

УДК 621.396.96 - 621.397

ЦИФРОВАЯ СИСТЕМА ГИБРИДНОГО ПОЗИЦИОНИРОВАНИЯ ГРУППЫ МОБИЛЬНЫХ РОБОТИЗИРОВАННЫХ АППАРАТОВ НА ОСНОВЕ ВЗАИМНЫХ РАДИОИЗМЕРЕНИЙ ДАЛЬНОСТЕЙ И ИЗОБРАЖЕНИЙ ОКРУЖАЮЩЕЙ ОБСТАНОВКИ

*Кириллов С.Н., д.т.н., профессор, зав. кафедрой Рязанского государственного радиотехнического университета,
e-mail: kirillov.lab@mail.ru;*

*Бердников В.М., к.т.н., доцент Рязанского государственного радиотехнического университета,
e-mail: bvm1986@mail.ru;*

*Акопов Э.В., инженер Рязанского государственного радиотехнического университета,
e-mail: akop78195@gmail.com.*

DIGITAL HYBRID POSITIONING SYSTEM OF MOBILE ROBOTS GROUP BASED ON RADIORANGING AND IMAGES OF ENVIRONMENT

Kirrilov S.N., Berdnikov V.M., Akopov E.V.

The hybrid positioning system of mobile robots group based on radioranging between robots and videoinformation from cameras. It shown, that achieved the ability to estimate coordinate robots with error around 3 m in rugged terrain and urban area.

Key words: videopositioning, group positioning, area relief contours.

Ключевые слова: видеопозиционирование, групповое позиционирование, контуры рельефа местности.

Введение

Позиционирование наземных мобильных роботизированных аппаратов (МРА) является актуальной проблемой ввиду интенсивного развития различных приложений робототехники [1, 2]. При этом возникают особенности управления, навигации и позиционирования МРА в составе группы [3].

Широко используемые решения на основе спутниковой навигации уязвимы к действию помех, особенно в городской местности, где появляются такие дополнительные отрицательные факторы как многолучевость и затенение сигналов высокими зданиями. В целях позиционирования группы МРА актуально применение взаимных измерений дальностей по радиоканалу между разными МРА группы, на основе которых можно получить информацию о расположении роботов относительно друг друга, однако при этом требуется последующая привязка к окружающей местности.

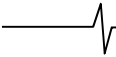
В урбанизированной среде перспективными являются системы позиционирования и навигации на основе видеоданных, благодаря таким преимуществам как автономность, малозаметность, а также относительно невысокая стоимость элементов [1, 2]. В работах [4, 5, 6] предложена система видеопозиционирования (СВП) на основе анализе контуров рельефа местности, выделяемых на изображениях окружающей среды и цифровой карты (ЦК) с указанием высот представленных на ней объектов. В этом случае основная информация, используемая при позиционировании, получается при выделении на изображениях, сигнатур контуров рельефа местности (СКРМ), представляющих векторы отсчетов $A_w(n)$, $n = 1, N$ (N – число столбцов изображения),

Представлена система гибридного позиционирования группы мобильных роботизированных аппаратов в городской местности, использующая взаимные радиоизмерения дальностей совместно с видеоинформацией, полученной из бортовых видеокамер МРА. Показано, что достигается возможность определения координат группы МРА с ошибкой порядка 3 м в слабопересеченной и урбанизированной местности.

каждый элемент которых равен номеру строки содержащей искомый контур в соответствующем столбце n -того изображения [4]. При этом наличие мешающих объектов (МО), заслоняющих и искажающих контур зданий, приводит к появлению направлений, в которых невозможно определить положения контура. Для учета данного фактора вводятся сигнатуры неинформативных направлений (СНН), которые представляют вектор отсчетов $X(n)$, $n = 1, N$, (N – число столбцов изображения), причем $X(n) = 1$ в случае если n -той строке искомый контур не обнаружен $X(n) = 0$ в противном случае [5].

Преимущества данной СВП заключаются в использовании сигнатур, представляющих информацию об объектах сцены и занимающих небольшой объем, что позволит увеличить быстродействие и уменьшит требуемый размер памяти при реализации. При этом обеспечивается точность определения координат в городской местности до одного метра без использования данных по азимуту. Однако эффективное функционирование такой СВП требует выполнения условий удовлетворительной видимости, а также наличия объектов застройки в наблюдаемой сцене [4].

В известных работах не рассмотрены методы видеопозиционирования при наличии дополнительной информации о взаимном расположении объектов в группе, полученных на основе радиоизмерений. При этом получение информации о контурах рельефа местности в нескольких разных точках с учетом взаимного расположе-



ния их друг относительно друга, позволит использовать меньший объем данных на каждом МРА. Это приведет к возможности работы при больших значениях доли неинформативных направлений, то есть к повышению устойчивости к действию МО и к возможности работы системы на открытой местности с редко встречающимися объектами застройки.

Цель работы: разработка алгоритма позиционирования группы МРА на основе данных об ориентации и взаимном расположении МРА, использующего информацию о контурах рельефа местности на изображениях окружающей обстановки, полученных из всех объектов группы.

Постановка задачи

Ставится задача определения координат $x_{1...W}, y_{1...W}$ МРА группы из W роботов на основе следующих исходных данных:

- расстояний между роботами группы $d_{1...W}^1, d_{1...W}^2, \dots, d_{1...W}^w, \dots, d_{1...W}^W$ (d_k^w между w -тым и k -тым роботом, измеренные при использовании радиоканала),
- изображений $I_{1...W}$, полученных при помощи одной или нескольких видеокамер;

- углов ориентации роботов (азимут $\varphi_{1...W}$, угол возвышения $\theta_{1...W}$, крен $\psi_{1...W}$), которые могут быть получены при использовании магнитометра и инклинометра.

Представленные данные передаются с каждого МРА на пункт управления, где производятся вычисления, направленные на решение данной задачи и включающее два этапа:

- Определение взаимного расположения роботов группы относительно друг друга на основе радиоизмерений (оценка относительных координат $\tilde{u}_{1...W}, \tilde{v}_{1...W}$ роботов внутри группы при использовании оценки расстояний $d_{1...W}^1, d_{1...W}^2, \dots, d_{1...W}^w, \dots, d_{1...W}^W$).

- Оценка абсолютных координат МРА группы ($x_{1...W}, y_{1...W}$) на основе информации о контурах рельефа местности, полученной из изображений, данных о ориентации МРА и цифровой карты местности. Выделение информации о контурах рельефа местности также может производиться на каждом МРА.

При оценке относительных координат МРА группы за точку отсчета принимается положение одного из роботов, где выбираются направления осей координат u и v (рис. 1) формирующих групповую систему координат $uO'v$.

Оценки ($\tilde{u}_{1...W}, \tilde{v}_{1...W}$) производится при решении системы нелинейных уравнений, связывающей расстояния между МРА и их относительные координаты:

$$\begin{cases} \sqrt{(\tilde{x}_1 - \tilde{x}_2)^2 + (\tilde{y}_1 - \tilde{y}_2)^2} = d_1^2, \\ \dots \\ \sqrt{(\tilde{x}_w - \tilde{x}_k)^2 + (\tilde{y}_w - \tilde{y}_k)^2} = d_k^w, \\ \dots \\ \sqrt{(\tilde{x}_W - \tilde{x}_{W-1})^2 + (\tilde{y}_W - \tilde{y}_{W-1})^2} = d_{W-1}^W; \end{cases} \quad (1)$$

при использовании численных методов Ньютона, Левенберга-Марквардта и др.

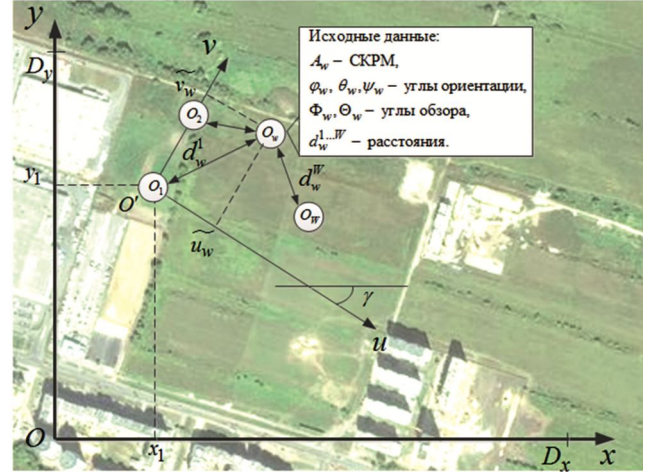


Рис. 1. Групповая и локальная система координат

Связь между абсолютными и относительными координатами группы определяется выражением:

$$\begin{cases} x_w = x_1 + u_w \cdot \cos(\gamma) + v_w \cdot \sin(\gamma), \\ y_w = y_1 - u_w \cdot \sin(\gamma) + v_w \cdot \cos(\gamma); \end{cases} \quad (2)$$

где (x_w, y_w) – координаты w -того робота в абсолютной системе координат xOy , γ – угол между xOy и $uO'v$;

(u_w, v_w) – истинные значения относительных координат w -того робота, причем

$$\begin{aligned} \tilde{u}_{1...W} &= u_{1...W} + \delta_{1...W}^u, \\ \tilde{v}_{1...W} &= v_{1...W} + \delta_{1...W}^v; \end{aligned} \quad (3)$$

где $\delta_{1...W}^u, \delta_{1...W}^v$ – ошибки оценки относительных координат. Таким образом, на этапе оценка абсолютных координат МРА требуется оценить координаты одного из роботов (x_1, y_1) и ориентацию группы в азимутальной плоскости (угол между локальной и групповой системами координат γ). Параметр γ может быть известен, на пример, при наличии прямой видимости, позволяющей измерить азимут между любыми двумя МРА группы.

Выделение сигнатур изображений местности

Выделение сигнатур $A_{1...W}$ и $X_{1...W}$ включает распознавание участков изображения $I_{1...W}$ содержащих границу зданий и фона бинаризацию и устранение мелких мешающих объектов путем морфологической обработки [5].

При этом для оценки эффективности алгоритма выделения контуров рельефа местности (АВКРМ) используются следующие параметры:

- угловая ошибка вычисления сигнатур, вычисляемая по формуле:

$$D_s = \frac{1}{N} \cdot \sum_{n=0}^{N-1} |A_w - A_w^t| \cdot (1 - X(n)), \quad (4)$$

где A_w^t – неискаженная СКРМ;

- доля неинформативных направлений (нормированная суммарная величина неиспользуемых столбцов изображения, возникающих в результате наличия крупных МО на изображениях):

$$S_X = \left(\sum_{n=0}^{N-1} X_n \right) / N. \quad (5)$$

В работах [4, 5, 6] рассмотрен АВКРМ, обеспечивающий получение СКРМ и СНН, имеющая угловую ошибку вычисления сигнатур, $D_S < 2,5^\circ$ при доле неинформативных направлений $S_X < 0,25$.

Оценка абсолютных координат группы

Оценка абсолютных координат группы включает определение глобального минимума функции, имеющей смысл меры различия между выделенными СКРМ и эталонными сигнатурами, построенными при использовании цифровой карты (ЦК) на основании данных об ориентации МРА.

Вид данной функции при использовании квадрата евклидовой меры представлен формулой:

$$f(x_{1...W}, y_{1...W}) = \sum_{w=1}^W \sum_{n=1}^N \left(\frac{A_w(n) - B(n, x_w, y_w, \Phi_w, \Theta_w, \varphi, \theta, \psi)}{A_w(n)^2 + (B(n, x_w, y_w, \Phi_w, \Theta_w, \varphi, \theta, \psi))^2} \right)^2; \quad (6)$$

$$\tilde{x}_{1...W}, \tilde{y}_{1...W} = \arg \min_{x_{1...W}, y_{1...W}} (f(x_{1...W}, y_{1...W})).$$

где $B(n, x_w, y_w, \Phi_w, \Theta_w, \varphi, \theta, \psi)$ – эталонная сигнатура, построенная на основе ЦК с учетом параметров видеоканера Φ_w, Θ_w и ориентации φ, θ, ψ в предположении о нахождении w -того робота в точке с координатами; $\tilde{x}_{1...W}, \tilde{y}_{1...W}$ – оценки координат роботов группы на выходе алгоритма позиционирования.

При этом оценка минимума меры различия, представленной в формуле 6), эквивалентна определению максимума целевой функции $g(x_{1...W}, y_{1...W})$ имеющей смысл нормированного коэффициента корреляции (НКК) между СКРЗ и ЭС:

$$g(x_{1...W}, y_{1...W}) = \sum_{w=1}^W \sum_{n=1}^N \left(\frac{A_w(n) \cdot B(n, x_w, y_w, \Phi_w, \Theta_w, \varphi, \theta, \psi)}{A_w(n)^2 + (B(n, x_w, y_w, \Phi_w, \Theta_w, \varphi, \theta, \psi))^2} \right) \times (1 - X_w(n)); \quad (7)$$

$$\tilde{x}_{1...W}, \tilde{y}_{1...W} = \arg \max_{x_{1...W}, y_{1...W}} (g(x_{1...W}, y_{1...W})).$$

С учетом связей между величинами $\tilde{x}_{1...W}, \tilde{y}_{1...W}$ и $\tilde{u}_{1...W}, \tilde{v}_{1...W}$ приведенных в формуле (2) можно записать:

$$g(x, y, \gamma) = \sum_{w=1}^W \sum_{n=1}^N \frac{A_w(n) \cdot B(n, x, y, \gamma) \cdot (1 - X_w(n))}{A_w(n)^2 + B(n, x, y, \gamma)^2}, \quad (8)$$

$$(\tilde{x}_1, \tilde{y}_1, \tilde{\gamma}) = \arg \max_{x, y, \gamma} (g(x, y, \gamma));$$

$$B(n, x, \gamma) = B(n, x + \tilde{u}_w \cdot \cos(\gamma) + \tilde{v}_w \cdot \sin(\gamma), y - \tilde{u}_w \cdot \sin(\gamma) + \tilde{v}_w \cdot \cos(\gamma), \Phi_w, \Theta_w, \varphi, \theta, \psi).$$

где $\tilde{x}_1, \tilde{y}_1, \tilde{\gamma}$ – оценки соответствующих величин x_1, y_1, γ , в последующим используемые для вычисления $(\tilde{x}_{1...W}, \tilde{y}_{1...W})$.

На рис. 2 представлены формы нормированной к максимуму целевой функции (ЦФ) для различных значений доли неинформативных направлений S_X , при ис-

пользовании сигнатур, выделенных на одном МРА ($W = 1$, рис. 2 а, б), при использовании сигнатур полученных с четырех МРА ($W = 4$), при неизвестной (рис. 2 в, г) и известной (рис. 2 д, е) ориентации группы γ .

Данные зависимости получены для зоны соответствующей окраине города, представляющей открытую местность с небольшой плотностью застройки (менее 10 %) и высотными зданиями.

Как можно увидеть, целевая функция (ЦФ) является кусочно-гладкой и имеет множество побочных локальных максимумов (ПЛМ), число которого пропорционально площади зоны неопределенности. При увеличении доли неинформативных направлений (рис. 2 а-б) наблюдается рост уровня ПЛМ, из-за чего появляется вероятность ложного позиционирования P_L (принятия одного из ПЛМ за глобальный максимум) при действии погрешностей АВКРМ. В то же время использование сигнатур из нескольких точек приводит к уменьшению количества и уровня ПЛМ (рис. 2 в), что позволяет производить позиционирование в этом случае при больших значениях доли неинформативных направлений S_X (рис. 2 г). Формы ЦФ, приведенные на рис. 2 в, г иллюстрируют, дополнительное улучшение свойств функции $g(x, y, \gamma)$, обеспечиваемое применением данных об ориентации группы (известном значении γ).

На основе исследования ЦФ для различных условий функционирования, предлагается производить определение глобального максимума в два этапа [4]:

– Поиск максимума ЦФ, на котором производится вычисление $g(x, y, \gamma)$ для каждой точки трехмерного множества Ω , включающего значения (x, y, γ) , взятые с шагом соответственно $(\Delta_x, \Delta_y, \Delta_\gamma)$

$$(x, y, \gamma) \in \Omega : x \in \{0, \Delta_x, 2 \cdot \Delta_x, \dots, D_x\}, \quad (9)$$

$$y \in \{0, \Delta_y, 2 \cdot \Delta_y, \dots, D_y\}, \gamma \in \{0, \Delta_\gamma, 2 \cdot \Delta_\gamma, \dots, 2 \cdot \pi\};$$

в результате чего определяется точка (x_p, y_p, γ_p) , принимаемая в качестве начального приближения для следующего этапа. При этом величины $(\Delta_x, \Delta_y, \Delta_\gamma)$ определяются шириной лепестка ЦФ, соответствующего полезному максимуму и необходимостью уменьшения вычислительных затрат (числа вычислений ЦФ).

– Определение глобального максимума функции $g(x, y, \gamma)$, которое производится методом градиентного спуска по параметрам (x, y, γ) в окрестности точки (x_p, y_p, γ_p) , полученной на предыдущем этапе. Таким образом, определяются значения $\tilde{x}_1, \tilde{y}_1, \tilde{\gamma}$, на основе которых по формуле (2) вычисляются оценки координат роботов группы на выходе предложенного алгоритма позиционирования $\tilde{x}_{1...W}, \tilde{y}_{1...W}$.

Таким образом, структурная схема варианта предложенной системы позиционирования группы МРА [7], включающая получение и обработку информации на МРА группы, передачу информации по радиоканалу, оценку относительных и абсолютных плановых координат группы на пункте управления, приведена на рис. 3.

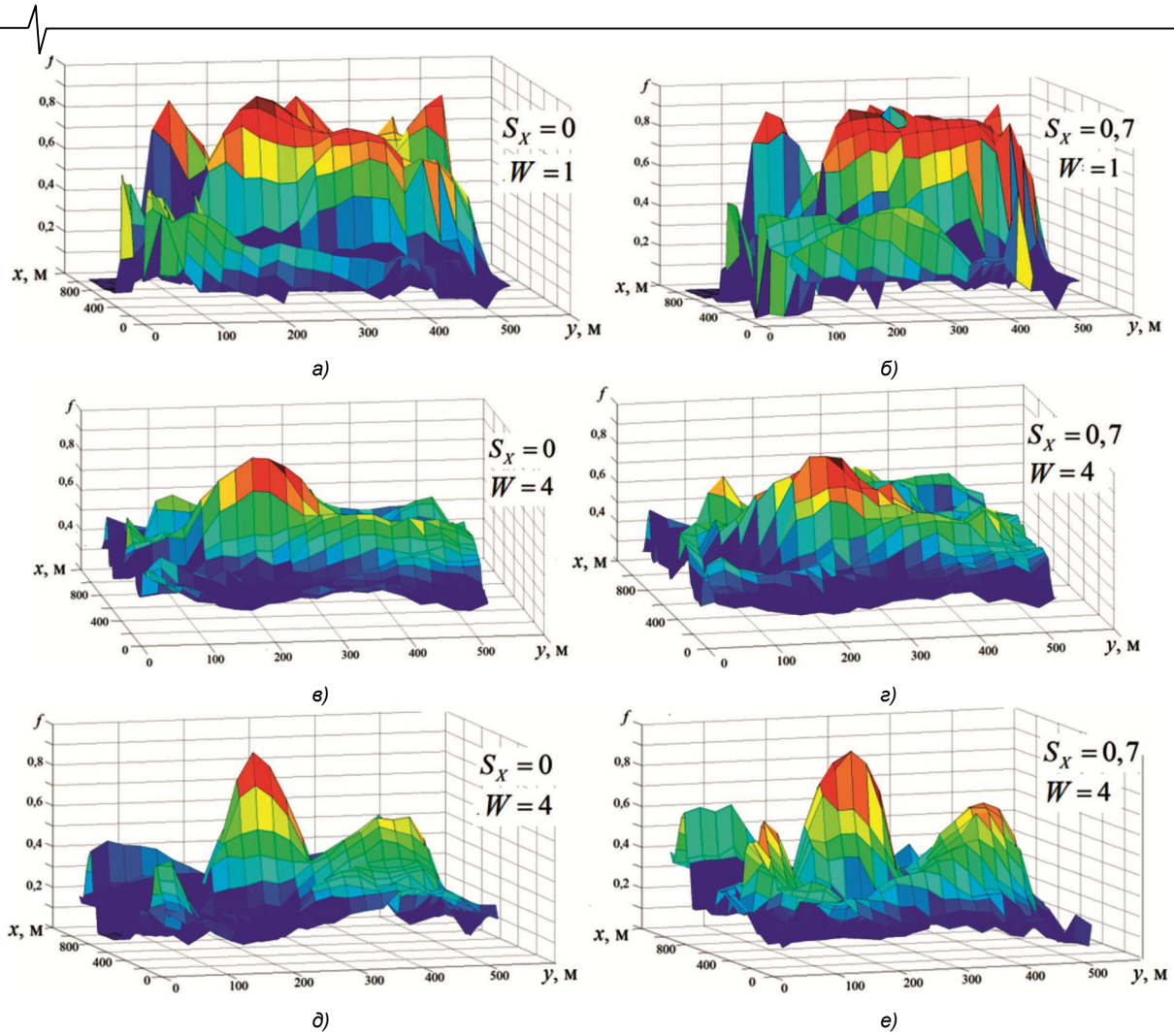


Рис. 2. Формы целевой функции

которая отражает действие погрешности СВКРМ.

Математическая модель искажающей сигнатуры определена эвристически и имеет вид:

$$A_w^H(n) = \Pi \left(\sum_{l=0}^K H_w(n+l) \right), \quad \Pi(a) = \begin{cases} a, & \text{при } a > 0, \\ 0, & \text{при } a \leq 0; \end{cases} \quad (11)$$

где H_w – вектор, содержащий значения случайной величины с нормальным распределением, K – величина окна усреднения, которая выбиралась в зависимости от характера действия МО и составляла $K = (0.01 \dots 0.05) \cdot N$.

В связи со случайным характером зависимости факта появления МО от азимута математическая модель СНН имеет вид:

$$X_n = \text{sign}(A_k^p - S_H); \quad n = 0, N-1; \quad k = \text{mod}(n, \delta); \quad (12)$$

где δ – минимальный азимутальный размер МО (в данном исследовании принято $\delta = 6^\circ$), A_k^p – случайный процесс с равномерным распределением от 0 до 1, S_H – имитируемая доля неинформативных направлений.

Экспериментальное исследование характеристик алгоритма

Для определения эффективности предполагаемого алгоритма исследовались зависимости вероятности ложного позиционирования (ЛП) от количества исполь-

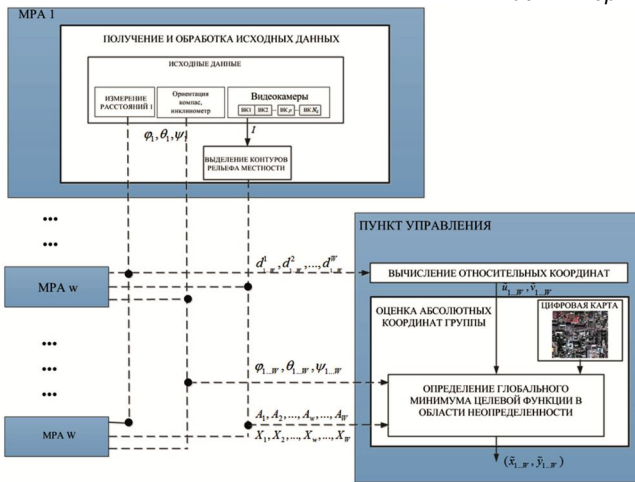


Рис. 3. Структурная схема СВГ группы МРА

Модели сигнатур изображений окружающей обстановки

В качестве модели СКРМ w -того робота $A_w(n)$, полученной в определенной точке x_w, y_w при заданных углах обзора видеокамер и ориентации МРА использовались искаженные ЭС, построенные с учетом указанных аргументов:

$$A_w = B(x_w, y_w, \Phi_w, \Theta_w, \varphi_w, \theta_w, \psi_w) + A_w^H, \quad (10)$$

где A_H – искажающая сигнатура из N элементов,

зуемых роботов в группе при разных значениях доли неинформативных направлений S_X . Эксперимент проводился путем имитационного моделирования при использовании цифровой карты участка городской местности площадью $S = 1,2$ кв. км. При этом исследовался набор из $N_T = 10$ конфигураций группы, включающей до $W = 4$ роботов, расположение которых выбиралось случайно, но таким образом, что расстоянию между любыми двумя удовлетворяло условию $d = 30...70$ м.

Рассматривался алгоритм поиска максимума с величинами шага по координатам и углу соответственно $\Delta_x = 12$ м, $\Delta_\gamma = 10^\circ$. При этом средняя ошибка вычисления СВКРМ принималась равной $D_S = 2,5^\circ$. Вероятность ЛП рассчитывалась как относительная частота появления значения ошибки на выходе поиска максимума ЦФ, превышающего величину шага по координатам, причем величина статистики в каждом случае составляла $N_{CT} = 500$.

На рис. 4 представлены экспериментально полученные зависимости вероятности ЛП P_L от количества используемых роботов в группе при значениях доли неинформативных направлений $S_X = \{0,7; 0,8; 0,9\}$, в случае отсутствия данных об ориентации группы γ и при их наличии (штриховые линии).

