

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

4/2007

НАУЧНО - ТЕХНИЧЕСКИЙ ЖУРНАЛ

- АДАПТИВНАЯ ФИЛЬТРАЦИЯ
- АНАЛИЗ И ОБРАБОТКА РЕЧИ
- ЦОС В РАДИОЛОКАЦИИ
- ПРОГРАММНАЯ РЕАЛИЗАЦИЯ НА DSP



ISSN 1684-2634



9 771684 263005 >

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

Научно-технический журнал
№ 4/2007

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. Мишенков,
А.А.Петровский, Ю.Н. Прохоров,
А.Н. Соловьев, Ю.Г. Сосулин,
В.В. Шахгильдян, 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.dspsa.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. Андреехина
Дизайн: И.Е. Артюхина

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

Подписано в печать 18.01.08
Формат 60х90/8.

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

Заказ № . 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:

Джиган В.И., Плетнева И.Д.

**Алгоритмы адаптивной фильтрации на основе QR-
разложения для антенных решеток систем
цифровой связи 2**

Соловьева Е.Б. Жеребцов А.П.

**Синтез полиномиальных фильтров импульсных
помех, поражающих речевые сигналы 8**

Витязев В.В., Витязев С.В.

**Методы синтеза узкополосного адаптивного КИХ-
фильтра на основе многоскоростной обработки .. 15**

Сахаров А.В.

**Построение статистической модели речевого
трафика при монологе 21**

Леонович А.А.

Проблемы распознавания слитной речи 25

Ашинянц Р.А., Морозова Т.Ю.

**Регуляризация алгоритма фильтрации Калмана-
Бьюси при плохой обусловленности корреляционной
матрицы шума 29**

Зайцев Г.В.

**Фильтрация радиолокационных сигналов при
использовании различных математических
представлений функции 33**

Хвостиков А.С., Щетинин В.С.

**Диагностирование процессов резания с помощью
вейвлет-анализа сигнала акустической эмиссии ... 40**

Бартенев В.Г., Бартенев Г.В.

**Программа адаптивного LMS компенсатора для
комплексных выборок наблюдений 44**

Нагулин Н.Е., Нагулин С.Н.

**Алгоритм параллельного вычисления быстрого
преобразования Фурье для транспьютероподобных
сигнальных процессоров 48**

Кошелев В.И., Белокуров В.А.

**Реализация метода Гуда-Томаса для доплеровской
фильтрации сигналов 51**

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

АЛГОРИТМЫ АДАПТИВНОЙ ФИЛЬТРАЦИИ НА ОСНОВЕ QR-РАЗЛОЖЕНИЯ ДЛЯ АНТЕННЫХ РЕШЕТОК СИСТЕМ ЦИФРОВОЙ СВЯЗИ

Джиган В.И., Плетнева И.Д.

Введение

Сегодня адаптивные антенные решетки (ААР) [1] широко используются в системах беспроводной связи [2]. В случае различной пространственной ориентации источника полезного сигнала и источников помех такие решетки обеспечивают повышение качества приема информации за счет подавления помех, находящихся в одной полосе частот с полезным сигналом. Вычислительная сложность, т.е. число арифметических операций, требуемых для выполнения одной итерации алгоритмов адаптивной фильтрации, лежащих в основе алгоритмов управления ААР, всегда пропорциональна числу весовых коэффициентов (антенн) N . Так как в системах связи в основном применяются антенные решетки с небольшим числом антенн, то не только простые (с линейной вычислительной сложностью), но и сложные адаптивные алгоритмы (с квадратичной вычислительной сложностью) [3] уже вполне реализуемы на современных цифровых сигнальных процессорах (ЦСП).

Использование алгоритмов адаптивной фильтрации обычно требует наличия тренировочных последовательностей. Если такие последовательности стандартом связи не предусмотрены, то можно использовать алгоритмы на основе линейно-ограниченного критерия [4] или на основе критерия Constant Modulus (CM) [5]. Недостатками линейно-ограниченных алгоритмов являются необходимость знания направления на источник полезного сигнала (в координатах, связанных с антенной решеткой), большая вычислительная сложность (обусловленная матрицей ограничений), а также чувствительность к ошибкам между фактическим и аналитическим (задаваемым ограничениями) направлениями на источник полезного сигнала. Алгоритмы адаптивной фильтрации на основе CM-критерия лишены указанных недостатков.

Вначале адаптивные алгоритмы на основе CM-критерия в основном использовались для построения «слепых», т.е. не требующих тренировочных последовательностей выравнивателей (эквалайзеров) каналов связи. Позже такие алгоритмы стали применяться и в ААР [6]. На сегодняшний день в трудах IEEE опубликовано около 900 работ по различным аспектам адаптивной фильтрации на основе CM-критерия. Из них только около 250 работ (в основном в трудах научно-технических конференций) относятся к ААР, что характеризует привлекательность адаптивных алгоритмов на основе CM-критерия, а также актуальность исследова-

Рассматривается использование алгоритмов адаптивной фильтрации на основе обратного QR-разложения в адаптивных антенных решетках по критерию постоянства модуля огибающей сигналов. Приводятся вычислительные процедуры таких алгоритмов, использующих вращения Гивенса с операциями и без операций извлечения квадратного корня. Результаты моделирования демонстрируют эффективность алгоритмов в задачах подавления внешних помех, а оценки требуемых вычислительных ресурсов – реализуемость этих алгоритмов на современных цифровых сигнальных процессорах.

ний, направленных на использование этих алгоритмов в ААР.

В настоящей работе рассматривается применение рекурсивных алгоритмов по критерию наименьших квадратов (Recursive Least Squares, RLS) на основе обратного QR-разложения в качестве алгоритмов управления ААР, функционирующих по CM-критерию.

Адаптивные алгоритмы на основе CM-критерия применяются для обработки сигналов с модуляцией типа Binary Phase Shift Keying (BPSK), Quadrature Phase Shift Keying (QPSK), а также с другими типами модуляции, характеризуемыми постоянным значением огибающей (модуля) информационных символов. Любой информационный символ a_i при таких видах модуляции обладает свойством $|a_i| = a_i^* a_i = S = \text{const}$, где значение S определяется типом модуляции и является известным на приемной стороне. В общем случае CM-критерий формулируется как

$$J(p, q) = E \left[\left| s^p - |y(k)|^q \right|^q \right], \quad (1)$$

а соответствующие адаптивные алгоритмы обозначаются как CM (p, q) . Здесь $E[\cdot]$ – операция усреднения, $y(k) = \mathbf{h}_N^H(k-1) \mathbf{x}_N(k)$ – комплексный выходной сигнал антенной решетки (рис. 1), $\mathbf{h}_N(k) = [h_1(k), \dots, h_n(k), \dots, h_N(k)]^T$ – вектор комплексных весовых коэффициентов, $\mathbf{x}_N(k) = [x_1(k), \dots, x_n(k), \dots, x_N(k)]^T$ – комплексный вектор пространственно-временных отсчетов сигналов, k – индекс дискретного времени, верхние индексы T и H обозначают соответственно транспонирование и эрмитово сопряжение (транспонирование и комплексное сопряжение) векторов и матриц, а верхний индекс $*$ обозначает операцию комплексного сопряжения переменной. Матрицы обозначаются полужирными прописными символами, а векторы – полужирными строчными символами. Нижний индекс N в обозначении матриц и векторов показывает, что рассматривается квадратная матрица с числом элементов $N \times N$ или вектор с числом элементов N .

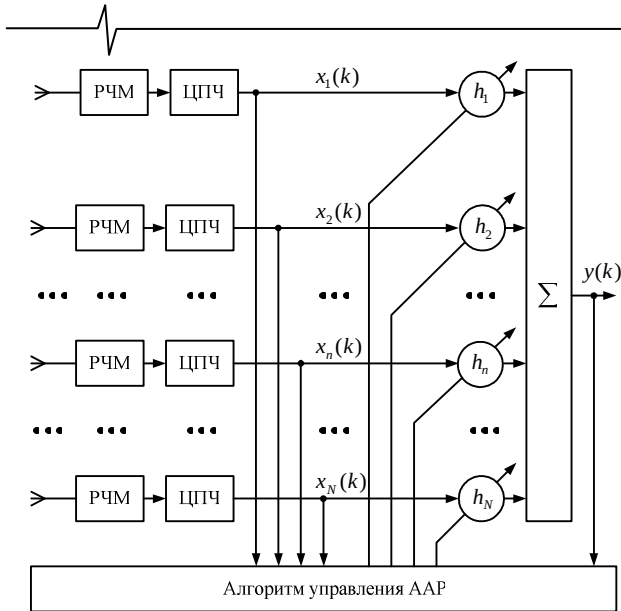


Рис. 1. Адаптивная антенная решетка без опорного сигнала: РЧМ – радиочастотный модуль, ЦПЧ – цифровой преобразователь частоты.

Критерий (1) является нелинейным, а потому использование простых градиентных алгоритмов в качестве алгоритмов адаптивной фильтрации часто приводит к локальным решениям задачи минимизации этого критерия. RLS-алгоритмы на основе леммы об обращении матрицы или QR-разложения также напрямую не могут быть использованы в этом случае.

В то же время, согласно [7], при $q = 2$, СМ-критерий (1) может быть преобразован в квадратичный критерий:

$$J'(k) = \sum_{i=1}^k \lambda^{k-i} |s^p - \mathbf{h}_N^H(k) \mathbf{z}_N(i)|^2, \quad (2)$$

где $\mathbf{z}_N(k) = \mathbf{x}_N(k) \mathbf{x}_N^H(k) \mathbf{h}_N(k-1) \mathbf{x}_N^H(k) \mathbf{h}_N(k-1)^{p-2}$, $(1 - 0.4/N) \leq \lambda < 1$ – параметр экспоненциального взвешивания, предназначенный для слежения в небольших пределах за медленно изменяющимися сигналами. Это позволяет применить любой из известных RLS-алгоритмов с квадратичной вычислительной сложностью $O(N^2)$ в качестве алгоритма поиска вектора весовых коэффициентов, минимизирующего функционал (2). В [7] для данной цели использован RLS-алгоритм на основе леммы об обращении матриц [8]. Соответствующие модификации RLS-алгоритмов на основе QR-разложения [9] для решения этой же задачи приведены в табл. 1 и табл. 2.

Если $p = 2$, то $\mathbf{z}_N(k) = \mathbf{x}_N(k) \mathbf{x}_N^H(k) \mathbf{h}_N(k-1)$. В этом случае вычисление вектора $\mathbf{z}_N(k)$ требует минимального числа арифметических операций, равного $2N$ умножениям и N сложениям, а длительность переходного процесса алгоритма адаптивной фильтрации минимальна, так как переменная $\mathbf{z}_N(k)$ обладает минимальной памятью [7]. Кроме того, в условиях низких отношений сигнал-шум (ОСШ) адаптивным фильтром обеспечивается меньшее значение среднеквадратичной ошибки при $p = 2$, чем при $p = 1$ [10].

Таблица 1.
СМ(2,2) RLS-алгоритм на основе обратного QR-разложения с использованием вращений Гивенса

Вычисления	Ссылки
Инициализация: $\mathbf{x}_N(0) = \mathbf{0}_N$, $\tilde{\mathbf{R}}_N^{-1}(0) = \sqrt{\delta^{-2}} \mathbf{I}_N$, $\mathbf{h}_N(0) = [\mathbf{1}, \mathbf{0}_{N-1}^T]^T$	(1.0)
For $k = 1, 2, \dots, K$	
$\mathbf{u}_N^{(0)H}(k) = \mathbf{0}_N^T$, $b_N^{(0)}(k) = 1$	(1.1)
$\mathbf{z}_N(k) = \mathbf{x}_N(k) \mathbf{x}_N^H(k) \mathbf{h}_N(k-1)$	(1.2)
For $i = 1, 2, \dots, N$	
$a_{N,i}(k) = \lambda^{-0.5} \tilde{\mathbf{R}}_N^{-H}(k-1) \begin{vmatrix} \mathbf{z}_N(k) \\ i, 1:i \end{vmatrix}$	(1.3)
$b_N^{(i)}(k) = \sqrt{[b_N^{(i-1)}(k)]^2 + a_{N,i}^*(k) a_{N,i}(k)}$	(1.4)
$s_{N,i}(k) = a_{N,i}^*(k) / b_N^{(i)}(k)$	(1.5)
$c_{N,i}(k) = b_N^{(i-1)}(k) / b_N^{(i)}(k)$	(1.6)
For $j = 1, 2, \dots, i$	
$\tilde{\mathbf{R}}_{N,i,j}^{-H}(k) = \lambda^{-0.5} c_{N,i}(k) \tilde{\mathbf{R}}_{N,i,j}^{-H}(k-1) - s_{N,i}^*(k) u_{N,j}^{(i-1)*}(k)$	(1.7)
$u_{N,j}^{(i)*}(k) = \lambda^{-0.5} s_{N,i}(k) \tilde{\mathbf{R}}_{N,i,j}^{-H}(k-1) + c_{N,i}(k) u_{N,j}^{(i-1)*}(k)$	(1.8)
End for j	
End for i	
$\mathbf{g}_N(k) = \mathbf{u}_N^{(N)}(k) / b_N^{(N)}(k)$	(1.9)
$\alpha_N(k) = s^2 - \mathbf{h}_N^H(k-1) \mathbf{z}_N(k)$	(1.10)
$\mathbf{h}_N(k) = \mathbf{h}_N(k-1) + \mathbf{g}_N(k) \alpha_N^*(k)$	(1.11)
End for k	

СМ(2,2) RLS-алгоритм на основе обратного QR-разложения
с использованием вращений Гивенса без операций извлечения квадратного корня

Вычисления	Ссылки
Инициализация: $\mathbf{x}_N(0) = \mathbf{0}_N, \bar{\mathbf{R}}_N^{-1}(0) = \sqrt{\delta^{-2} \mathbf{I}_N}, \mathbf{h}_N(0) = [1, \mathbf{0}_{N-1}^T]^T, \mathbf{K}_N^R(0) = \mathbf{I}_N$	(2.0)
For $k = 1, 2, \dots, K$	
$\bar{\mathbf{u}}_N^{(0)H}(k) = \mathbf{0}_N^T, K_N^{B(0)}(k) = 1$	(2.1)
$\mathbf{z}_N(k) = \mathbf{x}_N(k) \mathbf{x}_N^H(k) \mathbf{h}_N(k-1)$	(2.2)
For $i = 1, 2, \dots, N$	
$\bar{a}_{N,i}(k) = \bar{\mathbf{R}}_N^{-H}(k-1) \left \begin{smallmatrix} \mathbf{z}_N(k) \\ i, 1:i \end{smallmatrix} \right _{1:i}$	(2.3)
$K_N^{B(i)}(k) = K_N^{B(i-1)}(k) + \lambda^{-1} K_{N,i}^R(k-1) \bar{a}_{N,i}^*(k) \bar{a}_{N,i}(k)$	(2.4)
$\bar{s}_{N,i}(k) = \lambda^{-1} K_{N,i}^R(k-1) \bar{a}_{N,i}^*(k) / K_N^{B(i)}(k)$	(2.5)
$\bar{c}_{N,i}(k) = K_N^{B(i-1)}(k) / K_N^{B(i)}(k)$	(2.6)
For $j = 1, 2, \dots, i$	
$\bar{R}_{N,i,j}^{-H}(k) = \bar{R}_{N,i,j}^{-H}(k-1) - \bar{a}_{N,i}(k) \bar{u}_{N,j}^{(i-1)*}(k)$	(2.7)
$\bar{u}_{N,j}^{(i)*}(k) = \bar{s}_{N,i}(k) \bar{R}_{N,i,j}^{-H}(k-1) + \bar{c}_{N,i}(k) \bar{u}_{N,j}^{(i-1)*}(k)$	(2.8)
End for j	
$K_{N,i}^R(k) = \lambda^{-1} K_{N,i}^R(k-1) \bar{c}_{N,i}(k)$	(2.9)
End for i	
$\mathbf{g}_N(k) = \bar{\mathbf{u}}_N^{(N)}(k)$	(2.10)
$\alpha_N(k) = s^2 - \mathbf{h}_N^H(k-1) \mathbf{z}_N(k)$	(2.11)
$\mathbf{h}_N(k) = \mathbf{h}_N(k-1) + \mathbf{g}_N(k) \alpha_N^*(k)$	(2.12)
End for k	

В табл. 1 и табл. 2 символами $\mathbf{0}_N$ и $\mathbf{I}_N$ обозначены нуль-вектор и единичная матрица, соответственно. Векторы весовых коэффициентов инициализируются как $\mathbf{h}_N(0) = [1, \mathbf{0}_{N-1}^T]^T$ или $\mathbf{h}_N(0) = [1/N, \mathbf{0}_{N-1}^T]^T$. Это означает, что в начале работы алгоритмов (до первой итерации) только одна антенна принимает сигналы. Поскольку диаграмма направленности (ДН) антенного элемента ААР предполагается всенаправленной, то направление на источник полезного сигнала может быть неизвестным. В процессе адаптации формируется ДН с максимумом основного лепестка в направлении на источник полезного сигнала и с провалами в направлениях на источники помех.

Параметр δ^2 используется для начальной регуляризации обращения корреляционной матрицы

$$\mathbf{R}_N(k) = \sum_{i=1}^k \lambda^{k-i} \mathbf{z}_N(i) \mathbf{z}_N^H(i) = \tilde{\mathbf{R}}_N^H(k) \tilde{\mathbf{R}}_N(k),$$

где $\tilde{\mathbf{R}}_N(k)$ – верхняя треугольная матрица, значения которой $\tilde{\mathbf{R}}_N^H(k) = \mathbf{K}_N^{0.5R}(k) \bar{\mathbf{R}}_N^{-H}(k)$ (нижняя треугольная матрица) используются в алгоритмах. Параметр регуляризации выбирается из условия $\delta^2 \geq 0.01 \sigma_z^2$ [11], где σ_z^2 – дисперсия сигналов $z(k)$.

Нижние индексы i и j в обозначениях переменных указывают на номера используемых элементов векторов и матриц, обозначение $(\cdot)_{i,1:i}$ указывает на номера элементов строк матриц, а обозначение $(\cdot)_{1:i}$ – на номера элементов векторов.

В алгоритме (табл. 1) для определения вектора коэффициентов Калмана $\mathbf{g}_N(k)$, с помощью которых в уравнении (1.11) вычисляется вектор весовых коэффициентов $\mathbf{h}_N(k)$, используются фигурирующие в (1.9) переменные $\mathbf{u}_N(k)$ и $b_N(k)$. Эти переменные находят с помощью соотношения:

$$\mathbf{P}_{N+1}(k) \begin{bmatrix} \lambda^{0.5} \tilde{\mathbf{a}}_N(k) & \lambda^{0.5} \tilde{\mathbf{R}}_N^{-H}(k-1) \\ 1 & \mathbf{0}_N^T \end{bmatrix} = \begin{bmatrix} \mathbf{0}_N & \tilde{\mathbf{R}}_N^{-H}(k) \\ b_N(k) & \mathbf{u}_N^H(k) \end{bmatrix}, \quad (3)$$

где $\tilde{\mathbf{a}}_N(k) = \tilde{\mathbf{R}}_N^{-H}(k-1) \mathbf{z}_N(k)$,

$b_N(k) = \sqrt{1 + \lambda^{-1} \tilde{\mathbf{a}}_N^H(k) \tilde{\mathbf{a}}_N(k)}$, $\mathbf{P}_{N+1}(k)$ – матрица геометрических вращений Гивенса, определяемая уравнением:

$$\mathbf{P}_{N+1}(k) = \prod_{i=1}^N \mathbf{P}_{N+1}^{(i)}(k) = \prod_{i=1}^N \begin{bmatrix} \mathbf{I}_{i-1} & \mathbf{0}_{i-1} & \mathbf{O}_{i-1, N+1-i} & \mathbf{0}_{i-1} \\ \mathbf{0}_{i-1}^T & c_{N,i}(k) & \mathbf{0}_{N+1-i}^T & -s_{N,i}^*(k) \\ \mathbf{O}_{N+1-i, i-1} & \mathbf{0}_{N+1-i} & \mathbf{I}_{N+1-i} & \mathbf{0}_{N+1-i} \\ \mathbf{0}_{i-1}^T & s_{N,i}(k) & \mathbf{0}_{N+1-i}^T & c_{N,i}(k) \end{bmatrix}$$

Переменные $s_{N,i}(k)$ и $c_{N,i}(k)$ называются синусами и косинусами углов вращения [12] и удовлетворяют соотношениям $|c_{N,i}(k)|^2 + |s_{N,i}(k)|^2 = 1$. Здесь $\mathbf{O}_{n,m}$ – прямоугольная нулевая матрица с числом элементов, равным $n \times m$.

Чтобы избавиться от операций извлечения квадратного корня матрицы, векторы в уравнении (3), подобно алгоритмам на основе прямого QR-разложения [13], масштабируются как

$$\begin{bmatrix} \lambda^{-0.5} \mathbf{K}_N^{0.5R}(k-1) \bar{\mathbf{a}}_N(k) & \lambda^{-0.5} \mathbf{K}_N^{0.5R}(k-1) \bar{\mathbf{R}}_N^{-H}(k-1) \\ \mathbf{K}_N^{0.5B}(k) & \mathbf{0}_N^T \end{bmatrix}$$

и

$$\begin{bmatrix} \mathbf{0}_N^T & \mathbf{K}_N^{0.5R}(k) \bar{\mathbf{R}}_N^{-H}(k) \\ \mathbf{K}_N^{0.5B}(k) & \mathbf{K}_N^{0.5B}(k) \bar{\mathbf{u}}_N^H(k) \end{bmatrix}.$$

Результатом этого приема является алгоритм (табл. 2).

Алгоритмы (табл. 1 и табл. 2), а также алгоритм [7], математически эквивалентны друг другу и отличаются лишь вычислительной сложностью. Эквивалентность означает, что алгоритмы получены путем тождественных математических преобразований соотношений, обеспечивающих минимизацию функционала (2). Поэтому при реализации в арифметике с плавающей точкой, одинаковых параметрах адаптивного фильтра и одинаковых условиях функционирования, т.е. при обработке одинаковых сигналов, сравниваемые алгоритмы вычисляют на каждой итерации одинаковые показатели ААР (выходные сигналы, векторы весовых коэффициентов и др.). Эти показатели различаются лишь ошибками округления, так как в каждом из алгоритмов используется своя последовательность вычислений со своим числом арифметических операций. Эквивалентность обеспечивается на всем интервале наблюдения в случае, если алгоритмы функционируют устойчиво.

Алгоритм (табл. 1) содержит N операций извлечения квадратного корня, что обусловлено вычислением переменной $b_N(k)$ в уравнении (1.4) в течение N шагов $i=1, \dots, N$ на каждой k -й итерации алгоритма адаптивной фильтрации. Алгоритм (табл. 1) также содержит $3N^2 + 12N$ операций умножения, $1.5N^2 + 5.5N + 1$ операций сложения и N операций деления.

Вычислительная сложность алгоритма (табл. 2) равна $2.5N^2 + 12N$ операциям умножения, $1.5N^2 + 5.5N + 1$ операциям сложения и N операциям деления. Операции извлечения квадратного корня в алгоритме отсутствуют.

Для сравнения вычислительная сложность алгоритма [7] равна $4N^2 + 6N$ операциям умножения, $3N^2 + 5N + 2$ операциям сложения и одной операции деления.

В приведенных оценках сложности также учитывается число арифметических операций при вычислении вектора $\mathbf{z}_N(k) = \mathbf{x}_N(k) \mathbf{y}^*(k)$, куда входит и вычисление выходного сигнала антенной решетки $\mathbf{y}(k)$. В алгоритмах в основном используются комплексные арифметические операции.

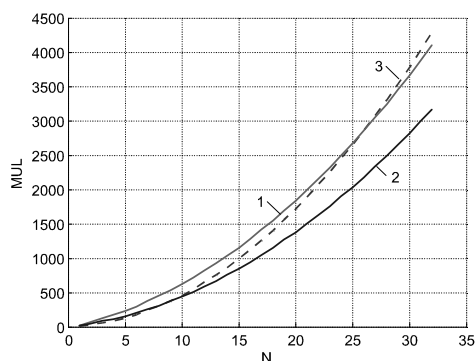


Рис. 2. Требуемое число умножений на одну итерацию: 1 – алгоритм (табл. 1), 2 – алгоритм (табл. 2), 3 – алгоритм [7].

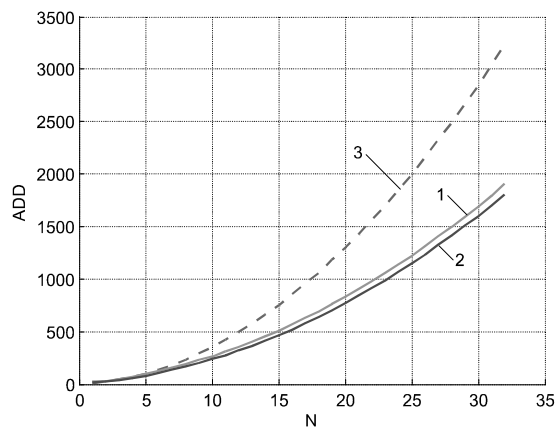


Рис. 3. Требуемое число сложений на одну итерацию: 1 – алгоритм (табл. 1), 2 – алгоритм (табл. 2), 3 – алгоритм [7].

Для сравнения вычислительной сложности рассмотренных алгоритмов необходимо представить операции извлечения квадратного корня и операции деления также в терминах умножений и сложений. В большинстве ЦСП используются программные реализации операций извлечения квадратного корня и деления. Для этого, как правило, применяются таблично-аналитические методы [14]. Например, для реализации операции извлечения квадратного корня в прикладных библиотеках СБИС сигнальных контроллеров отечественной серии «Мультикор» [15] требуется 13 умножений и 3 сложения, а для реализации операции деления – 7 умножений и 3 сложения.

Полная вычислительная сложность рассматриваемых алгоритмов в терминах умножений (MUL) и сложений (ADD) приведена на рис. 2 и рис. 3 для случаев $N=2, \dots, 32$. При малых N число арифметических операций, необходимое для выполнения одной итерации алгоритмов, определяется не только квадратичной, но и линейной составляющими вычислительной сложности.

Алгоритм (табл. 2) характеризуется минимальной вычислительной сложностью. Это обусловлено отсутствием в нем операций извлечения квадратного корня. При $N=8$ алгоритм требует 312 комплексных умножений и 165 комплексных сложений на одну итерацию. Поскольку в сигнальных контроллерах серии «Мультикор» арифметические операции однократные, то одна итерация алгоритма занимает $K_T = 4 \cdot 312 + 2 \cdot 312 + 2 \cdot 165$ тактов центрального процессора, в течение которых выполняется эквивалентное число действительных арифметических операций.

СБИС 1892ВМЗТ («Мультикор-12», МС-12) серии «Мультикор» с тактовой частотой центрального процессора 100 МГц и архитектурой SISD (Single Instruction Single Data) может выполнять примерно $100 \cdot 10^6 / K_T \approx 45 \cdot 10^3$ итераций за 1 секунду алгоритма (табл. 2). Если итерации следуют с частотой смены информационных символов, то, например, при модуляции QPSK-4 антенную решетку можно использовать в системах связи со скоростью передачи данных до $2 \cdot 45 \cdot 10^3 = 90$ кбит/с. Здесь множитель 2 означает, что при данной модуляции один символ содержит 2 информационных бита.

При использовании СБИС 1892ВМ2Т («Мультикор-24», МС-24) с тактовой частотой центрального процессора 100 МГц и архитектурой SIMD (Single Instruction Multiple Data) число итераций алгоритма за одну секунду примерно удваивается, что обеспечивает возможность обрабатывать информацию со скоростью в 2 раза большей.

Разрабатываемая в настоящее время СБИС МС-0428 («MultiForce») серии «Мультикор» при выполнении арифметических операций в формате с плавающей точкой характеризуется примерно в 33 раза большей производительностью по сравнению со СБИС 1892ВМ3Т, что в случае $N=8$ позволяет строить цифровую ААР на основе алгоритма (табл. 2) для приема информации на скорости до $33 \cdot 90 \cdot 10^3 \approx 2.97$ Мбит/с.

Приведенные цифры позволяют оценивать максимальную производительность ААР. На практике алгоритмы могут обеспечивать обработку информации на несколько меньшей скорости, так как в приведенных оценках отсутствуют операции, необходимые для организации вычислительной процедуры алгоритма (циклы, передача данных и т.п.). Точные значения предельной производительности можно получить только при реализации алгоритма на рассмотренных микроконтроллерах или других ЦСП.

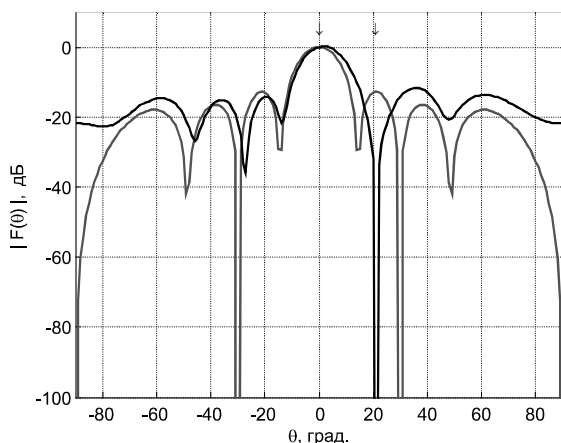


Рис. 4. Результаты моделирования: $\theta_s = 0^\circ$, $\theta_{J1} = 21^\circ$

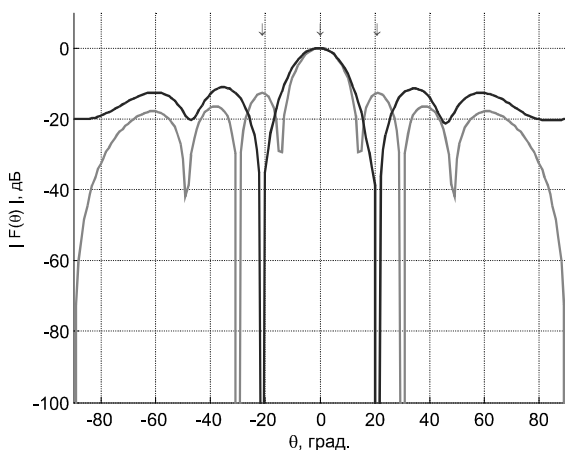


Рис. 5. Результаты моделирования:

$$\theta_s = 0^\circ, \theta_{J1} = 21^\circ, \theta_{J2} = -21^\circ$$

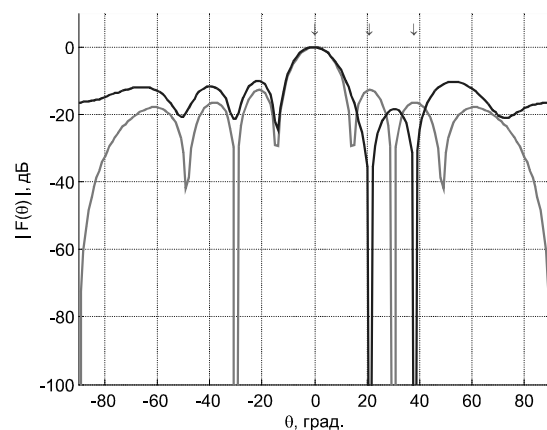


Рис. 6. Результаты моделирования: $\theta_s = 0^\circ$, $\theta_{J1} = 21^\circ$, $\theta_{J2} = 38^\circ$

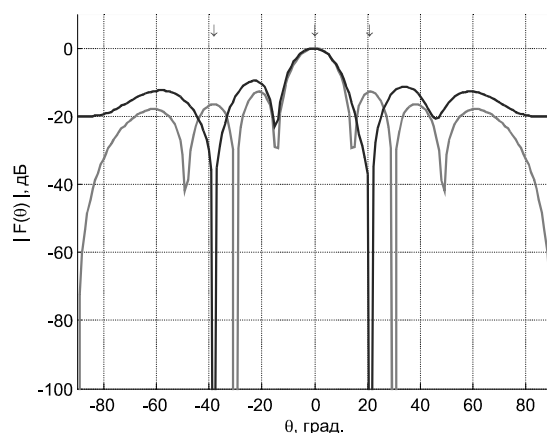


Рис. 7. Результаты моделирования: $\theta_s = 0^\circ$, $\theta_{J1} = 21^\circ$, $\theta_{J2} = -38^\circ$

Моделирование (рис. 4 – рис. 7) подтверждает работоспособность алгоритмов (табл. 1 и табл. 2). Из рисунков следует, что в процессе адаптации в направлении на источник полезного сигнала $\theta_s = 0^\circ$ формируется ДН (темная линия) с уровнем, равным ДН синфазной антенной решетки с ориентацией основного луча в направлении $\theta_s = 0^\circ$ (серая линия), а в направлениях на источники помех $\theta_j = \pm 21^\circ$ и $\theta_j = \pm 38^\circ$ формируются провалы. На рис. 4 – рис. 7 показаны ДН ААР в установленном режиме. Указанные направления на источники помех совпадали с направлениями максимумов первых четырех боковых лепестков ДН синфазной антенной решетки. ДН антенных элементов предполагались все-направленными. Расстояние между элементами решетки равнялось половине длины несущего колебания. Рассматривалась несимметричная антенная решетка.

Моделирование проводилось в информационной полосе частот. В качестве полезного сигнала использовался сигнал QPSK-4 с $|a_i| = 1$. В качестве сигналов источников помех использовался белый шум. Отношения сигнал-помеха для каждой из помех на входе антенного элемента равнялось -20 дБ. При моделировании были использованы следующие параметры адаптивных алгоритмов: $N = 8$, $\delta^2 = 0.01$, $\lambda = 0.9999$. Все вычисления проводились в арифметике с плавающей точкой.

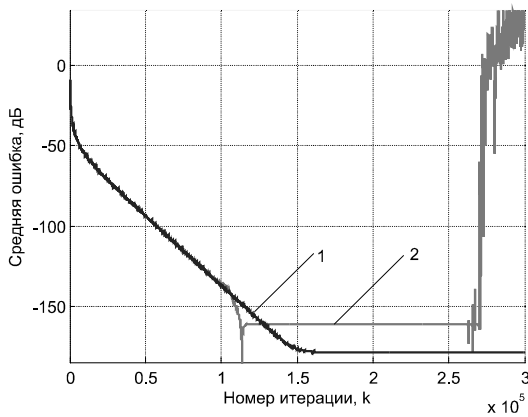


Рис. 8. Переходный процесс: 1 – алгоритмы (табл. 1 и табл. 2), 2 алгоритм [7]

Известно, что RLS-алгоритмы на основе QR-разложения характеризуются повышенной устойчивостью по сравнению с алгоритмами на основе леммы об обращении матриц. Это справедливо и для рассмотренных модификаций алгоритмов, что подтверждается результатами моделирования на рис. 8. Значения, представленные на графиках, определялись путем усреднения параметра $|\alpha_N(k)|$ на скользящем окне в 256 отсчетов. Алгоритмы (табл. 1 и табл. 2) на интервале наблюдения оставались устойчивыми, тогда как алгоритм [7] начинал расходиться после примерно 250 тыс. итераций. Кроме того, в установившемся режиме значение ошибок в алгоритмах (табл. 1 и табл. 2) меньше, чем в алгоритме [7], что обусловлено меньшим числом арифметических операций в алгоритмах (табл. 1 и табл. 2) по сравнению с алгоритмом [7].

Таким образом, рис. 8 подтверждает целесообразность использования алгоритмов (табл. 1 и табл. 2) для управления ААР систем цифровой связи с модуляцией, удовлетворяющей СМ-критерию, а рис. 4 – рис. 7 подтверждают эффективность этих алгоритмов. Материалы статьи могут представлять интерес для специалистов в области ААР и разработчиков систем беспроводной связи.

Литература

1. Монзинго Р.А., Миллер Т.У. Адаптивные антенные решетки. Введение в теорию. Пер. с англ. Челпанова В.Г., Лексаченко В.А. – М.: Радио и связь, 1986. – 448 с.
2. Godara L.C. Application of antenna arrays to mobile communications. II. Beam-forming and direction-of-arrival considera-

- tions // Proceedings of the IEEE. – 1997. – Vol. 85. – №8. – P. 1195–1245.
3. Джиган В.И. Прикладная библиотека адаптивных алгоритмов // Электроника: Наука, Технологии, Бизнес. – 2006. – №1. – С. 60–65.
4. Frost O.L. An algorithm for linearly constrained adaptive array processing // Proceedings of the IEEE. – 1972. – Vol. 60. – №8. – P. 926–935.
5. Treichler J., Larimore M. New processing techniques based on the constant modulus adaptive algorithm // IEEE Trans. Acoustics, Speech, and Signal Processing. – 1985. – Vol. 33. – №2. – P. 420–431.
6. Ohgane T., Shimura T., Matsuzawa N., Sasaoka H. An implementation of a CMA adaptive array for high speed GMSK transmission in mobile communications // IEEE Trans. Vehicular Technology. – 1993. – Vol. 42. – №3. – P. 282–288.
7. Chen Y., Le-Ngoc T., Champagne B., Xu C. Recursive least squares constant modulus algorithm for blind adaptive array // IEEE Trans. Signal Processing. – 2004. – Vol. 52. – №5. – P. 1452–1456.
8. Giordano A.A., Hsu F.M. Least square estimation with application to digital signal processing. – Canada, Toronto: John Wiley and Sons, Inc., 1985. – 412 p.
9. Джиган В.И. Многоканальные RLS- и быстрые RLS-алгоритмы адаптивной фильтрации // Успехи современной радиоэлектроники. – 2004. – №11. – С. 48–77.
10. Bellanger M. Which constant modulus criterion is better for blind adaptive filtering: CM(1,2) or CM(2,2)? // Proceedings of the IEEE International Conf. on Acoustics, Speech, and Signal Processing. – 2005. – Vol. 4. – P. IV/29–IV/32.
11. Джиган В.И. Условия эквивалентности рекурсивных алгоритмов адаптивной фильтрации по критерию наименьших квадратов // Телекоммуникации. – 2006. – №6. – С. 6–11.
12. Голуб Дж., Ван Лоун Ч. Матричные вычисления. Пер. с англ. под ред. Воеводина В.В. – М.: Мир, 1999. – 548 с.
13. Hsieh S.F., Liu K.J.R., Yao K. A unified square-root-free approach for QRD based recursive least squares estimation // IEEE Trans. Signal Processing. – 1993. – Vol. 41. – №3. – P. 1405–1409.
14. Солохина Т., Александров Ю., Петричкович Я. Сигнальные контроллеры компании «ЭЛВИС»: первая линейка отечественных DSP // Электроника: Наука, Технология, Бизнес. – 2005. – №7. – С. 70–77.
15. Wang L.-K., Schulte M.J. Decimal floating-point division using Newton-Raphson iteration // Proceedings of the 15-th IEEE International Conference on Application-Specific Systems, Architectures and Processors. – 2004. – P. 84–95.

