕСКОГО ВЕЙВЛЕТ-АНАЛИЗА С АДАПТИВНОЙ ФИЛЬТРАЦИЕЙ

Макаров В.А., Павлов А.Н., Туницын А.Н.

Введение

Центральная нервная система (ЦНС) живых организмов непрерывно получает и обрабатывает сенсорную информацию, поступающую посредством взаимодействия с внешним миром. Исследование механизмов кодирования, представления и обработки данной информации ЦНС является одной из актуальнейших задач современного естествознания. Оптические, звуковые, тактильные стимулы кодируются соответствующими рецепторами в последовательности электрических импульсов (потенциалов действия или спайков), которые могут быть зарегистрированы экспериментально «*in-vivo*» с помощью микроэлектрода, погруженного в нервную ткань [1]. При анализе кооперативной динамики нейронных ансамблей и изучении генерируемого нейронами информационного кода спайки рассматриваются в качестве стереотипных событий. Это означает, что важна не форма отдельных импульсов, а время их генерации нейронами. Следует отметить то обстоятельство, что при осуществлении внеклеточной записи электрического потенциала микроэлектрод регистрирует сигнал не только от одной клетки, вблизи которой он находится, но и от соседних нейронов, расположенных в некоторой локальной области. В результате, полученные экспериментальные данные представляют собой суперпозицию электрической активности некоторого нейронного ансамбля (рис. 1). Поэтому исследователь должен установить связь отдельных потенциалов действия с различными клетками, генерирующими соответствующие сигналы. Эта задача называется сортировкой или классификацией спайков [2].

При решении задачи классификации традиционно считается, что каждый нейрон генерирует спайки одной и той же формы и амплитуды, в то же время сигналы разных клеток имеют индивидуальные особенности (хотя их потенциалы действия могут быть очень похожими). Проводя сравнение потенциалов действия по форме, можно разделить спайки соседних нейронов с определенной степенью точности. На практике данная задача является чрезвычайно сложной из-за наличия существенного фонового шума, который вносит искажения, усложняя идентификацию сигналов разных клеток; наличия вариабельности формы потенциалов действия, генерируемых нейронами; слабых различий форм спайков соседних клеток и т.д. В частности, на рис. 1 видны, по крайней мере, две группы спайков (с меньшей - А и большей - В амплитудами). Для того чтобы провести выделение сигналов только одного нейрона, в общем случае необходимо применять специальные методы анализа данных.

Предлагается новый метод сортировки нейронных спайков – параметрический вейвлет-анализ с адаптивной фильтрацией, который обеспечивает ошибку разделения потенциалов действия, близкую к теоретическому минимуму. Продемонстрирована зависимость качества сортировки нейронных спайков на основе вейвлет-преобразования от статистики экспериментального шума. Возможности предлагаемого подхода иллюстрируются как на тестовых, так и на реальных электрофизиологических сигналах.

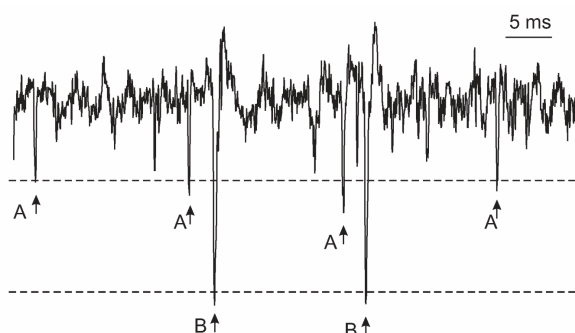


Рис. 1. Пример экспериментальной записи внеклеточного потенциала. Спайки достаточно большой (по сравнению с шумом) амплитуды обозначены стрелками. Наиболее простой способ разделения спайков – пороговая (амплитудная) сортировка. Спайки близкие по амплитуде ассоциируются с одним источником (нейроном).

В настоящее время методы сортировки спайков включают 4 независимых этапа, а именно: 1) фильтрацию экспериментальной записи внеклеточного сигнала; 2) определение всех спайков, которые требуется отсортировать (потенциалов действия, превышающих уровень фонового шума); 3) нахождение индивидуальных особенностей в формах потенциалов действия; 4) разделение спайков на группы (кластеры) в пространстве характеристик и установление соответствия полученных групп с различными нейронами. Наиболее сложными и неоднозначными из этих этапов являются первый и третий.

В рамках простейшего, но достаточно часто используемого на практике подхода, фильтрация внеклеточного потенциала осуществляется с помощью фильтра высоких частот (ФВЧ) с частотой среза примерно 300 Гц, после чего проводится визуальный анализ форм спайков в сочетании с различными вариантами пороговой (амплитудной) сортировки (рис. 1 спайки типа А и В). Данная процедура является субъективной, трудоемкой и зачастую неэффективной (в рамках типичного эксперимента можно зарегистрировать более 10^4 спайков). Сортировать спайки по амплитуде уда-

ется лишь при условии, что микроэлектрод расположен очень близко к какому-то одному нейрону – амплитуда генерируемых им сигналов будет существенно больше чем у других клеток, и в этом случае можно надежно отделить, по крайней мере, одну группу спайков. Если амплитуда спайков, генерируемых разными нейронами, является сопоставимой, то пороговое детектирование не позволяет обеспечить качественное разделение сигналов нервных клеток.

Значительно более «мощным» методом является анализ главных компонент (АГК) [2]. В рамках этого подхода также проводится фильтрация внеклеточного электрического потенциала (ФВЧ, 300 Гц), и затем масштабные коэффициенты нескольких первых главных компонент используются в качестве характеристик формы спайков для последующей сортировки. Недавно для решения задачи сортировки был предложен метод, основанный на расчете коэффициентов вейвлет-преобразования (ВП) [3]. В соответствии с работами [3-6], использование теории вейвлетов может быть эффективнее применения классического алгоритма АГК. Однако разделение спайков на основе коэффициентов ВП имеет ряд недостатков, среди которых следует отметить произвольность выбора базисного вейвлета и сложность автоматического нахождения коэффициентов, отражающих наиболее существенные различия форм потенциалов действия разных нейронов [13].

В рамках всех вышеупомянутых методов на первом этапе проводится стандартная процедура фильтрации с применением ФВЧ, которая не учитывает ни статистику шума, присутствующего в конкретной записи внеклеточного электрического потенциала, ни особенности формы анализируемых спайков. Шум, неизбежно присутствующий при детектировании сигнала микроэлектродом, имеет различную природу: от флуктуаций в самих электронных приборах до физиологических процессов, происходящих в клетках [7], и слабых смещений самого электрода [8,9]. Характеристики шума могут варьироваться от одного эксперимента к другому. Такие классические методы, как пороговое детектирование и анализ главных компонент, известны уже давно, и для них существуют рекомендации по подбору параметров фильтров. Этого нельзя сказать о методе, основанном на вейвлет-преобразовании. В данной работе мы показываем, что эффективность вейвлетного метода можно существенно улучшить путем включения адаптивной фильтрации непосредственно в процесс выявления основных особенностей формы спайков. Основываясь на этом заключении, далее мы предлагаем новый метод классификации нейронных потенциалов действия – параметрический вейвлет-анализ с адаптивной фильтрацией (ПВАФ).

Влияние статистики шума на качество сортировки нейронных спайков

Для демонстрации влияния статистики фонового шума на качество классификации потенциалов действия был проведен следующий эксперимент. Из реальных записей внеклеточного потенциала были выбраны две различные формы спайков, с помощью которых на следующем шаге формировались последовательности, содержащие по 1000 потенциалов действия каждая. Для моделирования эффекта присутствия фонового шума осуществлялось добавление к каждой получившейся последовательности случайного процесса с определенной центральной частотой и шириной

полосы $\Delta f=700$ Гц (полученного путем полосовой фильтрации Пуассоновского процесса). Следующим этапом являлся анализ смоделированных данных и проведение сортировки различными методами для оценки их погрешности (отношения числа ошибочно классифицированных спайков к общему числу анализируемых потенциалов действия).

На рис. 2,а показаны зависимости величины ошибки сортировки от центральной частоты добавляемого цветного шума. Как следует из этого рисунка, метод АГК приводит к большим погрешностям в области низких частот или для "медленного" (по сравнению с динамикой нейрона) шума, однако эффективность этого метода растет с ростом центральной частоты шума. Напротив, применение метода, использующего вейвлет-преобразование, целесообразно при доминировании «медленного» шума (в окрестности 2 кГц для рассматриваемого примера наблюдается значительный рост ошибки классификации). В данном случае мы рассматриваем так называемый «Вейвлетный Классификатор Спайков» (ВКС) [3], результаты которого качественно соответствуют другим алгоритмам, использующим теорию вейвлетов [6]. Таким образом, статистика шума (в частности его центральная частота) существенно влияет на погрешность сортировки, однако ее эффект отличается для методов АГК и вейвлет-преобразования. Таким образом, адекватная предварительная фильтрация внеклеточного потенциала может улучшить результаты классификации спайков.

На рис. 2,б показаны зависимости ошибки классификации от частоты среза фильтра низких частот (ФНЧ), примененного к анализируемым данным. Наглядно видно, что в случае вейвлет-преобразования ошибка классификации минимальна, когда частота среза фильтра составляет примерно 2.2 кГц и далее с ростом частоты она быстро возрастает. Для метода анализа главных компонент при увеличении частоты фильтра ошибка постепенно уменьшается и выше 2 кГц практически остается постоянной. Следовательно, низкочастотная фильтрация слабо влияет на метод АГК и оказывает значительно более сильное влияние на вейвлет-анализ, в рамках которого совершенно необходимо обеспечить подходящий выбор частоты среза фильтра.

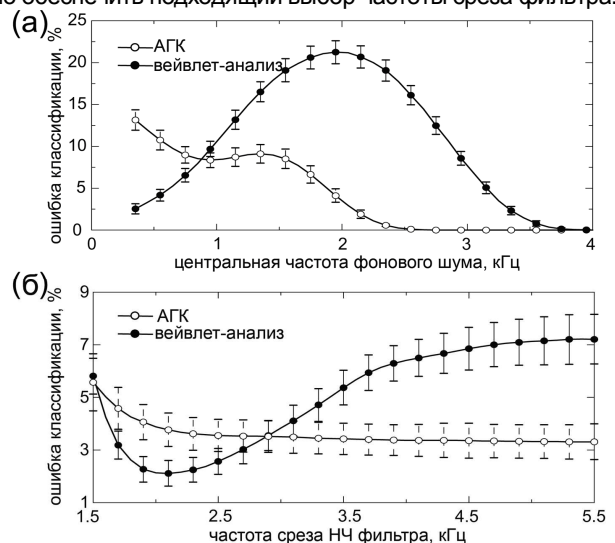
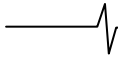


Рис. 2. Зависимость величины ошибки сортировки от (а) центральной частоты фонового шума (шум в полосе частот 700 Гц) и (б) частоты среза ФНЧ, применяемого для предварительной фильтрации внеклеточного потенциала.



Теоретические основы предлагаемого метода сортировки

Рассмотрим задачу сортировки спайков с помощью вейвлет-анализа более подробно. Предположим, что проводится исследование сигнала, содержащего $(N+M)$ спайков, соответствующих динамике двух разных нейронов А и В. Обозначим первоначальные (в отсутствие шума) формы спайков как $w_A(t)$ и $w_B(t)$. Следовательно, для регистрируемых в ходе эксперимента потенциалов действия можно записать следующее выражение:

$$s_i(t) = \xi_i(t) + \begin{cases} w_A(t), & i = 1, 2, \dots, N \\ w_B(t), & i = N + 1, \dots, N + M \end{cases} \quad (1)$$

где $\xi_i(t)$ – некоррелированные источники цветного шума, которые в первом приближении имеют одинаковую статистику. В наиболее общем виде непрерывное вейвлет-преобразование спайков $s_i(t)$ можно записать следующим образом:

$$W_i(a, b) = \frac{1}{\sqrt{a}} \int_0^T s_i(t) \psi_{a,b}(t) dt \quad (2)$$

где T – длительность спайка (обычно 1-3 мс),

$\psi_{a,b}(t) = \psi\left(\frac{t-b}{a}\right)$ – перемасштабируемая и

смещенная вдоль оси времени базисная функция (вейвлет), b и a – параметры временной локализации и масштаба, соответственно.

Применяя преобразование (2) с произвольными фиксированными значениями параметров (a, b) к спайкам (1), получим:

$$W_i(a, b) = \eta_i + \begin{cases} W_A, & i = 1, 2, \dots, N \\ W_B, & i = N + 1, \dots, N + M, \end{cases} \quad (3)$$

где

$$\eta_i(a, b) = \frac{1}{\sqrt{a}} \int_0^T \xi_i \psi_{a,b} dt, \quad (4)$$

$$W_{A,B}(a, b) = \frac{1}{\sqrt{a}} \int_0^T w_{A,B} \psi_{a,b} dt. \quad (5)$$

Появление первого слагаемого в правой части уравнения (3) связано с наличием экспериментального шума. Второе слагаемое – коэффициенты вейвлет-преобразования, несущие информацию о форме каждого спайка (без учета шума).

Полученные коэффициенты W_i могут быть использованы для сортировки спайков. При этом важно выбрать такие параметры (a, b) , которые позволят разделить спайки по группам (кластерам) с наименьшей ошибкой. Успешно решить данную задачу можно при условии, что распределение коэффициентов W_i является бимодальным. В этом случае можно установить пороговое значение $W_{th} \in [W_A, W_B]$, считая, что спайки, для которых $W_i < W_{th}$, соответствуют нейрону А, а остальные – нейрону В (рис. 3а).

Предположим, что измерительный шум имеет распределение, близкое к нормальному [10] со стандартным отклонением σ . Основанием для такого предположения служит центральная предельная теорема. Обозначим половину расстояния между незашумленными спайками в пространстве вейвлет-коэффициентов символом \bar{W} :

$$\bar{W} = \frac{W_B - W_A}{2} = \frac{1}{2\sqrt{a}} \int_0^T (w_B - w_A) \psi_{a,b} dt. \quad (6)$$

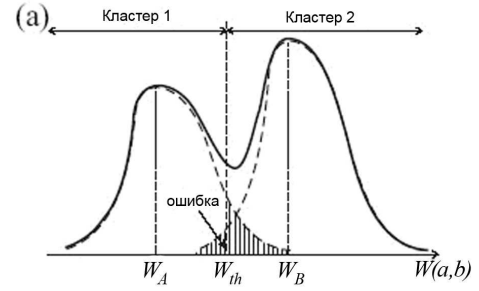


Рис. 3. (а) Пример распределения вейвлет-коэффициентов W_i для набора зашумленных спайков двух нейронов. Пунктирные линии обозначают распределения для одиночных нейронов, (б) Минимальная величина ошибки как функция параметра разделения кластеров для нескольких значений отношения количества спайков генерируемых нейронами γ .

Без потери общности можно считать, что $W_B = -W_A = \bar{W}$. Тогда распределение коэффициентов W_i будет выглядеть следующим образом:

$$h = \frac{M}{\sqrt{2\pi}\sigma} \left(\gamma e^{-\frac{(W+\bar{W})^2}{2\sigma^2}} + e^{-\frac{(W-\bar{W})^2}{2\sigma^2}} \right), \quad (7)$$

где $\gamma = N/M$ – отношение числа спайков двух нейронов.

Минимальное число ошибочно идентифицированных спайков достигается при:

$$W_{th} = \frac{\sigma^2}{2\bar{W}} \ln \gamma. \quad (8)$$

Отметим, что оптимальная величина порога W_{th} в общем случае ($\gamma \neq 1$) не совпадает с минимумом распределения (рис. 3а). Теоретически минимальная величина ошибки составляет:

$$R_{\min} = \frac{\gamma \operatorname{erfc}\left(\Delta + \frac{\ln \gamma}{4\Delta}\right) + \operatorname{erfc}\left(\Delta - \frac{\ln \gamma}{4\Delta}\right)}{2(1 + \gamma)}, \quad (9)$$

где erfc – функция ошибок, $\Delta = \frac{\bar{W}}{\sqrt{2}\sigma}$ – параметр разделяемости кластеров.

Таким образом, минимальная ошибка классификации представляет собой функцию, зависящую от двух величин – γ и Δ , и уменьшающуюся с ростом Δ (рис. 3б). В эксперименте γ является фиксированной величиной, и мы можем варьировать лишь значение Δ . Путем подбора параметров (a, b) вейвлет-преобразования можно достигнуть максимума \bar{W} , что в свою очередь

(при постоянном значении σ) приводит к увеличению Δ . Однако, как мы покажем далее, параметр масштаба a нетривиальным образом влияет на Δ .

Экспериментальный шум $\xi(t)$ в некоторой ограниченной полосе частот Ω_{noise} может быть представлен суммой гармоник:

$$\xi_i = \sum_{\Omega_{noise}} A(\omega_k) \cos(\omega_k t + \phi_{ki}), \quad (10)$$

где ω_k и ϕ_{ki} - частота и фаза соответствующей гармоники, $A(\omega_k)$ определяет спектр мощности шума. Используя вейвлет Хаара [6], можно получить следующие коэффициенты вейвлет-преобразования экспериментального шума (10):

$$\eta_i = -\frac{4}{\sqrt{a}} \sum_k \frac{A(\omega_k)}{\omega_k} \sin \phi_{ki} \sin^2 \frac{a\omega_k}{4}. \quad (11)$$

Таким образом, статистические свойства случайного процесса η не зависят от параметра локализации b , и стандартное отклонение для измерительного шума имеет вид:

$$\sigma^2(a, \Omega_{noise}) = \frac{8}{a} \sum_k \frac{A^2(\omega_k)}{\omega_k^2} \sin^4 \frac{a\omega_k}{4}. \quad (12)$$

Соответственно параметр разделяемости кластеров Δ нетривиальным образом зависит от параметра a , формы спайков и спектральных характеристик экспериментального шума. Одним из способов изменения спектра шума является фильтрация сигнала. Обозначив через f_c частоту среза НЧ-фильтра, мы можем свести задачу улучшения качества классификации спайков к поиску оптимального набора трех параметров (a, b, f_c) , максимизирующих величину Δ и, следовательно, минимизирующих ошибку классификации (рис. 3б).

Отметим, что рассматриваемая постановка задачи является более общей, чем для стандартных методов, основанных на поиске оптимальных коэффициентов вейвлет-преобразования. Учитывая особенности шума в каждом конкретном эксперименте и проводя соответствующую фильтрацию данных, в рамках предлагаемого параметрического подхода можно добиться более качественной сортировки потенциалов действия. Известные методы, основанные на расчете вейвлет-преобразования, обычно предполагают простое сравнение экспериментальных распределений коэффициентов, тогда как наш подход является параметрическим.

Практическая реализация метода ПВАФ и тестовые примеры

Для оптимизации разделяемости кластеров Δ необходимо оценить вейвлет коэффициент \bar{W} и спектр шума $A^2(\omega)$. Однако, при проведении анализа экспериментальных сигналов нам неизвестна ни форма незашумленных спайков, ни спектр шума. Для получения этих сведений предлагается следующий алгоритм:

1. *Получение формы незашумленных спайков.* Применяя стандартный метод сортировки спайков, например анализ главных компонент, можно найти максимумы функции плотности распределения масштабных коэффициентов первых главных компонент

и средние формы спайков в окрестности каждого максимума. Эти средние формы могут рассматриваться в качестве $w_{A,B}(t)$.

2. *Нахождение спектра шума $P(\omega)$.* В простейшем случае (относительно низкой частоте спайков), в качестве аппроксимации $P(\omega)$ можно рассмотреть общий спектр всего внеклеточного сигнала.

3. *Нахождение оптимального набора параметров (a^*, b^*, f_c^*) , максимизирующих величину Δ .* Используя результаты 1 и 2 шагов, необходимо: а) провести оценку \bar{W} ; б) вычислить $A^2(\omega) = P(\omega)H^2(\omega)$, где $H(\omega)$ - характеристика фильтра, в) провести оценку величины параметра разделения кластеров $\Delta = \bar{W} / \sqrt{2}\sigma$. Найти максимум функции $\Delta(a, b, f_c)$.

4. *Фильтрация спайков с частотой среза фильтра f_c^* и расчет $W_i(a^*, b^*)$ для всех спайков.*

5. *Классификация спайков в соответствии с полученными коэффициентами W_i .*

Предлагаемый метод может быть особенно эффективен для больших массивов данных. Он также позволяет использовать сразу несколько наборов характерных особенностей спайков, по которым их можно разделить. При наличии спайков более двух нейронов метод позволяет улучшить разрешение между любыми двумя перекрывающимися «облаками» точек в пространстве характеристик стандартных алгоритмов.

На рис. 4 показан пример сортировки спайков различными методами. Тестовые данные содержали 1300 спайков типа А и 1000 спайков типа В (рис. 4,а). На рис. 4,б представлены гистограммы характеристик спайков, вычисленных с помощью АГК и стандартного вейвлет-метода (ВКС). Метод анализа главных компонент демонстрирует сильное перекрытие распределений характеристик двух типов спайков и величину ошибки $\approx 5.5\%$. Для более качественного решения задачи классификации с применением вейвлет-анализа было проведено исследование всех пар (a, b) для снижения величины ошибки до минимума. При этом метод ВКС показал сопоставимый с АГК результат – величина ошибки равнялась $\approx 4.5\%$. Использование метода ПВАФ включает этап оценки параметра разделяемости кластеров Δ . Видно, что данный параметр имеет ярко выраженный пик на частоте $f_c = 100$ Гц (рис. 4,в). При фильтрации данных с такой частотой разница между спайками в пространстве вейвлет-коэффициентов идентифицируется наиболее четко. Таким образом, предложенный в нашей работе параметрический метод позволил снизить величину ошибки до $\approx 0.7\%$ (рис. 4,г), что приближается к теоретическому минимуму. Это подтверждает гипотезу о том, что для методов вейвлет-анализа фильтрация играет очень важную роль, и подбором характеристик фильтра непосредственно под конкретные формы нейронных спайков можно существенно повысить качество решения задачи сортировки. Отметим еще раз, что в отличие от классических подходов мы рассматриваем частоту среза фильтра как наиболее важный параметр, который необходимо подстраивать индивидуально для каждого экспериментального сигнала внеклеточного потенциала.

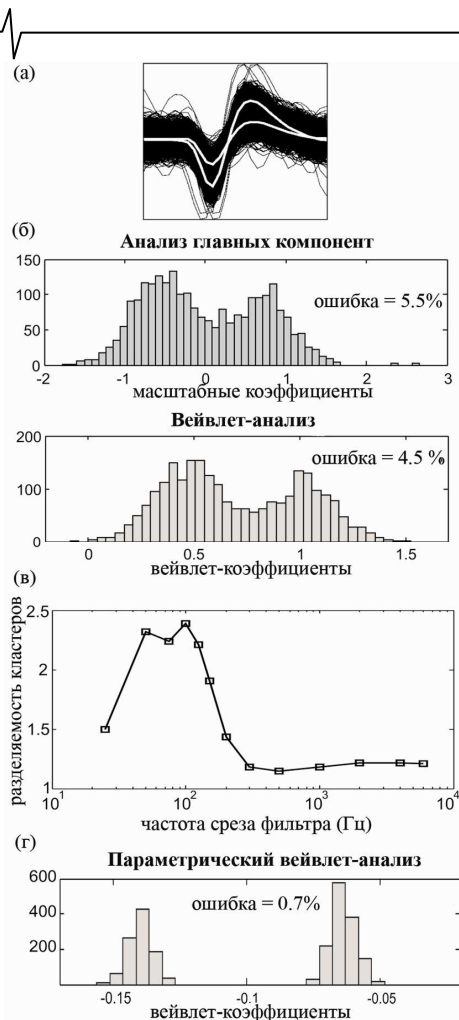


Рис. 4. Тестовый пример классификации спайков различными методами. (а) Суперпозиция всех экспериментальных спайков (1300 типа А и 1000 типа В). Белые линии – формы незашумленных спайков. (б) Гистограммы вычисленных характеристик спайков для метода анализа главных компонент и обычного (ВКС) вейвлет-анализа. (в) Зависимость параметра разделяемости кластеров от частоты среза фильтра низких частот. Для максимального значения $\Delta = 2.39$ теоретический минимум величины ошибки составляет 0.3%. (г) Гистограмма вейвлетных коэффициентов для метода ПВАФ, который позволяет существенно снизить ошибку классификации, приближаясь к теоретическому минимуму.

Далее аналогичная тестовая процедура была проведена для сортировки спайков, которые имели различия на малых временных масштабах, т.е. для случая, когда вейвлет-анализ оказывается предпочтительнее метода АГК [11]. При этом мы снова убедились в том, что алгоритм ПВАФ демонстрирует значительно более качественную классификацию по сравнению с методами АГК и ВКС (таб. 1).

Наряду с тестовыми данными были проанализированы реальные электрофизиологические записи внеклеточного потенциала. Они были сделаны с помощью тетродных микроэлектродов, дизайн которых позволял осуществлять запись сигналов одного и того же нейрона по двум каналам [12]. В отдельных, достаточно редких, случаях два канала наряду со значительной активностью большого числа разных клеток могли демонстрировать хорошо различимые спайки одного и того же нейрона. Среди значительного количества экспериментальных записей мы отобрали 2 сигнала, где лучше всего выполнялись эти условия. Для отобранных экспериментальных данных проводилась сортировка спайков «вручную» с высокой надежностью, достигаемой рассмотрением двух каналов одновременно, что позволяло успешно решать задачу идентификации в «спорных» случаях. Затем, используя эту информацию для контроля работы различных методов, мы оценивали их ошибки идентификации (анализируя записи только одного канала). Результаты, представленные в таблице 1, свидетельствуют в пользу предлагаемого подхода ПВАФ.

Заключение

Разработка алгоритмов оптимальной классификации нейронных спайков является важной задачей с точки зрения поиска решения проблемы кодирования и обработки информации живыми системами, поскольку без проведения надежной сортировки сигналов различных клеток крайне осложняются все последующие исследования информационного кода, генерируемого малыми нейронными ансамблями. Традиционные методы решения задачи разделения спайков, такие как пороговое детектирование, анализ геометрических особенностей и алгоритм АГК, имеют ряд значительных недостатков, в частности, связанных с большими ошибками идентификации форм потенциалов действия при высоком уровне фонового шума. Применение теории вейвлет-преобразования для сортировки спайков является одним из новых направлений исследований, где в данный момент существует значительное количество открытых вопросов. Вейвлеты не всегда обеспечивают более качественную идентификацию и сортировку спайковых сигналов, и чтобы пользоваться данным инструментом, необходимо вначале детально изучить его возможности и ограничения [11,13].

Мы показали, что эффективность методов классификации в значительной степени зависит от статистики экспериментального шума. Традиционный подход к фильтрации внеклеточных сигналов состоит в использовании фиксированных значений параметров фильтра без учета особенностей формы сигналов и характеристик шума. В ходе проведенных исследований нами предложен новый параметрический метод сортировки нейронных спайков (параметрический вейвлет-анализ с адаптивной фильтрацией).

Таблица 1.

Величина ошибки классификации для разных методов сортировки спайков.

Вид данных	Анализ главных компонент (АГК)	Стандартный вейвлет-анализ (ВКС)	Параметрический вейвлет-анализ (ПВАФ)
Тестовые №1	5.5 %	4.5 %	0.7 %
Тестовые №2	28.0 %	5.5 %	1.7 %
Экспериментальные №1	11.1 %	7.0 %	3.4 %
Экспериментальные №2	12.2 %	10.1 %	6.8 %

Принципиальной особенностью метода является подстройка характеристик фильтра под индивидуальные особенности формы потенциалов действия. Новый алгоритм включает процедуру фильтрации непосредственно в качестве составной части методики выбора оптимальных параметров для разделения спайков по группам.