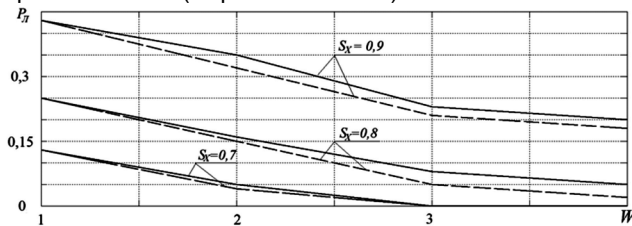


Рис. 4. Зависимости вероятности ложного позиционирования от количества используемых роботов

На основании данных зависимостей можно сделать вывод, что наличие информации об ориентации группы приводит к некоторому уменьшению вероятности ЛП, в то время как увеличение числа роботов, в которых производится получение сигнатур, приводит к значительному уменьшению этого показателя. Также можно

сделать вывод, что использование сигнатур из нескольких точек позволит компенсировать повышение доли неинформативных направлений S_X . Например, рост числа используемых роботов в группе от $W = 1$ до $W = 2$ обеспечивает увеличение допустимого значения S_X на 0,1 при фиксированном уровне вероятности ЛП равном $P_L = 0,13$. Среднеквадратическое значение ошибки вычисления координат роботов группы оставалось в пределах 3 м.

Литература

1. Zhang W., Kosecka J. "Image based localization in urban environments," in Proceedings of the International Symposium on 3D Data Processing, Visualization, and Transmission, 2006.
2. Leung K.Y.K., Clark C.M., Huissoon J.P. localization in urban environment by matching ground level video images with an aerial image, USA IEEE international Conference on Robotics and Automation Pasadena, CA, 2008 – P. 551-556
3. Щербатов И.А., Проталинский И.О., Проталинский О.М. Управление группой роботов: компонентный подход – Информатика и системы управления. 2015. № 1. – С. 93-104.
4. Кириллов С.Н., Бердников В.М., Акопов Э.В. Алгоритм видеопозиционирования объектов в городской местности на основе цифровой карты – Вестник Рязанского государственного радиотехнического университета. 2014. № 47. – С. 3-9.
5. Кириллов С.Н., Бердников В.М., Акопов Э.В. Алгоритм обработки видеоинформации в интересах позиционирования объектов в городской местности, Вестник РГРТУ, 2016, №55, ISSN 1995-4565, – С. 104-110
6. Кириллов С.Н., Бердников В.М., Косткин И.В., Акопов Э.В., Елютин А.В. Комплекс позиционирования роботов на основе систем технического зрения и автономной видео- и радионавигации. Радиотехника – 2016. – № 8. – С. 110-116.
7. Заявка на патент № 2016145192 от 17.11.2016 «Способ определения местоположения группы наземных роботов на основе комплексирования видеоданных и взаимных радиоизмерений дальности» Кириллов С.Н., Бердников В.М., Акопов Э.В.

ДИНАМИЧЕСКИЕ МОДЕЛИ ГЕОМЕТРИЧЕСКОЙ ОБРАБОТКИ ИЗОБРАЖЕНИЙ В СИСТЕМАХ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ

Егошкин Н.А., к.т.н., в.н.с. НИИ «Фотон» Рязанского государственного радиотехнического университета,
e-mail: foton@rsreu.ru.

DYNAMIC MODELS OF GEOMETRIC IMAGE PROCESSING IN EARTH REMOTE SENSING SYSTEMS

Egoshkin N.A.

The problem of improving the quality of geometric image processing models in Earth remote sensing systems is considered. It is shown that it is advisable to use analytical models and image analysis data together with reference information. Dynamic geometric models are proposed. Methods of dynamic systems identification are used for the proposed models construction and analysis. Optimal filters of geometric models correction with reference data usage in the cases of real-time and post-processing are constructed. The approach to correcting periodic distortions in geometry by studying the readings of thermal sensors is investigated. The proposed approaches are tested on the data from the geostationary satellite «Electro-L» No. 2.

Key words: remote sensing, geometric models, geodesic binding, dynamic models, system identification.

Ключевые слова: дистанционное зондирование, геометрические модели, геодезическая привязка, динамические модели, идентификация динамических систем.

Введение

Современные системы дистанционного зондирования Земли (ДЗЗ) содержат достаточно сложную съёмочную аппаратуру, а формируемые ими изображения земли нуждаются в геометрической обработке перед использованием потребителями. Основу геометрической обработки составляют модели процесса съёмки в виде пары функций

$$\mathbf{z} = \mathbf{F}(\boldsymbol{\mu}), \quad t = T(\boldsymbol{\mu}), \quad (1)$$

которые устанавливают соответствие между координатами пикселей $\boldsymbol{\mu} = (m, n)$ изображений на выходе съёмочной аппаратуры, некими эталонными координатами $\mathbf{z} = (x, y)$ на земной поверхности и моментом времени наблюдения t .

Обычно функция $\mathbf{F}$ строится достаточно сложным многоэтапным образом [1]. Формально можно записать:

$$\mathbf{F}(\boldsymbol{\mu}) = \mathbf{G}(\boldsymbol{\Psi}(t), \mathbf{P}(\boldsymbol{\mu})), \quad (2)$$

где $\mathbf{G}$ – оператор фундаментального описания процесса прохождения излучения от поверхности Земли на трассе спутник-земля, внутри съёмочной аппаратуры и его преобразования в регистрируемый сигнал; $\boldsymbol{\Psi}(t)$ и $\mathbf{P}(\boldsymbol{\mu})$ – вектор-функции, характеризующие работу конкретной системы ДЗЗ, причем $\boldsymbol{\Psi}(t)$ является функцией времени, а $\mathbf{P}(\boldsymbol{\mu})$ – координат $\boldsymbol{\mu}$. Как правило, вектор $\boldsymbol{\Psi}(t)$ включает пространственное положение, кватернионы ориентации спутника и установочные кватернионы [1], а также углы поворота сканирующих зеркал [2]. Вектор $\mathbf{P}(\boldsymbol{\mu})$ описывает элементы внутреннего ориентирования съёмочной аппаратуры, характеристики подстилающей поверхности и трассы прохождения излучения

Рассмотрена задача повышения качества моделей геометрической обработки изображений от систем ДЗЗ. Показано, что целесообразно совместное использование аналитических моделей и данных анализа изображений с привлечением опорной информации. Предложены динамические геометрические модели, для построения и анализа которых использованы методы идентификации динамических систем. Построены оптимальные фильтры для коррекции геометрических моделей с привлечением опоры для случаев реального времени и пост-обработки. Исследован подход к коррекции периодических искажений геометрии путем привлечения показаний термодатчиков. Предложенные подходы апробированы на данных от геостационарного спутника «Электрон-Л» №2.

(например, рельеф местности и тропосферные задержки для радиолокационной съёмки).

На процесс съёмки влияет множество факторов, в том числе случайных и трудно предсказуемых, поэтому функции $\boldsymbol{\Psi}(t)$, $\mathbf{P}(\boldsymbol{\mu})$ и $\mathbf{F}$ являются случайными. Обычно под моделью съёмки (1) понимаются описание конкретной реализации.

На борту спутника имеется аппаратура, позволяющая измерять координаты и ориентацию спутника, положение сканирующих зеркал и другие параметры. В результате обработки результатов этих измерений и другой служебной информации со спутника формируются оценки $\hat{\boldsymbol{\Psi}}(t)$ и $\hat{\mathbf{P}}(\boldsymbol{\mu})$, а также оценка функции $t = \hat{T}(\boldsymbol{\mu})$. Подставляя оценки в (2) получаем модель съёмки вида $\hat{\mathbf{z}} = \hat{\mathbf{F}}(\boldsymbol{\mu})$. Нормализация изображений перед выдачей потребителям заключается в трансформации исходного растра $b(m, n)$ в растр $b(\hat{\mathbf{z}})$ по закону $\hat{\mathbf{z}} = \hat{\mathbf{F}}(\boldsymbol{\mu})$.

Качество модели $\hat{\mathbf{F}}(\boldsymbol{\mu})$ и изображения $b(\hat{\mathbf{z}})$ можно охарактеризовать функцией ошибки $\mathbf{s}(\boldsymbol{\mu}) = \hat{\mathbf{F}}(\boldsymbol{\mu}) - \mathbf{F}(\boldsymbol{\mu})$, где $\mathbf{F}(\boldsymbol{\mu})$ – «истинная» модель съёмки. Качество модели зависит от достоверности описания $\mathbf{G}$ и точности оценок $\hat{\boldsymbol{\Psi}}(t)$, $\hat{\mathbf{P}}(\boldsymbol{\mu})$. Если описание $\mathbf{G}$ достаточно точ-

ное, ошибки $s(\mu)$ можно считать следствием погрешностей в оценках «истинных» функций $\Psi(t)$, $P(\mu)$. Это позволяет на основе (2) определять априорное качество моделей через характеристики ошибок оценок $\hat{\Psi}(t)$, $\hat{P}(\mu)$, $\hat{T}(\mu)$. Методы измерений и соответствующая бортовая аппаратура постоянно совершенствуются, ошибки измерений снижаются, и следовательно качество моделей растет. Однако не все важные факторы доступны прямому измерению. Современные знания позволяют неограниченно повышать точность описания G , но при этом возрастает и количество необходимых исходных данных, в том числе не доступных измерению.

Другой путь анализа качества модели – сравнить $b(\hat{z})$ с некоторой опорной информацией, например, картой местности или другими снимкам. Пусть функция $u(\hat{z})$ ставит в соответствие объекту с координатами $\hat{z}$ на изображении $b(\hat{z})$ отклонение u измеренных координат объекта от опорных. Обозначим $u(\mu) = u(\hat{F}(\mu))$. Тогда

$$u(\mu) = s(\mu) + \Delta u(\mu), \quad (3)$$

где $\Delta u(\mu)$ – ошибка, обусловленная измерения координат и погрешностями опоры.

Привлечение измерений $u(\mu)$ вносит принципиально новый момент в построение геометрических моделей: проявляется возможность улучшить качество модели (и соответственно геометрической обработки) без совершенствования бортовой аппаратуры, для этого достаточно по $u(\mu)$ как можно точнее оценить $s(\mu)$. Конечно, возможности такого улучшения не безграничны (из-за ошибок измерений), кроме того получения измерений $u(\mu)$ достаточно трудоемкая операция. Но вообще на интервале жизненного цикла космической системы фактически невозможно добиться высокого геометрического качества, если не привлекать измерения (3).

Таким образом, возникают следующие вопросы:

- а) Как использовать измерения $u(\mu)$ для повышения качества геометрических моделей;
- б) Как оптимальным образом организовать трудоемкое вычисление $u(\mu)$.

Широко распространенный вариант решения а) состоит в калибровке съёмочной аппаратуры, цель которой – путем анализа $u(\mu)$ повысить точность оценок элементов внутреннего ориентирования и установочных кватернионов [3], [4]. При этом предполагаются, что данные параметры не зависят от времени, и соответственно действие шумов измерений $\Delta u(\mu)$ может быть скомпенсировано за счет использования достаточно представительной статистической выборки. Калибровка очень важна, но таким образом нельзя устранить динамические составляющие погрешностей при формировании отдельных изображений и в ходе жизненного цикла системы ДЗЗ.

Другой часто используемый подход – оценивать поправки к конструктивным углам для каждого изображения независимо [5], [6]. Это позволяет учесть случайное изменение ошибок геометрии во времени, однако

при этом точность моделей становится ограниченной шумами измерений $\Delta u(\mu)$. Иногда для снижения шумов применяют сглаживание измерений $u(\mu)$ по времени или пространству, однако параметры сглаживания выбираются интуитивно авторами алгоритмов.

В данной статье рассматривается подход к оптимальному решению поставленных вопросов за счет более адекватного описания динамической природы ошибок $s(\mu)$, $\Delta u(\mu)$ на основе методов идентификации динамических систем [7].

Динамическая геометрическая модель

Рассмотрим ошибку $s(\mu)$ как функцию координат μ . С другой стороны, если $s(\mu)$ связана только с неправильным оцениванием $D(t)$, то более естественно рассмотреть ошибку как функцию времени $s(t)$. Следовательно, первым делом в соответствии с априорной физической причиной возникновения ошибок следует определиться с выбором некой независимой переменной. Обозначим t как векторную переменную. Её связь с координатами μ однозначна, обозначим её как $t = t(\mu)$. Тогда вместо $s(\mu)$ и $u(\mu)$ можно рассматривать соответствующие $s(t)$ и $u(t)$. Заметим, что если в вектор t входит время, можно единообразным образом рассматривать изображения, формируемые в ходе жизненного цикла космической системы.

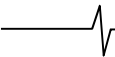
Компоненты вектора $z = (x, y)$ можно считать независимыми, так как механизмы развертки датчика в направлении осей x и y обычно совершенно различны (например, в одном направлении развертка за счет полета спутника, в другом – за счет опроса чувствительных элементов линейки фотоприёмников). Поэтому для простоты далее будем рассматривать независимо отдельные компоненты вектора z , обозначая соответствующие компоненты векторов скалярными символами.

Нахождение $u(\mu)$ удобно производить следующим образом: плоскость μ разбивается регулярным образом на фрагменты одинакового размера и для центра каждого находится значение u [1]. В координатах t этот процесс соответствует дискретному изменению t , следовательно, можно определить $t(\mu)$ так, чтобы $u(t)$ оказалась заданной в узлах целочисленной решетки. Далее считать компоненты t целочисленными.

Пусть необходимые процедуры калибровки съёмочной аппаратуры проведены и отклонения $s(\mu)$ достаточно малы. Как правило, значения $s(\mu)$ сильно коррелированы в пространстве или во времени, т.е. значения $s(t)$ сильно коррелированы для близких значений t . Считая, что переходные процессы после запуска спутника закончились, запишем $s(t)$ в виде стационарного случайного процесса:

$$s(t) = H(t) \otimes e(t), \quad (4)$$

где $\otimes$ – оператор свертки, $e(t)$ – последовательность независимых случайных величин с нулевым математическим ожиданием и одинаковой дисперсией $D[e]$,



$H(t)$ – импульсная характеристика формирующего фильтра. Несмотря на нулевое математическое ожидание $e(t)$, формула (4) позволяет адекватно описывать систематические ошибки модели (1), так как срок активного существования спутника конечен.

Процесс измерений (3) также можно записать как

$$u(t) = s(t) + I(t) \otimes \varepsilon(t), \quad (5)$$

где $\varepsilon(t)$ – последовательность независимых случайных величин с нулевым математическим ожиданием и одинаковой дисперсией $D[\varepsilon]$, $I(t)$ – импульсная характеристика формирующего фильтра.

Если $H(t)$, $I(t)$, $D[e]$, $D[\varepsilon]$ известны, то наилучшая оценка в плане минимума среднеквадратической ошибки обеспечивается фильтром Винера:

$$\hat{s}(t) = W(t) \otimes u(t),$$

$$W(\omega) = \frac{|H(\omega)|^2 D[e]}{|H(\omega)|^2 D[e] + |I(\omega)|^2 D[\varepsilon]}, \quad (6)$$

где $W(t)$ – импульсная характеристика фильтра Винера, $W(\omega)$, $H(\omega)$, $I(\omega)$ – спектры Фурье одноименных функций аргумента t .

Для использования (6) должны быть известны $H(t)$, $I(t)$, $D[e]$, $D[\varepsilon]$, а кроме того фильтр $W(t)$ непричинный, что может затруднить обработку данных в реальном времени. Рассмотрим далее задачу оценивания исходных для (6) данных и проблему обработки в реальном времени.

Предсказатели и идентификация передаточных характеристик

Потфрагментное вычисление $u(t)$ предполагает наличие определенного порядка обработки фрагментов. Вообще выбор порядка достаточно произволен. Однако если данные поступают в реальном времени, свобода в выборе порядка снижается, так как целесообразно обрабатывать информацию по мере её поступления. Так или иначе, значения t оказываются неким образом упорядоченными.

Для определенности зададим на t словарное упорядочивание: $\mathbf{t} = (t_1, t_2, \dots) > \mathbf{r} = (r_1, r_2, \dots) \Leftrightarrow t_i > r_i$, $t_{i+1} = r_{i+1}$, $t_{i+2} = r_{i+2}$, ... Это соответствует последовательной обработке фрагментов вдоль одной из координат, например, вдоль строки. После полной обработки строки осуществляют переход к следующей строке и так далее. Такое упорядочивание также легко согласуется с обработкой в реальном времени, если вектор $\mathbf{t} = (t_1, t_2, \dots, t)$ содержит время t в последней компоненте.

Обычно для вычисления $u(t)$ для каждого t по порядку полезно иметь некоторое максимально точное начальное приближение $v(t)$. Поставим задачу рассчитать его только по «предыдущим» значениям $u(t')$, $t' < t$. Простейший вариант организации такого прогнозирования – использовать линейный пространственно-

инвариантный причинный фильтр:

$$v(t) = \sum_{d>0} P(d) u(t-d), \quad (7)$$

где $P(d)$ – импульсная характеристика фильтра, $\mathbf{0} = (0, 0, \dots, 0)$.

Будем считать импульсный отклик $P(d)$ конечным. Тогда по методу наименьших квадратов легко найти оценку отклика из условия $\sum_{t \in \Theta} |u(t) - v(t)|^2 = \min$, где суммирование осуществляется в пределах области Θ , в которой уже известны значения $u(t)$, а $v(t)$ задана в соответствии с (7). Это процесс идентификации динамической системы.

Как только характеристика фильтра $P(d)$ оценена, можно эффективным образом использовать прогноз (7) для максимального эффективного определения $u(t)$.

На рис. 1 показаны измеренные отклонения координат $u(t)$ для геостационарного спутника «Электро-Л» № 2 (инфракрасный модуль, пространственное разрешение 4 км): сверху отклонение с запада на восток, снизу отклонение с юга на север в км. В качестве опоры использовались контурные точки диска Земли [8]. По оси абсцисс – переменная t , в качестве которой выбран номер сеанса съёмки, которые следуют периодически каждые 30 минут.

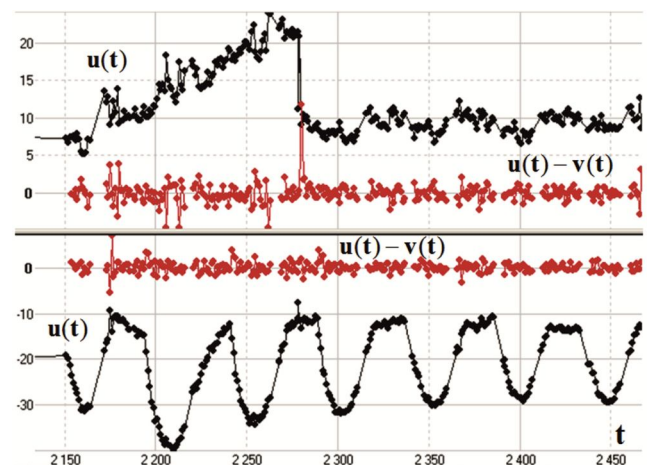


Рис. 1. Измеренные отклонение координат для спутника «Электро-Л» №2 и ошибка предсказания

Можно видеть, что отклонения координат $u(t)$ значительны (десятки километров) и сильно коррелированы по времени, а в направлении юг-север имеется периодичность, очевидно связанная с периодом обращения спутника. Также на рис. 1 показаны графики ошибки предсказания $u(t) - v(t)$. Их величина значительно меньше, что позволяет сократить время геодезической привязки и повысить её точность.

Согласно [7], из (4) и (5) следует, что ошибка предсказания

$$\Delta u(t) = u(t) - v(t) = e(t) + \varepsilon(t). \quad (8)$$

Пусть $I(t) = \delta(t)$, где $\delta(t)$ – дельта-функция, т.е. ошибки измерений $u(t)$ не коррелированы. Это выпол-

няется, если например фрагменты при корреляционно-экстремальном отождествлении не перекрываются. Тогда $u(t) = s(t) + \varepsilon(t)$. Можно показать, что $v(t) = s(t) - e(t)$. Поэтому оптимальная в плане минимума дисперсии оценка

$$\hat{s}(t) = \alpha v(t) + (1 - \alpha) u(t), \quad \alpha = D[\varepsilon] / (D[e] + D[\varepsilon]). \quad (9)$$

Формула (9) позволяет осуществлять уточнение геометрической модели в реальном времени. Из (8) и независимости $e(t)$ и $\varepsilon(t)$ следует $D[e] + D[\varepsilon] = D[\Delta u]$. Поэтому необходимую для использования (9) оценку $D[e] + D[\varepsilon]$ легко получить анализом невязок $\Delta u(t)$. Что касается $D[\varepsilon]$, то её нельзя выделить из исходных данных: с использованием только опорной информации принципиально невозможно различить ошибки геометрической модели и ошибки опоры. На практике можно принять $D[\varepsilon]$ равной ошибке корреляционного отождествления (около 0.3 пикселя без использования методов субпиксельного уточнения [1]).

Из (4) и (5) следует, что $u(t) = H(t) \otimes e(t) + I(t) \otimes \varepsilon(t)$. Можно показать, что существует фильтр с импульсной характеристикой $L(t)$, такой что $u(t) = L(t) \otimes \xi(t)$, где $\xi(t)$ – последовательность случайных величин с нулевым математическим ожиданием и известной дисперсией $D[\xi] = D[e] + D[\varepsilon]$.

Согласно [7], оптимальный предсказатель имеет вид $P = 1 - L^{-1}$ и, соответственно, $L = (1 - P)^{-1}$, где операции обращения определены в частотной области, т.е. например $L^{-1}(\omega) = (L(\omega))^{-1}$. Легко показать, что частотная характеристика удовлетворяет $|L(\omega)|^2 D[\xi] = |H(\omega)|^2 D[e] + |I(\omega)|^2 D[\varepsilon]$. Это выражение – знаменатель в (6). Для нахождения числителя (6) достаточно положить $|I(\omega)|^2 = 1$ и задать некоторую априорную величину $D[\varepsilon]$. В результате не причинный фильтр Винера (6) оказывается полностью определенным. Его целесообразно использовать для обработки полностью доступных изображений, например, от низкоорбитальных спутников (данные от которых, как правило, становятся доступными с большой задержкой из-за необходимости ожидания пролета над центром приёма).

Заметим, что спектр $|L(\omega)|^2$ можно было бы оценить непосредственно по данным $u(t)$ стандартными методами спектрального оценивания [7] без построения предсказателя P . Однако, во-первых, предсказатель необходим для более эффективного нахождения $u(t)$. Во-вторых, представление через предсказатель оказывается более экономным в ряде важных для практики случаев, например, когда в спектре $L(\omega)$ присутствует ряд резонансных частот (соответствующих периодами вращения сканирующих зеркал, спутника по орбите и прочее). В-третьих, согласно [7], предсказатель является даже более фундаментальным описанием динамической системы, чем модели (4), (5).

На практике для отдельных t измерения $u(t)$ могут отсутствовать, например, если сюжет оказался не информативен (облачность, недостаточная освещенность солнцем) или съёмка не велась. В этом случае для простоты прогнозирования можно получить отсутствующие измерения интерполяцией. Другой вариант – по аналогии с (7) построить семейство предсказателей, которые не используют отсутствующие данные, и применять в каждом случае наиболее подходящий предсказатель. При фильтрации в реальном времени по (9) при отсутствии измерения $u(t)$ принимается $\alpha = 1$, т.е. $\hat{s}(t) = v(t)$.

Пример на рис. 1 показывает, что если в спектре H имеется выраженная резонансная частота, то может быть полезным разбить время t на две компоненты: номер периода и время в пределах периода. Соответственно вектор t становится двухмерным. Тогда прогнозирование может оказаться более эффективным, если двухмерная импульсная характеристика P окажется более компактной, чем у одномерного предсказателя.

Наличие периодичности может свидетельствовать о влиянии на геометрию дополнительного фактора, не учтенного в G . Например, этим фактором может быть изменение теплового режима работы съемочной аппаратуры, зависящего от взаимного положения спутника, Земли и Солнца. Если у аппаратуры имеются термодатчики (ТД), то их показания целесообразно использовать для прогнозирования. Для этого достаточно на вход фильтра P подать показания температур наряду с «предшествующими» значениям $u(t)$.

На рис. 2 показан пример использования термодатчиков для коррекции суточных колебаний геометрической модели спутника «Электро-Л» №2. На двух графиках сверху показаны компоненты $u(t)$ и $v(t)$. Нижний график – показания термодатчиков (всего их более 16). Для предсказания использовались температуры одного из термодатчиков.

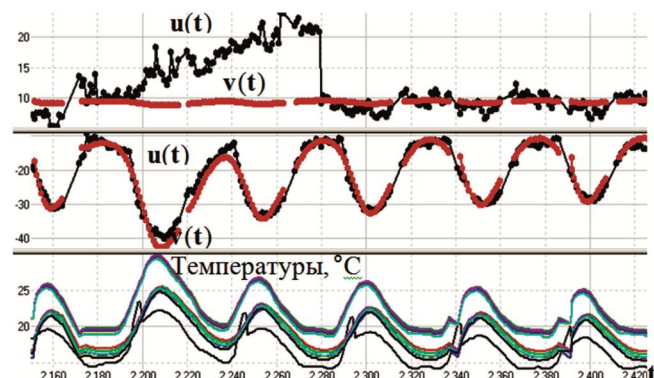


Рис. 2. Измеренные отклонение координат для спутника «Электро-Л» №2, температуры и предсказание с использованием одного термодатчика

Можно видеть, что использование показаний термодатчиков позволяет в значительной степени предсказывать отклонения геометрии, однако не полностью. В табл. 1 показаны среднеквадратического отклонения (СКО) предсказания для различных вариантов построения фильтра P . Оценка СКО выполнялась на выборке более чем из 1500 изображений.