СИНТЕЗ ПОЛИНОМИАЛЬНЫХ ФИЛЬТРОВ ИМПУЛЬСНЫХ ПОМЕХ, ПОРАЖАЮЩИХ РЕЧЕВЫЕ СИГНАЛЫ

Соловьева Е.Б., Жеребцов А.П.

Введение

В областях электротехники, радиотехники, электроники, электромагнитной совместимости нередко решается задача фильтрации импульсных помех. Импульсные помехи возникают при коммутации различных электротехнических и электронных устройств, при механических повреждениях поверхности устройств хранения информации, при работе двигателей внутреннего сгорания, под влиянием различных атмосферных явлений и т.д. Для улучшения качества восстановления и распознаваемости сигналов применяются методы борьбы с импульсными помехами.

Классическим методом подавления импульсных помех является медианная фильтрация [1, 2, 3]. Известно, что медианные фильтры (МФ) обладают недостатком: они вносят существенные искажения на некоторых участках сигнала, не пораженных импульсной помехой. МФ считаются неоптимальными, поскольку не используют информацию о статистических свойствах сигналов и помех. В результате разработка методов фильтрации импульсных помех, обеспечивающих высокое качество восстановления сигналов, является актуальной задачей.

В статье предлагается синтезировать полиномиальные фильтры импульсных помех во временной и частотной областях на основе метода расщепления сигналов [4, 5]. Этот метод обладает важными достоинствами:

- статистические свойства сигналов и помех учитываются автоматически в процессе синтеза (“обучения”) фильтра;
- по сравнению с математическим аппаратом функциональных рядов и полиномов Вольтерры [6, 7] метод расщепления дает более простую (адаптированную к заданному классу входных сигналов) полиномиальную модель фильтра;
- в отличие от ряда Вольтерры многочлен расщепленных сигналов свободен от проблемы сходимости, что позволяет на его основе синтезировать существенно нелинейные устройства (при этом точность обработки сигналов повышается с увеличением степени полиномиальной модели);
- многочлен расщепленных сигналов содержит линейно входящие параметры, поэтому параметры модели определяются как глобально оптимальное решение задачи аппроксимации оператора устройства в равномерной и среднеквадратичной метриках [4].

В данной работе метод расщепления сигналов применяется для синтеза нерекурсивных цифровых фильтров импульсных помех; сравниваются временной и частотный подходы к построению нелинейных операторов

Рассмотрены методы синтеза нелинейных нерекурсивных цифровых фильтров импульсных помех на основе теории расщепления сигналов во временной и частотной областях. Исследовано применение таких фильтров для обработки речевых сигналов, пораженных импульсными помехами. Выполнено сравнение результатов обработки сигналов предложенными фильтрами, медианными фильтрами и фильтром Вольтерры.

цифровых фильтров; оценивается качество фильтрации, выполненной предложенными способами, медианным фильтром и фильтром Вольтерры.

Синтез нелинейных нерекурсивных разностных уравнений методом расщепления

Задача синтеза цифровых фильтров импульсных помех решается в рамках принципа “черного ящика” на основе теории расщепления сигналов. Согласно этой теории оператор F_s нелинейного фильтра описывается композицией двух операторов: оператора F_p расщепителя и оператора P нелинейного безынерционного преобразователя [4, 5].

Оператор-расщепитель F_p отображает скалярный сигнал $x(n, \vec{a})$, $n \in I_n$, $\vec{a} \in G_a$ в векторный $\vec{x}_p(n, \vec{a})$,

$$\vec{x}_p(n, \vec{a}) = F_p[x(n, \vec{a})] = [x_{p1}(n, \vec{a}), x_{p2}(n, \vec{a}), \dots, x_{pm}(n, \vec{a})],$$

где

$$x_{p1}(n, \vec{a}) = F_{p1}[x(n, \vec{a})],$$

$$x_{p2}(n, \vec{a}) = F_{p2}[x(n, \vec{a})], \dots,$$

$x_{pm}(n, \vec{a}) = F_{pm}[x(n, \vec{a})]$, I_n – длительность финитного или периодического входного сигнала $x(n, \vec{a})$, n – дискретная нормированная переменная времени, $\vec{a} = [a_1, a_2, \dots, a_j]$ – вектор параметров из множества G_a сигнала $x(n, \vec{a})$, при выполнении следующих условий:

- векторные сигналы не исчезают, т. е.

$$\vec{x}_p(n, \vec{a}) \neq 0$$
для всех $n \in I_n$, $\vec{a} \in G_a$;
- фазовые портреты расщепленных сигналов не пересекаются, не касаются и не самопересекаются, т. е. при любых $\vec{a}_1 \neq \vec{a}_2$, $\vec{a}_1 \in G_a$, $\vec{a}_2 \in G_a$, $n_1 \neq n_2$, $n_1 \in I_n$, $n_2 \in I_n$ имеет место неравенство

$$\vec{x}_p(n_1, \vec{a}_1) \neq \vec{x}_p(n_2, \vec{a}_2).$$

Расщепление реализуют линейные, нелинейные, стационарные и нестационарные преобразователи сигналов [4].

Оператор P нелинейного безынерционного преобразователя отображает векторный сигнал $\vec{x}_p(n, \vec{a})$ в скалярный сигнал $y(n, \vec{a})$. Обычно такие операторы описываются многомерными многочленами

$$y(n, \vec{a}) = P[\vec{x}_p(n, \vec{a})] = \sum_{j_1=0}^{J_1} \sum_{j_2=0}^{J_2} \dots \sum_{j_m=0}^{J_m} C_{j_1, j_2, \dots, j_m} [x_{p1}(n, \vec{a})]^{j_1} \times \dots \times [x_{p2}(n, \vec{a})]^{j_2} \dots [x_{pm}(n, \vec{a})]^{j_m}, \quad (1)$$

хотя возможны и другие формы представления [4].

Многомерный полином (1) степени p ($p = J_1 + J_2 + \dots + J_m$) для всех $n \in I_n$, $\vec{a} \in G_a$ удовлетворяет условию

$$\|y^o(n, \vec{a}) - y(n, \vec{a})\| \leq \varepsilon,$$

где ε ($\varepsilon > 0$) – заданная погрешность аппроксимации оператора F_s нелинейного идеального фильтра, $y^o(n, \vec{a})$ – выходной сигнал идеального фильтра,

$$y^o(n, \vec{a}) = F_s[x(n, \vec{a})].$$

Функциональная схема нелинейного нерекурсивного цифрового фильтра (ННЦФ), реализующая преобразование (1), изображена на рис. 1.

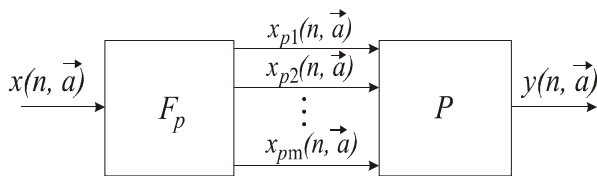


Рис. 1. Функциональная схема нелинейного цифрового фильтра.

Следует отметить, что количество коэффициентов $C_{j_1, j_2, \dots, j_m}$ в многочлене (1) равно

$$\sum_{i=1}^p R_{m+i-1}^i = \sum_{i=1}^p \frac{(m+i-1)!}{i!(m-1)!},$$

где R_{m+i-1}^i – число сочетаний из $(m+i-1)$ элементов по i элементов. При большом числе m каналов расщепления в многочлене (1) задача аппроксимации оператора F_s имеет высокую размерность, порождает плохую обусловленность ее решения и требует существенных вычислительных затрат.

Синтезируем ННЦФ, выделяющий речевой сигнал из смеси сигнала с импульсной помехой на основе метода расщепления.

Фильтрация импульсных помех речевых сигналов во временной области

Речевой сигнал, используемый для „обучения“ ННЦФ в рассматриваемом примере, имел длительность 35 секунд (280 000 отсчетов) и частоту дискретизации 8 кГц. Он состоял из разных фраз четырех дикторов (двух мужчин и двух женщин). Фразы отличались уровнями громкости, задаваемыми как 1, 0.75, 0.5, 0.25 (пропорционально указанным значениям нормировались мгновенные значения каждой дикторской речи, причем уровню 1 соответствовал диапазон речевого сигнала $[-0.5, 0.5]$).

Для исследования свойств синтезированного фильтра применялся речевой сигнал длительностью 20 секунд (160 000 отсчетов), отличающийся от „обучаемого“ и содержащий разные фразы мужской и женской дикторской речи с уровнями 0.8 и 0.4 соответственно.

Значения импульсной помехи формировались как случайные числа, распределенные равномерно в диапазоне $[-0.5, 0.5]$. Моменты появления помехи выбирались согласно следующему правилу [1, 2]. Если в момент времени n генератор случайных чисел с равномерным законом распределения в диапазоне $[0, 1]$ дает число меньшее, заданного порога α (в нашем случае $\alpha = 0.01$), то в этот момент времени действует импульсная помеха, в противном случае она отсутствует. Таким образом, вероятность появления помехи в текущий момент времени n равна α ; вероятность ее отсутствия – $(1-\alpha)$; вероятность того, что импульсная помеха появится через η временных отсчетов, равна $\alpha \cdot (1-\alpha)^{\eta-1}$ (геометрическая функция распределения для переменной η). Действовало также дополнительное ограничение: расстояние между соседними помехами – не менее 5 отсчетов речевого сигнала.

Расщепитель ННЦФ следует строить с минимальным числом каналов расщепления, поскольку в этом случае упрощается оператор синтезируемого устройства [4]. Исследования показали, что данному свойству удовлетворяет вектор расщепленных сигналов

$$\vec{x}_p(n, \vec{a}) = [x(n, \vec{a}), x(n-1, \vec{a}), x(n-2, \vec{a}), x(n-3, \vec{a}), x(n-4, \vec{a})].$$

Применение МФ [1, 2] и двусторонних предсказателей [3], использующих для получения выходного сигнала в текущий момент времени n предыдущие и последующие отсчеты воздействия, породило гипотезу о возможном учете данного свойства при расщеплении искаженного импульсной помехой речевого сигнала. Результаты проверки гипотезы представлены на рис. 2.

На рис. 2 изображены зависимости среднеквадратичной погрешности $\varepsilon(\xi)$, вычисленной по формуле

$$\varepsilon(\xi) = \frac{1}{Q-3} \sqrt{\sum_{n=4}^Q (y^o(n, \vec{a}) - y(n, \xi, \vec{a}))^2}, \quad (2)$$

где $y^o(n, \vec{a})$ – желаемая реакция фильтра (неискаженный речевой сигнал) длиной Q ($Q = 160\,000$) отсчетов, $y(n, \xi, \vec{a})$ – выходной сигнал модели (1) нелинейного фильтра при векторе расщепленных сигналов

$$\vec{x}_p(n, \xi, \vec{a}) = [x(n, \xi, \vec{a}), x(n+\xi-1, \vec{a}), x(n+\xi-2, \vec{a}), x(n+\xi-3, \vec{a}), x(n+\xi-4, \vec{a})],$$

ξ – количество последующих отсчетов относительно момента времени n в векторах расщепленных сигналов, сформированных согласно рис. 3. Номера кривых на рис. 2 соответствуют степени p полиномиальной модели (1).

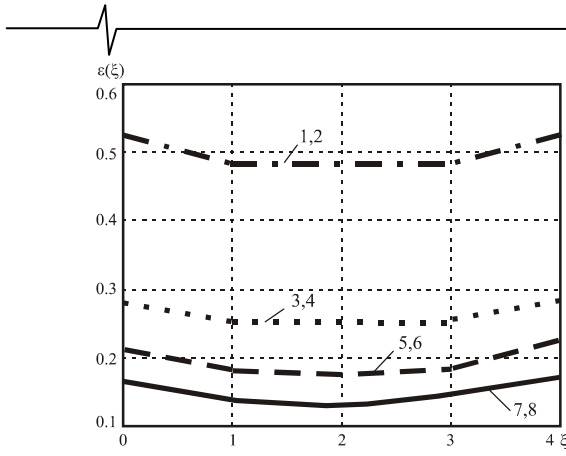


Рис. 2. Зависимости среднеквадратичной погрешности от число отсчетов

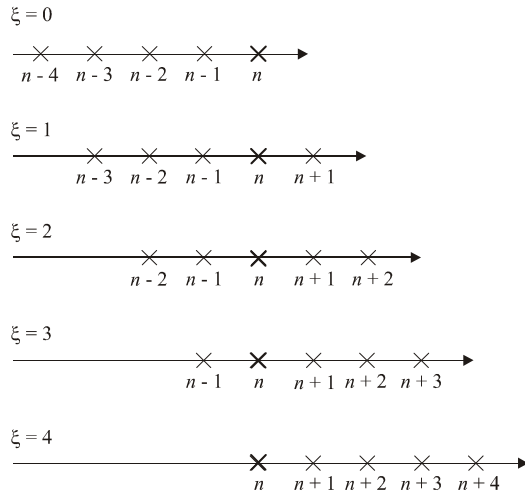


Рис. 3. Формирование векторов расщепленных сигналов

Анализ рис. 2 показывает, что

– способ формирования расщепленных сигналов влияет на точность фильтрации. Среднеквадратичная погрешность фильтрации минимальна при векторе расщепленных сигналов, содержащем одинаковое количество предыдущих и последующих отсчетов ($\xi = 2$) по отношению к текущему моменту времени n ;

– слагаемые четных степеней многочлена расщепленных сигналов не влияют на среднеквадратичную погрешность фильтрации, поэтому их можно исключить из модели (1).

В результате модель ННЦФ, подавляющего импульсные помехи в речевых сигналах, имеет вид

$$\begin{aligned} y(n, 2, \vec{a}) &= y(n, \vec{a}) = \\ &= \sum_{r=2k-1}^p \sum_{j_1=0}^{J_1} \sum_{j_2=0}^{J_2} \dots \sum_{j_5=0}^{J_5} C_{j_1, j_2, \dots, j_5} x^{j_1}(n-2, \vec{a}) x^{j_2}(n-1, \vec{a}) x^{j_3}(n, \vec{a}) \times \\ &\times x^{j_4}(n+1, \vec{a}) x^{j_5}(n+2, \vec{a}), \end{aligned} \quad (3)$$

$$j_1 + j_2 + \dots + j_5 = r, \quad k = 1, 2, \dots; \quad J_1 + J_2 + \dots + J_5 = p.$$

Известно, что полином хорошо моделирует гладкие функции [9]. Применим это свойство для операторов, реализуя его на каскадном соединении МФ и ННЦФ. Размер апертюры МФ равен 3 отсчетам (исследования показали, что такой МФ вносит наименьшие искажения в обрабатываемый сигнал). Устройство в виде каскадного

соединения МФ и ННЦФ будем называть комбинированным ННЦФ (КННЦФ).

Среднеквадратичные погрешности восстановления речевого сигнала, вычисленные по формуле (2) ($\xi = 2$) при фильтрациях ННЦФ и КННЦФ, указаны в табл. 1.

Таблица 1
Среднеквадратичная погрешность фильтрации ННЦФ и КННЦФ

Степень p	ННЦФ	КННЦФ
3	0.2501	0.2379
5	0.1757	0.2387
7	0.1297	0.2415

Из табл. 1 видно, что “сглаживание” искаженных речевых сигналов, выполненное медианным фильтром в КННЦФ, снижает точность фильтрации. Речь на выходе КННЦФ становится глухой с неестественным звучанием. Таким образом, синтез КННЦФ в виде каскадного соединения МФ и ННЦФ нецелесообразен.

Сравним результаты подавления импульсных помех ННЦФ, МФ и фильтром Вольтерры [8], модель которого имеет вид

$$\begin{aligned} y(n) &= \\ &= \sum_{r=2k-1}^p \sum_{j_1=0}^{J_1} \sum_{j_2=0}^{J_2} \dots \sum_{j_5=0}^{J_5} h_{j_1, j_2, \dots, j_5} x^{j_1}(n) x^{j_2}(n-1) x^{j_3}(n-2) x^{j_4}(n-3) x^{j_5}(n-4), \quad (4) \\ j_1 + j_2 + \dots + j_5 &= r, \quad k = 1, 2, \dots; \quad J_1 + J_2 + \dots + J_5 = p. \end{aligned}$$

Зависимости среднеквадратичной погрешности $\varepsilon(p)$ от степени полиномиальных моделей представлены на рис. 4. Кривые 1, 2, 3 получены медианной фильтрацией с апертюрами длиной 3, 5, 7 соответственно, кривая 4 – фильтрацией Вольтерры (4), кривая 5 – ННЦФ с моделью (3).

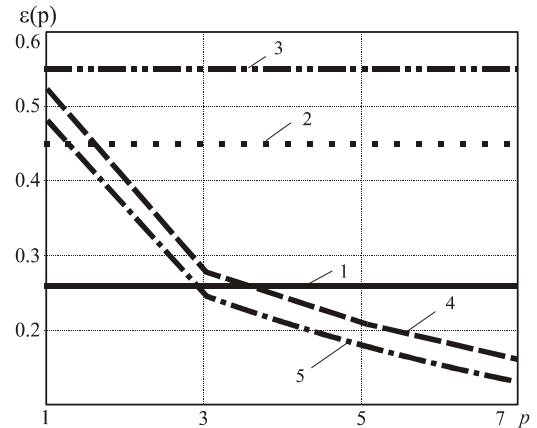


Рис. 4. Зависимости среднеквадратичной погрешности от степени полиномиальных моделей.

Анализ рис. 4 показывает, что фильтр, синтезированный методом расщепления, дает более точное восстановление сигналов по сравнению с его аналогами. С увеличением степени p полиномиальной модели погрешность фильтрации уменьшается.

При прослушивании выходных сигналов нелинейных фильтров наблюдаются следующие результаты. Речь на выходе МФ, хотя и не содержит импульсных помех, ста-

новится глухой и обладает неестественным звучанием. После фильтрации Вольтерры и обработки ННЦФ речевые сигналы имеют естественное звучание, в них присутствуют редкие остаточные импульсные помехи малой амплитуды, воспринимаемые как слабый шорох или потрескивание. На выходе ННЦФ остаточных помех значительно меньше, чем на выходе фильтра Вольтерры.

Для полиномиальных моделей высокой степени экспериментально замечено появление плохой обусловленности решения задачи аппроксимации нелинейного оператора фильтра. Указанная проблема частично снимается при синтезе фильтра в частотной области с использованием спектров входных и выходных сигналов [10].

Синтез полиномиальных фильтров спектральным методом расщепления

В спектральном методе расщепления входными и выходными сигналами фильтра являются дискретные преобразования Фурье (ДПФ) фрагментов реализаций соответствующих входных и выходных сигналов. Выбор длины (числа отсчетов) фрагментов связан с их расщеплением на множестве сигналов. Если фрагменты различны (расщеплены), то их векторы-спектры также различны и, следовательно, расщеплены.

Из расщепленных векторов-спектров входных сигналов можно конструировать векторы-спектры желаемых выходных сигналов. При этом используется теорема об аппроксимации нелинейных операторов в комплексной области [4]. Согласно данной теореме аппроксимирующий многочлен строится с учетом всех (комплексных и комплексно-сопряженных) составляющих векторов-спектров входных сигналов.

Так, если сигнал состоит из $N = 2M$ отсчетов (именно этот случай для определенности будем рассматривать), то вектор расщепленных сигналов

$$\begin{aligned} \bar{X}_p(\bar{a}) &= F_p[X(\bar{a})] = \\ &= [X(-(M-1), \bar{a}), X(-(M-2), \bar{a}), \dots, X(M, \bar{a})], \\ \bar{a} &\in G_a, \end{aligned}$$

включает отсчеты спектра $X(\bar{a})$, а аппроксимирующий многочлен имеет вид

$$\begin{aligned} P_k[\bar{X}_p(\bar{a})] &= Y(k, \bar{a}) = \\ &= \sum_{j_1}^{J_1} \sum_{j_2}^{J_2} \dots \sum_{j_m}^{J_m} C_{j_1, j_2, \dots, j_m}^{(k)} [X(-(M-1), \bar{a})]^{j_1} [X(-(M-2), \bar{a})]^{j_2} \dots \\ &\dots [X(M, \bar{a})]^{j_m}, \end{aligned} \quad (5)$$

где $k \in [-(M-1), M]$, m – число каналов расщепления.

Согласно теории расщепления, для $k \in [-(M-1), M]$, $\bar{a} \in G_a$ и $\varepsilon > 0$ существует полином степени $P(p = J_1 + J_2 + \dots + J_m)$ такой, что

$$\|Y^o(k, \bar{a}) - Y(k, \bar{a})\| \leq \varepsilon, \quad (6)$$

где $Y^o(k, \bar{a})$ – k -я составляющая дискретного спектра $[Y^o(-(M-1), \bar{a}), Y^o(-(M-2), \bar{a}), \dots, Y^o(M, \bar{a})]$ желаемого выходного сигнала $y^o(n, \bar{a})$ фильтра.

Представим полином (5) в удобном для формирования степенных составляющих многомерного многочлена виде

$$Y(k, \bar{a}) = \sum_{\ell=1}^p \sum_{i_1=-(M-1)}^M \sum_{i_2=i_1}^M \dots \sum_{i_\ell=i_{\ell-1}}^M B_{i_1, i_2, \dots, i_\ell}^{(k)} \prod_{r=1}^{\ell} X(i_r), \quad (7)$$

где коэффициенты

$$B_{i_1, i_2, \dots, i_\ell}^{(k)} = C_{\sum_{\alpha=1}^{\ell} \delta(q - (i_\alpha + (M-1)))}^{(k)}, \quad q \in [0, m-1], \quad \delta(q) -$$

цифровой единичный импульс. Таким образом, нижний индекс коэффициента $C^{(k)}$ формируется последовательностью целых чисел при сложении единичных функций вида $\delta(q - (i_\alpha + (M-1)))$. Такая последовательность состоит из m отсчетов. Отсчет равен 1 в точке $q - (i_\alpha + (M-1)) = 0$. В остальных $m-1$ точках отсчеты равны 0.

Согласно известному свойству нелинейного преобразования [11], спектральная составляющая $Y(k, \bar{a})$ на частоте k выходного сигнала формируется с учетом слагаемых, содержащих k -ю спектральную составляющую воздействия и произведения спектральных составляющих входного сигнала на частотах, удовлетворяющих условию $k = i_1 + i_2 + \dots + i_\ell$. Применение данного условия позволяет существенно уменьшить число членов многомерной суммы в выражении (7).

На основе указанного свойства нелинейного преобразования, а также свойства симметрии [12] $Y(-k, \bar{a}) = Y^*(k, \bar{a})$, $0 < k \leq (M-1)$, где $*$ – знак комплексного сопряжения, сформируем $M+1$ уравнений

$$Y(k, \bar{a}) = \sum_{\ell=1}^p \sum_{i_1=-(M-1)}^M \sum_{i_2=i_1}^M \dots \sum_{i_\ell=i_{\ell-1}}^M B_{i_1, i_2, \dots, i_\ell}^{(k)} \prod_{r=1}^{\ell} X(i_r), \quad (8)$$

$i_1 + i_2 + \dots + i_\ell = k$

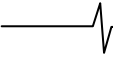
где k меняется от 0 до M .

Известно, что свойство периодичности ДПФ приводит к появлению областей наложения в спектрах степенных составляющих выходного сигнала [10, 12]. Для учета эффекта наложения спектров (УНС) введем дополнительные члены в многомерную сумму (8). В итоге полиномиальная модель нелинейного оператора фильтра для $k \in [0, M]$ имеет вид

$$\begin{aligned} Y(k, \bar{a}) &= \sum_{\ell=1}^p \sum_{i_1=-(M-1)}^M \sum_{i_2=i_1}^M \dots \sum_{i_\ell=i_{\ell-1}}^M B_{i_1, i_2, \dots, i_\ell}^{(k)} \prod_{r=1}^{\ell} X(i_r) + \\ &\quad + \sum_{\gamma=1}^{[\ell/2]} \sum_{i_1=-(M-1)}^M \sum_{i_2=i_1}^M \dots \sum_{i_\ell=i_{\ell-1}}^M B_{i_1, i_2, \dots, i_\ell}^{(k)} \prod_{r=1}^{\ell} X^*(i_r) + \\ &\quad + \left(\sum_{\gamma=1}^{[(\ell-1)/2]} \sum_{i_1=-(M-1)}^M \sum_{i_2=i_1}^M \dots \sum_{i_\ell=i_{\ell-1}}^M B_{i_1, i_2, \dots, i_\ell}^{(k)} \prod_{r=1}^{\ell} X(i_r) \right) \delta_1(\ell-3), \end{aligned} \quad (9)$$

$i_1 + i_2 + \dots + i_\ell = \gamma N - k$

где $[s]$ – наименьшее целое, такое что $[s] \leq s$, $\delta_1(\ell)$ – цифровой единичный скачок.



Если высокочастотные составляющие спектра входных сигналов фильтра малы, то эффект наложения спектров проявляется слабо. В этом случае при вычислении операторов $P_k[\bar{X}_p(\bar{a})]$, $k \in [0, M]$ указанный эффект можно не учитывать, либо учитывать не полностью (уменьшить число членов в многомерных суммах выражения (9) путем понижения верхнего предела переменной γ).

Параметры полинома (9) при $k \in [0, M]$ определяются из полученного при решении аппроксимационной задачи (6) в метрике L_2 выражения [13]

$$\vec{B}^{(k)} = \left\{ E \left[\left[\bar{X}^{(k)} \right]^* \left[\bar{X}^{(k)} \right]^t \right] \right\}^{-1} E \left[\left[\bar{X}^{(k)} \right]^* Y(k, \bar{a}) \right],$$

где $E \left[\left[\bar{X}^{(k)} \right]^* \left[\bar{X}^{(k)} \right]^t \right]$ – эрмитова матрица, состоящая

из спектральных моментов различных порядков, $E[\bullet]$ – оператор математического ожидания, $\vec{B}^{(k)}$ – вектор-столбец оценок параметров оператора $P_k[\bar{X}_p(\bar{a})]$,

$$\left[\bar{X}^{(k)} \right]^t = \begin{bmatrix} \left[\bar{X}^{(k)} \right]^t \Big|_{\sum_{\alpha=1}^{\ell} i_{\alpha}=k} & \left[\bar{X}^{(k)} \right]^t \Big|_{\substack{\sum_{\alpha=1}^{\ell} i_{\alpha}=\gamma N-k, \\ \gamma \in [1; \lfloor \ell/2 \rfloor]}} & \left[\bar{X}^{(k)} \right]^t \Big|_{\substack{\sum_{\alpha=1}^{\ell} i_{\alpha}=\gamma N+k, \\ \gamma \in [1; \lfloor (\ell-1)/2 \rfloor]}} \end{bmatrix},$$

$$\ell \in [1; p]$$

Здесь внутренние векторы-элементы содержат произведения спектральных составляющих воздействия на частотах, удовлетворяющих указанным условиям. Число сомножителей в произведении определяется значением ℓ . Длины векторов $\bar{X}^{(k)}$ и $\vec{B}^{(k)}$ зависят от того, какие слагаемые содержатся в выражении (9).

Функциональная схема нелинейного фильтра, синтезированного в частотной области, изображена рис. 5. На этом рисунке блок ДПФ(N) осуществляет ДПФ фрагментов входного сигнала длиной N ; блоки $P_0, P_1, \dots, P_{M-1}, P_M$ – нелинейные безынерционные преобразователи, вычисляющие отдельные спектральные составляющие выходного сигнала; БКС – блок комплексного сопряжения, формирующий сопряженные спектральные компоненты; ОДПФ(N) – блок обратного преобразования Фурье.

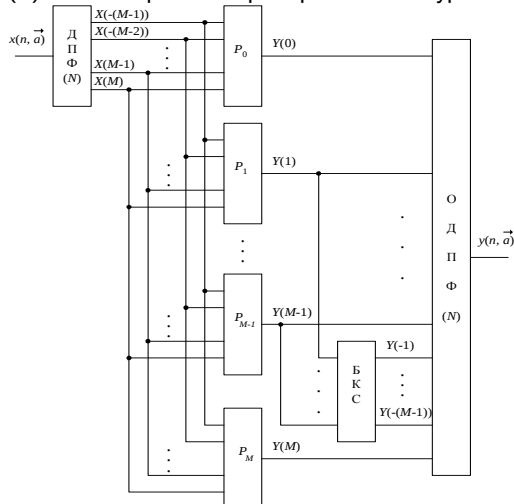


Рис. 5. Функциональная схема нелинейного фильтра, синтезированного в частотной области

Поскольку синтез фильтра выполняется в частотной области, синтезированное устройство будем называть частотным нелинейным нерекурсивным цифровым фильтром (ЧННЦФ).

В табл. 2 указано общее количество параметров многомерных многочленов $P_k[\bar{X}_p(\bar{a})]$, $k \in [-(M-1), M]$ различных степеней при использовании описанных ранее способов их сокращения. В качестве примера число каналов расщепления $m = 6$.

Из табл. 2 видно, что при синтезе ЧННЦФ спектральным методом расщепления размерность аппроксимационной задачи (6) существенно сокращается.

Таким образом, синтез полиномиальных фильтров в частотной области характеризуется разбиением общей задачи аппроксимации нелинейного оператора фильтра, сформулированной во временной области, на несколько аппроксимационных задач меньшей размерности, решаемых в частотной области. При таком подходе удастся практически снять проблему плохой обусловленности решения задачи аппроксимации высокой размерности.

Частотная фильтрация импульсных помех речевых сигналов.

Применим спектральный метод расщепления для синтеза фильтра импульсных помех на том же классе сигналов, на котором проводилась фильтрация во временной области. Согласно методу расщепления, временная и частотная модели нелинейного фильтра взаимосвязаны, поэтому синтез ЧННЦФ выполним с учетом свойств модели (3) ННЦФ.

Отметим, что спектральное преобразование упрощается, если ДПФ подвергаются фрагменты с четным количеством отсчетов [12]. Поэтому при синтезе ЧННЦФ используем расщепленные фрагменты длиной $N = 6$ (а не $N = 5$, как в модели (3)).

При частотной фильтрации выходной сигнал фильтра можно сформировать последовательным и блочным способами. Последовательный способ предполагает, что из N отсчетов, полученных ОДПФ (рис. 5), в выходном сигнале фильтра остается лишь один отсчет, соответствующий моменту времени n . При этом фрагменты воздействия формируются путем смещения окна длиной $N = 6$ вдоль входного сигнала с шагом в 1 отсчет. Блочный способ предполагает, что выходной сигнал ЧННЦФ содержит весь фрагмент, полученный ОДПФ. В этом случае фрагменты воздействия составлены при смещении окна длиной $N = 6$ вдоль входного сигнала с шагом в N отсчетов.

В табл. 3 указаны значения среднеквадратичной погрешности ε при блочном и последовательном способах формирования выходного сигнала ЧННЦФ 5-й степени. Модель ЧННЦФ строилась с учетом эффекта наложения спектров. Значения ξ при $N = 6$ задавались аналогично рис. 3.

Таблица 3

Среднеквадратичная погрешность ε при разных способах формирования выходного сигнала ЧННЦФ 5-й степени

Блочный способ	Последовательный способ					
	$\xi = 0$	$\xi = 1$	$\xi = 2$	$\xi = 3$	$\xi = 4$	$\xi = 5$
	0.2185	0.2502	0.2014	0.1982	0.1961	0.2000
						0.2426

Таблица 2

Число параметров нелинейных операторов, формирующих спектр выходного сигнала

Степень p	Число параметров		
	на основе полинома (7)	при использовании свойства нелинейного преобразования	
		без УНС (полином (8))	с УНС (полином (9))
1	24	4	4
2	108	16	17
3	332	40	55
4	836	85	154
5	1844	162	379

Из табл. 3 видно, что более точная фильтрация выполняется при последовательном способе формирования выходного сигнала ЧННЦФ, когда вектор расщепленных сигналов составлен из приблизительно равного числа предыдущих и последующих отсчетов по отношению к n -му отсчету выходного сигнала фильтра.

В дальнейших расчетах расщепленные фрагменты входного сигнала ЧННЦФ составлялись при $\xi = 2$, как показано на рис. 6.

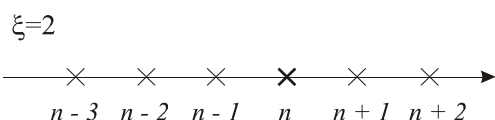


Рис. 6. Расщепленный фрагмент входного сигнала при $\xi = 2$.

Влияние эффекта наложения спектров на точность фильтрации при разной степени ЧННЦФ отражают данные, представленные в табл. 4.

Таблица 4

Погрешность ε при фильтрации ЧННЦФ с УНС и без УНС

Степень p	без УНС	с УНС
3	0.2775	0.2609
5	0.2299	0.1982
7	0.1923	0.1544

Анализ табл. 4 показывает, что учет эффекта наложения спектров в модели (9) ЧННЦФ повышает точность фильтрации импульсных помех речевых сигналов.

На рис. 7 представлены зависимости $\varepsilon(p)$, полученные МФ с апертурой длиной 3 отсчета (наиболее точный вариант МФ согласно рис. 4) (кривая 1), фильтром Вольтерры (4) (кривая 2) и ЧННЦФ с УНС (кривая 3).

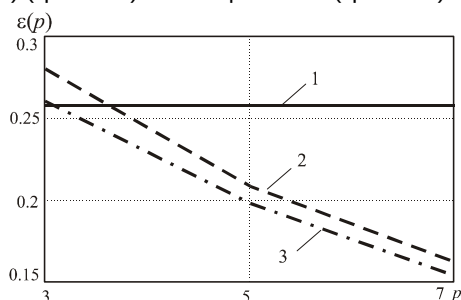


Рис. 7. Зависимости среднеквадратичной погрешности, полученные МФ, фильтром Вольтерры и ЧННЦФ с УНС.

Наглядное представление результатов фильтрации дают рис. 8, 9. На рис. 8, а, 9, а изображены огибающие фрагментов речевого сигнала с импульсными помехами. На рис. 8, б, 9, б показаны неискаженный сигнал (кривая 1) и результаты его обработки МФ длиной 3 (кривая 2), фильтром Вольтерры 5-й степени (кривая 3) и ЧННЦФ 5-й степени с учетом эффекта наложения (кривая 4).

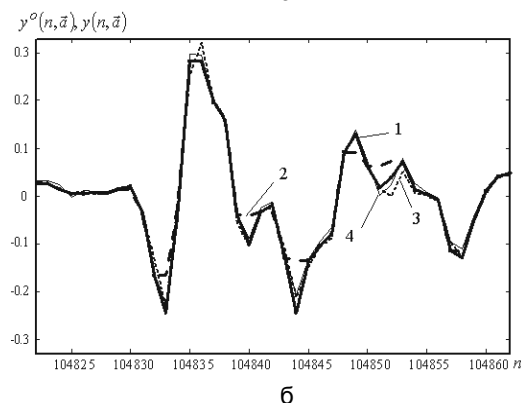
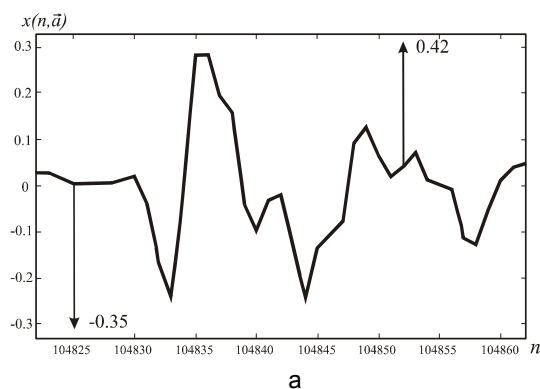
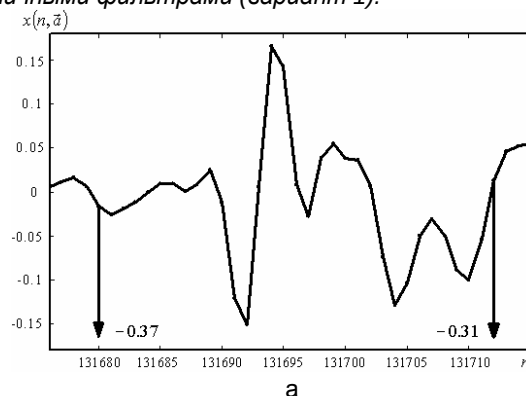


Рис. 8. Огибающие фрагментов речевого сигнала с импульсными помехами и результаты его обработки различными фильтрами (вариант 1).



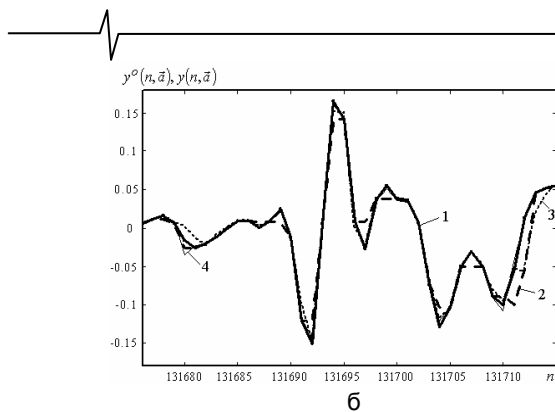


Рис. 9. Огибающие фрагментов речевого сигнала с импульсными помехами и результаты его обработки различными фильтрами (вариант 2).

Из рис. 7, 8, 9 видно, что ЧННЦФ обеспечивает более высокую точность фильтрации по сравнению с его аналогами.

Заключение

Рассмотренный метод синтеза фильтров импульсных помех основан на принципе “обучения с учителем”. Параметры модели фильтра определяются в результате решения задачи аппроксимации нелинейного оператора фильтра по соотношению вход-выход. В процессе “обучения” фильтра статистические свойства сигналов и помех учитываются автоматически. Поскольку модель в форме многочлена расщепленных сигналов линейна относительно своих параметров, вычисленные при синтезе фильтра параметры являются глобально оптимальными в равномерной и среднеквадратичной метриках.