На основе тестовых исследований и реальных экспериментальных данных продемонстрировано преимущество разработанного подхода по сравнению с известными методами, такими как АГК и ВКС. Установлено, что предлагаемый нами подход позволяет снизить ошибку разделения потенциалов действия до значения, близкого к теоретическому минимуму.

Проводимые исследования были поддержаны Министерством Образования и Науки РФ по программе «Развитие научного потенциала высшей школы (2006-2008 гг.)».

Литература

1. Kandel E.R., Schwartz J.H., and Jessell T.M. Principles of Neural Science. (4 ed.). McGrawHill, New York, 2000.
2. Lewicki M. A review of methods for spike sorting: the detection and classification of neural potentials // Net. Com. Neu. Sys. 1998. Vol. 9. P. R53.
3. Letelier J., and Weber P. Spike sorting based on discrete wavelet transform coefficients // J. Neurosci. Methods. 2000. Vol. 101. P. 93.
4. Hulata E., Segev R., Shapira Y., Benveniste M., and Ben-Jacob E. Detection and sorting of neural spikes using wavelet packets // Phys. Rev. Lett. 2000. Vol. 85. P. 4637.
5. Hulata E., Segev R., and Ben-Jacob E. A method for spike sorting and detection based on wavelet packets and Shannon's mutual information // J. Neurosci. Methods. 2002. V. 117. P. 1.
6. Quian Quiroga R., Nadasdy Z., and Ben-Shaul Y. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering // Neural Computation. 2004. Vol. 16. P. 1661.
7. Fee M.S., Mitra P.P., and Kleinfeld D. Variability of extracellular spike waveforms of cortical neurons // J. Neurophysiol. 1996. Vol. 76. P. 3823.
8. Snider R.K., and Bonds A.B. Classification of non-stationary neural signals // J. Neurosci. Methods. 1998. Vol. 84. P. 155.
9. Pouzat C., Mazor O., Laurent G. Using noise signature to optimize spike-sorting and to assess neuronal classification quality // J. Neurosci. Methods. 2002. Vol. 122. P. 43.
10. Gardiner C.W. Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences. (2 ed.) Springer-Verlag, Berlin, Heidelberg, 1990.
11. Думский Д.В., Павлов А.Н., Тупицын А.Н., Макаров В.А. Классификация нейронных потенциалов действия с помощью вейвлет-преобразования // Изв. вузов, Прикладная нелинейная динамика. 2005. No. 5-6. С. 77.
12. Harris K., Henze D., Csicsvari J., Hirase H., Buzsaki G. Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements // J. Neurophysiol. 2000. Vol. 84. P. 401.
13. Pavlov A.N., Makarov V.A., Makarova I., Panetsos F., Sorting of neural spikes: When wavelet based methods outperform principal component analysis // Natural Computing (Springer), vol. 6, pp. 269-281 (2007).

Уважаемые коллеги!

С первого полугодия 2009 года вы можете подписаться на наш журнал в любом отделении связи. Подписной индекс издания по каталогу ОАО «Роспечать» - 82185.

Для тех, кто не успел оформить подписку на второе полугодие 2008 года через ОАО «Роспечать», сохраняется возможность приобретения журналов непосредственно в редакции по адресу: 107031, г. Москва, Рождественка, 6/9\20, стр. 1, Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова, или оформить Заказ в соответствии с требованиями, выставленными на сайте журнала: www.dsps.ru.

Справки по телефонам: (495) 621-71-08, 621-06-10.

Факс: (495) 621-16-39.

E-mail: nto.popov@mtu-net.ru

tor@rgta.ryazan.ru

ЦИФРОВАЯ ФИЛЬТРАЦИЯ СИГНАЛОВ В ПРОМЫШЛЕННЫХ СИСТЕМАХ УПРАВЛЕНИЯ

Прохоренков А.М., Качала Н.М.

Введение

Постановка задачи

В информационно-измерительных комплексах автоматизированных систем управления технологическими процессами (АСУ ТП) приходится решать задачи измерения параметров зашумленных сигналов по одной реализации в реальном времени. Для выделения полезного сигнала на фоне помехи применяют различные алгоритмы фильтрации. Однако применение многих из них в измерительных системах АСУ ТП ограничено, поскольку особенностями данных систем являются разнообразие детерминированных составляющих наблюдаемых сигналов, изменение во времени принятых для расчета фильтров исходных предпосылок и отсутствие достаточного количества информации о свойствах объектов управления и условиях их функционирования.

В этой ситуации актуальной становится задача синтеза алгоритмов обработки случайных процессов, позволяющих найти наилучшую оценку полезной детерминированной компоненты на основании измерения наблюдаемого сигнала.

Процедуре синтеза новых алгоритмов всегда предшествует анализ применимости известных методов для решения той или иной задачи. Наиболее широкое применение для решения задач фильтрации в различных областях научных исследований и прикладных сферах нашли фильтры Калмана. На практике также часто используется простой по структуре фильтр скользящего среднего. Необходимо с учетом особенностей промышленных систем управления исследовать указанные методы фильтрации для включения их в процедуру обработки случайных процессов.

Случайные процессы, протекающие в объектах систем управления, можно представить как результат совместного действия детерминированного полезного сигнала и стационарной помехи. Различают следующие модели сигналов [1]:

$$\text{аддитивная модель } y(t) = x_1(t) + v(t); \quad (1)$$

$$\text{мультипликативная модель } y(t) = x_2(t)v(t); \quad (2)$$

$$\text{аддитивно-мультипликативная модель } y(t) = x_1(t) + x_2(t)v(t), \quad (3)$$

где $x_1(t)$, $x_2(t)$ – детерминированные функции времени, $v(t)$ – стационарный случайный процесс с нулевым математическим ожиданием $m_v = 0$ и постоянной дисперсией D_v .

Модель схемы оценки детерминированной составляющей x_1 аддитивного процесса (1) представлена на рис. 1.

Представлены результаты исследования качества фильтрации синусоидального, линейного и экспоненциального зашумленных сигналов фильтром Калмана и фильтром скользящего среднего. Предложена процедура обработки случайных процессов с использованием фильтра скользящего среднего, в которой фильтрации предшествует классификация наблюдаемого процесса, позволяющая определить класс процесса (стационарный, нестационарный), вид процесса (аддитивный, мультипликативный, аддитивно-мультипликативный) и тип детерминированных составляющих случайных процессов.

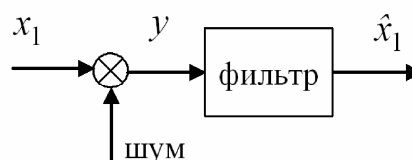


Рис. 1.

В соответствие с принятой схемой оценки (рис. 1) модель наблюдаемой системы может быть описана конечноразностными уравнениями:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}_{k+1/k} \mathbf{x}_k \\ \mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \end{cases} \quad (4)$$

В уравнениях приняты следующие условные обозначения: \mathbf{x}_k – вектор состояния системы, \mathbf{y}_k – вектор измерения системы в момент времени $t = t_k$; \mathbf{v}_k – гауссовский белый шум измерения со статистическими характеристиками: математическое ожидание $M[\mathbf{v}_k] = 0$ и $cov(\mathbf{v}_k, \mathbf{v}_j) = \mathbf{R}_k \delta_{kj}$ (\mathbf{R}_k – положительно определенная симметричная матрица интенсивности); $\mathbf{A}_{k+1/k}$ – переходная матрица системы; \mathbf{H}_k – матрица измерений. В модели (4) предполагается, что векторы \mathbf{x}_k и \mathbf{v}_k взаимно некоррелированные.

Для системы уравнений (4) необходимо получить оценки вектора состояния \mathbf{x} с помощью фильтра Калмана и фильтра скользящего среднего для синусоидального, линейного и экспоненциального зашумленных сигналов.

Фильтрация во временной области дает оценку математического ожидания случайного процесса. Обычно наблюдаемый сигнал представляется в виде аддитивной модели (1), которая не всегда адекватна реальному процессу. Ошибка выбора модели процесса приводит к возникновению дополнительной методической погрешности при измерениях математического ожидания.

В силу этого возникает задача анализа погрешности оценки математического ожидания измеряемого процесса, вызванной неправильной классификацией наблюдаемого процесса.

С учетом принятых допущений математическое ожидание случайных процессов, представленных моделями (1 и 3), в момент времени t_k определяется детерминированной составляющей $x_{1,k}$. В качестве меры точности оценки $\hat{x}_{1,k}$ предлагается использовать величину среднеквадратической ошибки:

$$\xi_k = \sqrt{M[(\hat{x}_{1,k} - x_{1,k})^2]} \quad (5)$$

Для оценки малости величины среднеквадратической погрешности (5) представим ее в долях от оцениваемого параметра:

$$\delta_k = \frac{\xi_k}{x_{1,k}}$$

Нормированная среднеквадратическая ошибка (НСКО) в зависимости от значения $x_{1,k}$ изменяется в широких пределах вплоть до $\delta_k = \infty$ при $x_{1,k} = 0$. Поэтому при нормировании предлагается использовать максимальное значение истинного математического ожидания, взятого по модулю $|x_{1max}|$:

$$\delta_k = \frac{\xi_k}{|x_{1max}|} \quad (6)$$

Качество фильтрации сигналов, описываемых моделями (1, 3), а также величина погрешности оценки математического ожидания измеряемого процесса, исследовались для следующих законов изменения аддитивных (индекс а) и мультипликативных (индекс м) составляющих:

$$x_1 = 1 + \sin(2\pi f_a t), f_a = 1 \text{ Гц};$$

$$x_1 = 2 - \exp(-b_a t), b_a = 2;$$

$$x_1 = 1 + a_a t, a_a = 0,5;$$

$$x_2 = a_m t, a_m = 0,5;$$

$$x_2 = \sin(2\pi f_m t), f_m = 4 \text{ Гц};$$

$$x_2 = \exp(-b_m t), b_m = 4.$$

При моделировании длина реализации процессов была 1,5 сек., интервал дискретизации $\Delta t = 0,02$ сек., дисперсия шума $D_v = 0,1$.

Фильтр Калмана

Формулы дискретного фильтра Калмана [2] для случая линейных уравнений состояния и измерения (4) имеют вид:

$$\begin{cases} x(k/k) = x(k/k-1) + K_k(y_k - H_k x(k/k-1)) \\ x(k/k-1) = A(k/k-1)x(k-1/k-1) \\ P(k/k-1) = A(k/k-1)P(k-1/k-1)A^T(k/k-1) \\ K_k = P(k/k-1)H_k^T [H_k P(k/k-1)H_k^T + R_k]^{-1} \\ P(k/k) = (I - K_k H_k)P(k/k-1) \end{cases} \quad (7)$$

здесь $P(k/k-1)$ – корреляционная матрица вектора состояния, K_k – коэффициент усиления фильтра.

Рекуррентная форма фильтра (7) легко реализуется с помощью цифровых вычислительных машин в реальном масштабе времени.

Нормированная среднеквадратическая ошибка (6) для аддитивной модели процессов с разными детерминированными составляющими приведена на рис. 2. Расчеты показали, что НСКО практически не зависит от вида детерминированной составляющей. Это объясняется тем, что коэффициент усиления K и корреляционная матрица P вектора состояния не зависят от поступающих значений вектора измерений y .

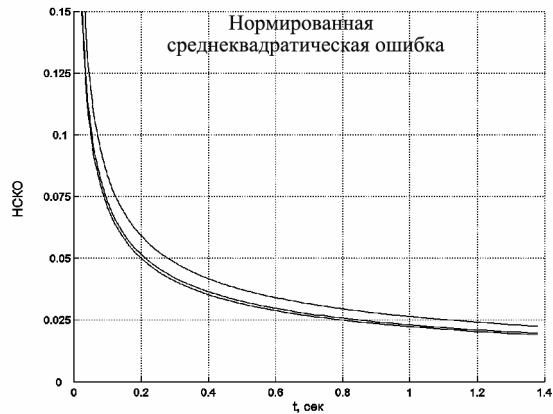


Рис. 2.

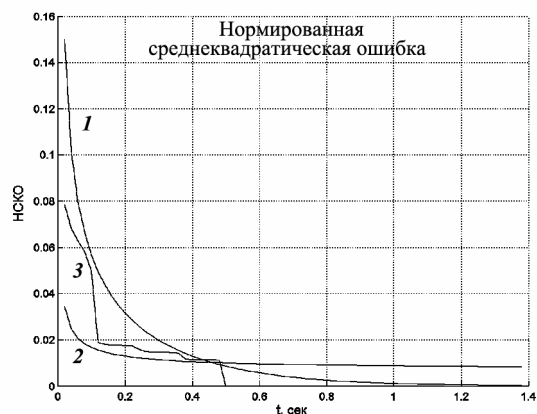


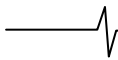
Рис. 3.

Дисперсия шума v существенно влияет на НСКО: увеличение дисперсии шума на порядок приводит к росту на порядок и нормированной среднеквадратической ошибки.

В алгоритме фильтра Калмана предполагается, что помеха представляет собой белый шум. Существуют решения задачи оптимальной фильтрации для случая «окрашенного шума» [3]. Дисперсия аддитивно-мультипликативного процесса (3) зависит от детерминированной составляющей x_2 , не является постоянной и рассчитывается по формуле:

$$D_y(t_k) = x_2^2(t_k)D_v.$$

В этом случае сходимость процесса оценивания зависит от функции x_2 и дисперсии шума. Результаты моделирования для аддитивно-мультипликативной модели сигнала, аддитивная составляющая которого изменяется по экспоненциальному закону, представлены на рис. 3. Кривая 1 характеризует НСКО при экспонен-



циальной мультипликативной составляющей, 2 – линейной, 3 – периодической составляющей. Если мультипликативная составляющая изменяется по периодическому закону, то наблюдаются резкие изменения коэффициента усиления фильтра, а к середине интервала наблюдения он становится равным нулю.

Для системы, заданной уравнениями (4), были проведены расчеты с фильтром Калмана второго порядка. Для аддитивной модели фильтр второго порядка для выбранных условий моделирования не дал улучшения оценок \mathbf{X} . В случае же аддитивно-мультипликативной модели, элементы матрицы ошибок обращаются в бесконечность, и матрица коэффициентов усиления становится вырожденной.

Применение метода рекуррентной динамической фильтрации, использующей квадратичный критерий качества, оправдано при использовании модели измерительной помехи с априорно известной ограниченной дисперсией. В реальных условиях эксплуатации промышленных систем наблюдается отклонение параметров процессов от расчетных, поэтому точность оценок ухудшается. В работе рассматривалась идеальная модель измерения (рис. 1). Как правило, вектор состояния \mathbf{X} является зашумленным. При этом модель сигнала на выходе объекта управления и параметры шума априори неизвестны. Перечисленные причины ограничивают применение фильтра Калмана в виде (7).

Оператор скользящего среднего

В алгоритме скользящего среднего оценка x_1 может быть получена при вычислении по рекуррентной формуле:

$$\hat{x}_1(k) = \hat{x}_1(k-1) + \frac{1}{c} [y(k+h) - y(k-1-h)],$$

где c – интервал усреднения (сглаживания), $h = (c-1)/2$.

Величина среднего значения квадрата ошибки оценки \hat{x}_1 может быть представлена в виде:

$$\begin{aligned} \xi_j^2 &= M \left[(\hat{x}_{1,j} - x_{1,j})^2 \right] = M \left[\left(\frac{1}{c} \sum_{k=-h}^h y_{j+k} - x_{1,j} \right)^2 \right] = \\ &= M \left[(\bar{y}_j - x_{1,j})^2 \right] = M \left[(\bar{y}_j - M[\bar{y}_j] + M[\bar{y}_j] - x_{1,j})^2 \right] = \\ &= M \left[(\bar{y}_j - M[\bar{y}_j])^2 \right] + M \left[(M[\bar{y}_j] - x_{1,j})^2 \right] \end{aligned}$$

Первое слагаемое в данном выражении равно дисперсии оценки выборочного среднего D_j и характери-

зует долю случайной составляющей в значении ошибки. Второе слагаемое дает квадрат смещения оценки от истинного значения и определяет систематическую составляющую ошибки Δ_j^2 . Тогда среднеквадратическая ошибка может быть определена как:

$$\xi_j = \sqrt{D_j + \Delta_j^2} \quad (8)$$

Математическое ожидание выборочного среднего для аддитивной и аддитивно-мультипликативной моделей можно представить в виде:

$$M[\bar{y}_j] = M \left[\frac{1}{c} \sum_{k=-h}^h y_{j+k} \right] = \frac{1}{c} \sum_{k=-h}^h x_{1,j+k}.$$

Систематическая составляющая ошибки для обеих рассматриваемых моделей (1, 3) рассчитывается в соответствии со следующей формулой:

$$\Delta_j^2 = M \left[(M[\bar{y}_j] - x_{1,j})^2 \right] = \left(\frac{1}{c} \sum_{k=-h}^h x_{1,j+k} - x_{1,j} \right)^2.$$

При принятых условиях, относительно шумовой компоненты v , дисперсия выборочного среднего для аддитивной модели определяется как

$$D[\bar{y}_j] = M \left[(\bar{y}_j - M[\bar{y}_j])^2 \right] = \frac{D_v}{c},$$

для аддитивно-мультипликативной модели имеет вид

$$D[\bar{y}_j] = \frac{D_v}{c^2} \sum_{k=-h}^h x_{2,j+k}^2.$$

В табл. 1 приведены значения НСКО (6) оценки x_1 , полученные с помощью оператора скользящего среднего для тех же условий, что при расчете по алгоритму Калмана.

Как показали результаты расчетов, фильтр Калмана дает для аддитивной модели меньшее значение НСКО, чем алгоритм скользящего среднего. Однако фильтр скользящего среднего менее чувствителен к изменению дисперсии наблюдаемого сигнала. Кроме того, проведенные исследования показали эффективность оператора скользящего среднего при фильтрации аддитивно-мультипликативных процессов. Следует также отметить, что алгоритм его реализации более простой, чем для фильтра Калмана. В связи с этим задачу анализа влияния ошибочной классификации наблюдаемого процесса на точность оценки математического ожидания аддитивно-мультипликативного сигнала будем решать для этого оператора.

Таблица 1

Модель процесса	x_1	x_2	Интервал усреднения	δ_{max}
Аддитивная	$x_1 = a_a t$	–	15	0,11
	$x_1 = 1 + \sin(2\pi f_a t)$	–	9	0,059
	$x_1 = 2 - \exp(-b_a t)$	–	15	0,041
Аддитивно-мультипликативный процесс	$x_1 = 1 + \sin(2\pi f_a t)$	$x_2 = a_m t$	9	0,042
	$x_1 = 1 + \sin(2\pi f_a t)$	$x_2 = \sin(2\pi f_m t)$	7	0,047
	$x_1 = 1 + \sin(2\pi f_a t)$	$x_2 = 1 - \exp(-b_m t)$	9	0,056

Анализ результатов ошибочной классификации случайного процесса

Пусть относительно исследуемого процесса выдвинута гипотеза о принадлежности его к классу аддитивных процессов (1), в то время как в действительности процесс относится к классу аддитивно-мультипликативных процессов (3).

Для оценки влияния ошибочной классификации наблюдаемого процесса и вида детерминированных составляющих на НСКО (8) проводилось моделирование процессов при тех же законах изменения детерминированных составляющих, что и исследовались алгоритмы фильтрации.

В вычислительном эксперименте длина реализации составляла 200 отсчетов, величина интервала усреднения изменялась в пределах от 3 до 199, интервал дискретизации – $\Delta t = 0,005$ сек. Результаты моделирования представлены на рис. 4.

Анализ полученных результатов показывает, что при принятых условиях моделирования НСКО оценки математического ожидания для аддитивной модели процесса (кривые 1) меньше, чем для аддитивно-мультипликативной модели (кривые 2). При этом разность в значениях ошибки колеблется от 0.03 до 0.3 в зависимости от интервала усреднения.

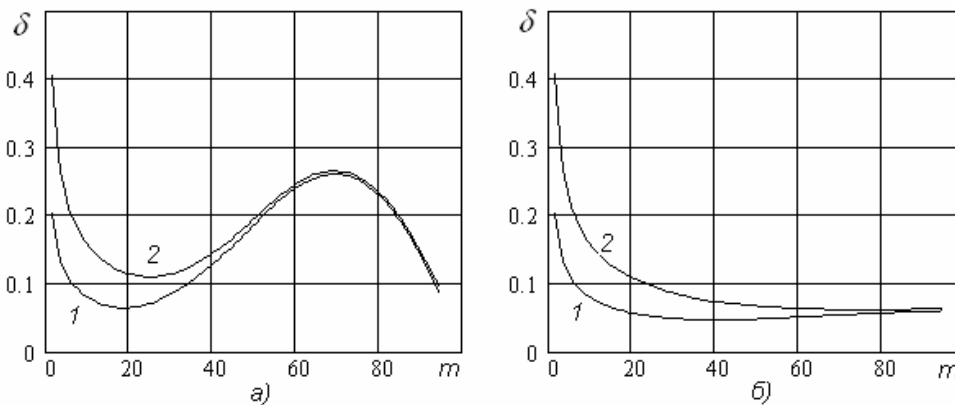


Рис. 4.

Погрешность результатов измерения влияет на качество управления технологическими процессами. Неучтенная погрешность, вызванная ошибкой классификации наблюдаемого процесса, будет снижать эффективность работы информационно-измерительных и управляющих систем.

Как показывает анализ зависимостей, представленных на рис. 4, а, для процесса, математическое ожидание которого изменяется по периодическому закону, можно указать оптимальный интервал сглаживания. Если закон изменения математического ожидания наблюдаемого процесса изменяется монотонно (рис. 4, б), то НСКО уменьшается с увеличением интервала сглаживания. В связи с этим рекомендации по выбору окна усреднения могут быть даны, исходя из вида аддитивной детерминированной составляющей наблюдаемого процесса.

С целью исключения погрешностей результатов измерений, вызванных ошибочным предположением о виде исследуемого процесса и необоснованным выбором интервала сглаживания, необходимо, чтобы процедуре скользящего среднего предшествовала процедура классификации случайного процесса.

Предлагаемая процедура обработки случайных процессов

В существующей практике анализа случайных процессов одним из первых этапов является этап выявления класса случайного процесса (стационарный или нестационарный). Выводы о классе процесса делаются, как правило, на основе проверки статистических гипотез о наличии тренда [4]. В отличие от сложившейся практики предлагается включить в информационно-измерительную систему блок, в котором реализуется процедура классификации случайного процесса (блок 5 на рис. 5), позволяющая выявить характер нестационарности процесса и вид детерминированных составляющих. В результате этой классификации процесс, заданный вектором информативных признаков P_i , соотносится с одним из наперед заданных классов d_j случайных процессов:

$$P_i = \{p_{i,1}, p_{i,2}, \dots, p_{i,n}\} \rightarrow y \in \{d_1, d_2, \dots, d_m\}.$$

Первым этапом процедуры обработки случайных процессов должен быть этап формирования вектора информативных признаков.

В условиях ограниченной априорной информации о свойствах исследуемого процесса решение о классе про-

цесса следует принимать по результатам применения совокупности непараметрических критериев. Как известно, при использовании непараметрических критериев в качестве нулевой гипотезы H_0 принимается гипотеза, что все наблюдения независимы и процесс является случайным, альтернативной является гипотеза H_1 – временной ряд содержит детерминированные составляющие. В связи с существованием только двух возможных исходов предлагается получить обобщенный классификационный показатель на основе байесовской процедуры для бинарных признаков [5]. В блоке 1 (рис. 5) выполняется расчет непараметрических критериев, в блоке 2 реализуется алгоритм получения обобщенного классификационного показателя по совокупности критериев.

Одним из методов анализа временных рядов, ориентированным на выявление детерминированных компонент случайного процесса, является метод Херста [6]. По значению показателя Херста можно оценить, является ли наблюдаемый процесс «чистым» шумом или он содержит монотонную или периодическую составляющую. Расчет показателя Херста выполняется в блоке 3.

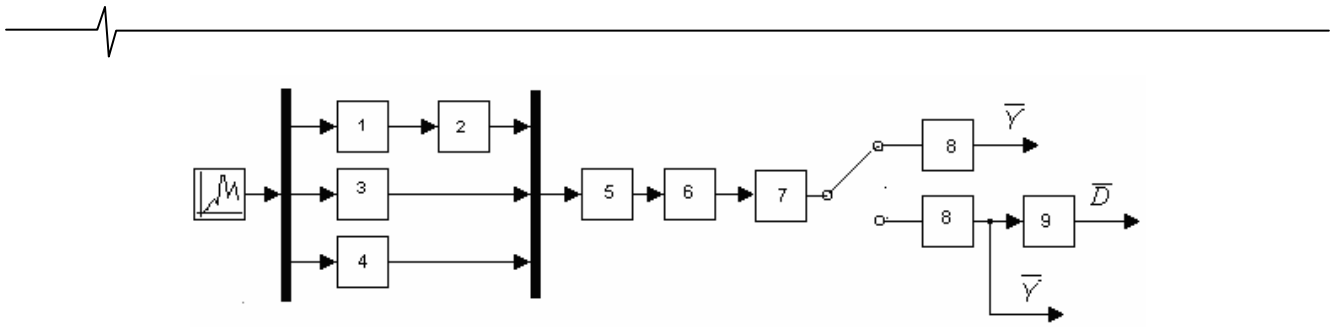


Рис.5. Структура информационной системы обработки случайного процесса

Важной характеристикой случайного процесса является корреляционная функция. Случайный временной ряд, не содержащий детерминированных составляющих, при различных лагах имеет близкие к нулю коэффициенты корреляции. Если во временном ряде присутствует некоторая детерминированная составляющая, то найдутся лаги, для которых значения коэффициентов автокорреляции выйдут за пределы 95 % доверительного интервала [7]. В блоке 4 рассчитывается и анализируется функция автокорреляции.

Таким образом, вектор информативных признаков предлагается формировать из обобщенного показателя по совокупности непараметрических критериев, показателя Херста и числа лагов, превышающих доверительные границы автокорреляционной функции.

В силу наличия неопределенности исходных данных для установления класса наблюдаемого процесса предлагается использовать методы нечеткой логики. Целью блока процедуры нечеткой классификации является определение класса процесса (стационарный, нестационарный), вида процесса (аддитивный, мультипликативный, аддитивно-мультипликативный) и типа детерминированной составляющей.

Далее в блоке 6 определяются параметры детерминированной аддитивной составляющей.

Для периодической составляющей в блоке 6 оценка периода T в момент времени n определяется по максимуму функции:

$$R(n, T) = \sum_{i=0}^{N-1} y_{n+i} y_{n-T+i}$$

При постановке задачи допускается, что помеха U может иметь закон распределения отличный от нормального. Тогда для определения параметров линейной и экспоненциальной составляющих можно рекомендовать использование метода непараметрической регрессии, основанного на рангах наблюдения [7]. Применение этого же метода допустимо и для экспоненциальной составляющей после ее предварительного логарифмирования.

В блоке 7 осуществляется выбор интервала усреднения. Для сигналов, математическое ожидание которых изменяется по периодическому закону, предлагается использовать полученные в результате моделирования зависимости интервала усреднения от интервала дискретизации и периода наблюдаемого сигнала.

В случае линейной и монотонной аддитивных составляющих в блоке 7 для оценки интервала усреднения можно хранить базу данных номограмм, отражающих зависимость среднеквадратической ошибки ξ от параметров детерминированных составляющих и интервала усреднения. Задавшись значением ξ и зная оценки парамет-

ров детерминированной составляющей и доверительные интервалы для оценок, можно получить допустимый интервал усреднения.