Таблица 1. Сравнение эффективности различных предсказателей

Число входов предсказателя	СКО предсказания, км	
	запад-восток	юг-север
–	2.4	8.2
1	0.8	1.2
2	0.8	1.0
3	0.8	0.9
1 на ТД	1.5	1.7
16 на ТД	1.5	1.1
1 + 16 на ТД	0.8	0.7

Анализа табл. 1 показывает, что в данном случае даже самый простой предсказатель оказывается достаточно эффективным. Использование показаний термодатчиков также полезно. Но лучше всего использовать комбинированный предсказатель – как на основе предыдущих значений, так и учитывающих температуры.

Заключение

По результатам выполненных исследований можно сделать следующие выводы.

1. Для повышения качества геометрических моделей целесообразно опираться на анализ изображений с привлечением опорной информации. При этом необходимо учитывать динамическую природу ошибок и применять методы идентификации динамических систем для их оптимальной компенсации.

2. Целесообразно ввести упорядочивание исходных данных по времени или в пространстве с учетом возможных физических причин появления динамики. Это позволяет использовать предсказатель как наиболее общее описание динамической системы и позволяет при необходимости вести обработку данных в реальном времени.

3. Построены оптимальные фильтры коррекции геометрических моделей для реального и не реального времени, основанные на идентификации передаточной функции предсказателя.

4. Показано, что периодическое изменение температурного режима спутника целесообразно учитывать в гео-

метрических моделях, используя показания термодатчиков в качестве исходных данных для предсказателя.

5. Предложенные подходы апробированы на данных от геостационарного спутника «Электро-Л» № 2. В результате скорость анализа инфракрасных изображений увеличилась в 8 раз, а точность геодезической привязки – в 3 раза.

Литература

1. Современные технологии обработки данных дистанционного зондирования Земли / Под. ред. В.В. Еремеева. М.: Физматлит. 2015. – 460 с.

2. Егошкин Н.А., Еремеев В.В., Москвитин А.Э. Повышение точности коррекции изображений на основе фильтрации измерений углового положения сканирующего зеркала глобального наблюдения земли // Вестник РГРТУ. 2010. №3. – С. 7-11.

3. Gektin U.M., Egoshkin N.A., Eremeev V.V., Kuznetsov A.E., Kochergin A.M. Prelaunch photogrammetric calibration of russian satellite Elektro-I imagery instruments // «The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences». Prague, Czech Republic, 2016. vol. XLI-B6.

4. Ахметов Р.Н., Еремеев В.В., Кузнецов А.Е., Мятлов Г.Н., Пошехонов В.И., Стратилатов Н.Р. Высокоточная геодезическая привязка изображений земной поверхности от КА «Ресурс-П» // Исследование земли из космоса. 2017. №1. – С. 44-53.

5. Воронин А.А., Егошкин Н.А., Еремеев В.В., Москвитин И.В. Геометрическая обработка данных космических систем глобального наблюдения земли // Вестник РГРТУ. 2009. №1.

6. Гектин Ю.М., Егошкин Н.А., Еремеев В.В., Кузнецов А.Е., Москвитин И.В., Смелянский М.Б. Базовое геометрическое обеспечение систем наблюдения земли с геостационарных и высокоэллиптических орбит // Исследование Земли из космоса. 2016. №6. – С. 68-81.

7. Льюинг Л. Идентификация систем. Теория для пользователя. М.: Наука, 1999. – 432 с.

8. Еремеев В.В., Егошкин Н.А., Москвитин А.Э. Координатная привязка изображений от геостационарных спутников по контурным точкам диска Земли // Вестник РГРТУ. 2007. № 22. – С. 10-17.

УДК 004.932

КОМПЛЕКСИРОВАНИЕ ЦИФРОВЫХ МОДЕЛЕЙ РЕЛЬЕФА С ЦЕЛЬЮ ПОВЫШЕНИЯ ТОЧНОСТИ ОПОРНОЙ ИНФОРМАЦИИ О ВЫСОТЕ ОБЪЕКТОВ ЗЕМНОЙ ПОВЕРХНОСТИ

Егошкин Н.А., к.т.н., в.н.с. НИИ «Фотон» Рязанского государственного радиотехнического университета (РГРТУ), e-mail: foton@rsreu.ru;

Ушенкин В.А., м.н.с. НИИ «Фотон» РГРТУ, e-mail: foton@rsreu.ru.

DEM FUSION TO IMPROVE THE ACCURACY OF REFERENCE EARTH SURFACE ELEVATION DATA

Egoshkin N.A., Ushenkin V.A.

Artifacts in existing global DEMs are analyzed in this paper. The problem of reference elevation data accuracy improving by collation and combination of existing DEMs, obtained from different sources and containing different errors, is considered. The algorithm of DEM fusion designed to improve the accuracy of elevation data is proposed. The examples of the algorithm results are given.

Keywords: fusion, DEM, interferometric processing, stereophotogrammetric processing, reference data.

Ключевые слова: комплексирование, цифровая модель рельефа, интерферометрическая обработка, стереофотограмметрическая обработка, опорная информация.

Введение

Опорная информация о высоте рельефа земной поверхности требуется при решении широкого класса задач, связанных с обработкой данных дистанционного зондирования Земли. Априорное знание рельефа требуется при геодезической привязке и геометрической коррекции спутниковых снимков, а также при построении цифровых моделей рельефа и карт смещений земной поверхности путем интерферометрической обработки радиолокационных изображений [1].

В настоящее время в качестве опорной информации о высоте используется ряд глобальных низкодетальных цифровых моделей рельефа (ЦМР), содержащих только высоты естественных неживых элементов ландшафта, и цифровых моделей местности (ЦММ), включающих также искусственные строения и растительность. В дальнейшем для простоты изложения все глобальные низкодетальные трехмерные модели будем называть ЦМР, подразумевая, что часть из них на самом деле является ЦММ.

Все глобальные ЦМР получены одним из двух способов: путем стереофотограмметрической [2] или интерферометрической [1] обработки материалов оптической или радиолокационной космической съемки Земли. Глобальный характер указанных ЦМР заключается в том, что они покрывают большую часть земной суши. При этом для каждой из ЦМР имеется ряд регионов, в которых нет данных о высоте. На стереофотограмметрических ЦМР пропуски вызваны наличием облачности и засветкой льдов, снега и песков при оптической съемке. На интерферометрических ЦМР пропуски вызваны низкой когерентностью сигнала радиолокационных изображений на льдах, в песках и лесах. Кроме того, накло-

Анализируются на наличие артефактов существующие глобальные низкодетальные цифровые модели рельефа земной поверхности. Рассматривается задача повышения точности опорной информации о высоте объектов земной поверхности за счет сопоставления и объединения существующих цифровых моделей рельефа, полученных из различных источников и содержащих разные ошибки. Предлагается алгоритм комплексирования цифровых моделей рельефа, направленный на повышение точности данных о высоте. Приводятся примеры результатов предложенного алгоритма.

нение орбиты космических аппаратов не позволяет получить данные для построения ЦМР в полярных и приполярных областях.

Перечисленные причины возникновения пропусков могут также приводить к ошибкам высоты, если они не были выявлены при формировании ЦМР. Дополнительными источниками ошибок являются:

1) невыявленные участки водной поверхности и прочие малоконтрастные объекты, на которых могут возникнуть значительные ошибки корреляционного совмещения при стереофотограмметрической обработке;

2) неправильное развертывание фазы при интерферометрической обработке [3].

Ошибки высоты и пропуски на глобальных ЦМР располагаются на локальных участках. Локальный характер ошибок обусловлен тем, что глобальная ЦМР формируется в виде мозаики, состоящей из ЦМР, построенных по одной паре маршрутов. Поэтому даже на интерферометрических ЦМР, в которых возможно значительное распространение ошибки из-за неправильного развертывания фазы, это распространение ограничивается краями маршрута. Благодаря тому, что все ЦМР получены из разных источников двумя принципиально отличающимися способами (интерферометрией и стереофотограмметрией), расположение ошибок и пропусков на них в целом не совпадает. Следовательно, появляется возможность повышения точности данных о высоте объектов участка земной поверхности путем комплексирования нескольких ЦМР.

Цель настоящей работы – построение алгоритма комплексирования ЦМР, который для каждой точки земной поверхности сопоставляет имеющиеся различные данные о высоте и выбирает из них наиболее достоверные, тем самым повышая среднюю точность результата комплексирования по сравнению с каждой из исходных ЦМР.

Анализ существующих глобальных ЦМР

Среди всех существующих глобальных ЦМР можно выделить три, отличающиеся наилучшей детальностью и точностью высоты: Alos World 3D 30, SRTM1 версии 3 и ASTER GDEM версии 2. Указанные ЦМР имеют проекцию пикселя в одну угловую секунду, что соответствует примерно 30 м на экваторе. Пространственное разрешение указанных ЦМР лежит в диапазоне от 30 до 90 м, а среднеквадратичная ошибка высоты – от 5 до 20 м.

Alos World 3D 30 представляет собой ЦММ, полученную путем стереофотограмметрической обработки троек оптических изображений с космического аппарата Alos, снятых при визировании вперед, в надири и назад [4]. За счет такой организации съемки на ЦММ не образуются пропусков, вызванных затенением. Исходная версия Alos World 3D имеет пространственное разрешение 5 м и распространяется Японским космическим агентством на коммерческой основе, а в открытом доступе начиная с 2016 г. представлена версия с пространственным разрешением, закругленным до 30 м. Среднеквадратичная погрешность высоты на Alos World 3D 30 составляет 5 м. Очень редко на ЦММ встречаются грубые ошибки высоты на невыявленной облачности или водной поверхности, а также ошибочное выявление водной поверхности. ЦММ покрывает всю сушу от 82° ю.ш. до 82° с.ш., за исключением регионов с постоянной облачностью и ледниками.

ЦМР SRTM1 получена путем интерферометрии с жесткой базой в результате топографической миссии шаттла «Индевор» в 2000 г. [5]. Пространственное разрешение ЦМР колеблется в различных регионах от 40 до 60 м. При этом из-за особенностей геометрии интерферометрической радиолокационной съемки может происходить значительное снижение детальности на горных склонах. Среднеквадратичная погрешность высоты на SRTM1 составляет 10 м, причем большая часть этой погрешности вызвана шумами, не полностью подавленными при фильтрации интерферограммы. В первых версиях ЦМР наблюдалось значительное количество ошибок из-за неправильного развертывания фазы в горной местности, однако в третьей версии практически все они скорректированы по данным из альтернативных источников. ЦМР покрывает земную поверхность от 56° ю.ш. до 60° с.ш.

ЦММ ASTER GDEM получена путем стереофотограмметрической обработки пар оптических изображений, снятых в ближнем инфракрасном диапазоне космическим аппаратом Terra [6]. Пространственное разрешение исходных изображений составляло всего 20 м, из-за чего на ЦММ наблюдается сильный шум, вызванный субпиксельной погрешностью корреляционного совмещения исходных стереопар изображений. При этом среднеквадратичная погрешность высоты составляет 20 м. Кроме того, при формировании ЦММ была некачественно построена маска облачности и водной поверхности, что привело к локаль-

ным ошибкам высоты величиной до нескольких сотен метров, особенно часто встречающимся в приполярных и полярных областях. Пространственное разрешение ASTER GDEM версии 2 составляет порядка 80 м. ЦММ покрывает земную поверхность от 83° ю.ш. до 83° с.ш., за исключением небольших участков с облачностью и ледниками.

Комплексирование цифровых модулей

Ключевая идея комплексирования

Поскольку все существующие глобальные ЦМР характеризуются разной точностью высоты и пространственным разрешением, среди них можно выбрать наиболее точную и детальную в качестве базовой. Результат комплексирования совпадает с базовой ЦМР во всех участках, кроме тех, где выявлены локальные ошибки путем сравнения с другими ЦМР. Участки с локальными ошибками следует заполнять данными из других источников. При этом необходимо обеспечить отсутствие нестыковок на границе этих участков, вызванных несогласованностью данных на разных ЦМР.

Можно выделить два основных типа локальных ошибок:

- 1) ошибочное выделение водной поверхности на склоне рельефа с назначением фиксированной высоты;
- 2) ошибки большой величины из-за невыявления отдельного облака, участка ненадежной корреляции или неправильного развертывания фазы.

Ошибки первого типа могут быть выявлены на основе следующего правила: водная поверхность не может быть выше, чем берег. Области с ошибками второго типа, как правило, имеют ярко выраженные границы, на которых наблюдаются резкие изменения высоты величиной от нескольких десятков метров до нескольких километров. Это является основным признаком для их выявления. Однако при этом следует также учесть, что аналогичным признаком обладают высотные здания, которые присутствуют на ЦММ, но отсутствуют на ЦМР.

Критерии выявления ошибок на базовой ЦМР

Критерий выявления ошибочно выделенной водной поверхности можно записать следующим образом:

$$N_{\sigma\sigma} > \varepsilon_{\sigma\sigma} N_{\sigma}, \quad (1)$$

где N_{σ} – количество пикселей ЦМР, расположенных вдоль границы (берега) выделенной водной поверхности, $N_{\sigma\sigma}$ – количество пикселей ЦМР, расположенных вдоль берега, в которых высота воды более чем на σ_h больше высоты берега, где σ_h – среднеквадратичная ошибка высоты на ЦМР, $\varepsilon_{\sigma\sigma}$ – порог, значительно меньший единицы.

Области с ошибками, связанными с облачностью, ненадежной корреляцией или неправильным развертыванием фазы, выявляются в два этапа.

На первом этапе находятся пиксели базовой ЦМР, значение высоты в которых значительно отличается от аналогичных значений на альтернативной ЦМР. При этом допустимая степень отличия выбирается в зависимости от типа рельефа. Для равнинной местности значения высоты должны быть максимально близки, а в горной местности следует допустить сильные расхож-

дения, которые могут быть связаны как с разной детальностью сравниваемых ЦМР, так и с погрешностями их геодезической привязки:

$$h_{\text{баз}mn} - h_{mn} \notin [-\Delta h_{\text{ниж}}, \Delta h_{\text{верх}}] \cup [-\varepsilon_{Gh} G_{h_{\text{max}mn}}, \varepsilon_{Gh} G_{h_{\text{max}mn}}], \quad (2)$$

где $h_{\text{баз}}$ – значение высоты на базовой ЦМР, h – значение высоты на альтернативной ЦМР, $\Delta h_{\text{ниж}}$ и $\Delta h_{\text{верх}}$ – пороги допустимых отличий высоты для равнинной местности, $G_{h_{\text{max}mn}}$ – максимум модуля градиента высоты в малой окрестности mn -го пикселя альтернативной ЦМР, ε_{Gh} – коэффициент учета градиента высоты.

В результате применения критерия (2) образуется некоторое количество кандидатов в области с ошибками высоты, часть из которых отбраковывается на втором этапе выявления. Области с ошибками высоты признаются только те из кандидатов, которые удовлетворяют следующему условию:

$$\frac{1}{N_{zp}} \sum_{i=1}^{N_{zp}} |G_{h_{zp} \text{баз}i}| - |G_{h_{zp}i}| > \varepsilon_{zp}, \quad (3)$$

где $G_{h_{zp} \text{баз}i}$ и $G_{h_{zp}i}$ – градиент высоты в i -м пикселе границы области-кандидата, направленный по внешней нормали к линии границы, на базовой и альтернативной ЦМР, N_{zp} – периметр области-кандидата в пикселях, ε_{zp} – неотрицательный порог, подбираемый эмпирическим путем.

Если кандидат не удовлетворяет условию (3), то предполагается, что в данной области ошибка содержится не на базовой ЦМР, а на альтернативной ЦМР.

Заполнение пропусков и областей с ошибками на базовой ЦМР

Пропуски и выделенные ошибочные области на базовой ЦМР требуют заполнения по альтернативной методике с обеспечением стыковки значений высоты на границах этих областей. Стыковку можно обеспечить с использованием корректирующей высотной поверхности $h_{\text{корр}}$, применяемой к значениям высоты на альтернативной ЦМР. В пикселях с корректными значениями высоты на базовой ЦМР корректирующая поверхность определяется как $h_{\text{корр}mn} = h_{\text{баз}mn} - h_{mn}$.

В выделенных ошибочных областях и пропусках на базовой ЦМР значения корректирующей поверхности определяются с помощью интерполяции по методу обратных взвешенных расстояний [7]:

$$h_{\text{корр}mn} = \frac{\sum_k \sum_l h_{\text{корр}m+k, n+l} \omega_{\text{корр}m+k, n+l} \left(\sqrt{k^2 + l^2} \right)^{-p}}{\sum_k \sum_l \omega_{\text{корр}m+k, n+l} \left(\sqrt{k^2 + l^2} \right)^{-p} + \varepsilon}, \quad (4)$$

где $0 < \sqrt{k^2 + l^2} \leq D$; k, l – целые; $\omega_{\text{корр}mn} = 1$ в пикселях с корректными значениями высоты на базовой ЦМР и $\omega_{\text{корр}mn} = 0$ в ошибочных областях и пропусках на

базовой ЦМР; D и p – радиус и показатель фильтра обратных расстояний, ε – малая величина, используемая для регуляризации.

В результате пропуска и ошибочные области на базовой ЦМР заполняются значениями $h_{mn} + h_{\text{корр}mn}$.

Алгоритм комплексирования нескольких ЦМР

При описании критериев выявления ошибок и заполнения ошибочных областей рассматривался простейший частный случай, когда комплексировались только две ЦМР. В общем случае комплексирование осуществляется следующим образом.

Комплексируемые ЦМР упорядочиваются по интегральному критерию точности и детальности. Сначала в качестве альтернативной выбирается худшая ЦМР, а в качестве базовой – предпоследняя по качеству, и осуществляется их комплексирование. Затем комплексировются ЦМР, третья от конца по качеству, выбранная как базовая, и результат объединения двух худших ЦМР. Последовательное комплексирование осуществляется до тех пор, пока в качестве базовой не будет выбрана лучшая ЦМР.

Экспериментальные исследования

С учетом имеющихся в настоящее время глобальных низкодетальных ЦМР и их точности и детальности, на практике может быть реализовано комплексирование следующих из них: ALOS World 3D 30, SRTM1 версии 3, ASTER GDEM версии 2 и GMTED2010. ЦМР упорядочены в порядке убывания качества. GMTED2010 имеет низкое пространственное разрешение, варьирующееся в различных регионах от 250 м до нескольких километров, но покрывает абсолютно всю сушу на Земле, поэтому заполнение пропусков на ней не требуется.

Ввиду отсутствия точных сведений о высоте рельефа качество результатов комплексирования ЦМР оценивалось на основе визуального выявления ошибок путем сопоставления ЦМР и спутниковых снимков на тот же регион. Оценка производилась не по всему земному шару, а по небольшой выборке отдельных регионов. Примеры таких регионов представлены на рис. 1–3.

Результаты оценки показали, что в пределах покрытия ЦМР SRTM все крупные ошибки на ALOS World 3D устраняются при комплексировании, а остаются лишь небольшие ошибки, которые трудно отличить от сюжета или временных изменений рельефа. Количество ошибочных пикселей на результате комплексирования в среднем оказывается примерно на три десятичных порядка меньше, чем на базовой ЦМР. Среднее значение абсолютной погрешности высоты в ошибочных пикселях уменьшается после комплексирования примерно в 12 раз.

В полярных и приполярных регионах точность результата комплексирования снижается. Это связано с тем, что ALOS World 3D приходится сравнивать с более грубыми ЦМР ASTER GDEM и GMTED2010. Следовательно, возрастает вероятность содержания ошибки одновременно на всех сравниваемых ЦМР, что иллюстрирует рис. 3.

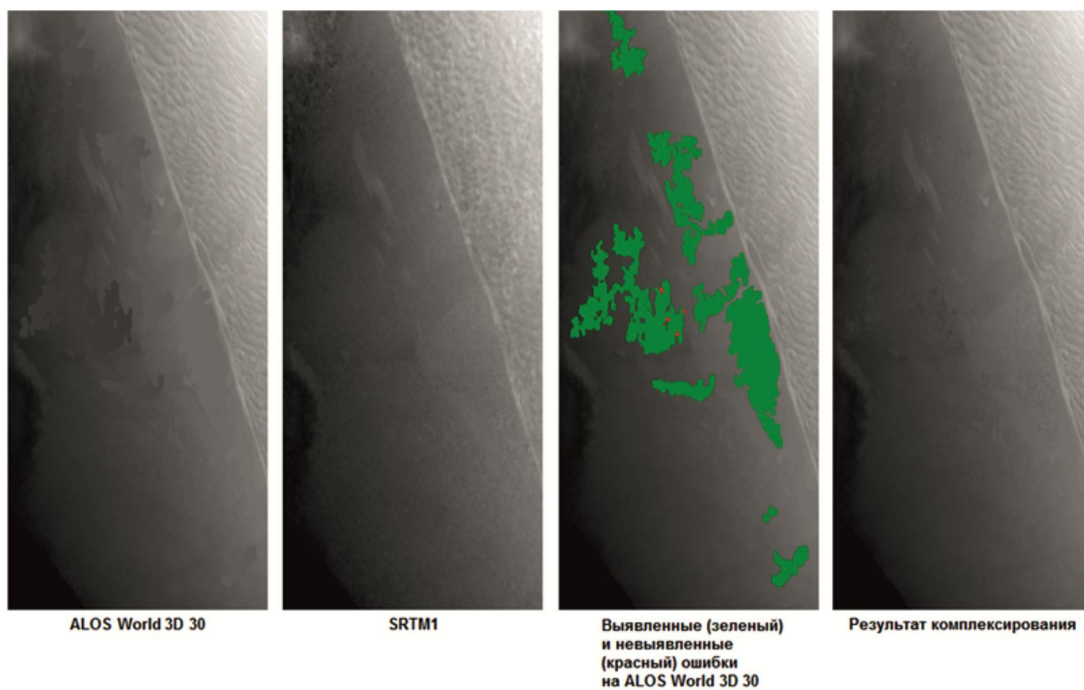


Рис. 1. Пример выявления и устранения при комплексировании ошибочно примененной маски водной поверхности

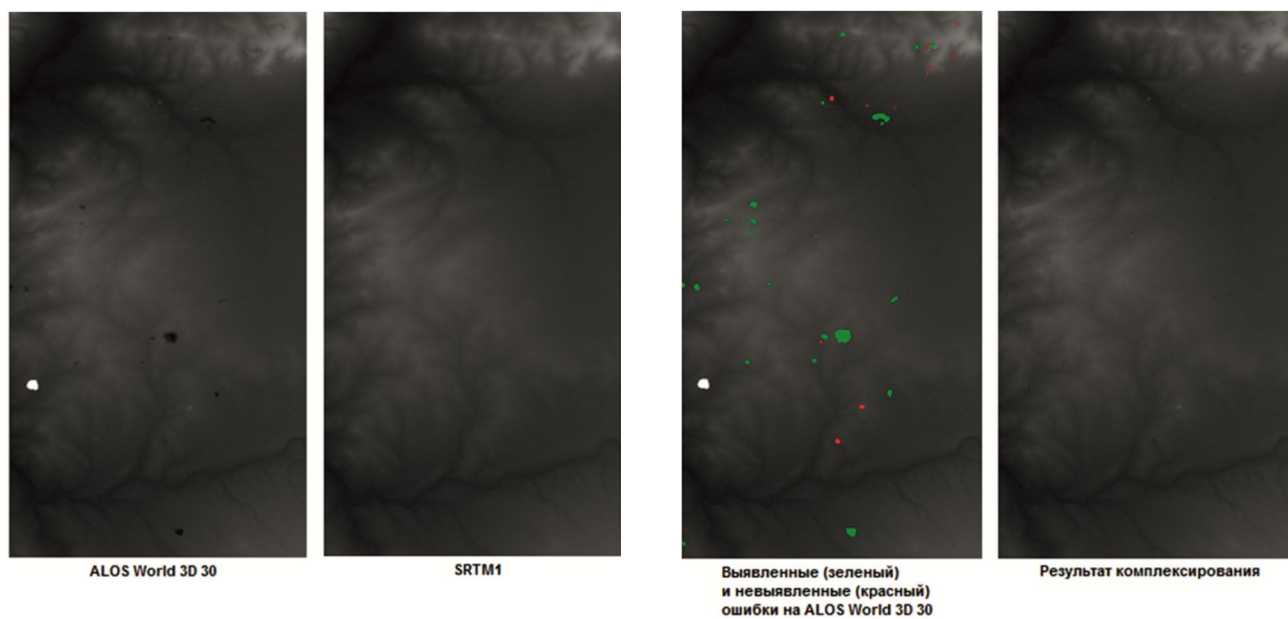


Рис. 2. Пример выявления и устранения при комплексировании крупных артефактов, вызванных наличием облачности при съемке

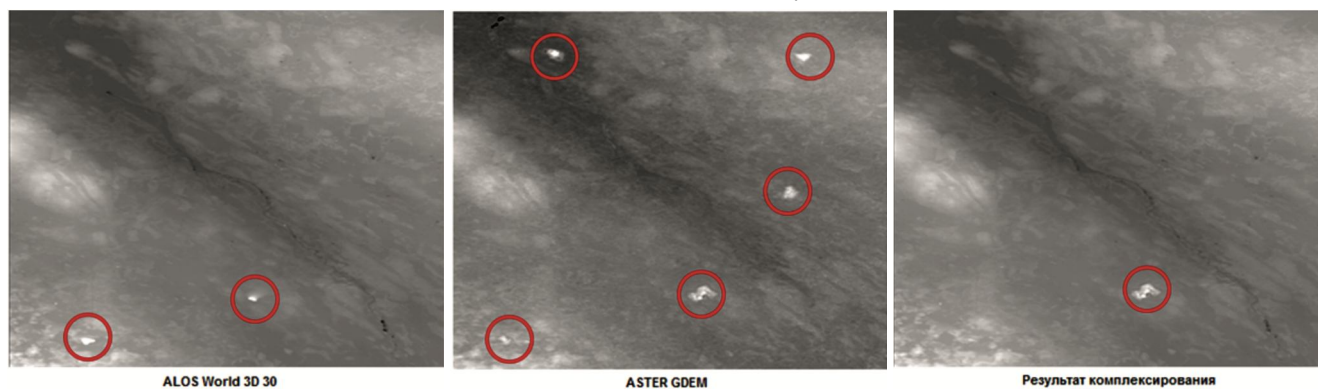


Рис. 3. Пример остаточной ошибки на результате комплексирования

Заключение

Предложенный алгоритм комплексирования ЦМР позволяет получить опорные данные о высоте рельефа для всей земной поверхности, предоставляющие для каждого локального участка наиболее точную из имеющейся информации. Результат комплексирования может содержать редкие локальные ошибки, либо неотделимые в автоматическом режиме от сюжета, либо вызванные тем, что на всех комплексируемых ЦМР в данном локальном участке нет правильных значений высоты.