Замечательным свойством фильтров, синтезированных методом расщепления, является их инвариантность к входным сигналам при условии, что статистические свойства сигналов и помех сохраняются. Таким образом, фильтр, синтезированный на испытательных сигналах, удовлетворительно работает на других сигналах и помехах с близкими статистическими характеристиками. Для повышения точности фильтрации следует увеличить степень полиномиальной модели фильтра.

При синтезе ННЦФ импульсных помех речевых сигналов установлено, что

- моделью ННЦФ является многомерный полином нечетной степени;
- для расщепления сигналов следует использовать 5 каналов расщепления;
- наименьшая среднеквадратичная погрешность фильтрации достигается, если расщепленные сигналы содержат равное количество предыдущих и последующих отсчетов по отношению к текущему моменту времени n ;
- комбинированный фильтр в виде каскадного соединения МФ и ННЦФ не улучшает качество фильтрации импульсных помех, поражающих речевые сигналы;
- ННЦФ дает наименьшую среднеквадратичную погрешность фильтрации по сравнению с МФ и фильтром Вольтерры.

Качество результатов работы ННЦФ и ЧННЦФ практически одинаково.

На практике для синтеза нелинейного фильтра следует применять спектральный метод расщепления, поскольку в нем общая задача аппроксимации оператора фильтра, поставленная во временной области, разбивается на несколько аппроксимационных задач существенно меньших размерностей, решаемых в частотной области. Таким образом, удается практически снять вычислительные проблемы (плохую обусловленность, шумы арифметики), типичные для синтеза ННЦФ.

При синтезе ЧННЦФ импульсных помех речевых сигналов установлено, что

- для удобства ДПФ следует выбирать 6 (четное число) каналов расщепления;
- как и в случае ННЦФ, следует использовать многомерные полиномы нечетной степени;
- наименьшую среднеквадратичную погрешность дает некаузальный частотный фильтр при обработке ДПФ вектора расщепленных сигналов с приблизительно равным числом предыдущих и последующих отсчетов относительно текущего момента времени n ;
- для повышения точности фильтрации следует учитывать эффект наложения в спектрах степенных составляющих выходного сигнала ЧННЦФ;
- блочный (фрагментарный) способ формирования выходного сигнала частотного фильтра уступает по точности обработки последовательному способу вычисления реакции фильтра.

Прослушивание выходных сигналов фильтров дает следующие результаты. Обработанный МФ речевой сигнал не содержит импульсных помех, однако, он существенно искажен (речь становится глухой с неестественным звучанием).

После фильтрации Вольтерры и обработки ЧННЦФ (или ННЦФ) в речевом сигнале присутствуют редкие остаточные импульсные помехи малой амплитуды, которые воспринимаются как шорох или потрескивание, при этом выходной речевой сигнал имеет естественное звучание. ЧННЦФ формирует значительно меньше остаточных помех по сравнению с фильтром Вольтерры.

Литература

1. Тьюки Дж. Анализ результатов наблюдения. – М.: Мир, 1981.
2. Быстрые алгоритмы в цифровой обработке изображений / Т.С. Хуанг, Дж.-О. Эклунд, Г.Дж. Нуссбаумер и др.; Под ред. Т.С. Хуанга. – М.: Радио и связь, 1984.
3. Vaseghi S.V. Advanced Digital Signal Processing and Noise Reduction. – N.Y.: Wiley, 2000.
4. Ланнэ А.А. Нелинейные динамические системы: синтез, оптимизация и идентификация. – Л.: ВАС, 1985.
5. Ланнэ А.А. Нелинейные полиномиальные цифровые фильтры // Цифровая обработка сигналов. – 1999. – №1. – С.18–26.
6. Пупков К.А., Капалин В.И., Ющенко А.С. Функциональные ряды в теории нелинейных систем – М.: Наука, 1976.
7. Данилов Л.В. Ряды Вольтерра-Пикара в теории нелинейных электрических цепей. – М.: Радио и связь, 1987.
8. Башарин С.А., Соловьева Е.Б. Моделирование и анализ нелинейных электрических цепей: Учеб. пособие. – СПб.: СПбГЭТУ, 1999.
9. Schetzen M. The Volterra and Wiener theory of nonlinear systems. – N. Y.: Wiley, 1980.
10. Ланнэ А.А., Соловьева Е.Б. Нелинейная фильтрация импульсных помех методом расщепления // Изв. вузов. Радиоэлектроника. – 1999. – Т.42, №7. – С.3–17.
11. Rugh W.J. Nonlinear System Theory. The Volterra/Wiener Approach. – The Johns Hopkins University Press, 1981.
12. Основы цифровой обработки сигналов: Курс лекций/ А. И. Солонина, Д.А. Улахович, С.М. Арбузов, Е.Б. Соловьева/ Изд. 2-е испр. и перераб. – СПб.: БХВ-Петербург, 2005.
13. Kim K.I., Powers E.J. Digital method of modeling quadratically nonlinear systems with a general random input // IEEE Trans. ASSP. – 1988. – V.36, №11. – P.1758–1769.

УДК 621.372

МЕТОДЫ СИНТЕЗА УЗКОПОЛОСНОГО АДАПТИВНОГО КИХ-ФИЛЬТРА НА ОСНОВЕ МНОГОСКОРОСТНОЙ ОБРАБОТКИ

Витязев В.В., Витязев С.В.

Введение

Методы и алгоритмы цифровой адаптивной фильтрации являются эффективным инструментом решения многих проблем в области современных телекоммуникаций, в значительной степени ориентированных на цифровые технологии передачи информации по традиционным аналоговым каналам связи как проводным, так и беспроводным. Это – борьба с замираниями и межсимвольной интерференцией, эхо-компенсация и шумоподавление, сжатие и закрытие речи и изображений. При этом, как правило, синтез адаптивного фильтра ведется в классе КИХ-цепей, не имеющих обратных связей (в рабочем режиме, по окончании обучения), а, следовательно, отличающихся абсолютной устойчивостью и стабильностью характеристик. Вместе с тем, с увеличением порядка КИХ-фильтра, который может достигать сотни и тысячи единиц, значительно возрастают вычислительные затраты, память фильтра, требуемая разрядность представления данных, затягивается процесс обучения, ухудшаются качественные показатели работы.

Необходимость значительного повышения порядка адаптивного КИХ-фильтра является следствием того, что моделируемая в процессе обучения динамическая система (например, канал связи) может иметь ярко выраженные перепады функции передачи в частотной области, которые принимают форму глубоких «замираний» или «резонансов» на определенных участках частот. С тем, чтобы максимально близко подстроиться к таким частотным перепадам, требуется соответствующее разрешение по частоте, а это достигается только одним способом – адекватным увеличением длины импульсной характеристики адаптивного фильтра, моделирующего заданную динамическую систему.

Известно [1,2], что в основе современных методов построения систем цифровой частотной селекции сигналов и, в частности, узкополосных КИХ-фильтров, лежит фундаментальная идея использования вторичной дискретизации (децимации) и интерполяции обрабатываемого сигнала с целью многократного уменьшения вычислительных затрат и памяти данных. Многоскоростная обработка сигналов (МОС) получила широкое применение не только для обычной частотной селекции и построения банков фильтров [3-5], но и эффективно используется во всех практических приложениях, там, где возможно «принудительное» понижение и повышение частоты дискретизации. Одной из таких сфер приложений, в первую очередь, стали телекоммуникационные и радиотехнические системы. Сочетание адаптив-

Рассматривается проблема адаптивной фильтрации узкополосных процессов на фоне широкополосного шума. С целью значительной минимизации вычислительных затрат и памяти данных и коэффициентов синтезируемых КИХ-фильтров предлагается использование многоскоростной обработки на основе эффекта децимации обрабатываемого сигнала и импульсной характеристики с последующей их интерполяцией.

ной и многоскоростной обработки сигналов дало новое решение традиционных проблем и стимулировало дальнейшие исследования в этой области. [6-11].

Постановка задачи

Рассматривается задача построения узкополосного адаптивного КИХ-фильтра при следующих исходных условиях моделирования. Обучающий сигнал $s(n)$ – узкополосный случайный процесс, формируемый узкополосной динамической системой (моделируемым фильтром), на вход которой подается дискретный белый шум с нулевым математическим ожиданием и заданным (условно единичным) уровнем дисперсии. Подобный процесс может иметь место, например, при диагностике и идентификации динамических объектов и систем с ярко выраженными «резонирующими» свойствами их частотных характеристик. Предполагается, что на обучающий сигнал $s(n)$ в процессе его формирования и измерения аддитивно накладывается шум $p(n)$, равномерно распределенный во всем диапазоне рабочих частот. Таким образом, используемый для автоматической настройки адаптивного фильтра обучающий сигнал $d(n)$ представляет собой сумму полезной составляющей $s(n)$ и шума $p(n)$:

$$d(n) = s(n) + p(n).$$

Если диапазон рабочих частот обучающего сигнала $d(n)$ значительно превышает эффективную ширину спектральной плотности мощности полезной составляющей $s(n)$, что может быть обусловлено как изменением в широких пределах центральной частоты спектрального положения сигнала $s(n)$, так и широкополосным шумом $p(n)$, то обучающий сигнал $d(n)$ – широкополосный случайный процесс со всеми вытекающими последствиями.

Вынужденное повышение частоты дискретизации обучающего сигнала $d(n)$, значительно превышающей эффективную ширину спектра его полезной составляющей $s(n)$, приводит к многократному увеличению требуемого порядка N моделирующего адаптивного КИХ-фильтра, который, как известно [2], определяется прежде всего отношением частоты дискретизации входного

сигнала к эффективной ширине полосы пропускания. Увеличение порядка адаптивного фильтра – это не только значительный рост вычислительных затрат, особенно на этапе обучения при использовании быстрых алгоритмов адаптации, таких как метод РНК, но и уменьшение скорости и точности настройки [12].

В предлагаемой вашему вниманию работе исследуются методы построения моделирующего адаптивного КИХ-фильтра, в предположении, что моделируемая динамическая система также относится к классу КИХ-цепей, т.е. имеет конечную импульсную характеристику. В более общем случае, моделируемая система может относиться к классу БИХ-цепей, а, следовательно, должна предполагать использование адекватного класса моделирующих фильтров. Вместе с тем, возможно (с точностью до постоянной задержки обучающего сигнала) построить адаптивный КИХ-фильтр, приближающийся по своим избирательным свойствам к моделируемой БИХ-системе, но – это отдельная задача.

Общая структурная схема подключения адаптивного фильтра к идентифицируемой динамической системе представлена на рис.1. Здесь: $x(n)$ - тестирующий сигнал, $y(n)$ - сигнал на выходе адаптивного фильтра (АФ), $\varepsilon(n)$ - ошибка обучения.

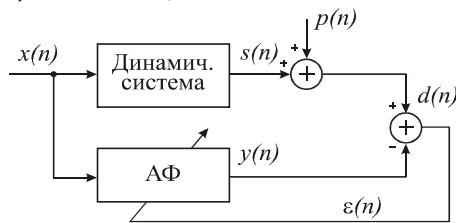


Рис. 1. Схема подключения адаптивного фильтра к идентифицируемой системе.

Поскольку тестирующий сигнал $x(n)$ представляет собой случайный процесс, то и все выходные сигналы: $s(n)$, $d(n)$ и ошибка обучения $\varepsilon(n)$ являются случайными процессами. Поэтому задача фактически заключается в поиске такого оптимального вектора весовых коэффициентов АФ $H_{opt} = \{h_{0opt}, h_{1opt}, h_{2opt}, \dots, h_{N-1opt}\}$, при котором среднеквадратичная ошибка (СКО) обучения $\delta = E[\varepsilon^2(n)]$ достигает минимального значения, т.е.

$$\delta = E[\varepsilon^2(n), H] \rightarrow \min_H.$$

Если априорно известны статистические характеристики тестирующего сигнала $x(n)$ такие, как автокорреляционная матрица R и вектор взаимной корреляции P тестирующего $x(n)$ и обучающего $s(n)$ сигналов, то оптимальный вектор весовых коэффициентов для квадратичной рабочей функции принимает вид []:

$$H_{opt} = R^{-1} P.$$

Однако, в реальных условиях работы, когда отсутствует подобная информация, как правило, используют градиентные методы поиска оптимального решения, такие как МНК, РНК и БМНК (метод наименьших квадратов, рекурсивный МНК и быстрый МНК на основе прямого и обратного БПФ с адаптацией в частотной области) [13]. Все эти методы хорошо исследованы и дают прекрасные результаты при относительно небольшом порядке N вектора весовых коэффициентов и достаточном отноше-

нии сигнал/шум на выходе идентифицируемой динамической системы. Вместе с тем, увеличение порядка N приводит либо к недопустимо длительному процессу обучения при пропорциональном росте вычислительных затрат (для МНК), либо требует многократного увеличения вычислительных затрат при высокой чувствительности к шумовой составляющей в обучающем сигнале (для РНК). В некотором смысле альтернативой является метод быстрого МНК, позволяющий при больших значениях порядка фильтра N значительно уменьшить объем вычислительных затрат путем использования блочной обработки и перехода к адаптации в частотной области. Но, механизм блочной обработки – это дополнительные затраты памяти данных и свои проблемы с обеспечением необходимой скорости сходимости алгоритма адаптации [13]. Дело в том, что все перечисленные классические методы являются общими и не учитывают специфику узкополосности обучающего сигнала $s(n)$ и соответственно сигнала $y(n)$ на выходе адаптивного фильтра.

Задача заключается в поиске альтернативных методов адаптации узкополосного КИХ-фильтра, минимизирующих ошибку обучения $\varepsilon(n)$ при существенно меньшем объеме вычислительных затрат и при этом достигающих большой скорости адаптации и точности настройки с учетом специфических особенностей узкополосных фильтров. В основе нового эффективного решения проблемы повышения качественных характеристик процесса обучения, и прежде всего многократного уменьшения вычислительных затрат, лежит применение многоскоростной обработки [14].

Метод адаптивной многоскоростной фильтрации узкополосных процессов

Многоскоростная фильтрация предполагает, что в процессе обработки сигнала может использоваться переход на другую частоту дискретизации, понижающую – при децимации, или повышающую – при интерполяции. Децимация и интерполяция являются тем технологическим инструментом, с помощью которого, при прочих равных условиях, удастся многократно уменьшить вычислительные и аппаратные затраты на реализацию фильтров частотной селекции, а в случае адаптивной фильтрации обеспечить устойчивую и более точную работу в условиях воздействия шумов при одновременном увеличении скорости сходимости (в зависимости от применяемого в сочетании с многоскоростной обработкой алгоритма адаптации).

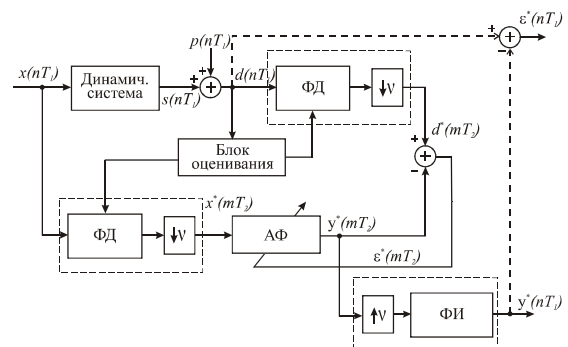


Рис. 2. Схема двухэтапного прямого моделирования динамической системы

На рис. 2 представлена общая структурная схема прямого моделирования динамической системы с использованием узкополосного адаптивного фильтра и МОС с использованием адаптивной многоскоростной фильтрации.

В соответствии с данной схемой, адаптация к «резонирующим» узкополосным участкам спектра обучающего сигнала $d(nT_1)$ с выхода динамической системы выполняется в два этапа. На первом этапе, в режиме адаптации без обратной связи блоком оценивания, на основе цифрового спектрального анализа, производится поиск «резонирующих» участков спектра обучающего сигнала $d(nT_1)$. При их обнаружении блок оценивания настраивает параметры полосового фильтра-дециматора (центральную частоту и полосу пропускания) на соответствующую полосу частот. На втором этапе, в режиме адаптации с обратной связью, после понижения частоты дискретизации выполняется настройка на обучающий сигнал $d^*(mT_2)$. С целью восстановления исходной частоты дискретизации выходного сигнала $y^*(nT_1)$ применяется фильтр-интерполятор. Для окончательной оценки точности применяемого метода адаптации и сравнения с обычными классическими алгоритмами может быть вычислена ошибка обучения $\varepsilon^*(nT_1)$.

Рассматриваемый метод построения адаптивного КИХ-фильтра отличается, прежде всего, значительным уменьшением вычислительных затрат, так как порядок АФ после децимации сигнала на его входе уменьшается в V раз. Кроме того, в V раз уменьшается и требуемая скорость обработки прореженного входного потока данных. Таким образом, приведенные к периоду дискретизации T_1 общие вычислительные затраты уменьшаются в V^2 раз. Не менее важным фактором является значительное повышение соотношения сигнал/шум для обучающего сигнала $d^*(mT_2)$ на выходе фильтра-дециматора, что является следствием его сглаживающих свойств. Известны и другие преимущества эффективного сочетания многоскоростной и адаптивной фильтрации [14]. Вместе с тем, не надо забывать и о недостатках, свойственных фильтрам-дециматорам. Это, прежде всего, появление специфического практически неустраняемого дальнейшей обработкой шума децимации, обусловленного неидеальностью частотной характеристики фильтра в зоне непрозрачности. Во-вторых, это дополнительная задержка, которую вносит всякий КИХ-фильтр (чтобы ее минимизировать, необходимо расширение полосы пропускания фильтра-дециматора, а значит – уменьшение коэффициента децимации V и, как следствие, увеличение всех последующих затрат на адаптацию и фильтрацию). И, наконец, если полоса частот спектра обучающего сигнала $s(nT_1)$ не ограничена его эффективной шириной, то встает вопрос об оптимальном выборе параметров фильтра-дециматора с позиции минимизации СКО, обусловленного не только шумом $p(nT_1)$, но и частичным усечением спектральной плотности мощности полезной составляющей $s(nT_1)$. Поэтому представляется целесообразным в этих условиях исключить из общей схемы дополнительные фильтры-дециматоры и вести процесс адаптации, используя децимацию непосредственно на выходе адаптивного КИХ-фильтра по схеме, приведенной на рис. 3.

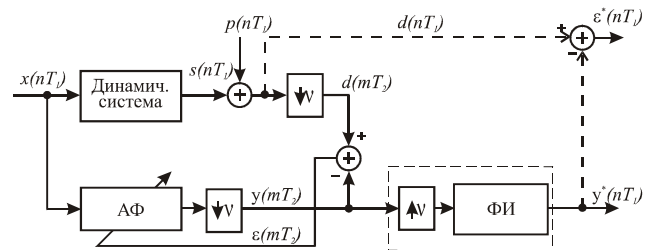


Рис.3. Схема одноэтапного прямого моделирования динамической системы с использованием многоскоростной адаптивной фильтрации.

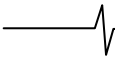
В отличие от схемы двухэтапного моделирования, представленной на рис. 2, схема одноэтапного прямого моделирования динамической системы предполагает, что ошибка обучения $\varepsilon(mT_2)$ вычисляется с периодом $T_2 = VT_1$, т.е. в V раз реже, что потенциально уменьшает приведенные вычислительные затраты в V раз. Кроме того, уменьшается влияние шума $p(nT_1)$ на сам процесс обучения, при условии, что период дискретизации T_2 значительно превышает интервал корреляции его отсчетов. Вместе с тем, с увеличением коэффициента децимации V и, соответственно, периода формирования ошибки обучения $\varepsilon(mT_2)$ может иметь место шум децимации, связанный с потерями высокочастотных составляющих сигнала обучения $s(nT_1)$. Поэтому выбор оптимального значения коэффициента децимации V , минимизирующего суммарную ошибку, является в данном варианте построения адаптивной многоскоростной системы задачей первостепенной важности.

Решение поставленной задачи может быть получено путем моделирования системы и оценки минимально достижимой ошибки обучения при различных значениях коэффициента децимации V в окрестности начального значения, определяемого отношением частоты дискретизации входного тестирующего сигнала $x(nT_1)$ к эффективной ширине спектра сигнала $s(nT_1)$ на выходе динамической системы.

Описанные выше одно- и двухэтапная схемы прямого моделирования динамической системы с использованием многоскоростной адаптивной обработки не лишены одного общего недостатка: роста шума децимации с увеличением коэффициента децимации V и с уменьшением частотной избирательности моделирующей системы за пределами зоны подавления высокочастотных составляющих обучающего сигнала. Представляется целесообразным, как и для цифровых систем частотной селекции [2], рассмотреть альтернативный подход к проектированию узкополосных адаптивных фильтров, использующих вместо многоскоростной обработки децимацию и интерполяцию импульсной характеристики фильтра.

Метод синтеза узкополосного адаптивного фильтра на основе децимации и интерполяции импульсной характеристики

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, вы-



деляющего из множества спектральных составляющих на выходе ЦГФ полосу частот обучающего сигнала. На рис. 4 представлен один из вариантов построения структуры узкополосного адаптивного фильтра в форме последовательного соединения двух фильтров: адаптивного гребенчатого фильтра (АГФ) и цифрового сглаживающего фильтра (ЦСФ).

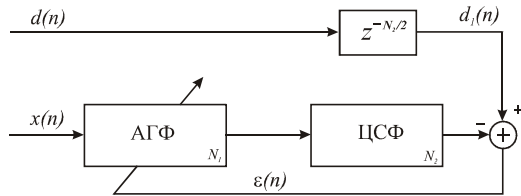


Рис. 4. Вариант построения адаптивного узкополосного фильтра на основе децимации и интерполяции импульсной характеристики

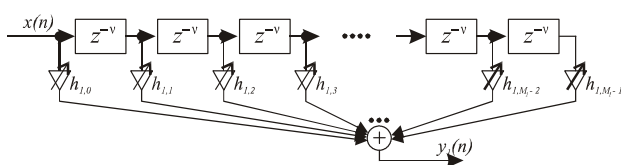


Рис. 5. Прямая форма гребенчатого КИХ-фильтра с перестраиваемыми весовыми коэффициентами.

Адаптация к обучающему сигналу $d(n)$ выполняется в два этапа. На первом этапе обработки используется КИХ-фильтр с прореженной в V раз импульсной характеристикой, реализуемый по структуре, приведенной на рис. 5. Если предположить, что между каждой парой соседних весовых коэффициентов (отсчетов импульсной характеристики) $h_{1,i}, i = 0, M_1 - 1$, равномерно расположены по $v - 1$ нулей, то частотные характеристики подобного фильтра будут носить периодический характер, что и оправдывает его название – гребенчатый фильтр. Отметим, что в силу строгой периодичности частотной характеристики АГФ настройка его параметров в процессе обучения будет осуществляться таким образом, чтобы максимально адаптироваться к мощной узкополосной составляющей (при достаточном соотношении сигнал/шум), оставляя почти без внимания все остальные боковые полосы частот. В частности, если реализуется низкочастотная узкополосная фильтрация, то период дискретизации $T_2 = vT_1$ импульсной характеристики ЦГФ будет значительно превышать интервал корреляции отсчетов всех частотных составляющих обучающего сигнала $d(n)$, лежащих за основной полосой в низкочастотной области. В общем случае частотное положение основной полосы спектра обучающего сигнала может быть неизвестно. Поэтому возникает неопределенность: какую из V частотных составляющих на выходе АГФ должен выделить выходной сглаживающий фильтр? Эту задачу и решает ЦСФ, подавляя все остальные боковые составляющие. Предполагается, что ЦСФ настраивается на заданную полосу частот по результатам предварительного спектрального анализа.

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. Но, очевидно, и другое – метод эффективно работает только при достаточном отношении сигнал/шум, когда можно пренебречь влиянием на процесс адаптации всех частотных составляющих за полосой частот полезной составляющей обучающего сигнала.

С целью увеличения соотношения сигнал/шум на входе АГФ целесообразно изменить порядок включения фильтров, поставив первым полосовой сглаживающий фильтр, выделяющий по априорной информации (или полученной по результатам предварительного спектрального анализа) основную полосу частот обучающего сигнала и подавляющую все мешающие частотные составляющие (аналогично дециматору на входе последующего адаптивного фильтра, см. рис. 2).

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.

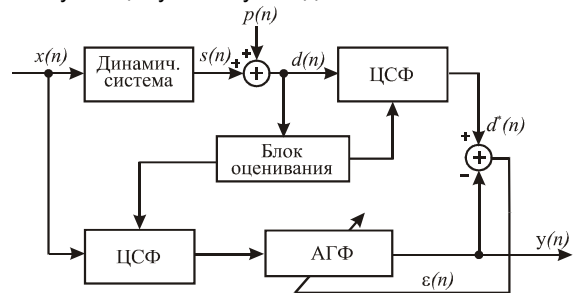


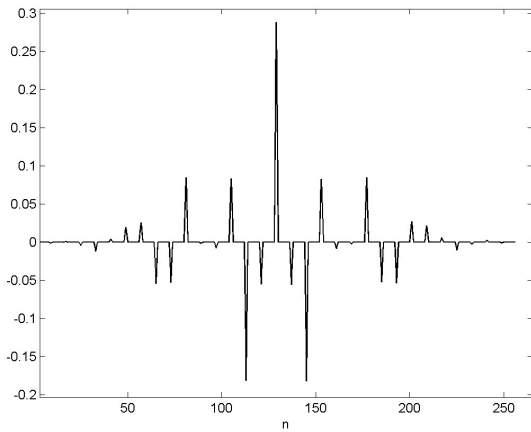
Рис. 6. Схема двухэтапного прямого моделирования динамической системы с использованием адаптивного гребенчатого фильтра.

Моделирование адаптивного узкополосного КИХ-фильтра

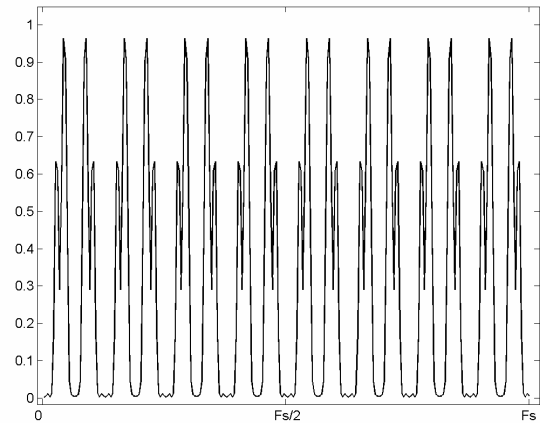
Моделирование процесса обучения с применением децимации импульсной характеристики адаптивного фильтра проводилось при следующих исходных условиях. Моделируемая динамическая система – КИХ-фильтр заданного порядка N ($N = 256$). Функция передачи (АЧХ) моделируемой системы носит узкополосный «резонирующий» характер в произвольно заданной полосе частот (приблизительно, от 270 кГц до 300 кГц при частоте дискретизации 1 МГц) и однозначно определяется M ($M = 8$) дискретными отсчетами, равномерно распределенными с шагом по частоте $\Delta\omega = 2\pi / N$ (вне указанной полосы частот соответствующие отсчеты принимают нулевое значение). Фазочастотная характеристика считается линейной.

На вход моделируемой динамической системы и моделирующего адаптивного фильтра подается случайный сигнал $x(n)$, представляющий собой дискретный белый шум с нулевым математическим ожиданием и дисперсией, принимающей условно единичное значение. К полезному сигналу $s(n)$ на выходе моделируемой системы аддитивно добавляется шум $p(n)$ – дискретный белый с нулевым математическим ожиданием и дисперсией от 0,001 до 0,1 в относительных единицах.

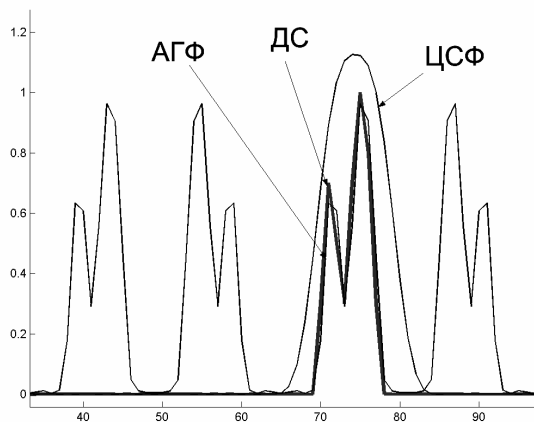
Для моделирования работы адаптивного КИХ-фильтра использовался метод МНК. Коэффициент децимации импульсной характеристики $v = 1..16$. Результаты моделирования для случая $v = 8$ представлены на рис. 7.



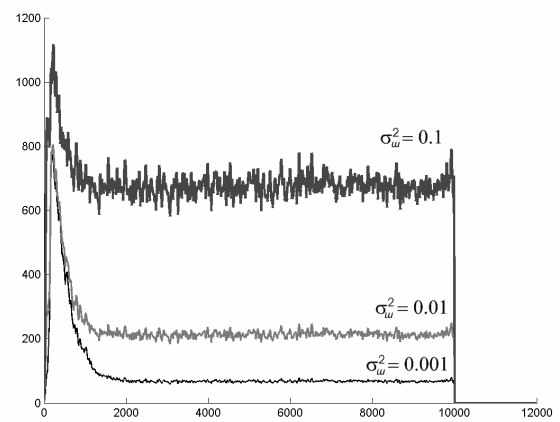
а)



б)



в)



г)

Рис. 7. Результаты моделирования работы адаптивного гребенчатого фильтра.

На рис. 7 а) показана импульсная характеристика адаптивного гребенчатого фильтра по завершению процесса настройки. По графику хорошо видно, что большинство коэффициентов фильтра равно нулю. Частотная характеристика (АЧХ) АГФ после настройки показана на рис. 7 б). На интервале от 0-й частоты до частоты Найквиста ($F_s/2$) укладывается 4 пары копий АЧХ моделируемой динамической системы. На рис. 7 в) выделен интервал частот в окрестности интересующей нас полосы пропускания моделируемой системы. АЧХ АГФ достаточно точно повторяет форму АЧХ динамической системы (ДС), но включает и набор копий АЧХ ДС, появившихся в результате прореживания импульсной характеристики. Цифровой сглаживающий фильтр (ЦСФ) устраняет влияние этих боковых полос на результат работы моделирующей системы. На рис. 7 г) представлены кривые среднеквадратических отклонений (СКО) сигнала ошибки, построенные для случаев дисперсии шума $p(n)$ равной 0,1, 0,01 и 0,001 по результатам 100 наблюдений.

Заключение

Анализ результатов проведенных исследований потенциальной возможности и эффективности применения многоскоростной обработки сигналов в задачах адаптивной фильтрации узкополосных процессов на фоне широкополосного шума позволяет сделать следующие выводы:

1. Предложенные методы синтеза структуры узкополосного адаптивного КИХ-фильтра на основе децимации и интерполяции обрабатываемого сигнала или его импульсной характеристики позволяют многократно уменьшить приведенные вычислительные затраты, а также требуемую емкость памяти данных (для методов на основе децимации обрабатываемого сигнала) и памяти коэффициентов (для методов на основе децимации импульсной характеристики), сохраняя высокую скорость и точность настройки.

2. Дальнейшие исследования должны быть направлены на разработку модификаций предложенных методов с использованием комплексного представления обрабатываемых данных. Это снимет установленное в работе ограничение на выбор значения коэффициента децимации вещественной импульсной характеристики адаптивного гребенчатого фильтра, обусловленное двухполосным представлением частотной характеристики соответствующего ему базового НЧФ (фильтра-прототипа).

3. Необходимы дополнительные исследования, связанные с разработкой работоспособных структур узкополосного адаптивного фильтра на базе гребенчатого фильтра, настраиваемого по алгоритмам МНК и РНК, и подстраиваемого сглаживающего фильтра, с последующим сравнительным анализом их эффективности.

Литература

1. Crochiere R.E., Rabiner L. Multirate Digital Signal Processing. Prentice Hall. Englewood Cliffs. – NJ, 1983.
2. Витязев В.В. Цифровая частотная селекция сигналов. М.: Радио и связь, 1993. – 240 с.
3. Vaidyanathan P.P. Multirate Systems and Filter Banks. Prentice Hall. Englewood Cliffs. – NJ, 1993.
4. Зайцев А.А. Методы построения банков цифровых фильтров: тематический обзор // Цифровая обработка сигналов, 2003, № 1, с. 2-10.
5. Витязев В.В., Зайцев А.А. Основы многоскоростной обработки сигналов: Учебное пособие, Ч.1 и Ч.2. Рязан. госу- дарст. радиотехн. ун-т, Рязань, 2005, - 124 с., 2006, - 104 с.
6. Линович А.Ю., Витязев В.В. Субполосная адаптивная фильтрация в задачах обратного моделирования // Циф- ровая обработка сигналов, 2004, № 1, с. 41-48.
7. Линович А.Ю. Применение методов частотно-временной декомпозиции при решении задач обратного моделирова- ния // Цифровая обработка сигналов, 2005, № 3, с. 28-37.
8. Витязев В.В., Колодыко Г.Н., Витязев С.В. Способы и ал- горитмы формирования радиолокационного изображения в режиме доплеровского облучения луча // Цифровая об- работка сигналов, 2006, № 3, с. 31-41.
9. Кузнецов Е.П., Витязев В.В. Цифровая обработка сигна- лов в задачах эхо-компенсации // Цифровая обработка сигналов, 2006, № 4, с. 20-28.
10. Витязев В.В., Колодыко Г.Н., Витязев С.В. Селекция на- земных движущихся целей на основе многоскоростной адаптивной обработки траекторного сигнала // Цифровая обработка сигналов, 2007, № 1, с. 41-50.
11. Кузнецов Е.П. Методы и алгоритмы адаптивной эхо- компенсации: сравнительный анализ эффективности приме- нения // Цифровая обработка сигналов, 2007, № 2, с. 26-34.
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., Бодров К.А., Иванов С.В. Адаптивная много- скоростная фильтрация узкополосных процессов // Циф- ровая обработка сигналов и ее применение – DSPA'1999. Труды 1-й Международной научно-технической конферен- ции, М.:, 1999, Т. 1, с. 150-160.

Уважаемые коллеги!

Предлагаем вам принять участие в формировании тематических выпусков журнала «Цифровая обра- ботка сигналов» и размещению рекламы продукции (услуг) Вашей фирмы на его страницах. В случае положи- тельного 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 2008 году планируется выпуск 4-х номеров журнала (тираж до 1000 экз.). Журнал будет распростра- няться по подписке через агентство «Роспечать» в России, СНГ и странах Балтии (индекс 82185), а также на Выставках: «Цифровая обработка сигналов и ее применение – DSPA'2008», «ЕвроElectronica», «СвязьЭкспо- комм», «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 следую- щие 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 ка- ждом 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.dsra.ru) в течение всего года, что привлечет внимание к продукции (ус- лугам) Вашей фирмы новых участников на рынке DSP-технологий (ежедневно фиксируется до 100 и более по- сещений сайта www.dsra.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
д.т.н., профессор Витязев Владимир Викторович

Предложения прошу направлять по адресу: E-mail: tor@rgtra.ryazan.ru или info@dsra.ru

УДК 519.224

ПОСТРОЕНИЕ СТАТИСТИЧЕСКОЙ МОДЕЛИ РЕЧЕВОГО ТРАФИКА ПРИ МОНОЛОГЕ

Сахаров А.В.

Введение

IP-телефония – новая технология, получившая широкое распространение во всем мире, и в настоящее время продолжает динамично развиваться. Свидетельством перспективности данной технологии может служить тот факт, что многие операторы сотовой и стационарной связи заявили о предоставлении этой услуги, а всемирно-известная компания AT&T, уже осуществляющая передачу речи по IP-сетям, объявила о долгосрочном плане перевода всего своего трафика дальней связи на платформу IP.

Создание любой более или менее разветвленной пакетной сети начинается с определения ее топологии и характера нагрузки, создаваемой каждым рабочим местом, а также требований, предъявляемых приложениями к пропускной способности, задержкам и вариации задержки пакетов при передаче. Важным этапом при построении сети является имитационное моделирование нагрузки на разных ее участках, что позволяет оптимизировать размеры буферов и, как следствие, снизить вероятность заторов в сети. Трафик в пакетных сетях хорошо изучен и достаточно точно описывается самоподобными случайными процессами с параметром Херста 0,65-0,8 [1]. Большинство клиент-серверных приложений, служба электронной почты и многие другие программы не критичны к задержкам и тем более к вариации задержки. Не имеет решающего значения, придут ли данные немного ранее или немного позднее. Поврежденные данные всегда можно запросить повторно. Здесь главное – сам факт доставки данных. Совсем иначе обстоит дело в случае с IP-телефонией. Речевой трафик не похож ни на один другой вид трафика. Приложения IP-телефонии предъявляют к сети ряд специфических требований (поскольку изначально сеть проектировалась исключительно для передачи данных), описываемых ниже. Качество передачи речи считается хорошим, согласно рекомендации G.114 союза ITU-T, если суммарная задержка при передаче сигнала в одну сторону не превышает 150 мс. Удовлетворительным считается качество при задержке до 450 мс. Пакеты с кодированной речью, задержавшиеся на время

Предлагается новый подход к построению статистической модели трафика IP-телефонии при монологе с учетом уровня кодируемого сигнала. Введение данного параметра в модель позволяет получать на выходе трафик с разной битовой скоростью, что наиболее полно отражает реальную ситуацию. Проведено моделирование битовой скорости. Сравнение модельных и опытных данных показало достаточно высокую степень корректности полученной модели.

более данного порогового значения, отбрасываются и считаются потерянными. Понятно, что никакой речи о повторной передаче данных здесь не может быть, поскольку это вызовет еще большие задержки. Еще одной проблемой при передаче речи поверх протокола IP является вариация задержки данных или, другими словами, джиттер. Блоки с речью генерируются кодеком через равные промежутки времени. Задержка при передаче является случайной величиной и, таким образом, пакеты приходят к месту назначения с интервалами отличными от исходных. Джиттер является причиной «дрожания» речи. С ним призваны бороться джиттер-буферы, восстанавливающие изначальные интервалы и порядок следования пакетов. Необходимая ширина канала для передачи речи довольно небольшая и, в случае подключения к сети Ethernet, колеблется в пределах 13 – 21 Кбит/с. В ряду с описанными отличиями трафик IP-телефонии не является самоподобным [3]. Все это говорит о необходимости учета при проектировании компьютерных сетей исключительных требований, предъявляемых приложениями IP-телефонии к передаче речи. Специфические особенности речевого трафика, такие как низкая битовая скорость и последствия работы детектора речевой активности (VAD – voice activity detector), отбрасывающего блоки, не содержащие речь, приводят к выводу о необходимости создания особой модели, учитывающей все характерные перечисленные свойства трафика IP-телефонии.