В блоке 8 реализуется алгоритм скользящего среднего. Оператор скользящего среднего одновременно с вычислением оценок математического ожидания x_1 допускает выполнение операции центрирования. По результатам операции центрирования можно судить о качестве оценок математического ожидания исследуемого процесса.

Наблюдаемые процессы могут относиться к классу процессов нестационарных по дисперсии и корреляционной функции. В этом случае необходимо получить оценки дисперсии, что предполагает включение в систему обработки случайного процесса блока нормирования (блок 9). Степень отличия оценки дисперсии от единицы после нормирования будет говорить о качестве выделения мультипликативной детерминированной составляющей исследуемого процесса.

Заключение

В статье предложен алгоритм обработки случайных процессов с использованием фильтра скользящего среднего. С целью повышения точности вероятностных характеристик измеряемых сигналов фильтрации предшествует процедура классификации наблюдаемого процесса. В зависимости от класса процесса, вида процесса, типа и параметров детерминированных составляющих реализуются автоматическая настройка интервала усреднения и выбор алгоритма обработки, предусматривающего блок центрирования или блок нормирования, или последовательное включение обоих блоков.

Литература

1. Харкевич А. А. Борьба с помехами. – М.: Наука, 1965. – 275 с.
2. Бесекерский В.А. Цифровые автоматические системы. – М.: Наука, 1976. – 576 с.
3. Статистическая динамика и идентификация систем автоматического управления / Под ред. К.А. Пупкова, Н.Д. Егупова – М.: Издательство МГТУ им. Н.Э. Баумана, 2004. – 640 с.
4. Кендалл М., Стьюарт А. Многомерный статистический анализ и временные ряды / Пер. с англ. – М.: Наука, 1976. – 736 с.
5. Афифи А., Эйзен С. Статистический анализ: Подход с использованием ЭВМ / Пер. с англ. – М.: Мир, 1982. – 488 с.
6. Федер Е. Фракталы / Пер. с англ. – М.: Мир, 1991. – 254 с.
7. Тюрин Ю. Н., Макаров А. А. Статистический анализ данных на компьютере / Под ред. В. Э. Фигурнова – М.: ИНФРА-М, 1998. – 528 с.

УДК 621.391

ДЕКОМПОЗИЦИЯ НА ЭМПИРИЧЕСКИЕ МОДЫ И ЕЕ ИСПОЛЬЗОВАНИЕ ПРИ АНАЛИЗЕ ДРОБНОГО БРОУНОВСКОГО ДВИЖЕНИЯ

Клионский Д.М., Орешко Н.И., Геппенер В.В.

Введение

Большинство сигналов, рассматриваемых в рамках решения различных прикладных задач, являются по своей природе нестационарными. Стационарность сигналов, обычно рассматриваемая на практике в широком смысле, означает, что некоторые характеристики не зависят от времени, а некоторые зависят только от величины интервала между конкретной парой временных отсчетов. Среди упомянутых характеристик одними из наиболее информативных являются начальные и центральные моменты различных порядков, значения автокорреляционной функции (АКФ), а также значения мгновенной частоты, обычно вычисляемой на основе преобразования Гильберта.

Широко используемый классический анализ Фурье на основе аппроксимации сигнала конечным набором дискретных гармоник предназначен для обработки стационарных сигналов в силу очевидных особенностей используемых базисных функций (гармонических функций). Но у них при этом отсутствует точная временная локализация отдельных особенностей (при вычислении коэффициентов разложения сигнала по Фурье-базису используется весь сигнал, что негативно сказывается на временной локализации) и, кроме того, для нестационарного сигнала может потребоваться довольно большое число гармоник для обеспечения требуемой точности представления. Энергия сигнала при этом рассеивается по широкой полосе частот, что затрудняет выявление основных диапазонов ее концентрации. Для повышения точности и достоверности анализа нестационарных сигналов необходим особый подход, обладающий свойством *адаптивности* к конкретному рассматриваемому сигналу. Другими словами, разложение по базису с целью дальнейшего анализа должно проводиться с учетом *локальных особенностей* (таких как, например, экстремумы и нули) и *внутренней структуры* (наличия амплитудной, частотной и фазовой модуляций, шумов и помех) каждого конкретного сигнала. Этим важнейшим требованиям отвечает алгоритм *декомпозиции на эмпирические моды* (ДЭМ, *EMD – Empirical Mode Decomposition*) [1-4], являющийся весьма перспективным при решении таких задач, как очистка сигналов от шумов различной природы, выделение тренда среднего значения, анализ степени хаотичности и регулярности сигналов на основе оценивания показателя Херста, экстраполяция сигналов, частотно-временной анализ на основе *преобразования Гильберта-Гуанга* (*HHT – Hilbert-Huang Transform*). Наряду с этим, не менее важным является возможность проведения мультиразрешающего анализа сигналов (представление в виде набора компонент с различным временным разрешением), что позволяет выявить отдельные компо-

Представлен принципиально новый подход к анализу нестационарных сигналов сложной внутренней структуры – алгоритм декомпозиции на эмпирические моды. Описаны идеи, заложенные в его основу, а также основные преимущества по сравнению с существующими методами анализа нестационарных сигналов. В качестве примера рассмотрена задача анализа дробного броуновского движения, в частности, оценивание его показателя Херста.

ненты, входящие в состав сложного сигнала, классифицировать их в частотной области и сконструировать на их основе банк фильтров. Основным преимуществом подхода ДЭМ является адаптивность всех используемых методов анализа на его основе, что дает возможность обеспечить более высокую точность решения вышеперечисленных задач. Следует также отметить, что в силу адаптивной природы появляются дополнительные возможности для физической интерпретации полученных базисных функций в различных предметных областях [1].

Выше упомянутые свойства ДЭМ позволяют также применить эту стремительно развивающуюся технологию для «интеллектуального» анализа сигналов (сегментации, кластер-анализа, классификации на основе разложения и спектра Гильберта) [4]. Безусловно, требуют дальнейшего исследования новые возможности ДЭМ при работе с различными классами нестационарных сигналов.

Понятие эмпирической и характеристической мод. Основные свойства

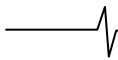
Из названия алгоритма следует, что *эмпирические моды* (*ЭМ, Empirical mode, Intrinsic mode function*) представляют собой те компоненты, на которые осуществляется декомпозиция (разложение) исходного сигнала. В общем случае данные функции могут иметь произвольную форму, быть заданы аналитически (точной формулой) или численно (в виде набора дискретных отсчетов), однако они непременно должны удовлетворять 2-м необходимым условиям [1-4]:

1) На всей длительности функции общее число экстремумов должно быть равно общему числу нулей с точностью до 1:

$$N_{\max} + N_{\min} = N_{\text{zero}} \pm 1, \quad (1)$$

где N_{\max} , N_{\min} , N_{zero} - соответственно, общее число максимумов, минимумов и нулей на всей длительности сигнала, не считая начальный и конечный отсчеты (которые в редких случаях могут быть единственными экстремумами сигнала).

2) Полусумма 2-х огибающих функции – верхней, ин-терполирующей весь найденный набор локальных макси-



мумов и нижней, интерполирующей весь найденный набор локальных минимумов, не должна превышать некоторого малого значения η , определяемого машинной точностью \mathcal{E} и погрешностями, связанными с получением, преобразованием и передачей данных. При интерполяции обеих огибающих используются в основном *кубические сплайны* [11,12], причина выбора которых будет пояснена далее. Точного равенства нулю данной полусуммы на практике добиться невозможно в силу ряда объективных причин: точность определения экстремумов, вычислительные погрешности (связанные с особенностями машинной арифметики с плавающей точкой), плохая обусловленность систем уравнений, на основе которых определяются коэффициенты сплайнов, а также влияние самого исходного сигнала (например, краевые эффекты – сильные осцилляции интерполирующей функции на краях). Аналитически данное условие записывается как

$$\frac{U(k) + L(k)}{2} \leq \eta, \quad k = \overline{1, N}, \quad (2)$$

где $U(k)$ и $L(k)$ - значения верхней и нижней огибающих сигнала, сформированных на основе интерполяции кубическими сплайнами (k - номер отсчета сигнала), η - выбранное пороговое значение, N - общее количество отсчетов сигнала. Данная полусумма есть не что иное, как *локальное среднее* сигнала, которое должно быть меньше либо равно пороговому значению η в каждый дискретный момент времени.

Данные условия требуют некоторых комментариев. Первое из них, по сути, является требованием узкополосности ЭМ. Данная мера узкополосности ν выражается следующим образом [10]:

$$\nu^2 = \frac{(N_{\max} + N_{\min})^2 - N_{\text{zero}}^2}{N_{\text{zero}}^2}. \quad (3)$$

Очевидно, эта величина равна нулю, если число экстремумов равняется числу нулей, и она близка к нулю, если между ними есть отличие на 1. Именно это и утверждается в первой части определения ЭМ.

Второе условие также имеет физическую интерпретацию. Из него следует, что ЭМ является стационарной функцией по отношению к вышеопределенному локальному среднему значению, которое неизменно и меньше порогового значения η в любой момент времени. Кроме того, для достижения этого условия ЭМ должна иметь положительные значения в точках максимумов и отрицательные значения в точках минимумов (должна соблюдаться симметрия относительно оси времени), т.к. иначе для отдель-

ных моментов времени условие может нарушиться. Наконец, важно отметить, что ЭМ в общем случае обладает одновременно и амплитудной и частотной модуляциями. Закон амплитудной модуляции может быть установлен из огибающих, полученных интерполяцией экстремумов или преобразованием Гильберта, а закон частотной модуляции устанавливается путем вычисления мгновенной частоты.

Понятие огибающей, введенное для определения ЭМ, рассматривается с учетом выбранного вида интерполяции. Как было указано, в данном алгоритме используется кубическая сплайн-интерполяция. У сплайнов в общем и кубических сплайнов в частности есть ряд преимуществ [11] по сравнению с другими функциями, обладающими высокой степенью гладкости, наличие которой существенно повышает точность при извлечении ЭМ из сигнала.

Во-первых, по сравнению с полиномами у сплайнов нет эффекта “раскачивания” [11] – сильных осцилляций интерполирующей функции в случае, если данные заведомо имеют не полиномиальную природу. Кроме того, кубические сплайны являются непрерывными (подразумевается равенство значений в узлах стыков) и дважды дифференцируемыми. Последнее означает, что сам сплайн не имеет острых углов (определенность 1-й производной), и в любой точке определен радиус его кривизны (определенность 2-й производной), что дает возможность оценивать интенсивность осцилляций самой моды. Но самым важным является то, что кубический сплайн минимизирует осциллирующее поведение функции – т.е. из всего многообразия дважды дифференцируемых, непрерывных на некотором интервале $[a; b]$ функций $f(k)$, интерполирующих заданную совокупность точек $\{(x_k, y_k)\}_{k=1}^N$, кубический сплайн меньше всего осциллирует. Математически данное утверждение можно записать в следующем виде:

$$\int_a^b [S''(x)]^2 dx \leq \int_a^b [f''(x)]^2 dx, \quad (4)$$

при условии $S'(a) = f'(a), S'(b) = f'(b)$ где $S(x)$ - кубический сплайн, $f(x)$ - произвольная дважды дифференцируемая функция. Условие записано для аналоговых (непрерывных) функций, но оно также верно и для дискретных. Аналогичное условие справедливо и для сплайнов более высоких порядков (там рассматривается множество n раз дифференцируемых функций, где $n+1$ - порядок сплайна). Но их применение сопряжено с неоправданно большими для данного алгоритма временными затратами на вычисление коэффициентов.

Ниже (рис. 1) приведены примеры ЭМ с интерполяцией огибающих кубическими сплайнами:

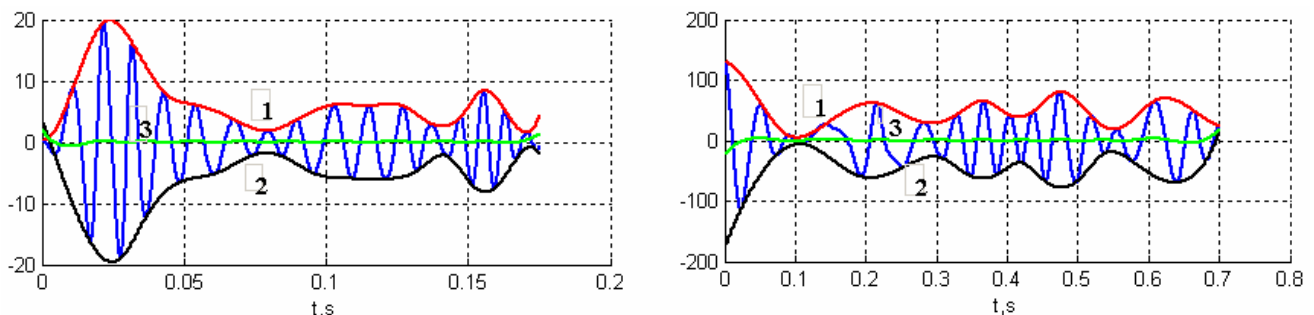


Рис. 1. Примеры ЭМ. 1 - верхняя огибающая, 2 - нижняя огибающая, 3 - их полусумма

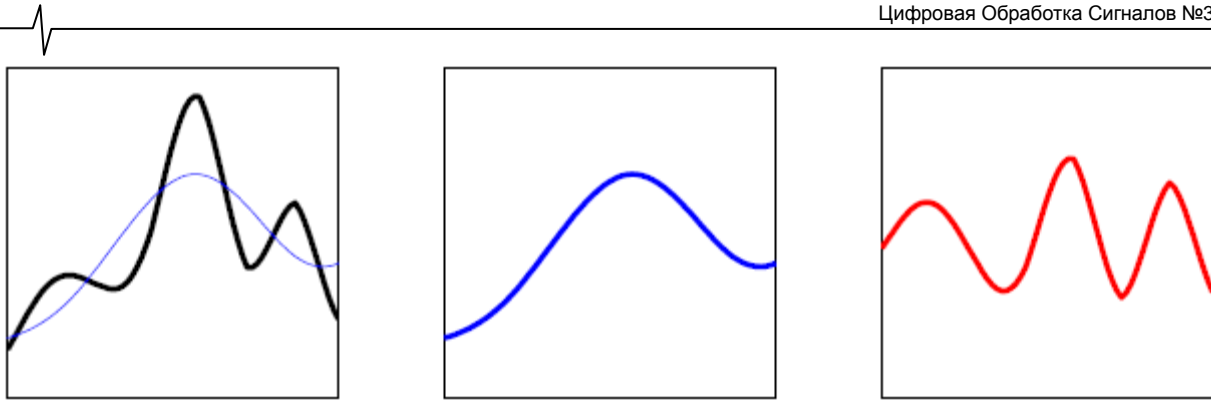


Рис. 2. Иллюстрация представления сигнала (слева) в виде суммы низкочастотной, слабо осциллирующей компоненты (в центре) и высокочастотной, заметно осциллирующей (справа)

Характеристические моды (ХМ, *Characteristic mode*), в отличие от ЭМ, всегда имеют строгое и точное аналитическое выражение, но они довольно редко встречаются на практике. При этом оба необходимых условия, сформулированные для ЭМ, остаются в силе. Подобные модели можно использовать при аппроксимации ЭМ с помощью ХМ, что иногда облегчает дальнейший анализ и обработку. Некоторые примеры ХМ приведены ниже:

- 1) $s_1(k) = A \cos(w_0 k + \varphi)$ - гармонический сигнал с известными A, w_0, φ .
- 2) $s_2(k) = A \cos(w_0 k + 0.5 \mu k^2 + \varphi)$ - ЛЧМ-сигнал с известными A, w_0, μ, φ .
- 3) $s_3(k) = e^{-ak^2} \cos(w_0 k + \varphi)$ - Гауссов радиоимпульс с известными a, w_0, φ .

Алгоритм ДЭМ. Процесс отсеивания

Алгоритм ДЭМ представляет любой сигнал с конечным числом отсчетов (конечной энергией) в виде набора ЭМ. В основе алгоритма лежит важная и фундаментальная идея, основанная на том, что на каждом этапе, т.е. при извлечении очередной моды, сигнал (вернее, текущий остаток) представляется в виде суммы быстро осциллирующей компоненты, которой является сама извлеченная ЭМ, и медленно осциллирующей, которой является результирующий остаток. Последний затем подвергается дальнейшему разложению, если у него имеется хотя бы один минимум и хотя бы один максимум.

Сказанное иллюстрируется графически на рис. 2.

Основные этапы алгоритма, в соответствии со схемой на рис. 3, описаны далее.

Шаг 1. Рассматривается текущий остаток $r_p(k)$ ($r_1(k)$ - первый остаток, который и есть сам исходный сигнал $s(k)$). Определяются его экстремумы и формируются 2 набора:

$$\{M_i\}, i=1,2,3,\dots; \{m_i\}, i=1,2,3,\dots, \quad (5)$$

где $\{M_i\}$ и $\{m_i\}$ - набор максимумов и минимумов соответственно;

Если экстремумы отсутствуют в текущем остатке, то возможно применение операции численного дифференцирования с помощью конечноразностных формул для их порождения. По окончании работы всего алгоритма производится возврат к начальному

множеству значений путем интегрирования каждой ЭМ на основе численных методов [11,12].

Далее по найденным экстремумам строятся две огибающие с помощью интерполяции кубическими сплайнами:

$$U_j(k) = f_U(M_i, k), L_j(k) = f_L(m_i, k), \quad (6)$$

где $U_j(k)$ и $L_j(k)$ - верхняя и нижняя огибающие, построенные, соответственно, по найденным локальным максимумам и минимумам, j - номер итерации отсеивания (см. ниже).

Необходимо отметить, что построенные огибающие должны заключать в себе весь текущий остаток, т.е. любой отсчет должен удовлетворять условию:

$$L_j(k) \leq r_p(k) \leq U_j(k), \quad \forall k. \quad (7)$$

Однако в некоторых случаях наблюдаются так называемые "всплески" - явление, при котором какой-либо отсчет превышает значения верхней огибающей или же уступает значению нижней огибающей. Также необходимо учитывать влияние начального и конечного отсчетов на интенсивность осцилляций на краях (краевые эффекты). Для уменьшения этих осцилляций используются методы зеркального отображения экстремумов, ближайших к краевым отсчетам [1], что позволяет в некоторой степени подавить или сгладить колебания на краях.

После построения огибающих вычисляется их полусумма - локальное (мгновенное) среднее значение, зависящее от времени:

$$e_j(k) = 0.5 \cdot [U_j(k) + L_j(k)]. \quad (8)$$

Затем выполняется переход к шагу 2.

Шаг 2. Найденное среднее значение вычитается из текущего остатка, и полученный результат $h_j(k)$ оказывается "кандидатом" на то, чтобы стать очередной модой:

$$h_j(k) = r_p(k) \text{ при } j=1, \quad (9)$$

$$h_j(k) = h_{j-1}(k) - e_j(k) \text{ при } j > 1, \quad (10)$$

где p - номер извлекаемой ЭМ.

При этом требуется проверить два необходимых условия отнесения эмпирической функции $h(k)$ к классу ЭМ. Если оба условия выполняются, то осуществляется переход к шагу 3.

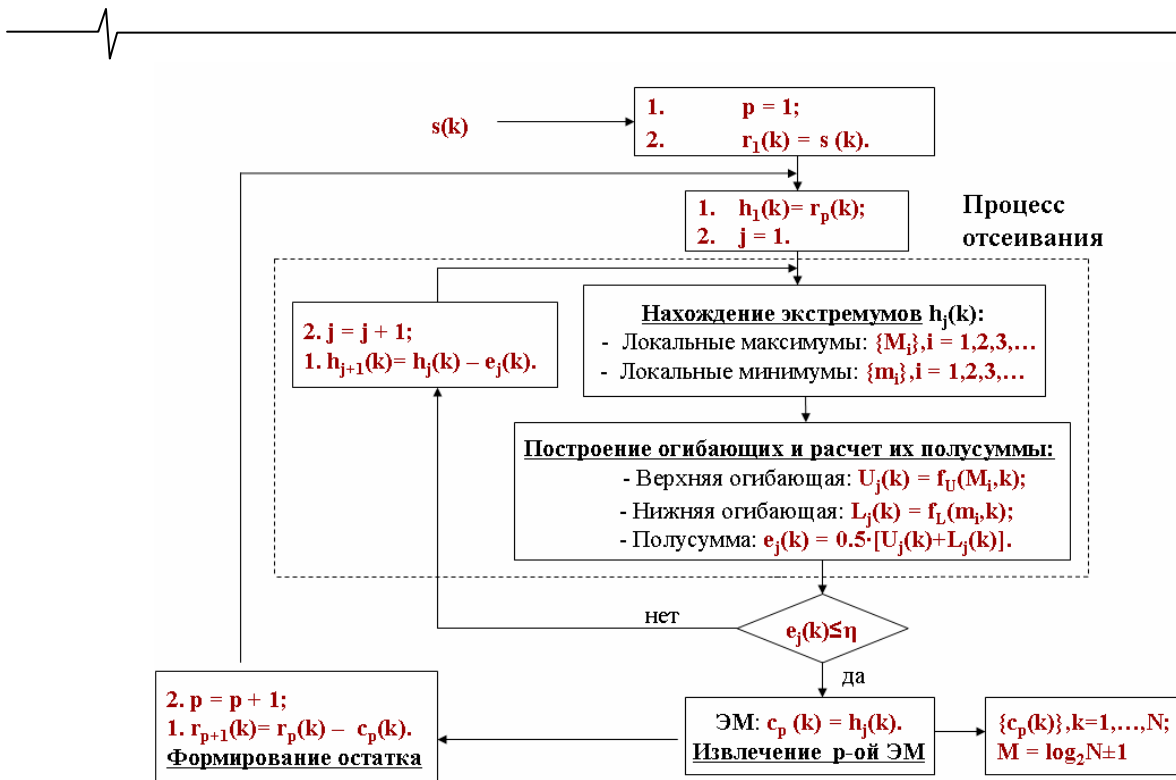


Рис. 3. Графическое представление основных этапов алгоритма ДЭМ

Если какое-либо из них нарушено, то осуществляется возврат к шагу 1, но теперь уже в качестве текущего остатка выступает полученный на 2-м шаге результат. Тем самым начинается так называемый *процесс отсеивания* (*sifting process*) [1-4], который может быть записан в следующем виде:

$$\begin{aligned} h_1(k) &= r_p(k), h_2(k) = h_1(k) - e_1(k), \\ h_{j+1}(k) &= h_j(k) - e_j(k), \dots c_p(k) = h_{iter}(k), \end{aligned} \quad (11)$$

где $e_j(k)$ - среднее значение функции на j -ой итерации процесса отсеивания, $h_j(k)$ - текущий результат на j -ой итерации отсеивания, $iter$ - общее число итераций для данной ЭМ.

На итерации с номером $iter$ процесс отсеивания для извлечения очередной моды прекращается и выполняется переход к шагу 3.

Шаг 3. После извлечения ЭМ в ее окончательном виде осуществляется ее вычитание из текущего остатка для формирования нового (для обновления остатка):

$$r_{p+1}(k) = r_p(k) - c_p(k), \quad (12)$$

где $c_p(k)$ - полученная ЭМ; $r_p(k)$ - текущий остаток, $r_{p+1}(k)$ - новый остаток.

Шаг 4. Далее осуществляется переход к шагу 1, где в качестве функции, из которой будут извлекаться ЭМ с более высокими номерами, выступит тот остаток, который был получен на 3-м шаге, т.е. $r_{p+1}(k)$.

Основные этапы представленного алгоритма требуют некоторых комментариев. В настоящее время он еще не имеет такой стройной и глубоко изученной теоретической базы, которой обладает классический анализ Фурье или вейвлет-анализ. Алгоритм имеет эмпирическую природу. Например, строго не доказана его сходимость при произвольном способе интерполяции огибающих. Однако, во-первых, она вытекает из того обстоятельства, что на каждом этапе происходит выделе-

ние компонент таким образом, что удаляется локальное среднее значение, и в итоге остается результирующий низкочастотный монотонный остаток, отражающий общую тенденцию в сигнале (тренд среднего). Он не будет иметь достаточного числа экстремумов, и понятие огибающих (либо верхней, либо нижней, либо обеих) для него потеряет смысл - это будет означать завершение работы алгоритма. Кроме того, возможно введение ограничения на число извлекаемых ЭМ, например, с использованием *информационных критериев* (Акаике, Риссанена и др.) [5]. Последние, как известно, находят оптимальное значение порядка модели, доставляющее экстремум (максимум или минимум в зависимости от конкретного вида) выражению, определяемому выбранным критерием. Он обычно содержит два слагаемых, одно из которых возрастает с увеличением порядка модели, а другое, наоборот, убывает. Применительно к данной задаче в качестве порядка модели будет выступать число ЭМ в разложении, а в качестве дисперсии ошибки - дисперсия сигнала ошибки приближения исходного сигнала выбранным числом ЭМ. Данный подход направлен на то, чтобы сохранить лишь наиболее информативные компоненты, относящиеся к полезному сигналу, и исключить из рассмотрения те, что являются шумовыми. Наиболее часто рассматриваемые информационные критерии имеют вид:

$$F(q, D) = N \ln D + g(q, N), \quad (13)$$

где N - число отсчетов в сигнале, q - порядок модели (число ЭМ), D - дисперсия ошибки восстановления сигнала числом ЭМ, равным q . Широко используемыми являются

$$ICA(q, D) = N \ln D + 2q, \quad (14)$$

критерий Акаике, где $g(q, N) = g(q) = 2q$, и

$$MDL(q, D) = N \ln D + q \ln N, \quad (15)$$

критерий Риссанена (критерий минимума длины описания), где $g(q, N) = q \ln N$.

При отыскании оптимального числа ЭМ разумно начать с тех мод, что имеют самые большие номера, т.е. являются самыми низкочастотными, и наращивать это число в направлении высокочастотных мод, которые являются шумовыми и нежелательными с точки зрения результирующего отношения сигнал-шум.

Исходя из практических результатов, была выведена эмпирическая формула для оценки общего числа ЭМ в разложении произвольного сигнала:

$$M = \log_2 N \pm 1, \quad (16)$$

где M - общее число ЭМ, N - общее число отсчетов сигнала.

Ключевая роль во всем алгоритме отводится уже упоминавшемуся процессу отсеивания (на рис. 3 он обозначен пунктиром). Он имеет итерационную природу и предназначен для извлечения конкретной ЭМ, удовлетворяющей двум необходимым условиям из ее определения. Смысл названия "отсеивание" состоит в том, что отсеиваются (отбраковываются) все те "кандидаты" на роль ЭМ, которые ими на самом деле не являются по причине невыполнения одного из двух требуемых условий.

Всю процедуру разложения можно рассматривать как два вложенных друг в друга цикла: внешний отвечает за извлечение очередной моды, а внутренний – за процесс отсеивания для очередной моды. Отсеивание применяется для каждой ЭМ в отдельности, однако число итераций может отличаться, причем порой весьма существенно. В силу итерационного характера невозможно заранее предсказать, сколько именно вычислений потребуется для каждой ЭМ. Для более точного выполнения отсеивания были специально разработаны различные критерии его остановки, наиболее широко используемые из которых приведены ниже.