Литература

1. Ушенкин В.А., Егошкин Н.А. Использование априорной информации при интерферометрической обработке высокодетальной радиолокационной информации // Вестник Самарского государственного аэрокосмического университета имени академика С.П. Королёва (национального исследовательского университета). 2016. Т. 15. № 2. – С. 208–219.
2. Кузнецов А.Е., Пошехонов В.И. Информационная технология стереофотограмметрической обработки видеоданных от многоматричных сканирующих устройств // Цифровая обработка сигналов. 2010. № 3. – С. 44–49.
3. Егошкин Н.А., Ушенкин В.А. Интерферометрическая обработка радиолокационной информации на основе комбинации методов развертывания фазы // Вестник Рязанского государственного радиотехнического университета. 2015. № 54-2. – С. 21–31.
4. Tadono T., Nagai H., Ishida H., Oda F., Naito S., Minakawa K., Iwamoto H. Generation of the 30 m-mesh global digital surface model by Alos PRISM // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2016. Vol. 41. Part B4. – P. 157–162.
5. Farr T.G., Kobrick M. The Shuttle Radar Topography Mission // RTO Meeting Proceedings 61: Space-Based Observation Technology. 2000. – P. 33-1–33-3.
6. Cuartero A., Felicísimo A.M., Ariza F.J. Accuracy of DEM generation from Terra-ASTER stereo data // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2004. Vol. 35. Part B2. – P. 559–563.
7. Shepard D. A two-dimensional interpolation function for irregularly-spaced data // Proc. of the 1968 ACM National Conference. 1968. P. 517–524.

АЛГОРИТМЫ ФОРМИРОВАНИЯ ИЗОБРАЖЕНИЙ ОБЪЕКТОВ В РАДИОМЕТРЕ С ДВУМЯ АНТЕННАМИ

*Клочко В.К., д.т.н., зав. кафедрой Рязанского государственного радиотехнического университета (РГРТУ),
e-mail: klochkovk@mail.ru;*

Гудков С.М., инженер-конструктор 2 кат. Рязанской радиоэлектронной компании (РРК);

Кошелев А.А., ведущий инженер-программист РРК.

ALGORITHMS OF OBJECTS IMAGES FORMATION IN THE RADIOMETER WITH TWO ANTENNAS

Klochko V.K., Gudkov S.M., Koshelev A. A.

Algorithms of joint array processing of radiometric observation in the radiometer with two antennas accepting signals in two frequency ranges are offered. Application of algorithms allows to create two image matrixes in the form of the selected segments of objects image with the best spatial resolution in this mode of observation array processing. Amplitudes of the selected segments contain information on the objects radio brightness temperature in two frequency ranges. Operability of algorithms is confirmed by results of a full-scale experiment.

Key words: radiometer, radiometric image, segmentation, Winer's filter, resolution capability, radio brightness temperature.

Ключевые слова: радиометр, радиометрическое изображение, сегментация, фильтр Винера, разрешающая способность.

Введение

Радиометр активно применяется для дистанционного измерения температуры объектов, излучающих электромагнитные волны в разных частотных диапазонах. Также радиометр является удобным всепогодным средством наблюдения за наземными объектами в режиме сканирования антенны [1, 2]. В результате сканирования формируется матрица радиометрического наблюдения (в дальнейшем – матрица наблюдений). Амплитуды элементов матрицы несут информацию о температуре излучающих элементов земной поверхности в данном частотном диапазоне, а в совокупности дают изображение объектов на местности. Объектом изучения в данной работе является сканирующий радиометр с двумя совмещенными антеннами, принимающими сигналы в двух разных частотных диапазонах. Одна антенна, имеет широкую диаграмму направленности (ДНА), вторая антенна – более узкую ДНА. Наличие двух антенн определяется необходимостью исследования излучающих свойств объектов в разных частотных диапазонах.

При одновременном сканировании антенн по азимуту съем данных осуществляется с определенным шагом дискретизации, определяющим количество элементов в строке формируемых матриц наблюдения. Переход к другой строке производится, как правило, изменением угла места на величину большую, чем шаг дискретизации по азимуту. Этим достигается увеличение скорости формирования матриц изображения в двух каналах первичной обработки принимаемых сигналов. Две матрицы получаются с пропусками строк (прореженные вдоль строк) и имеют одинаковые размеры, но отличаются пространственным разрешением (в дальнейшем –

Предложены алгоритмы совместной обработки матриц радиометрического наблюдения в радиометре с двумя антеннами, принимающими сигналы в двух частотных диапазонах. Применение алгоритмов позволяет формировать две матрицы изображения в виде выделенных сегментов изображения объектов с наилучшим пространственным разрешением в данном режиме обработки матриц наблюдения. Амплитуды выделенных сегментов содержат информацию о температуре объектов в двух частотных диапазонах. Работоспособность алгоритмов подтверждается результатами натурного эксперимента.

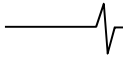
просто разрешением), зависящим от ширины ДНА. Пропущенные строки учитываются при синтезе алгоритмов восстановления изображений или заполняются методом интерполяции [3].

Разрешение матрицы наблюдения с широкой ДНА в несколько раз хуже, чем разрешение матрицы с узкой ДНА. Поэтому возникает необходимость повысить разрешение матрицы с широкой ДНА до разрешения матрицы с узкой ДНА, сохранив температурные характеристики частотных диапазонов. Повысить разрешение можно за счет оптического изображения контролируемого участка местности, как эталонного изображения [4]. Однако это не всегда возможно из-за плохих погодных условий и темного времени суток. Для радиометра с двумя антеннами эталонное изображение можно получить путем обработки матрицы наблюдения, сформированной для узкой ДНА.

Цель работы – разработка алгоритмов повышения разрешения матрицы наблюдения, полученной для широкой ДНА, до разрешения матрицы, полученной для узкой ДНА, с сохранением температурных характеристик объектов в двух частотных диапазонах.

Модель наблюдений и постановка задачи

Модель наблюдений в радиометре с двумя антеннами на выходе двух каналов первичной обработки имеет вид двумерной свертки:



$$y_q(i, j) = \sum_{i_1=-m_q}^{m_q} \sum_{j_1=-n_q}^{n_q} \alpha_q(i_1, j_1) x_q(i - i_1, j - j_1) + p_q(i, j), \quad (1)$$

$$q = \overline{1, Q}, \quad i_q = \overline{m_q + 1, h, M - m_q}, \quad j_q = \overline{n_q + 1, h, N - n_q},$$

где i и j – номера строки и столбца матрицы наблюдений $Y_q = \{y_q(i, j)\}$, полученной в q -м канале в процессе сканирования антенной радиометра по пространству; $y_q(i, j)$ – i -й, j -й элемент матрицы, представляющий амплитуду сигнала, принятого при i -м, j -м положении антенны и прошедшего тракт первичной обработки; $2m_q + 1$ и $2n_q + 1$ – ширина ДНА по азимуту и углу места q -й антенны (на уровне 0,5 мощности) в количестве строк и столбцов, причем $m_2 < m_1$ и $n_2 < n_1$, то есть первая антенна имеет широкую ДНА, а вторая – узкую; $\alpha_q(i, j)$, $i = \overline{-m_q, m_q}$, $j = \overline{-n_q, n_q}$ – нормированная аппаратная функция q -го канала (АФ), описывающая действие антенны и тракта первичной обработки на искомое изображение; $X_q = \{x_q(i, j)\}$, $i = \overline{1, M}$, $j = \overline{1, N}$ – матрица искомого изображения в M строк и N столбцов в q -м канале, элементы которой $x_q(i, j)$ представляют радиояркостную температуру i -го, j -го излучающего элемента поверхности в q -м диапазоне длин волн; h – шаг сканирования по углу места в q -х каналах в количестве элементов дискретизации, не превышающих ширину узкой ДНА (при $h > 1$ возникает пропуск строк искомой матрицы изображения; шаг сканирования по азимуту составляет 1 элемент дискретизации); $p_q(i, j)$ – шум аппаратуры в виде белого шума; Q – количество каналов ($Q = 1$ или 2) соответственно количеству антенн радиометра.

Формирование матриц Y_q , $q = \overline{1, Q}$, осуществляется одновременно, причем разрешение в матрице Y_2 лучше, чем в матрице Y_1 , в m_1/m_2 раз по строкам и n_1/n_2 по столбцам.

Задача 1 в рамках модели (1) без дополнительных операций обработки матриц наблюдения Y_q заключается в нахождении оценок $X_1^* = \{x_1^*(i, j)\}$ и $X_2^* = \{x_2^*(i, j)\}$ искомых матриц X_1 и X_2 с лучшим разрешением матрицы Y_2 и температурными характеристиками двух частотных диапазонов.

В рамках модели (1) задача заключается в нахождении оценок $X_1^* = \{x_1^*(i, j)\}$ и $X_2^* = \{x_2^*(i, j)\}$ искомых матриц X_1 и X_2 с пространственным разрешением матрицы Y_2 (с более узкой ДНА) и сохранением температурных характеристик двух частотных диапазонов.

Дополнительно повысить разрешение в матрице Y_2 (а также Y_1) можно с помощью алгоритмов восстановления изображений [3], действующих как оператор восстановления R на матрицы Y_q : $Z_q = R[Y_q] =$

$= \{z_q(i, j)\}$, $q = \overline{1, Q}$. Результатом являются оценки Z_q искомых изображений X_q с повышенным в несколько раз разрешением (примерно в 3 раза).

Если модель (1) представить в матричной форме, то восстановление осуществляется матричным методом, оптимальным в смысле метода наименьших квадратов, с учетом пропущенных строк [3]. Вид свертки (1) удобен для перехода в область пространственных частот и восстановления изображения с помощью фильтра Винера с заполнением пропущенных строк методом интерполяции [3]. Ограничимся рассмотрением второго подхода к восстановлению изображений с помощью фильтра Винера.

Оценки $z_q(i, j)$, полученные в результате восстановления, удовлетворяют модели, аналогичной (1):

$$z_q(i, j) = \sum_{i_1=-k_q}^{k_q} \sum_{j_1=-l_q}^{l_q} \beta_q(i_1, j_1) x_q(i - i_1, j - j_1) + \varepsilon_q(i, j), \quad (2)$$

$$q = \overline{1, Q}$$

где $i = \overline{1, M}$, $j = \overline{1, N}$, $\beta_q(i, j)$, $i = \overline{-k_q, k_q}$, $j = \overline{-l_q, l_q}$ ($k_q < m_q$, $l_q < n_q$) – нормированная функция рассеяния в i -й, j -й точке (ФРТ), обусловленная ошибками восстановления; $\varepsilon_q(i, j)$ – остаточный шум. Причем из $m_2 < m_1$, $n_2 < n_1$ следует: $k_2 < k_1$, $l_2 < l_1$.

В силу ошибок восстановления разрешение изображений в матрицах Z_q не достигает разрешения искомых матриц X_q . Тем не менее, матрицу Z_2 с лучшим разрешением можно взять за эталон по отношению к матрице Z_1 .

В рамках модели (2) при наличии операций восстановления задача заключается в нахождении оценок $X_1^{**} = \{x_1^{**}(i, j)\}$ и $X_2^{**} = \{x_2^{**}(i, j)\}$ искомых матриц X_1 и X_2 с разрешением матрицы Z_2 и температурными характеристиками частотных диапазонов матриц X_1 и X_2 .

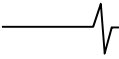
Решение поставленных задач

Алгоритм 1 повышения разрешения матрицы Y_1 без операций восстановления изображения заключается в следующем.

1. В результате сканирования антенны радиометра с широкой ДНА по азимуту и углу места зоны обзора формируется матрица изображения $Y_1 = \{y_1(i, j)\}$ с элементами $y_1(i, j)$, отвечающими модели (1).

2. Одновременно вторая антенна радиометра с узкой ДНА сканирует ту же зону обзора, в результате чего формируется вторая матрица изображения $Y_2 = \{y_2(i, j)\}$ модели (1).

3. Пропущенные при сканировании антенн строки матриц Y_1 и Y_2 заполняются методом интерполяции (линейной, биквадратной или бикубической) путем обработки элементов двух соседних наблюдаемых строк.



4. Матрица Y_2 разбивается на K непересекающихся однородных по амплитуде подобластей $D_1, D_2, \dots, D_K$ с помощью оператора сегментации [5]. В результате получается матрица $S = \{S(i, j)\}$, где $S(i, j)$ – номер сегмента, которому принадлежит i -й, j -й элемент матрицы Y_2 , поставленный в соответствие i -му, j -му элементу матрицы Y_1 .

5. Для каждого s -го сегмента вычисляется средняя радиометрическая амплитуда $\tilde{y}_1(s)$ на основе i -х, j -х элементов матрицы Y_1 с меткой s :

$$\tilde{y}_1(s) = (1/n_s) \sum_i \sum_j y_1(i, j), \quad i, j: S(i, j) = s,$$

где n_s – количество элементов с меткой s .

6. Всем i -м, j -м элементам матрицы Y_1 с меткой s присваивается амплитуда $\tilde{y}_1(s)$, нормированная коэффициентом μ_1 , и запоминается как $x_1^*(i, j) = \mu_1 \cdot \tilde{y}_1(s)$, где μ_1 учитывает связь амплитуды с температурой.

В результате формируется матрица $X_1^* = \{x_1^*(i, j)\}$ изображения с пространственным разрешением матрицы Y_2 и температурными характеристиками матрицы X_1 .

7. Все элементы матрицы Y_2 нормируются коэффициентом μ_2 и запоминаются как $x_2^*(i, j) = \mu_2 \cdot y_2(i, j)$. В результате формируется матрица $X_2^* = \{x_2^*(i, j)\}$ изображения с пространственным разрешением Y_2 и температурными характеристиками матрицы X_2 .

Алгоритм 2 повышения разрешения матрицы Y_1 с учетом операций восстановления изображений в Y_1 и Y_2 отличается от алгоритма 1 следующим.

1-3. Данные пункты совпадают с п.п. 1-3 алгоритма 1 при формировании матриц наблюдения Y_1 и Y_2 .

4. Полученные матрицы Y_1 и Y_2 подвергаются двумерному преобразованию Фурье и получаются спектральные матрицы $Y_{f1} = \{y_{f1}(i, j)\}$ и $Y_{f2} = \{y_{f2}(i, j)\}$.

5. Элементы матриц Y_{f1} и Y_{f2} умножаются на вычисленную передаточную функцию $w_f(i, j)$ восстанавливающего фильтра Винера [4], что приводит к спектральным матрицам оценок $Z_{f1} = \{z_{f1}(i, j)\}$ и $Z_{f2} = \{z_{f2}(i, j)\}$:

$$z_{f1}(i, j) = y_{f1}(i, j) \cdot w_f(i, j),$$

$$z_{f2}(i, j) = y_{f2}(i, j) \cdot w_f(i, j).$$

6. Матрицы Z_{f1} и Z_{f2} подвергаются обратному преобразованию Фурье: и получаются матрицы $Z_1 = \{z_1(i, j)\}$, $Z_2 = \{z_2(i, j)\}$ восстановленного изображения объектов в пространственной области.

7. Матрица Z_2 разбивается на сегменты аналогично п. 4 алгоритма 1.

8. Для каждого s -го сегмента вычисляется средняя радиометрическая амплитуда $\tilde{z}_1(s)$ аналогично п. 5 алгоритма 1.

9. Всем i -м, j -м элементам матрицы Z_1 с меткой s присваивается амплитуда $\tilde{z}_1(s)$, нормированная коэффициентом μ_1 и запоминается как $x_1^{**}(i, j) = \mu_1 \cdot \tilde{z}_1(s)$. В результате формируется матрица $X_1^{**} = \{x_1^{**}(i, j)\}$ изображения с пространственным разрешением матрицы Z_2 и температурными характеристиками матрицы X_1 .

10. Все элементы матрицы Z_2 нормируются коэффициентом μ_2 и запоминаются как $x_2^{**}(i, j) = \mu_2 \cdot z_2(i, j)$. В результате формируется матрица $X_2^{**} = \{x_2^{**}(i, j)\}$ изображения с пространственным разрешением Z_2 и температурными характеристиками X_2 .

Алгоритм 3 повышения разрешения матрицы Y_1 с учетом операций восстановления изображения только в матрице Y_2 отличается следующим.

С целью повышения быстродействия алгоритма 1 целесообразно использовать алгоритм 3, в котором операции восстановления осуществляются только над матрицей Y_2 при формировании матрицы Z_2 . 1-3. Данные пункты алгоритма 3 совпадают с п.п. 1-3 алгоритмов 1 и 2 при формировании матриц наблюдения Y_1 и Y_2 .

4-7. В данных пунктах операции выполняются только над матрицей Y_2 .

8-10. Данные пункты совпадают с п.п. 8-10 алгоритма 2 с тем отличием, что под символом Z_1 понимается Y_1 .

Результаты эксперимента

Натурные испытания проводились с применением радиометра с двумя антеннами: первая антенна с шириной ДНА в 3^0 принимала сигналы в 8 мм диапазоне длин волн при наблюдении объектов на местности на расстоянии 30 м, вторая антенна – с шириной ДНА в 1^0 , совмещенная с первой антенной, одновременно принимала сигналы в 3 мм диапазоне. Шаг сканирования по углу места и пропуск строк соответствовали 1^0 . На рис. 1 показано видео- изображение наблюдаемого участка местности с тремя объектами в виде щитов.



Рис. 1. Видеоизображение наблюдаемого участка местности

На рис. 2 слева направо – изображение в матрице Y_1 , соответствующее 8 мм диапазону, до и после фильтра Винера. На рис. 3 – изображение в матрице Y_2 , соответствующее 3 мм диапазону, до и после фильтра Винера. На рис. 4 – тепловое изображение в цвете (по трем уровням) после сегментации в матрице Y_2 и тепловое изображение в сегментах матрицы Y_1 , полученные алгоритмом 1 без фильтра Винера. На рис. 5 – тепловое изображение в цвете после сегментации в матрице Z_2 и тепловое изображение в сегментах матрицы Z_1 , полученные алгоритмами 2 и 3 с фильтром Винера.

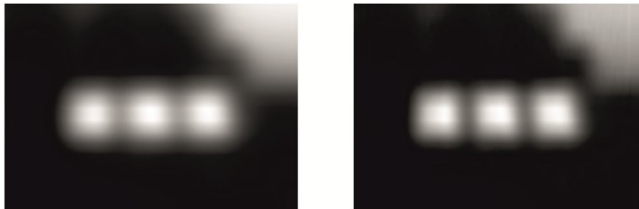


Рис. 2. Видеоизображение в матрице Y_1 , соответствующее 8 мм диапазону, до и после фильтра Винера

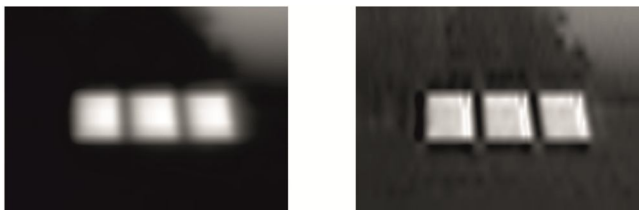


Рис. 3. Видеоизображение в матрице Y_2 , соответствующее 3 мм диапазону, до и после фильтра Винера



Рис. 4. Тепловое изображение в цвете (по трем уровням) после сегментации в матрице Y_2 и тепловое изображение в сегментах матрицы Y_1 , полученные алгоритмом 1 без фильтра Винера



Рис. 5. Тепловое изображение в цвете после сегментации в матрице Z_2 и тепловое изображение в сегментах матрицы Z_1 , полученные алгоритмами 2 и 3 с фильтром Винера

Заключение

Изображения объектов на рис. 4 и 5 в виде выделенных сегментов, полученные с помощью предлагаемых алгоритмов, соответствуют наилучшему разрешению в данном режиме обработки матриц наблюдения в сравнении с изображениями объектов на рис. 2 и 3, а амплитуды выделенных сегментов содержат информацию о температуре объектов в двух частотных зонах.

Литература

1. Николаев А.Г., Перцов С.В. Радиотеплолокация (пассивная радиолокация). М.: Сов. радио, 1964. 335 с.
2. Шарков Е.А. Радиотепловое дистанционное зондирование Земли: физические основы: в 2 т. / Т. 1. М.: ИКИ РАН, 2014. – 544 с.
3. Клочко В.К., Кузнецов В.П. Методы восстановления изображений и оценивания аппаратной функции по прореженной матрице наблюдений // Автометрия. 2016. Т. 52, № 56. – С. 12 – 21.
4. Клочко В.К., Макарова О.Н., Гудков С.М., Кошелев А.А. Алгоритм формирования теплового изображения объектов при радиометрическом наблюдении // Цифровая обработка сигналов. 2016. № 3. – С. 34 – 36.
5. Гонсалес Р., Вудс Р., Эддинс С. Цифровая обработка изображений в среде MATLAB. М.: Техносфера, 2006. – 616 с.

АДАПТИВНЫЕ РЕЖЕКТОРНЫЕ ФИЛЬТРЫ С ДЕЙСТВИТЕЛЬНЫМИ ВЕСОВЫМИ КОЭФФИЦИЕНТАМИ

Попов Д.И., д.т.н., профессор кафедры радиотехнических систем Рязанского государственного радиотехнического университета, e-mail: adop@mail.ru.

ADAPTIVE REJECTOR FILTERS WITH REAL WEIGHTS

Popov D.I.

The adaptation algorithms with real weights are synthesized and the relevant principles of the adaptive rejector filters (ARF) are considered. The region of the appropriate use of the ARF data determined. The analysis of the ARF efficiency conducted depending on the volume of training sample and clutter parameters.

Key words: adaptation, real weights, training sample set, clutter, rejector filters, rejection efficiency.

Ключевые слова: адаптация, действительные весовые коэффициенты, обучающая выборка, пассивные помехи, режекторные фильтры, эффективность режектирования.

Введение

Пассивные помехи со времен Второй мировой войны и до настоящего времени являются эффективным средством противорадиолокационной маскировки [1, 2]. С тех пор одной из актуальных в радиолокации является проблема обнаружения сигналов движущихся целей на фоне пассивных помех. Априорная неопределенность спектрально-корреляционных характеристик помехи, а также их неоднородность и нестационарность в зоне обзора существенно затрудняют реализацию предельных возможностей основной операции при обнаружении сигналов движущихся целей на фоне пассивных помех – режектирования спектральных составляющих помехи. Преодоление априорной неопределенности параметров помехи в соответствии с методологией адаптивного байесовского подхода основывается на методах адаптации к неизвестным корреляционным параметрам помехи – аргументу и модулю коэффициентов межпериодной корреляции – путем замены этих параметров их состоятельными оценками [3, 4], что приводит к алгоритмам адаптивного режектирования с комплексными весовыми коэффициентами и соответствующим адаптивным режекторным фильтрам (АРФ) [5]. Реализация данных АРФ в цифровом виде требует высокого быстродействия выполнения арифметических операций. Избежать указанных трудностей можно путем предварительной компенсации доплеровского сдвига фазы помехи. В работе [6] синтезированы алгоритмы оценивания и предложены принципы построения и структурные схемы автокомпенсаторов доплеровской фазы пассивных помех с прямой и обратной связью. Особенности адаптации к корреляционным свойствам помехи на выходе автокомпенсатора и последующего ее режектирования рассмотрены в работе [7]. Определенное упрощение процедуры адаптации достигается в АРФ каскадного типа [8].