В ходе телефонного разговора оба абонента попеременно создают нагрузку на сеть. Таким образом, телефонный разговор можно представить как диалог, состоящий из следующих друг за другом монологов разной продолжительности (рис. 1).

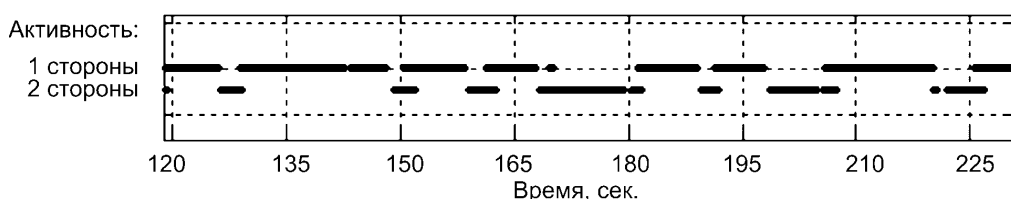


Рис. 1. Диаграмма активности участников телефонного разговора

Способ построения статистической модели речевого трафика

В данной статье предлагается способ построения статистической модели трафика при монологе как «on/off» последовательности. Такая последовательность состоит из следующих друг за другом интервалов с активной речью и пауз. Паузы являются результатом работы детектора речевой активности, входящего в состав вокодера. Блоки, не содержащие речь, не попадают в выходной поток (рис. 2в.) и, таким образом, уменьшается необходимая ширина канала для передачи без существенного ухудшения качества передаваемой речи.



Рис. 2. Этапы обработки речи на стороне источника (а) оцифрованный поток, (б) после разбиения на блоки, (в) поток, прошедший VAD т.е. без пауз, (г) сжатый поток, (д) поток RTP пакетов (по одному блоку на пакет)

Из выше сказанного следует, что построение статистической модели трафика сводится к определению продолжительностей интервалов с паузами и речью. Таким образом, необходимо подобрать функции плотностей распределения вероятностей и их параметры, которые наиболее точно описывают вероятность появления интервала с речью или паузами той или иной продолжительности.

Экспериментальными данными послужили 106 файлов, собранные в течение монолога. Речь кодировалась кодеком G.723.1 с битовой скоростью 6.4 Кбит/с., хорошо себя зарекомендовавшим и принятым производителями средств пакетной телефонии. Каждый файл содержит пакеты с кодированной речью. Продолжительность сбора информации каждого файла составила от 3 до 5 минут. Перехват пакетов производился на том же компьютере, на котором они и генерировались, что позволило снизить «дрожание» в следовании пакетов до порядка 1 мс [4].

Продолжительности интервалов с речью и паузами определялись на основе маркеров и временных меток, являющихся частью заголовка RTP (Real Time Transport Protocol – транспортный протокол реального масштаба времени). Он описан в RFC 1889. Маркером помечается первый из пакетов, ответственных за передачу участка с речью. Ключевым параметром в определении продолжительностей явилось значение временной метки, поскольку маркер иногда выставляется и среди подряд идущих блоков, что может быть объяснено наличием речи и паузы в блоке предшествующем маркеру. Подобное случалось редко и в расчет не бралось. Значение временной метки для последовательно идущих блоков отличается на количество отсчетов закодированных в одном блоке. Для кодека G.723.1 это значение равно 240. Величина временной метки увеличивается и в том случае, когда блок отбрасывается, как не содержащий речь [2]. Таким образом, разница в значениях меток более чем на 240 (т.е. 480, 720 и т.д.)

говорит о наличии паузы. Время первых и последних пакетов с речью фиксировалось, что позволило восстановить последовательность продолжительностей участков с речью и пауз.

В первую очередь необходимо определиться с набором распределений, среди которых будет выбрано одно, наиболее точно соответствующее экспериментальным данным. Было выбрано 6 распределений: гамма, Вейбулла, логнормальное, Пирсона 5 и 6 типа, бета распределения (рис. 4).

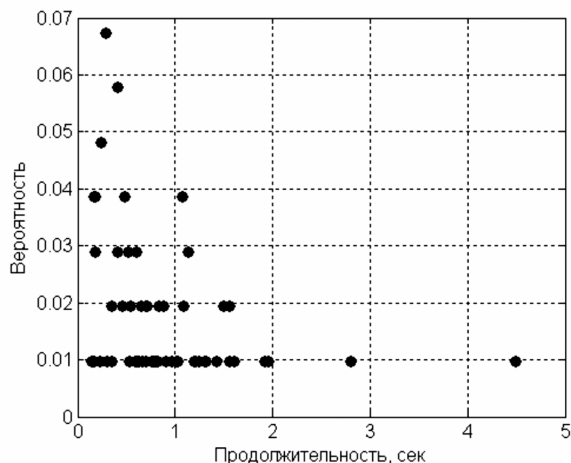


Рис. 3. Гистограмма плотности распределения вероятности продолжительности пауз

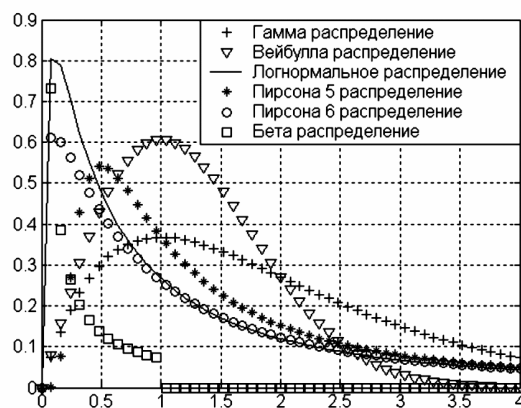


Рис. 4. Функции плотностей распределения вероятностей

Выбор производился на основе соответствия графиков функций плотностей распределения вероятностей и гистограмм плотностей распределения вероятностей экспериментальных данных (рис. 3).

Выбор наиболее подходящего распределения производился на основе критерия согласия Колмогорова-Смирнова. В начале, с помощью метода максимального правдоподобия, для каждого распределения на основе опытных данных подбирались его параметры. В этом методе предполагается независимость экспериментальных наблюдений $X_1, X_2, \dots, X_n$. В доказательство истинности данного предположения был подсчитан коэффициент корреляции, значение которого по абсолютной величине в среднем не превысило 0,23. Поскольку использование этого коэффициента не всегда оправдано, была проведена еще одна проверка. Неформальной оценкой независи-

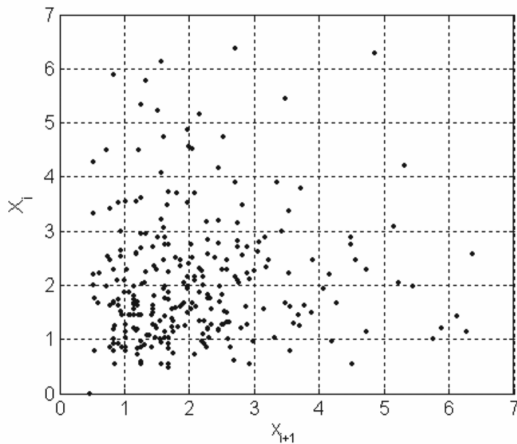


Рис. 5. Диаграмма разброса опытных данных

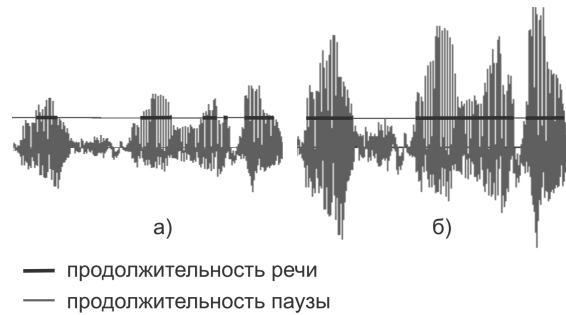


Рис. 6. Влияние амплитуды сигнала на продолжительность интервалов с речью и паузами а) сигнал с амплитудой A1 б) тот же сигнал с амплитудой A2, (A2 > A1).

Таблица 1

Результаты тестирования по критерию Колмогорова-Смирнова

Распределения:	Гамма	Вейбулла	Логнормальное	Пирсона 5 тип	Пирсона 6 тип	Бета
Речь	76	67	77	71	76	58
Пауза	101	101	106	96	106	85
Общий балл	177	168	183	167	182	143

мости данных, упорядоченных по времени их сбора, является диаграмма разброса (рис. 5). В случае, если величины $X_1, X_2, \dots, X_n$ являются зависимыми, то точки (X_i, X_{i+1}) стремятся расположиться вдоль линии с положительным или отрицательным уклоном в первом квадранте. В настоящем случае точки разбросаны хаотично, что подтверждает независимость данных.

Далее по критерию согласия Колмогорова-Смирнова с уровнем значимости $\alpha = 0,05$ проводилась проверка на соответствие распределений опытным данным. В случае положительного результата распределение «получало» один балл, иначе – ноль баллов. Результаты тестирования по всем 106 файлам представлены в таблице 1.

В итоге получается, что логнормальное распределение наиболее точно описывает как распределение продолжительностей интервалов с речью, так и с паузами.

Во время проведения анализа опытных данных было отмечено, что среднее число интервалов с речью и с паузами $\bar{N}_{on/off}$ составило 63 интервала на файл, при этом каждый файл содержит различное количество интервалов с речью и паузами $N_{on/off}$, варьирующееся от 2 до 160, что нельзя объяснить разницей только в продолжительности сбора информации в файл. Причем с уменьшением $N_{on/off}$ средняя продолжительность интервалов с речью $\bar{T}_{on}$ и паузами $\bar{T}_{off}$ увеличивается и уменьшается соответственно. Это говорит о том, что чувствительность микрофона (несмотря на активированные функции «настройка чувствительности во время вызова» и «автоматическое обнаружение тишины») или громкость говорящего менялись от монолога к монологу, что влияло на уровень кодируемого сигнала S (рис. 6).

Таким образом, получается, что создаваемая нагрузка на сеть и параметры распределения случайных величин T_{on} и T_{off} зависят от уровня S , который можно определить следующим образом:

$$S = \bar{N}_{on} / \bar{N}_{off}.$$

Поскольку случайная величина, распределенная по выбранному по результатам тестирования логнормальному распределению $LN(\mu, \sigma)$, принимает значения из области $[0, \infty)$, то к параметрам μ и σ данного распределения следует добавить максимальные и минимальные возможные значения случайной величины. Обозначим $T_{on}^{max}(S), T_{off}^{max}(S), T_{on}^{min}(S), T_{off}^{min}(S)$ – максимальные и минимальные продолжительности интервалов с речью и паузами соответственно. Из выше сказанного следует:

$$T_{on} \sim LN(\mu_{on}(S), \sigma_{on}(S)), \quad T_{on} \in [T_{on}^{min}(S), T_{on}^{max}(S)];$$

$$T_{off} \sim LN(\mu_{off}(S), \sigma_{off}(S)), \quad T_{off} \in [T_{off}^{min}(S), T_{off}^{max}(S)],$$

где функции

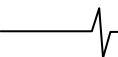
$$\mu_{on}(S), \sigma_{on}(S), T_{on}^{min}(S), T_{on}^{max}(S),$$

$$\mu_{off}(S), \sigma_{off}(S), T_{off}^{min}(S), T_{off}^{max}(S)$$

были получены в результате аппроксимации экспериментальных данных в смысле наименьших квадратов полиномами до третьей степени включительно (рис. 7):

$$T_{on}^{max}(S) = \begin{cases} 0.05 \times S^3 - 0.87 \times S^2 + 5.96 \times S - 1.73, & S \in [0.72, 15.29] \\ -0.0002 \times S^2 + 0.45 \times S + 42.99, & S \in [15.29, 910.3], \end{cases}$$

$$T_{on}^{min}(S) = -3e - 7 \times S^3 + 0.0002 \times S^2 - 0.002 \times S + 0.54,$$



$$T_{off}^{max}(S) = 6/\sqrt{S},$$

$$T_{off}^{min}(S) = 2e - 10 \times S^3 - 4e - 7 \times S^2 + 0.0002 \times S + 0.09,$$

$$\mu_{on}(S) = \begin{cases} 0.0005 \times S^3 - 0.02 \times S^2 + 0.29 \times S - 0.37, & S \in [0.72, 25.8] \\ 1e - 9 \times S^3 - 2e - 5 \times S^2 + 0.01 \times S + 1.63, & S \in [25.8, 910.3] \end{cases}$$

$$\sigma_{on}(S) = \begin{cases} 0.002 \times S^3 - 0.04 \times S^2 + 0.20 \times S + 0.32, & S \in [0.72, 10.69] \\ 2e - 9 \times S^3 - 2e - 6 \times S^2 + 0.0006 \times S + 1.12, & S \in [10.69, 910.3] \end{cases}$$

$$\mu_{off}(S) = \begin{cases} -0.005 \times S^3 + 0.09 \times S^2 - 0.57 \times S + 0.09, & S \in [0.72, 10.29] \\ -3e - 9 \times S^3 + 5e - 6 \times S^2 - 0.003 \times S - 1.19, & S \in [10.29, 910.3] \end{cases}$$

$$\sigma_{off}(S) = \begin{cases} -0.0002 \times S^3 + 0.007 \times S^2 - 0.08 \times S + 1.03, & S \in [0.72, 27.56] \\ -1e - 9 \times S^3 + 2e - 6 \times S^2 - 0.001 \times S + 0.65, & S \in [27.56, 910.3] \end{cases}$$

По причине громоздкости коэффициенты были округлены до первых значащих цифр.

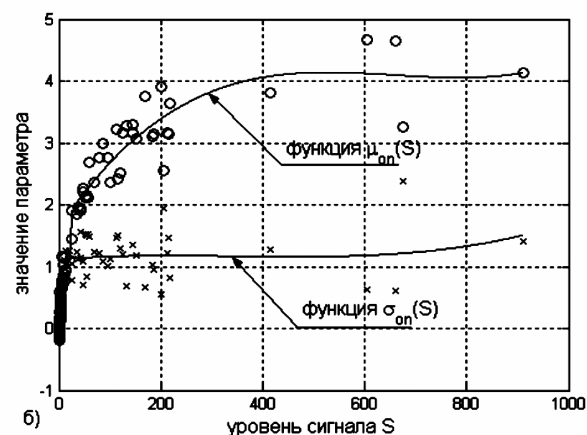
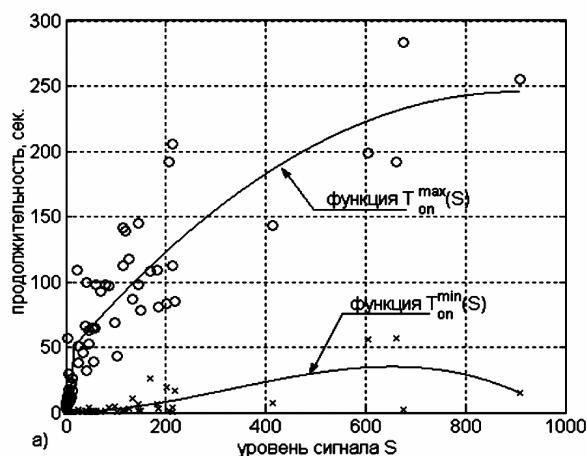


Рис. 7. Аппроксимация экспериментальных данных
а) максимальное и минимальное значения интервала речи при заданном S – функции $T_{on}^{min}(S)$ и $T_{on}^{max}(S)$
б) параметры логнормального распределения интервала речи при заданном S – функции $\mu_{on}(S)$ и $\sigma_{on}(S)$

Значение переменной S изменяется в пределах от 0.72 до 910.3. $S = 0.72$ соответствует нижнему граничному уровню слышимости и минимальному объему передаваемых данных. В противоположность этому значение $S = 910.3$ соответствует практически полному отсутствию пауз и, как следствие, максимально возможной нагрузки на сеть. При отсутствии постороннего шума и ручной настройки чувствительности микрофона (присутствие сигнала только в присутствии речи) значение S составляет в среднем, как правило, 3.6.

В соответствии с полученными результатами в среде моделирования J-sim было проведено моделирование нагрузки, создаваемой на сеть при различных значениях уровня кодируемого сигнала S . При значении $S = 3.6$ ошибка составила 0.2%. Максимальная ошибка 7.4% была получена при значении $S = 0.72$. Значительная разница в значениях ошибки объясняется крутизной кривых, полученных при аппроксимации, и малым количеством экспериментальных данных в данной области. При необходимости положение можно поправить дополнительными измерениями.

Заключение

Построенная статистическая модель трафика с достаточно высокой степенью точности соответствует опытным данным и может быть полезна при оптимизации IP-сети, осуществляющей передачу речи. Логичным продолжением данной работы является построение статистической модели трафика при диалоге.

Литература

1. A. Estepa, R. Estepa, and J. Vozmediano A New Approach for VoIP Traffic Characterization, IEEE communications letters, vol. 8, no. 10, october 2004
2. Leland W. E., Taqqu M. S., Willinger W., Wilson D. V. On the self-similar nature of Ethernet traffic, IEEE Transactions on networking. – Vol. 12. – 1994. - № 1. – P. 2-15.
3. RFC 1889 A Transport Protocol for Real-Time Applications H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, January 1996
4. Сахаров А. В. Анализ частотных свойств трафика. Сборник докладов Всероссийской научно-технической конференции, посвященной 60 летию Победы в Великой Отечественной войне и 110 летию изобретения радио А. С. Поповым – Нижний Новгород 2005
5. Сахаров А.В. Задержка в стеке протоколов UDP/IP/ETHERNET. Сборник трудов по работе 2-го Международного форума «Актуальные проблемы современной науки», Самара 2006

УДК 621.865.8: 534.78

ПРОБЛЕМЫ РАСПОЗНАВАНИЯ СЛИТНОЙ РЕЧИ

Леонovich А.А.

Введение

Голосовой интерфейс это тема, которая на протяжении последних пятидесяти лет привлекает внимание ученых и инженеров всего мира. Голосовой интерфейс на языке пользователя – это наилучшее решение, поскольку речь – это наиболее натуральная, удобная, эффективная и экономичная форма человеческого взаимодействия.

Речевой ввод объединяет множество различных технологий и приложений. В некоторых случаях важно не понимание лингвистического содержания, а идентификация говорящего или языка, на котором происходит общение. Идентификация говорящего в свою очередь может включать в себя получение некоторых специфических параметров, определяющих данного пользователя: пола, возраста или местности проживания по характерному диалекту.

Однако в основе технологии речевого ввода в первую очередь лежит процесс распознавания речи. В соответствии с общепринятым определением, распознавание речи – это процесс преобразования акустического сигнала, полученного с микрофона или телефона, в набор слов. Для таких приложений как командные системы, системы ввода данных или обработки документации, распознанные слова могут сами по себе быть конечным результатом. Однако в некоторых случаях, полученные после распознавания данные могут подвергаться дальнейшей лингвистической обработке для получения форматированного текста либо для достижения понимания речи машиной.

Одним из наиболее сложных аспектов в разработке систем машинного распознавания речи является широкая междисциплинарность задачи. При работе над такими системами затрагиваются вопросы теории обработки сигналов, математического анализа, лингвистики, теории коммуникаций, физиологии и в некоторых случаях психологии. И для построения успешной системы распознавания необходимо рассмотреть такой круг дисциплин, который один человек охватить не в состоянии. Следовательно, для разработчика становится особенно важно понимание основ речевого распознавания без необходимости быть экспертом в каждой из сторон проблемы.

Распознавание речи - это не простая задача, особенно усложненная тем, что существует множество источников вариативности, связанных с речевым сигналом:

1. Акустическая реализация фонем, как наименьших звуковых единиц, составляющих слово, в большой степени зависит от контекста, в котором фонема появляется.

Описываются характеристики систем распознавания речи и их классификация. Обсуждаются методы сегментации и проблемы распознавания речи. Приводятся сравнительные характеристики методов сегментации.

В потоке речи звуки видоизменяются под влиянием соседних фонем, причем иногда звуки накладываются друг на друга или вообще выпадают: *говорить* — [г'вар'ит'].

2. Акустические изменения могут быть вызваны влиянием окружающей среды, а так же характеристиками и позицией приемника речевого сигнала.

3. Физическое и эмоциональное состояние диктора, темп или качество произношения также могут вносить свой вклад в изменчивость речи.

4. Различия в социолингвистическом окружении, диалекте и объеме речевого тракта способствуют возникновению междикторской вариативности.

Классификация систем распознавания речи

Системы распознавания речи могут характеризоваться множеством параметров:

- по типу речи различают системы распознавания изолированных слов и слитной речи;
- по стилю речи: речь может быть спонтанной, либо зачитанной;
- по типу диктора: дикторозависимые и дикторонезависимые;
- по размеру словаря: с маленьким (< 100 слов) или большим словарем (> 10 000 слов).

Системы распознавания изолированных слов требуют от пользователя внесения кратких пауз в речи между словами. Такие системы давно себя зарекомендовали на коммерческом рынке подобных продуктов и показывают достаточно высокие результаты, в частности, в составе командных систем [1]. Однако, данный способ речевого ввода уже является не натуральным для человека, так как требует искусственного изменения потока речи. Это приводит к замедлению процесса ввода информации и быстрой усталости диктора. Соответственно, системы распознавания слитной речи являются более предпочтительными, хотя создание подобных систем - задача во много раз более сложная.

Особые условия на процесс распознавания слитной речи накладывает ее стилистика – это может быть зачитанный текст научной статьи, художественное произведение или часть спонтанного диалога. В данном случае обработка спонтанной речи будет требовать значительно больших усилий и ресурсов, так как такая речь носит произвольный характер, часто не подчиняется правилам языка и полна эмоциональной окраски и междометий [2].



Некоторым системам перед началом работы с новым пользователем необходимо пройти предварительный этап обучения, т.е. каждый диктор должен предоставить образцы своей речи для дальнейшей работы. Это делает подобные системы дикторозависимыми, что в некоторой степени ограничивает и усложняет их использование. Но в то же время, это позволяет повысить качество распознавания, а в ряде случаев данное ограничение может выступать в качестве меры безопасности, осуществляя контроль доступа.

При проектировании систем распознавания речи важно учитывать область применения: либо это будет система диктовки текста, работающая со словарем, содержащим несколько тысяч записей, либо командная система, обрабатывающая несколько десятков слов. Распознавание для систем с большим словарем, многие из слов которого могут звучать достаточно схоже, требует применения особых комплексных алгоритмов поиска и обработки информации.

Сейчас можно сказать, что наиболее привлекательным, с точки зрения пользователя, становится создание дикторонезависимой системы распознавания слитной речи, работающей с большим словарем, при этом рассчитанной на обработку спонтанной речи. Однако, для разработчика данная проблема является наиболее трудной в силу вышеизложенного. Но не смотря на это, создание подобной системы в настоящее время является первостепенной задачей в теории речевого ввода.

Речевая единица сегментации

Деление непрерывной речи на элементарные единицы это одна из наиболее сложных задач в процессе распознавания слитной речи. Обычно, данная проблема подразделяется на две независимые подзадачи:

- преобразование речевого сигнала в строку дискретных минимальных единиц речи (фонем) с последующей классификацией;
- разделение полученной строки на значимые сегменты (слова, или в более общем смысле, лексические единицы).

Для большинства языков, первая задача уже является достаточно трудной, в частности, из-за коартикуляции, когда происходит взаимовлияние соседних звуков при произнесении. Звуки накладываются, создавая переходные участки как внутри слов, так и на стыках смежных, что становится особенно проблематичным для второй подзадачи. Примером могут служить фразы: *Про силу вы ли говорили?* – *Просил, увы ли говорили...* Поскольку в нормальной разговорной речи не существует заметных пауз между словами, становится достаточно не просто правильно выделить лексемы.

Сейчас большинство систем распознавания слитной речи с большим словарем для моделирования элементарной акустической единицы используют дифон или его контекстно-зависимый вариант – трифон [3-5]. Такой выбор обусловлен тем, что дифон – это звуковая единица, имеющая протяженность от середины одного звука до середины последующего. Считается, что речевой сигнал содержит стационарные участки звуков, независимые от влияния соседних, т.е. не подверженные коартикуляционному эффекту. В середине такого участка и проводится граница дифона. Следовательно дифон сохраняет информацию, хранящуюся в переходном участке между фоне-

мами, которая, как было доказано, является полезной [6].

Однако, универсальность использования дифонов для представления речи также можно поставить под вопрос, поскольку эти единицы часто не отражают все сложности речевого сигнала. Во-первых, коартикуляционные эффекты обычно широко растянуты по времени и соответствующие временные зависимости не могут быть переданы дифонами, которые несут характер кратковременного сегмента. Во-вторых, само использование дифонов основано на упрощении, что слово состоит из набора фонем. В рамках такого представления различия в произношении могут быть выражены только заменой, вставкой или удалением фонем. Более того, такое описание ограничивает возможности эффективного использования фонетических зависимостей более высокого порядка, например, относящихся к слоговой структуре слова [7].

Одним из возможных решений данной проблемы может быть выбор акустической единицы, которая содержит спектральную и временную информацию в себе. Наиболее очевидными кандидатами для этих целей кажутся слово и слог. В частности, правильность использования слога подтверждается исследованиями в области фонетики и восприятия речи человеком [8,9]. У слога существует так называемая *произносительная неделимость*, доказанная экспериментально – как бы ни была замедлена речь, она никогда не распадается на промежутки меньшие чем слог.

Тем не менее, выбор более длинной единицы сегментации речевого сигнала для распознавания речи не является чем-то новым: модели, использующие слово как минимальный сегмент, широко используются в приложениях с ограниченным словарем, таких как распознаватели цифр или набора команд. Использование слоговых моделей также предлагалось ранее [10,11]. Но подобный подход приводит к возникновению новой проблемы при работе с системами с большим словарем – это недостаток данных для обучения системы, создания начальных эталонов. С повышением длины элементарной единицы речи возрастает количество вариантов данного элемента. И в случае выбора слова в качестве такой меры, число вариаций увеличивается до недопустимых пределов. Применительно к русскому языку, существуют разработки, в которых применен метод деления речи на морфемы, что значительно уменьшает размер словаря. Однако, вследствие того, что морфема не является акустической единицей, данный способ требует дополнительных усилий в плане построения морфем из сегментов, полученных после обработки речевого сигнала, что в свою очередь снижает эффективность алгоритма [12]. Существуют разработки, которые пытаются преодолеть данную проблему, путем совместного иерархического использования нескольких видов речевых сегментов: слов, слогов и фонем [13,14]. И все же, даже в случае применения таких иерархических систем, задача однозначной сегментации речевого сигнала все еще окончательно не решена.

Методы сегментации речевого сигнала

Прежде чем непосредственно приступить к распознаванию речи, в первую очередь необходимо создать базу эталонов речевых единиц. И здесь снова встает задача сегментации речевого сигнала. Самым первым и наиболее простым способом получения элементарных единиц явля-

лась ручная обработка записанных фраз. Эксперты выполняли сегментацию основываясь на спектрограммах, кривых энергии, интонациях и других приемах, используемых в речевом анализе. Этот способ обладает некоторыми преимуществами перед автоматическими методами – опытные лингвисты могут с большой точностью, анализируя многие факторы, определить границы сегментов. Но в то же время, такая процедура очень трудоемка и требовательна к ресурсам, что делает данный метод применимым только в ограниченных случаях.

Одним из основных направлений в области распознавания слитной речи является применение аппарата скрытого Марковского моделирования (СММ), где речевой сигнал представляется набором состояний с некоторыми вероятностями перехода между ними. Отсюда возник метод автоматической сегментации речи.

Рассмотрим схему системы распознавания слитной речи на основе СММ, работающей со звуковыми единицами, длиной меньше слова (рис. 1). Этап построения модели слова по сути является некоторым вариантом сегментации – здесь происходит подбор эталонов для входного вектора признаков. Соответственно, выполнив некоторые модификации для уменьшения вычислительной сложности полного процесса распознавания, данный процесс можно применить для автоматического поиска границ элементарных единиц речи [15].

Однако, классический процедура, основанная на СММ, требует полного транскрибирования входного потока речи, другими словами необходимо пройти полный процесс речевого распознавания.

Другой класс алгоритмов рассматривает речь только с позиций обработки акустических сигналов. В простейшем случае, происходит разделение речевого тракта на сегменты, где в качестве границ выступают паузы [16]. Различные характеристики служат для определения границ данной речевой единицы в речевом тракте: энергетические параметры [17-19], резонансные частоты и частота

основного тона [20], а также разнообразная просодическая информация [21].

Существуют и иные альтернативные подходы к сегментации речи, не обязательно ограниченные техникой речевой обработки, но и применяющие обобщенные статистические алгоритмы. Среди них: нейросетевой метод [22, 23], статистическое моделирование [24] и динамическое программирование [25,26] (таблицу 1).

Сегментация речевых баз данных – одна из наиболее актуальных задач. В этом случае соответствующая транскрипция обычно известна заранее, и быстрые алгоритмы обработки не являются необходимыми. С другой стороны, если сегментация является частью процесса распознавания речи, то при отсутствии какой либо информации о лексическом содержании сигнала, необходима обработка в реальном времени. Следовательно, характеристики подходящего алгоритма должны определяться исходя из задач итогового приложения.

Заключение

Следует сказать, что при всем многообразии существующих подходов и методов добиться максимально высоких показателей так и не удалось. Хотя исследования и эксперименты в области речевого распознавания проводятся уже не один десяток лет, ясно, что все еще остается множество нерешенных проблем. Так и не решена задача построения распознавателя, обладающего характеристиками на уровне человеческого восприятия речи.

Однако уже сейчас можно обозначить первостепенные проблемы в теории распознавания речи и, в частности, для ее непрерывного варианта:

1. Выбор наиболее подходящей речевой единицы, для сегментации и представления речевого потока. Это особенно актуально для систем, работающих с большим словарем.
2. Построение автоматического метода поиска границ сегментов речевых единиц в потоке слитной речи. Существующие техники все еще в большой степени уступают способу ручной сегментации.

Таблица 1

Характеристики методов сегментации речи (по отношению к ручной разметке, которая считалась эталонной)

Метод сегментации	Метод СММ [15]	Метод спектр. характеристик [18]	Нейросетевой метод [23]	Метод стат. моделирования [24]	Байесовский метод предсказаний [27]	Совместный метод [28]
% правильных границ (допуск ± 15 мсек)	85.5%	85%	83%	84.5%	-	-
% правильных границ (допуск ± 20 мсек)	-	-	94.19%	89.51%	76.21%	90.22%

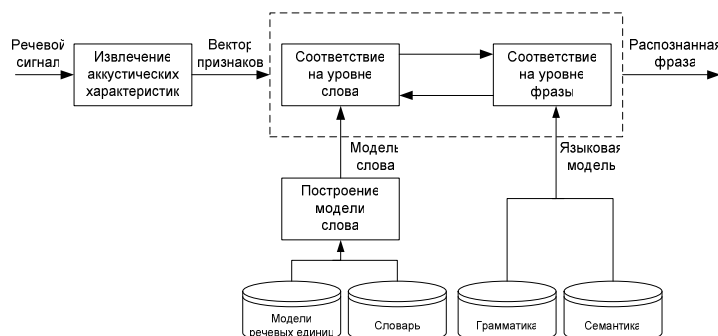


Рис. 1 Схема системы распознавания слитной речи



В рамках работ в области сегментации речи проводятся эксперименты в плане применения аппарата вейлет-преобразования для поиска границ речевых сегментов [30]. Последние результаты показали, что наиболее удобной единицей сегментации является слог, так как информация о фонемах распределена на всем протяжении данной единицы. Полученные данные хорошо согласуются с выводами исследователей кафедры фонетики Санкт-Петербургского Университета [6,7]. Это, в свою очередь, позволило сделать вывод о целесообразности проведения дальнейших экспериментов именно в направлении слоговой сегментации непрерывного потока речи, что сейчас и осуществляется.

Литературы

1. Леонович А.А. Современные технологии распознавания речи // Международный семинар Диалог'2005 по компьютерной лингвистике и ее приложениям. Электронная публикация на сайте конференции. www.dialog-21.ru
2. Weintraub M., Taussig K., Hunnicke-Smith K. & Snodgrass A. Effect of Speaking Style on LVCSR Performance. Proc. Int. Conf. on Spoken Language Proc., supplement, 1996.
3. Watanabe, S., Sako, A., Nakamura, A. Automatic determination of acoustic model topology using variational Bayesian estimation and clustering for large vocabulary continuous speech recognition. Audio, Speech and Language Processing, IEEE Transactions on. Volume 14, Issue 3, May 2006, pp. 855-872.
4. Sakti, S., Markov, K., Nakamura, S. Incorporation of Pentaphone-Context Dependency Based on Hybrid HMM/Bn Acoustic Modeling Framework. Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. Volume 1, 14-19 May 2006, pp. 1177-1180.
5. Jurafsky D., Ward W., Jianping Z., Herold K., Xiuyang Y., and Sen Z. What kind of pronunciation variation is hard for triphones to model? Proc. ICASSP-2001, Salt Lake City, Utah, USA, May 8-11. 2001, vol. 1, pp. 577-580.
6. Бондарко Л.В., Кузнецов В. И., Скрелин П.А., Шалонина К. Б. Звуковая система русского языка в свете задач компилятивного синтеза // Бюллетень фонетического фонда русского языка. № 6, май 1997.
7. Белявский В. М., Светозарова Н. Д. Слоговая фонетика и три фонетики Л.В. Щербы. Статья, расположенная по адресу <http://www.auditech.ru/doc/cherba.htm>
8. Schiller N.O., Meyer A.S., and Levelt W.J.M. The syllabic structure of spoken words: Evidence from the syllabification of intervocalic consonants. Language & Speech, vol. 40, 1997, pp. 103-140.
9. Зиндер Л.П. Общая фонетика. М., 1979.
10. Greenberg S. Speaking in shorthand – a syllable-centric perspective for understanding pronunciation variation. Speech Communication, vol. 29, pp. 159-176, 1999.
11. Jones R.J., Downey S., and Mason J.S. Continuous speech recognition using syllables," in Proc. Eurospeech-97, Rhodes, Greece, Sept 22-25. 1997, vol. 3, pp. 1171-1174.
12. Ронжин А.Л., Карпов А.А., Ли И.В. Система автоматического распознавания русской речи SIRIUS.// Санкт-Петербургский институт информатики и автоматизации РАН.
13. Hämmäläinen, A., de Veth, J., and Boves, L. Longer-Length Acoustic Units for Continuous Speech Recognition, in Proc. EUSIPCO-2005, Antalya, Turkey, Sep 4-8, 2005.
14. Messina, R. and Jouvett D. Context dependent "long units" for speech recognition, in Proc. ICSLP-2004, Jeju Island, Korea , Oct 4-8, 2004, pp. 645-648.
15. Brugnara F., Falavigna D., and Omologo M. Automatic segmentation and labeling of speech based on hidden Markov models. Speech Communication, vol. 12, no. 4, pp. 357-370, 1993.
16. Pfeiffer S. Pause Concepts for audio Segmentation at Different Semantic Levels. ACM Multimedia, 2001, pp. 187-193.
17. Milone D.H., Merelo J.J., Rufiner H.L. Evolutionary algorithm for speech segmentation. Evolutionary Computation, 2002. CEC '02, vol. 2, 12-17 May 2002, pp. 1115-1120.
18. Сорокин В.Н., Цыплихин А.И. Сегментация и распознавание гласных. Информационные процессы, том 4, № 2, 2004, стр. 202-220.
19. Ермоленко Т.В., Шевчук В.В. Алгоритмы сегментации с применением быстрого вейвлет-преобразования. // Международный семинар Диалог'2003 по компьютерной лингвистике и ее приложениям. Электронная публикация на сайте конференции. www.dialog-21.ru
20. Wendt C., Petropulu A.P. Pitch determination and speech segmentation using the discrete wavelet transform. Circuits and Systems, 1996. ISCAS '96, 'Connecting the World', vol. 2, 12-15 May 1996, pp. 45 – 48.
21. Dong Wang, Lie Lu, Hong-Jiang Zhang. Speech segmentation without speech recognition. Multimedia and Expo, 2003. ICME '03. Proceedings, vol. 1, 6-9 July 2003, pp. 405-408.
22. Vorstermans A., Martens J.-P. and Van Coile B. Automatic segmentation and labeling of multi-lingual speech data. Speech Communication, vol. 19, pp. 271-293, 1996.
23. Toledano D.T. Neural network boundary refining for automatic speech segmentation. Acoustics, Speech, and Signal Processing, 2000. ICASSP '00, vol. 6, 5-9 June 2000, pp. 3438-3441.
24. Pauws S., Kamp Y. and Willens L. A hierarchical method of automatic segmentation for synthesis applications. Speech Communications, vol. 20, pp. 207-220, 1996.
25. Bajwa R.S., Owens R.M., Kelliher T.P. Simultaneous speech segmentation and phoneme recognition using dynamic programming. Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings, vol. 6, 7-10 May 1996, pp. 3213-3216.
26. Sharma, M., Mammone R. "Blind" speech segmentation: automatic segmentation of speech without linguistic knowledge. Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on, vol. 2, 3-6 Oct. 1996, pp. 1237-1240.
27. Ming Liu, Huang T.S. A Bayesian Predictive Method for Automatic Speech Segmentation. Pattern Recognition, 2006. ICPR 2006. 18th International Conference on, vol. 4, 20-24 Aug. 2006, pp. 290-293.
28. Runqiang Yan, Yiqing Zu, Yisheng Zhu. Automatic Speech Segmentation Combining an HMM-Based Approach and Recurrence Trend Analysis. Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings, vol. 1, 14-19 May 2006, pp. 797-800.
29. Rabiner L. and Juang B.-H. Fundamentals of Speech Recognition, Prentice Hall, Englewood Cliffs, New Jersey, 1993.
30. Леонович А.А. Выбор вейвлет базиса для алгоритма автоматической сегментации речи. Издательство УРСС. Коллектив авторов. Первая Международная конференция «Системный анализ и информационные технологии» САИТ-2005 (г., Переславль-Залесский, Россия): Труды конференции. 2005.