Процесс отсеивания для очередной моды прекращается, если величина показателя нормированной квадратичной разности между значениями эмпирической функции $h(k)$ на двух последовательных итерациях отсеивания SD выходит за границы интервала $[0.2; 0.3]$. SD определяется как

$$SD = \sum_{k=1}^N \frac{[h_j(k) - h_{j-1}(k)]^2}{h_{j-1}^2(k) + \mu}, \quad (17)$$

где j и $j-1$ – номера двух последовательных итераций процесса отсеивания; $h_j(k)$ и $h_{j-1}(k)$ – значения эмпирической функции на двух последовательных итерациях отсеивания, μ - маленькое число (обычно $\mu = 0.00001$), введенное на случай равенства $h_{j-1}(k)$ нулю в отдельные моменты времени. Указанный интервал $[0.2; 0.3]$ получен в работе [1]. В общем случае пороговое значение показателя SD может быть изменено. Главным недостатком данного критерия является то, что он никак не учитывает определение ЭМ, а именно ни условие равенства числа нулей и экстремумов, ни требование симметричности огибающих. Следовательно, вполне возможен вариант, когда результатом процесса отсеивания окажется функция, не удовлетворяющая одному из необходимых условий, что даст неверный результат в целом. Однако в большинстве случаев,

как показывают эксперименты, этого не происходит, и полученные по результатам процесса отсеивания функции являются истинными ЭМ.

Процесс отсеивания для очередной моды прекращается, если число экстремумов и нулей на текущей и предыдущей итерациях совпадают, и эти числа между собой одинаковы либо отличаются на 1. Если при этом полусумма двух огибающих функции превосходит пороговое значение η (т.е. не выполняется второе условие из определения ЭМ), то число итераций может быть увеличено.

После завершения алгоритма можно провести реконструкцию (восстановление) исходного сигнала с помощью выделенных мод. Для этого используются следующие формулы:

$$s(k) = \sum_{i=1}^{M-1} c_i(k) + r_M(k) \quad (18)$$

при использовании всех ЭМ (как высокочастотных, так и низкочастотных) и

$$s(k) = \sum_{i \in I} c_i(k) + r_M(k), \quad (19)$$

в случае исключения некоторых ЭМ из рассмотрения, например, в результате применения информационных критериев. В этом случае верхняя граница в суммировании в (19) не указывается, а индекс суммирования i принадлежит индексному множеству I , которое упорядочивается в зависимости от включения той или иной ЭМ.

Алгоритм ДЭМ может быть применен для решения широкого класса задач, часть из которых уже была упомянута в начале данной статьи. Далее внимание будет уделено одному из важных и наименее освещенных в литературе практических приложений – анализу на основе ДЭМ дробного броуновского движения и, в частности, оцениванию показателя Херста.

Показатель Херста дробного броуновского движения

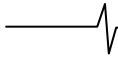
Показатель Херста [8,9] является важной характеристикой сигналов, используемой для оценки степени их хаотичности (близости по своим свойствам к шумовым сигналам) и регулярности. Согласно определению

$$R/S = (N/2)^H, \quad (20)$$

где R - размах сигнала (разница между минимальным и максимальным значениями), S - стандартное отклонение сигнала, N - число отсчетов в сигнале, H - значение показателя Херста.

Существует несколько способов оценивания значения показателя Херста [6-9], однако все они сильно зависят от выбора начальных параметров соответствующих алгоритмов. В данной статье предлагается новый метод оценивания показателя Херста на основе ДЭМ.

Дробное броуновское движение (ДБД) является распространенной моделью для многих прикладных задач. Например, оно широко используется в экономике (при анализе рынков, колебаний курсов валют на биржах) [8], гидрологии [9] и др. научных направлениях. Смысл названия "дробное" состоит в том, что энергетический спектр такого сигнала описывается обратным степенным законом вида $f^{-\nu}$, где ν - целое число. Согласно



определению, ДБД с показателем Херста H - гауссовский сигнал $x_H(k)$ с нулевым средним значением, для которого выполняются следующие условия:

$$x_H(0) = 0, \quad (21)$$

$$E\{[x_H(k) - x_H(k - \tau)]^2\} = \sigma^2 \tau^{2H}, \quad (22)$$

где E - оператор математического ожидания, σ^2 - дисперсия $x_H(k)$. ДБД является обобщением обычного броуновского движения (гауссовского сигнала с независимыми приращениями) со значением $H = 0.5$, производная которого есть белый гауссовский шум.

Определим АКФ сигнала, удовлетворяющего данным свойствам. Для этого необходимо учесть тот факт, что при задержке на $\tau = k$ справедливо соотношение

$$E\{[x_H(k)]^2\} = \sigma^2 k^{2H}. \quad (23)$$

Подставив в выражение для АКФ $\tau = k - u$ и используя вышеприведенное соотношение и свойство линейности оператора математического ожидания E , можно получить следующее соотношение для $R_1(k, \tau)$ (АКФ):

$$R_1(k, \tau = k - u) = E\{x_H(k) \cdot x_H(u)\} = 0.5 \cdot \sigma^2 [k^{2H} + u^{2H} - (k - u)^{2H}] \quad (24)$$

Можно показать, используя приведенные выше формулы и физический смысл величины H , что ее значения лежат в пределах $[0; 1]$. При этом диапазон значений $[0; 0.5]$ соответствует так называемым *антиперсистентным* сигналам или сигналам с *неустойчивым трендом* (называемым также *сигналами с кратковременной памятью*). Последнее означает, что сигнал стремится вернуться к своему среднему значению, и, следовательно, на всей длительности наблюдается чередование участков монотонного возрастания и убывания тренда. Чем ближе значение H к нулю, тем устойчивее общая тенденция к чередованию. Сигналы со значениями H из диапазона $[0.5; 1]$ соответствуют так называемым *персистентным* сигналам или сигналам с *устойчивым трендом*, (называемым также *сигналами с долговременной памятью*) для которых общая тенденция является долговременной и сохраняется на более длительном временном интервале, чем у антиперсистентных сигналов. Чем ближе значение H к 1, тем более устойчивой эта тенденция является.

Данные предположения основываются на том, что сигналы с кратко- и долговременной памятью имеют разные особенности АКФ. Сигнал с кратковременной памятью должен иметь абсолютно суммируемую АКФ (строгое доказательство имеется для реализации бесконечной длины, в практических задачах рассматривается ее конечный фрагмент) [9]:

$$\sum_{u=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} |R(k, u)| < \infty, \quad (25)$$

в то время как для сигнала с долговременной памятью вышеприведенный ряд является расходящимся, т.е. АКФ не суммируема абсолютно. Зная выражение для $R_1(k, \tau)$, непосредственной подстановкой можно убедиться, что для $H \in (0; 0.5]$ АКФ суммируема абсолютно (сигнал с кратковременной памятью), а для $H \in [0.5; 1]$ данное свойство не выполняется (сигнал с долговременной памятью).

Функция $R_1(k, \tau)$ зависит не только от относительного сдвига τ , но и от выбранного отсчета k , и поэтому ДБД является нестационарным сигналом. При этом данный тип сигналов обладает свойством *статистического самоподобия* [7], проявляющимся при масштабировании:

$$E\{x_H(nk)x_H(nu)\} = E\{n^H x_H(k)n^H x_H(u)\}, \quad (26)$$

при $n > 0$, что может быть выведено аналогично АКФ ДБД. Поскольку оба сигнала $x_H(nk)$ и $n^H x_H(k)$ являются гауссовскими, имеют одинаковые среднее значение и ковариацию (n^H - константа, не влияющая на закон распределения), то они обладают свойством самоподобия.

В силу нестационарности ДБД для него нельзя определить энергетический спектр (который будет использоваться в дальнейшем при оценке значений H). Поэтому далее переходят к рассмотрению приращений данного сигнала, которые являются стационарным сигналом. Разностный сигнал ДБД (РДБД) первого порядка определяется как

$$z_H(k) = x_H(k) - x_H(k - 1). \quad (27)$$

Для такого сигнала можно аналогично определить АКФ $R_2(\tau)$, зависящую на этот раз только от величины сдвига τ :

$$R_2(\tau) = 0.5 \cdot \sigma^2 [(\tau + 1)^{2H} + (\tau - 1)^{2H} - 2\tau^{2H}]. \quad (28)$$

Согласно теореме Винера-Хинчина, энергетический спектр $S_H(f)$ РДБД определяется как ДПФ от АКФ $R_2(\tau)$:

$$S_H(f) = C\sigma^2 |e^{i2\pi f} - 1|^2 \sum_{k=-\infty}^{\infty} \frac{1}{|f + k|^{2H+1}} = C\sigma^2 4 \sin^2(\pi f) \sum_{k=-\infty}^{\infty} \frac{1}{|f + k|^{2H+1}}, \quad (29)$$

где $|f| \leq 0.5$ - частота, нормированная к частоте Найквиста. При $H = 0.5$ (что соответствует белому гауссовскому шуму, т.к. $R_2(\tau) = 0$ при $\forall \tau$) сумма приобретает вид

$$\sum_{k=-\infty}^{\infty} \frac{1}{|f + k|^{2H+1}} = \sum_{k=-\infty}^{\infty} \frac{1}{(f + k)^2} = \frac{1}{4 \sin^2(\pi f)}, \quad (30)$$

$$S_H(f) = C\sigma^2 = const. \quad (31)$$

Исследуем значения и характер энергетического спектра для трех различных диапазонов значений показателя Херста H :

1) $H = 0.5$; $S_H(f) = C\sigma^2 = const$, что соответствует равномерному энергетическому спектру в полосе частот $|f| \leq 0.5$.

2) $0 < H < 0.5$; $S_H(0) = 0$, т.е. значение энергетического спектра равно нулю на нулевой частоте и увеличивается по мере приближения к частоте Найквиста. Таким образом, энергетический спектр имеет высокочастотный характер.

3) $0.5 < H < 1$; $S_H(0) = \infty$, т.е. значение энергетического спектра равно бесконечности на нулевой частоте и уменьшается по мере приближения к частоте Найквиста. Таким образом, энергетический спектр имеет низкочастотный характер.

На рис. 4 показано типичное поведение энергетических спектров ЭМ, соответствующих РДБД с разными

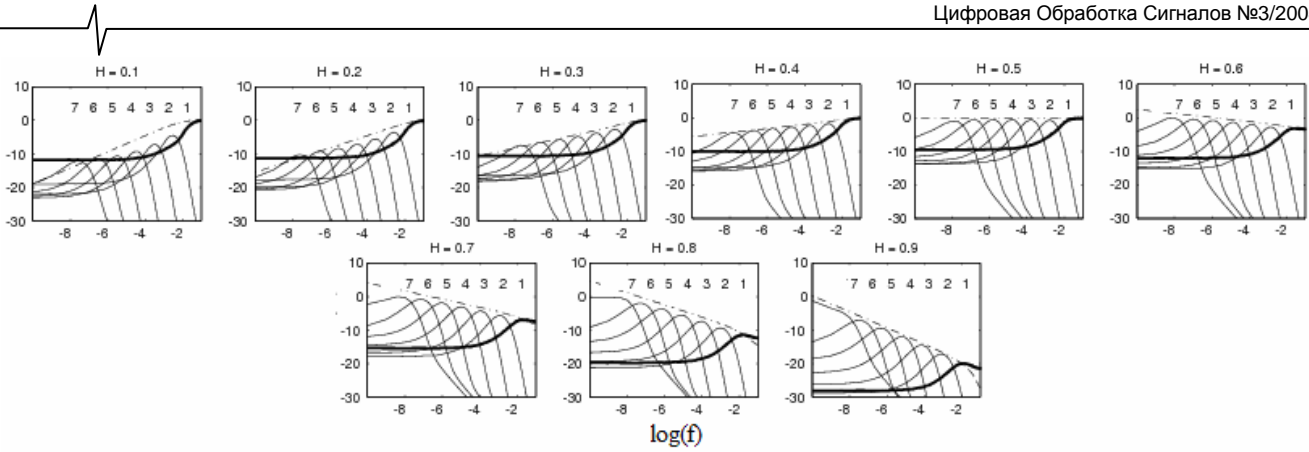


Рис. 4. Энергетические спектры ЭМ, полученных в результате декомпозиции РДБД с разными значениями H (на основе 5000 независимых реализаций длиной по 512 отсчетов)

значениями H . Из данной картины можно сделать следующие выводы:

Во-первых, первая мода РДБД, показанная на рис. 4 жирной линией, независимо от значения H , имеет ярко выраженный высокочастотный характер. При этом максимальное затухание в области низких частот составляет 10 дБ (по отношению к затуханию на частоте Найквиста) для $H < 0.6$, и увеличивается при больших значениях показателя Херста (при $H \geq 0.6$). Все остальные моды по структуре частотного спектра носят полосовой характер.

Во-вторых, с увеличением значения H от 0.1 до 0.9 энергетический спектр последней моды, представляющей собой либо константу, либо тренд среднего значения, постепенно меняет свой характер с полосового на низкочастотный.

В-третьих, огибающая семейства энергетических спектров также имеет определенную тенденцию. Для значений H , находящихся в диапазоне $0 < H < 0.5$, она представляет собой прямую с наклоном, уменьшающимся при приближении H к 0.5. Значению $H = 0.5$ соответствует нулевой наклон. Наконец, при изменении H в пределах $0.5 < H < 1$ огибающая имеет отрицательный наклон, увеличивающийся по абсолютной величине при стремлении H к 1.

В-четвертых, при рассмотрении ЭМ, имеющих полосовой характер спектра (всех, кроме первой и последней), можно проверить, что их ширина полосы частот уменьшается в 2 раза с увеличением номера. На этом основана интерпретация всего разложения как *диадического банка фильтров* [3].

Построение данных семейств спектров ЭМ, соответствующих РДБД, может использоваться для приближенной оценки показателя Херста в силу вышеизложенных свойств и соотношений, а именно на основе поведения огибающей семейства энергетических спектров.

Оценивание показателя Херста на основе метода ДЭМ

Основываясь на том факте, что ДЭМ можно интерпретировать как диадический банк фильтров, можно сделать вывод, что эффективная ширина Фурье-спектра каждой моды уменьшается в число раз, равное степени 2. Исходя из этого, энергетический спектр очередной моды определяется как

$$S_{p',H}(f) = 2^{\alpha(p'-p)} S_{p,H}(2^{p'-p} f), \tag{32}$$

где $\alpha = 2H + 1$ - экспериментально определенная константа, p, p' - номера ЭМ, причем $p' > p \geq 2$. Эта рекуррентная формула определяет энергетический спектр очередной ЭМ на основе известных спектров предыдущих ЭМ.

Зная энергетические спектры всех ЭМ, можно определить их дисперсии, причем для удобства дальнейших вычислений лучше представить выражение для дисперсий ЭМ в рекуррентном виде:

$$V_H[p'] = \int_{-1/2}^{1/2} S_{p',H}(f) df = 2^{\alpha(p'-p)} \int_{-1/2}^{1/2} S_{p,H}(2^{p'-p} f) df = 2^{\alpha(p'-p)} V_H[p]. \tag{33}$$

В вышеприведенной формуле использовалась теорема Винера-Хинчина, а также то, что интегрирование ведется по симметричному промежутку частот, нормированных к частоте Найквиста. Из нее можно получить в явном виде выражение для дисперсии p -ой моды как функции ее номера:

$$V_H[p] = C \cdot 2^{2Hp}, \tag{34}$$

где C - некоторая константа. Справедливость формулы устанавливается путем проверки рекуррентного соотношения (33). После логарифмирования данного выражения по основанию 2 получается окончательный результат:

$$\log_2 V_H[p] = \log_2 C + 2Hp. \tag{35}$$

Из этого выражения непосредственно вытекает способ экспериментальной оценки показателя Херста. Выполняется ДЭМ для данной реализации ДБД, рассчитываются значения дисперсий каждой моды и их двоичные логарифмы откладываются на графике. После этого на основе регрессии методом наименьших квадратов (МНК) строится линейная зависимость и определяется ее угловой коэффициент. Тогда само значение H будет найдено следующим образом:

$$H = \frac{|k_H|}{2}, \tag{36}$$

где k_H - угловой коэффициент прямой регрессии. Знак модуля нужен потому, что в отдельных случаях линеаризованная зависимость логарифма дисперсии ЭМ от ее номера может иметь отрицательный наклон, а показатель Херста - априорно положительная величина.

Альтернативный способ оценивания показателя Херста с использованием среднего периода ЭМ иллюстрируется следующей формулой:

$$V_H[p] = C' \cdot (T_H[p])^{2H}, \quad (37)$$

где C' - константа, $T_H[p'] = 2^{p'-p} T_H[p]$ для $p' > p \geq 2$, $T_H[k]$ - средний период p -ой ЭМ, определяемый как расстояние между первым и последним нулем (или максимумом, или минимумом) функции, деленное на число нулей (или максимумов, или минимумов). Рекуррентная формула, связывающая средние периоды двух ЭМ, основывается на интерпретации ДЭМ как диадического банка фильтров, и, следовательно, средний период ЭМ связан с ее номером как

$$T_H[p] = 2^p. \quad (38)$$

Из описанных процедур следует, что для оценки значения H необходимо определение дисперсии каждой из мод. Это делается, в основном, по двум формулам, дающим оценку по методу наименьших квадратов и по медиане абсолютных отклонений [7]. Первая имеет вид:

$$V_H[p] = \frac{1}{N-1} \sum_{k=1}^N c_p^2(k). \quad (39)$$

Выражение записано в предположении о центрированности ЭМ, которое обычно выполняется с высокой точностью. Главным преимуществом оценки является ее несмещенность, состоятельность и эффективность. Вторая оценка (40) существует в двух незначительно отличающихся модификациях, приведенных выше. Ее использование объясняется широкой распространенностью гауссовых сигналов, к числу которых относятся ДБД (данная оценка строго применима только к гауссовым сигналам) и высокой робастностью. Применение этой оценки имеет смысл только для начальных уровней разложения, которые являются высокочастотными, сильно осциллирующими и содержат значительную часть шумовой составляющей.

$$\sigma_{1H}[p] = \frac{\text{median}\{c_p - \text{median}(c_p)\}}{0.6745};$$

$$\sigma_{2H}[p] = \frac{\text{median}\{c_p\}}{0.6745}; \quad V_{1,2H}[p] = \sigma_{1,2H}^2[p] \quad (40)$$

Генерации ДБД с заданным значением H может осуществляться несколькими возможными способами. Один из наиболее известных - алгоритм Абри-Селлана-Мейера, использующий для этого биортогональные вейвлеты [6]. В качестве примера работы алгоритма рассмотрим две реализации сигналов ДБД (рис. 5), имеющих показатели Херста, соответственно, $H_1 = 0.3$ и $H_2 = 0.6$. Из рис. 5 видно, что реализация ДБД с $H_2 = 0.6$ является более регулярной и с более устойчивым трендом, чем та, у которой $H_1 = 0.3$.

После применения алгоритма ДЭМ к данным сигналам необходимо выбрать те ЭМ, на основании которых будет осуществляться оценка значения H . В качестве критерия отбора целесообразно использовать *выборочные средние значения* c_i всех выделенных ЭМ, так называемые компоненты с большими номерами, относящиеся к числу трендовых, могут существенно исказить экспериментально построенные зависимости (за счет большого разброса значений в случае монотонных трендовых компонент и нулевого разброса значений в случае констант), что в итоге приведет к неточной оценке показателя Херста. В данном случае к трендовым модам будут относиться те моды, среднее значение которых по модулю больше 0.04 (при таком пороге достигается наибольшая точность оценки). Не следует также забывать, что первая ЭМ не используется при оценивании в силу особенностей ее спектра, который имеет ярко выраженный высокочастотный характер (рис. 4) в отличие от остальных компонент, имеющих полосовой спектр. В таблицах 1 и 2 жирным выделены те ЭМ, которые участвуют в оценивании.

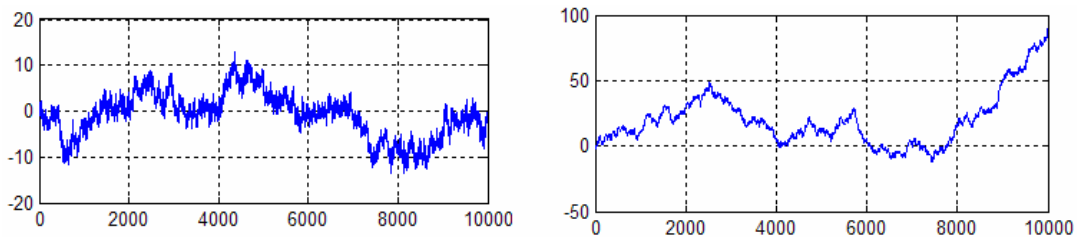


Рис. 5. Две реализации ДБД с показателями Херста $H_1 = 0.3$ (слева) и $H_2 = 0.6$ (справа)

Таблица 1.

Средние значения ЭМ для ДБД с $H_1 = 0.3$.

№ ЭМ (i)	1	2	3	4	5	6	7	8	9	10	11	12
$\bar{c}_i \cdot 10^3$	-0.9	0.8	0.6	-5.2	7.2	15.8	-41.9	-16.1	0.9	6.4	-580.8	-312.2

Таблица 2.

Средние значения ЭМ для ДБД с $H_2 = 0.6$.

№ ЭМ (i)	1	2	3	4	5	6	7	8	9	10
$\bar{c}_i \cdot 10^3$	-0.1	1.1	4.3	-0.4	3.1	-6.4	-17.9	0.5	-2200.7	22038.4

Таблица 3.

Сравнения оценок для сигнала ДБД с $H_1 = 0.3$

Название оценки	На основе ДЭМ-1	На основе ДЭМ-2	На основе биортогональных вейвлетов	На основе R/S анализа
Величина	0.3038	0.2958	0.3010	0.3146

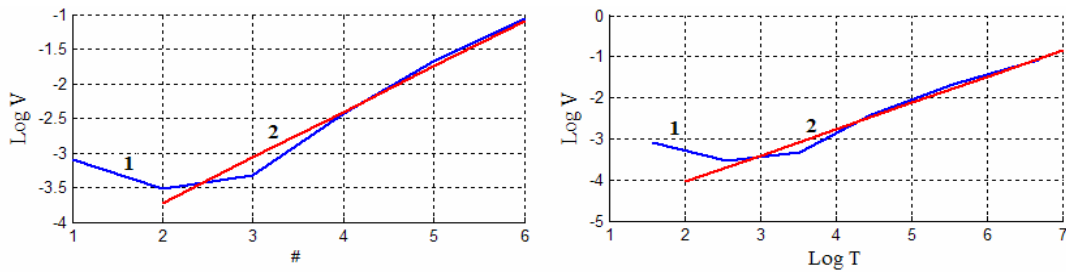


Рис. 6. Зависимость логарифма дисперсии от номера ЭМ (слева) и от среднего периода ЭМ (справа) для ДБД с $H_1 = 0.3$ 1 – экспериментальные значения, 2 – регрессия по МНК

Таблица 4.

Сравнения оценок для сигнала ДБД с $H_2 = 0.6$

Название оценки	На основе ДЭМ-1	На основе ДЭМ-2	На основе биортогональных вейвлетов	На основе R / S анализа
Величина	0.6996	0.6021	0.6454	0.5932

Далее, после определения соответствующих ЭМ (со 2-й по 6-ю для ДБД с $H_1 = 0.3$ и со 2-й по 8-ю для ДБД с $H_2 = 0.6$) строится зависимость логарифма дисперсии ЭМ от их номера (первая оценка) и логарифма дисперсии ЭМ от среднего периода (вторая оценка), как было описано выше. Выше (рис. 6) приведены графики, на основании которых определяются две оценки показателя Херста для сигнала ДБД с $H_1 = 0.3$ (для сигнала с $H_2 = 0.6$ все делается аналогично).

В таблицах 3 и 4 приведены найденные значения оценок H для 2-х рассмотренных реализаций ДБД и здесь же приведены результаты сравнения двух оценок на основе ДЭМ с другими оценками, полученными в результате использования аналогичного ДЭМ разложения сигнала с помощью биортогональных вейвлетов и R / S анализа – наиболее распространенных на сегодняшний день альтернативных способов оценки показателя Херста.

Как видно из полученных результатов, оценка по методу ДЭМ в первом случае незначительно уступает по точности оценке на основе вейвлет-преобразования, однако это справедливо лишь при верном выборе типа вейвлет-базиса, что довольно затруднено на практике в связи с их большим многообразием и необходимостью тщательного и порой затратного перебора. К тому же, правильный выбор соответствующего вейвлет-базиса для обеспечения высокой точности во многом зависит от алгоритма генерации самого ДБД, который зачастую неизвестен. В данных примерах использовался алгоритм генерации на основе биортогональных вейвлетов, что и объясняет столь высокую точность результатов на их основе. Другие варианты (вейвлет-базис Хаара, Добеши и пр.) дают неудовлетворительные результаты (абсолютная погрешность оценивания может достигать 0.1, что неприемлемо). Кстати, даже на основе биортогональных вейвлетов оценка показателя Херста для реализации с $H_2 = 0.6$ проигрывает по точности оценке на основе ДЭМ. Имеет смысл также сравнивать представленные результаты с результатами, полученными на основании R / S анализа [8], который применяется непосредственно к самой реализации во временной области, что часто обеспечивает высокую точность оценивания.

Заключение

В статье представлен принципиально новый подход к анализу нестационарных сигналов сложной внутренней структуры – алгоритм декомпозиции на эмпирические моды. Описаны идеи, заложенные в его основу, а также основные

преимущества по сравнению с существующими методами анализа нестационарных сигналов. В качестве примера рассмотрена задача анализа дробного броуновского движения, в частности, оценивание его показателя Херста. В результате проведенных экспериментов на модельных сигналах достигнута очень высокая точность результатов. Основное достоинство предлагаемого подхода в том, что по сравнению с существующими альтернативными методами (вейвлет-анализ, R/S - анализ) устранен весьма существенный недостаток - необходимость выбора начальных параметров соответствующих алгоритмов оценивания (таких как вид вейвлет-базиса, длина окна во временной области при использовании R/S - анализа). Полученные результаты позволят в дальнейшем применять ДЭМ для оценки показателя Херста реальных сигналов – в медицине (кардиологии, неврологии), экономике, гидрологии и других областях.

Литература

- Huang N., Shen S. Hilbert-Huang Transform and Its Applications, World Scientific, 2005.
- Huang N.E., etc. The Empirical Mode Decomposition and The Hilbert Spectrum for Non-linear and Non-stationary Time Series Analysis, Proc. Royal Soc. London, Vol. 454, pp. 903-995, 1998.
- Flandrin P., Rilling G., Gonsalves P. Empirical mode decomposition as a filter bank, IEEE Signal Process. Lett., Vol. 11, pp. 112-114, 2004.
- Klionsky D.M, Geppener V.V. Empirical mode decomposition in segmentation of slowly and fast changing non-stationary signals, Proceedings of PRIA Conference, Yoshkar-Ola, Russia, 2007.
- Марпл-мл. С.Л. Цифровой спектральный анализ и его приложения: Пер. с англ., изд. "Мир", Москва, 1990.
- Смоленцев Н. К. Основы теории вейвлетов. Вейвлеты в Matlab, изд. "ДМК", Москва, 2008.
- Малла С. Вейвлеты в обработке сигналов, изд. "МИП", Москва, 2005.
- Петерс Эдгар Э. Фрактальный анализ финансовых рынков, изд. "Интернет-Трейдинг", Москва, 2004.
- McLeod H. Time Series Modelling of Water Resources and Environmental Systems, Elsevier, Amsterdam, 1994.
- Longuet-Higgins M.S. The statistical analysis of random moving surface, Proc. R. Soc. Lond., 1957.
- Мэтьюз Д. Г., Финк К. Д. Численные методы. Использование Matlab, изд. дом "Вильямс", Москва-Санкт-Петербург-Киев, 2001.
- Вержбицкий В.М. Основы численных методов, изд. "Высшая школа", Москва, 2002.

НОВЫЙ СИММЕТРИЧНЫЙ АЛГОРИТМ ШИФРОВАНИЯ

Вишневский К.П., Чижиков В.И., Барышенский Д.С., Жучкова В.В.