Поиски вариантов упрощения процедуры адаптации сохраняют свою актуальность. Одним из вариантов

Синтезированы алгоритмы адаптации с действительными весовыми коэффициентами и рассмотрены принципы построения соответствующих адаптивных режекторных фильтров (АРФ). Определена область целесообразного применения данных АРФ. Проведен анализ эффективности АРФ в зависимости от объема обучающей выборки и параметров помехи.

упрощения является переход от комплексных весовых коэффициентов к действительным. Представляет интерес синтез и анализ АРФ с действительными весовыми коэффициентами и определение области целесообразного их использования.

Синтез адаптивных режекторных фильтров

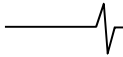
Синтез АРФ с действительными весовыми коэффициентами состоит в определении оптимального по критерию эффективности выделения сигнала на фоне поступающей помехи вектора весовых коэффициентов $\{G_k\}$ при ограничивающем условии, согласно которому его проекции являются действительными величинами, т.е. $G_k = g_k$. Аналогично [5] в качестве критерия эффективности используем коэффициент улучшения отношения сигнал/помеха, который для нерекурсивного АРФ рассматриваемого типа m -го порядка принимает вид

$$\mu(\varphi) = \frac{\sum_{j,k=0}^m g_j g_k R_{jk}^c}{\left(\sum_{j,k=0}^m g_j g_k R_{jk}^n + \lambda \sum_{j=0}^m g_j^2 \right)},$$

где R_{jk}^c , R_{jk}^n – элементы корреляционных матриц сигнала и помехи соответственно, причем $R_{jk} = \rho_{jk} e^{i(j-k)\theta}$;

$\rho_{jk} = \rho[(j-k)T]$ – коэффициенты межпериодной корреляции; θ – доплеровский сдвиг фазы за период повторения T для сигнала ($\theta_c = \varphi$) или помехи ($\theta_n = \psi$); $\lambda = \sigma_{ш}^2 / \sigma_n^2$ – отношение собственной шум/пассивная помеха на входе фильтра.

Учитывая инвариантность параметров АРФ к корреляционным характеристикам сигнала и предполагая равномерное распределение величины φ в интервале $(-\pi, \pi)$, соответствующем интервалу однозначности доплеровской частоты $(-1/2T, 1/2T)$ [2], найдем



$$\mu = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mu(\varphi) d\varphi =$$

$$= \sum_{j,k=0}^m g_j g_k \tilde{R}_{jk}^c / \left(\sum_{j,k=0}^m g_j g_k R_{jk}^n + \lambda \sum_{j=0}^m g_j^2 \right),$$

где $\tilde{R}_{jk}^c = \rho_{jk}^c \text{sinc}[(j-k)\pi]$ – преобразованные в результате интегрирования элементы корреляционной матрицы сигнала.

С учетом свойства sinc-функции

$$\text{sinc}[(j-k)\pi] = \frac{\sin[(j-k)\pi]}{(j-k)\pi} = \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases}$$

окончательно получаем

$$\mu = \sum_{j,k=0}^m g_j^2 / \left(\sum_{j,k=0}^m g_j g_k R_{jk}^n + \lambda \sum_{j=0}^m g_j^2 \right) = \left[\left(\sum_{j,k=0}^m g_j g_k r_{jk} / \sum_{j=0}^m g_j^2 \right) + \lambda \right]^{-1}, \quad (1)$$

где $r_{jk} = \text{Re } R_{jk}^n = \rho_{jk}^n \cos(j-k)\psi$.

Оптимальный вектор $\{g_k\}$ для АРФ невысоких порядков ($m \leq 3$) найдем из условия

$$\mu_{\max} \rightarrow \max_{\{g_k\}} \mu.$$

В соответствии с экстремальными свойствами характеристических (собственных) чисел матриц [9] минимум отношения Релея в выражении (1) является наименьшее собственное число $\alpha_{\min}$ матрицы $[r_{jk}]$. При этом

$\mu_{\max} = (\alpha_{\min} + \lambda)^{-1}$, а оптимальный вектор $\{g_k\}$ не зависит от λ и определяется как собственный вектор матрицы $[r_{jk}]$, соответствующий $\alpha_{\min}$, из системы уравнений

$$\sum_{k=0}^m (r_{jk} - \alpha_{\min} \delta_{jk}) g_k = 0, \quad j = \overline{0, m},$$

где δ_{jk} – символ Кронекера, $\alpha_{\min}$ – наименьший корень характеристического уравнения

$$\det[r_{jk} - \alpha_{\min} \delta_{jk}] = 0.$$

Преодоление априорной неопределенности параметров помехи основывается на адаптивном байесовском подходе [3], в соответствии с которым неизвестные величины r_{jk} заменяются их состоятельными оценками $\hat{r}_{jk}$.

Оптимальные весовые коэффициенты $\hat{g}_k$ определяются при условии $\hat{g}_0 = g_0 = 1$ оценками $\hat{r}_{1k}$ в соответствии с конкретными в зависимости от порядка фильтра адаптивными алгоритмами. В частности, при $m=1$ получаем $\hat{g}_1 = g_1 = -1$. С учетом свойства симметрии коэффициентов $\hat{g}_k = (-1)^m \hat{g}_{m-k}$ найдем конкретный вид адаптивных алгоритмов для других порядков АРФ канонического типа:

$$\text{при } m=2 \quad g_0 = g_2 = 1, \quad \hat{g}_1 = -\frac{2\hat{r}_{12}}{1 - \hat{\alpha}_{\min}},$$

$$\text{где } \hat{\alpha}_{\min} = 1 + \frac{\hat{r}_{13}}{2} \left(1 - \sqrt{1 + \left(\frac{\hat{r}_{12}}{\hat{r}_{13}} \right)^2} \right);$$

при $m=3 \quad g_0 = -g_3 = 1, \quad \hat{g}_1 = -\hat{g}_2 =$

$$-\left(1 + \frac{\hat{r}_{12} - \hat{r}_{14}}{1 - \hat{r}_{13} - \hat{\alpha}_{\min}} \right) \cong -\left(1 + \frac{\hat{r}_{12} - \hat{r}_{14}}{1 - \hat{r}_{13}} \right),$$

где величина $\hat{\alpha}_{\min}$ по сравнению с $r_{jj} = 1$ пренебрежимо мала. Это позволяет не учитывать ее, что практически не влияет на эффективность режектирования помехи.

С учетом ограничения $g_0 = 1$, исключая тривиальное решение $\{g_k\} = 0$, весовые коэффициенты могут быть определены в соответствии с модифицированным критерием

$$\gamma = \sum_{j,k=0}^m g_j g_k r_{jk}, \quad (2)$$

который представляет собой упрощенный вариант критерия (1).

Оптимальный по критерию (2) вектор $\{g_k\}$, соответствующий условию $\gamma_{\min} \rightarrow \min_{\{g_k\}} \gamma$, определяется из

уравнения

$$\nabla_{\{g_k\}} \left(\sum_{j,k=0}^m g_j g_k r_{jk} \right) = 0,$$

где $\nabla_{\{g_k\}} = \left\{ \frac{\partial}{\partial g_1}, \dots, \frac{\partial}{\partial g_s} \right\}$ – оператор градиента по варьируемым компонентам вектора $\{g_k\}$, $s = m/2$ или $(m-1)/2$ соответственно при четном или нечетном m .

Используя оценочные значения коэффициентов $\hat{r}_{1k}$, с учетом симметрии весовых коэффициентов $\hat{g}_k$ для адаптивных алгоритмов при $m \leq 3$ получаем:

при $m=1 \quad g_0 = 1, \quad \hat{g}_1 = -\hat{r}_{12};$

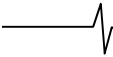
при $m=2 \quad g_0 = g_2 = 1, \quad \hat{g}_1 = -2\hat{r}_{12};$

при $m=3 \quad g_0 = -g_3 = 1, \quad \hat{g}_1 = -\hat{g}_2 = -\frac{\hat{r}_{12} - \hat{r}_{13}}{1 - \hat{r}_{12}}.$

Принципы построения адаптивных режекторных фильтров

Структуру АРФ удобно синтезировать в канонической форме в соответствии с системной функцией (8) в z -плоскости работы [5] при $e^{i\psi} = 1$. Исключив из приведенной в [5] на рис. 2 структурной схемы блок оценивания доплеровской фазы помехи и соответствующие комплексные множители, приходим к АРФ с действительными весовыми коэффициентами. Фильтры высоких порядков ($m > 2$) могут быть также реализованы в виде каскадного включения звеньев 1-го и 2-го порядков. Аналогично работе [8] адаптация последующего звена осуществляется по оценкам остатков помехи на выходе предыдущего звена.

Оценки $\hat{r}_{1k}$ вычисляются по цифровым отсчетам по-



мехи $U_{jl} = x_{jl} + iy_{jl}$, поступающим с n смежных элементов разрешения по дальности, в соответствии с предложенным в работе [7] для получения оценок максимального правдоподобия (ОМП) действительной части комплексного коэффициента корреляции алгоритмом $\hat{r}_{1k} = \text{Re } X_{1k} / (a_1 a_k)^{1/2}$, $k = \overline{2, m}$, (3) где

$$\begin{aligned} \text{Re } X_{1k} &= \text{Re} \sum_{l=1}^{n+1} U_{jl}^* U_{j+k-1, l} = \\ &= \sum_{l=1}^{n+1} (x_{jl} x_{j+k-1, l} + y_{jl} y_{j+k-1, l}), \\ a_1 &= \sum_{l=1}^{n+1} |U_{jl}|^2 = \sum_{l=1}^{n+1} (x_{jl}^2 + y_{jl}^2), \\ a_k &= \sum_{l=1}^{n+1} |U_{j+k-1, l}|^2 = \sum_{l=1}^{n+1} (x_{j+k-1, l}^2 + y_{j+k-1, l}^2). \end{aligned}$$

где n – объем обучающей выборки, определяемый числом независимых отсчетов со смежных элементов разрешения по дальности, за исключением отсчета с номером $l = n/2 + 1$.

Так как адаптивную обработку на основе поступающих оценок для уменьшения ошибок рассогласования следует осуществлять после предварительной задержки исходных данных для среднего элемента разрешения временного строба, образующего обучающую выборку, то целесообразно исключить из обучающей выборки соответствующий этому элементу отсчет U_{jl} с номером

$l = n/2 + 1$. Тогда в случае сигнала, соизмеримого с помехой, или разрывной помехи при обработке элемента разрешения, содержащего сигнал, ослабление или подавление сигнала за счет его влияния на используемые оценки не происходит.

Структурная схема измерителя оценки $\hat{r}_{12}$ в соответствии с алгоритмом (3) приведена на рис. 1, где $ЗУ_T$ – запоминающее на период повторения T устройство, Н – накопитель, осуществляющий скользящее по дальности накопление; БО – блок объединения, вычисляющий сумму квадратов проекций; Д – делитель. При $\psi \neq 0$ данный измеритель вычисляет действительные части комплексных коэффициентов корреляции.

В целом структура синтезированного АРФ оказывается существенно проще структуры АРФ с комплексными весовыми коэффициентами и является ее упрощенным вариантом. Адаптация к модулю и аргументу коэффициентов корреляции, входящих в оценки $\hat{r}_{1k}$, осуществляется совместно путем оптимизации взаимного положения нулей системной функции АРФ без их общего поворота на угол ψ , что приводит к деформациям АЧХ и потерям в эффективности по сравнению с инвариантными к величине ψ АРФ с комплексными весовыми коэффициентами.

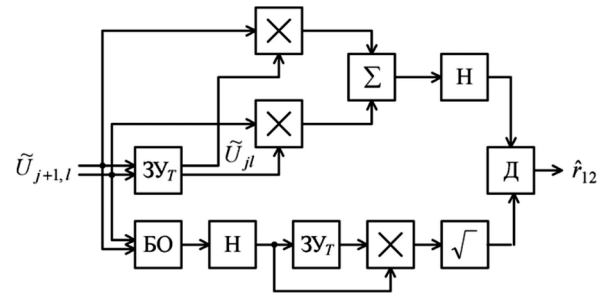


Рис. 1. Структурная схема измерителя оценки коэффициента корреляции помехи

Анализ адаптивных режекторных фильтров

Эффективность рассматриваемого типа АРФ 2-го порядка ($m = 2$) характеризуется приведенными на рис. 2 зависимостями коэффициента улучшения отношения сигнал/помеха μ от доплеровского сдвига фазы ψ помехи для различных значений нормированной ширины β гауссовского спектра помехи. Сплошные кривые

соответствуют $\lambda = 0$, штриховые – $\lambda = 10^{-4}$. Как видно, эффективность данных АРФ падает с ростом сдвига фазы ψ помехи, приводя к существенным потерям по отношению к АРФ с комплексными весовыми коэффициентами (штрихпунктирные прямые). Только в области малых доплеровских сдвигов фазы помехи эффективность АРФ обоих типов сближается. С расширением спектра помехи эта область увеличивается, а потери уменьшаются. Наличие некоррелированного шума снижает эффективность АРФ в целом и величину рассматриваемых потерь, особенно при узкополосной помехе и возрастании порядка АРФ.

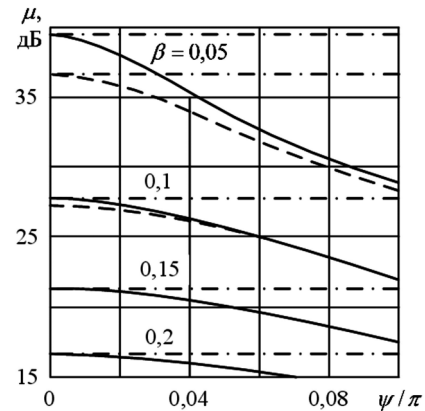


Рис. 2. Зависимости эффективности АРФ от доплеровской фазы помехи

Таким образом, применение АРФ с действительными весовыми коэффициентами оправдано при ограниченной и сравнительно малой в зависимости от порядка фильтра и ожидаемых параметров помехи величине ее доплеровского сдвига фазы.

Проведем анализ эффективности АРФ в зависимости от параметров помехи и погрешностей адаптации (объема обучающей выборки n). Выходная величина АРФ с действительными весовыми коэффициентами $\hat{g}_k$ имеет вид

$$V = \sum_{k=0}^m \hat{g}_k U_{m-k+1}.$$

Учитывая независимость оценок параметров помехи и отсчетов обрабатываемого элемента разрешения, исключаемых из обучающей выборки [7], для дисперсии помехи с шумом на выходе АРФ находим

$$\sigma_V^2 = \overline{VV^*} / 2 = \sigma_n^2 \sum_{j,k=0}^m \overline{\hat{g}_j \hat{g}_k} r_{jk} + \sigma_{\text{ш}}^2 \sum_{j=0}^m \overline{\hat{g}_j^2}. \quad (4)$$

Коэффициенты $\hat{g}_j = g_j(\{\hat{r}_{il}\})$, т.е. определяются вектором оценок $\{\hat{r}_{il}\}$, где $l = \overline{2, m+1}$ для критерия (1) и $l = \overline{2, m}$ для критерия (2). Для перехода от текущих значений оценок $\hat{r}_{il}$ к параметрам их распределений произведем соответствующие усреднения в (4), используя асимптотические свойства оценок максимального правдоподобия (ОМП). Разложив функцию $g_j(\{\hat{r}_{il}\})$ в многомерный ряд Тейлора в окрестностях точки $\{r_{il}\}$ и ограничившись ввиду малости $\{\hat{r}_{il} - r_{il}\}$ двумя первыми членами, получим

$$\hat{g}_j = g_j(\{\hat{r}_{il}\}) = g_j(\{r_{il}\}) + \sum_{l=2}^{m+1} (\hat{r}_{il} - r_{il}) \frac{\partial g_j(\{r_{il}\})}{\partial r_{il}}.$$

С учетом $g_j(\{r_{il}\}) = g_j$ и асимптотической нормальности распределения вектора оценок $\{\hat{r}_{il}\}$ с вектором средних $\{r_{il}\}$ найдем

$$\begin{aligned} \overline{\hat{g}_j \hat{g}_k} &= \overline{\left[g_j + \sum_{l=2}^{m+1} (\hat{r}_{il} - r_{il}) \frac{\partial g_j}{\partial r_{il}} \right] \left[g_k + \sum_{p=2}^{m+1} (\hat{r}_{ip} - r_{ip}) \frac{\partial g_k}{\partial r_{ip}} \right]} = \\ &= g_j g_k + \sum_{l,p=2}^{m+1} K_{l-1,p-1} \frac{\partial g_j}{\partial r_{il}} \frac{\partial g_k}{\partial r_{ip}}, \end{aligned} \quad (5)$$

где $K_{l-1,p-1} = \overline{(\hat{r}_{il} - r_{il})(\hat{r}_{ip} - r_{ip})}$ – элементы ковариационной матрицы оценок коэффициентов корреляции, определяемые с учетом асимптотической эффективности ОМП как элементы матрицы, обратной информационной матрице Фишера [3]. (Выражение для соответствующих элементов матрицы Фишера получено в работе [10].)

При известной форме огибающей корреляционной функции помехи оценивается только один коэффициент корреляции $r_{12} = r$, а матрица Фишера содержит единственный элемент. Величина, обратная этому элементу, является дисперсией $\sigma_{\hat{r}}^2$ оценки $\hat{r}$.

Весовые коэффициенты АРФ теперь определяются зависимостями $\hat{g}_j = g_j(\hat{r})$, разложение которых в окрестности r принимает вид $\hat{g}_j = g_j + (\hat{r} - r)g'_j$, что приводит к упрощению выражения (5) в виде

$$\overline{\hat{g}_j \hat{g}_k} = g_j g_k + \sigma_{\hat{r}}^2 g'_j g'_k. \quad (6)$$

Используя в соотношении (4) выражение (5) или (6), получим соответствующие варианты окончательной формулы. При использовании (6) имеем

$$\frac{\sigma_V^2}{\sigma^2} = \sum_{j,k=0}^m g_j g_k r_{jk} + \lambda \sum_{j=0}^m g_j^2 +$$

$$+ \sigma_{\hat{r}}^2 \left\{ \sum_{j,k=1}^{m-1} g'_j g'_k r_{jk} + \lambda \sum_{j=1}^{m-1} g_j'^2 \right\}. \quad (7)$$

Ввиду асимптотической эффективности ОМП коэффициента корреляции $\hat{r}$ для определения дисперсии $\sigma_{\hat{r}}^2$ используем выражение Крамера–Рао, в соответствии с которым:

$$\sigma_{\hat{r}}^2 = - \left[\partial^2 \ln P(\{U_{jl}\} / r) / \partial r^2 \right]^{-1}.$$

Функция правдоподобия при этом имеет вид

$$P(\{U_{jl}\} / r) = (2\pi)^{-2n} \det^n [W_{jk}] \times$$

$$\times \exp \left\{ -\frac{1}{2} \sum_{l=1}^n \sum_{j,k=1}^2 W_{jk} U_{jl}^* U_{kl} \right\},$$

где W_{jk} – элементы матрицы $[W_{jk}]$, обратной матрице помехи $[R_{jk}]$, т.е. $[W_{jk}] = [R_{jk}]^{-1}$.

Учитывая, что $R_{jk} = \rho_{jk} e^{i(j-k)\psi} = \text{Re} R_{jk} + i \text{Im} R_{jk}$ и обозначая $\rho_{12} = \rho$ и $\text{Re} R_{12} = r_{12} = r$, в результате соответствующих вычислений получаем

$$\begin{aligned} \sigma_{\hat{r}}^2 &= \frac{\det^2 [R_{jk} + \lambda \delta_{jk}]}{2n(\det [R_{jk} + \lambda \delta_{jk}] + 2r^2)} = \\ &= \frac{[(1+\lambda)^2 - \rho^2]^2}{2n[(1+\lambda)^2 - \rho^2 + 2r^2]}. \end{aligned}$$

В целом эффективность АРФ характеризуется усредненным по доплеровской фазе сигнала коэффициентом улучшения отношения сигнал/помеха (1), который с учетом ошибок адаптации в соответствии с выражением (7) имеет вид

$$\mu = \sum_{j=0}^m g_j^2 \left/ \frac{\sigma_V^2}{\sigma^2} \right.$$

Погрешности оценивания неизвестных параметров помехи приводят к потерям $\Delta\mu$ по отношению к предельной эффективности АРФ. Расчеты показывают, что влияние параметров помехи на величину потерь оказывается несущественным. Анализ влияния ошибок адаптации показывает, что эффективность АРФ приближается к предельной при сравнительно небольшом объеме обучающей выборки. В частности, при $n \geq 4$ величина потерь составляет доли децибела.

Заключение

Синтезированы алгоритмы адаптации с действительными весовыми коэффициентами и рассмотрены принципы построения соответствующих АРФ, адаптирующихся к действительной части комплексных коэффициентов корреляции пассивной помехи. Исследование эффективности АРФ показало целесообразность их применения при ограниченной и сравнительно малой в зависимости от порядка фильтра и ожидаемых параметров помехи величине ее доплеровской скорости. Проведенный анализ АРФ устанавливает связь между эффективностью АРФ и объемом обучающей выборки в зависимости от порядка и структуры АРФ и параметров поме-

хи. Показано, что эффективность АРФ приближается к предельной при сравнительно небольшом объеме обучающей выборки.

Литература

1. Степанов Ю.Г. Противорадиолокационная маскировка. – М.: Сов. радио, 1968. – 144 с.
2. Richards M.A., Scheer J.A., Holm W.A. (Eds.). Principles of Modern Radar: Basic Principles. New York: SciTech Publishing, IET, Edison. 2010. – 924 p.
3. Репин В.Г., Тартаковский Г.П. Статистический синтез при априорной неопределенности и адаптация информационных систем. М.: Сов. радио. 1977. – 432 с.
4. Попов Д.И. Оптимальная обработка многочастотных сигналов // Известия вузов России. Радиоэлектроника. – 2013. – Вып. 1. – С. 32–39.
5. Попов Д.И. Адаптивные режекторные фильтры с комплексными весовыми коэффициентами // Вестник Концерна ПВО «Алмаз – Антей». – 2015. – № 2 (14). – С. 21-26.
6. Попов Д.И. Автокомпенсация доплеровской фазы пассивных помех // Цифровая обработка сигналов. – 2009. – № 2. – С. 30–33.
7. Попов Д.И. Адаптивное подавление пассивных помех // Цифровая обработка сигналов. – 2014. – № 4. – С. 32-37.
8. Попов Д.И. Адаптивные режекторные фильтры каскадного типа // Цифровая обработка сигналов. – 2016. – № 2. – С. 53-56.
9. Гантмахер Ф.Р. Теория матриц. М.: Наука, 1988. – 552 с.
10. Попов Д.И. Анализ адаптивных систем обнаружения сигналов на фоне пассивных помех // Радиотехника. – 2016. – № 11. – С. 74-79.

УДК 621.396.669

ОБРАБОТКА И КОДИРОВАНИЕ РЕЧЕВЫХ СИГНАЛОВ С ПОМОЩЬЮ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

*Кириллов С.Н., зав. кафедрой РУС РГРТУ, д.т.н., профессор, e-mail: kirillov.lab@mail.ru;
Сазонова Е.С., магистрант РГРТУ, e-mail: katrinik94@mail.ru.*

PROCESSING AND CODING VOICE SIGNAL USING ARTIFICIAL NEURAL NETWORKS

Kirillov S.N., Sazonova E.S.

The algorithm of noise reduction in speech signals, and the theoretical and practical aspects of design and optimization structures of neural network implementations predictors of speech signals. Managed to achieve a decrease in the level of dispersion of acoustic noise dozens of time. Justify bath implementation advantages predictors based on artificial neural networks compared with known predictors based on non-recursive FIR filters. In proved possible to reduce the prediction order of 10 % to 60 % at the same error, and the prediction error reduction from 15 % to 70 % at the same manner.

Key words: artificial neural network, perceptron, linear regression network, FIR filter, predictor, voice signals, algorithm of noise reduction.

Ключевые слова: искусственные нейронные сети, персептрон, линейная регрессионная сеть, КИХ-фильтр, предсказатель, речевые сигналы, алгоритмы шумоподавления.

Введение

В настоящее время разработка новых алгоритмов обработки и передачи речевых сигналов (РС) в телекоммуникационных системах вызывает значительный интерес. При этом РС, с которыми приходится иметь дело на практике, всегда в той или иной степени подвержены действию акустических помех (АП). В тех случаях, когда АП имеют значительную интенсивность, их наличие может существенно исказить результаты обработки, анализа или распознавания речи. В целом ряде других случаев, например, при анализе зашумленных РС в криминалистических целях или восстановлении аудиозаписей в архивах, задача уменьшения негативного воздействия АП на РС носит самостоятельный характер и является единственной целью работы. Поэтому разработка алгоритмов снижающих влияние АП на РС является весьма актуальным направлением исследований.

Применение известных алгоритмов подавления АП при цифровой обработке РС, осуществляется в случае наличия параллельного канала при включенном на входе АЦП, который позволяет представить непрерывный поток данных в цифровом виде, для его дальнейшей обработки на ЭВМ. Данный факт является существенным техническим и экономическим недостатком применения таких алгоритмов [4]. Кроме того известно, что в настоящее время для микропроцессоров наступает так называемый «технологический предел», заключающийся в том, что они достигли максимального уровня повышения быстродействия.