УДК 62-52

РЕГУЛЯРИЗАЦИЯ АЛГОРИТМА ФИЛЬТРАЦИИ КАЛМАНА-БЬЮСИ ПРИ ПЛОХОЙ ОБУСЛОВЛЕННОСТИ КОРРЕЛЯЦИОННОЙ МАТРИЦЫ ШУМА

Ашинянц Р.А., Морозова Т.Ю.

Введение

Успешное решение задач анализа и синтеза систем автоматического управления, повышение эффективности их работы связано с получением возможно более полной информации о сигналах, воздействующих на объект управления. В большинстве случаев точное описание входных воздействий затруднено недоступностью их измерения. В тех же случаях, когда возможно измерение, для описания необходимо учитывать искажения, вносимые измерительными устройствами, внешними помехами. С аналогичной ситуацией сталкиваются и в тех случаях, когда для успешного ведения некоторого процесса необходимы знания об определенных компонентах вектора состояния объекта, которые в силу их физической природы или специфики процесса не измеримы.

В указанных случаях возникает задача определения оптимальных в некотором смысле оценок полезных сигналов (входных воздействий, компонент вектора состояния объекта) косвенным путем: по искаженным помехами измерениям входных сигналов.

Широкое применение в решении задач восстановления полезных труднодоступных сигналов получили оптимальные фильтры Калмана [1], [2]. Прямое применение методов оптимальной фильтрации Калмана для решения задач восстановления сигналов затруднено тем обстоятельством, что чаще всего мы находимся в условиях недостаточности априорной информации, когда неизвестны некоторые параметры процесса и характеристики шумов измерения. В то время как реализация алгоритмов фильтрации Калмана связана с необходимостью полной априорной информации о полезном сигнале и статистике воздействий.

Возможно ли, используя основные преимущества метода, уменьшить, если не устранить, влияние тех факторов, которые определяют недостатки? В частности, возможно ли определение весовых коэффициентов фильтра сделать независимым от решения уравнения для дисперсий и возможно ли устранение расходимости оценок фильтрации при плохой обусловленности или нулевой матрице интенсивностей помех измерения?

Следует обратить внимание, что цель работы – получение эффективных, несмещенных оценок и алгоритма оценивания, который был бы доступен для инженерной практики.

Постановка задачи

Пусть линейная динамическая система (объект) описывается системой дифференциальных уравнений n -го порядка с постоянными коэффициентами

Представлен обобщенный алгоритм фильтрации Калмана - Бьюси, который применим как для случая отсутствия помех, так и при зашумленных измерениях.

$$\frac{dx}{dt} = Fx(t) + Gu(t), \quad x(t_0) = x_0, \quad (1)$$

где $x(t)$ – n – мерный вектор координат состояний; $u(t)$ – l –мерный вектор воздействий; F , G – матрицы коэффициентов размером соответственно $[n \times n]$ и $[n \times l]$; x_0 – вектор начальных условий.

Выходной сигнал $y(t)$ определяется линейной комбинацией координат вектора состояния

$$y(t) = Hx(t),$$

где $y(t)$ – m – мерный вектор, H – матрица наблюдений размера $[m \times n]$.

Наблюдаемый в момент времени сигнал $z(t)$ представляет сумму выходного сигнала и помехи

$$z(t) = y(t) + v(t), \quad (2)$$

где $v(t)$ – m – мерный вектор шума интенсивности R .

Тогда оптимальная оценка состояний системы в смысле минимума среднеквадратической ошибки определяется решением системы дифференциальных уравнений

$$\frac{d\hat{x}}{dt} = F\hat{x}(t) + K(t)[z(t) - H\hat{x}(t)], \quad \hat{x}(t_0) = 0 \quad (3)$$

В алгоритмах фильтрации Калмана оптимальные весовые коэффициенты $K(t)$ системы фильтрации (3) определяются решением уравнений для дисперсий

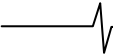
$$\frac{dP}{dt} = FP(t) + P(t)F^T(t) - P(t)H^T R^{-1}HP(t) + GQG^T \quad (4)$$

$$K(t) = P(t)H^T R^{-1}, \quad (5)$$

где $P(t)$ – матрица дисперсий ошибок фильтрации, R – матрица интенсивностей помехи измерения, H – матрица наблюдений, Q – матрица интенсивностей белого шума.

Предположим, что матрица интенсивностей R помех измерения плохо обусловлена или нулевая. В этом случае имеет место некорректная постановка задачи. Задача состоит в отыскании такого процесса $\hat{x}^\alpha(t)$, который был бы близок в некотором смысле процессу $x(t)$ при плохо обусловленной матрице R .

Пусть линейный динамический объект описывается системой дифференциальных уравнений с постоянными коэффициентами (1). При выполнении условия (2) будем считать заданной структуру системы восстановле-



ния (3). Откажемся от решения системы уравнений (4) с целью определения элементов матрицы $K(t)$. Тогда система (3) не замкнута. Естественно, что при некотором значении элементов матрицы $K(t)$ достигается минимум среднеквадратической ошибки.

Необходимо получить алгоритм настройки параметров восстановления состояний системы (3), чтобы достигался минимум среднеквадратической ошибки.

Известно, что невыполнение условия теоремы Калмана об управляемости может привести к неустойчивости системы фильтрации. При этом дисперсия ошибки восстановления при $t \rightarrow \infty$ стремится к нулю, и расходимость оценок фильтрации обуславливается накоплением ошибок за счет моделирующей системы. Однако это не единственная причина расходимости оценок фильтрации. Одной из причин расходимости оценок, известной как проблема расходимости фильтра Калмана, является неадекватность модели реальной наблюдаемой системе, которая объясняется грубой аппроксимацией характеристик реальной системы и преднамеренным уменьшением числа наблюдаемых координат. Кроме того, очень часто в модели принимают входное воздействие и шум измерения независимыми белыми шумами, в то время как они являются взаимно коррелированными. Такая проблема рассматривалась в [3].

Отметим еще одну проблему расходимости оценок фильтра Калмана. Так же как и в работе [4], примем входное воздействие и шум измерения независимыми белыми шумами. Поскольку при определении матрицы коэффициентов усиления фильтра Калмана применяется операция обращения матрицы интенсивностей помехи измерения R в формуле (5), то необходимо решать задачу повышения устойчивости алгоритма обращения матрицы R .

Решение задачи методом регуляризации

При решении данной задачи воспользуемся методами регуляризации некорректно поставленных задач [5], [6].

В выражении (5) умножим обе части справа на R , получим

$$KR = PH^T \quad (6)$$

Введем обозначения

$$K^T = z - \text{матрица } [m \times n],$$

$$HP^T = u - \text{матрица } [m \times n], \quad (7)$$

$$R^T = A - \text{матрица } [m \times m].$$

Уравнение (6) принимает вид

$$Az = u. \quad (8)$$

Представим матрицы z и u в уравнении (8) в виде

$$z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{bmatrix}, \quad u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix},$$

где $z_1, z_2, \dots, z_m, u_1, u_2, \dots, u_m$ - строки матрицы z и u соответственно.

Пусть решения z системы (8) принадлежат метрическому пространству $\bar{Z}$, а свободные члены u линейному

подпространству $\bar{U}$. Предположим, что $u^0 \in \bar{U}$ и элементы матрицы A определены с точностью δ : $\tilde{u} = \tilde{u}^\delta, \tilde{A} = \tilde{A}^\delta$.

Устойчивое решение уравнения (8) будет получено, если

$$\rho_u(u^0, \tilde{u}^\delta) \leq \delta, \quad \rho_z(z^0, \tilde{z}) < \varepsilon,$$

где $\rho_z(z^0, \tilde{z})$ - расстояние, мера уклонения на линейном пространстве, $\varepsilon > 0$; ε - любое число, представляющее заданную точность решения z .

Поскольку задача определения z по уравнению (8) поставлена некорректно, $\tilde{z}^\delta$ не может быть определено как точное решение уравнения

$$A\tilde{z}^\delta = \tilde{u} \quad (9)$$

Меру отклонения определим с помощью норм

$$\|A - \tilde{A}\| \leq \delta, \quad \|u - \tilde{u}\| \leq \delta$$

В качестве норм u и z примем

$$\|u\| = \sqrt{u_1^2 + u_2^2 + \dots + u_m^2}, \quad \|z\| = \sqrt{z_1^2 + z_2^2 + \dots + z_m^2},$$

где u_i^2, z_i^2 - скалярные квадраты.

Для матрицы A

$$\|A\| = \sqrt{\sum_{ij} a_{ij}^2}$$

Решение $\tilde{z}^\delta$ уравнения (9) получим с помощью параметрического функционала $M^\alpha[\tilde{u}, \tilde{A}, \delta]$, в котором параметр α связан с точностью δ [5]. При этом решения $\tilde{z}^\delta$ дают приближения к точному решению: z^0 $\|z^0 - \tilde{z}^\delta\| \leq \varepsilon$, если $\|u - \tilde{u}\| \leq \delta, \|A - \tilde{A}\| \leq \delta$.

Рассмотрим параметрический функционал, введенный в [5]

$$M^\alpha[z, \tilde{u}, \tilde{A}] = \|\tilde{A}z - \tilde{u}\|^2 + \alpha\|z\|^2. \quad (10)$$

Используя методы вариационного исчисления, для заданных $\tilde{A}, \tilde{u}$ находим единственный элемент z^α , реализующий минимум функционала (10).

В развернутом виде функционал (10) можно представить в виде

$$\sum_j \left(\sum_{i=1}^m a_{ij} z_j - u_i \right)^2 + \alpha \sum_j z_j^2 = M^\alpha, \quad i=1,2,\dots,m, \quad j=1,2,\dots,n \quad (11)$$

Приравнявая частные производные функционала (11) по z_j нулю, получим $m \times n$ уравнений относительно z_j :

$$\sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} z_j - u_i \right) a_{is} + \alpha z_s = 0, \quad s=1,2,\dots,n$$

После элементарных преобразований получим регуляризованные уравнения в матричной форме

$$[A^T A + \alpha I]^T z = A^T u. \quad (12)$$

Учитывая обозначения (7), уравнение (12) принимает окончательный вид

$$K^\alpha = P^\alpha H^T R^T [R^T R + \alpha I]^{-1}, \quad (13)$$

где I - единичная матрица $[m \times m]$. Сопоставляя уравнения (5) и (4), последнее можно переписать следующим образом:

$$\frac{dP}{dt} = FP + PF^T - KHP + GQG^T, \quad P(t_0) = P_0$$

Подставим в это уравнение значение K^α из (13)

$$\frac{dP^\alpha}{dt} = FP^\alpha + P^\alpha F^T - P^\alpha H^T R^T [RR^T + \alpha I]^{-1} HP^\alpha + GQG^T. \quad (14)$$

Для стационарного случая (F, H, R, G, Q - матрицы с постоянными элементами) установившееся решение задачи оценки восстановления координат исходного процесса может быть получено приравняв нулю правой части уравнения (14) и решением системы алгебраических уравнений второй степени.

Уравнение (14) представляет регуляризованное уравнение дисперсий ошибок фильтрации Калмана-Бьюси.

Рассмотрим вопрос о выборе параметра α . В основной теореме работы [5] доказано, что для любого $\varepsilon > 0$ можно найти такое δ_0 , что норма разности регуляризованного и точного решений окажутся меньше ε , если заданная точность $\delta \leq \delta_0$. При этом параметр регуляризации α удовлетворяет условию

$$\frac{\delta^2}{\varepsilon(\delta)} \leq \alpha \leq \alpha_0(\delta), \quad (15)$$

где $\varepsilon(\delta)$ и $\alpha_0(\delta)$ какие-либо убывающие функции δ , стремящиеся к нулю при $\delta \rightarrow 0$.

Отсюда следует, что при увеличении точности исходных данных регуляризованное решение стремится к точному.

Выбор параметра α , удовлетворяющий условию (15), очень широк.

Другой способ определения параметра α заключается в следующем. Среди множества решений, соответствующих различным значениям параметра α , можно найти такое, которое минимизирует функционал

$$f(\alpha, \delta) = \left\| \delta - \|Az^\alpha - \tilde{u}\| \right\|,$$

где z^α - элемент, реализующий минимум функционала (10).

Приведенные способы не исчерпывают возможности выбора параметра регуляризации α . Различные методы выбора можно найти в работах [7], [8]. Заметим, что метод регуляризации в достаточной мере трудоемкий, нерегулярный. Так, каждый раз при изменении точности исходных данных необходимо заново отыскивать параметр регуляризации.

Заметим, что полученное выражение (13) для матрицы коэффициентов усиления системы восстановления имеет частное применение. Существенным ограничением в применении этой формулы является отличие от нуля матрицы интенсивностей шумов измерения $R \neq 0$. Здесь следует оговорить и вид матрицы наблюдений H .

Рассмотрим случай, когда матрица H имеет диагональный вид.

Рассмотрим двумерный объект с матрицей наблюдений

$$H = \begin{bmatrix} h_{11} & 0 \\ 0 & h_{22} \end{bmatrix}.$$

Это означает, что наблюдаются две координаты. В этом случае отсутствие шумов измерения ($z_1 = h_{11}x_1, z_2 = h_{22}x_2, R = 0$) вообще снимает про-

блему фильтрации. Однако, если матрица наблюдений имеет более общий вид $[m \times n]$ и матрица интенсивностей нулевая, уравнения для оценок Калмана теряют смысл. Предположим, что матрица наблюдений имеет вид $H = [1 \ 0 \ 0]$, то есть наблюдается одна координата трехмерного объекта, при этом $R=0$. Восстановление неизмеримых координат по точному измерению выходной координаты означает, что необходимо восстановить ее первую и вторую производные. Дифференцирование же приведет к еще большим ошибкам. Следовательно, выражение (13) применимо только при наличии шумов измерения.

Однако можно получить регуляризованный алгоритм фильтрации, который применим как для случая отсутствия помех ($R=0$), так и при зашумленных измерениях ($R \neq 0$). Для этого вместо известного функционала (10) рассмотрим функционал вида

$$M^\alpha = \|Az - u\|^2 + \alpha \sum_{i=1}^n z_i^T A z_i \quad (16)$$

или в развернутом виде

$$M^\alpha = \sum_j \left(\sum_i a_{ij} z_i - u_j \right)^2 + \alpha \sum_i \sum_j a_{ij} z_i z_j.$$

Так же как и в предыдущем случае, приравнявая частные производные по z_j нулю, получим $m \times n$ уравнений вида

$$\sum_j \left(\sum_i a_{ij} z_i - u_j \right) a_{is} + \alpha \sum_i (a_{is} + a_{si}) z_i = 0, \quad s = 1, 2, \dots, m.$$

Учитывая, что $a_{ij} = a_{ji}$, получим уравнение в матричной форме

$$(A^T A + \alpha A^T) z^\alpha = A^T u$$

или в окончательном виде

$$(A + \alpha I) z^\alpha = u. \quad (17)$$

В обозначениях (7) уравнение (17) принимает вид

$$K^\alpha = P^\alpha H^T (R + \alpha I)^{-1} \quad (18)$$

Оптимальная оценка вектора состояний процесса в смысле минимума среднеквадратической ошибки при любой матрице R шума измерения определяется из уравнения

$$\frac{d\hat{x}^\alpha}{dt} = F\hat{x}^\alpha + P^\alpha H^T (R + \alpha I)^{-1} [z - H\hat{x}^\alpha], \quad (19)$$

а матрица дисперсий удовлетворяет уравнению

$$\frac{dP^\alpha}{dt} = FP^\alpha + P^\alpha F^T - P^\alpha H^T (R + \alpha I)^{-1} HP^\alpha + GQG^T. \quad (20)$$

При $\alpha \rightarrow 0$ $\hat{x}^\alpha(t)$ приближается к точной оценке.

Заметим, что, по существу, добавление к диагональным элементам матрицы R величины α означает введение в уравнения наблюдений дополнительного шума с интенсивностью α .

Примеры оценивания координат исходного процесса

Рассмотрим примеры, в которых при различной информации о шуме необходимо получить оценки координат исходного процесса, коэффициенты фильтрации Калмана – Бьюси и элементы матрицы дисперсий ошибок фильтрации.

Пусть процесс описывается системой дифференциальных уравнений

$$\begin{cases} \frac{dx_1}{dt} = -ax_1 + x_2, \\ \frac{dx_2}{dt} = -bx_2 + x_3, \\ \frac{dx_3}{dt} = u. \end{cases} \quad (21)$$

Наблюдаются координаты x_1 и x_2 в присутствии шумов V_1 и V_2

$$\begin{cases} z_1 = h_{11}x_1 + v_1, \\ z_2 = h_{22}x_2 + v_2. \end{cases}$$

Необходимо восстановить координаты исходного процесса, то есть получить их оценки, если V_1 и V_2 коррелированные белые шумы.

Имеем

$$H = \begin{bmatrix} h_{11} & 0 \\ 0 & h_{22} \end{bmatrix}, \quad R = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}, \quad F = \begin{bmatrix} -a & 1 & 0 \\ 0 & -b & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad G = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Система восстановления описывается уравнениями (19):

$$\begin{aligned} \frac{d\hat{x}_1}{dt} &= -a\hat{x}_1 + \hat{x}_2 + k_{11}(z_1 - h_{11}\hat{x}_1) + k_{12}(z_2 - h_{22}\hat{x}_2), \\ \frac{d\hat{x}_2}{dt} &= -b\hat{x}_2 + \hat{x}_3 + k_{21}(z_1 - h_{11}\hat{x}_1) + k_{22}(z_2 - h_{22}\hat{x}_2), \\ \frac{d\hat{x}_3}{dt} &= k_{31}(z_1 - h_{11}\hat{x}_1) + k_{32}(z_2 - h_{22}\hat{x}_2). \end{aligned}$$

В соответствии с (18) и (20) имеем систему уравнений

$$\begin{aligned} \frac{dp_{11}}{dt} &= -2ap_{11} + 2p_{12} - \\ &- \frac{1}{\beta}(h_{11}^2 p_{11}^2 (r_{22} + \alpha) - 2h_{11}h_{22}p_{12}r_{12} + h_{22}^2 (r_{11} + \alpha)p_{12}^2), \\ \frac{dp_{12}}{dt} &= -(a+b)p_{12} + p_{22} + p_{13} - \\ &- \frac{1}{\beta}(h_{11}^2 (r_{22} + \alpha)p_{11}p_{12} - h_{11}h_{22}r_{12}(p_{12}^2 + p_{11}p_{22}) + h_{22}^2 (r_{11} + \alpha)p_{11}p_{22}), \\ \frac{dp_{13}}{dt} &= -ap_{13} + p_{23} - \\ &- \frac{1}{\beta}(h_{11}^2 p_{11}p_{13}(r_{22} + \alpha) - h_{11}h_{22}r_{12}(p_{13} + p_{23}) + h_{22}^2 p_{22}p_{13}(r_{11} + \alpha)), \\ \frac{dp_{22}}{dt} &= -2bp_{22} + 2p_{23} - \\ &- \frac{1}{\beta}(h_{11}^2 p_{12}^2 (r_{22} + \alpha) - 2h_{11}h_{22}p_{12}p_{22}r_{12} + h_{22}^2 p_{22}^2 (r_{11} + \alpha)), \\ \frac{dp_{23}}{dt} &= -bp_{23} + p_{13} - \\ &- \frac{1}{\beta}(h_{11}^2 p_{12}p_{13}(r_{22} + \alpha) - h_{11}h_{22}r_{12}(p_{13}p_{22} + p_{12}p_{23}) + h_{22}^2 p_{22}p_{23}(r_{11} + \alpha)), \\ \frac{dp_{33}}{dt} &= -\frac{1}{\beta}(h_{11}^2 p_{13}^2 (r_{22} + \alpha) - 2h_{11}h_{22}p_{13}p_{23}r_{12} + h_{22}^2 p_{23}^2 (r_{11} + \alpha)) + Q, \\ k_{11} &= \frac{1}{\beta}(p_{11}h_{11}(r_{22} + \alpha) - p_{12}h_{22}r_{12}), \quad k_{12} = \frac{1}{\beta}(p_{12}h_{22}(r_{11} + \alpha) - p_{11}h_{11}r_{12}), \\ k_{21} &= \frac{1}{\beta}(p_{12}(r_{22} + \alpha)h_{11} - p_{12}h_{22}r_{12}), \quad k_{22} = \frac{1}{\beta}(p_{22}h_{22}(r_{11} + \alpha) - p_{12}h_{11}r_{12}), \\ k_{31} &= \frac{1}{\beta}(p_{13}h_{11}(r_{22} + \alpha) - p_{32}h_{22}r_{12}), \quad k_{32} = \frac{1}{\beta}(p_{23}h_{22}(r_{11} + \alpha) - p_{13}h_{11}r_{12}), \end{aligned}$$

где $\beta = (r_{11} + \alpha)(r_{22} + \alpha) - r_{12}^2$.

Пусть процесс описывается той же системой уравнений (21). Наблюдается координата x_1 $z = x_1$, то есть $H = [1 \ 0 \ 0]$, $R = 0$. Необходимо получить оценки $\hat{x}_2$ и $\hat{x}_3$.

Очевидно, оптимальные оценки $\hat{x}_2$ и $\hat{x}_3$ совпадают с первой и второй производными измеряемого сигнала x_1 , вычисление которых связано с большими ошибками. Применяя (19), (20) и (18), получим

$$\begin{aligned} \frac{d\hat{x}_1}{dt} &= -a\hat{x}_1 + \hat{x}_2 + k_{11}(z - \hat{x}_1), \quad \frac{d\hat{x}_2}{dt} = -b\hat{x}_2 + \hat{x}_3 + k_{21}(z - \hat{x}_1), \quad \frac{d\hat{x}_3}{dt} = k_{31}(z - \hat{x}_1), \\ \frac{dp_{11}}{dt} &= -2ap_{11} + 2p_{12} - \frac{p_{11}^2}{\alpha}, \quad \frac{dp_{12}}{dt} = -(a+b)p_{12} + p_{22} - \frac{p_{11}p_{12}}{\alpha}, \\ \frac{dp_{13}}{dt} &= -ap_{13} + p_{13} - \frac{p_{11}p_{13}}{\alpha}, \quad \frac{dp_{22}}{dt} = -2bp_{22} + 2p_{23} - \frac{p_{12}^2}{\alpha}, \\ \frac{dp_{23}}{dt} &= -bp_{23} + p_{13} - \frac{p_{13}p_{12}}{\alpha}, \quad \frac{dp_{33}}{dt} = -\frac{p_{13}^2}{\alpha} + Q, \\ k_{11} &= \frac{p_{11}}{\alpha}, \quad k_{21} = \frac{p_{12}}{\alpha}, \quad k_{31} = \frac{p_{13}}{\alpha}. \end{aligned}$$

Поскольку $v = 0$, то очевидно, что $P = 0$. Тогда при $\alpha \rightarrow 0$ оценки $\hat{x}_2^\alpha$ и $\hat{x}_3^\alpha$ приближаются к истинным значениям.

Заключение

Таким образом, метод регуляризации не определяет однозначно вид регуляризирующего функционала. Решения одной и той же задачи при выборе различного вида функционалов могут существенно отличаться друг от друга. Одна и та же, казалось бы, часто употребляемая форма функционала приводит к успешным результатам при решении одних задач и к нереализуемым (неуспешным) при решении других.

Значения параметров регуляризации задаются, исходя из дополнительных соображений. Причем для одной и той же задачи можно предложить различные условия выбора параметра регуляризации, конкретный же выбор требует дополнительных исследований.

В заключение отметим, что в статье рассмотрена задача обобщения фильтра Калмана на случай отсутствия шумов ($R = 0$). При этом использовался математический подход, основанный на приведении исходного уравнения (6) к матричному виду (8) и использовании функционала (16) для решения уравнения (8).

Литература

1. Ioffe B.L., Sargent R.W. The design of an in-line control schema for a tubular catalytic reactor. Trans. Inst. Chem. Eng., 1972, №5.
2. Гулько Ф.Б., Новосельцева Ж.А., Смирнов Н.А. О корректности решения задач фильтрации и восстановления сигналов на АВМ. В сб. «Теория автоматического управления», Тр. Всесоюзного совещания по автоматическому управлению, «Наука», 1972.
3. Лифшиц Н.А., Виноградов В.Н., Голубев Г.А. Линейная фильтрация при особенной матрице интенсивностей помех. Изв. АН СССР, Техническая кибернетика, 1969, №3.
4. Калман Р.Е., Бьюси Р.С. Новые результаты в линейной фильтрации и теории предсказания. Техническая механика, т.83, серия Д, №1, 1961.
5. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. М.: Наука. 1986.
6. Петров Ю.П. Корректные, некорректные и промежуточные задачи с приложениями. СПб.: Политехника, 2003.
7. Морозов В.А. О регуляризации некорректно поставленных задач и выборе параметра регуляризации. ЖВМ и МФ, 1967, №2.
8. Арсенин В.Я., Иванов В.В. Восстановление формы сигнала, свободной от искажений, обусловленных аппаратурой и каналом передач. «Измерительная техника», 1969, №1.

УДК 621.396.96

ФИЛЬТРАЦИЯ РАДИОЛОКАЦИОННЫХ СИГНАЛОВ ПРИ ИСПОЛЬЗОВАНИИ РАЗЛИЧНЫХ МАТЕМАТИЧЕСКИХ ПРЕДСТАВЛЕНИЙ ФУНКЦИЙ

Зайцев Г.В.

Введение

Широкое распространение цифровой обработки радиолокационных сигналов [1] требует развития и использования разнообразных методов обработки, адекватных решаемой задаче. Одним из принципиальных вопросов, отражающихся на выборе эффективных алгоритмов обработки и во многом определяющих ее качество, является способ построения аналого-цифрового преобразователя (АЦП).

Упрощенная блок-схема устройства частотно-временной обработки радиолокационных сигналов приведена на рис. 1. В функции устройства аналоговой обработки входит усиление принятого сигнала, понижение несущей частоты и полосовая фильтрация с целью выделения анализируемой области частот. Отфильтрованный сигнал преобразуется в цифровую форму, и основная обработка выполняется в процессоре сигналов.

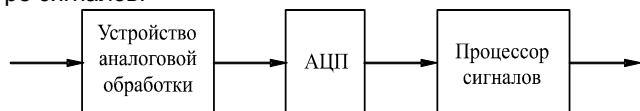


Рис. 1. Упрощенная блок-схема устройства обработки сигнала.

Традиционно АЦП строится по квадратурной схеме [2] (рис. 2а). Однако разбаланс квадратурных каналов по фазе и по амплитуде приводит к искажению спектра сигналов и существенному снижению характеристик тракта, в связи с чем, во многих случаях целесообразен переход к одноканальному построению АЦП [3] без использования квадратурных каналов (рис. 2б). При этом возможен два подхода к последующей цифровой обработке сигнала: первый состоит в формировании квадратурных каналов в цифровом виде после АЦП, второй – в непосредственной обработке сигнала, получаемого с выхода одноканального АЦП. Второй подход является более последовательным и позволяет в ряде случаев снизить требования по производительности к цифровым устройствам.

Возможность аналого-цифрового преобразования сигнала и его обработки без использования квадратурных каналов следует из теоремы Котельникова. Цифровые системы, реализующие этот метод, были использованы нами, начиная с конца 70-х годов прошлого века, для нескольких типов радиолокационных станций, успешно прошедших испытания и подтвердивших эффективность заложенных решений. Однако до сих пор до-

Показано, что результаты фильтрации радиолокационных сигналов при их различном математическом представлении (квадратурном, действительном, гильбертовом) могут быть сделаны практически идентичными путем корректного выбора несущей частоты. Приведено теоретическое обоснование метода цифровой обработки, состоящего в одноканальном (без квадратуры) аналого-цифровом преобразовании сигнала и непосредственной фильтрации получаемого действительного цифрового сигнала. Определены нижние границы для необходимой несущей частоты и частоты дискретизации. Показано, что описываемый метод сохраняет фазовые соотношения. Описывается эффект зависимости результата фильтрации от начальной фазы входного сигнала. Производится сравнение описываемого метода со стандартным, использующим квадратуру.

вольно широко распространено мнение об «ущербности» и не качественности такой обработки. В связи с этим, в данной работе дается обоснование корректности обработки радиолокационных сигналов без использования квадратурных каналов и описываются особенности такой обработки.

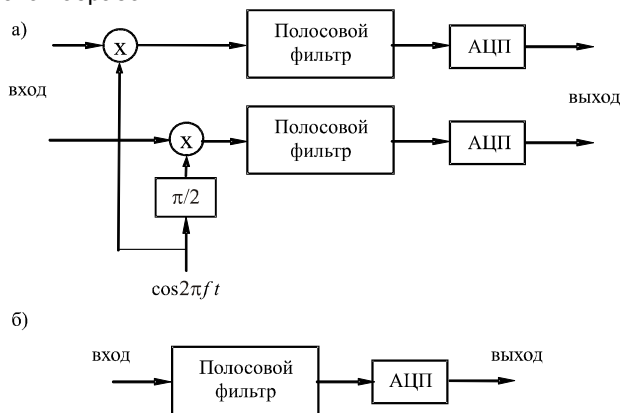


Рис. 2. Варианты построения АЦП.

Различные методы построения АЦП соответствуют в теоретическом плане различным способам представления сигналов. Можно выделить три основных способа представления: действительное, квадратурное и гильбертово [4,5]. Квадратурное представление сигналов соответствует квадратурной схеме АЦП, действительное – одноканальной схеме АЦП, гильбертово удобно при теоретическом анализе.

Перечисленные способы математического представления сигналов описываются во втором разделе. В третьем разделе рассматривается фильтрация аналоговых сигналов. Показывается, что при корректном выборе несущей частоты результаты фильтрации идентичны для всех перечисленных представлений. В четвертом разделе рассматриваются аналогичные условия идентичности для случая цифровых сигналов.

Способы представления сигналов и их взаимосвязь

Сигналы во временной области будем обозначать строчными буквами, их спектры – соответствующими прописными буквами, а нижние индексы r, q и h – для обозначения действительного, квадратурного и гильбертова представления соответственно. В случае отсутствия нижнего индекса представление может быть произвольным.

При действительном представлении радиолокационный сигнал имеет вид

$$s_r(t) = a(t) \cos[2\pi f_0 t + \theta(t)], \quad (1)$$

где $a(t)$ – амплитуда, $\theta(t)$ – фаза, f_0 – несущая частота сигнала. Спектр $S_r(f)$ сигнала (1) в силу действительности $s(t)$ обладает свойством комплексно-сопряженной симметрии: $S_r(f) = S_r^*(-f)$. Радиолокационный сигнал (1) характеризуется тем, что функции $a(t)$ и $\theta(t)$ изменяются медленно по сравнению с функцией $\cos(2\pi f_0 t)$.

При квадратурном представлении сигналу (1) соответствует комплексный сигнал

$$s_q(t) = u(t) \exp(2\pi j f_0 t), \quad (2)$$

где

$$u(t) = a(t) \exp[j\theta(t)]. \quad (3)$$

Выражение (1) можно представить в виде

$$s_r(t) = p(t) \cos(2\pi f_0 t) - q(t) \sin(2\pi f_0 t), \quad (4)$$

где

$$p(t) = a(t) \cos \theta(t); \quad q(t) = a(t) \sin \theta(t) \quad (5)$$

– квадратурные составляющие сигнала.

Из (4) следует, что квадратурные составляющие можно получить путем умножения сигнала (1) на функции $2\cos(2\pi f_0 t)$ и $-2\sin(2\pi f_0 t)$ и последующей низкочастотной фильтрации, что соответствует схеме квадратурного АЦП (рис. 2а).

Сравнивая (3) и (5), а также (1) и (2), нетрудно получить связь между указанными представлениями:

$$u(t) = p(t) + jq(t), \quad (6)$$

$$s_r(t) = 0,5[s_q(t) + s_q^*(t)]. \quad (7)$$

Из формул (2), (7) следуют выражения для спектров $S_q(f)$ и $S_r(f)$ сигналов (1) и (2) через спектр $U(f)$ ком-

плексной огибающей (3) (здесь и далее предполагается, что интегралы Фурье, также, как и другие интегралы, сходятся; это всегда имеет место при рассмотрении практических случаев):

$$S_q(f) = U(f - f_0), \quad (8)$$

$$\begin{aligned} S_r(f) &= 0,5[S_q(f) + S_q^*(-f)] = \\ &= 0,5[U(f - f_0) + U^*(-f - f_0)] \end{aligned} \quad (9)$$

Гильбертово представление сигнала определяется выражением

$$s_h(t) = s_r(t) + j\hat{s}_r(t), \quad (10)$$

где

$$\hat{s}_r(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s_r(\tau)}{t - \tau} d\tau. \quad (11)$$

При этом спектр сигнала (10) равен

$$S_h(f) = \begin{cases} 2S_r(f), & f > 0, \\ S_r(f), & f = 0, \\ 0, & f < 0, \end{cases} \quad (12)$$

Рис. 3 иллюстрирует различия между спектрами S_r , S_q , S_h . На рисунке вид гипотетической функции $U(f)$ выбран из соображений наглядности; спектр $U(f - f_0)$ и его отдельные части заштрихованы вертикальными линиями, а спектр $U^*(-f - f_0)$ – равномерными точками. Заметим, что для случая финитного спектра $U(f)$, равного нулю при $|f| > A$, спектры S_r , S_q , S_h совпадают в области положительных частот с точностью до постоянного множителя, если $f_0 > A$. В общем случае величина различия спектров определяется “хвостами” функции $U(f)$ при $f < -f_0$.

Оптимальная фильтрация при различных представлениях сигналов

Оптимальная линейная фильтрация радиолокационного сигнала сводится к вычислению корреляционно-взвешенного интеграла [1,2,4,6]:

$$z(\tau) = \int_{-\infty}^{\infty} x(t)h(\tau - t)dt = \int_{-\infty}^{\infty} X(f)H(f)\exp(2\pi j f \tau)df, \quad (13)$$

где $x(t)$ – принятый сигнал, $h(t)$ – импульсная характеристика оптимального фильтра, а $X(f)$ и $H(f)$ – спектры этих функций, причем $x(t)$ и $h(t)$ должны иметь, строго говоря, гильбертово представление.

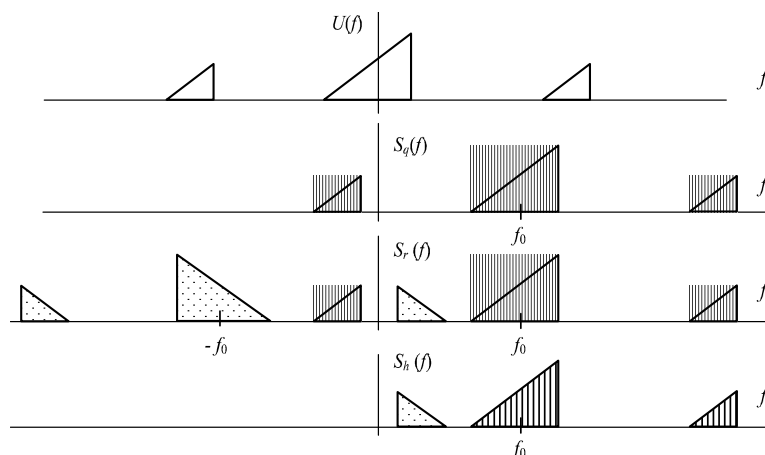


Рис. 3. Различия между спектрами S_r , S_q , S_h .

Пусть принимаемый сигнал $x_r(t)$ соответствует излучаемому (1) и отличается от него, как это обычно имеет место на практике, некоторой задержкой, амплитудой, доплеровским сдвигом F_1 и начальной фазой α . Тогда, выбирая соответствующим образом начало отсчета времени и масштабный коэффициент, имеем:

$$x_r(t) \cong a(t) \cos[2\pi(f_0 + F_1)t + \theta(t) + \alpha]. \quad (14)$$

Для спектров различных представлений принимаемого сигнала аналогично (8) и (9) получаем

$$X_q(f) = U(f - f_0 - F_1) \exp(j\alpha),$$

$$X_r(f) = 0.5[U(f - f_0 - F_1) \exp(j\alpha) + U^*(-f - f_0 - F_1) \exp(-j\alpha)], \quad (15)$$

а спектр $X_h(f)$ выражается через $X_r(f)$ согласно формуле (12).

Для эффективной оптимальной фильтрации сигнала на несущей частоте f_0 спектр $H(f)$ также должен быть сосредоточен у частоты f_0 . Поэтому импульсная характеристика фильтра также может быть представлена в виде, аналогичном (14):

$$h_r(t) = b(t) \cos[2\pi(f_0 + F_2)t + \psi(t)], \quad (16)$$

где F_2 – доплеровская частота, на которую настроен фильтр. При многоканальном приеме используется набор фильтров с различными частотами F_2 , перекрывающими необходимый доплеровский диапазон. Заметим, что $h(t)$ не учитывает начальную фазу α принимаемого сигнала (14), так как ее учет, как правило, не представляется возможным.

Обозначим символом $V(f)$ спектр комплексной огибающей $v(t) = b(t) \exp(j\psi(t))$ функции (16). Тогда спектры различных представлений $h(t)$ имеют вид

$$H_q(f) = V(f - f_0 - F_2),$$

$$H_r(f) = 0.5[V(f - f_0 - F_2) + V^*(-f - f_0 - F_2)], \quad (17)$$

а спектр $H_h(f)$ может быть получен по формуле (12).

В частности, для согласованного фильтра [4,5,6] $H(f) = S^*(f)$, $h(t) = s^*(-t)$.