Введение

При шифровании большого объема данных в реальном времени требуются быстрые алгоритмы. Для этих целей подходят поточные шифры. Они, как правило, похожи на одноразовый шифр-блокнот. Его суть заключается в сложении по модулю 2 битов потока ключей с битами сообщения. Однако вместо фиксированного потока ключей, который в ранних поточных шифрах и являлся секретным ключом, в современных системах поток ключей генерируется из короткого основного ключа с помощью однозначно определенных детерминированных алгоритмов. Над развитием разных способов их комбинирования с целью создания быстрых поточных шифров работают многие лаборатории.

Используемые на практике симметричные и несимметричные криптосистемы должны удовлетворять определенным требованиям:

- 1) знание использованного алгоритма не должно снижать надежность шифрования;
- 2) зашифрованный текст не может быть прочитан без знания ключа;
- 3) каждый ключ из многообразия ключей должен обеспечивать достаточную надежность;
- 4) изменение длины ключа не должно приводить к изменению алгоритма шифрования;
- 5) если известны зашифрованный и открытый тексты, то число операций, необходимых для определения ключа, не должно быть меньше полного числа возможных ключей;
- 6) дешифрование текста путем перебора всех возможных ключей должно выходить далеко за пределы возможностей компьютеров, используемых на практике;
- 7) не должно быть легко устанавливаемой зависимости между последовательно используемыми ключами;
- 8) исходный и зашифрованный тексты должны быть одинакового размера;
- 9) алгоритм может быть реализован аппаратно.

Следует подчеркнуть, что для обратного преобразования зашифрованного текста часто используются два термина: расшифровывание и дешифрование. При расшифровывании ключ считается известным, а при дешифровании ключ неизвестен. Однако мы не будем делать различия между ними.

Согласно теореме К. Шеннона [1] совершенной секретностью обладает традиционная система шифрования Г. Вернама или одноразовый шифр-блокнот. В ней для пересылки битов необходим ключ, содержащий тоже количество двоичных знаков, что и сообщение. Каждый бит посылаемого текста складывается по модулю 2 с соответствующим битом ключа и получается криптограмма. Любой ключ разрешается использовать только один раз. Следовательно, распределение ключей в этой системе является основной проблемой [2].

Предложен новый симметричный способ шифрования большого объема данных в реальном времени. Он является наиболее быстрым среди уже известных симметричных методов. Важное преимущество данного подхода состоит в возможности изменения длины ключа без изменения алгоритма шифрования. Ключ может быть любой длины.

В принципе любой способ шифрования сообщения, использующего битовое представление информации, можно формально рассматривать как систему Г. Вернама. Действительно, пусть M – сообщение, C – криптограмма и K – ключ. Тогда $C = M \oplus K$, $K = C \oplus M$, $M = C \oplus K$, т.е. знание C и M позволяет рассматривать любой способ шифрования как одноразовый шифр-блокнот.

В данной статье мы предлагаем измененный способ реализации этой системы шифрования. Основным элементом – ключ генерируется случайным образом. Причем его длина меньше длины сообщения ($K \leq M$). Само сообщение перед шифрованием зашумляется. Это способствует уменьшению длины ключа. Кроме того, на основе операции сложения по модулю 2 образуется информационная свертка. Она играет основную роль в предлагаемом способе шифрования, который мы назвали SEA (симметричный алгоритм шифрования).

Концепция нового способа шифрования

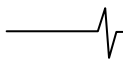
Принцип работы симметричного алгоритма шифрования (SEA) основан на применении информационной свертки – специальным образом организованной процедуры поглощения битов.

Для ее демонстрации рассмотрим процедуру информационного свертывания на примере произвольной битовой строки. Свертка данных происходит следующим образом: берутся два смежных входных бита и на основе их значений устанавливается значение выходного бита. Далее процесс повторяется, т.е. рассматриваются следующие два смежных бита. Важной особенностью информационной свертки является возможность обратного преобразования при условии применения симметричного правила преобразования. В данном случае (битовая строка) используется следующая таблица преобразований (табл. 1). Очевидно, что эта процедура для двух битов является сложением по модулю два ($A \oplus B$).

Таблица 1.

Таблица преобразований значений для 2 битов

Входные значения	Выходное значение
0 0	0
0 1	1
1 0	1
1 1	0



Как видно из таблицы, преобразование является симметричным. На рис. 1 показан принцип работы информационной свертки: после каждого этапа свертывания размер данных уменьшается на 1 бит.

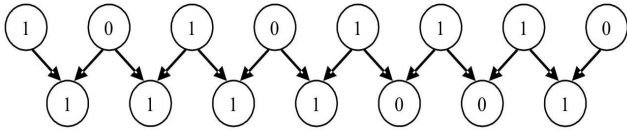


Рис. 1. Информационная свертка

При обратном процессе получения информационной развертки размер данных увеличивается на 1 бит. Однако в результате развертки образуются две битовых строки, свертка которых даст исходную строку (рис. 2). Данное свойство информационной свертки можно использовать для шифрования данных. Рассмотрим два вида шифрования с фиксированным ключом и с генерируемым ключом в процессе свертки.

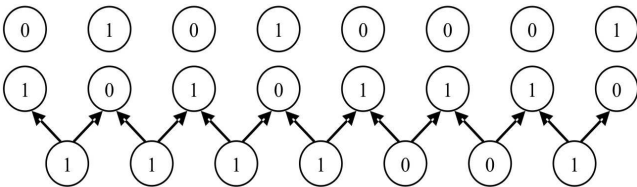


Рис. 2. Информационная развертка

Шифрование данных с созданием ключа производится путем циклической свертки данных (рис. 3), в результате чего зашифрованные данные имеют уменьшенный размер. Дешифрование происходит во время циклической развертки с выбором выходного варианта таким образом, что бы он совпадал с ключом.

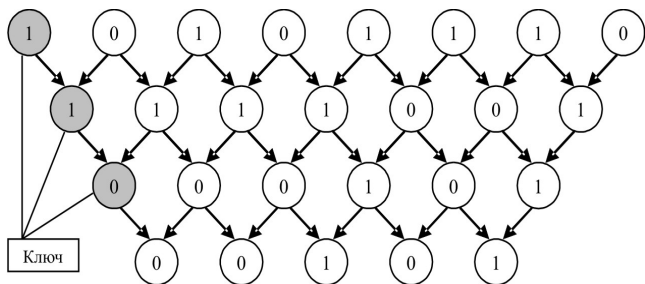


Рис. 3. Шифрование с генерацией ключа

Шифрование данных с фиксированным ключом, несколько отличается от рассмотренного выше метода. Данные так же шифруются с применением циклической свертки, однако на каждом этапе к строке прибавляется очередной бит ключа. В результате такого шифрования размер данных остается неизменным. Дешифрование происходит во время циклической развертки с выбором выходного варианта таким образом, что бы он совпадал с ключом. При этом после каждого шага развертки происходит отбрасывание ключевого бита.

Иногда в процессе шифрования/дешифрования выгоднее использовать не битовые строки, а байтовые строки. В этом случае преобразование байтов A и B проводится по формуле: $C = A \text{ XOR } B$.

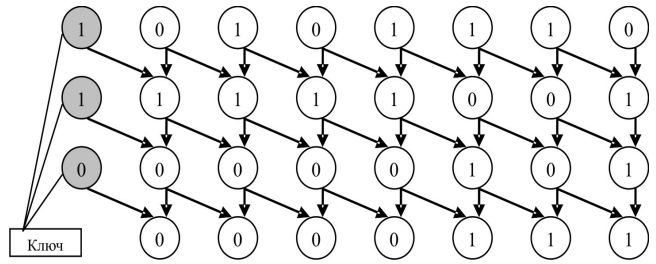


Рис. 4. Шифрование с фиксированным ключом

Для повышения криптостойкости метода, можно использовать простой алгоритм зашумления: $M_0 = \text{Key}$, $N_i = M_i \text{ XOR } M_i$, $M_i = M_i \cdot Z_i^3$, $Z_i = r \cdot k$, где N_i – текущий байт информации, M_i и Z_i – рекурсивные параметры зашумления, k – произвольный коэффициент.

Сравнение различных способов шифрования

Для сравнения различных способов шифрования возьмем следующий рисунок 5: черный круг на белом фоне. При шифровании размер рисунка сохраняется.

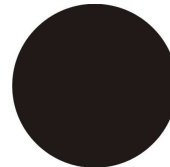


Рис. 5. Исходный объект для шифрования

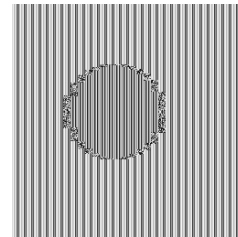


Рис. 6. Метод 3DES

Зашифрованный методом 3DES текст посредством формата raw показан на рис. 6.

Аналогично зашифрованные методами DES, BLOWFISH, RIJNDAEL и SEA тексты показаны на рис. 7, 8, 9 и 10 соответственно.

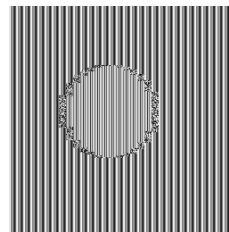


Рис. 7. Метод DES

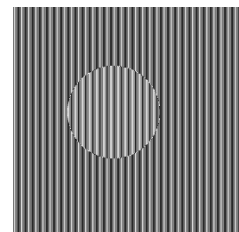


Рис. 8. Метод BLOWFISH

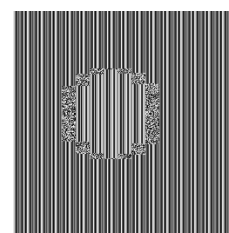


Рис. 9. Метод RIJNDAEL

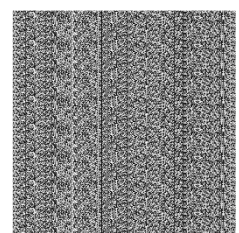
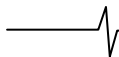


Рис. 10. Метод SEA



Скорости шифрования при использовании различных методов приведены в таблице 2. При сравнении в предлагаемом алгоритме SEA длина ключа составляла 128 байтов.

Таблица 2.
Сравнение скорости работы различных алгоритмов

Алгоритм	Скорость на P4 2.6 ГГц
SEA	247 Мбит/сек
Blowfish	128 Мбит/сек
Rijndael	54 Мбит/сек
3DES	24 Мбит/сек
DES	21 Мбит/сек

Заключение

Предложенный способ шифрования SEA (Symmetric Encryption Algorithm) является наиболее быстрым среди уже известных симметричных методов. При

аппаратной реализации шифрование может происходить со скоростью N тактов на блок, где N длина ключа. Блок же может достигать больших размеров (мегабайт и более). Самым важным преимуществом SEA перед другими симметричными способами является возможность изменять длину ключа без изменения алгоритма шифрования. Она может быть любой. Естественно, что скорость шифрования зависит от длины ключа. Поскольку длина ключей не велика, то для их распространения можно использовать известные методы.

Литература

1. Введение в криптографию. Под общей редакцией Яценко В. В. – М.: МЦНМО: “Черо”, 1999. – 272 с.
2. Насыпный В. В. Одноразовое шифрование с открытым распределением ключей // Открытые системы. 2004, №1. С. 66–69.

НОВЫЕ КНИГИ

Цифровая обработка сигналов. Практическое руководство для инженеров и научных работников (+CD) / Стивен Смит; пер. с англ. А. Ю. Линовича, С.В. Витязева., И.С. Гусинского. – М. : Додэка-XXI, 2008. — 720 с. : ил. — (Серия «Схемотехника»).

ISBN 978-5-94120-145-7

Заказ: www.dokabooks.ru

В книге изложены основы теории цифровой обработки сигналов. Акцент сделан на доступности изложения материала и объяснении методов и алгоритмов так, как они понимаются при практическом использовании. Цель книги — практический подход к цифровой обработке сигналов, позволяющий преодолеть барьер сложной математики и абстрактной теории, характерных для традиционных учебников. Изложение материала сопровождается большим количеством примеров, иллюстраций и текстов программ (которые вы также можете найти на прилагаемом CD). Книга предназначена научным работникам и инженерам, желающим применять методы цифровой обработки в различных технических сферах. Рекомендуется аспирантам и студентам, изучающим цифровую обработку сигналов.

ЦОС В УНИВЕРСИТЕТАХ

В течение ряда лет, начиная с первого выпуска в 1999 году, редакция журнала систематически отслеживает появление на мировом и российском рынках новых перспективных DSP-технологий и знакомит читателей с тенденциями и направлениями развития цифровых сигнальных процессоров и инструментальных средств проектирования систем ЦОС на их основе. За прошедшие годы неоднократно публиковались статьи, отражающие развитие DSP-технологий таких мировых лидеров в этой области, как компании Texas Instruments Inc. и Analog Devices Inc. Что особенно приятно отметить – редакция журнала всегда откликалась одной из первых на новые разработки в области DSP-технологий отечественных фирм: НТЦ «Модуль», ГУП НПЦ «ЭЛВИС», ЗАО «Инструментальные системы», ЗАО «АВТЭК», ЗАО «СКАН Инжиниринг-телеком» и др.

Отвечая целям и задачам укрепления и расширения сотрудничества с ведущими мировыми и отечественными производителями современной элементной базы DSP-технологий и средств проектирования систем ЦОС, с одной стороны, и российскими вузами, с другой стороны, редакция журнала открыла новую рубрику: «ЦОС в университетах». В рамках новой рубрики найдут отражение вопросы организации учебного процесса и учебно-методического обеспечения по широкому спектру дисциплин, связанных с применением ЦОС и DSP-технологий, текущая информация об университетских программах, семинарах и конкурсах фирм-производителей, новых разработках научных лабораторий российских вузов.

Приглашаем к сотрудничеству все заинтересованные организации и творческие коллективы преподавателей и сотрудников российских вузов и стран СНГ, работающих в области ЦОС и DSP-технологий.

Зам. Главного редактора, профессор В. В. Витязев

УДК 621.397

ЧАСТОТНЫЕ СВОЙСТВА ДВУМЕРНЫХ РЕКУРСИВНЫХ ЦИФРОВЫХ ФИЛЬТРОВ ВТОРОГО ПОРЯДКА С СИММЕТРИЧНЫМИ КОЭФФИЦИЕНТАМИ

Балусов И.Л., Приоров А.Л., Хрящев В.В.

Введение

Для эффективной цифровой обработки двумерных сигналов необходимо знание частотных свойств соответствующих цифровых фильтров различного порядка. Именно эти свойства являются одними из важнейших характеристик системы, определяющих возможность ее применения для конкретных задач обработки сигналов [1]. Поэтому получение заданных частотных характеристик является одним из исходных условий в задаче построения систем обработки информации, в том числе и в задаче синтеза цифровых фильтров. Возникает и обратная задача – определения влияния параметров системы на ее частотную избирательность.

Двумерная рекурсивная цифровая система (фильтр) второго порядка с симметричными коэффициентами определяется следующим разностным уравнением с постоянными коэффициентами:

$$y(m,n) = b_{10}y(m-1,n) + b_{01}y(m,n-1) + b_{11}y(m-1,n-1) + b_{20}y(m-2,n) + b_{02}y(m,n-2) + x(m,n), \quad (1)$$

где $x(m,n)$ и $y(m,n)$ – входной и выходной двумерные сигналы соответственно, а b_{ij} – коэффициенты (параметры) фильтра. Передаточная функция исследуемой системы (1) записывается следующим образом:

$$H(z_1, z_2) = \frac{1}{1 - b_{10}z_1^{-1} - b_{01}z_2^{-1} - b_{11}z_1^{-1}z_2^{-1} - b_{20}z_1^{-2} - b_{02}z_2^{-2}} \quad (2)$$

$$= \frac{1}{1 - a(z_1^{-1} + z_2^{-1}) - b(z_1^{-2} + z_2^{-2}) - cz_1^{-1}z_2^{-1}}.$$

Здесь $b_{01} = b_{10} = a$, $b_{20} = b_{02} = b$, $b_{11} = c$. Анализируя (2), можно установить, что условия устойчивости для рассматриваемой системы имеют вид:

$$\begin{cases} c < 1 - 2a - 2b, \\ c < 1 + 2a - 2b, \\ c > 2b - 1, \\ c > -2b - 1, \\ c < 1 + 2b. \end{cases} \quad (3)$$

Фигура устойчивости системы, для которой условия (3) выполняются, изображена на рис. 1.

Частотную характеристику двумерного фильтра найдем, сделав подстановку $z_1 = e^{j\omega_1}$, $z_2 = e^{j\omega_2}$ в передаточную функцию (2)

$$H(e^{j\omega_1}, e^{j\omega_2}) = \frac{1}{1 - a(e^{-j\omega_1} + e^{-j\omega_2}) - b(e^{-j2\omega_1} + e^{-j2\omega_2}) - ce^{-j(\omega_1 + \omega_2)}}. \quad (4)$$

Объектом исследования является двумерный рекурсивный цифровой фильтр второго порядка с симметричными коэффициентами. Определены области существования двумерных фильтров нижних и верхних частот с монотонными амплитудно-частотными характеристиками. Установлено максимальное подавление сигнала, возможное в фильтрах нижних и верхних частот. Получено уравнение линии среза. В пространстве параметров определены и классифицированы области существования двумерных полосовых и режекторных фильтров. Приведены примеры амплитудно-частотных характеристик полученных фильтров.

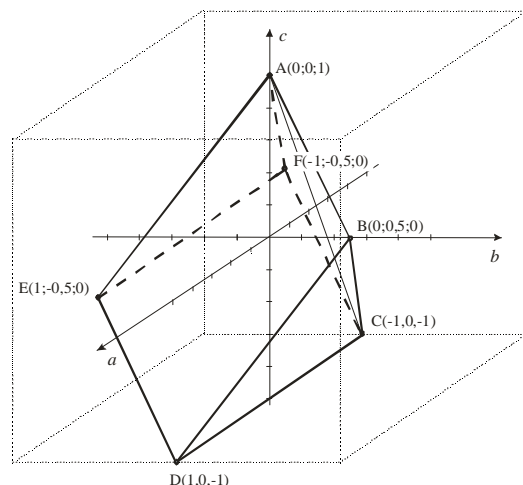


Рис. 1. Фигура устойчивости двумерного рекурсивного цифрового фильтра второго порядка с симметричными коэффициентами

Введем обозначение для квадрата модуля частотной характеристики (4)

$$|H(\omega_1, \omega_2)|^2 = \frac{1}{f(\omega_1, \omega_2)}, \quad (5)$$

где

$$f(\omega_1, \omega_2) = (1 - a(\cos\omega_1 + \cos\omega_2) - c\cos(\omega_1 + \omega_2) - b(\cos 2\omega_1 + \cos 2\omega_2))^2 + (a(\sin\omega_1 + \sin\omega_2) + c\sin(\omega_1 + \omega_2) + b(\sin 2\omega_1 + \sin 2\omega_2))^2. \quad (6)$$

Необходимо исследовать форму амплитудно-частотной характеристики (АЧХ) системы. Анализ частотных свойств с использованием выражений (5-6) будем проводить методами дифференциальной геометрии в пространстве [2], ранее использованными при исследовании двумерных цифровых систем первого порядка [3].

Методика исследований

Предлагаемая методика исследования частотных характеристик двумерного цифрового фильтра основана на том, что его АЧХ представляется поверхностью в трехмерном пространстве $(\omega_1, \omega_2, |H|)$, имеющей свои экстремальные точки. Для анализа поведения поверхности в окрестности этих точек применим аппарат дифференциальной геометрии в пространстве [1], позволяющий определить, является ли точка локальным максимумом или минимумом или это какой-то другой тип точки.

Пусть $|H|^2$ – поверхность квадрата модуля частотной характеристики (рис. 2) и $\vec{r} = \vec{r}(\omega_1, \omega_2, |H|^2)$ – ее радиус-вектор. Выберем на поверхности $|H|^2$ некоторую точку X_0 и рассмотрим плоскость π_0 , которая касается поверхности $|H|^2$ в этой точке, а также точку S_0 , в которой поверхности $|H|^2$ касается плоскость π_1 .

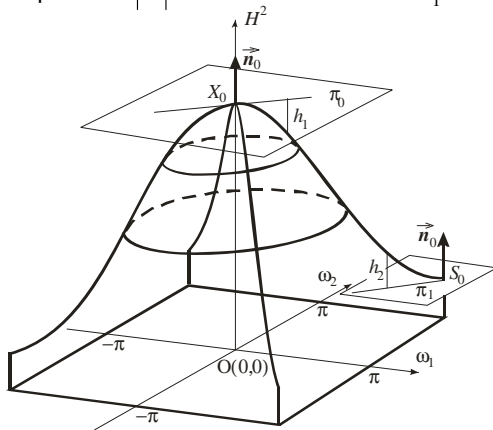


Рис. 2. Графическая иллюстрация метода исследования двумерной АЧХ

О типе фильтра (фильтр нижних частот (ФНЧ), фильтр верхних частот (ФВЧ), ...) можно судить по знаку отклонений h_1 и h_2 поверхности квадрата модуля его частотной характеристики (ЧХ) от касательных плоскостей π_0 и π_1 в точках $X_0(0,0)$ и $S_0(\pi, \pi)$ соответственно. Далее вместо отклонений h_1 и h_2 используется отклонение h , которое находится одинаково для обоих случаев.

Отклонение $h = h_1$ произвольной точки X поверхности $|H|^2$ от плоскости π_0 определим формулой

$$h = (\vec{r}(X) - \vec{r}(X_0), \vec{n}_0),$$

где $\vec{n}_0 = \frac{\vec{r}_{\omega_1} \times \vec{r}_{\omega_2}}{|\vec{r}_{\omega_1} \times \vec{r}_{\omega_2}|}$ – единичный вектор нормали

к поверхности в точке X_0 ; $\vec{r}_{\omega_1} = (1, 0, H_{\omega_1}^2)$, $\vec{r}_{\omega_2} = (0, 1, H_{\omega_2}^2)$ – производные радиуса-вектора по соответствующим координатам (для плоскости π_1 имеем $h = h_2$). Это отклонение, взятое по абсолютной величине, равно расстоянию от точки X до плоскости π_0 . Отклонение положительно, если точка X и конец вектора \vec{n}_0 лежат по одну сторону от плоскости π_0 и отрицательно, если эти точки лежат по разные стороны от плоскости π_0 . Разложим функцию $\vec{r}(X)$ в окрестности точки X_0 в ряд Тейлора с точностью до квадратичных членов, тогда разность $(\vec{r}(X) - \vec{r}(X_0))$ можно записать следующим образом:

$$\begin{aligned} \vec{r}(X) - \vec{r}(X_0) = & \vec{r}_{\omega_1}(X_0)(\omega_1 - \omega_{10}) + \vec{r}_{\omega_2}(X_0)(\omega_2 - \omega_{20}) + \\ & + \frac{1}{2} \{ \vec{r}_{\omega_1 \omega_1}(X_0)(\omega_1 - \omega_{10})^2 + 2\vec{r}_{\omega_1 \omega_2}(X_0)(\omega_1 - \omega_{10})(\omega_2 - \omega_{20}) + \\ & + \vec{r}_{\omega_2 \omega_2}(X_0)(\omega_2 - \omega_{20})^2 \} + o(\rho^2) \end{aligned} \quad (7)$$

$$\text{где } \rho = \sqrt{(\omega_1 - \omega_{10})^2 + (\omega_2 - \omega_{20})^2}.$$

Умножив обе части равенства (7) на \vec{n}_0 , получим

$$\begin{aligned} h = & (\vec{r}_{\omega_1}, \vec{n}_0)(\omega_1 - \omega_{10}) + (\vec{r}_{\omega_2}, \vec{n}_0)(\omega_2 - \omega_{20}) + \\ & + \frac{1}{2} \{ (\vec{r}_{\omega_1 \omega_1}, \vec{n}_0)(\omega_1 - \omega_{10})^2 + 2(\vec{r}_{\omega_1 \omega_2}, \vec{n}_0) \times \\ & \times (\omega_1 - \omega_{10})(\omega_2 - \omega_{20}) + (\vec{r}_{\omega_2 \omega_2}, \vec{n}_0)(\omega_2 - \omega_{20})^2 \} + o(\rho^2) = \\ = & \frac{1}{2} \{ L(X_0)d\omega_1^2 + 2M(X_0)d\omega_1 d\omega_2 + N(X_0)d\omega_2^2 \} + o(\rho^2), \end{aligned}$$

где $d\omega_1 = \omega_1 - \omega_{10}$, $d\omega_2 = \omega_2 - \omega_{20}$, а

$$L = \frac{(\vec{r}_{\omega_1 \omega_1}, \vec{r}_{\omega_1}, \vec{r}_{\omega_2})}{\sqrt{EG - F^2}}, \quad M = \frac{(\vec{r}_{\omega_1 \omega_2}, \vec{r}_{\omega_1}, \vec{r}_{\omega_2})}{\sqrt{EG - F^2}},$$

$$N = \frac{(\vec{r}_{\omega_2 \omega_2}, \vec{r}_{\omega_1}, \vec{r}_{\omega_2})}{\sqrt{EG - F^2}} -$$

коэффициенты второй квадратичной формы поверхности, причем $E = r_{\omega_1}^2$, $G = r_{\omega_2}^2$, $F = (\vec{r}_{\omega_1}, \vec{r}_{\omega_2})$, – коэффициенты первой квадратичной формы.

Для определения знака отклонения h проведем анализ знакоопределенности квадратичной формы. Составим матрицу квадратичной формы

$$A(X_0) = \begin{bmatrix} L(X_0) & M(X_0) \\ M(X_0) & N(X_0) \end{bmatrix}.$$

Согласно критерию Сильвестра, для положительной определенности требуется положительность главных миноров матрицы $A(X_0)$, т.е. в нашем случае

$$\begin{cases} L(X_0) > 0; \\ L(X_0)N(X_0) - M^2(X_0) > 0. \end{cases} \quad (8)$$

Для отрицательной определенности необходимо чередование знаков главных миноров матрицы $A(X_0)$, причем первый минор должен быть меньше нуля. В нашем случае имеем

$$\begin{cases} L(X_0) < 0; \\ L(X_0)N(X_0) - M^2(X_0) > 0. \end{cases} \quad (9)$$

При выполнении условий (8) или (9) отклонение h сохраняет знак (совпадающий со знаком второй квадратичной формы) для всех достаточно малых значений ρ независимо от выбора направления на поверхности.

Исследование частотных свойств

Первоначально найдем экстремальные точки поверхности. Частотная характеристика двумерной цифровой системы периодична с одинаковыми периодами по обоим частотам, равными 2π . Рассматриваемая система имеет действительную импульсную характеристику, что уменьшает интервал однозначности вдвое по каждой частоте. Рассмотрим квадрат амплитудно-частотной характеристики. Из условий симметрии коэффициентов достаточно рассмотреть поведение функ-

ции (4) на осях ω_1 и ω_2 , а также на диагоналях $\omega_1 = \omega_2$ и $\omega_1 = -\omega_2$ в верхней полуплоскости частотной плоскости (ω_1, ω_2) .