Одним из способов решения данной проблемы может быть использование новой элементной базы, на-

Предложен алгоритм подавления акустических помех (АП) в кодерах речи, а так же показана возможность реализации предсказателя речевых сигналов (РС) в системах кодирования речи на основе искусственных нейронных сетей (ИНС). Удалось достигнуть значительного снижения уровня дисперсии акустического шума на выходе системы кодирования РС, при изменении отношения сигнал-шум от 7 до 23 дБ, без снижения субъективной оценки качества речи по шкале MOS. Показаны преимущества реализации предсказателей РС на базе ИНС по сравнению с известными предсказателями на основе КИХ-фильтров. Доказана возможность уменьшения порядка предсказания от 10 % до 60 % при той же ошибке и снижения ошибки предсказания от 15 % до 70 % при том же порядке в случае реализации предсказателей на основе ИНС. Достигнуто увеличение субъективной оценки качества РС по шкале MOS до 0,5 балла, что является существенным преимуществом систем кодирования на основе ИНС.

пример, на основе искусственных нейронных сетей (ИНС). Нейронные сети представляют собой весьма перспективную вычислительную технологию, дающую новые подходы к исследованию различных динамических задач. Свойство толерантности, присущее для ИНС, позволяет находить решения, робастные к различным видам искажений. Способность к моделированию нелинейных процессов к работе с зашумленными данными и адаптивность [5] дают возможность применению ИНС при решении широкого класса задач [6, 7].

Помимо обработки РС в усовершенствовании также нуждается процесс кодирования РС. Одним из основных способов первичного кодирования является алгоритм адаптивной дифференциальной импульсно-кодовой модуляции (АДИКМ). Для реализации линейного предсказателя в системах АДИКМ чаще всего используются нерекурсивные КИХ-фильтры, точность предсказания которых ограничена и определяется порядком фильтра и точностью оценки коэффициентов КИХ-фильтра b_k .

Для увеличения точности предсказания могут быть использованы ИНС, которые широко применяются в ситуациях, когда невозможно точно установить зависимость между входными и выходными данными, как например, при задаче предсказания отсчетов РС.

Таким образом, **целью работы является** разработка и исследование алгоритма подавления АП, а также анализ возможности реализации предсказателя РС в кодерах речи на базе ИНС.

Анализ статистических характеристик АП

Акустический сигнал, поступающий через микрофон на вход системы цифровой обработки, практически всегда содержит в себе не только РС, но и различного рода АП. Отрицательное влияние АП на РС проявляется в уменьшении разборчивости и ухудшении качественных характеристик речи. Особенно сильно данный эффект проявляется в цифровых системах обработки речи, так как приводит к дополнительным нелинейным искажениям РС.

При экспериментальных исследованиях анализировалось влияние АП на РС, создаваемых автотранспортом вблизи дороги, а также АП внутри автомобилей марок ВАЗ (Lada) 2112, ВАЗ (Lada) 2190, HyundaiSonata. Запись АП проводилась с частотой дискретизации 40 кГц. В дальнейшем, для корректного наложения данных АП на исходный РС, осуществлялась фильтрация записанных аудио файлов в полосе частот 0,3-3,4 кГц, и децимация отсчетов до частоты дискретизации 8 кГц.

Проводился анализ спектральных и статистических характеристик данных АП. Примеры частотных спектров АП представлены на рис. 1.

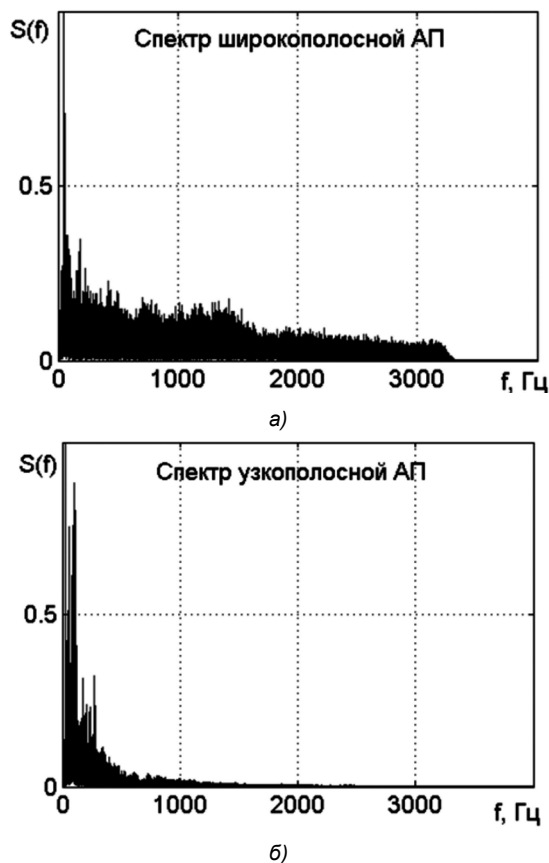


Рис. 1. Частотные спектры АП, используемых в экспериментальных исследованиях

Показано, что АП, создаваемые автотранспортом около дороги, являются широкополосными, так как их спектр относительно равномерно распределен в полосе частот 0,3-3,4 кГц, а АП внутри автомобиля – узкополосными, при этом их спектр расположен в полосе частот 0,3-1,5 кГц.

На рис. 2 приведены функции плотности вероятности (ФПВ) исследуемых АП.

Анализ полученных результатов по критерию согласия χ^2 показал, что данные ФПВ исследуемых АП могут быть аппроксимированы нормальным законом распределения.

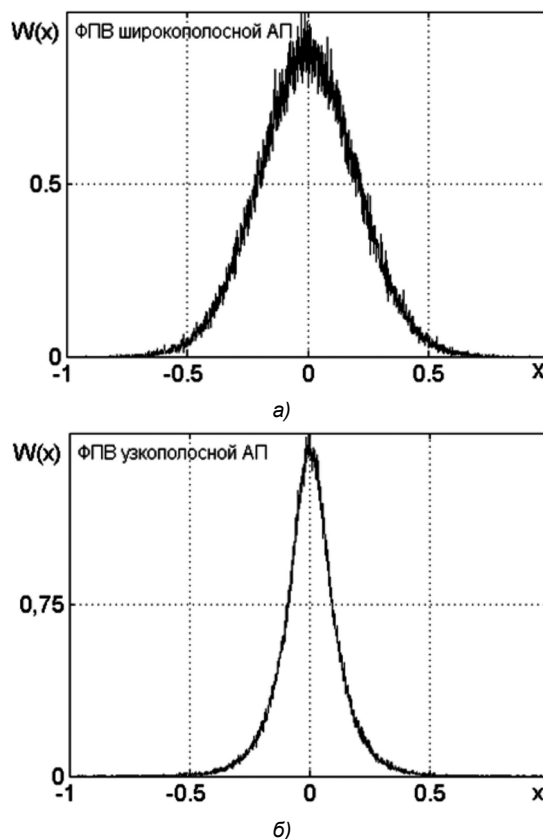


Рис. 2. ФПВ исследуемых АП

Структура проектируемой ИНС

Для решения какой-либо задачи с применением ИНС следует, прежде всего, спроектировать структуру сети, адекватную поставленной задаче. Для реализации алгоритма подавления АП была выбрана нейронная сеть типа «многослойный персептрон». Данный тип архитектуры нейронной сети является классической многослойной сетью с полными последовательными связями нейронов. При обучении ИНС, за счёт своего внутреннего строения, выявляются закономерности и связи входных и выходных образов. Таким образом, ИНС типа «многослойный персептрон» позволяет наиболее точно аппроксимировать выходные данные при обучении [12]. Обучение ИНС проводилось по методу Левенберга-Марквардта. Данный метод может быть представлен как комбинация методов наискорейшего спуска и Гаусса-Ньютона, которые являются примером способа быстрой оптимизации обучения [13]. Главными достоинствами данного алгоритма являются высокая скорость обучения и отсутствие необходимости в указании критериев остановки обу-

чения [14]. В процессе проектирования была проведена оптимизация структуры ИНС по критерию минимума среднеквадратического отклонения (СКО) обучения. При оптимизации происходило изменение количества скрытых слоев, количества нейронов в слоях и наклона сигмоидальной функции активации нейронов.

Оптимальная структура ИНС включала в себя:

- количество входов – 1;
- количество выходов – 1;
- количество скрытых слоев – 2;
- количество нейронов в первом скрытом слое – 10;
- количество нейронов во втором скрытом слое – 10;
- вид активационной функции – сигмоидальная.

Уровень подавления АП существенно зависит от параметров ИНС. В связи с этим критерием перестройки параметров ИНС было выбрано значение СКО АП – σ , что позволяло подавить примерно 68,2 % помех, при условии, что их плотность распределения вероятности хорошо аппроксимируется нормальным распределением. Таким образом, был введен порог подавления ИНС $p = \sigma$. Для определения значения порога подавления необходимо в паузах речи оценивать дисперсию АП.

Таким образом, структурная схема алгоритма подавления АП на основе ИНС имеет вид показанный на рис. 3.



Рис. 3. Структурная схема

нейросетевой реализации алгоритма подавления АП

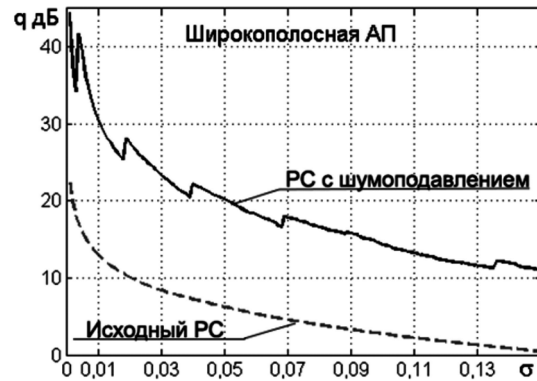
Для проведения экспериментального исследования нейросетевой реализации алгоритма подавления АП был использован РС содержащий в себе акустически взвешенные фразы представленные в ГОСТ Р 51061-97. АП накладывались на РС с различной дисперсией шума. При исследованиях было показано, что погрешность квантования обученной ИНС при отсутствии АП не превышала погрешности квантования стандартного кодера G.711. Исследовались зависимости ОСШ q зашумленного РС от уровня СКО АП σ . Для этого на исходный РС аддитивно накладывались различные АП с уровнем нормированного СКО от 0,01 до 0,12. Нормировка осуществлялась к СКО РС. После чего для сравнения зависимостей ОСШ от уровня СКО АП зашумленные данные РС пропускались через нейросетевую реализацию алгоритма шумоподавления, приведенную на рис. 3 и через стандартный кодер с μ -компандированием без шумоподавления, результаты сравнения приведены на рис. 4.

Из анализа рис. 4 следует, что нейросетевая реализация алгоритма шумоподавления позволяет увеличить ОСШ от 7 до 23 дБ.

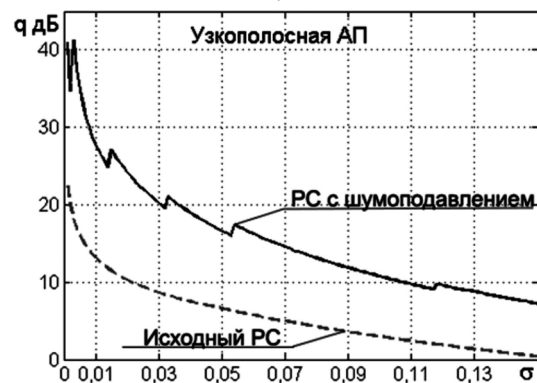
Соответствующие выбросы на рис. 4 связаны с дискретной перестройкой порога подавления АП. Можно отметить, что степень подавления АП не существенно

зависит от типа АП и его спектральных характеристик.

Для оценки дисперсии $D_{\text{дкш}}$ декодированного шума АП с различным уровнем дисперсии (от 0 до 0,01) проходили через схему, представленную на рис. 3. В данном случае АП не накладывались на РС для более точного определения его параметров. В результате экспериментальных исследований было отмечено снижение уровня дисперсии шума $D_{\text{ш}}$ на выходе проектируемого устройства по сравнению со стандартным кодером источника сообщения в 10 – 100 раз.



а)



б)

Рис. 4. Зависимости ОСШ декодированного сигнала при нейросетевой реализации кодера РС по сравнению со стандартным кодером

Для получения субъективной оценки качества речи по методу MOS [18] была произведена запись РС шестью дикторами, которые начитывали акустически взвешенные фразы, прописанные в ГОСТ Р 51061-97. После этого каждая АП аддитивно накладывалась на РС с различным ОСШ – от 0 до 40 дБ. Полученные РС кодировались стандартным кодером источника сообщения и с помощью нейросетевой реализации перестраиваемого кодера. Десять auditors производили субъективную оценку прослушиваемых РС по 5-ти бальной шкале [19]:

- 5 – понимание речи без малейшего напряжения внимания;
- 4 – понимание речи без затруднений;
- 3 – понимание речи с напряжением внимания без переспросов и повторений;
- 2 – понимание речи с некоторым напряжением внимания, редкими переспросами и повторениями;
- 1 – понимание речи с большим напряжением внимания, частыми переспросами и повторениями.

Оценки auditors складывались, а затем находилось среднее значение субъективной оценки по методу MOS.

В табл. 1 и табл. 2 приведены средние значения субъективных оценок по шкале MOS в зависимости от ОСШ.

Таблица 1. Среднее значение субъективных оценок для широкополосных АП

Среднее значение субъективных оценок для широкополосных АП	ОСШ, дБ				
	0	10	20	30	40
Стандартный кодер	1,1	2,3	3,2	4,2	4,8
Кодер на основе ИНС	1,2	2,6	3,8	4,3	4,8

Таблица 2. Среднее значение субъективных оценок для узкополосных АП

Среднее значение субъективных оценок для узкополосных АП	ОСШ, дБ				
	0	10	20	30	40
Стандартный кодер	1,1	2,1	3,1	3,9	4,6
Кодер на основе ИНС	1,1	2,3	3,4	4,2	4,6

Таким образом, удалось достигнуть увеличения субъективной оценки качества РС на 0,2-0,5 балла по шкале MOS.

Архитектуры ИНС, используемые в качестве предсказателей РС, и их анализ

Известно [8], что для задач регрессии рекомендуется использовать такие типы архитектур ИНС, как многослойный перцептрон и линейная регрессионная сеть.

Многослойный перцептрон является классической многослойной сетью с полными последовательными связями нейронов с сигмоидальной функцией активации.

Линейная регрессионная сеть является частным случаем перцептрона, где в качестве активационной функции используется линейная функция с ограничениями, определяемая минимальными и максимальными значениями массива выходных данных. Для обеспечения динамики ИНС между скрытыми слоями добавлены обратные связи, позволяющие более точно устанавливать зависимости между входными и выходными данными [9].

Анализ ИНС типа многослойный перцептрон

Схема обучения ИНС типа многослойный перцептрон приведена на рис. 5.

В качестве входного массива данных был использован акустически взвешенные РС с частотой дискретизации 8 кГц, соответствующие ГОСТ Р 51061-97.



Рис. 5. Схема обучения ИНС типа многослойный перцептрон

В качестве выходного массива данных были выбраны выходные значения фильтра предсказателя соответствующего порядка. Обучение многослойного перцептрона проводилось по алгоритму Левенберга-Марквардта и обратного распространения ошибок, а также квази-Ньютоновским методом обратного распространения. Зависимости среднеквадратичной ошибки предска-

зания σ от порядка фильтра N изображены на рис. 6.



Рис. 6. Зависимости среднеквадратичной ошибки предсказания от порядка предсказателя для ИНС типа многослойный перцептрон

Анализ полученных зависимостей показал, что лучшие результаты достигаются при обучении по методу обратного распространения.

На рис. 7 изображена схема эксперимента.

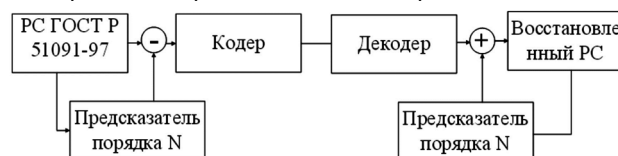


Рис. 7. Схема эксперимента

Эксперимент соответствует неадаптивному методу кодирования РС ДИКМ без передачи коэффициентов предсказания. Рассматривались два варианта реализации предсказателя на приемной стороне:

- с обучением ИНС как на передающей, так и на принимающей стороне, при этом значения векторов смещения и весов ИНС-предсказателей на приемной и передающей сторонах различны;
- с обучением ИНС только на передающей стороне, значения векторов смещения и весов ИНС-предсказателя на приемной стороне полностью соответствуют значениям векторов смещения и весов ИНС-предсказателя на передающей.

На рис. 8 изображены зависимости среднеквадратичной ошибки предсказания σ от порядка фильтра N при наличии и отсутствии обучения ИНС-предсказателя на приемной стороне.

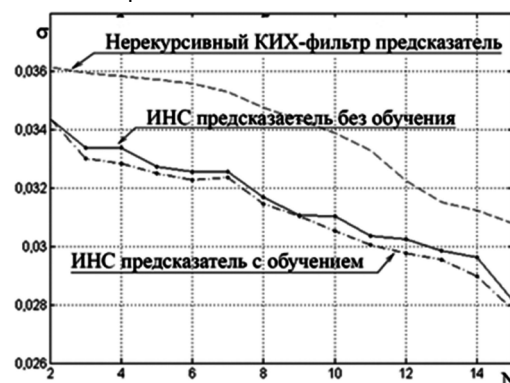


Рис. 8. Зависимости среднеквадратичной ошибки предсказания от порядка предсказателя для ИНС типа многослойный перцептрон

Из анализа рис. 8 следует, что существенных разли-

чий между схемами с обучением и без обучения ИНС-предсказателя на приемной стороне не наблюдается. Таким образом, можно использовать схему без обучения на приемной стороне, чтобы не усложнять нейросетевую реализацию предсказателя. Также очевидно преимущество предсказателя, спроектированного на базе ИНС типа многослойный персептрон, перед известным нерекурсивным КИХ-фильтром предсказателем, что позволяет уменьшить порядок предсказателя от 20 % до 50 % при той же ошибке предсказания.

Анализ ИНС типа линейная регрессионная сеть

В отличие от многослойного персептрона линейная регрессионная сеть не нуждается в обучении, так как самоорганизуется и адаптируется в зависимости от исходных входных и выходных массивов данных. Таким образом, данная сеть может быть использована для реализации адаптивного предсказателя. Эксперимент проводился в соответствии с рис. 7, при этом адаптация КИХ-фильтра осуществлялась по алгоритму Левенсона-Дарбина. В данном случае для реализации схемы эксперимента с помощью предсказателя на основе н КИХ-фильтра необходимо создать низкоскоростной канал для передачи коэффициентов предсказания. В связи с этим было рассмотрено две реализации организации предсказателя на базе линейно-регрессионной ИНС:

- с передачей коэффициентов весов и векторов смещения ИНС после адаптации на передающей стороне;

- с адаптацией ИНС на приемной стороне.

Зависимости среднеквадратичной ошибки предсказания от порядка предсказателя для линейной регрессионной ИНС приведены на рис. 9. Из анализа рис. 9 можно сделать вывод, что адаптивный ИНС-предсказатель обеспечивает более высокие результаты по сравнению с известным предсказателем на основе КИХ-фильтра, при этом не требуя организации дополнительного низкоскоростного канала для передачи коэффициентов предсказания, что является его существенным преимуществом.

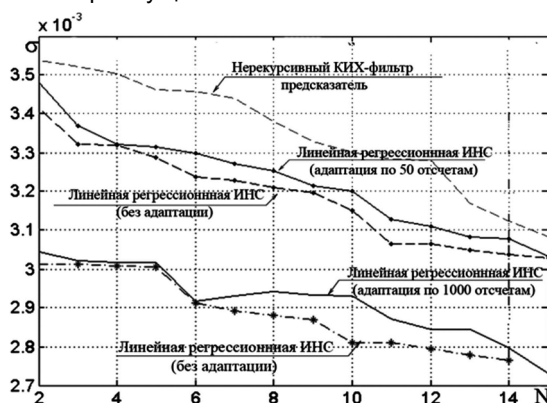


Рис. 9. Зависимости среднеквадратичной ошибки предсказания от порядка предсказателя для линейной регрессионной ИНС

На рис. 10 приведены зависимости среднеквадратичной ошибки предсказания от порядка предсказателя для 50, 150, 500 и 1000 отсчетов адаптации РС.

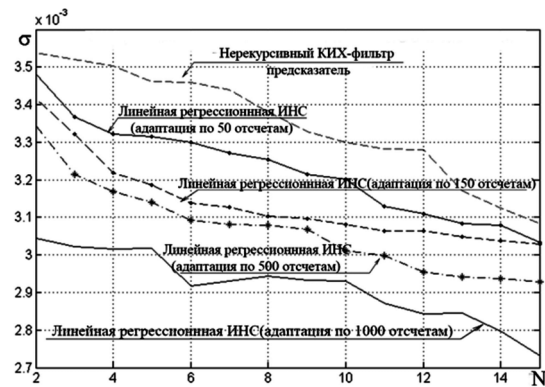


Рис. 10. Зависимости среднеквадратичной ошибки предсказания от порядка предсказателя для линейной регрессионной ИНС

Из анализа зависимостей следует, что ошибка предсказания линейной регрессионной сети существенно зависит от количества отсчетов, по которым происходит адаптация. Наилучшие результаты достигнуты при адаптации по 1000 отсчетам, что при частоте дискретизации 8 кГц составляет 125 мс. Таким образом, можно сделать вывод, что линейная регрессионная сеть может быть применена в качестве адаптивного предсказателя. Помимо этого, ИНС данной архитектуры при условии адаптации на приёмной стороне позволяет снизить ошибку предсказания от 15 % до 70 % при том же порядке или уменьшить порядок предсказания на 10 %–60 % при той же ошибке.

Анализ показателей качества речи при нейросетевых реализаций предсказателей в кодерах РС

Далеко не всегда качество восстановленного РС может быть охарактеризовано значениями СКО отсчетов, в связи с этим необходимо проводить анализ субъективной оценки качества речи.

Для получения субъективной оценки качества РС по методу MOS [18] были произведены записи речи шести дикторов, которые начитывали акустически взвешенные фразы, прописанные в ГОСТ Р 51061-97. Полученные РС кодировались стандартным кодером источника сообщения с предсказателем на основе КИХ-фильтра, а также с помощью нейросетевой реализации фильтра-предсказателя. Десять аудиторов производили субъективную оценку прослушиваемых РС по 5 бальной шкале [19].



Рис. 11. Зависимости оценки качества речи по шкале MOS от порядка предсказателя

Оценки аудиторов складывались, а затем находилось среднее значение субъективной оценки по методу

MOS. На рис. 11 и рис. 12 изображены зависимости субъективной оценки качества речи по шкале MOS от порядка предсказателя.

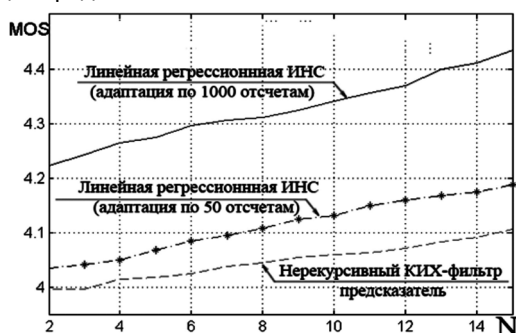


Рис. 12. Зависимости оценки качества речи по шкале MOS от порядка предсказателя

Из анализа зависимостей видно, что предсказатель на базе ИНС позволяет увеличить оценку качества речи по шкале MOS до 0,5 балла.

Заключение

Проведенные исследования показали эффективность применения нейросетевой реализации алгоритма подавления АП. Удалось достигнуть снижения уровня дисперсии помех на выходе кодера в десятки раз, по сравнению со стандартным кодером; при этом увеличилось отношение сигнал-помеха на 7 - 23 дБ. Однако субъективная оценка качества речи по шкале MOS не снижалась. Отмечено преимущество нейросетевых предсказателей перед известным предсказателем на основе КИХ-фильтра. При реализации неадаптивного предсказателя на базе персептрона возможно уменьшить порядок предсказателя на 20% - 50% при той же ошибке предсказания, а применение в качестве адаптивного предсказателя ИНС линейно-регрессионной архитектуры позволяет снизить ошибку предсказания от 15 % до 70 % при том же порядке или уменьшить порядок предсказания от 10 % до 60 % при той же ошибке предсказания. Также отмечено увеличение субъективной оценки качества речи по шкале MOS до 0,5 балла, что является существенным преимуществом данных систем.

Литература

1. Gibak Kim, Phillips C Loizou. Why do speech-enhancement algorithms not improve speech intelligibility // Processing of ICASSP-2010. Vol. 1. – P. 397–400.
2. Phillips C Loizou. Speech enhancement theory and practice: 1st ed. Boca Raton, FL.: CRC, 2007. Releases Taylor & Francis.
3. Benesty J., Chen J., Huang Y., Cohen I. Noise Reduction in Speech Processing // Springer-Verlag, 2009.
4. Злобин В.К., Григоренко Д.В., Ручкин В.Н., Романчук В.А. Кластеризация и восстанавливаемость нейросетевых систем обработки данных // Известия ТулГУ. Технические науки. 2013. Вып. 9. Ч.2 – С. 125-133.
5. Данилин С.Н., Макаров М.В., Щаников С.А. Проектирование технических средств с нейросетевой архитектурой при искажении шумами входной информации. 24-я Международная Крымская конференция «СВЧ-техника и телекоммуникационные технологии»: материалы конф.: в 2 т. – Севастополь, 2014.