Рассмотрим различия в выходном эффекте фильтра (13) при использовании различных представлений для сигнала $x(t)$ и импульсной характеристики фильтра $h(t)$. В случае гильбертовых представлений для обеих функций, подставив в (13) выражения для спектров из формул (12), (15), (17), получим:

$$z_{hh} = \exp(j\alpha)(J_1 + J_3) + \exp(-j\alpha)(J_2 + J_4), \quad (18)$$

где интегралы J_1, J_2, J_3, J_4 имеют вид

$$J_1 = \int_0^\infty U(f - f_0 - F_1) V(f - f_0 - F_2) \exp(2\pi j f \tau) df = \quad (19)$$

$$= \exp[2\pi j(f_0 + F_2)\tau] \int_{-f_0 - F_2}^\infty V(f) U(f - F) \exp(2\pi j f \tau) df;$$

$$J_2 = \int_0^\infty U^*(-f - f_0 - F_1) V^*(-f - f_0 - F_2) \exp(2\pi j f \tau) df = \quad (20)$$

$$= \exp[-2\pi j(f_0 + F_2)\tau] \int_{-\infty}^{-f_0 - F_2} V^*(f) U^*(f - F) \exp(-2\pi j f \tau) df;$$

$$J_3 = \int_0^\infty U(f - f_0 - F_1) V^*(-f - f_0 - F_2) \exp(2\pi j f \tau) df; \quad (21)$$

$$J_4 = \int_0^\infty U^*(-f - f_0 - F_1) V(f - f_0 - F_2) \exp(2\pi j f \tau) df, \quad (22)$$

а величина F определяется выражением $F = F_1 - F_2$.

Вид функций, стоящих под знаком интеграла в J_1, J_2, J_3, J_4 для гипотетического случая, удобного для наглядной иллюстрации, приведен на рис. 4а, 4б, 4в соответственно.

Как нетрудно убедиться, интеграл J_1 представляет собой основной член в отклике фильтра, и при отсутствии рассогласования параметров ($F = 0, \tau = 0$) для случая согласованной фильтрации примерно равен энергии сигнала

$$\int_{-f_0}^\infty U(f) U^*(f) df.$$

Аналогичным образом интеграл J_2 для этого случая дает энергию “хвостов” спектра сигнала при $f < -f_0$ ($F = 0, \tau = 0$). Интегралы J_3, J_4 представляют собой комбинационные члены, и из рис. 4б, 4в наглядно видно, что при достаточно большой частоте f_0 они должны быть малы. Ниже этот вопрос рассматривается подробнее.

Выражения, аналогичные (18), нетрудно получить и для других представлений принятого сигнала и импульсной характеристики фильтра. Результаты этих вычислений сведены в таблицу 1. При этом используются следующие обозначения.

Если $J = \int_a^b g(x) dx$ –

некоторый интеграл, то символом $\tilde{J}$ обозначен интеграл

$$\tilde{J} = \int_{-\infty}^\infty g(x) dx - J.$$

Рассмотрим выражения таблицы 1, отбросив на время клетку, соответствующую действительному представлению обеих функций. Все остальные выражения имеют, с точностью до постоянного множителя, в качестве основного члена интеграл $\exp(j\alpha) J_1$ и отличаются линейной комбинацией интегралов с индексами 2,3,4.

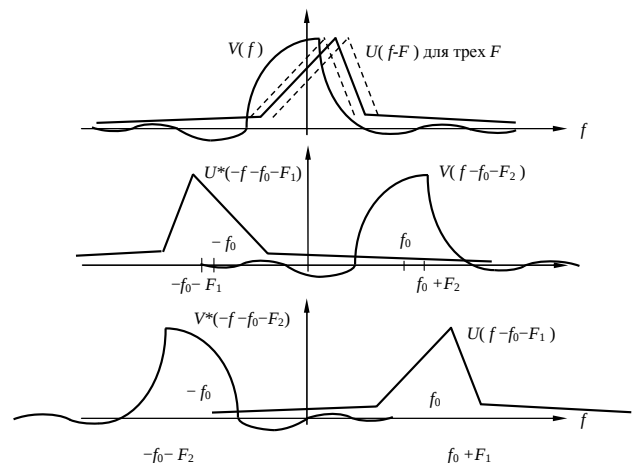


Рис. 4. Взаимное расположение спектров подинтегральных функций.

Результаты фильтрации для различных представлений сигнала и фильтра

Представление фильтра	Представление сигнала		
	гильбертово	квадратурное	действительное
гильбертово	$\exp(j\alpha)(J_1 + J_3) + \exp(-j\alpha)(J_2 + J_4)$	$\exp(j\alpha)(J_1 + J_3)$	$0,5\{\exp(j\alpha)(J_1 + J_3) + \exp(-j\alpha)(J_2 + J_4)\}$
квадратурное	$\exp(j\alpha)J_1 + \exp(-j\alpha)J_4$	$\exp(j\alpha)(J_1 + J_2^*)$	$0,5\{\exp(j\alpha)(J_1 + J_2^*) + \exp(-j\alpha)(J_4 + \tilde{J}_4)\}$
действительное	$0,5\{\exp(j\alpha)(J_1 + J_3) + \exp(-j\alpha)(J_2 + J_4)\}$	$0,5\exp(j\alpha)(J_1 + J_2^* + J_3 + \tilde{J}_3)$	$0,25\{\exp(j\alpha)(J_1 + J_2^* + J_3 + \tilde{J}_3) + \exp(-j\alpha) \cdot (J_1^* + J_2 + J_4 + \tilde{J}_4)\}$

Постоянные множители 0,5 и 0,25 в таблице отражают тот факт, что энергия действительного сигнала (1) вдвое меньше, чем комплексного (2). Однако при этом отношение «сигнал/шум» сохраняется постоянным, так как и мощность шумов в двухканальном варианте также выше. В связи с этим при анализе различий выражений таблицы 1 указанные коэффициенты можно не рассматривать. Покажем, что выбором достаточно большой величины f_0 каждый из интегралов с индексами 2, 3, 4, а значит и любые их линейные комбинации, могут быть сделаны произвольно малыми. Рассмотрим, например, интеграл J_3 . В силу неравенства Коши-Буняковского имеем:

$$|J_3|^2 \leq \int_0^\infty |U(f - f_0 - F_1)|^2 df \cdot \int_0^\infty |V(-f - f_0 - F_2)|^2 df. \quad (23)$$

Для физически реализуемых сигнала и фильтра и устойчивого фильтра энергия функций $x(t)$ и $h(t)$ ограничена. Поэтому

$$\int_{-\infty}^\infty |U(f)|^2 df = c_1 < \infty, \quad (24)$$

$$\int_{-\infty}^\infty |V(f)|^2 df = c_2 < \infty. \quad (25)$$

Оценивая первый интеграл в (23) с помощью (24) и произведя замену переменных во втором, имеем

$$|J_3|^2 \leq c_1 \int_{-\infty}^{-f_0 - F_2} |V(f)|^2 df. \quad (26)$$

Но из (25) следует, что для любого сколь угодно малого ε найдется такое A , что при $f_0 > A$ выполняется неравенство

$$\int_{-\infty}^{-\beta} |V(f)|^2 df < \varepsilon.$$

Тогда из (26) имеем, $|J_3| \leq \sqrt{c_1 \varepsilon}$ при $f_0 > A - F_2$. Так как ε произвольно мало, а величина F_2 ограничена, то доказываемое утверждение для интеграла J_3 справедливо. Аналогично производится доказательство и для других интегралов.

Таким образом, различия в выходном эффекте оптимального фильтра (13) для различных представлений сигнала и фильтра (кроме действительного представле-

ния обеих функций) пренебрежимо малы при достаточно большой величине несущей частоты f_0 , а соответствующие алгоритмы фильтрации эквивалентны.

Для случая действительного представления, как сигнала, так и фильтра, основной полезный член в выходной реакции z_{rr} равен $0,5 \operatorname{re}[\exp(j\alpha) J_1]$ и имеет несущую вида $\cos(2\pi f t)$ вместо $\exp(2\pi j f t)$. В силу случайности начальной фазы α это приводит к тому, что в любой точке (t, F) выходной эффект может быть равен нулю, и при вычислении функции z_{rr} в отдельных точках применять такой способ нецелесообразно. Если же функция z_{rr} рассчитывается на отрезке времени, то при несущей, существенно превосходящей полосу сигнала, огибающая функции z_{rr} «обрисовывается», тем не менее, хорошо. При малых значениях несущей указанный способ может привести к существенным потерям.

Во многих радиолокационных приложениях функции $x(t)$ и $h(t)$ в (13) можно считать функциями с финитным спектром [7]. Например, при обработке принимаемых радиолокационных сигналов входные цепи приемного устройства всегда имеют конечную полосу. Поэтому сигнал с финитным спектром является хорошей моделью для большинства приложений. Из приведенных выше формул нетрудно определить требования к несущей частоте f_0 для того, чтобы результаты фильтрации (13) были идентичны для всех представлений функций $x(t)$ и $h(t)$ с финитными спектрами. Рассмотрим наиболее интересный с практической точки зрения случай.

Пусть спектр $U(f)$ равен нулю вне интервала частот (a, b) , $V(f)$ — вне интервала (c, d) , а доплеровский диапазон соответствует интервалу (E, G) . Будем считать, для определенности, что несущая частота положительна $f_0 > 0$, а интервалы (a, b) и (c, d) содержат начало координат. Определим условия обращения в нуль интегралов таблицы 1 с индексами 2, 3, 4 при любых частотах F_1 и F_2 из доплеровского диапазона. Для интеграла J_2 получаем, что функции $V(f)$ и $U(f - F)$, стоящие под интегралом в (20), не равны нулю, соответственно, на интервалах (c, d) и $(a + F, b + F)$. Для того, чтобы интеграл (20) был равен нулю, достаточно, чтобы хотя бы одна из левых границ этих интервалов была бы больше $-f_0 - F_2$, откуда получаем условие

$$f_0 \geq \min(-a, -c) - E. \quad (27)$$

Интегралы $J_3, J_4, \tilde{J}_3, \tilde{J}_4$ обращаются в нуль, если области ненулевых значений подынтегральных функций в (21), (22) не пересекаются. Отсюда нетрудно получить условие

$$f_0 \geq 0,5(-c-a) - E. \quad (28)$$

Так как справедливо соотношение $0,5(y+z) \geq \min(y, z)$, то условие (28) жестче, чем (27). В частности, для наиболее интересного с практической точки зрения случая равенства интервалов ненулевых значений функций $U(f)$ и $V(f)$ из (28) получим

$$f_0 \geq -a - E. \quad (29)$$

Это выражение показывает, что результаты фильтрации идентичны при всех представлениях функций $x(t)$ и $h(t)$ уже при достаточно низкой частоте f_0 , примерно равной половине ширины спектра сигнала плюс половина доплеровского диапазона. Для таких низких частот понятия фазы и амплитуды сигнала, как правило, теряют смысл. Однако представление сигнала в виде (14) остается полезным, так как сигнал именно такой формы получается после пропускания сигнала вида (14) с высокочастотной несущей (для которого это представление является классическим) через смеситель, понижающий несущую до указанной величины.

Следующее существенное замечание по выражениям таблицы 1 состоит в том, что начальная фаза α фильтруемого сигнала линейно входит в фазу выходного эффекта фильтра, имеющую в области главного лепестка вид $\arg(J_1) + \alpha$. Это обстоятельство является существенным при использовании моноимпульсных методов пеленгации.

На основании анализа таблицы 1 можно сделать еще один важный вывод. Для каждого выражения результатов фильтрации во второй и четвертой колонке таблицы присутствуют два члена: один с фазовым множителем $\exp(j\alpha)$, другой – с множителем $\exp(-j\alpha)$. Так как интегралы при этих множителях, вообще говоря, ненулевые, то это означает, что модуль отклика фильтра зависит от начальной фазы сигнала α . Действительно, модуль выражения $\exp(j\alpha)A + \exp(-j\alpha)B$ меняется от $|A| + |B|$ до $||A| - |B||$ при изменении фазы α от 0 до 2π . Таким образом, наблюдается зависимость величины модуля отклика фильтра от фазы входного сигнала. В области главного лепестка функции неопределенности сигнала справедливо соотношение $|J_1| \gg |J_k|, |\tilde{J}_1| \gg |\tilde{J}_k|, k = 2, 3, 4$, и эта зависимость незначительна. В области же боковых лепестков эта зависимость в процентном отношении может быть значительной. Из таблицы 1 следует, что в случае использования квадратурного представления сигнала такой зависимости нет.

Цифровая оптимальная фильтрация

Перейдем к рассмотрению дискретного случая. При цифровой обработке из обрабатываемой функции $y(t)$ берутся выборки через интервал $\delta = 1/F_s$, где F_s – частота дискретизации. Для обозначения полученных при этом дискретных сигналов и функций от них будем использовать те же символы, что и для соответствующих функций непрерывного времени, но с добавлением индекса s . Тогда процесс дискретизации функции $y(t)$ дает функцию $y_s(k)$ дискретного аргумента k :

$$y_s(k) = y(k\delta), k = -\infty, \dots, -1, 0, 1, \dots, \infty.$$

Спектр $Y_s(f)$ этой функции определяется выражением

$$Y_s(f) = \sum_{k=-\infty}^{\infty} y_s(k) \exp(-2\pi j k \delta f). \quad (30)$$

Дискретизация приводит к наложению участков спектра $Y(f)$ [1], а именно, спектры $Y_s(f)$ и $Y(f)$ связаны соотношением

$$Y_s(f) = \frac{1}{\delta} \sum_{k=-\infty}^{\infty} Y(f + kF_s). \quad (31)$$

Нетрудно видеть, что функция $Y_s(f)$ является периодической с периодом F_s , а ряд (30) является ее рядом Фурье, и, следовательно, справедливо выражение

$$y_s(k) = \delta \int_0^{F_s} Y_s(f) \exp(2\pi j k \delta f) df. \quad (32)$$

Для дискретных функций вместо интеграла (13) вычисляется сумма

$$z_s(\tau) = \sum_{k=-\infty}^{\infty} x(k\delta)h(\tau - k\delta). \quad (33)$$

Здесь индекс s имеет несколько другой смысл, чем в (30), однако в данной статье это не приводит к недоразумениям. Заметим, что хотя функция (33) относится к случаю дискретных сигналов, она является функцией непрерывной переменной τ и может быть вычислена для любого τ . При цифровой обработке эта функция обычно вычисляется с шагом δ по переменной τ . При этом интегральная свертка (13) заменяется дискретной:

$$z_s(m\delta) = \sum_{k=-\infty}^{\infty} x_s(k)h_s(m-k) = \sum_{k=-\infty}^{\infty} x_s(m-k)h_s(k). \quad (34)$$

Определим связь между результатами цифровой (34) и аналоговой (13) фильтрации. При этом ограничимся для простоты наиболее часто используемым на практике случаем квадратурного представления импульсной характеристики фильтра. Тогда вместо (16) имеем

$$h_q(t) = b(t) \exp[2\pi j (f_0 + F_2)t + j\psi(t)].$$

Подставив это выражение в (13), получим для случая непрерывных переменных:

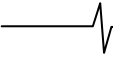
$$z(\tau, F) = \int_{-\infty}^{\infty} h(t)x(\tau-t, f_0 + F_1)dt = \int_{-\infty}^{\infty} x(\tau-t, f_0 + F_1)b(t) \exp[2\pi j (f_0 + F_2)t + j\psi(t)]df, \quad (35)$$

где в аргументы функций x и z в явном виде внесена несущая $f_0 + F_1$ и относительный доплеровский сдвиг $F = F_1 - F_2$ соответственно. Вид функции x приведен для действительного случая в (14), а для других представлений может быть получен по формулам (2) – (11). Произведя в (35) замену переменных $f_1 = f_0 + F_1$, получим

$$z(\tau, F) = \int_{-\infty}^{\infty} w(t, \tau, f_1) \exp(-2\pi j Ft) dt, \quad (36)$$

где

$$w(t, \tau, f_1) = x(\tau-t, f_1)b(t) \exp[2\pi j f_1 t + j\psi(t)].$$



Выражение (36) можно рассматривать как преобразование Фурье функции $w(t, \tau, f_1)$ по переменной t , а (33) – как соответствующее дискретное преобразование. В связи с этим из общей формулы (31), связывающей интегральное и дискретное преобразования Фурье, получим

$$z_s(\tau, F) = \frac{1}{\delta} \sum_{k=-\infty}^{\infty} z(\tau, F + kF_s). \quad (37)$$

Таким образом, результаты цифровой и аналоговой фильтрации связаны таким же соотношением, как и спектры соответствующих функций дискретного и непрерывного аргументов. Для случая согласованной фильтрации и квадратурного представления сигнала это соотношение получено и подробно проанализировано в [1].

Выразим теперь функцию $z_s(m\delta)$ (34) через спектры дискретизированных функций $x_s(k)$ и $h_s(k)$ аналогично тому, как это сделано в (13) для функций непрерывного аргумента. Подставив в (34) выражения, аналогичные (32), для $x_s(k)$ и $h_s(m-k)$ через спектры этих функций и производя несложные преобразования, получим:

$$z_s(m\delta) = \delta \int_0^{F_s} X_s(f) H_s(f) \exp(2\pi j m \delta f) df. \quad (38)$$

Прежде, чем перейти к анализу полученных соотношений, сделаем одно обобщающее замечание. При вычислении функции (38) шаг δ_1 по переменной τ , вообще говоря, может не совпадать с шагом δ по переменной t . Другими словами, функции $x(t)$ и $h(t)$ могут иметь различные частоты дискретизации. Практический смысл могут иметь случаи, когда эти частоты кратны. Можно показать, что и в этом случае формулы (34), (37), (38) остаются справедливыми, если произвести следующие изменения обозначений. Пусть $\delta = L\delta_0$, $\delta_1 = M\delta_0$, где L и M – взаимно-простые целые числа. Приведем последовательности выборок из функций $x(t)$ и $h(t)$ к одной частоте дискретизации $F_{s1} = 1/\delta_0$, вставив $L-1$ (соответственно $M-1$) нулевых отсчетов между каждыми двумя отсчетами функции $x(t)$ (соответственно $h(t)$). Для полученных новых функций $\bar{x}_s(k)$ и $\bar{h}_s(k)$ справедливы соотношения (34), (37), (38) при замене в (34), (38) δ на δ_0 , F_s на F_{s1} , а в (37) – δ на $LM\delta_0$ и F_s на $1/LM\delta_0$.

Вернемся к анализу соотношений (37), (38), позволяющих сделать вывод о необходимой частоте дискретизации. Проанализируем случай сигналов с финитным спектром. Пусть, как и в п.3, спектры $U(f)$ и $V(f)$ не равны нулю на интервалах (a, b) и (c, d) соответственно, а доплеровский диапазон определяется интервалом (E, G) .

Рассмотрим сначала случай квадратурного представления сигнала и фильтра. Из сравнения (38) с (13) следует, что результаты аналоговой и цифровой фильтрации совпадают при $\tau = l \cdot \delta$, $l = 0, \pm 1, \pm 2, \dots$, если интервал $(A, A + F_s)$ при некотором A перекрывает области ненулевых значений функций $X_q(f)$ и $H_q(f)$. При этом не происходит наложения отрезков спектров $X_q(f)$ и $H_q(f)$ при дискретизации. Рассматривая выражения для спектров $X_q(f)$ и $H_q(f)$ второго раздела, нетрудно получить, что указанное условие эквивалентно неравенству

$$F_s \geq \max_{F_1, F_2 \in (E, G)} \{ \max(b + F_1, d + F_2) - \min(a + F_1, c + F_2) \} \quad (39)$$

Несколько завышенная оценка правой части этой формулы дает

$$F_s \geq \max(b, d) - \min(a, c) + G - E.$$

При равенстве интервалов (a, b) и (c, d) условие (39) упрощается:

$$F_s \geq W_c + \Delta F. \quad (40)$$

где $W_c = b - a$ – ширина спектра сигнала, а $\Delta F = G - E$ – величина доплеровского диапазона. Заметим, однако, что в общем случае минимально необходимая частота дискретизации может быть меньше, чем величина, задаваемая формулой (40). Например, при $b + E > d + G$ и $a + G < c + E$ формула (39) дает $F_s \geq b - a = W_c$.

Определим аналогичные условия для случая действительного представления сигнала. При этом спектр $X(f)$ определяется формулой (15) и содержит две области ненулевых значений. Для того, чтобы в данном случае интегралы (38) и (13) совпадали при $\tau = l \cdot \delta$, $l = 0, \pm 1, \pm 2, \dots$, достаточно, чтобы при дискретизации спектр $U^*(-f - f_0 - F_1)$ в (15) не наложился ни на спектр $U(f - f_0 - F_1)$, ни на спектр $H_q(f) = V(f - f_0 - F_2)$. Выразим эти условия в виде неравенств, воспользовавшись рис. 5, где изображен отрезок частотной оси около частоты f_0 . Сплошными линиями показаны спектры $U(f - f_0 - F_1)$ и $V(f - f_0 - F_2)$, а пунктирными линиями – наложения спектра $U^*(-f - f_0 - F_1)$ при дискретизации. На частотной оси отмечены точки, кратные $F_s/2$. Одинаковыми цифрами в кружках показаны размеры отрезков, дающих рассматриваемые ниже неравенства.

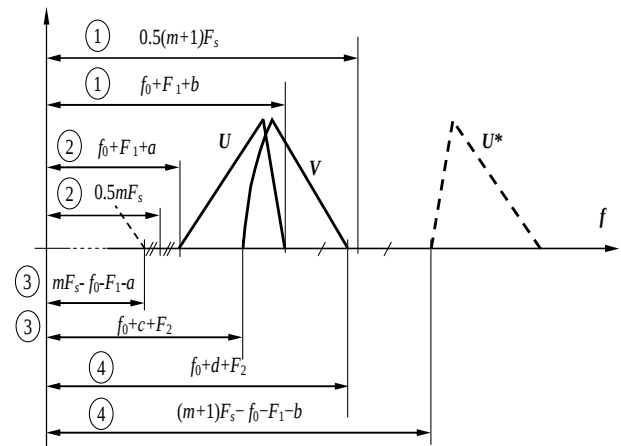


Рис. 5. Взаимное расположение спектров для действительного представления

Для того, чтобы спектры $U(f - f_0 - F_1)$ и $U^*(-f - f_0 - F_1)$ не наложился при дискретизации, необходимо, чтобы область ненулевого спектра $U(f - f_0 - F_1)$ для некоторого целого m располагалась внутри отрезка $[mF_s/2, (m+1)F_s/2]$, откуда получим (см. отрезки, обозначенные кружками с цифрами 1 и 2 на рис. 5):

$$\begin{cases} 2(f_0 + F_1 + b) \leq (m+1)F_s \\ mF_s \leq 2(f_0 + F_1 + a) \end{cases}$$

Для усиления этих неравенств положим в первом из них $F_1 = G$, а во втором $F_1 = E$:

$$\begin{cases} 2(f_0 + G + b) \leq (m+1)F_s \\ mF_s \leq 2(f_0 + E + a) \end{cases} \quad (41)$$

Складывая эти неравенства, получим:

$$F_s \geq 2(W_c + \Delta F). \quad (42)$$

Аналогично нетрудно получить условия, гарантирующие отсутствие наложения спектров $U^*(-f - f_0 - F_1)$ и $V(f - f_0 - F_2)$ (отрезки с цифрами 3 и 4 на рис. 5):

$$\begin{cases} 2(f_0 + G) + b + d \leq (m+1)F_s \\ mF_s \leq 2(f_0 + E) + a + c \end{cases} \quad (43)$$

Сложение этих неравенств дает условие:

$$F_s \geq W_c + W_f + 2\Delta F. \quad (44)$$

где $W_f = d - c$ – ширина спектра фильтра.

Объединяя (42) и (44) в одно неравенство, имеем окончательное условие для частоты дискретизации:

$$F_s \geq 2\Delta F + W_c + \max(W_c, W_f). \quad (45)$$

Сравнивая условия (45) и (40), получим, что в случае действительного представления сигнала минимальная частота дискретизации должна быть примерно вдвое выше, чем в случае квадратурного представления.

Условие (45) не исчерпывает неравенств (41) и (43). При обработке действительного сигнала несущая частота f_0 не может быть выбрана произвольно, а должна быть связана с частотой дискретизации F_s указанными соотношениями. Действительно, если, например, частоты f_0 и F_s кратны ($f_0 = kF_s$), то при дискретизации происходит наложение спектров $U(f - f_0)$ и $U^*(-f - f_0)$, что недопустимо. Объединяя четыре условия (41) и (43) в два, получим соотношения, которым должны удовлетворять частоты f_0 и F_s :

$$\begin{cases} 2(f_0 + G) + b + \max(b, d) \leq (m+1)F_s \\ mF_s \leq 2(f_0 + E) + a + \min(a, c) \end{cases} \quad (46)$$

При равных интервалах (a, b) и (c, d), а также при симметрии этих интервалов и интервала доплеровских частот относительно нуля, условия (46) упрощаются:

$$0,5(mF_s + \Delta F + W_c) \leq f_0 \leq 0,5[(m+1)F_s - \Delta F - W_c].$$

В литературе обычно предлагается частоту f_0 выбирать в середине интервала $[mF_s/2, (m+1)F_s/2]$. Тогда

$$f_0 = 0,25(2m+1)F_s,$$

а условия (46) приобретают вид

$$\begin{cases} F_s \geq 4G + 2b + 2\max(b, d) \\ F_s \geq -4E - 2a - 2\min(a, c) \end{cases} \quad (47)$$

При симметрии всех интервалов и равенстве интервалов (a, b) и (c, d) условия (47) являются частным случаем условия (44).

Таким образом, цифровая обработка действительного сигнала дает точный результат при соблюдении условий (45) и (46).

Аналогичным образом могут быть рассмотрены и другие случаи представления сигналов с финитными спектрами.

Для случая нефинитных спектров результаты дискретной фильтрации несколько отличаются от результатов аналоговой фильтрации за счет эффекта наложения участков спектра при дискретизации. Однако эти отличия, как можно показать, имеют одинаковый порядок величины для случаев квадратурного и действительного представления сигналов, если частота дискретизации и несущая частота выбраны в

соответствии с рассмотренными соотношениями.

Итак, цифровая фильтрация для действительного и для квадратурного представлений сигнала дает эквивалентные результаты, если частота дискретизации в первом случае примерно вдвое выше. При этом вдвое увеличивается также и количество отсчетов сигнала за фиксированное время наблюдения. Однако это не только не приводит к увеличению объема вычислений, но часто даже сокращает его. Например, при фильтрации сигнала во временной области путем прямого вычисления суммы (34) каждое слагаемое требует для комплексных отсчетов сигнала четыре операции умножения и две операции сложения двух действительных чисел, а для действительных отсчетов – только 2 операции умножения. При фильтрации в частотной области [1] основная операция – быстрое преобразование Фурье – требует одинакового объема вычислений для комплексного сигнала из N отсчетов и для действительного сигнала из $2N$ отсчетов и т.д.

Заключение

Для любого представления радиолокационного сигнала и импульсной характеристики фильтра (кроме действительного представления обеих функций) результаты аналоговой фильтрации практически идентичны при выборе не слишком низкой несущей частоты. Для случая сигналов с финитным спектром несущая частота должна быть не менее половины ширины спектра плюс половина доплеровского диапазона. При этом в любом случае фильтрация сохраняет фазовые соотношения во входном сигнале, что дает возможность реализовать моноимпульсный метод пеленгации. При использовании представления входного сигнала, отличного от квадратурного, следует обращать внимание на зависимость боковых лепестков результата фильтрации от фазы входного сигнала.

Для цифровой фильтрации переход к одноканальной схеме аналого-цифрового преобразователя и дальнейшая обработка сигнала без использования квадратурных каналов являются вполне корректным методом построения систем цифровой фильтрации. При этом снимается довольно сложная техническая задача нейтрализации паразитного влияния разбалансов квадратурных каналов АЦП. Платой за это является повышение примерно вдвое частоты дискретизации сигнала. Повышается также и общее количество отсчетов сигнала, однако это не требует увеличения производительности цифровых устройств, так как операции с действительными числами проще.

Литература

1. Дж.Х. Маклеллан, Р. Дж. Пурди, Применение цифровой обработки сигналов в радиолокации, в сб. "Применение цифровой обработки сигналов", под ред. Э. Оппенгейма, М., Мир, 1980.
2. С.З. Кузьмин, Основы проектирования систем цифровой обработки радиолокационной информации, М., Радио и связь, 1986.
3. Д.Ю. Бобров, А.П. Доброжанский, Г.В. Зайцев и др., Цифровая обработка сигналов в многофункциональных РЛС, Цифровая обработка сигналов, № 4, 2001, стр. 2-12.
4. Справочник по радиолокации под ред. М. Скольника, т.1, М., Сов. Радио, 1976.
5. Ч. Кук, М. Бернфельд, Радиолокационные сигналы, М., Сов. Радио, 1971.
6. Я.Д. Ширман, В.Н. Манжос, Теория и техника обработки радиолокационной информации на фоне помех, М., Радио и связь, 1981.
7. Я.И. Хургин, В.П. Яковлев, Финитные функции в физике и технике, М., Наука, 19

ДИАГНОСТИРОВАНИЕ ПРОЦЕССОВ РЕЗАНИЯ С ПОМОЩЬЮ ВЕЙВЛЕТ-АНАЛИЗА СИГНАЛА АКУСТИЧЕСКОЙ ЭМИССИИ

Хвостиков А.С., Щетинин В.С.

Создание на базе станков с ЧПУ автоматизированных производств с управлением от ЭВМ ставит задачу активного технологического диагностирования процесса обработки с целью его автоматизации, контроля и его управления. Поэтому проблема создания надежного аппарата технологического диагностирования процесса обработки является актуальной. Традиционное использование для этой цели силы и температуры резания в ряде случаев оказывается неэффективным или неадекватным. [5].

Применение диагностик методами «Акустической эмиссии» позволяет получить информацию о таких важнейших параметрах обработки, как износ инструмента, качество поверхностного слоя деталей и др.

Существующие системы автоматической диагностики состояния процесса резания обладают целым рядом существенных недостатков и не удовлетворяют в полной мере требованиям современного гибкого автоматизированного производства, так как не обеспечивают функционирование с высокой степенью точности и достоверности в реальном масштабе времени [7].

Экспериментальные исследования [1, 2, 3, 4 и др.] по диагностированию процессов резания проводятся при расположении датчика в непосредственной близости от зоны резания, а на производстве установка датчика в непосредственной близости от зоны резания нецелесообразна, вследствие негативного воздействия стружки, которая может повредить датчик и кабели, связывающие его с принимающей аппаратурой, а также других тепловых и механических воздействий. При этом нельзя не упомянуть, что при прохождении сигнала даже на незначительном удалении датчика от места съема, информационный сигнал существенно изменяется. При этом в автоматизированном производстве становится проблематична смена инструмента, а при удалении датчика от зоны резания встает вопрос о достоверности снимаемого сигнала.

Для вращающегося инструмента проблема съема сигнала встает более остро. Удаление от зоны резания к стационарному узлу станка добавляет в сигнал акустической эмиссии большое количество шумоподобных включений, зависящих от упругих, тепловых и других процессов, возникающих на расстоянии от зоны резания до места съема сигнала акустической эмиссии, а также вызывает появление составляющих от подвижных элементов и узлов, например, подшипников. Все эти процессы, в настоящее время, невозможно описать ни математически, ни моделями и, как следствие, не возможно по принимаемому сигналу судить о процессах, происходящих в зоне резания.

Наиболее эффективным и простым методом диагностирования процесса резания является частотный и частотно-

Исследуется эффективность частотно-временного анализа сигналов виброакустической эмиссии применительно к диагностированию процессов металлообработки на станках с ЧПУ автоматизированных производств. Показано преимущество многомасштабного вейвлет-анализа по отношению к частотно-временному анализу Фурье. Результаты исследований иллюстрируются примерами.

временной анализ сигналов виброакустической эмиссии (ВАЭ). При анализе Фурье сигнал раскладывается на единичные составляющие. Разложенный сигнал представляет собой составляющие, параметры частоты которого зачастую отличаются от своего истинного значения – смещены в пределах интервала определения частоты с помощью преобразования Фурье. Причем на интервале определения частоты может находиться не одна, а две или несколько составляющих спектра мощности. Кроме этого частота колебания инструмента за рассматриваемый период времени не остается постоянной, а изменяется вокруг некоторой величины.

При прохождении сигнала ВАЭ через систему измерения информативность сигнала резко снижается. На рис. 1-3 изображен частотно-временной анализ спектров мощности сигналов ВАЭ, снятых с державки в непосредственной близости от зоны резания, а на рис. 4-6 - снятых при аналогичных условиях, но с волнового приемника расположенного на оборотной стороне резцедержателя.

Исходя из спектров мощности сигналов ВАЭ, снятых на значительном удалении (рис. 4–6), можно сказать о существенной потере информативности: увеличения шумовых составляющих, невозможность определения периодичности воздействия сил резания, существенное искажение сигнала.

Для систем с большим числом звеньев на расстоянии от зоны резания до места съема сигнала не возможно определить степень износа режущего инструмента с помощью частотно-временного анализа.

При анализе вейвлет-спектров многомасштабного вейвлет-анализа, построенных по данным, содержащим шумы и выпадения, следует учитывать, что искажения сосредоточены в основном в высокочастотной части спектра. Игнорируя эту область, мы получаем вейвлет-спектр, который слабо зависит от числа ячеек гистограммы и ее искажений [6].

Многомасштабный вейвлет-анализ обладает способностью выделять из сигнала компоненты разного масштаба. Это часто используют для того, чтобы разделить исходные данные на составляющие (аналогично тому, что происходит при фильтрации с помощью преобразования Фурье). Проблема тесно связана с двумя другими: шумоподавлением и определением параметров сигнала по результатам наблюдения.

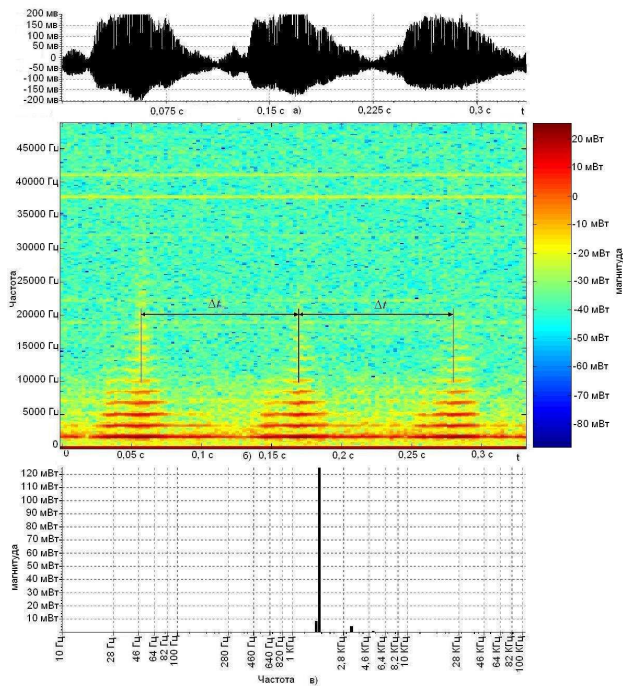
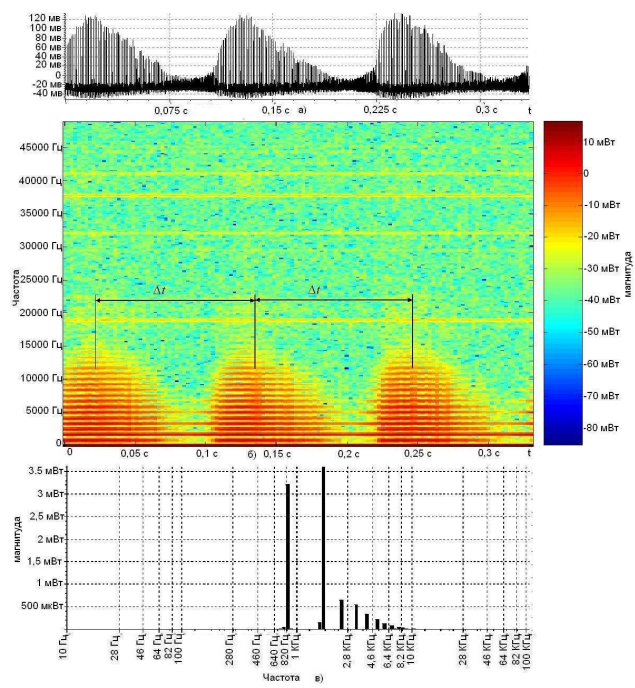


Рис. 1. Частотно-временной анализ и б) спектр мощности в) сигнала ВАЭ а), излучаемой в процессе точения жесткой заготовки, снятой в непосредственной близости от зоны резания, износ инструмента $h_3=0$, шероховатость $R_a=1.7$ мкм, показатель Ляпунова $\lambda_1=-0,005887$.



Рису. 2. Частотно-временной анализ и б) спектр мощности в) сигнала ВАЭ а), излучаемой в процессе точения жесткой заготовки, снятой в непосредственной близости от зоны резания, износ инструмента $h_3=0.15$ мм, шероховатость $R_a=2.1$ мкм, показатель Ляпунова $\lambda_1=-0,442315$.

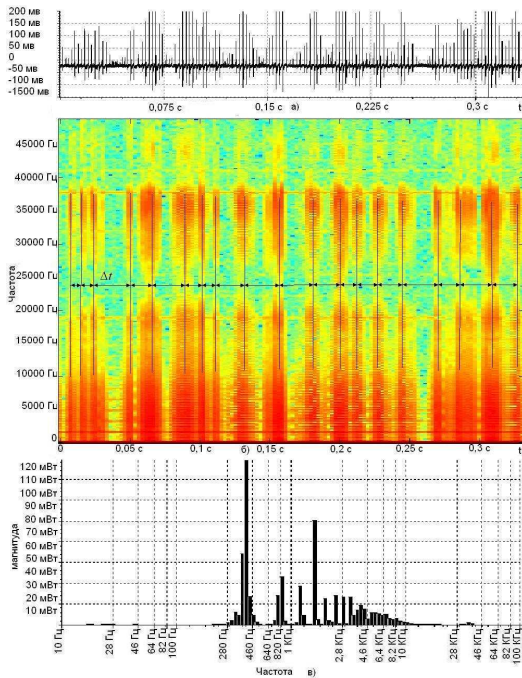


Рис. 3. Частотно-временной анализ и б) спектр мощности в) сигнала ВАЭ а), излучаемой в процессе точения жесткой заготовки, снятой в непосредственной близости от зоны резания, износ инструмента $h_3=0.5$ мм, шероховатость $R_a=4.0$ мкм, показатель Ляпунова $\lambda_1=-0,668861$.