а) Пусть $\omega_1 = 0$, тогда

$$f(\omega_1, \omega_2) = f(\omega_2) = ((1-a-b) - (a+c)\cos\omega_2 - b\cos 2\omega_2)^2 + ((a+c)\sin\omega_2 + b\sin 2\omega_2)^2 \quad (10)$$

Найдем нули производной $\frac{df(\omega_2)}{d\omega_2}$:

$$\frac{df(\omega_2)}{d\omega_2} = \sin\omega_2(4b(1-a-b)\cos\omega_2 + 2(a+c)(1-a-2b)).$$

Видно, что $\frac{df(\omega_2)}{d\omega_2} = 0$ при $\sin(\omega_2) = 0$, откуда

$$\omega_{2,1,2} = 0, \pi \text{ на рассматриваемом отрезке, а из выражения } (a+c)(1-a-2b) + 2b(1-a-b)\cos\omega_2 = 0 \text{ получаем}$$

$$\cos\omega_2 = \frac{(a+c)(a+2b-1)}{2b(1-a-b)}, \quad (11)$$

причем должно выполняться условие

$$\left| \frac{(a+c)(a+2b-1)}{2b(1-a-b)} \right| \leq 1. \quad (12)$$

б) Пусть $\omega_2 = 0$, тогда

$$f(\omega_1, \omega_2) = f(\omega_1) = ((1-a-b) - (a+c)\cos\omega_1 - b\cos 2\omega_1)^2 + ((a+c)\sin\omega_1 + b\sin 2\omega_1)^2 \quad (13)$$

$$\frac{df(\omega_1)}{d\omega_1} = 2\sin\omega_1((a+c)(1-a-2b) + 4b\cos\omega_1).$$

Здесь $\frac{df(\omega_1)}{d\omega_1} = 0$ при $\omega_{1,1,2} = 0, \pm\pi$ и

$$(a+c)(1-a-2b) + 4b\cos\omega_1 = 0, \text{ откуда}$$

$$\cos\omega_1 = \frac{(a+c)(a+2b-1)}{4b}, \quad (14)$$

при этом

$$\left| \frac{(a+c)(a+2b-1)}{4b} \right| \leq 1. \quad (15)$$

в) Пусть $\omega_1 = \omega_2 = \omega_+$, тогда

$$f(\omega_1, \omega_2) = f(\omega_+) = (1-2a\cos\omega_+ - (2b+c)\cos 2\omega_+)^2 + (2a\sin\omega_+ + (2b+c)\sin 2\omega_+)^2 \quad (16)$$

$$\frac{df(\omega_+)}{d\omega_+} = 2\sin\omega_+(2a(1-2b-c) + 4(2b+c)\cos\omega_+).$$

Найдем нули полученной функции:

$$\frac{df(\omega_+)}{d\omega_+} = 0 \text{ при } \sin\omega_+ = 0, \text{ откуда } \omega_+ = 0, \pm\pi \text{ и}$$

$$2a(1-2b-c) + 4(2b+c)\cos\omega_+ = 0, \text{ откуда}$$

$$\cos\omega_+ = \frac{a(2b+c-1)}{2(2b+c)}, \quad (17)$$

при условии

$$\left| \frac{a(2b+c-1)}{2(2b+c)} \right| \leq 1. \quad (18)$$

г) Пусть $\omega_2 = -\omega_1 = \omega_-$, тогда

$$f(\omega_1, \omega_2) = f(\omega_-) = (1-2a\cos\omega_- - 2b\cos 2\omega_-)^2. \quad (19)$$

Найдем нули функции:

$$\frac{df(\omega_-)}{d\omega_-} = 4\sin\omega_-(a+4b\cos\omega_-)(1-c) - 2a\cos\omega_- - 2b\cos 2\omega_-$$

$$\frac{df(\omega_-)}{d\omega_-} = 0 \text{ при}$$

$$1) \sin\omega_- = 0, \text{ откуда } \omega_- = 0, \pi;$$

$$2) a+4b\cos\omega_- = 0, \text{ откуда}$$

$$\cos\omega_- = -\frac{a}{4b}, \quad (20)$$

при условии

$$\left| \frac{a}{4b} \right| \leq 1; \quad (21)$$

$$3) (1-c) - 2a\cos\omega_- - 2b\cos 2\omega_- = 0.$$

Решив это уравнение, получим

$$\cos\omega_{-1,2} = \frac{-a \pm \sqrt{a^2 + 4b(1+2b-c)}}{4|b|}, \quad (22)$$

при выполнении условий

$$\left\{ \begin{aligned} \left| \frac{-a \pm \sqrt{a^2 + 4b(1+2b-c)}}{4|b|} \right| \leq 1. \\ a^2 + 4b(1+2b-c) \geq 0 \end{aligned} \right. \quad (23)$$

Таким образом, анализируя выражения (10-23), можно установить, что рассматриваемая система имеет 13 экстремумов по направлениям осей и диагоналей (рис. 3).

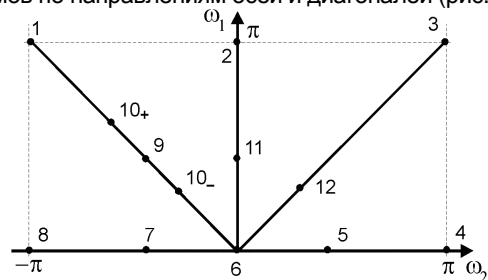


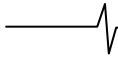
Рис. 3. Экстремумы амплитудно-частотной характеристики двумерного фильтра

Система является монотонной при отсутствии экстремумов у функций (10), (12), (16), (19). С учетом (20-23) этого можно достичь при выполнении следующих условий:

$$\left\{ \begin{aligned} \left| \frac{(a+c)(a+2b-1)}{2b(1-a-b)} \right| > 1; \\ \left| \frac{(a+c)(a+2b-1)}{4b} \right| > 1; \\ \left| \frac{a}{4b} \right| > 1; \\ \left| \frac{-a \pm \sqrt{a^2 + 4b(1+2b-c)}}{4|b|} \right| > 1; \\ a^2 + 4b(1+2b-c) > 0. \end{aligned} \right. \quad (24)$$

Фильтры нижних и верхних частот

Далее определим области существования фильтров нижних и верхних частот с монотонными характеристиками. Для этого достаточно рассмотреть отклонения произвольных точек



от касательных плоскостей, проведенных только в точках $(0,0)$ и (π, π) , причем рассматривать будем достаточно малые окрестности точек по всем направлениям от точки касания. Для этого, как уже отмечалось, необходимо вычислить коэффициенты второй квадратичной формы поверхности.

Найдем производные $\frac{\partial H^2}{\partial \omega_1}$ и $\frac{\partial H^2}{\partial \omega_2}$:

$$\frac{\partial H^2}{\partial \omega_1} = -\frac{1}{f^2(\omega_1, \omega_2)} \frac{\partial f(\omega_1, \omega_2)}{\partial \omega_1}, \quad (25)$$

$$\frac{\partial H^2}{\partial \omega_2} = -\frac{1}{f^2(\omega_1, \omega_2)} \frac{\partial f(\omega_1, \omega_2)}{\partial \omega_2}. \quad (26)$$

В точках $(0,0)$ и (π, π) производные (25-26) равны нулю. Следовательно, в этих точках $\mathbf{r}_{\omega} = (1, 0, 0)$ и $\mathbf{r}_{\omega} = (0, 1, 0)$. Тогда коэффициенты второй квадратичной формы можно записать следующим образом:

$$L(0,0) = \frac{\partial^2 H^2}{\partial \omega_1^2} \Big|_{(0,0)}, \quad M(0,0) = \frac{\partial^2 H^2}{\partial \omega_1 \partial \omega_2} \Big|_{(0,0)}, \quad N(0,0) = \frac{\partial^2 H^2}{\partial \omega_2^2} \Big|_{(0,0)},$$

$$L(\pi, \pi) = \frac{\partial^2 H^2}{\partial \omega_1^2} \Big|_{(\pi, \pi)}, \quad M(\pi, \pi) = \frac{\partial^2 H^2}{\partial \omega_1 \partial \omega_2} \Big|_{(\pi, \pi)},$$

$$N(\pi, \pi) = \frac{\partial^2 H^2}{\partial \omega_2^2} \Big|_{(\pi, \pi)}$$

Получим выражения для вторых производных квадрата модуля ЧХ:

$$\frac{\partial^2 H}{\partial \omega_1^2} = \frac{1}{f^2(\omega_1, \omega_2)} \left(\frac{2}{f(\omega_1, \omega_2)} \left(\frac{\partial f(\omega_1, \omega_2)}{\partial \omega_1} \right)^2 - \frac{\partial^2 f(\omega_1, \omega_2)}{\partial \omega_1^2} \right), \quad (27)$$

$$\frac{\partial^2 H}{\partial \omega_2^2} = \frac{1}{f^2(\omega_1, \omega_2)} \left(\frac{2}{f(\omega_1, \omega_2)} \left(\frac{\partial f(\omega_1, \omega_2)}{\partial \omega_2} \right)^2 - \frac{\partial^2 f(\omega_1, \omega_2)}{\partial \omega_2^2} \right), \quad (28)$$

$$\frac{\partial^2 H}{\partial \omega_1 \partial \omega_2} = \frac{1}{f^2(\omega_1, \omega_2)} \left(\frac{2}{f(\omega_1, \omega_2)} \frac{\partial f(\omega_1, \omega_2)}{\partial \omega_1} \frac{\partial f(\omega_1, \omega_2)}{\partial \omega_2} - \frac{\partial^2 f(\omega_1, \omega_2)}{\partial \omega_1 \partial \omega_2} \right), \quad (29)$$

В точках $(0,0)$ и (π, π) производные (27-29) имеют значения

$$\frac{\partial^2 H}{\partial \omega_1^2} \Big|_{(0,0)} = \frac{\partial^2 H}{\partial \omega_2^2} \Big|_{(0,0)} = \frac{2\{(1-2(a+b)-c) \times (c+a+4b) + (a+c+2b)^2\}}{(1-2(a+b)-c)^4},$$

$$\frac{\partial^2 H}{\partial \omega_1^2} \Big|_{(\pi, \pi)} = \frac{\partial^2 H}{\partial \omega_2^2} \Big|_{(\pi, \pi)} = \frac{2\{(1+2(a-b)-c) \times (c-a+4b) + (c-a+2b)^2\}}{(1+2(a-b)-c)^4},$$

$$\frac{\partial^2 H}{\partial \omega_1 \partial \omega_2} \Big|_{(0,0)} = -\frac{2c}{(1-2(a+b)-c)^3},$$

$$\frac{\partial^2 H}{\partial \omega_1 \partial \omega_2} \Big|_{(\pi, \pi)} = -\frac{2c\{(1+2(a-b)+c)\}}{(1+2(a-b)-c)^4}.$$

Получим выражения для положительности и отрицательности отклонений поверхности квадрата модуля ЧХ

от касательных плоскостей в точках $(0,0)$ и (π, π) . В точке $(0,0)$ условие положительности имеет вид:

$$\begin{cases} (1-2(a+b)-c) \times (c+a+4b) + \\ + (a+c+2b)^2 < 0, \\ [(1-2(a+b)-c) \times (c+a+4b) + \\ + (a+c+2b)^2]^2 - [1-2(a+b)-c]^2 > 0, \end{cases} \quad (30)$$

а условие отрицательности

$$\begin{cases} (1-2(a+b)-c) \times (c+a+4b) + (a+c+2b)^2 > 0, \\ [(1-2(a+b)-c) \times (c+a+4b) + (a+c+2b)^2]^2 - [1-2(a+b)-c]^2 > 0. \end{cases} \quad (31)$$

В точке (π, π) условие положительности имеет вид:

$$\begin{cases} (1+2(a-b)-c) \times (c-a+4b) + (c-a+2b)^2 < 0, \\ [(1+2(a-b)-c) \times (c-a+4b) + (c-a+2b)^2]^2 - [1+2(a-b)+c]^2 > 0, \end{cases} \quad (32)$$

а условие отрицательности

$$\begin{cases} (1+2(a-b)-c) \times (c-a+4b) + (c-a+2b)^2 > 0, \\ [(1+2(a-b)-c) \times (c-a+4b) + (c-a+2b)^2]^2 - [1+2(a-b)+c]^2 > 0. \end{cases} \quad (33)$$

Для того чтобы система была фильтром нижних частот, необходимо чтобы в точке $(0,0)$ отклонение было отрицательно, а в точке (π, π) – положительно. Это достигается при выполнении условий (31) и (32).

Для того чтобы система была фильтром верхних частот, необходимо чтобы в точке $(0,0)$ отклонение было положительно, а в точке (π, π) – отрицательно. Это достигается при выполнении условий (30) и (33).

Решим уравнения (30)-(33) совместно с условиями устойчивости (3) и условиями монотонности (24). Получим области коэффициентов, соответствующие устойчивым фильтрам верхних и нижних частот с монотонными АЧХ.

В случае ФНЧ максимальное подавление входного сигнала будет осуществляться на частоте (π, π) , достигая значения:

$$D_{\max} = 20 \lg \left(\frac{1-2a-c-2b}{1+2a-c+2b} \right). \quad (34)$$

В случае ФВЧ максимальное подавление входного сигнала будет осуществляться на частоте $(0,0)$ и определяться количественно в виде:

$$D_{\max} = 20 \lg \left(\frac{1+2a-c-2b}{1-2a-c+2b} \right). \quad (35)$$

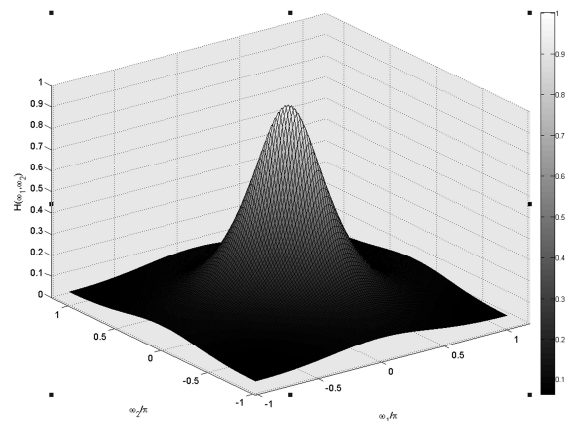


Рис. 4. Типичный вид АЧХ двумерного фильтра нижних частот

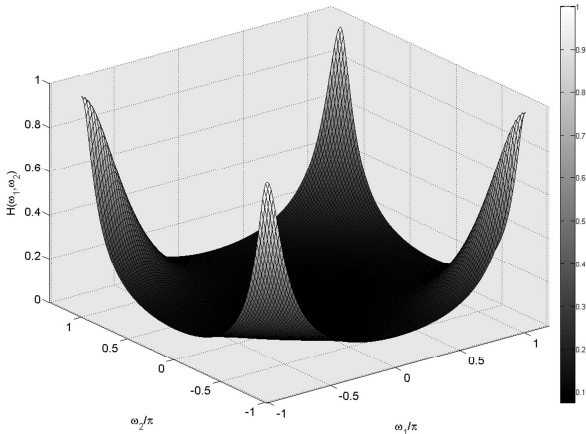


Рис. 5. Типичный вид АЧХ двумерного фильтра верхних частот

Типичный вид АЧХ фильтра нижних частот приведен на рис. 4. Он соответствует следующим параметрам фильтра: $a = 0.49, b = -0.06, c = 0.01, D_{\max} = 53$ дБ.

Типичный вид АЧХ фильтра верхних частот приведен на рис. 5. Он соответствует следующим параметрам фильтра: $a = -0.46, b = -0.04, c = 0.01$. При этом, исходя из (35), $D_{\max} = 50$ дБ.

Селектирующие свойства одномерных цифровых фильтров нижних и верхних частот принято характеризовать частотой среза, соответствующей затуханию АЧХ на уровне 3 дБ. Для двумерного случая введем понятие линии среза, точки которой имеют координаты $(\omega_{c1}, \omega_{c2})$. Найдем ее уравнение для случая низкочастотного фильтра.

Свое максимальное значение АЧХ принимает в точке с координатами $(0,0)$

$$|H|_{\max}^2 = |H(0,0)|^2 = \frac{1}{f(0,0)} = \frac{1}{(1-2a-c-2b)^2}. \quad (36)$$

Проведем нормировку квадрата АЧХ на максимальное значение (36) и найдем уравнение линии, на которой квадрат АЧХ спадает в два раза, что соответствует спаду самой АЧХ на 3 дБ

$$\frac{|H|^2(\omega_{c1}, \omega_{c2})}{|H|_{\max}^2} = \frac{(1-2a-c-2b)^2}{f(\omega_{c1}, \omega_{c2})} = \frac{1}{2}. \quad (37)$$

Используя (6), из (37) получим:

$$(1-a(\cos\omega_{c1} + \cos\omega_{c2}) - c \cos(\omega_{c1} + \omega_{c2}) - b(\cos 2\omega_{c1} + \cos 2\omega_{c2}))^2 + (a(\sin\omega_{c1} + \sin\omega_{c2}) + c \sin(\omega_{c1} + \omega_{c2}) + b(\sin 2\omega_{c1} + \sin 2\omega_{c2}))^2 = 2(1-2a-c-2b)^2. \quad (38)$$

Уравнение (38) достаточно сложно и неудобно для использования, так как необходимо определять количественные значения ω_{c1} и ω_{c2} . Аппроксимируем его более удобной кривой, используя разложения в ряд Тейлора вблизи точки $(0,0)$ функций $\sin \omega$ и $\cos \omega$.

После некоторых математических выкладок получим уравнение линии среза системы для случая фильтра нижних частот в следующем виде:

$$A\omega_{c1}^2 + C\omega_{c2}^2 + 2B\omega_{c1}\omega_{c2} = R^2, \quad (39)$$

где

$$R^2 = \frac{1-2a-c-2b}{a+4b},$$

$$A = C = (1-2a-c-2b)(a+c+4b) + (a+c+2b)^2, \quad (40)$$

$$2B = 2(1-2a-c-2b)c + (a+c+2b)^2.$$

Проведем аналогичные выкладки для случая фильтра верхних частот. Свое максимальное значение АЧХ принимает в точке с координатами (π, π)

$$|H|_{\max}^2 = |H(\pi, \pi)|^2 = \frac{1}{f(\pi, \pi)} = \frac{1}{(1+2a-c-2b)^2}. \quad (41)$$

Проведем нормировку квадрата АЧХ на максимальное значение (41) и найдем уравнение линии, на которой квадрат АЧХ спадает в два раза, что соответствует спаду самой АЧХ на 3 дБ

$$\frac{|H(\omega_{c1}, \omega_{c2})|^2}{|H|_{\max}^2} = \frac{(1+2a-c-2b)^2}{f(\omega_{c1}, \omega_{c2})} = \frac{1}{2}. \quad (42)$$

Используя (6), из (42) получим:

$$(1-a(\cos\omega_{c1} + \cos\omega_{c2}) - c \cos(\omega_{c1} + \omega_{c2}) - b(\cos 2\omega_{c1} + \cos 2\omega_{c2}))^2 + (a(\sin\omega_{c1} + \sin\omega_{c2}) + c \sin(\omega_{c1} + \omega_{c2}) + b(\sin 2\omega_{c1} + \sin 2\omega_{c2}))^2 = 2(1+2a-c-2b)^2. \quad (43)$$

Используя разложения в ряд Тейлора вблизи точки (π, π) функций $\sin \omega$ и $\cos \omega$, получим уравнение линии среза двумерного рекурсивного цифрового фильтра верхних частот в следующем виде, аналогичном (43), но с учетом, что

$$R^2 = \frac{1+2a-c-2b}{-a+4b},$$

$$A = C = (1+2a-c-2b)(-a+c+4b) + (-a+c-2b)^2, \quad (44)$$

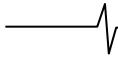
$$2B = 2(1+2a-c-2b)c + 2(-a+c+2b)^2.$$

Выражение (39) представляет собой уравнение эллипса. Анализируя его и (32), можно установить, что линии среза двумерных рекурсивных цифровых фильтров нижних и верхних частот – окружности. Равенства (40) и (44) позволяют количественно оценить значения ω_c по обоим частотным координатам.

Диагональный режекторный фильтр

В рассматриваемой системе АЧХ имеет экстремальные точки только на осях частот ω_1 и ω_2 и на диагоналях $\omega_1 = \omega_2$ и $\omega_1 = -\omega_2$. Под диагональным двумерным режекторным фильтром будем понимать фильтр, у которого области пропускания существуют только на диагоналях $\omega_1 = \omega_2$ и $\omega_1 = -\omega_2$ частотной плоскости, причем обязательно разделенные областями непропускания на интервале однозначности. Определим математические ограничения на коэффициенты исследуемой системы для получения областей существования диагонального режекторного фильтра.

Необходимо получить области пропускания в окрестностях точек (π, π) , $(0,0)$ и $(-\pi, \pi)$, разделенные между собой областями непропускания. В силу симметричности системы $|H(\pi, \pi)| = |H(-\pi, \pi)|$, будем рассматривать только точку с координатами (π, π) . В окрестностях точек (π, π) и $(0,0)$ необходимо, чтобы отклонение любой точки АЧХ от касательных плоскостей, проведенных к ее поверх-



ности в этих точках, было отрицательно, что достигается выполнением условий (26) и (28). При этом значения АЧХ в этих точках должны попадать в полосу пропускания, т.е. должны превышать значение $\frac{1}{\sqrt{2}}$ от максимального. Так как рассматривается симметричная система, то возможны два различных случая, меняющие все математические выкладки. Первый – наличие максимума АЧХ в точке $(0,0)$, и второй – наличие максимума АЧХ в точке (π, π) . Рассмотрим области подавления сигнала на диагоналях $\omega_1 = \omega_2$ и $\omega_1 = -\omega_2$.

Максимальное подавление сигнала на диагонали $\omega_1 = \omega_2$ будет происходить на частоте

$$\omega = \pm \arccos \frac{a(c + 2b - 1)}{2(c + 2b)},$$

и равно

$$D_{\max 1} = 20 \lg \sqrt{[1 - 2a\omega - (c + 2b)(2\omega^2 - 1)]^2 + 4(a + c\omega + 2b\omega)^2(1 - \omega^2)}.$$

Максимальное подавление сигнала на диагонали $\omega_1 = -\omega_2$ будет происходить на частоте

$$\omega = \pm \arccos \left(-\frac{a}{4b} \right),$$

и равно $D_{\max 2} = 20 \lg \{1 - 2a\omega - c - 2b - 2b(2\omega^2 - 1)\}$.

Пример типичной АЧХ диагонального режекторного фильтра приведен на рис. 6. Она соответствует следующим параметрам: $a = 0.01$, $b = 0.23$, $c = 0.35$.

Осевой режекторный фильтр

Введем понятие осевого режекторного фильтра. Под осевым режекторным фильтром будем понимать фильтр, у которого области пропускания существуют только на осях ω_1 и ω_2 частотной плоскости, причем обязательно разделенные областями непропускания. Области пропускания в окрестностях точек $(\pi, 0)$, $(0, 0)$, $(0, -\pi)$, $(-\pi, 0)$ и $(0, \pi)$ разделены между собой областями непропускания. В окрестностях точек $(\pi, 0)$ и $(0, 0)$ необходимо, чтобы отклонение h любой точки АЧХ от касательных плоскостей, проведенных к ее поверхности в этих точках, было отрицательно. При этом значения АЧХ характеристики в этих точках должны попадать в полосу пропускания, т.е. превышать значение $\frac{1}{\sqrt{2}}$ от максимального. Так как рассматривается симметричная система, то $|H(\pi, 0)| = |H(0, \pi)| = |H(-\pi, 0)| = |H(0, -\pi)|$, поэтому возможны два различных случая.

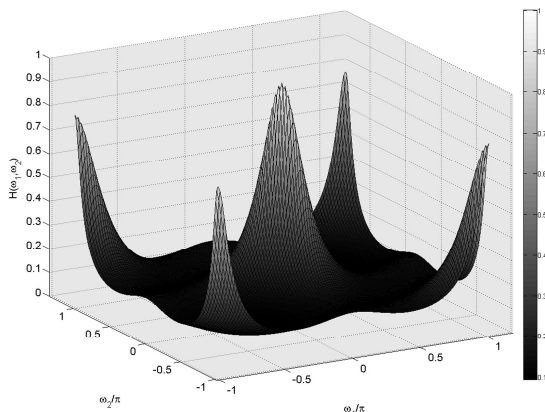


Рис. 6. Типичный вид АЧХ двумерного диагонального режекторного фильтра

Первый – наличие максимума АЧХ в точке $(0,0)$ и второй – наличие максимума АЧХ в точке $(\pi, 0)$. Получены математические ограничения на коэффициенты исследуемой системы для получения областей существования осевого режекторного фильтра.

Максимальное подавление сигнала по оси ω_1 совпадает с максимальным подавлением по оси ω_2 , реализуется в точке

$$\omega = \pm \arccos \frac{(a + c)(a + 2b - 1)}{4b(1 - a - b)}$$

и равно

$$D_{\max} = 20 \lg \sqrt{[1 - a - a\omega - c\omega - 2b^2\omega^2]^2 + (a + c + 2b\omega)^2(1 - \omega^2)}.$$

Пример типичной АЧХ осевого режекторного фильтра приведен на рис. 7. Она соответствует следующим параметрам: $a = 0.245$, $b = 0.33$, $c = 0.33$, $D_{\max} = 18.5$ дБ.

Как и в случае режекторных фильтров, понятие полосового фильтра однозначно для одномерной системы. Для двумерного случая существует некоторая свобода в его определении. Рассматривать будем только некоторые направления АЧХ, а именно: обе оси частот ω_1 и ω_2 и обе диагонали $\omega_1 = \omega_2$ и $\omega_1 = -\omega_2$.

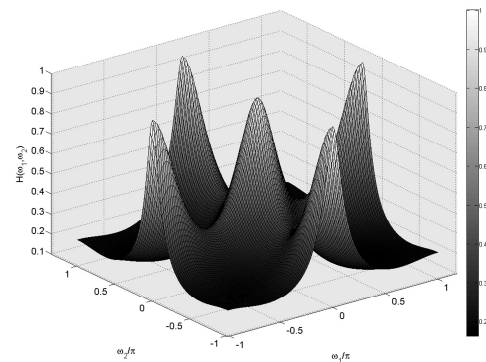


Рис. 7. Типичный вид АЧХ двумерного осевого режекторного фильтра

Диагональный полосовой фильтр

Введем понятие двумерного диагонального полосового фильтра. Под диагональным полосовым фильтром будем понимать фильтр, у которого области пропускания существуют только на диагоналях $\omega_1 = \omega_2$ и $\omega_1 = -\omega_2$ частотной плоскости, причем на интервале однозначности существует только одна область пропускания на каждой из диагоналей. Как и в случае диагональных режекторных фильтров, имеют место ограничения на коэффициенты исследуемой системы для получения областей существования диагонального полосового фильтра.

Необходимо получить области пропускания в окрестностях точек 5, 12, 11, 9 и 7 (рис. 3). В силу симметричности системы достаточно рассмотреть только точки 5, 12 и 9. В окрестностях точек (π, π) и $(0,0)$ необходимо, чтобы отклонение любой точки АЧХ от касательных плоскостей, проведенных к ее поверхности в этих точках, было положительно, что достигается выполнением условий (25) и (27). При этом значения АЧХ в этих точках не должны попадать в полосу пропускания, т.е. не должны превышать значение $\frac{1}{\sqrt{2}}$ от максимального. Так как рассматривается симметричная система, то возможны два различных случая. Первый – наличие максимума АЧХ в точке 12, и второй – наличие максимума АЧХ в точке 9.

Получена область существования диагональных полосовых фильтров. Ее форма в пространстве коэффициентов (рис. 1) достаточно сложна и может быть представлена в виде соответствующих сечений плоскостями $c = const$. Сечения позволяют оценить форму области и ее местоположение в пространстве коэффициентов. Определим резонансные частоты диагонального полосового фильтра. Первоначально рассмотрим диагональ $\omega_1 = \omega_2$. В этом случае

$$\omega_{p_1} = \pm \arccos \frac{a(c + 2b - 1)}{2(c + 2b)}.$$

На диагонали $\omega_1 = -\omega_2$ резонансная частота определяется следующим образом

$$\omega_{p_2} = \pm \arccos \left(-\frac{a}{4b} \right).$$

Пример типичной АЧХ диагонального полосового фильтра приведен на рис. 8. Она соответствует следующим параметрам: $a = 0.04$, $b = -0.275$, $c = -0.05$.