тирование технических средств с нейросетевой архитектурой при искажении шумами входной информации. 24-я Международная Крымская конференция «СВЧ-техника и телекоммуникационные технологии»: материалы конф.: в 2 т. – Севастополь, 2014.

6. Кириллов С.Н., Сазонова Е.С. Нейросетевая реализация предсказателя речевых сигналов в кодерах речи // Вестник РГРТУ № 56 – 2016 – 31-37 с.

7. Кириллов С.Н., Попова Е.С. Нейросетевая реализация кодера речевых сигналов адаптивного к уровню акустических шумов. Вестник РГРТУ, 2015 № 54, – 40-44 с.

8. Болодурина В.Н. Решетников М.Г. Таспаева. Применение и адаптация нейросетевых технологий в задаче идентификации динамических объектов. Программные продукты, системы и алгоритмы № 1, 2012 г.

9. Ciresan D., Meier U., Masci J and Schmidhuber J. Multi-column Deep Neural Network for Traffic Sign Classification. Neural Networks, Vol. 34, August 2012, pp. 333 – 338 И.П.

10. Карасев О.Е. Применение теории нечетких множеств для обработки видеoinформации в телекоммуникационных системах. VI Всероссийские научные Зворыкинские чтения: сб. тез. докл. Всероссийской межвузовской научной конференции. Муром.-Муром: Изд.-полиграфический центр МИ ВлГУ, 2014. – 791 с.

11. Еремин Д.М., Гарцев И.Б. Искусственные нейронные сети в интеллектуальных системах управления. – М.: МИРЭА, 2004. – 75 с.

12. Ясницкий Л.Н. Введение в искусственный интеллект. – М.: Издат. центр «Академия», 2005. – 176 с. – ISBN 5-7695-1958-4.

13. Martinetz M., Berkovich S., Schulten K. "Neural-gas" network for vector quantization and its application to time series prediction H Trans. Neural Networks, 1993. – Vol. 4.

14. Gill P. Murray W., Wright M. Practical Optimization. – N.Y.: Academic Press, 1987.

15. Sohn J., Kim N.S., and Sung W. A statistical model-based voice activity detection. IEEE Signal Processing Lett., 6 (1): 1–3, 1999.

16. Ephraim, Y. & Malah, D. Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator IEEE Trans Acoustics Speech and Signal Processing, 32(6):1109-1121, Dec 1984

17. Rainer Martin. Noise power spectral density estimation based on optimal smoothing and minimum statistics. IEEE Trans. Speech and Audio Processing, 9(5):504-512, July 2001.

18. Рекомендация МСЭ-Т Р.80/Р.800

19. ГОСТ Р 51061-97 Системы низкоскоростной передачи речи по цифровым каналам. Параметры качества речи и методы измерений. – Введ. 01.01.98. – М.: Госстандарт России

20. Строгонов А. Использование нейронных сетей для прогнозирования деградации выходных параметров ТТЛ ИС в системе MATLAB/Simulink // Компоненты и технологии. №1. 2006.

21. Дьяконов В.П., Абраменкова И.В., Круглов В.В. MATLAB 5.3.1 с пакетами расширений. М.: Нолидж. 2001.

О НОВОМ ЭТАПЕ РАЗВИТИЯ ОПТИМИЗАЦИОННОЙ ТЕОРИИ КОДИРОВАНИЯ

Золотарёв В.В., д.т.н., профессор кафедры ВПМ Рязанского государственного радиотехнического университета и Института космических исследований РАН, e-mail: zolotasd@yandex.ru.

ABOUT NEW STAGE OF DEVELOPING OF OPTIMIZING CODING THEORY

Zolotarev V.V.

The results obtained during the 25-year evolution of the error-correction coding optimization theory (OT) and multithreshold decoding (MTD) methods, which have been created on its basis, are presented. Research results on MTDs and other error-correction methods for binary and non-binary codes used to send messages over channels with binary, symbolic errors and erasures are presented. It is shown MTDs simply decode very long codes, which are the only ones capable of supporting the effective implementation of error correction at high channel noise levels. The methodological basis of the OT and the new paradigms for successful research into the theory and applied issues of error-correction coding are discussed. General conclusions are formulated on the study, and directions for further development of work on MTD are suggested.

Key words: communication systems, data storage systems, error-correction coding, computer simulation, self-orthogonal codes, concatenated codes, multithreshold decoders, coding gain, communication channel.

Ключевые слова: телекоммуникационные системы, системы хранения данных, помехоустойчивое кодирование, компьютерное моделирование, самоортогональные коды, каскадные коды, многопороговые декодеры, энергетический выигрыш кодирования, канал связи.

Введение

В 2015 году исполнилось 25 лет со дня защиты диссертации [1], в которой для очень простых по сегодняшним меркам кодов были доказаны многие основные результаты, которые позднее были систематизированы и представлены в полном объеме в Оптимизационной Теории помехоустойчивого кодирования [2-6, 10-13, 16-21]. Она развивалась на основе идей мажоритарного декодирования [8]. Оптимизационная Теория (ОТ) создала новые технологии итеративной коррекции ошибок декодирования, базовые методы которой были запатентованы ещё в 1972 г. [25]. Сейчас именно ОТ и только её парадигмы определяют практически все улучшения в характеристиках декодирования для классических моделей каналов.

В настоящее время все основные этапы создания, исследования и проектирования многопороговых декодеров (МПД) проводятся на основе специальных мощных оптимизационных процедур, эффективность и сложность которых быстро растут. При этом сложность самих методов МПД, разработанных на основе ОТ, остаётся минимальной, по-прежнему растущей всего лишь линейно с длиной кода. Но эффективность декодирования на базе МПД оказывается совпадающей с возможностями переборных, т.е. наилучших возможных методов даже для больших уровней искажений в канале связи. Отметим в связи с этим, что в современных сложных научных изысканиях невозможно переоценить значение теорий оптимизации как таковых. Уже весьма давно, например, считается общеизвестной истиной [27], что роль оптимизационных теорий в математике

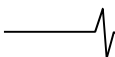
Рассмотрены новые результаты основных кластеров параметров алгоритмов декодирования, построенных на основе оптимизационной теории (ОТ) кодирования. Обсуждаются главные характеристики многопороговых декодеров для двоичных гауссовских, недвоичных, стирающих и других каналов. Указано на большие возможности блочного алгоритма Витерби. Подчеркиваются перспективы дивергентного кодирования. Отмечается группа алгоритмов прямого контроля метрики. Обсуждается ценность основных парадигм ОТ. Определяются основные пути развития теории кодирования на ближайшие годы и на перспективу.

столь же велика, как и роль математики вообще в науках.

Благодаря применению ОТ с увеличением числа итераций коррекции ошибок и появлением кодов со всё меньшей степенью подверженности размножению ошибок (РО) возможности МПД существенно растут при сохранении весьма небольшой сложности самого алгоритма. К настоящему времени характеристики многопороговых декодеров по сравнению с другими методами при всех практически интересных для техники связи параметрах систем уже стали существенно лучше [2-6]. Ниже сначала рассмотрены возможности МПД, соответствующих главным кодовым кластерам (типичным наборам параметров кодов и моделей каналов) [10-13, 22-26, 28-36], а затем состоится обсуждение представленных результатов.

Гауссовские каналы

Рассмотрим характеристики основных алгоритмов декодирования в гауссовском канале при кодовой скорости $R=1/2$, представленные на рис. 1. На нем показаны в традиционном виде зависимости вероятности ошибки на бит $P_b(e)$ различных алгоритмов декодирования как функции от уровня битовой энергетике канала E_b/N_0 в децибелах. Вертикаль $C=1/2$ отмечает уровень шума, при котором пропускная способность гауссовского кана-



ла равна кодовой скорости $C=R/1/2$. Пунктир P_0 отмечает вероятность ошибки при отсутствии кодирования. Граница АТ указывает на предельные возможности турбо кодов, которые, однако, до сих пор не могут быть воплощены в характеристики аппаратуры из-за сложности алгоритмов этого класса. Кривая VA:K7 отражает эффективность повсеместно применяемого алгоритма Витерби (AB) для свёрточных кодов с длиной кода $K=7$, а зависимость CC:VA*RS соответствует каскадной схеме на основе AB и кода Рида-Соломона [9, 38]. Кривая LDPC приведена для min-sum декодера кода с низкой плотностью проверок на четность (LDPC) стандарта DVB-S2 длиной 64800 битов, реализованного в НИИ Радио [46]. График TR представляет реальные возможности декодера для турбо кода длиной 3060 битов стандарта CDMA2000.

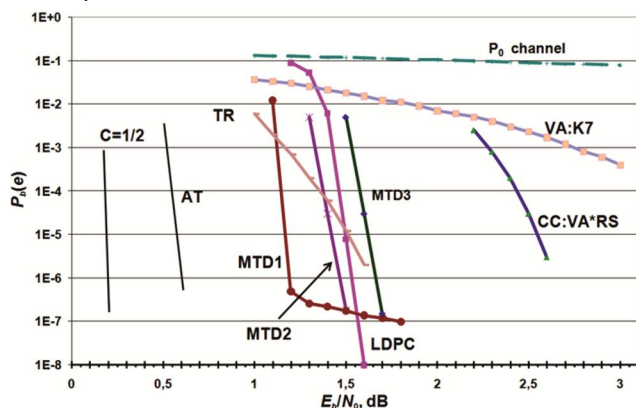


Рис. 1. Характеристики основных алгоритмов декодирования в канале с аддитивным белым гауссовским шумом (АБГШ) при кодовой скорости $R=1/2$

МПД алгоритм в свёрточном варианте реализации в двоичном гауссовском канале и 4-х битовом квантовании сигнала в демодуляторе показан на рис. 1 на графике MTD1. В настоящее время он практически столь же оптимально, как и переборные алгоритмы, декодирует длинные коды при очень низкой энергетике гауссовского канала $E_b/N_0=1,2$ дБ, когда до его пропускной способности C оказывается всего лишь около 1 дБ [3-4, 28]. Таким образом, МПД1 работает при таком уровне шума, когда мощность передатчика только на ~26 %, т.е. лишь на четверть превышает её уровень при $C=1/2$. Для работы декодера требуется не более $I=192$ итераций. Величина задержки декодирования для построенного кода с малым РО и с кодовым расстоянием $d=21$ составляет менее 8 Мбит кодовых символов. Объём собранной статистики для всех точек этого графика превышает $2,3 \cdot 10^9$ битов.

Далее, при незначительном снижении уровня шума до $E_b/N_0=1,5$ дБ для МПД уже нужны лишь $I=90$ итераций и задержка решения свёрточного декодера около 1 Мбит, как это показывает график MTD2. А при $E_b/N_0=1,8$ дБ обычный МПД декодер с 40 итерациями (кривая MTD3) оказывается лучше несравненно более сложной последовательной каскадной схемы алгоритма Витерби (AB) с декодером кода Рида-Соломона (РС), при всего лишь втрое большей задержке. Не в последнюю очередь надо помнить, что эта каскадная схема имеет ещё и на ~0,6 дБ (на 12,5 %) меньшую кодовую

скорость R , чем у рассматриваемых МПД алгоритмов, что ещё более наглядно показывает преимущества алгоритмов, созданных в соответствии с ОТ.

Детальное сравнение МПД декодеров и других основных алгоритмов для гауссовских каналов, которые относительно успешно развиваются исследователями кодирования у нас и в зарубежье, показывает, что к характеристикам МПД при $E_b/N_0 \sim 1,2$ дБ за последнее десятилетие даже не приблизились никакие алгоритмы типа LDPC, турбо или какие-либо ещё методы с сопоставимой сложностью декодирования. Реальные их характеристики за последнее десятилетие так и не преодолели условную границу энергетике $E_b/N_0 \sim 1,5$ дБ даже с использованием каскадных схем, что при такой близости к границе Шеннона является огромной разницей по сравнению с первым графиком для МПД. В настоящее время нет никаких указаний на то, что эта граница будет преодолена при разумном уровне сложности декодирования. Отметим здесь только, что и структура любых схем кодирования на базе свёрточных МПД декодеров при аппаратной реализации всегда гораздо проще, чем у других методов как на передающей, так и, особенно, на приёмной стороне.

Далее, МПД алгоритмы могут быть реализованы аппаратно так, что они становятся как бы однократной решающей схемой мгновенного действия. А это приводит к тому, что, как и все алгоритмы, МПД декодеры создают задержку решения, но совершенно не снижают скорости работы любого устройства, в составе которого они работают. В итоге получаем, что представленные выше реализации МПД алгоритмов обеспечивает в аппаратуре любую произвольно высокую производительность, что принципиально невозможно для других методов. Это происходит согласно [16, 21, 28] так, словно кодовая последовательность просто поступает в декодер и без какой-либо обработки с той же предельно высокой скоростью обычного сдвига данных в чипе, с какой она была введена в декодер, выходит из него, но уже почти без ошибок в информационной части принятого сообщения. Кроме того, и структура связей между ячейками в аппаратном МПД декодере много проще, чем у прочих алгоритмов. Он более, чем на 99 % состоит из простейшей памяти на длинных регистрах сдвига, что дополнительно облегчает его создание и отладку [38]. Такое свойство свёрточного МПД естественно назвать максимальной аппаратной теоретической производительностью. Поэтому, даже просто сопоставимого с алгоритмами МПД быстроедействие другие алгоритмы коррекции достичь не могут в принципе. Но при этом МПД декодеры, как было показано выше, ещё и работают при таком уровне шума, при котором неработоспособны прочие методы. А так как эта область уже очень близка к пропускной способности канала типа АБГШ, то разница с другими методами, составляющая несколько десятых децибел, оказывается принципиальной и на самом деле очень значимой, непреодолимой, определяющей и перспективы развития алгоритмов МПД в будущем. Наконец, отметим, что и абсолютная разница в числе итераций в пользу МПД там, где прочие алгоритмы всё-таки ещё работают, также весьма значительна. Причём, сложность каждой

итерации у декодеров МПД, в которых лишь суммируются небольшие целые числа, тоже существенно меньше. А в случае программной реализации методов МПД за счёт хранения вычисленных сумм на пороговом элементе (ПЭ) каждая итерация вообще превращается в простую проверку значений этих сумм, на что обычно требуется не более 1-2 операций сравнения [2-4]. Совокупность перечисленных преимуществ алгоритмов МПД в двоичных каналах с АБГШ перед прочими методами столь значима, что МПД декодеры как продукт ОТ к настоящему моменту стали абсолютными лидерами в мировом конкурсе прикладных достижений в одном из главных кластеров разработок систем кодирования как по совокупности параметров сложности, быстродействия и помехоустойчивости, так и по всем им в отдельности.

Указанное соотношение по эффективности и сложности реализации между различными декодерами полностью подтверждается и результатами работы демо-программ для различных декодеров, представленных на наших уникальных сетевых порталах [29], которые полностью соответствуют всем рассмотренным выше соотношениям свойств между алгоритмами. В частности, например, МПД алгоритм в программной демоверсии при весьма высоком уровне шума исправляет все ошибки до оптимального (переборного!) уровня всего за 10 итераций на ПК с процессором Core-i7 на тактовой частоте ~3 ГГц на скорости более 16 Мбит/с! И пока что не просматривается никаких путей для каких-либо других алгоритмов коррекции ошибок, чтобы хотя бы приблизиться по своим характеристикам простоты реализации, быстродействия и энергетической эффективности к возможностям МПД.

Символьные коды

В высшей степени ценным для ОТ, теории кодирования и различных прикладных цифровых систем передачи, обработки, хранения и восстановления данных стало открытие в 1984 году и к настоящему времени уже полное исследование нашей научной школой символьных кодов [1-4, 6, 12, 14-15, 18, 21, 22, 38-40, 47], реализация символьного многопорогового декодирования (QMПД) которых также чрезвычайно проста, как и в случае двоичных их аналогов.

Строго говоря, впервые эти недвоичные коды с мажоритарным декодированием рассмотрел Дж. Мессси (J. Massey), который доказал теоремы 1-4 для них в [8]. Но затем он очень негативно оценил возможности таких кодов в разделах 1.2, 6.2, 6.5, 6.6 и 8.2 этой же книги и больше уже никогда не занимался этой темой. При этом нам до сих пор неизвестны другие сколько-нибудь содержательные работы по мажоритарному декодированию недвоичных кодов, а тем более публикации по итеративным алгоритмам для них.

Рассмотрим возможности недвоичных кодов. На рис. 2 представлены характеристики декодеров кодов Рида-Соломона (РС) и QМПД при кодовой скорости $R=1/2$. По горизонтальной оси отложены вероятности ошибки на символ при различных размерах алфавита q , $q=2^8=256$ и $q=2^{16}=65536$. По вертикальной оси отклады-

ваются вероятности ошибки декодера на символ $P_d(s)$ для любых q . График P_0 показывает вероятность ошибки в симметричном недвоичном канале q СК. Кривая RS2⁸ даёт представление о возможностях кода РС с (n, k, d) параметрами (255, 128, 128), в котором размер символа соответствует одному байту. Далее пунктир RS-Su2⁸ соответствует нижней оценке возможностей довольно сложного декодера для этого же кода, предложенного Суданом [37]. График RS2¹⁶ относится к коду РС длины 65535, который ещё долго будет считаться очень сложным в реализации.

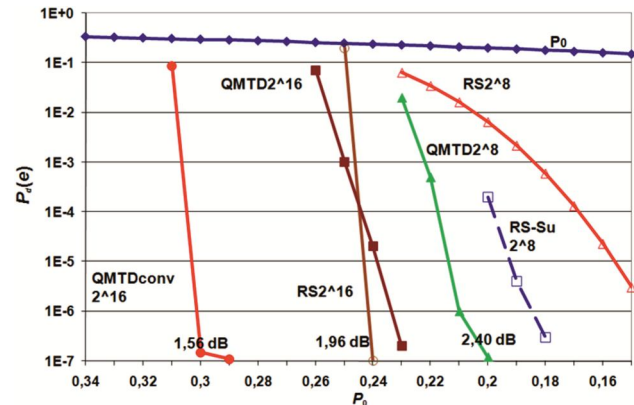


Рис. 2. Характеристики основных алгоритмов декодирования в q -ичном симметричном канале при кодовой скорости $R=1/2$

К настоящему времени для символьных кодов получены все основные оценки характеристик декодирования и детально проработана их общая теория как для блочных, так и для свёрточных кодов. Они могут полностью заменить коды РС во всех возможных приложениях, выигрывая у них и у других потенциальных конкурентов все конкурсы по достоверности и быстродействию [12, 38-40]. Причиной этого преимущества являются возможность выбора любой длины этих символьных кодов, которая совершенно не зависит от алфавита выбранного кода q . И, самое главное, символьные коды, как и двоичные, также обеспечивают фактически оптимальное декодирование (эквивалентное переборному!) при использовании простейших мажоритарных методов даже при достаточно большом уровне шума. Это оказалось возможным благодаря очень точной и удачной модификации недвоичного ПЭ по сравнению с двоичным [3-6, 12, 18, 19, 22, 28, 29, 38-40, 47]. При этом сложность символьного МПД (QMПД), как и в двоичном случае, оказывается теоретически минимальной, линейной от длины кода.

Полезность символьных кодов становится особенно важной и впечатляюще значимой, если вспомнить, что для недвоичных кодов фактически вообще невозможно создать сколько-нибудь работоспособный алгоритм Витерби, сложность которого в большинстве случаев будет порядка $\sim q^K$, где q – размер алфавита, K – длина кода в символах. Например, для кода с $R=1/2$ уже при $q \geq 32$ и $K \geq 7$ число отслеживаемых в таком АВ путей превысит миллиард, а характеристики будут слабее, чем у кода Рида-Соломона (РС).

Вспомним теперь, что за 50 лет развития недвоичных кодов РС, которые надолго и вполне заслуженно имели широкое поле приложений, ничего реально эффектив-

нее их для двоичных алфавитов придумано не было. Но реально применяемые коды РС для каскадных схем, оптических дисков и прочих систем относятся к коротким кодам и поэтому малоэффективны. Методы Судана [37, 47], позволяющие несколько снизить вероятность ошибки в кодах РС по сравнению со стандартными алгоритмами, также не решили проблему малой эффективности этих кодов даже за счёт весьма значительного усложнения процесса коррекции ошибок.

С другой стороны, большое число публикаций, диссертаций, в том числе и докторских, а также демопрограммы для символьных МПД свидетельствуют, что длинные символьные коды с простейшей реализацией и быстродействием до десятков Мбит/с эффективно работают при вероятностях ошибок, кратно превосходящих уровень допустимых вероятностей для кодов РС [1-4, 6, 14, 15, 28, 29].

В качестве важнейшего на текущий момент прикладного результата, который имеет серьёзную идеологическую поддержку со стороны ОТ, можно указать на недавнее достижение символьными свёрточными МПД с кодовой скоростью $R=1/2$ для $q=2^{16}$ особо высокой помехоустойчивости даже в q СК канале с вероятностью ошибки $P_0=0,3$, как показано на рис. 2 в виде графика QMTDconv2¹⁶ [47, 53]. Этот уровень шума недостижим при $R=1/2$ даже для кодов РС в поле $GF(2^{16})$, реализовать декодер для которых исключительно сложно, а из-за невысоких характеристик декодирования и бесполезно. Можно также указать в связи с этим на то, что скорости работы символьных МПД действительно очень велики, и благодаря крайне простой идее их реализации на самом деле могут достигать очень высоких значений [12, 14, 15, 18, 21, 22]. График QMTD2⁸ для символьного кода длины 8000 байтов при $q=256$ эффективнее короткого кода РС с декодером Судана [37]. Сложность программной версии QMTD2⁸ (для того же процессора) определяется скоростью декодирования, равной ~300 Кс/с (Кс – Килосимволы), которая, конечно, огромна. Менее чем за 1 час набирается статистика на более чем миллиард символов, что может составить до $3 \cdot 10^{10}$ битов. Читателям доступна также демопрограмма QMTD для малоизбыточного символьного кода с $R=0,95$ с простой инструкцией по настройке параметров и использованию [29]. Она работает на том же процессоре Core-i7 при очень высоком уровне шума для этого значения R со скоростью порядка 40 Мбит/с.

Символьные МПД на много лет вперёд полностью решили все проблемы цифрового мира по передаче и, особенно, хранению и восстановлению цифровых данных в любых информационных системах в диапазоне кодовых скоростей $R=0,5-0,98$.

Стирающие каналы

Обратимся ещё к одной важной области теории кодирования, которая создаёт и изучает декодеры для каналов со стираниями. Однако до применения к ней методов ОТ и алгоритмов МПД результаты для этих каналов были крайне скромны у всех методов [1, 2, 4, 8, 30, 31, 47]. Основной причиной этого было, конечно, гораздо большее внимание специалистов именно к ка-

налам с ошибками. Собственно, именно поэтому исследования декодеров для стирающих каналов были как бы в тени, а их характеристики весьма далеки от допускаемых теорией пределов эффективной работы этих алгоритмов.

Ограничимся только кратким рассмотрением тех характеристик МПД, которые показаны на рис. 3. Из последних результатов в этой области следует, что для восстановления стираний методами МПД [1-4, 30, 53, 54] при $R=1/2$ необходим свёрточный код с минимальным кодовым расстоянием $d \geq 21$ и особенно малым группированием своих неправильных решений при декодировании, т.е. с минимальной подверженностью эффекту РО. Соответствующий МПД алгоритм с таким кодом, представленный графиком MTD05, успешно восстанавливает поток стёртых (любых!) символов с вероятностью независимых стираний $P_{es}=0,48$ в канале до уровня вероятности оставшихся невосстановленных символов $P_s(s) \sim 3 \cdot 10^{-7}$. Таким образом, в канале с пропускной способностью $C=0,52$ МПД алгоритм успешно работает при отношении $R/C \sim 0,961$, что является абсолютно уникальным достижением для процедур восстановления стираний. При этом скорость работы соответствующей МПД демопрограммы при $I=90$ итерациях коррекции была равна 95 Кс/с. Другие быстрые методы восстановления стираний при $R=1/2$ для таких уровней искажений на входе декодера неизвестны. График MTD45 соответствует свёрточному коду с $R=4/5$, с которым МПД при $P_{es} \leq 0,19$ успешно работал при $R/C \approx 0,988$ [47]. Такой результат также крайне труден для повторения какими-либо другими алгоритмами коррекции стираний. Графики MTD05-s и MTD45-s для гораздо более простых алгоритмов этого типа также соответствуют той области искажений символов в канале, в которой декодеры для кодов РС совершенно неработоспособны [47, 53]. Основной вывод по этому разделу состоит в том, что и в стирающих каналах алгоритмы МПД тоже не имеют равных по простоте реализации и эффективности декодирования.