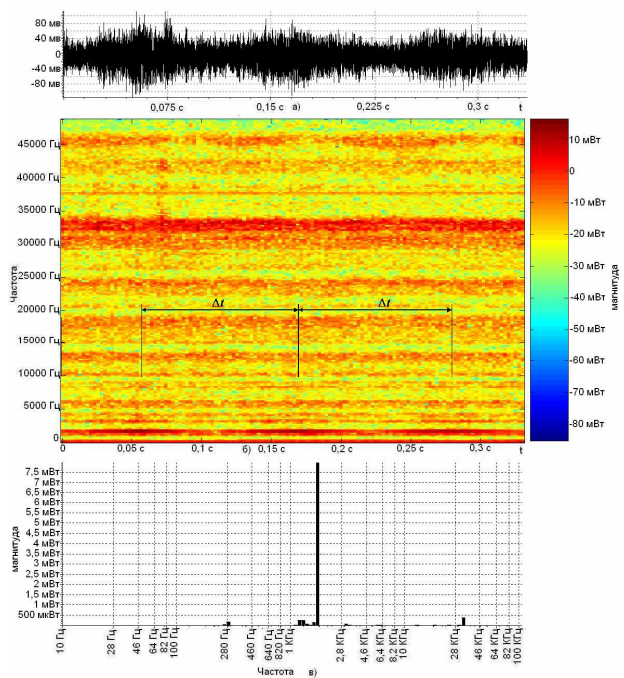


Рис. 4. Частотно-временной анализ и б) спектр мощности в) сигнала ВАЭ а), излучаемой в процессе точения жесткой заготовки, снятого с волнового приемника расположенного на оборотной стороне резцедержателя, износ инструмента $h_3=0$, шероховатость $R_a=1.7$ мкм, показатель Ляпунова $\lambda_1=-0,005887$.

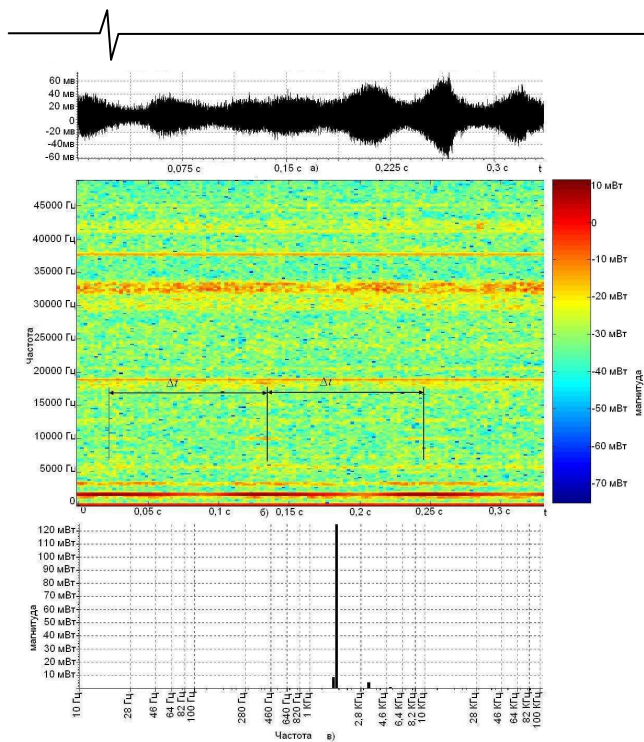


Рис. 5. Частотно-временной анализ и б) спектр мощности в) сигнала ВАЭ а), излучаемой в процессе точения жесткой заготовки, снятого с волнового приемника расположенного на оборотной стороне резцедержателя, износ инструмента $h_3=0.15$ мм, шероховатость $R_a=2.1$ мкм, показатель Ляпунова $\lambda_1=-0.442315$.

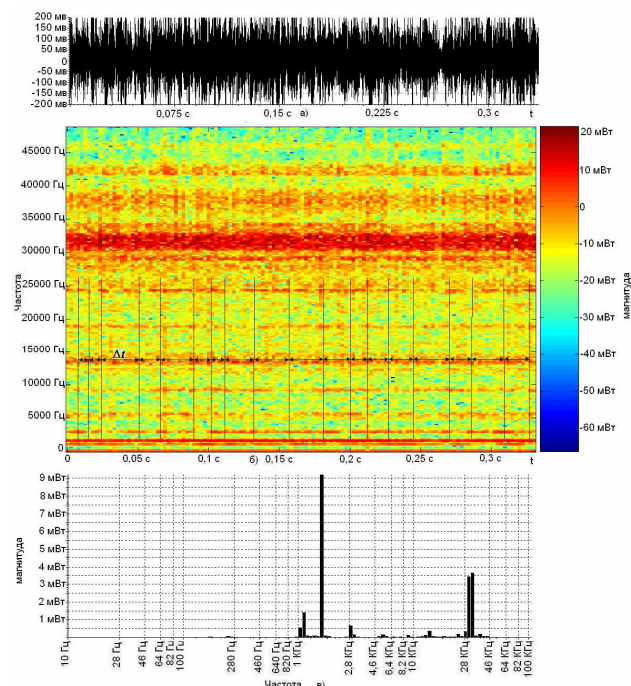


Рис. 6. Частотно-временной анализ и б) спектр мощности в) сигнала ВАЭ а), излучаемой в процессе точения жесткой заготовки, снятой в непосредственной близости от зоны резания, износ инструмента $h_3=0.5$ мм, шероховатость $R_a=4.0$ мкм, показатель Ляпунова $\lambda_1=-0.668861$.

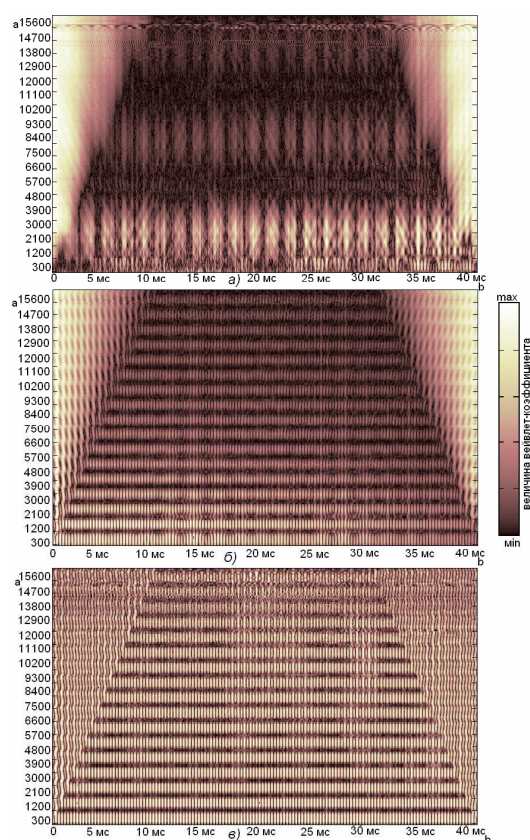


Рис. 7. Многомасштабный вейвлет-анализ сигналов ВАЭ, излучаемой в процессе точения жесткой заготовки и снятого с волнового приемника расположенного на оборотной стороне резцедержателя: а) – износ инструмента $h_3=0$; б) – $h_3=0.15$ мм; в) – $h_3=0.5$ мм

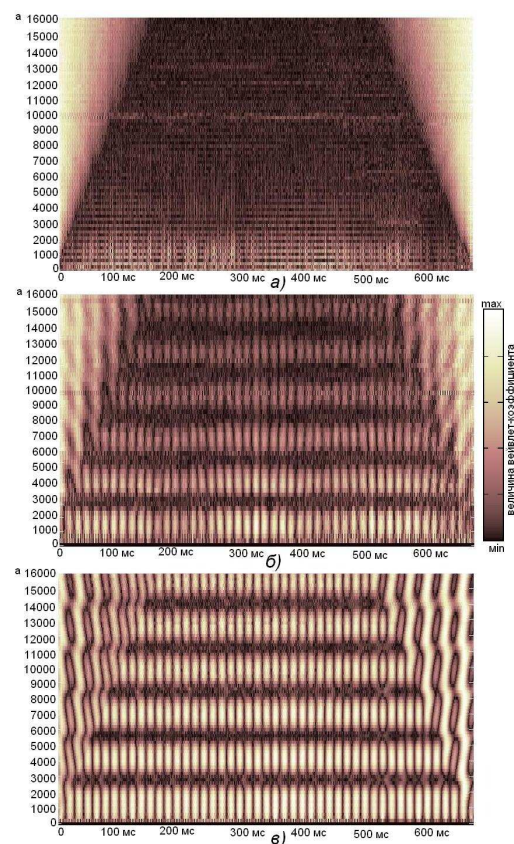


Рис. 8. Многомасштабный вейвлет-анализ сигнала ВАЭ, снятого с волнового приемника установленного на корпусе шпиндельной головки станка 675П концевой фрезой $\varnothing 30$ мм частоте вращения 320 об/мин, подаче 80 мм/мин а) износ $h_3=0$ б) $h_3=0.35$ мм в) $h_3=0.5$ мм.

При минимальном износе (рис. 7, а и 8, а) режущего инструмента сигнала ВАЭ яркие цвета наблюдаются только в начале временного интервала (вейвлет-анализ воспринимает начала сигнала как внезапное появление сигнала от нуля), в конце временного интервала (внезапный спад сигнала до нуля) и в нижней области вейвлет-спектра (рис. 7, а и 8, а).

Светлые участки вейвлет-спектра в нижней его области свидетельствуют о наличии сингулярностей малой интенсивности (рис. 7, а и 8, а). Колебание инструмента происходит наиболее устойчиво, свидетельствуя об оптимальных параметрах обработки. Поэтому наиболее целесообразно применение многомасштабного вейвлет-анализа.

При увеличении степени износа появляются светлые участки вейвлет-спектра в верхней его области, свидетельствуя о наличии сингулярностей большой интенсивности (масштаба) (рис. 7, в и 8, в). В данном случае колебание инструмента происходит наиболее хаотично, свидетельствуя о работе на пределе допустимых параметрах обработки.

Применение многомасштабного вейвлет-анализа возможно также для диагностирования вращающегося инструмента, что не было возможно ни с помощью частотного анализа колебания, ни с помощью какого либо другого метода. Исследование по диагностированию степени износа вращающегося инструмента проводилось на базе широкоуниверсального фрезерного станка при установке датчика на волновом приемнике, закрепленном на корпусе шпиндельной головки. Даже при расположении датчика на таком значительном удалении отчетливо видно появление ярких участков в верхней области вейвлет-спектра (рис. 8, в).

Литература

1. Биленко С.В. Самоорганизующиеся алгоритмы искусственного интеллекта в управлении процессами механической обработки. Диссертация на соискание ученой степени кандидата технических наук. Комсомольск-на-Амуре КНАГТУ 2000.
2. Кабалдин Ю.Г., Олейников А.И., Шпилев А.М., Бурков А.А. Математическое моделирование самоорганизующихся процессов в технологических системах обработки резанием. Владивосток: Дальнаука, 2000. 195с.
3. Кабалдин Ю.Г., Щетинин В.С. Шпилев А.М. Повышение точности обработки в технологических самоорганизующихся системах //Вестник машиностроения 99 №6 с 34-37
4. Кабалдин Ю.Г., Шпилев А.М. Самоорганизующиеся процессы в технологических системах обработки резанием. Диагностика управление. – Владивосток: Дальнаука. 1998. 298с.
5. Подураев В.Н. и др. Технологическая диагностика резания методом акустической эмиссии Подураев В.Н., Базаров А.А., Горелов В.А. – М.: Машиностроение, 1988. -56 с.: ил.
6. Щитов А.Б., Разработка численных методов и программ связанных с применением вейвлет анализа для моделирования и обработки экспериментальных данных. Диссертация на соискание ученой степени кандидата физико-математических наук. – Иваново. Ивановский Государственный университет. 125 с.
7. Отаров С.К. Многопараметрическая диагностика и управление процессом обработки на металлорежущих станках в условиях гибкого автоматизированного производства. Диссертация на соискание ученой степени кандидата технических наук. Ростов-на-Дону ДГТУ 1999.

ИНФОРМАЦИОННОЕ ПИСЬМО

Уважаемый коллега!

Приглашаем Вас принять участие в работе 10-ой Международной конференции и выставки «Цифровая обработка сигналов и ее применение – DSPA'2008», которая состоится в Москве 26-28 марта 2008 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

ПРОГРАММА АДАПТИВНОГО LMS КОМПЕНСАТОРА ДЛЯ КОМПЛЕКСНЫХ ВЫБОРОК НАБЛЮДЕНИЙ

Бартенев В.Г., Бартенев Г.В.

Введение

Одним из наиболее распространенных адаптивных алгоритмов является LMS (Least Mean Square) алгоритм [1], известный также как алгоритм Уидроу-Хоффа. Это наиболее простой и в то же время эффективный способ адаптивной обработки на основе метода наименьших квадратов (МНК). О распространенности этого алгоритма говорит хотя бы тот факт, что он используется почти в каждом сотовом телефоне для компенсации мешающих эхо-сигналов. Не трудно показать, что для комплексных выборок наблюдений LMS алгоритм можно представить в следующем виде

$$\dot{W}(n+1) = \dot{W}(n) + \mu \dot{p}(n) \left[\dot{d}(n)^T - \dot{W}(n)^T \dot{p}(n)^* \right]$$

где $\dot{d}(n)$ - входной обучающий вектор-сигнал; $\dot{p}(n)$ - вектор-сигнал на входе адаптивного КИХ-фильтра; n - номер итерации; $\dot{W}(n)$ - вектор весовых коэффициентов КИХ-фильтра, обновляемый на каждой n -ой итерации; точка сверху означает комплексное число, $(\cdot)^T$ - операцию транспонирования, а $(\cdot)^*$ - операцию комплексного сопряжения.

Именно это выражение и явилось основой для программирования LMS автокомпенсатора на языке MATLAB.

Эти программы приводятся далее в примерах моделирования двух типовых задач, для сигналов, представленных в комплексном виде.

Адаптивная фильтрация помех

Появление в последнее время целых семейств мощных DSP, производимых фирмами Texas Instruments и Analog Devices, а также супер ПЛИС компаний Xilinx и Altera, открыло новые горизонты в создании программируемых систем адаптивной компенсации помех. Проекты, реализованные на DSP, проникают в такие области как адаптивные фазированные антенные решетки, программируемые адаптивные компенсаторы помех, действующие по боковым лепесткам ДНА. И в том и в другом случае полезные сигналы и помехи обычно представляют в комплексном виде, поэтому и в том и другом случае на практике может найти применение рассматриваемый модифицированный LMS алгоритм адаптивной компенсации помех.

При создании таких сложных программируемых устройств на DSP или ПЛИС одной из важнейших проблем является предварительная проверка программируемых алгоритмов. Прежде чем программировать тот или иной алгоритм на DSP, предназначенный для работы в реальном времени, целесообразно этот алгоритм проверить, моделируя его работу с помощью MATLAB. Чтобы

Предложена программа адаптивного LMS алгоритма для комплексных выборок наблюдений. Рассмотрены два примера использования предложенной программы: для адаптивной фильтрации помех и для адаптивной идентификации параметров системы. Данные примеры подтверждают правильность работы программы. Статья имеет учебно-методическое значение.

проиллюстрировать сказанное, рассмотрим практический пример программы компенсации помехи, действующей как в основном, так и в дополнительном канале адаптивного компенсатора с модифицированным LMS алгоритмом. Итак, на входе автокомпенсатора два сигнала: в основном канале искаженный помеховым фоном полезный сигнал, а во вспомогательном канале только помеха (рис. 1).

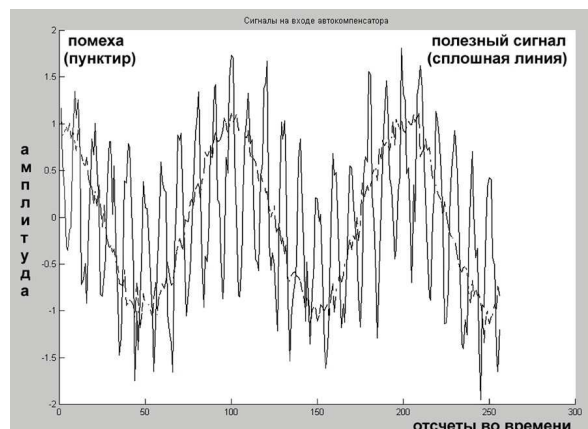


Рис.1 Временная диаграмма сигналов на входе автокомпенсатора

Спектры сигналов на входе автокомпенсатора представлены на рис.2.

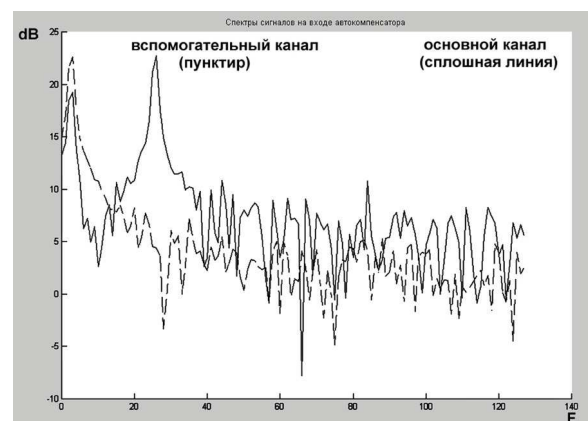


Рис.2 Спектральное представление сигналов на входе автокомпенсатора

Выберем порядок адаптивного КИХ фильтра равным 4 и произведем автокомпенсацию помехи. Сигнал ошибки на выходе автокомпенсатора (рис.3) после настройки весовых коэффициентов фактически и есть полезный сигнал. В подтверждение сказанного рассмотрим спектр сигнала ошибки на выходе (рис.4).

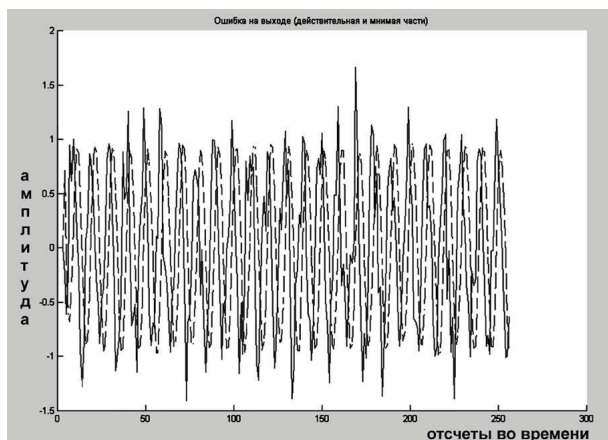


Рис.3 Сигнал ошибки на выходе автокомпенсатора.

Из рис.4 следует – спектральная составляющая помехи подавлена, а спектральная составляющая полезного сигнала присутствует.

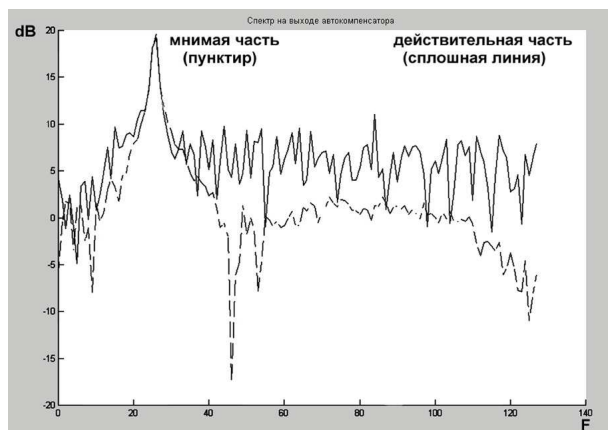


Рис.4 Спектры сигнала ошибки (для действительной и мнимой части).

Ниже приведена моделирующая программа для этого примера.

```
% Моделирование LMS алгоритма
% адаптивная фильтрация комплексных %сигналов
clear all
close all
hold off
sysorder = 4 ;%порядок КИХ фильтра с %перемен-
ными весовыми коэффициентами
N=256;%Число отсчетов на входе
n = 1:N;
%формируем помеху
p=exp(.02*pi*n*sqrt(-1));
%добавляем шум
inp = p+.25*randn(N,1);
%формируем сигнал
s=exp(.2*pi*n*sqrt(-1));
z= randn(N,1); %добавляем шум
% аддитивная смесь
```

```
d=0.5*inp+0.25*z+s';
%на основном входе автокомпенсатора
figure(1)
hold on
plot(real(d(1:256)),'R')
plot(real(inp(1:256)),'B')
D = fft(d',256);
f = 256*(0:127)/256;
INP = fft(inp',256);
f = 256*(0:127)/256;
title('Сигналы на входе автокомпенсатора') ;
figure(2)
hold on
plot(f,10*log10(abs(D(1:128))), 'R')
plot(f,10*log10(abs(INP(1:128))), 'B')
title('Спектры сигналов на входе автокомпенсатора') ;
totallength=size(d,1);
%вводим параметр-объем выборки
%ВСЕ ГОТОВО ЧТОБЫ НАЧАТЬ РАБОТУ
%АВТОКОМПЕНСАТОРА
w = zeros ( sysorder , 1 ) ;
%начальные весовые коэффициенты
%зададим равными нулю
for n = sysorder : N
%основной цикл работы по отсчетам
u = inp(n-1:n-sysorder+1) ;
%отсчеты на входе КИХ фильтра
y(n)= (w).' * conj(u);
%сама фильтрация адаптивным фильтром
e(n) = d(n)' - y(n) ;
%вычисление ошибки
mu=0.15;%
%ИТЕРАЦИИ КОГДА ПРОИЗВОДИТСЯ
%РАСЧЕТ ВЕСОВЫХ КОЭФФИЦИЕНТОВ
w = w + mu * u * e(n) ;
figure(3)
hold on
plot(real(e),'K')
plot(imag(e),'G')
title('Ошибка на выходе') ;
end %Закончили работу
% автокомпенсатора
ER = fft(real(e),256);
f = 256*(0:127)/256;
EI = fft(imag(e),256);
f = 256*(0:127)/256;
figure(4)
hold on
plot(f,10*log10(abs(ER(1:128))), 'K')
plot(f,10*log10(abs(EI(1:128))), 'G')
title('Спектр на выходе автокомпенсатора');
Итак, задав полезный сигнал в виде комплексной
синусоиды, мы его подвергли воздействию помехи, а
затем подали на один из входов автокомпенсатора.
На другой вход автокомпенсатора был подан сигнал
помехи, также представленный в комплексном виде. В
результате чего, уже через несколько итераций поме-
ха отфильтрована автокомпенсатором, а на выходе
остался только полезный сигнал и собственный шум,
что подтверждается спектром сигналов на выходе
автокомпенсатора.
```

Адаптивная идентификация параметров

В данном примере рассматривается задача идентификации параметров системы. Рассматривается случай, когда на входе действует комплексная выборка наблюдений, а в качестве системы рассматривается трехкратная схема череспериодного вычитания (ЧПВ), блок схема которой изображена рис. 5. Задача состоит в адаптивной идентификации параметров ЧПВ или, проще говоря, в нахождении весовых коэффициентов такого КИХ фильтра (рис 5), который был бы эквивалентен трехкратной схеме ЧПВ.

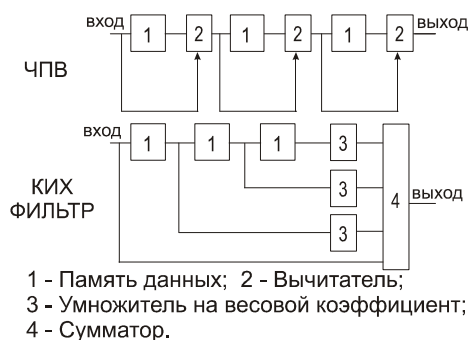


Рис.5 Блок-схема идентифицируемой системы ЧПВ и ее эквивалентный КИХ фильтр, весовые коэффициенты которого – искомые параметры.

Далее приводится текст программы с подробными комментариями.

```
%Моделирование LMS алгоритма для
%адаптивной идентификации параметров %трех-
кратной ЧПВ
clear all
close all
hold off
sysorder = 4 ;%порядок КИХ фильтра с %перемен-
ными весовыми коэффициентами
N=128;%Число отсчетов на входе
s = exp(sqrt(-1))*randn(N,1);
%на входе комплексный сигнал в виде %гауссова
шума
ns = randn(N,1); %собственный шум
%первая ступень вычитания
z1 = filter([1 -1],1,s);
%вторая ступень вычитания
z2 = filter([1 -1],1,z1);
%третья ступень вычитания
z3 = filter([1 -1],1,z2);
nn = ns/10;
%добавляем собственный шум в 10 раз
%ослабив
d = z3 + nn;% аддитивная смесь
figure(1)
hold on
plot(real(d(1:128)), 'R')
plot(real(s(1:128)), 'B')
D = fft(d,256);
f = 256*(0:127)/256;
S = fft(s,256);
f = 256*(0:127)/256;
title('Сигналы на входе автокомпенсатора');
figure(2)
```

```
hold on
plot(f,abs(D(1:128)), 'R')
plot(f,abs(S(1:128)), 'B')
title('Спектры сигналов на входе автокомпенсатора');
totallength=size(d,1);%вводим параметр-объем вы-
борки
w = zeros ( sysorder , 1 ) ;
%начальные весовые коэффициенты
%зададим равными нулю
for n = sysorder : N %основной цикл работы по отсчетам
u = s(n:-1:n-sysorder+1) ;
%отсчеты на входе фильтра
y(n)= (w).' *conj( u);
%сама фильтрация адаптивным фильтром
e(n) = d(n)' - y(n) ;
%вычисление ошибки
mu=0.15;%
w = w + mu * u * e(n) ;
%ЭТО ИТЕРАЦИИ КОГДА
%ПРОИЗВОДИТСЯ РАСЧЕТ ВЕСОВЫХ %
КОЭФИЦИЕНТОВ КИХ ФИЛЬТРА
figure(3)
hold on
title('Ошибка на выходе') ;
plot(real(e), 'K')
plot(imag(e), 'G')
hold on
end
%проверим как настроился
%своими весовыми коэффициентами
%автокомпенсатор
figure(4)
plot([1,-3,3,-1], 'rd')
hold on
plot(real(w), 'k*')
hold on
plot(imag(w), 'rx')
legend('истинные веса','оцененные веса')
title('Сравнение весовых коэффициентов фильтров') ;
grid
axis([0 4 -4 4])
```

Итак, задав помеху в виде белого шума, которая по-
дана на один их входов автокомпенсатора, на другой его
вход была подана та же помеха, но пропущенная через
идентифицируемую систему в виде трехкратной ЧПВ.
На рис.6 изображены спектры помехи до и после про-
пускания ее через трехкратную схему ЧПВ, т.е. спектры
сигналов на входе автокомпенсатора.

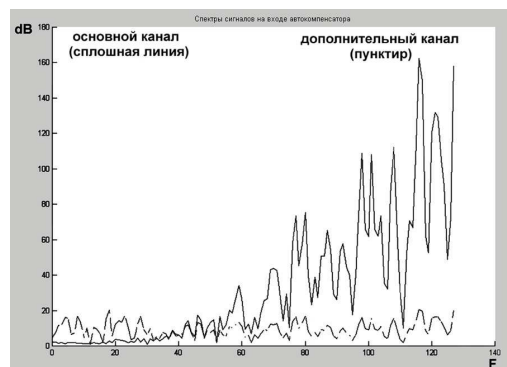


Рис.6 Спектр помехи на входе автокомпенсатора.

В результате уже через несколько итераций произошла настройка автокомпенсатора. На рис. 7 сигнал ошибки на выходе автокомпенсатора это подтверждает.

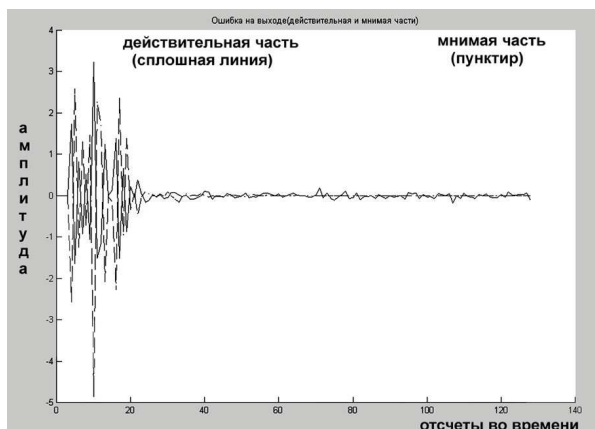


Рис.7 Сигнал ошибки автокомпенсатора (действительная и мнимая составляющие).

При этом, автокомпенсатор настроился так, что весовые коэффициенты его адаптивного КИХ фильтра практически соответствуют весовым коэффициентам идентифицируемого фильтра, т.е. 1, -3, 3, -1.

Заключение

Предложенная программа LMS алгоритма, для обработки сигналов, представленных в комплексном виде, с помощью моделирования в MATLAB прошла всестороннюю проверку. Результаты адаптивного подавления помехи действующей как в основном, так и в дополнительном канале, а также адаптивная идентификация параметров трехкратной череспериодной схемы вычитания подтверждают ее правильность.

Литература

1. Уидроу Б., Стирнз С. Адаптивная обработка сигналов.— М.: Радио и Связь, 1989.— 440с.

Пакет Matrox Inspector – мощный инструмент анализа и обработки изображений

Matrox Inspector - это пакет интерактивного программного обеспечения и библиотека прототипов от признанного лидера, канадской компании Matrox Imaging. Нацеленный как на разработчиков, так и на конечных пользователей (научные и промышленные приложения), пакет Matrox Inspector представляет собой интерактивное приложение под Windows® 2000/XP для захвата, обработки, анализа, отображения и архивирования изображений.

Приложение Matrox Inspector основано на программном обеспечении для разработки Matrox Imaging Library (MIL) и предлагает простой и удобный интерактивный интерфейс для доступа к широкому спектру оптимизированных функций для обработки изображений, вывода закономерностей, анализа пятен, выделения и анализа контуров, измерения, распознавания символов, считывания одно- и двумерных кодов, калибровки и компрессии изображений.

Пользователи библиотек MIL применяют Matrox Inspector в качестве вспомогательного инструмента, позволяющего облегчить и упростить разработку приложений для машинного зрения, анализа изображений и медицинской визуализации. Конечные пользователи, такие как научные работники, технический персонал и инженеры, могут использовать пакет Matrox Inspector для выполнения и автоматизации работ по измерению и улучшению качества изображений.

Ключевые особенности и возможности пакета Matrox Inspector:

- прямое получение изображений от различных источников видео с помощью фрейм- грабберов Matrox;
- загрузка и сохранение изображений, последовательностей изображений в различных форматах (включая AutoCAD® DXF);
- высокая точность и широкий спектр средств для обработки и анализа;
- калибровка изображений для исправления визуальных искажений и выполнение измерений в реальных единицах;
- обработка изображений "вживую" или из архива, включая последовательности;
- статистика слежения, включая скорость в результатах анализа;
- конфигурирование средств анализа для использования с приложениями Matrox Imaging Library (MIL);
- компенсирование изображений текстом и графикой;
- управление и обмен данными с приложениями Microsoft® Windows®; - создание процедур совместимых с Microsoft® VBA или "C" - совместимых.

Текущая версия Matrox Inspector 8.0 включает в себя два новых модуля: String Reader - для ориентированного на топологию оптического распознавания символов и Edge Finder - для выделения признаков. Среди прочих нововведений реализована поддержка цикличной обработки изображений без необходимости создания сценариев (для пакетного тестирования алгоритмов на тестовых последовательностях изображений), как предварительно сохраненных изображений, так и напрямую с камеры, с поддержкой ввода с нескольких камер или систем.

Приложение Matrox Inspector полностью совместимо со всеми аппаратными компонентами Matrox Imaging, включая Matrox Morphis, Solios и Helios. Приложение Matrox Inspector также доступно в версии, поддерживающей семейство Odyssey.

АЛГОРИТМ ПАРАЛЛЕЛЬНОГО ВЫЧИСЛЕНИЯ БЫСТРОГО ПРЕОБРАЗОВАНИЯ ФУРЬЕ ДЛЯ ТРАНСПЬЮТЕРОПОДОБНЫХ СИГНАЛЬНЫХ ПРОЦЕССОРОВ

Нагулин Н.Е., Нагулин С.Н.

Введение

Одним из основных способов повышения общей производительности устройства первичной обработки радиосигналов является распараллеливание вычислений путем использования нескольких одновременно работающих процессоров с упорядоченной системой связей между ними и специализация процессоров под достаточно узкий класс решаемых задач. Очевидно, что эффективность работы многопроцессорной системы существенно зависит при этом от степени соответствия классу решаемых задач структуры процессоров и системы межпроцессорных связей.

Для достижения высокой скорости первичной обработки радиосигналов путем распараллеливания вычислений фирмами Analog Devices, Texas Instruments были разработаны серии цифровых сигнальных процессоров (ЦСП) TMS320C4x и ADSP2106x, имеющих дополнительные каналы связи с другими процессорами. Дополнительные каналы межпроцессорной связи, названные в литературе линк-портами, аналогичны системе связей созданных ранее транспьютеров и позволяют осуществлять обмен данными по принципу «точка-точка» [1]. Соответственно семейство сигнальных процессоров, в состав которых входят линк-порты, часто называют транспьютероподобными процессорами [2].

С помощью линк-портов стало возможным построение многопроцессорных систем различных конфигураций. Однако недостатком линк-портов является сравнительно низкая скорость передачи по сравнению со скоростью передачи по общей параллельной шине. И это обстоятельство очень важно учитывать при распараллеливании вычислений на вышеуказанных процессорах для каждого конкретного алгоритма первичной обработки информации.

Одним из наиболее часто применяемых алгоритмов первичной обработки сигналов является быстрое преобразование Фурье (БПФ). Алгоритм БПФ широко используется для согласованной обработки широкополосных сигналов в частотной области, доплеровской фильтрации, корреляционной обработки, подавления различного рода помех и пр.

Внутренняя архитектура транспьютероподобного процессора в известной степени ориентирована на эффективную реализацию алгоритма БПФ в режиме автономной работы, что нельзя сказать о многопроцессорной системе параллельного вычисления БПФ на основе ЦСП. Целью настоящей работы является рассмотрение способов многопроцессорной реализации БПФ на транспьютероподобных процессорах с использованием линк-портов.

Многопроцессорная реализация БПФ: алгоритм параллельного вычисления

В случае алгоритма БПФ могут быть использованы следующие способы распараллеливания вычислений:

Рассматриваются варианты построения многопроцессорной системы на основе транспьютероподобных сигнальных процессоров серий ADSP2106x, ADSP 21160, TigerSharc (Analog Devices), TMS320C4x (Texas Instruments) для параллельного вычисления БПФ. Описана конфигурация связей и алгоритм обмена данными, при которых требуемая скорость обмена информацией между процессорами минимальна.

- организация многопроцессорной системы с обменом данными по общей параллельной шине;

- построение многопроцессорной системы со специальной системой связей между ними, обеспечивающей максимально эффективную реализацию БПФ.

Основным недостатком реализации БПФ путем использования многопроцессорной системы с обменом данными по общей шине является то, что эффективность распараллеливания вычислений ограничивается пропускной способностью общей шины. Действительно, поскольку для выполнения БПФ необходим обмен между всеми процессорами, то с увеличением числа процессоров должна пропорционально возрастать скорость обмена данными по общей шине. Однако реально скорость передачи данных по общей шине ограничена конструктивными параметрами и не может увеличиваться бесконечно, что в свою очередь приводит к ограничению быстрой реакции всей многопроцессорной системы.

Вычислительная эффективность многопроцессорного устройства со специальной системой связей находится в прямой зависимости от соответствия его структуры информационному графу, при помощи которого описывается решаемая задача или класс решаемых задач. Информационный граф стандартного алгоритма БПФ характеризуется наличием перекрестных связей между вершинами графа, причем вид связей зависит от номера итерации алгоритма БПФ [3]. Соответственно, при реализации стандартного алгоритма БПФ вычислительной системой, состоящей из нескольких процессоров, должна быть предусмотрена коммутационная структура, обеспечивающая в соответствии с алгоритмом БПФ переключение связей между процессорами.

В работе [4] предложена система фиксированных связей между процессорами при выполнении БПФ за счет использования алгоритма БПФ с однородной структурой на всех итерациях. При этом многопроцессорная система включает K однотипных процессоров, причем $K = 2^k$, $k = 1, 2, 3, \dots$. Оптимальная конфигурация связей, при которой скорость обмена информацией между процессорами минимальна, следует из графа алгоритма БПФ основания K и однородной структурой на всех итерациях. В соответствии со структурой данного алгоритма БПФ каждый процессор имеет $K - 1$ двунаправленных портов для обмена данными с дру-

гими процессорами, причем i -й порт j -го процессора связан с j -м портом i -го процессора.

Граф алгоритма БПФ с однородной структурой на всех итерациях для $K = 4$ и длины последовательности сигнала $N = 16$ представлен на рис. 1. Особенностью этого алгоритма является то, что порядок чтения и порядок записи промежуточных результатов вычислений не зависит от номера итерации. Вследствие однородности графа алгоритма БПФ связи между вычислительными модулями остаются фиксированными в процессе вычислений.

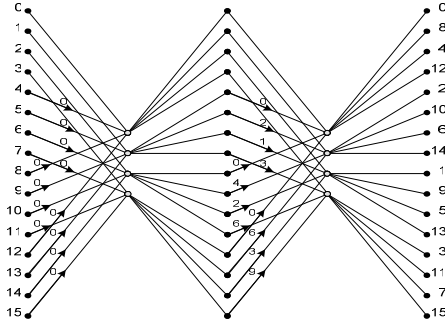


Рис. 1. Граф алгоритма БПФ основания K с однородной структурой на всех итерациях

Граф алгоритма БПФ основания 4 с однородной структурой на всех итерациях, представленный на рис.1, сформирован путем объединения базовых операций алгоритма БПФ основания 2 и однородной структурой на всех итерациях [3]. В отличие от стандартного алгоритма БПФ с основанием 4 [3], у данного алгоритма порядок выходных отсчетов является не четверично-инверсным, а двоично-инверсным. Это достигается путем установки соответствующих информационных связей при выполнении базовой операции алгоритма БПФ. На рис. 2а) показана блок-схема выполнения базовой операции БПФ с основанием 4, при которой формируется двоично-инверсный порядок выходных отсчетов, а ее графическое обозначение, используемое для построения информационного графа алгоритма БПФ, приведено на рис. 2б).