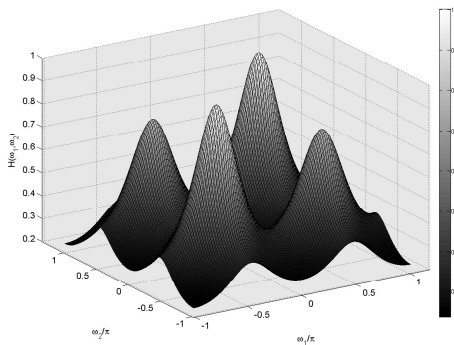


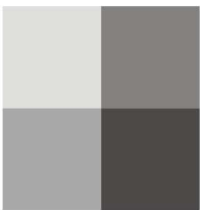
Рис. 8. Типичный вид АЧХ двумерного диагонального полосового фильтра

Заключение

В пространстве параметров двумерной рекурсивной цифровой системы второго порядка с симметричными коэффициентами получены аналитические ограничения на параметры, соответствующие фильтрам нижних и верхних частот. Найдены также аналитические ограничения на коэффициенты в указанных областях для отсутствия у этих фильтров пульсаций амплитудно-частотной характеристики. Предложена аппроксимация линии среза. В связи с тем, что АЧХ двумерной системы является поверхностью, понятие режекторных и полосовых фильтров в двумерном случае намного шире, чем в одномерном. Предложена классификация двумерных полосовых и режекторных фильтров. Введены понятия диагональных и осевых фильтров. Аналитически получены ограничения на параметры системы для существования диагональных полосовых и режекторных, осевых режекторных фильтров, определены резонансные частоты указанных фильтров.

Литература

1. Даджион Д., Мерсеро Р. Цифровая обработка многомерных сигналов. – М.: Мир, 1988.
2. Погорелов А.В. Дифференциальная геометрия. – М.: Наука, 1969.
3. Брюханов Ю.А., Приоров А.Л., Мясников Е.А., Калинин С.А. Частотные свойства двумерных рекурсивных цифровых систем первого порядка // Изв. высш. учеб. заведений. Радиоэлектроника. 1995. № 4. С. 26-30.



**MATROX
IMAGING**

Компания Matrox Imaging более 30 лет успешно поставляет свои ПК-ориентированные программные и аппаратные технологии разработчикам *систем машинного зрения, анализа объектов и изображений, цифрового видео наблюдения, систем визуализации в медицине и биологии, систем управления технологическими процессами.*

Линейка продуктов включает в себя: *фрейм-грабберы, процессоры обработки изображения, видео смарт-камеры, готовые системы машинного зрения, программное обеспечение для анализа изображения (библиотеки MIL). Штаб квартира в Монреале (Канада), офисы и представительства в 20 странах мира.*

Авторизованный представитель компании **Matrox Imaging** в России –
System Video Graphics Alliance (SVGA).

АЛГОРИТМ СИНТЕЗИРОВАНИЯ АПЕРТУРЫ ДЛЯ ПОЛУЧЕНИЯ РАДИОЛОКАЦИОННОГО ИЗОБРАЖЕНИЯ ПРИ МАЛОВЫСОТНОМ ПОЛЕТЕ

Заикин С.А., Мойбенко В.И., Парфиров В.А.

Введение

Задача формирования радиолокационного изображения (РЛИ) земной поверхности для обеспечения посадки при плохой видимости остается актуальной, хотя ей посвящен ряд работ [1,2]. Поскольку разрешение по азимуту в реальном луче в режиме маловысотного полета недостаточно, предложено использование синтезированной апертуры (СА) для радиовидения в передней зоне.

Особенностью режима СА в передней зоне является большое время синтеза, что в маловысотном полете (МВ) ведет к повышению влияния на качество РЛИ алгоритмов учета маневра. Если же синтезирование ведется при наблюдении в передней зоне сбоку, то более жесткими становятся требования и к алгоритмам учета фокусировки из-за значительной скорости и разницы скоростей линии визирования при наблюдении разных участков поверхности на небольших дальностях. Кроме того, при работе на небольших дальностях в РЛС используется значительная частота повторения импульсов, а сигнал от расположенной впереди поверхности, по которой строится РЛИ, отличается узкой полосой доплеровских частот. Это приводит к тому, что при фильтрации с помощью БПФ выполняются вычисления, большая часть результатов которых не требуется (тем более если точек фокусировки не одна, а несколько). В [3] предложены алгоритмы децимации для работы в таких условиях. Однако они ориентированы на режим доплеровского обужения луча, для которого учет маневра носителя и способ фокусировки не так существенны, и поэтому не рассматриваются. Эти вопросы частично отражены в последующих работах авторов, в частности в [4].

В настоящей работе предлагается двухступенчатый алгоритм с выполнением децимации с помощью фильтров первой ступени, отличающихся тем, что их главная задача - это учет маневра носителя РЛС. Последующий алгоритм ДПФ выполняется в узкой по

Рассматривается алгоритм синтеза апертуры, предназначенный для получения радиолокационного изображения поверхности в маловысотном полете на небольших дальностях, в условиях высокой неустойчивости траектории и высокой скорости линии визирования.

сравнению с частотой повторения полосе частот, чем достигается уменьшение объема вычислений. Этот алгоритм отличается возможностью учитывать фокусировку по большому числу точек на РЛИ практически без повышения объема вычислений, чем и достигается повышение качества фокусировки.

Алгоритм СА для получения РЛИ в режиме маловысотного полета

Структура алгоритма СА представлена на рис. 1. Навигационные данные, необходимые для формирования опорных функций, включают такие характеристики, как мгновенная скорость носителя, угловая скорость разворота, начальные значения угла визирования на поверхность, расстояния и некоторые другие.

Получение выборок отсчетов АЦП, поступающих на вход первой ступени, иллюстрируется рис.2. В каждом временном интервале наблюдения излучается N_{imp} зондирующих импульсов (ЗИ). В зоне приема сигнала между двумя соседними ЗИ регистрируются отсчеты АЦП $A(i)$, $i=1 \dots I_{str}$, количество I_{str} которых соответствует количеству анализируемых стробов дальности. Каждому i -му отсчету АЦП циклограммы соответствует своя выборка из N_{imp} элементов, по числу ЗИ в интервале наблюдения, формируемая из отсчетов АЦП с одинаковыми номерами в каждой зоне приема сигнала. На вход фильтра первой ступени поступают реальная и мнимая составляющие I_{str} таких выборок.

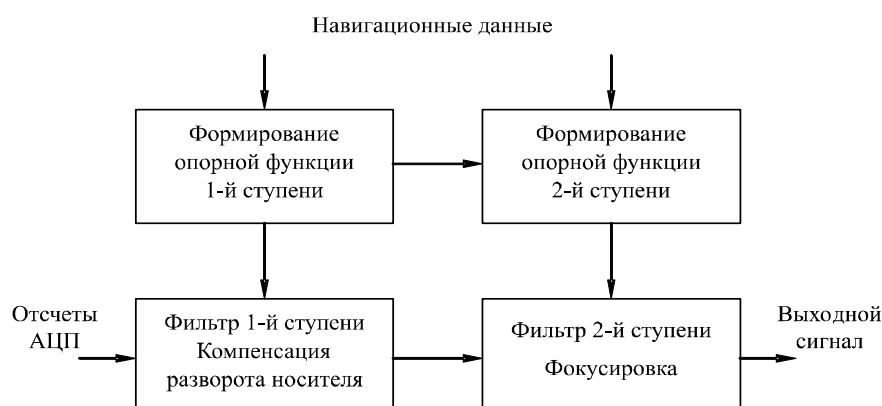


Рис. 1. Структура алгоритма СА

Фильтрация сигнала на первой ступени представляет собой комплексное умножение и суммирование принятых выборок входного накопленного массива на опорную функцию в соответствии с формулами:

$$\text{Re}l_{in} = \sum_{j=(n-1) \cdot K1+1}^{n \cdot K1} (\text{Re}i_{ij} \cdot \cos \varphi_j + \text{Im}i_{ij} \cdot \sin \varphi_j),$$

$$\text{Im}l_{in} = \sum_{j=(n-1) \cdot K1+1}^{n \cdot K1} (\text{Im}i_{ij} \cdot \cos \varphi_j - \text{Re}i_{ij} \cdot \sin \varphi_j),$$

$$i = \overline{1, I_{str}}; n = \overline{1, N} = \frac{N_{imp}}{K1},$$

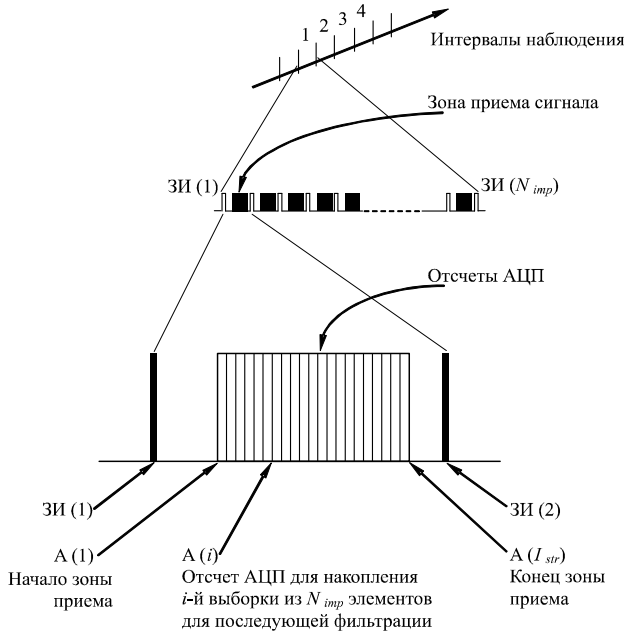


Рис. 2. Циклограмма получения отсчетов АЦП

На рис. 2: I_{str} - количество отсчетов АЦП (стробов дальности); N_{imp} - размер пачки – количество импульсов в выборке по i -му отсчету АЦП, равное количеству ЗИ за время наблюдения; $K1$ - количество импульсов, суммируемых на первой ступени по каждому отсчету (стробу дальности); N - количество импульсов на выходе первой ступени по каждому отсчету; $\text{Re}i_{ij}, \text{Im}i_{ij}, i = \overline{1, I_{str}}; j = \overline{1, N_{imp}}$ - соответственно реальная и мнимая составляющие сигнала на входе первой ступени по i -му отсчету АЦП для j -го импульса в пачке; $\text{Re}l_{in}, \text{Im}l_{in}, i = \overline{1, I_{str}}; n = \overline{1, N}$ - соответственно реальная и мнимая составляющие сигнала по i -му отсчету АЦП для n -го импульса на выходе первой ступени; $\varphi_j, j = \overline{1, N_{imp}}$ - фаза опорной функции для j -го импульса в пачке, град.

Фаза опорной функции для j -го импульса в пачке вычисляется в соответствии с уравнением

$$\varphi_j = \varphi_{j-1} + 2\pi \cdot Fd_j \cdot dt,$$

где Fd_j - частота Доплера для j -го импульса в пачке, Гц; dt - период повторения ЗИ, с.

Частота Доплера для j -го импульса в пачке определяется следующим образом:

$$Fd_j = \frac{2 \cdot V_j}{\lambda} \cdot \cos \beta_j, \quad (1)$$

где V_j - скорость носителя, соответствующая j -му импульсу, м/с; λ - длина волны несущей частоты ЗИ, м; β_j - азимут на точку фокусировки, изменяющийся вследствие разворота носителя, для j -го импульса в пачке, град.

Учет разворота носителя в алгоритме СА иллюстрируется на рис. 3. Как видно из рис. 3, справедливы следующие соотношения:

$$\beta_j = \beta_{j-1} + \Delta\beta_j,$$

$$\Delta\beta_j = \omega_{j-1} \cdot dt,$$

$$\beta_j = \beta_0 + \Delta\beta_{\Sigma j},$$

$$\Delta\beta_{\Sigma j} = \sum_{k=1}^{j-1} \omega_k \cdot dt,$$

где $\Delta\beta_j$ - приращение угла поворота на цель ДНА для j -го импульса в пачке вследствие разворота носителя в момент времени t_j , град; $\Delta\beta_{\Sigma j}$ - суммарное приращение положения на точку фокусировки ДНА для j -го импульса в пачке вследствие разворота носителя в момент времени t_j , град; ω_{j-1} - угловая скорость разворота носителя для $(j-1)$ -го импульса в момент времени t_{j-1} , град/с; β_0 - заданное начальное значение угла поворота на точку фокусировки ДНА в момент времени t_1 , град.

Так как тригонометрические функции при их программировании требуют большого объема вычислений, то используется рекуррентный способ вычисления значений $\cos \beta_j$ в формуле (1) с разложением тригонометрических функций в ряд Тейлора, справедливым по числу членов для малых углов w , в виде

$$\sin w = w - \frac{w^3}{6}, \quad (2)$$

$$\cos w = 1 - \frac{w^2}{2}. \quad (3)$$

Отсюда, с использованием тригонометрических формул сложения и с учетом малости $\Delta\beta_j$, следует:

$$\begin{aligned} \cos \beta_j &= \cos(\beta_{j-1} + \Delta\beta_j) = \\ &= \cos \beta_{j-1} \cdot \left(1 - \frac{\Delta\beta_j^2}{2}\right) - \sin \beta_{j-1} \cdot \left(\Delta\beta_j - \frac{\Delta\beta_j^3}{6}\right), \end{aligned}$$

$$\begin{aligned} \sin \beta_j &= \sin(\beta_{j-1} + \Delta\beta_j) = \\ &= \sin \beta_{j-1} \cdot \left(1 - \frac{\Delta\beta_j^2}{2}\right) + \cos \beta_{j-1} \cdot \left(\Delta\beta_j - \frac{\Delta\beta_j^3}{6}\right). \end{aligned}$$

Начальными условиями в момент времени t_1 для 1-го импульса пачки ($j=1$) на входе первой ступени являются:

$$\varphi_j = \varphi_1 = 0,$$

$$\beta_j = \beta_1 = \beta_0,$$

$$\cos \beta_j = \cos \beta_1 = \cos \beta_0,$$

$$\sin \beta_j = \sin \beta_1 = \sin \beta_0.$$

При выполнении вычислений на второй ступени алгоритма СА используются угол отклонения от первоначального курса вследствие разворота носителя для n -го импульса на входе второй ступени B_n , град и проекция скорости носителя на первоначальное направление W_n , м/с, определяемые по формулам

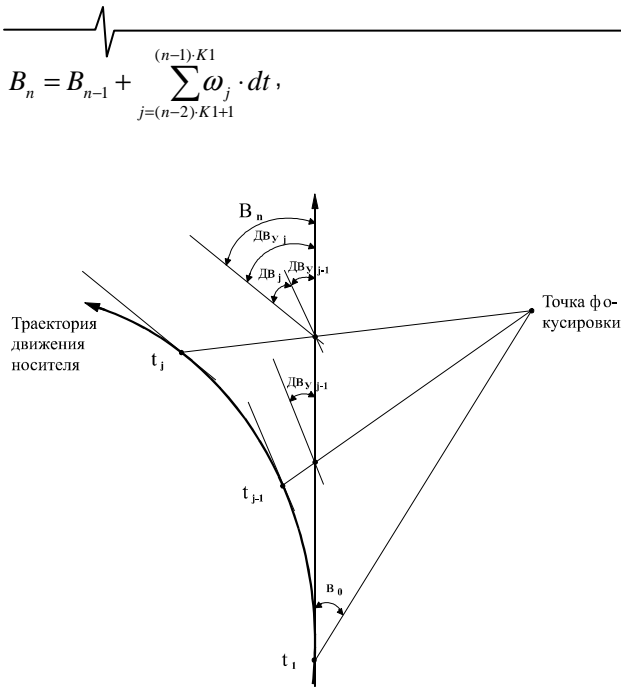


Рис. 3. Учет разворота носителя в первой ступени

$$W_n = V_{nK1} \cdot \cos B_n,$$

$$n = \overline{1, N}$$

с начальным условием $B_n = 0$ при $n=1$.

Фильтрация сигнала на второй ступени реализуется комплексным умножением переданных выборок выходного накопленного массива первой ступени на опорную функцию второй ступени в соответствии с формулами:

$$\text{Re}2_{im} = \sum_{n=(m-1)K2+1}^{mK2} (\text{Re}1_{in} \cdot \cos \psi_n + \text{Im}1_{in} \cdot \sin \psi_n),$$

$$\text{Im}2_{im} = \sum_{n=(m-1)K2+1}^{mK2} (\text{Im}1_{in} \cdot \cos \psi_n - \text{Re}1_{in} \cdot \sin \psi_n),$$

$$i = \overline{1, I_{str}}; m = \overline{1, M} = \frac{N}{K2} = \frac{N_{imp}}{K1 \cdot K2},$$

где I_{str} - количество отсчетов АЦП; N_{imp} - размер пачки первой ступени – количество импульсов в выборке по i -му отсчету, равное количеству ЗИ за время наблюдения; $K1$ - количество импульсов, суммируемых на первой ступени по каждому отсчету АЦП; $K2$ - количество импульсов, суммируемых на второй ступени по каждому отсчету АЦП; N - количество импульсов на выходе первой ступени по каждому отсчету АЦП; M - количество импульсов на выходе второй ступени по каждому отсчету АЦП; $\text{Re}2_{im}, \text{Im}2_{im}, i = \overline{1, I_{str}}; m = \overline{1, M}$ - соответственно реальная и мнимая составляющие сигнала на выходе второй ступени по i -му отсчету АЦП для m -го импульса; $\text{Re}1_{in}, \text{Im}1_{in}, i = \overline{1, I_{str}}; n = \overline{1, N}$ - соответственно реальная и мнимая составляющие сигнала по i -му отсчету АЦП для n -го импульса на выходе первой ступени; $\psi_n, n = \overline{1, N}$ - фаза опорной функции для n -го импульса, град.

Фаза опорной функции второй ступени включает составляющую, рассчитанную с учетом фокусировки, основанной на вычислении текущего значения скорости Доплера по углу между скоростью носителя и по-

ложением биссектрисы ДНА. Учет фокусировки в алгоритме СА иллюстрируется рис. 4.

Фаза опорной функции второй ступени для n -го импульса вычисляется в соответствии с уравнением

$$\psi_n = \psi_{n-1} + 2\pi \cdot FD_n \cdot dt,$$

где FD_n - частота Доплера для n -го импульса, Гц; dt - период повторения ЗИ, с.

Частота Доплера для n -го импульса определяется следующим уравнением:

$$FD_n = \frac{2 \cdot VD_n}{\lambda},$$

где VD_n - скорость Доплера для n -го импульса, м/с – проекция скорости носителя на ось, направленную на цель.

В соответствии с определением для скорости Доплера

$$\begin{aligned} VD_n &= V_{nK1} \cdot \cos(\gamma_n + B_n) = \\ &= V_{nK1} \cdot (\cos \gamma_n \cdot \cos B_n - \sin \gamma_n \cdot \sin B_n) \end{aligned} \quad (4)$$

$$\text{и } \gamma_n = \gamma_{n-1} + \Delta\gamma_n,$$

где V_{nK1} - скорость носителя, соответствующая $(n \cdot K1)$ -му импульсу первой ступени, м/с; B_n - угол отклонения от первоначального курса вследствие разворота носителя для n -го импульса на входе второй ступени, град; γ_n - угол направления на цель при движении по прямой первоначального курса носителя для n -го импульса на входе второй ступени, град.; $\Delta\gamma_n$ - изменение угла направления на цель вследствие движения по прямой для n -го импульса на входе второй ступени, град.

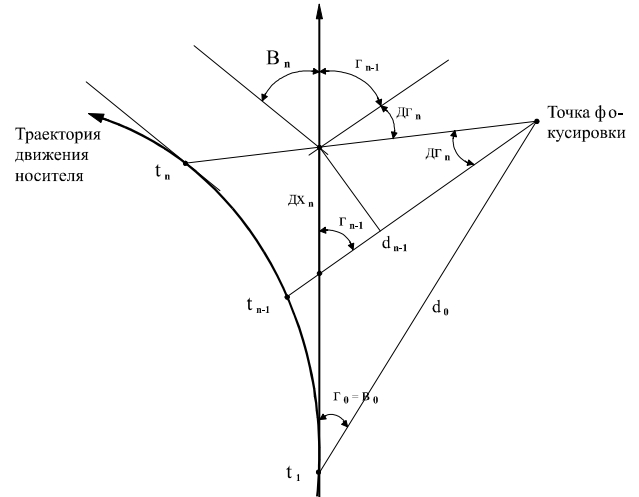


Рис. 4. Учет фокусировки во второй ступени

Как видно из рис. 4, справедливы следующие соотношения:

$$\Delta\gamma_n = \arctg \frac{\Delta x_n \cdot \sin \gamma_{n-1}}{d_{n-1} - \Delta x_n \cdot \cos \gamma_{n-1}}, \quad (5)$$

$$\Delta x_n = W_n \cdot K1 \cdot dt = V_{nK1} \cdot \cos B_n \cdot K1 \cdot dt,$$

$$d_n = \frac{d_{n-1} - \Delta x_n \cdot \cos \gamma_{n-1}}{\cos \Delta\gamma_n}, \quad (6)$$

где Δx_n - перемещение носителя за время между $K1$ отсчетами по прямой начального курса для n -го вы-

ходного импульса первой ступени, м; d_n - текущая дальность до цели при движении по прямой для n -го импульса, м; W_n - проекция скорости носителя на первоначальное направление, м/с.

С использованием формул (2) и (3), а также формулы

$$\text{arctg } w = w - \frac{w^3}{3} \quad \text{- разложения тригонометрических}$$

функций в ряд Тейлора для малых углов w применительно к углам B_n и $\Delta\gamma_n$, соотношения (4), (5), (6) принимают следующий вид:

$$VD_n = V_{n \cdot K1} \cdot \left(\cos \gamma_n \cdot \left(1 - \frac{B_n^2}{2} \right) - \sin \gamma_n \cdot \left(B_n - \frac{B_n^3}{6} \right) \right),$$

$$\Delta\gamma_n = \frac{\Delta x_n \cdot \sin \gamma_{n-1}}{d_{n-1} - \Delta x_n \cdot \cos \gamma_{n-1}} - \frac{\left(\frac{\Delta x_n \cdot \sin \gamma_{n-1}}{d_{n-1} - \Delta x_n \cdot \cos \gamma_{n-1}} \right)^3}{3},$$

$$\Delta x_n = V_{n \cdot K1} \cdot K1 \cdot dt \cdot \left(1 - \frac{B_n^2}{2} \right),$$

$$d_n = \frac{d_{n-1} - \Delta x_n \cdot \cos \gamma_{n-1}}{1 - \frac{\Delta\gamma_n^2}{2}}.$$

Рекуррентное вычисление значений величин $\cos \gamma_n$ и $\sin \gamma_n$ производится аналогично ранее приведенным соотношениям и имеет вид

$$\cos \gamma_n = \cos(\gamma_{n-1} + \Delta\gamma_n) = \cos \gamma_{n-1} \cdot \left(1 - \frac{\Delta\gamma_n^2}{2} \right) -$$

$$- \sin \gamma_{n-1} \cdot \left(\Delta\gamma_n - \frac{\Delta\gamma_n^3}{6} \right);$$

$$\sin \gamma_n = \sin(\gamma_{n-1} + \Delta\gamma_n) =$$

$$\sin \gamma_{n-1} \cdot \left(1 - \frac{\Delta\gamma_n^2}{2} \right) + \cos \gamma_{n-1} \cdot \left(\Delta\gamma_n - \frac{\Delta\gamma_n^3}{6} \right).$$

Начальными условиями в момент времени t_1 для 1-го импульса второй ступени ($n=1$) являются:

$$\psi_n = \psi_1 = 0, \quad B_n = B_1 = 0, \quad d_n = d_1 = d_0,$$

$$\gamma_n = \gamma_1 = \gamma_0 = \beta_0, \quad \cos \gamma_n = \cos \gamma_1 = \cos \gamma_0 = \cos \beta_0,$$

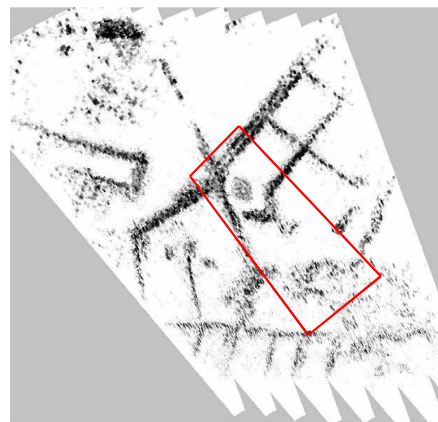
$$\sin \gamma_n = \sin \gamma_1 = \sin \gamma_0 = \sin \beta_0.$$

Здесь d_0 - начальное значение дальности до точки фокусировки.

Алгоритмы второй ступени, реализованные в виде 500..700 фильтров в полосе 50..70 Гц по сравнению с БПФ в полосе частот повторения 7..10 кГц дают выигрыш по числу фильтров на два порядка, к тому же реализация БПФ на 61440 точек затруднена. Это определяет выигрыш в объеме вычислений, но его более точная оценка затруднена тем, что выбор коэффициента децимации требует дополнительных исследований. С учетом разбиения РЛИ на значительное число точек фокусировки этот выигрыш значительно возрастает, но выбор числа точек фокусировки также требует дополнительных исследований (см. рис.5).

Проверка на моделях и результаты летных экспериментов подтверждают работоспособность приведенного алгоритма СА. Для примера на рис.5 показаны радиолокационное изображение и фотография соответствующего участка местности. РЛИ получено на дальности 0.3..3 км в зоне впереди сбоку, с помощью установленной на

борту вертолета дециметровой РЛС, что определяет большое время синтезирования, а из-за малых дальностей и высокую скорость и большую разницу в скорости линии визирования в пределах РЛИ. Тем самым РЛИ иллюстрирует работу алгоритма, когда учет фокусировки и маневра особенно важен.



Заключение

Для режима синтезирования апертуры в условиях маловысотного полета целесообразно использование рассмотренного алгоритма на основе многоскоростной фильтрации, в котором децимация выполняется с учетом нестабильности полета летательного аппарата, а дальнейшая фильтрация выполняется с учетом фокусировки в условиях большой скорости линии визирования на много точек поверхности в пределах РЛИ. Предлагаемый алгоритм позволяет также сократить объем вычислений тригонометрических функций при вычислении ДГФ.

Литература

1. Орлов М.С. Авиационная радиолокационная станция с синтезированной апертурой антенны и передним обзором земной поверхности // Радиотехника, 2003, № 1.
2. Кондратенков Г.С., Фролов А.Ю. Радиовидение в передней зоне обзора бортовой радиолокационной станции с синтезированной апертурой антенны // Радиотехника. 2004, № 1.
3. Витязев В.В., Колодько Г.Н., Витязев С.В. Способы и алгоритмы формирования радиолокационного изображения в режиме доплеровского облучения луча // Цифровая обработка сигналов, 2006, № 3.
4. Витязев В.В., Колодько Г.Н., Воронков Д.В. Формирование радиолокационного изображения в режиме фокусируемого синтезирования апертуры ДНА // Цифровая обработка сигналов. 2006. № 4. -С. 34-40.

КОМПЕНСАЦИЯ МЕЖСИМВОЛЬНЫХ ПОМЕХ ПОСРЕДСТВОМ ОБРАТНОЙ СВЯЗИ ПО РЕШЕНИЮ ДЛЯ СИГНАЛОВ С РЕШЕТЧАТОЙ КОДИРОВАННОЙ МОДУЛЯЦИЕЙ

Григоренко Д.В., Ручкин В.Н.