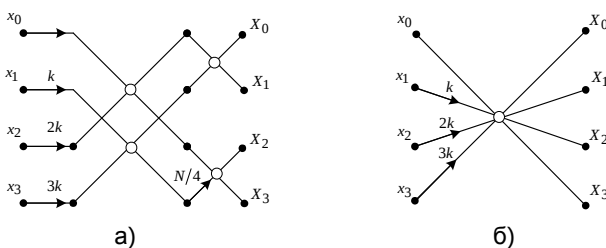


Рис. 2. Блок-схема выполнения базовой операции БПФ с основанием $K=4$ и ее графическое обозначение

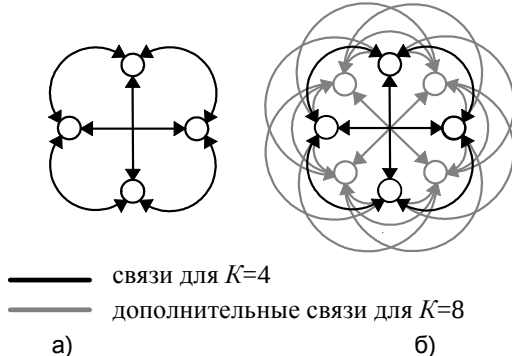


Рис. 3. Структура связей между процессорами для $K=4$ и $K=8$

Структура связей между процессорами для $K = 4$ и $K=8$ представлена соответственно на рис. 3а и рис. 3б.

Для выполнения БПФ внутренняя память каждого процессора разбивается на K блоков $M_1, M_2, \dots, M_K$. Соответственно, перед началом вычислений отсчеты N -точечной последовательности входного сигнала распределяются группами по N/K^2 отсчетов в блоках памяти $M_1, M_2, \dots, M_K$ всех процессоров так, что в i -й блок памяти j -го процессора ($i, j = 1, 2, \dots, K$) записываются отсчеты с номерами:

$$n = \frac{N}{K}(j-1) + \frac{N}{K^2}(i-1) + m, \quad m = 1, 2, \dots, N/K^2,$$

где $n = 0, 1, \dots, N-1$ - номер отсчета входного сигнала.

Если пронумеровать базовые операции у алгоритма БПФ сверху вниз (см. рис. 1 для $N=16$), то j -й процессор ($j = 1, 2, \dots, K$) будет выполнять базовые операции с номерами:

$$l = \frac{N}{K^2}(j-1) + m,$$

где l - номер базовой операции, $l, m = 1, 2, \dots, N/K^2$.

На первой итерации алгоритма БПФ из блоков памяти $1, 2, \dots, K$ процессоров считываются операнды на соответствующие выходные шины. В соответствии со связями между вычислительными модулями, приведенными на рис. 3, на входы j -го процессора по K шинам поступают K операндов из блоков памяти с номером j всех K процессоров для выполнения базовой операции БПФ с основанием K . Отметим при этом, что считывание из j -го блока памяти j -го процессора есть ничто иное, как считывание данных из внутренней памяти процессора.

При выполнении базовой операции алгоритма БПФ в соответствии с блок-схемой, представленной на рис. 2а для $K=4$, производится умножение промежуточных результатов вычислений на соответствующие значения весовых коэффициентов, которые рассчитываются заранее и хранятся в памяти процессоров.

За счет использования оперативной памяти дуплексного типа в каждом вычислительном модуле может быть организована конвейерная обработка таким образом, чтобы одновременно реализовывались следующие этапы вычислений:

- запись в регистры процессора входных операндов, поступающих из секционированной памяти всех K процессоров для выполнения $(l+1)$ -ой базовой операции;
- выполнение в арифметическом устройстве l -ой базовой операции;
- запись в памяти процессора результатов $(l-1)$ -ой базовой операции.

Из-за начальной задержки при конвейерной обработке информации, реализуемой в вычислительных модулях, начало записи в память процессоров результатов промежуточных вычислений задержано относительно начала чтения входных операндов для выполнения базовых операций БПФ на величину $2T_r = 2KT_w$, где T_r и T_w - соответственно периоды чтения и записи данных в память процессора.

Поскольку за время выполнения произвольной l -ой базовой операции в память процессора записывается K операндов, а из блоков памяти всех K процессоров для выполнения $(l+1)$ -ой итерации поступает только по одно-

му операнду, то, следовательно, темп передачи данных по внутренней шине приблизительно в K раз выше, чем темп передачи данных по шинам, связывающим процессоры.

В соответствии с графом алгоритма БПФ, представленного на рис. 1, запись результатов промежуточных вычислений производится последовательно по N/K^2 операндов в блоки внутренней памяти $M_1, M_2, \dots, M_K$ каждого процессора.

Для эффективного формирования весовых коэффициентов, представляющих собой значения комплексной экспоненты, целесообразно использовать следующие свойства, вытекающие из алгоритма БПФ основания K с нормальным порядком входных отсчетов и двоично-инверсным порядком выходных отсчетов:

- порядок следования весовых коэффициентов подчиняется правилу двоичной инверсии;
- на произвольной l -ой итерации ($l = 1, 2, \dots, \log_K N$) соседние базовые операции можно представить как K^{l-1} групп так, что в каждой группе базовые операции имеют одни и те же значения весовых множителей;
- на всех итерациях во всех вычислительных модулях выполняются базовые операции с одинаковыми весовыми коэффициентами (кроме последней итерации).

Заметим, что при данной структуре связей между процессорами, возможна реализация N -точечного БПФ как в режиме постоянного обмена данными между вычислительными модулями, так и путем автономного выполнения каждым вычислительным модулем БПФ размером N/K , с последующим выполнением последней итерации алгоритма БПФ с основанием K в режиме межпроцессорного обмена. Однако во втором случае необходимо перед началом вычислений БПФ выполнить прореживание последовательности отсчетов входного сигнала $x(n), n = 0, 1, \dots, N-1$ таким образом, чтобы в памяти каждого j -го процессора $j = 1, 2, \dots, K$ записались отсчеты входного сигнала с номерами:

$$n = Kp + j - 1, \quad p = 0, 1, \dots, N/K - 1. \quad (1)$$

Действительно, используя соотношение (1), преобразование Фурье входного сигнала можно представить следующим образом:

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N^{kn} = \sum_{s=0}^{K-1} W_N^{sk} \sum_{p=0}^{N/K-1} x(Kp+s) W_N^{Kpk} = \sum_{s=0}^{K-1} Y_s(k) W_N^{sk}, \quad (2)$$

$$k = 0, 1, \dots, N-1,$$

где $W_N^{kn} = \exp(-\frac{2\pi I kn}{N})$, $I = \sqrt{-1}$ - значения комплексной экспоненты.

С учетом свойств комплексной экспоненты выражение (2) можно переписать в виде:

$$X(k) = \sum_{s=0}^{K-1} Y_s(k) W_N^{sk}, \quad s = 0, 1, \dots, K-1, \quad (3)$$

где величины $Y_s(k) = \sum_{p=0}^{N/K-1} x_s(p) W_{N/K}^{pk}$, $k = 0, 1, \dots, N/K - 1$ -

преобразования Фурье последовательностей $x_s(p) = x(Kp + s)$, $p = 0, 1, \dots, N/K - 1$, $s = 0, 1, \dots, K-1$. При фиксированном k выражение (3) представляет собой базовую операцию алгоритма БПФ с основанием K [3].

Организация процесса параллельного вычисления БПФ на транспьютероподобных сигнальных процессорах

серий ADSP2106x, ADSP 21160, TigerSharc (Analog Devices), TMS320C4x (Texas Instruments) представлена на рис. 4. Как видно из рис. 4, для параллельного вычисления БПФ по вышеприведенному алгоритму система линк-портов у каждого процессора задействуется полностью. В первом процессоре (процессор I на рис. 4) для выполнения базовых операций БПФ данные оперативной памяти используются следующим образом: данные внутреннего блока памяти M_1 поступают на арифметическое устройство, а данные блоков $M_2 - M_4$ - на линк-порты процессора $L1 - L3$. Соответственно, во втором процессоре (процессор II на рис. 4) данные внутреннего блока памяти M_2 поступают на арифметическое устройство, а данные блоков M_1, M_3, M_4 - на линк-порты процессора $L1 - L3$ и т.д.

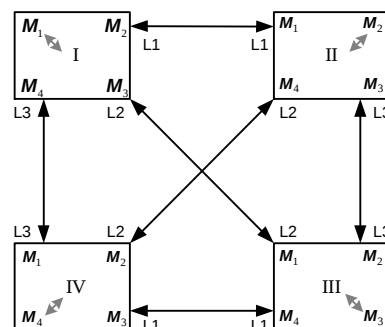


Рис. 4. Организация процесса параллельного вычисления БПФ на транспьютероподобных сигнальных процессорах

Указанная система организации вычислительного процесса БПФ может быть реализована в различных модулях ЦОС, построенных на базе транспьютероподобных сигнальных процессоров. Достаточно широкая номенклатура этих модулей поставляется отечественными и зарубежными производителями.

Заключение

Приведенный алгоритм параллельного вычисления БПФ обладает следующими преимуществами для реализации на транспьютероподобных сигнальных процессорах:

- позволяет использовать в полном объеме систему линк-портов на передачу данных между процессорами;
- без потери производительности темп передачи данных по линк-портам может быть приблизительно в K раз ниже, чем темп передачи данных по внутренней шине.
- поскольку каждый процессор выполняет одинаковый набор операций, алгоритм достаточно прост в реализации.

Литература

1. Руководство пользователя по сигнальным процессорам семейства SHARC ADSP-2106X. Пер. с англ. Бархатов А.В., Коновалов А.А., Петров М.Н. Санкт-Петербург, 2002.
2. Шпаковский Г.И. Параллельные микропроцессоры для цифровой обработки сигналов и медиа данных. - Мн.: БГУ, 2000.
3. Рабинер Л., Гоулд Б. Теория и применение цифровой обработки сигналов / Пер. с англ. Под ред. Ю. И. Александрова. - М.: Мир, 1978.
4. Зайцев Г.В., Нагулин Н.Е. Устройство для быстрого преобразования Фурье. Авторское свидетельство № 1304034, кл. G 06 F 15/332, 1986.

УДК 621.396.96

РЕАЛИЗАЦИЯ МЕТОДА ГУДА-ТОМАСА ДЛЯ ДОПЛЕРОВСКОЙ ФИЛЬТРАЦИИ СИГНАЛОВ

Кошелев В.И., Белокуров В.А.

Введение

Доплеровская фильтрация сигналов является одной из ключевых операций решаемых устройством приема и обработки радиолокационных сигналов, от правильного выбора параметров которой зависит обеспечение требуемых вероятностных характеристик обнаружения целей. Обеспечение этих характеристик для целей с заданной эффективной поверхностью рассеяния на предельной дальности может быть достигнуто при различных сочетаниях параметров РЛС, входящих в уравнение дальности радиолокационного обнаружения (УДРО). Анализ УДРО позволяет построить зависимость дальности обнаружения объекта различных параметров, в том числе от количества импульсов в пачке, используемых при доплеровской фильтрации. При изменении количества импульсов изменяется весь набор параметров, входящих в УДРО. Обычно при доплеровской фильтрации сигналов используется алгоритм быстрого преобразования Фурье (БПФ), поэтому разработчики стремятся выбирать размерность временной выборки (количество импульсов в пачке), равным целой степени 2. При заданной однозначности измерения скорости (в импульсно-доплеровских РЛС) или дальности (в когерентно-импульсных РЛС) объектов и фиксированном времени обзора лоцируемого пространства число импульсов пачки является зависимым от других параметров и при вычислении принимает произвольное значение.

Цель работы состоит в исследовании возможности вычислительно эффективной программной реализации алгоритма БПФ при произвольном числе импульсов пачки.

Выбор числа импульсов зондирующего сигнала

Как отмечено выше, начальным этапом проектирования радиолокационной станции является анализ УДРО [1]:

$$R_{\max} = \sqrt[4]{P_p \tau_{\text{имп}} \frac{G^2 \lambda^2 \sigma_c}{(4\pi)^3 q k N T L_p}}, \quad (1)$$

где $R_{\max}$ – максимальная дальность обнаружения цели; P_p – импульсная мощность передатчика; $\tau_{\text{имп}}$ – длительность импульса; G – коэффициент направленного действия антенны; λ – длина волны излучения; σ_c – ЭПР цели; q – пороговое отношение сигнал/шум по мощности; k – постоянная Больцмана; N – шумовая температура приёмника; T – абсолютная температура; L_p – суммарный коэффициент потерь. Как правило, ряд параметров, таких как, P_p , λ , σ_c , G , N , задаётся разработчику.

Статистический синтез устройства обработки радиолокационного сигнала производится в предположении

Рассмотрен вариант программной реализации быстрого алгоритма доплеровской фильтрации радиолокационного сигнала при произвольном числе импульсов в пачке на основе метода Гуда-Томаса. Его применение позволяет сократить вычислительные затраты по сравнению с широко применяемым алгоритмом Кули-Тьюки.

гауссовской статистики аддитивной смеси сигнала, помехи и шума плотности вероятности [1]:

$$W(\mathbf{X}) = (4\pi)^{-N} \det^{-1} \mathbf{B} \exp\left(-\frac{\mathbf{X}^H \mathbf{B}^{-1} \mathbf{X}}{2}\right),$$

где $\mathbf{X}$ – выборочный n -мерный вектор процесса; $\mathbf{B}$ и $\mathbf{B}^{-1}$ – прямая и обратная матрицы обрабатываемого процесса. Правило решения формируется на основе вычисления отношения правдоподобия (ОП) $L(\mathbf{X}) = W_1(\mathbf{X}/H_1)/W_0(\mathbf{X}/H_0)$, где $W_1()$, $W_0()$ – условные плотности вероятностей при гипотезах H_1 и H_0 , соответственно. Структура оптимальной по критерию максимума отношения правдоподобия обработки определяется после логарифмирования ОП: $\xi(\mathbf{X}) = \mathbf{X}^H \mathbf{Q} \mathbf{X}$, где $\mathbf{Q}$ – матрица обработки. При предположении слабых флуктуаций сигнала в пределах пачки, минимальная достаточная статистика приводится к виду (для случая обработки на фоне шумов или декоррелированной помехи): $\xi(\mathbf{X}) = \mathbf{X}^H \mathbf{S}$, где $\mathbf{S} = \{e^{ik\varphi_c}\}$ – вектор ожидаемого сигнала; φ_c – доплеровский набег фазы отражённого сигнала; $k=0, \dots, N-1$.

На практике параметр φ_c априорно неизвестен, поэтому обнаружение производится в многоканальном по доплеровской фазе φ_c сигнала устройстве. В качестве многоканального фильтра обработки используется процессор БПФ. Пороговое отношение сигнал/шум по мощности для данного случая определяется [1]:

$$q = \frac{1}{N} \left(\frac{\ln(F)}{\ln(D)} - 1 \right),$$

где D – вероятность правильного обнаружения цели (задаётся перед этапом проектирования); F – вероятность ложной тревоги на один канал обнаружения; N – число импульсов в пачке зондирующего сигнала, которое выбирается либо, исходя из необходимости обеспечения заданной дальности обнаружения либо, исходя из заданного времени обзора сектора сканирования. Традиционно в обоих случаях полученное число импульсов округлялось до ближайшей степени 2 (N_1), для выполнения БПФ по алгоритму Кули-Тьюки по основанию 2.



Однако округление количества импульсов в пачке до степени 2 приводит при $N_1 < N$ к некоторому снижению дальности обнаружения и ухудшению разрешения целей по скорости, а при $N_1 > N$ к возрастанию времени обзора сектора сканирования. Уменьшение периода повторения импульсов T_n для сохранения $N_1 = N$ приводит к эффекту неоднозначности на максимальной дальности обнаружения. Известный метод дополнения нулевыми отсчетами обрабатываемой выборки до степени 2 также приводит к ухудшению спектрального разрешения в каналах и другим негативным последствиям.

При разработке программного обеспечения РЛС необходимо учитывать дополнительные вычислительные затраты при реализации алгоритмов, устраняющих перечисленные недостатки. Известно [2], что число вычислительных операций, затрачиваемых при реализации БПФ по основанию 2 $V_{БПФ}$ составляет $V_{БПФ} = 2N \log_2(N) + 3N \log_2(N)$. При использовании корреляционно-фильтровой обработки число операций возрастает: $V_{обр} = N_R V_{БПФ}$, где N_R – число каналов по дальности. При обнаружении маневрирующих целей общее число операций: $V_{общ} = N_R V_{БПФ} M_{уск}$, где $M_{уск}$ – число каналов по ускорению.

Известны эффективные алгоритмы вычисления ДПФ произвольной длины. Одним из таких алгоритмов является алгоритм Винограда, который позволяет значительно (до 80%) сократить число умножений, по сравнению с алгоритмом Кули-Тьюки по основанию 2. Однако данный алгоритм обладает исключительно сложной нерегулярной структурой вычислений, вследствие чего, редко применяется на практике. Наиболее приемлемым с этой точки зрения является алгоритм вычисления БПФ Гуда-Томаса [3]. Ниже рассмотрен один из возможных вариантов его программной реализации на базе сигнального процессора ADSP-21062.

Программная реализация метода Гуда-Томаса на сигнальных процессорах

Как известно [3], алгоритм Гуда-Томаса представляет собой способ отображения линейной последовательности из $n=n_1 \cdot n_2$ целых чисел в двумерную таблицу $n_1 \times n_2$, преобразующего одномерное преобразование Фурье в двумерное преобразование Фурье. При этом числа n_1 , n_2 должны быть взаимно простыми. Способ переупорядочивания входной и выходной последовательностей определяется на основе китайской теоремы об остатках:

перестановки входных индексов:

$$i_1 = i \bmod n_1 \quad i = i_1 N_2 n_1 + i_2 N_1 n_2$$

$$i_2 = i \bmod n_2 \quad N_1 n_1 + N_2 n_2 = 1$$

перестановки выходных индексов определяются по правилу:

$$k_1 = N_2 k \bmod n_1, k = n_2 k_1 + n_1 k_2$$

$$k_2 = N_1 k \bmod n_2$$

Выражение для вычисления преобразования Фурье имеет вид:

$$Y_{k_1, k_2} = \sum_{i_1=0}^{n_1-1} \sum_{i_2=0}^{n_2-1} x_{i_1, i_2} \cdot e^{-j \frac{2\pi}{n_1} i_1 k_1} \cdot e^{-j \frac{2\pi}{n_2} i_2 k_2},$$

в котором задаются n_1 -точечное и n_2 -точечные преобразования Фурье.

С точки зрения повышения быстродействия для вычисления ДПФ по строкам и столбцам наиболее целесообразным является использование набора коротких ДПФ алгоритмов Винограда [3], и структурная схема алгоритма приобретает вид, представленный на рис.1.

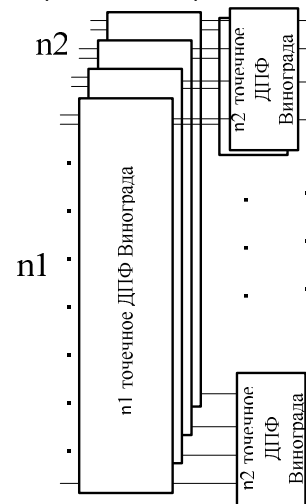


Рис. 1

Число вычислительных операций, затрачиваемых на вычисление коротких ДПФ алгоритмов, представлено в таблице 1.

Число операций, затрачиваемых на реализацию алгоритма Гуда-Томаса, составляет:

$$C = n_1 \cdot C_2 + n_2 \cdot C_1,$$

где C_1 и C_2 определяют число вычислительных операций, затрачиваемых на реализацию n_2 и n_1 точечных ДПФ.

Таблица 1

Порядок ДПФ	Число умножений	Число сложений	Число «сложений/вычитаний»
3	4	12	4
5	10	34	12
7	16	72	20
8	4	52	26
16	20	148	68

Таблица 2

БПФ Гуда-Томаса	Суммарное число операций	БПФ Кули-Тьюки	Суммарное число операций
15 точек	222	16 точек	320
40 точек	632	64 точки	1920
80 точек	1544	128 точки	4480

В таблице 2 представлены сравнительные данные с БПФ Кули-Тьюки по основанию 2:

При программной реализации алгоритма необходимо учитывать отсутствие в алгоритме Гуда-Томаса битреверсивной адресации. Также необходимо отметить возможность вычисления операции типа «сложение/вычитание», а также операции чтения и записи из памяти программы или памяти данных за один такт. При этом вычислительная эффективность по тактам будет отличаться от представленных в таблице 2 данных.

Рассмотрим пример $n=15$, $n_1=3$, $n_2=5$. Перестановки входных и выходных отсчётов будут определяться правилами:

$$\begin{bmatrix} 0 & 6 & 12 & 3 & 9 \\ 10 & 1 & 7 & 13 & 4 \\ 5 & 11 & 2 & 8 & 14 \end{bmatrix} \text{ и } \begin{bmatrix} 0 & 3 & 6 & 9 & 12 \\ 5 & 8 & 11 & 14 & 2 \\ 10 & 13 & 1 & 4 & 7 \end{bmatrix}, \text{ соответственно.}$$

Данные двумерные таблицы преобразуются в одномерные для циклов по n_1 и по n_2 , а также для заключительной переадресации:

.VAR input_3dft[15] = 0, 10, 5, 6, 1, 11, 12, 7, 2, 3, 13, 8, 9, 4, 14;

.VAR input_5dft[15] = 0, 6, 12, 3, 9, 10, 1, 7, 13, 4, 5, 11, 2, 8, 14;

.VAR output_dft[15] = 0, 5, 10, 3, 8, 13, 6, 11, 1, 9, 14, 4, 12, 2, 7.

Для повышения быстродействия тексты малых ДПФ Винограда размещаются в циклах.

```
lcntr = 5, do cikl_5iter_3dft until lce;
  f0 = dm(m1,i0);
  f1 = dm(m2,i0);
  f0 = f0 + f1, f1 = f0 - f1, f2 = dm(m0,i0);
  f2 = f2 + f0;
  f0 = f0 * f15, dm(m0,i2) = f2;
  f8 = dm(m1,i1);
  f9 = dm(m2,i1);
  f8 = f8 + f9, f9 = f8 - f9, f10 = dm(m0,i1);
  f10 = f10 + f8;
  f8 = f8 * f15, dm(m0,i3) = f10;
  f9 = f9 * f13;
  f1 = f1 * f14;
  f2 = f2 + f0;
```

```
f2 = f2 + f9, f9 = f2 - f9;
f10 = f10 + f8, dm(m2,i2) = f2;
f10 = f10 + f1, f1 = f10 - f1, dm(m1,i2) = f9;
dm(m2,i3) = f10;
dm(m1,i3) = f1;
r0 = dm(i4,m4);
r1 = dm(i4,m4);
r2 = dm(i4,m4);
m0 = r0;
m1 = r1;
```

cikl_5iter_3dft: m2 = r2;

Результаты вычисления коротких ДПФ Винограда размещаются в промежуточных буферах для действительной и мнимой частей, соответственно.

Заключительным этапом преобразования является выходная перестановка отсчётов.

// Перестановка выходных индексов

```
lcntr = 15, do label_out until lce;
  r0 = dm(i4,m4);
  f0 = dm(m0,i0);
  dm(i2,m2) = f0;
  f0 = dm(m0,i1);
  label_out: dm(i3,m3) = f0;
```

Время выполнения ДПФ на 15 точек составило 456 тактов, время вычисления БПФ Кули-Тьюки на 16 точек составило 740 тактов.

Таким образом, программная реализация алгоритма Гуда-Томаса имеет простую структуру и требует многократного использования коротких ДПФ Винограда. Целесообразным является использование коротких ДПФ Винограда на 4, 8, 16 точек, так как данные алгоритмы содержат значительное число операций «сложение/вычитание».

Литература

1. Бакулев П. А. Радиолокационные системы. М.: Радиотехника. 2007. 376 с.
2. Блейхут Р. Быстрые алгоритмы цифровой обработки сигналов. М.: Мир, 1989. 451с.
3. Нуссбаумер Г. Быстрое преобразование Фурье и алгоритмы вычисления свёрток. М.: Радио и связь, 1985. 248с.

«ЦОС В УНИВЕРСИТЕТАХ» - НОВАЯ РУБРИКА ЖУРНАЛА

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 «АВТЭКС», 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 открывает 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.

Новая рубрика открывается информацией об Университетской программе компании Texas Instruments Inc. (США), любезно предоставленной ее менеджером в странах Европы и Северной Африки Робертом Оуэном и координатором Программы в России и странах СНГ А. Петровским. 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 Владимир В. Витязев

УНИВЕРСИТЕТСКАЯ ПРОГРАММА КОМПАНИИ TEXAS INSTRUMENTS, часть 1

Компания Texas Instruments Incorporated – мировой лидер в области технологий цифровой обработки сигналов, аналоговых устройств и полупроводниковой техники – проводит на территории Европы, Ближнего Востока и Африки программу университетского образования, призванную обеспечить поддержку образовательных курсов и научно-исследовательских работ учебных заведений, посвященных передовым методам цифровой обработки сигналов и информационным технологиям реального времени.

Почему компания Texas Instruments уделяет большое внимание образованию? Во-первых, Texas Instruments заинтересована в том, чтобы научить следующее поколение инженеров, а пока еще студентов, работать с технологиями завтрашнего дня. Во-вторых, осуществляя поддержку научно-исследовательской деятельности ВУЗов, Texas Instruments способствует расширению сфер применения технологий реального времени и стимулирует инновации в этой области. В-третьих, программа компании позволяет в рамках научно-исследовательских проектов получать образцы готовых изделий, что способствует взаимодействию учебных заведений с промышленностью и внедрению новых наукоемких технологий в производство. При решении всех этих задач компания Texas Instruments опирается на свою продукцию, укрепляя, тем самым, собственные позиции на мировом рынке.

Рассмотрим конкретные механизмы реализации Университетской программы компании Texas Instruments. К ним относятся:

- специальные цены для ВУЗов (University Pricing – университетский прайс-лист);

- возможность общения со специалистами Texas Instruments и своими коллегами;
- обучающие программы (повышение квалификаций преподавателей);
- специальные обучающие диски (Teaching ROMs), материалы семинаров, конференций, книги, статьи и др.;
- техническая поддержка от специалистов Texas Instruments;
- оборудование для реализации научно-исследовательских проектов, подготовки диссертаций или написания дипломных работ;
- цифровые и аналоговые средства для организации учебных лабораторий.

Предоставляя все это университетам, компания Texas Instruments ожидает от них только одного: максимально широкого использования современных технологий обработки сигналов реального времени на базе продуктов Texas Instruments в учебном процессе и научно-исследовательских работах.

В настоящий момент в Университетской программе Texas Instruments задействовано более 1200 технических университетов в Европе. В России насчитывается около 50 лабораторий, оборудованных при содействии компании TI в различных ВУЗах страны, и мы надеемся размещать информацию о них на страницах журнала или сайта www.dsps.ru/cosvuz.

Остановимся на каждом перечисленном пункте программы Texas Instruments более подробно.

Университетский прайс-лист. Аппаратные и программные средства разработки компании Texas Instruments могут быть приобретены по специальным ценам. Перечень

предлагаемых инструментов (не все продукты Texas Instruments попадают в данную категорию) и их стоимость по Университетской программе можно найти на сайте: www.ti.com/pricelist. Например, среда разработки и отладки программного обеспечения Code Composer Studio стоимостью более трех с половиной тысяч долларов может быть приобретена в рамках Университетской программы по цене девятьсот долларов (приводятся ориентировочные цены), а стартовый набор разработчика на процессоре C6713 стоимостью триста девяносто пять долларов – по цене триста пятьдесят пять долларов (приводятся ориентировочные цены). При рассмотрении вопроса о поставке оборудования в тот или иной ВУЗ компания TI всегда готова к поиску компромиссных решений, особенно в тех ситуациях, когда заметна самостоятельная активная работа ВУЗа в соответствующем направлении.

Оплата и приобретение продукции Texas Instruments по Университетской программе осуществляется через авторизованных дистрибьюторов компании в России. Наиболее зарекомендовали себя в работе с ВУЗами следующие фирмы: ЗАО «Сканти-рус» (www.scanti.ru) и «Компэл» (www.compel.ru).

Обучение. Компания Texas Instruments дает возможность преподавателям и сотрудникам ВУЗов повысить свою квалификацию в области технологий реального времени. С этой целью организуются практические семинары, на которых участникам предоставляется возможность поработать с аппаратными и программными средствами Texas Instruments под руководством специалистов. Такие семинары проводятся компанией достаточно часто в различных городах мира: Париже, Мюнхене, Милане, Тель-Авиве, Москве. В России подобные мероприятия проводились уже неоднократно и начинают принимать периодический характер. Узнать о семинарах можно на сайте Texas Instruments: www.ti.com/europe/workshops, а так же в новостях журнала «Цифровая обработка сигналов»: www.dsps.ru.

Семинары (workshops) дают возможность получить практические навыки работы с инструментальными средствами Texas Instruments. Это могут быть или однодневные технические семинары, или четырехдневные семинары по проектированию. Они охватывают весьма широкую тематику, начиная от программирования DSP и микроконтроллеров конкретных платформ и заканчивая аналоговой техникой и вопросами проектирования систем обработки сигналов в различных приложениях.

Семинары проводятся компанией не только для университетов, но и для представителей коммерческих организаций. Однако для ВУЗов участие в подобных мероприятиях возможно с 50%-ой скидкой. Кроме того, для представителей ВУЗов всегда выделяются бесплатные места.

Очень эффективным источником технической и любой другой информации, касающейся продукции Texas Instruments, является Европейский центр поддержки. В случае каких-либо затруднений или вопросов, связанных с использованием продуктов компании, всегда есть возможность позвонить по телефону или отправить письмо по электронной почте eric@ti.com и получить квалифицированную помощь технических специалистов Texas Instruments. Чрезвычайно важным для Российского пользовате-

ля является то, что телефонный звонок в Европейский центр поддержки Texas Instruments может быть осуществлен по московскому номеру +7 (495) 981 0701, и разговор ведется на русском языке. Преподавателям и сотрудникам ВУЗов не следует давать эту контактную информацию студентам, чтобы на специалистов центра не обрушился поток ученических вопросов. Следует соблюдать правило: студент задает вопросы преподавателю, преподаватель, при необходимости консультируется со специалистами Texas Instruments.

Предлагаемые средства. В рамках Университетской программы компания Texas Instruments предлагает все свои базовые аппаратные и программные инструментальные средства разработки. В области цифровых сигнальных процессоров это, в первую очередь, стартовые наборы разработчика (DSK) стоимостью не более 500 долларов. Такие наборы являются идеальным решением для быстрой организации лабораторных курсов по различным дисциплинам, связанным с обработкой сигналов. В семействе процессоров C5000, ориентированных на телекоммуникационные приложения, предлагаются наборы DSK на базе процессоров C5416 и C5510. В семействе C6000 – наиболее производительном классе DSP компании Texas Instruments – предлагаются DSK C6713 и C6416. В семействе сигнальных контроллеров C2000, используемых в системах управления, доступны DSK LF2407 и F2812.

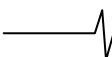
Более сложными модулями, стоящими по несколько тысяч долларов, и ориентированными на научно-исследовательские проекты, являются оценочные модули EVM. В семействе C5000 – это плата на базе DSP C5509 со специальными инструментами по оптимизации потребления энергии. В семействе C6000 предлагается целый ряд плат, ориентированных на обработку видео и аудиоданных. В семействе C2000 предлагается плата на базе процессора F2812.

Кроме перечисленных инструментов еще одним аппаратным средством разработки являются эмуляторы, позволяющие работать с процессором, встроенным в конечный продукт (плату собственного изготовления). Стоимость экземпляров лежит в диапазоне одной-двух тысяч долларов.

Совместно с перечисленными средствами или независимо от них поставляется программная среда разработки и отладки приложений на базе DSP Texas Instruments под названием Code Composer Studio.

Остановимся на некоторых из перечисленных средств более подробно.

Стартовый набор разработчика на процессоре TMS320C2812 предназначен для реализации приложений цифрового управления. За 335 долларов* вы приобретаете плату с 32-битным сигнальным контроллером с фиксированной точкой, работающим на частоте 150 МГц и имеющим внутреннее ОЗУ 18 К слов и внутреннее ПЗУ Flash объемом 128 К слов. Последнее позволяет плате выполнять ваши программы, будучи автономным устройством. На плате также имеется микросхема внешней памяти ОЗУ объемом 64 К слов и интерфейс JTAG, обеспечивающий связь платы с персональным компьютером для полнофункциональной разработки, отладки и оптимизации программного обеспечения в среде Code Composer Studio, входящей в состав набора. Имеющиеся стандартные ин-



терфейсы дают возможность подключать к плате внешние аналоговые и цифровые схемы и быстро создавать на ее основе законченную систему обработки сигналов, демонстрирующую эффективность современных технологий (в рамках лабораторных курсов) или решающую требуемые технические задачи (в рамках дипломного проектирования или исследовательских работ).

Стартовые наборы на процессорах TMS320C6713 (плавающая точка) и TMS320C6416 (фиксированная точка) сохраняют принципы организации набора C2812, но относятся к самому мощному классу процессоров TI и ориентированы, в первую очередь, на обработку аудиосигналов. Рабочие частоты устройств – 225 МГц и 1 ГГц, соответственно (имеется также стартовый набор разработчика на процессоре TMS320C6455 с тактовой частотой 1.2 ГГц). Стоимость плат – 355 и 445 долларов*. На платах размещен кодек, реализующий аналого-цифровое и цифро-аналоговое преобразования. Его выводы заведены на стандартные разъемы аудиосигналов. Благодаря этому очень просто становится организовать типовую схему цифровой обработки сигнала: аналоговый вход, АЦП, обработка на цифровом процессоре, ЦАП, аналоговый выход, работающие в реальном времени и наглядно демонстрирующие функционирование системы. Типовым случаем является подключение к аналоговому входу платы микрофона или звуковой платы персонального компьютера, а к выходу – колонок. В этом случае становится очень удобно воспринимать на слух те изменения, которые вносит ваша программа в исходный сигнал – произносимую в микрофон речь или проигрываемое на компьютере музыкальное произведение. При этом все входные, промежуточные и результирующие сигналы и их спектры могут быть графически отображены в окнах среды Code Composer Studio.

Относительно новым устройством, предлагаемым компаниями TI в качестве стартового набора разработчика, является платформа разработки систем цифрового видео на процессоре TMS320C6437. Стоимость данной платы составляет 495 долларов*. Ее уникальность заключается в наличии не только аудио, но и стандартных видеоразъемов. Таким образом, плата становится ориентированной на обработку видеосигналов и изображений, а это дает стартовым наборам разработчика новый уровень функциональности и демонстративности. В состав платформы входит широкий набор типового программного обеспечения обработки звуковых, речевых и видеосигналов, в частности кодеки H.264, MPEG-4 и G.711.

Отдельно выделим программную среду разработки и отладки программного обеспечения DSP Code Composer Studio. Она может быть приобретена отдельно и дать огромные возможности по обучению разработке программ для DSP. Входящие в ее состав программы-симуляторы позволяют полноценно имитировать работу с процессором любого семейства, не имея аппаратного обеспечения. Можно сказать, что среды Code Composer Studio может быть вполне достаточно для глубокого изучения архитектуры процессора и обучения работе с ним. Аппаратные средства, такие как DSK, становятся необходимы лишь для наглядности работы системы ЦОС и тогда, когда необходимо иметь дело с реальными аналоговыми сигналами. В рамках Университетской программы TI возможно приобретение интегрированной среды разработки Code Composer Studio с необходимым количеством авторизаций для лаборатории.

К другим программным инструментальным средствам разработки, заслуживающим особого внимания, хотелось бы отнести связь мощной и популярной системы МАТЛАБ со средой Code Composer Studio и шаблоны программного обеспечения Reference Frameworks. Связь МАТЛАБ и CCS дает уникальные возможности перехода от математики и исследований алгоритма к реализации системы на аппаратной платформе, например, на плате DSK. В частности, специальный демонстрационный диск с системой МАТЛАБ, поддерживающей взаимодействие с Code Composer Studio, прилагается к стартовому набору разработчика на процессоре TMS320C6713. Заметим, что данный инструмент, несмотря на всю свою эффективность и привлекательность, не является «волшебной палочкой» превращающей любой алгоритм в эффективно работающий проект на цифровом сигнальном процессоре. Во-первых, чтобы хорошо освоить этот инструмент понадобится немало времени. А во-вторых, все равно его основной задачей останется ускорение процесса разработки программного обеспечения; помощь в создании программных кодов, но не замена этих процедур.

Шаблоны программного обеспечения, так же как и другие типовые программные модули от компании Texas Instruments, могут оказать существенную помощь при начальном ознакомлении с программированием DSP. Они закладывают фундамент, на основе которого можно создавать собственные программы, опираясь на надежный, проверенный и эффективный код.

На этом мы заканчиваем первую часть статьи, посвященной обзору Университетской программы компании Texas Instruments Inc. Вторая часть готовится к публикации в следующем номере журнала «Цифровая обработка сигналов». В ней мы постараемся осветить еще ряд средств, доступных для приобретения ВУЗами в рамках университетской программы, в частности, среди аналоговых устройств и микроконтроллеров; расскажем об обучающих дисках, доступных для преподавателей и представителей ВУЗов; о вознаграждении, на которое могут рассчитывать наиболее активные и продуктивные студенты, а также подведем итог изложенному материалу – ответим на вопрос: как подключиться к Университетской программе TI и что конкретно дает участие в этой программе.

Настоящее информационное сообщение подготовлено на основе материалов доклада менеджера Университетской программы TI в странах Восточной и Западной Европы и Северной Африки Роберта Оуэна (rcwo@ti.com) и координатора Университетской программы TI в России и странах СНГ Алексея Петровского (petrovsky@bsuir.by). Еще раз обращаем ваше внимание на открытие нового раздела сайта журнала «Цифровая обработка сигналов» под названием «ЦОС в университетах». На его страницах мы намерены давать информацию об университетских программах фирм-производителей оборудования ЦОС и рассказывать о научных и учебных лабораториях, имеющих и появляющихся в российских ВУЗах. Все заинтересованные организации приглашаем к сотрудничеству.

** указанные в статье цены являются ориентировочными, взятыми с сайта компании TI в декабре 2007 года. Чтобы узнать цены, действующие на текущий момент на территории Российской Федерации, обращайтесь представителям TI в России.*