Введение

В рекомендации V.32bis Международного телекоммуникационного союза [1] для повышения помехоустойчивости модема используется решетчатое кодирование (РК) с квадратурной амплитудной модуляцией (КАМ). РК с КАМ уменьшает вероятность появления ошибок при высоких скоростях передачи данных по телефонному каналу [2,3]. Такой же вид РК и созвездия используются в факс-протоколе модуляции V.17 и в протоколе связи по четырехпроводным выделенным каналам связи – протокол V.33.

Обычно телефонный канал имеет равномерную амплитудную частотную характеристику с плавным спадом на краях [3]. Для таких каналов применяют линейный адаптивный эквалайзер, подавляющий межсимвольную интерференцию сигнала с РК. Однако, если в полосе пропускания канала имеется сильное затухание или существенные узкополосные помехи, линейный эквалайзер не эффективен из-за повышения уровня шумов на корректируемых частотах. В этом случае для сигналов без РК можно использовать эквалайзер с ОСР [5].

Декодирование решетчатых кодов осуществляется с помощью алгоритма Витерби [2]. Простое каскадное соединение эквалайзера с ОСР и алгоритма Витерби невозможно использовать из-за задержки при декодировании [1,3].

Целью данной работы является разработка алгоритма компенсации межсимвольных помех с помощью ОСР для РК сигналов, обеспечивающего высокие скорости передачи данных даже при сильных амплитудных искажениях.

Декодирование с компенсацией межсимвольных помех посредством ОСР

Проведем анализ алгоритма декодирования сигналов с РК модуляцией без ОСР, состоящий из двух основных этапов [1,6]:

1. Поиск точки y_k в каждом подмножестве, которая является ближайшей к принятой точке x_k . Данный поиск проводится среди нескольких точек в эталонном созвездии, ближайших к принятой точке. Этим точкам y_k соответствуют: переход (ветвь) S_k в S_{k-1} в решетчатой диаграмме, биты данных и проверочный бит, определяемый исходным состоянием S_{k-1} . В результате отбрасываются параллельные ветви между исходными S_{k-1} и конечными состояниями S_k .

2. Поиск кратчайшего пути в решетчатой диаграмме с помощью алгоритма Витерби. Длина каждой из ветвей, из которых состоят пути, равна квадрату расстояния между найденной на предыдущем этапе точкой y_k и принятой точкой x_k . Суть алгоритма заключается в том, что среди

Рассмотрен вопрос о совместной реализации решетчатого кодирования и алгоритма компенсации межсимвольных помех с помощью обратной связи по решению (ОСР). На примере модема V.32bis показана эффективность такой реализации.

всех путей, входящих в одно состояние, выбирается кратчайший, остальные пути отбрасываются. В результате такого отбора «выживают» только кратчайшие пути, заканчивающиеся в возможных состояниях S_k к определенному моменту времени k .

Рассмотрим возможность использования ОСР в алгоритме декодирования сигналов с РК модуляцией. В состоянии S_{k-1} входит только один выживший путь, состоящий из $k-1$ ветвей. Поскольку ветвям соответствуют определенные точки y_m , для каждого состояния S_{k-1} можно применить компенсацию с ОСР:

$$x'_k = x_k + \sum_{i=0}^{N-1} b_i y_{k-1-i}, \quad (1)$$

где $\{b_i\}$ - коэффициенты фильтра обратной связи по решению.

Скорректированная точка x'_i может оказаться ближе к другим точкам в эталонном созвездии y_i , чем принятая точка x_i ; точки y_i могут быть различными для разных состояний S_m . Поэтому коррекцию (1) и поиск наилучших точек y_k на первом этапе необходимо проводить для каждого S_{k-1} . На первом этапе декодирования с ОСР квадрат расстояний между точками рассчитывается для каждого состояния S_{k-1} . Таким образом, компенсацию межсимвольных помех посредством ОСР можно использовать для каждого из отдельно выживших путей.

Для стандарта V.32bis определены восемь состояний (подмножества A-G). Следовательно, ОСР потребует 8 операций коррекции (1) и нахождения ближайших точек y_k , что допустимо при реализации на современных сигнальных процессорах.

Фазовая автоподстройка частоты с петлей, управляемой решениями

Настройка эквалайзера с использованием декодированных с задержкой точек (комплексных отсчетов) не вызывает серьезных затруднений. Однако такая задержка недопустима в петле обратной связи фазовой автоподстройки частоты (ФАПЧ) [1,7], отслеживающей сдвиг частоты и фазовое дрожание [1]. ФАПЧ с задержкой является устойчивой только при узкой полосе частот петлевого фильтра и не отслеживает быстрые изменения фазы в телефонном канале. Таким образом, для устойчивой и быстрой работы

ФАПЧ требуются мгновенные эталонные отсчеты, которые можно получить различными способами.

Например, в качестве опорных можно использовать точки созвездия, ближайшие к принятым точкам x_k [1], или точки y'_k после промежуточного декодирования. Предлагается реализовать промежуточное декодирование без задержки по критерию максимального правдоподобия следующим образом.

Для всех «выживших» путей на текущий дискретный момент времени k выбирается кратчайший путь, которому соответствует декодированная точка y'_k . Использование точки y'_k при работе ФАПЧ обеспечивает более высокую помехоустойчивость и отсутствие ложных нулей дискриминационной характеристики. Следует отметить, что выбор ближайших точек в качестве опорных приводит к увеличению числа ошибок и ложных нулей. В процессе работы модема точки y'_k можно использовать в качестве эталонных точек в петле ФАПЧ.

Фазовая ошибка несущей при использовании ОСР имеет вид:

$$\hat{\Delta\phi} = \arg \left(x_k \left(y'_k - \sum_{i=0}^{N-1} b_i y_{k-i} \right)^* \right), \quad (2)$$

где $x_k = z_k \exp j\phi$ - принятая точка, z_k - искаженный переданный отсчет, $\Delta\phi$ - поворот фазы. Корректирующая добавка в правой части (1) имеет нулевой поворот фазы, поэтому в (2) эта добавка из отсчета (1) перенесена к эталонному отсчету y'_k .

Результаты моделирования

На рис.1 показана зависимость вероятности символьной ошибки от отношения сигнал-шум при разных вариантах декодирования. Результаты получены имитационным моделированием для канала с дискретным временем без межсимвольных помех. Наименьшая вероятность ошибки получена для декодирования Витерби с задержкой в 31 такт (рис. 1.2, 1.5, 1.8). Декодирование путем простого поиска ближайшей эталонной точки созвездия к принятой точке дает наибольшую вероятность ошибки (рис. 1.1, 1.4, 1.7, 1.10). Неполное декодирование Витерби без задержки обеспечивает на порядок меньшую вероятность ошибки (рис. 1.3, 1.6, 1.9).

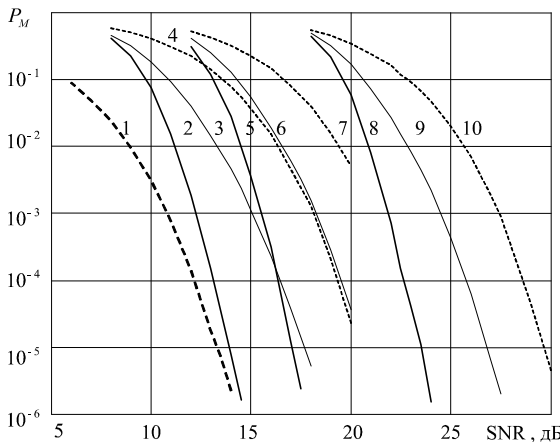


Рис.1. Зависимость вероятности символьной ошибки от отношения сигнал-шум для: 1) скорости 4800 бит/сек (ФМн-4 без решетчатого кодирования); 2) скорости 7200 бит/сек (КАМ-16 с решетчатым кодированием); 3) скорости 7200 бит/сек (КАМ-16 с решетчатым кодированием) и декодирования без задержки; 4) скорости 9600 бит/сек (КАМ-32 с решетчатым кодированием) и декодирования с задержкой 31 ; 5) скорости 9600 бит/сек (КАМ-32 с решетчатым кодированием) и декодирования без задержки; 6) скорости 9600 бит/сек (КАМ-32 с решетчатым кодированием) и декодирования с задержкой 31; 7) скорости 14400 бит/сек (КАМ-128 с решетчатым кодированием) и декодирования с задержкой 31; 8) скорости 14400 бит/сек (КАМ-128 с решетчатым кодированием) и декодирования без задержки; 9) скорости 14400 бит/сек (КАМ-128 с решетчатым кодированием) и декодирования с задержкой 31; 10) скорости 14400 бит/сек (КАМ-128 с решетчатым кодированием) и декодирования без задержки.

7200 бит/сек (КАМ-16 с решетчатым кодированием) и декодирования без задержки; 4) КАМ-16 сигнала без решетчатого кодирования (скорость 9600 бит/сек), 5) скорости 9600 бит/сек (КАМ-32 с решетчатым кодированием) и декодирования с задержкой 31; 6) скорости 9600 бит/сек (КАМ-32 с решетчатым кодированием) и декодирования без задержки; 7) КАМ-32 сигнала без решетчатого кодирования; 8) скорости 14400 бит/сек (КАМ-128 с решетчатым кодированием) и декодирования с задержкой 31; 9) скорости 14400 бит/сек (КАМ-128 с решетчатым кодированием) и декодирования без задержки; 10) КАМ-128 сигнала без решетчатого кодирования.

Для иллюстрации эффективности ОСР используем канал с импульсной характеристикой $\{0,407; 0,815; 0,407\}$ [3]. Рассмотрим влияние искажений в таком канале на вероятность ошибки для сигналов с РК модуляцией. По приведенным в [5] формулам рассчитаем оптимальные значения весовых коэффициентов обратной связи $\{b_i\} = \{-1,903; -0,906\}$ и коэффициентов входного фильтра при условии нулевых межсимвольных помех. На рис.2 показана зависимость вероятности ошибок от отношения сигнал-шум при использовании ОСР. При декодировании Витерби (рис.2.2 и 2.4) обеспечивается на 8 дБ более высокая помехоустойчивость, чем при декодировании по ближайшим точкам (рис.2.8 и 2.9). Увеличение скорости передачи приводит к снижению преимущества полного декодирования (рис.1.2, 1.4, 1.6) по сравнению с декодированием без задержки (рис.1.3, 1.5, 1.7). Однако при полном декодировании наблюдаются отдельные сплошные пачки ошибок, а при частичном – отдельные ошибки, распределенные равномерно. Пачки ошибок исправляются при автоматическом запросе повторения передачи, а отдельные ошибки усредняются в системе ФАПЧ.

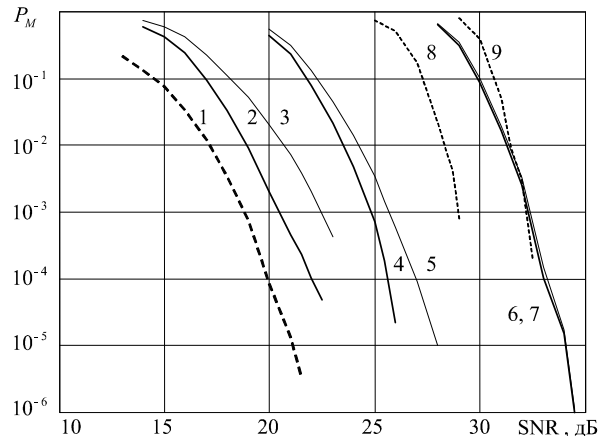
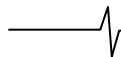


Рис.2. Зависимость вероятности символьной ошибки от отношения сигнал-шум при наличии искажений для: 1) скорости 4800 бит/сек (ФМн-4 без решетчатого кодирования); 2) скорости 7200 бит/сек (КАМ-16 с решетчатым кодированием) и декодирования с задержкой 31; 3) скорости 7200 бит/сек (КАМ-16 с решетчатым кодированием) и декодирования без задержки; 4) скорости 9600 бит/сек (КАМ-32 с решетчатым кодированием) и декодирования с задержкой 31; 5) скорости 9600 бит/сек (КАМ-32 с решетчатым кодированием) и декодирования без задержки; 6) скорости 9600 бит/сек (КАМ-32 с решетчатым кодированием) и декодирования с задержкой 31; 7) скорости 14400 бит/сек (КАМ-128 с решетчатым кодированием) и декодирования с задержкой 31; 8) скорости 14400 бит/сек (КАМ-128 с решетчатым кодированием) и декодирования без задержки; 9) скорости 14400 бит/сек (КАМ-128 с решетчатым кодированием) и декодирования с задержкой 31; 10) скорости 14400 бит/сек (КАМ-128 с решетчатым кодированием) и декодирования без задержки.



рованием) и декодирования без задержки; 6) скорости 14400 бит/сек (КАМ-128 с решетчатым кодированием) и декодирования с задержкой 31; 7) скорости 14400 бит/сек (КАМ-128 с решетчатым кодированием) и декодирования без задержки; 8) КАМ-16 сигнала без решетчатого кодирования; 9) КАМ-32 сигнала без решетчатого кодирования.

Заключение

Разработан алгоритм декодирования РК кодов совместно с ОСР, позволяющий увеличить помехоустойчивость в каналах с резкими провалами в частотной характеристике и узкополосными помехами. Показано, что частично декодированные с ОСР отсчеты могут быть использованы в системе ФАПЧ. Проведено имитационное моделирование работы декодера при наличии спектральных нулей и показано, что решетчатое кодирование на 8 дБ улучшает помехоустойчивость при ОСР.

Литература

1. John A.C. Bingham. «The theory and practice of modern design», издательство John Wiley and Sons, 605 Third Ave., New York, NY 10158. (A WILEY-INTERSCIENCE PUBLICATION. 1988. 454 pp.).
2. Скляр Б. Цифровая связь. Теоретические основы и практическое применение: Пер. с англ. 2-е изд. испр. М.: Издат. дом «Вильямс», 2004. 1099с.
3. Прокис Д. Цифровая связь. Пер. с англ. / Под ред. Д.Д. Кловского. М.: Радио и связь, 2000. 800 с.
4. Форни Дж. Алгоритм Витерби // ТИИЭР. 1973. №3. С.12-25.
5. Бельфиоре К.А., Парк Дж. Х. Компенсация посредством решающей обратной связи // ТИИЭР, 1979, т.67. №8. С.67-83.
6. Кочетков А.А. Оптимизация алгоритма декодирования Витерби для треллис-модулированных сигналов // Проблемы передачи и обработки информации в сетях и системах телекоммуникаций: Материалы 11-й Международной науч.-техн. конф. Рязань: РГРТА, 2002. С.124-126.
7. Астапкович К.Ф. и др. Цифровой модем-9600 для каналов ТЧ // Электросвязь. 1984. №3. С.15-18.

ИНФОРМАЦИОННОЕ ПИСЬМО

Уважаемый коллега!

*Приглашаем Вас принять участие в работе 11-ой Международной конференции и выставки
«Цифровая обработка сигналов и ее применение – DSPA'2009»,
которая состоится в Москве 24-26 марта 2009 года*

ОРГАНИЗАТОРЫ:

- Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова
- IEEE Signal Processing Society
- Российская секция IEEE
- Институт радиотехники и электроники РАН
- Институт проблем управления РАН
- Институт проблем передачи информации РАН
- Московский научно-исследовательский телевизионный институт
- Компания AUTEX Ltd. (АВТЭКС)

ОСНОВНЫЕ НАПРАВЛЕНИЯ РАБОТЫ КОНФЕРЕНЦИИ:

- Теория сигналов и систем
- Теория и методы ЦОС
- Цифровая обработка многомерных сигналов
- Цифровая обработка речевых и звуковых сигналов
- Цифровая обработка изображений
- ЦОС в системах телекоммуникаций
- ЦОС в радиотехнических системах
- ЦОС в системах управления и робототехники
- Цифровая обработка измерительной информации
- Нейрокомпьютерная обработка сигналов и изображений
- Цифровое телевидение и телевидение
- Цифровое радиовещание
- ЦОС в системах защиты информации
- Проектирование и техническая реализация систем ЦОС
- ЦОС в открытых системах
- Проблемы подготовки специалистов в области ЦОС

ФИКСИРОВАННАЯ ИЛИ ПЛАВАЮЩАЯ ТОЧКА? ТЕПЕРЬ ЭТО НЕ ВОПРОС

Разработчики цифровых систем управления должны по достоинству оценить революционный прорыв в подходе к проектированию приложений. На смену эпохе программирования на ассемблере, скрупулезной оптимизации кодов, работающих в формате с фиксированной точкой, и их поддержке в конечных изделиях пришла новая альтернатива – использование процессоров с плавающей точкой в классе цифровых сигнальных контроллеров, позволяя передовым технологиям в приложениях промышленного управления больше не искать компромисса между производительностью системы и ее ценой.

15 лет назад компания Texas Instruments разработала первый цифровой сигнальный процессор (ЦСП) со встроенной флеш-памятью и набором периферии для задач управления, включая аналого-цифровой преобразователь (АЦП) и модуль широтно-импульсной модуляции (ШИМ). Это было ответом на требования рынка промышленного оборудования, где появилась необходимость в управляющих модулях, совмещающих уровень интеграции микроконтроллеров (МК) и вычислительную производительность ЦСП. Так появился первый цифровой сигнальный контроллер (ЦСК) TMS320F240 семейства C2000™, работающий в формате с фиксированной точкой на частоте 20 МГц. С тех пор семейство существенно расширилось, демонстрируя повышение интеграции (сверхбыстродействующие АЦП, ШИМ высокого разрешения, улучшенные порты подключений) и увеличение производительности (рабочие частоты до 150 МГц), позволяя разработчикам выбирать требуемое конкретным приложением сочетание мощности и цены. О номенклатуре изделий в семействе C2000™ можно узнать на сайте производителя: www.ti.com/C2000.

Разработчики следующего поколения столкнутся с новыми требованиями к вычислительной элементной базе, делающими выбор цифрового управляющего модуля еще более важным для успеха конечного изделия. Сегодня промышленность нацелена на большую дифференциацию приложений, высокую «интеллектуальность» вычислительных систем и широкие возможности обмена информацией. Экономические и социальные условия требуют снижения энергозатрат, повышения надежности, большего динамического диапазона при одновременном снижении цены и сокращении времени выпуска готовых изделий.

Важность вопроса экономии электроэнергии отражается на любом оборудовании от устройств широкого потребления до промышленных установок. На работу двигателя расходуется до двух третей всей электроэнергии промышленного потребления. При этом большинство двигателей могло бы работать более эффективно, если бы использовало переменную скорость вращения. Однако сейчас лишь один из двадцати двигателей включает цепь цифрового управления скоростью вращения. А ведь двигатели с переменной скоростью способны ежегодно экономить энергию, вырабатываемую десятью электростанциями.

После этого становится очевидным, что цифровое управление играет существенную роль в вопросах энергосбережения и сохранения окружающей среды и ее ресурсов.

Другая перспективная сфера использования цифровых систем управления – это альтернативные источники энергии – солнечные батареи и ветряные турбины. Также как и электродвигатели эти виды систем нуждаются в использовании передовых технологий управления.

Таким образом, цифровое управление позволяет решать соответствующие задачи с повышенным качеством в широком диапазоне изменяющихся условий с минимальным расходом электроэнергии.

Проектирование подобных систем обычно начинается с разработки алгоритма управления и моделирования его работы в системе МАТЛАБ. При этом ядро алгоритма реализуется в виде программы на языке Си или Си++ и работает с сигналами в формате с плавающей точкой. До сегодняшнего дня последующая разработка, связанная с переходом от абстрактной математики к конкретной аппаратуре, требовала принятия важного решения. Первый вариант – это использование процессора с плавающей точкой, так сказать, «большого цифрового мозга», с ограниченной интеграцией периферии, требующей подключения дополнительного микроконтроллера, управления им и использования внешних периферийных модулей. Данное решение подходит для ряда передовых устройств, выпускаемых в ограниченном количестве когда стоимость системы не является решающим фактором.

В большинстве же случаев приходится использовать второй вариант. Требование снижения цены конечного оборудования заставляет адаптировать программу к реализации на 16-разрядных или 32-разрядных процессорах с фиксированной точкой, которые и используются в большинстве систем. Этот процесс адаптации алгоритма, рассчитанного на формат с плавающей точкой под архитектуру устройства, работающего в формате с фиксированной точкой, является достаточно сложным и требует от разработчика определенных знаний и усилий. На доработку программных кодов уходят многие дни и недели. Кроме того, программа, разработанная для процессора с фиксированной точкой, оказывается плохо соотносима с исходным алгоритмом. В результате отладка, исследования, последующая модернизация и даже документирование таких программных проектов оказываются весьма неудобными.

Теперь, однако, разработчики могут выбрать другой подход, не требующий от них идти на какие-либо компромиссы. В семействе цифровых сигнальных контроллеров C2000™ компании Texas Instruments появилась новая линейка процессоров TMS320F283x, работающих в формате с плавающей точкой, обеспечивая повышенную производительность и упрощение процесса разработки программного обеспечения для цифровых систем управления. Кроме того, использование арифметики с плавающей точкой позволяет сокращать вычислительные затраты и экономить память, а также реализовывать более сложные алгоритмы, более эффективные с позиции функциональности и экономичности.

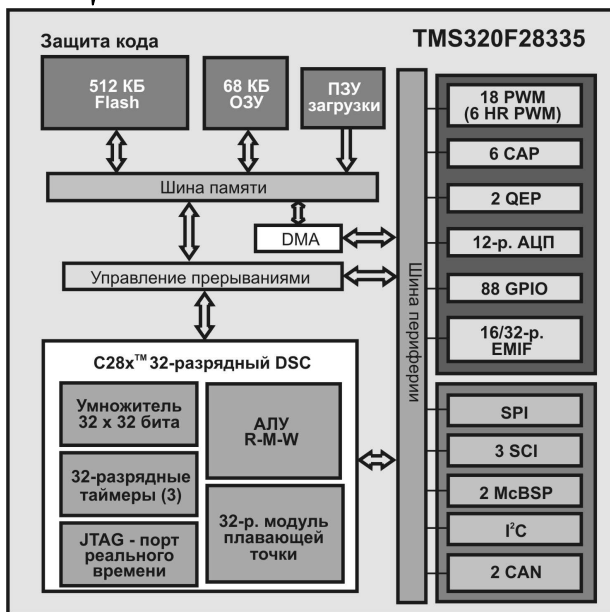


Рис. 1. Архитектура процессоров TMS320C283x

На рис. 1 проиллюстрирована архитектура процессоров TMS320C283x. 32-разрядное ядро цифрового сигнального контроллера (DSC) включает арифметико-логическое устройство (АЛУ), организованное по принципу «чтение-модификация-запись» (R-M-W), умножитель 32 на 32 бита и сопроцессор вычислений в формате с плавающей точкой, а также таймеры общего назначения и внутрисхемный эмулятор JTAG. Процессор включает загрузочное ПЗУ, 512 КБ ПЗУ Flash и 68 КБ ОЗУ на кристалле. Обращает на себя внимание развитая периферия устройства: блоки широтно-импульсной модуляции (PWM), в том числе с повышенной разрешающей способностью (HR PWM); модули захвата событий (CAP); модуль квадратурного кодирования для работы с датчиками перемещения (QEP); 12-разрядный АЦП (ADC); ввод-вывод общего назначения (GPIO); интерфейс внешней памяти (EMIF); последовательный интерфейс SPI; последовательный связной интерфейс (SCI – аналог UART); многоканальные буферизированные последовательные порты (McBSP); межсхемный интерфейс I²C; асинхронный последовательный связной интерфейс CAN.

Процессоры TMS320F283x предоставляют уникальную возможность наравне с плавающей точкой использовать и формат с фиксированной точкой. Это позволяет в пределах одного программного кода обеспечить одновременно и широкий динамический диапазон, и для некоторых наиболее важных переменных высокую точность представления чисел, характерную для формата с фиксированной точкой. Кроме того, разработчик может создавать один единственный исходный код с использованием библиотеки IQMath™ (www.ti.com/C2000sigproclib) и затем на этапе компиляции задавать режим фиксированной или плавающей точки. Это оказывается очень удобно, так как дает возможность повторно использовать ранее созданные программы и легко переходить от одних представителей семейства C2000™ к другим.

Использование цифровых сигнальных контроллеров с плавающей точкой существенно ускоряет выпуск готовых изделий, позволяя быстро создавать прототип программного проекта, использовать средства автоматической генерации машинных кодов, легко повторять переход от конечной аппаратной платформы назад в среду моделирования и вновь

возвращаться к конечным изделиям, отлаживая, оптимизируя или модернизируя программный продукт.

Выпуск серии процессоров с плавающей точкой TMS320F283x начинается с трех первых моделей, отличающихся объемом внутрикристалльной памяти и составом периферийных узлов.

Одновременно с процессорами TMS320F283x компания Texas Instruments предлагает линейку процессоров с фиксированной точкой TMS320F282x, которые полностью совместимы с F283x по кодам и разъемам. Это позволяет вести быструю разработку программного обеспечения, использующего формат с плавающей точкой, а затем перекомпилировать тот же исходный код для работы с фиксированной точкой. Дополнительная информация может быть найдена на сайте www.ti.com/f2823x.

Чтобы быстро ознакомиться с новой технологией, компания Texas Instruments предлагает приобрести стартовый набор F28335 eZdsp по относительно невысокой цене. Этот набор предоставляет полный ряд инструментальных средств, включая плату конечного изделия с выведенной и доступной для подключения периферией, порт JTAG, обеспечивающий связь платы с персональным компьютером и предоставляющий возможности обмена данными с работающим устройством в реальном масштабе времени, интегрированную среду разработки и отладки программного обеспечения Code Composer Studio. Стартовый набор F28335 eZdsp предлагают все официальные дистрибьюторы компании Texas Instruments. Также его можно заказать через электронный магазин (eStore) компании. Набор включен в Университетскую программу компании Texas Instruments и может быть приобретен образовательными учреждениями на льготных условиях. О механизмах реализации Университетской программы TI мы рассказывали на страницах предыдущих выпусков журнала «Цифровая обработка сигналов» (№4 за 2007 год, стр. 54 и №1 за 2008 год, стр. 68). Данная информация размещается также на страницах сайта журнала в разделе «ЦОС в университетах»: <http://www.dsps.ru/cosvuz>. Примеры программного обеспечения для платы F28335 eZdsp могут быть бесплатно загружены из сети Интернет: <http://www.ti.com/c2000getstarted>. В качестве поддержки разработчика компания TI предлагает обучающие видеоролики, свободно доступные по адресу: <http://www.ti.com/lit/mp4/sprc540>. Можно заказать или скачать руководство по освоению платформы C2000 и процессоров C283x «Getting Started with DSP» («Начиная работать с ЦСП»), комплектующее бесплатной демоверсией среды Code Composer Studio, работать с которой можно в течение 120 дней (<http://www.ti.com/freetools>).

Итак, разработчикам цифровых систем управления предоставлена возможность коренным образом изменить традиционные подходы к проектированию новых устройств. С появлением сигнальных контроллеров с плавающей точкой TMS320F2833x необходимость в скрупулезной отладке и оптимизации программных кодов, поддерживающих формат с фиксированной точкой, исчезает. Теперь разработчики смогут по достоинству оценить все удобство и эффективность современных способов разработки приложений, ключевым из которых, является автоматическая генерация кодов программ.

Материалы статьи предоставлены менеджером в странах Европы и Северной Африки Университетской программы компании Texas Instruments Inc. Робертом Оуэном.