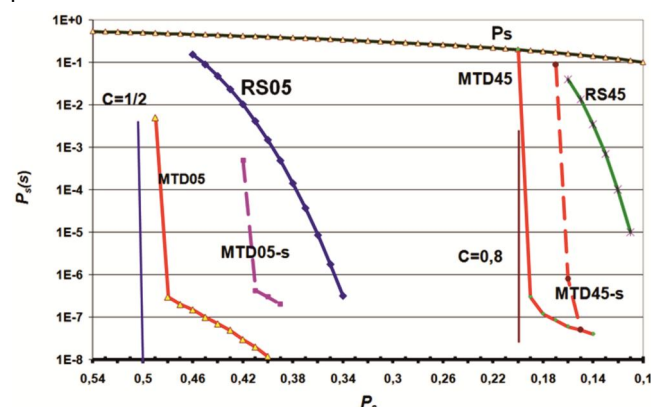


Рис. 3. Характеристики МПД и кодов РС в каналах со стираниями

Специальные приложения ОТ

Отметим кратко другие результаты ОТ, которые оказались важными для теории кодирования и её приложений.

Во-первых, это масштабные исследования различ-

ных высокоскоростных алгоритмов МПД при кодовой скорости $R \sim 0,8$ и выше. Все свойства МПД и их соотношение с возможностями других алгоритмов, сохраняются и в этой области кодовых скоростей [1-4, 13-15, 22, 28, 38-40].

Например, в [3-5, 35, 36] рассмотрены методы высокоскоростного МПД декодирования при $R=4/5$ для оптических каналов при гауссовской модели канала, полученные нашей научной школой и зарубежными авторами, которых мы консультировали. Результаты российских исследователей были существенно лучше по энергетической эффективности и простоте реализации МПД алгоритмов. С учётом теоретически максимально возможной скорости декодирования МПД наши методы имеют существенные преимущества.

Применение методов МПД декодирования оказывается особенно полезным при передаче и хранении очень больших массивов данных, в том числе и на борту спутников систем дистанционного зондирования Земли (ДЗЗ). Вероятность нарастающих со временем искажений в таких системах памяти может быть весьма высокой, а необходимая скорость передачи этих данных с орбиты часто также очень велика, что приводит к необходимости использования там только МПД алгоритмов. Рабочий макет простого МПД декодера на ПЛИС Altera для спутникового канала на скорость более 1 Гбит/с, созданный в ИКИ РАН, представлен на рис. 4. Подобный ему чип с МПД декодером был ранее разработан в НИИ Радио. Такие характеристики высокоскоростного декодирования для других алгоритмов, видимо, будут ещё очень долго недостижимы.

Многопороговый декодер (МПД) для спутниковых и космических каналов связи, повышает кпд их использования в 3 -10 раз, в том числе для ДЗЗ.

МАКЕТ МПД для каналов на 2,08 Гбит/с

The multithreshold decoder (MTD) for satellite and Space channels, raises efficiency of their usage in 3-10 times, including channels up to 2Gb/s

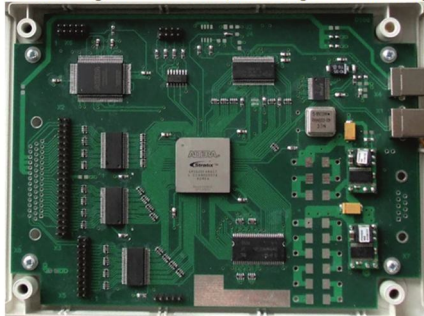


Рис. 4. Действующий макет МПД декодера на ПЛИС Altera для гауссовского канала на информационную скорость 1 Гбит/с

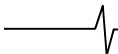
Далее можно указать на совсем небольшую сложность реализации МПД алгоритмов в канале типа ДСК, с которых и начались много лет назад исследования МПД [51]. В качестве наглядного примера высокой эффективности МПД в этой области укажем на важнейшую фундаментальную и одновременно сложнейшую техническую задачу декодирования для флеш памяти. Специфика этой проблемы состоит в том, что вероятность ошибки на бит для таких систем должна быть не хуже $P_b(e) \sim 10^{-15}$, а в скором времени требования к достоверности решений соответствующих декодеров ещё более усилятся. В исходных массивах такой памяти вероят-

ность ненадёжного хранения символов может быть на уровне $p_e \sim 10^{-3}$ и даже существенно большей. Найденные специалистами нашей научной школы технические решения на основе МПД алгоритмов позволили довольно просто решить эту проблему [4, 21, 32, 47]. При этом использовался код с МПД декодированием при $R=3/4$, что гарантированно позволяет обеспечить требуемые уровни достоверности даже при вероятности ненадёжного хранения $p_e \sim 10^{-2}$. Столь же простые и одновременно эффективные результаты по сверхдостоверному декодированию в условиях применения малонадёжной памяти у других разработчиков нам неизвестны.

Основные ресурсы ОТ

Ограничимся кратким обсуждением тех наиболее важных теоретических, математических и технологических ресурсов, которые образуют пространство, совокупность знаний, понятий и методов решения задач, называемых обычно парадигмами той или иной отрасли науки, в нашем случае - теории кодирования.

Отметим сначала главные парадигмы (результаты, методы и технологии) классической теории кодирования, длительное время связываемой с алгебраической теорией. Конечно же, она имела гораздо более широкое поле развития. Однако даже после появления алгоритма Витерби (AB) основные направления развития классической теории включали как важнейший аспект развития разработку алгебраических методов декодирования (в дискретных полях). Но хотя классы новых кодов множились, а декодеры понемногу становились более простыми, долгое время исправление ошибок за пределами половины кодового расстояния было для алгебраических кодов неразрешимой задачей. Поэтому наилучшим научным результатом, широко применяемым в технике декодирования, уже 40 лет является каскадная схема АВ с кодами РС. Хорошими примерами эвристического подхода, выросшего из классической теории, стали в последние десятилетия турбо и низкоплотностные (LDPC) коды, которые позволили немного приблизиться к пропускной способности канала. Однако с началом нового тысячелетия это движение фактически остановилось, хотя возможности элементной базы и теории увеличивались. Дело в том, что эти методы излишне сложны и, например, многие вопросы их применения для сверточных кодов и больших скоростей всё ещё не решены. Поэтому поиски новых методов в рамках алгебраической теории продолжаются. Одним из промежуточных итогов этого поиска стали полярные коды, к которым сразу обратилось большое число специалистов по кодам [45]. Но до сих пор реальные результаты разработки таких декодеров представлены очень невнятно, а иногда и весьма странно. Анализ этого направления по имеющимся работам был нами сделан в [7]. Более содержательные данные по этим кодам, если они появятся, позволят сделать более точные прогнозы по их перспективам. А сейчас весьма слабые результаты для полярных кодов компенсируются обращением к декодированию списками и созданию крайне сложных каскадных схем для этих кодов, что пока полностью исключает их из группы перспективных алгоритмов. Не лишне напомнить, что ана-



огичные подходы «переключения интереса» несколько ранее вынуждены были использовать специалисты по алгебраическим методам и последовательным алгоритмам, возможности которых даже теоретически были очень далеки от границы Шеннона. Однако значительно улучшить эти методы не удалось. Весьма интересное направление 80-х годов [41] тоже не получило тогда хорошего импульса для своего развития. Возможно, это было связано с тем, что это направление было в большей степени сориентировано на алгебраические коды со слабыми характеристиками. Возможно, последним единственным успешным результатом алгебраической теории следует считать алгоритмы Судана [37], которые немного улучшили эффективность декодирования кодов РС и послужили развитию дискретной математики. К сожалению, эти методы тоже оказались весьма сложными, а их реальная эффективность декодирования выросла незначительно. Этими подходами и исчерпывается пока реальный прогресс классической теории, хотя предложенный здесь комментарий, конечно, должен быть гораздо более детальным.

Определим важнейшие парадигмы Оптимизационной Теории, которые позволяют сейчас постоянно получать принципиально новые результаты практически для всех групп кодовых кластеров, некоторые основные типы которых мы рассматривали выше.

ОТ началась около 45 лет назад с появлением метода особого итеративного декодирования на основе мажоритарных процедур [25, 26] и Основной Теоремы многопорогового декодирования (ОТМТД), философскую целеполагающую роль которой невозможно переоценить [1-6]. Этот метод сразу позволил в двоичных симметричных каналах (ДСК) при использовании как блочных, так и свёрточных кодов исправлять число ошибок в несколько раз большее, чем половина кодового расстояния [2-4]. При этом, что было абсолютно неожиданным для специалистов, ОТМТД гарантировала при каждом изменении декодируемых символов строгое приближение к оптимальному решению, которое обычно достигается только переборными методами, известным элегантным примером которых является АВ для свёрточных кодов. Дальнейшим развитием ОТ стало распространение возможностей МПД на все прочие каналы с независимыми искажениями.

Вторым фундаментальным результатом ОТ стала всесторонняя разработка теории эффекта размножения ошибок (РО), без которой значимость ОТМТД была бы весьма ограниченной. А совместное синергетическое воздействие этих двух фундаментальных теоретических результатов на эффективность МПД оказалась столь огромным, что при линейной от длины кода сложности этим алгоритмом сразу оказалось возможным исправлять на фактически оптимальном уровне эффективности почти все ошибки в блочных и свёрточных кодах длины $n=10^4$ - 10^5 битов и более. Напомним, что реальная длина кода для переборного АВ даже в будущем, видимо, не будет более 30 битов, чем и определяются все преимущества МПД, который остался алгоритмом с теоретически минимальной сложностью.

Дальнейший прогресс ОТ поддерживался созданием

нескольких поколений высокопроизводительного программного обеспечения для разработки и исследований различных типов МПД и средств поиска кодов с малым уровнем РО, что позволило постепенно приближать уровень практически оптимального (эквивалентного переборному!) декодирования на основе МПД к границе Шеннона.

Следующими очень мощными методами улучшения характеристик МПД для всех каналов стали открытые нами ещё в 1986 г. параллельные методы каскадирования, технологии улучшенного способа назначения порогов, весов проверок и подбора кодов, а также методы ускоренного вычисления сумм на пороговых элементах [2-4]. Особое влияние на идеологию ОТ оказало выделение в особую группу декодеров с прямым контролем метрик (ДПКМ) методов МПД и АВ, которые точно измеряют расстояние до принятого сообщения, что и позволяет сейчас только этим алгоритмам успешно работать при больших уровнях шума. Другие методы с такими важнейшими свойствами не известны.

Для сохранения простоты реализации алгоритмов МПД мы также пересмотрели принципы каскадирования, известные нам, и нашли очень простые решения, которые без усложнения этих декодеров многократно повышают эффективность каскадных схем [1-5, 14, 15]. Но эта тема заслуживает отдельных публикаций.

Принципиальное значение, решающим образом меняющее соотношение свойств эффективности и сложности между различными алгоритмами, имеет открытие нашей научной школой блочных АВ (БАВ) со сложностью, как и в случае свёрточных кодов, порядка 2^K , K – длина кода, тогда как до недавнего времени считалось, что сложность таких методов имеет порядок, более близкий к 2^{2K} [48]. БАВ также относится к группе ДПКМ, что очень сильно уменьшает потенциальные перспективы полярных кодов, методов Чейза и других алгоритмов для блочных кодов как в гауссовских каналах, так и в ДСК [17, 43, 44, 52]. Очевидно, что большое число проблем эффективного декодирования блочных кодов успешно решают простые каскадные схемы с использованием БАВ.

Ещё больший простор в разработке методов декодирования на основе ОТ открылся после точной формулировки принципов дивергентного декодирования [24, 29, 42, 49, 50]. Этот подход чётко определил отдельный класс простых некаскадных методов увеличения кодового расстояния используемых кодов, который дополнительно приблизил рабочие характеристики алгоритмов ОТ к пропускной способности канала С. Они использовались и в МПД алгоритмах, представленных на рис. 1-3, что особенно сильно улучшило работу МПД при экстремально высоких уровнях шума. Важно, что при этом можно одновременно использовать несколько различных методов коррекции ошибок.

ОТ успешно осваивается мировым научно-техническим сообществом. Наши крупнейшие в России сетевые порталы по ОТ и МПД алгоритмам www.mtdbest.ru (РГРТУ) и www.mtdbest.iki.rssi.ru (ИКИ РАН) посещают за год до 100'000 тысяч читателей из более 80 стран мира и переписывают около 20 Гб информации по тео-

рии и технологиям нашей научной школы [50], включая множество демопрограмм по всем известным лучшим алгоритмам декодирования. Их можно переписать на свои ПК и сразу непосредственно заняться исследованиями по технике декодирования согласно прилагаемым к ним инструкциям. Такое масштабное внимание мировой науки к разработкам российской научной школы является уникальным и свидетельствует о серьёзных достижениях нашей теории кодирования.

Рассмотренные методы, свойства и возможности кодов и алгоритмов в рамках ОТ, как видно из их описания, все без исключения позволяющие улучшить характеристики декодирования простейшими способами, образуют очень мощное перспективное интеллектуальное пространство развития методов декодирования. Эти новые парадигмы ОТ уже позволили создать методы для разных кодовых кластеров, которые эффективно работают непосредственно вблизи границы Шеннона. Развитие всех этих методов быстро переносится на новые кодовые структуры.

Опыт общения нашей научной школы с коллегами позволил взглянуть на ОТ ещё с одной стороны. Большинство из них справедливо указывает на то, что все модификации МПД и наши различные версии АВ можно выделить как новый особый образ мышления и как отдельное совершенно неожиданное открытие заново всей теории кодирования, для которой сейчас наступила новая эра развития. Но в такой же степени к таким открытиям следует отнести и отдельно: ОТПД, теорию РО, символьные коды, параллельное каскадирование, а также блоковые АВ и дивергентное кодирование, всестороннее развитие которых позволит в самом ближайшем будущем ещё более упростить алгоритмы декодирования и обеспечить их эффективность в новых кодовых схемах.

Заключение

В данной работе нашей научной школой представлено достижение для основных типов каналов непосредственной близости границы Шеннона на основе довольно простых алгоритмов для основных классов моделей каналов. Как в физике достижение скорости света материальными телами невозможно, так и совсем небольшое расстояние (никогда не достижение!) рабочей области алгоритмов МПД до пропускной способности S рассматривавшихся каналов, что следует из графиков на рис. 1-3, является особенно наглядным свидетельством состоятельности ОТ. А огромное преимущество этих методов по эффективности и одновременно по сложности реализации перед прочими алгоритмами коррекции ошибок позволяет утверждать, что поставленная 70 лет назад Шенноном проблема решена, причём, и на вполне приемлемом технологическом уровне, что подтверждает и большое число патентов по алгоритмам ОТ и МПД, лишь частично представленных в этой статье [16-21, 25, 28, 54].

Важность, успешность и перспективность ОТ можно сравнить только с появлением квантовой механики в самом начале прошлого века. Но тогда физика как наука не столь много значила в жизни людей. Тем не менее,

результаты Планка, Шредингера и целой плеяды других великих физиков были высоко оценены научным сообществом, и они заслуженно стали Нобелевскими лауреатами. Достижение, фактически в одиночку, аналогичных по масштабу результатов нашей научной школой, конечно, имеет значительно большее значение, так как мы решили важнейшую и чрезвычайно сложную проблему простого достижения высокой достоверности цифровых потоков в период взрывного развития нашей информационной цифровой цивилизации, создав все возможности для применения наших простых и понятных методов кодирования в цифровой технике и науке будущего.

Таким образом, изложенные в статье данные свидетельствуют об успешном решении главной научной и технологической проблемы всей нашей информационной цифровой цивилизации – создании обширных классов простых методов достижения произвольно высокой достоверности передачи, хранения и восстановления цифровой информации. Решение этой сложнейшей проблемы позволяет переключить усилия инженеров и учёных на решение новых важных задач нашего цифрового мира.

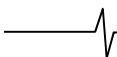
Наши благодарности

Исследования ОТ в течение прошедших 45 лет её эволюции с момента первых публикаций нашей научной школы поддерживали: МФТИ, ИППИ РАН, концерн «Созвездие», Совет по кибернетике АН СССР, ОНТИ РАН, РГРТУ, НИИ Радио, МНИТИ, ИКИ РАН. Методы МПД тестировали ООО «ОРТ», НПО им. С.А. Лавочкина, а также ряд организаций и предприятий отрасли связи. Ценность финансовой поддержки разработки методов МПД со стороны РАН, РГРТУ, РФФИ и других источников также была исключительно велика (гранты РФФИ14-07-00859, 14-07-00824, 08-07-00078 и 05-07-00024). Мы полагаем, что число наших последователей и сторонников, как и раньше, будет расти, а новые сферы исследований и разработок с использованием парадигм нашей «квантовой теории» в области помехоустойчивого кодирования, т.е. ОТ, технологий МПД и новых видов АВ декодирования и каскадных схем будет быстро расширяться.

Большинство ссылок, относящихся к этой статье, как и многие другие материалы об ОТ и МПД алгоритмам, в том числе множество демопрограмм, можно найти на наших сетевых порталах [29].

Литература

1. Золотарёв В.В. Субоптимальные алгоритмы многопорогового декодирования. Докторская диссертация. М., 1990, 278 с.
 2. Золотарёв В.В. Теория и алгоритмы многопорогового декодирования. М.: Радио и связь, Горячая линия – Телеком, 2006, второе издание - 2014.
 3. Золотарёв В.В., Зубарев Ю.Б., Овечкин Г.В. Многопороговые декодеры и оптимизационная теория кодирования // Под редакцией академика РАН В.К. Левина. М.: Горячая линия – Телеком, 2012, 238 с.
 4. Zolotarev V., Zubarev Y., Ovechkin G. Optimization Coding Theory and Multithreshold Algorithms. Geneva, ITU, 2015, 159 p.
- (e-book reference: <http://www.itu.int/pub/S-GEN-OCTMA-2015>).



5. Золотарёв В.В., Овечкин Г.В. Применение многопороговых методов декодирования помехоустойчивых кодов в высокоскоростных системах передачи данных // Электросвязь. М., 2014, № 12, с. 10-14.
6. Кузнецов Н.А., Золотарёв В.В., Овечкин Г.В., Овечкин П.В. Недвоичные многопороговые декодеры и другие методы коррекции ошибок в символьной информации // Радиотехника. М., 2010, № 6, Вып.141, с. 4–9.
7. Золотарёв В.В., Овечкин Г.В. О сопоставлении новых методов помехоустойчивого кодирования // 18 Международная конференция «Цифровая обработка сигналов и её применение – DSPA 2016». М., 2016, Т.1, с. 59-65.
8. Месси Дж. Пороговое декодирование // Пер. с англ. Ю.Л. Сагаловича под ред. Э.Л. Блоха. М.: Мир, 1966. 208 с.
9. Кларк Дж., Кэйн Дж. Кодирование с исправлением ошибок в системах цифровой связи. М.: Радио и связь, 1987, 391 с.
10. Золотарёв В.В., Назиров Р.Р., Никифоров А.В., Чулков И.В. Новые возможности многопорогового декодирования по высокодостоверной передаче данных ДЗЗ // Современные проблемы дистанционного зондирования земли из космоса. Сборник научных статей. М.: ООО «Азбука-2000», 2009, Вып.6, Т.1. с.167-173.
11. Золотарёв В.В., Назиров Р.Р., Чулков И.В., Овечкин Г.В. Алгоритмы МПД // Российский космос. М., 2009, № 1, с. 60-63.
12. Золотарёв В.В., Овечкин Г.В., Назиров Р.Р., Овечкин П.В., Чулков И.В. Эффективное недвоичное многопороговое декодирование помехоустойчивых кодов для систем дистанционного зондирования Земли // Современные проблемы дистанционного зондирования Земли из космоса, Сборник статей. М., 2010, Т.7, №2, с. 269-274.
13. Золотарёв В.В., Чулков И.В. Малоизбыточное кодирование для высокоскоростных каналов // 11-я Всероссийская открытая конференция «Современные проблемы дистанционного зондирования Земли из космоса». М., 2013.
14. Овечкин Г.В. Теория каскадного декодирования линейных кодов для цифровых радиоканалов на основе многопороговых алгоритмов. Докторская диссертация. Рязань, РГРТУ, 2011, 301 с.
15. Овечкин П.В. Разработка алгоритмов повышения эффективности недвоичных многопороговых декодеров в системах передачи и хранения больших объемов информации. Кандидатская диссертация. Рязань, РГРТУ, 2009, 131 с.
16. Золотарёв В.В., Овечкин Г.В., Сатыбалдина Д.Ж., Ташатов Н.Н., Адамова А.Д. Способ мягкого многопорогового декодирования помехоустойчивого кода. Удостоверение автора (патент Республики Казахстан) №93989 от 15.10.2014.
17. Золотарёв В.В., Овечкин П.В. Способ кодирования и декодирования блочного кода с использованием алгоритма Витерби. Патент на изобретение РФ №2608872 от 25.01.2017.
18. Золотарёв В.В., Овечкин Г.В. Способ работы символьного порогового элемента в символьном мажоритарном декодере. Патент на изобретение РФ № 2573741 от 22.12.2015.
19. Золотарёв В.В.. Способ декодирования помехоустойчивого кода. Патент на изобретение РФ №2557454 от 25.06.2015.
20. Золотарёв В.В., Овечкин Г.В. Устройство многопорогового декодирования линейных кодов для гауссовских каналов. Патент на полезную модель №44215 от 27.02.2005.
21. Золотарёв В.В.. Высокоскоростное устройство многопорогового декодирования линейных кодов. Патент на полезную модель №44216 от 27.02.2005.
22. Zolotarev V.V., Averin S.V. Non-Binary Multithreshold Decoders with Almost Optimal Performance // 9-th ISCTA'07. UK, Ambleside, 2007.
23. Averin S.V., Ovechkin G.V., Zolotarev V.V. Algorithm of Multithreshold Decoding for Self-Orthogonal Codes over Gaussian Channels // 11-th ISCTA'09. UK, Ambleside, 2009.
24. Золотарёв В.В., Овечкин Г.В. Дивергентное кодирование сверточных кодов // 18-я Международная научно-техническая конференция «Проблемы передачи и обработки информации в сетях и системах телекоммуникаций». Рязань, 2015, с.27-32.
25. Золотарёв В.В. Устройство для декодирования линейных сверточных кодов. А.с. СССР № 492878 от 25.11.1975 с приоритетом от 31.07.1972.
26. Самойленко С.И., Давыдов А.А., Золотарёв В.В., Третьякова Е.Л. Вычислительные сети. М.: Наука, 1981, 278 с.
27. Стрекаловский А.С. Частное сообщение, 2016.
28. Золотарёв В.В. Способ декодирования помехоустойчивого кода. Патент на изобретение РФ №2377722 от 27.12.2009.
29. Ресурсы www.mtdbest.ru и www.mtdbest.iki.rssi.ru.
30. Золотарёв В.В. Многопороговое декодирование в стирающих каналах // Вопросы радиоэлектроники. Серия ЭВТ. М., 1983, Вып.10, с.67-70.
31. Гринченко Н.Н., Золотарёв В.В., Овечкин Г.В., Овечкин П.В. Применение многопорогового декодера в каналах со стираниями // Труды НТОРЭС им. А.С. Попова, 2006. с. 338-340.
32. Овечкин Г.В., Золотарев В.В., Федиев В.С. Повышение достоверности хранения цифровых данных на флеш-памяти // 6-я Международная научно-техническая конференция «Космонавтика. Радиоэлектроника. Геоинформатика». Рязань, 2013. с.201-203.
33. Овечкин Г.В., Као В.Т. Многопороговые декодеры для гауссовских каналов // 19-я Всероссийская научно-техн. конф. «Новые информационные технологии в научных исследованиях и в образовании». Рязань, РГРТУ, 2014, с. 121-122.
34. Гринченко Н.Н., Као В.Т., Овечкин Г.В. Повышение эффективности многопорогового декодера // 17-я Международная конференция «Цифровая обработка сигналов и ее применение – DSPA 2015». М., 2015.
35. Zolotarev V., Ovechkin G., Satybaldina D., Tashatov N., Adamova A., Mishin V. Effective multithreshold decoder for optical and other data transmission systems // Latest trends on Communications: Proceedings of the 18th International Conference on Communications (part of CSCC'14). Santorini Island, Greece, 2014, pp.152-156.
36. Zolotarev V., Ovechkin G., Satybaldina D., Tashatov N., Adamova A., Mishin V. Efficiency multithreshold decoders for

self-orthogonal block codes for optical channels // International Journal of Circuits, Systems and Signal Processing. 2014, vol. 8, pp. 487-495.

37. Sudan M. Decoding of Reed Solomon codes beyond the error-correction bound // Journal of Complexity. 1997, vol.13, pp.180-193.

38. Золотарёв В.В., Овечкин Г.В. Помехоустойчивое кодирование. Методы и алгоритмы. Справочник. М.: Горячая линия – Телеком, 2004, 126 с.

39. Золотарёв В.В. Многопороговое декодирование в неавтономных каналах // Вопросы радиоэлектроники, серия ЭВТ. 1984, Вып. 12, с.73-76.

40. Золотарёв В.В. Алгоритмы коррекции символьных данных в вычислительных сетях // В сб.: «Вопросы кибернетики», ВК-105, АН СССР, Научный совет по комплексной проблеме «Кибернетика». М., 1985, с. 54-62.

41. Блох Э.Л., Зяблов В.В. Обобщённые каскадные коды. М.: Связь, 1976.

42. Золотарёв В.В., Овечкин Г.В., Ташатов Н.Н. Применение принципа дивергенции при декодировании сверточных кодов // III Международная научно-практическая конференция «Информационная безопасность в свете Стратегии Казахстан-2050». 2015, с.158-164.

43. Золотарёв В.В., Назиров Р.Р. Блоковая модификация алгоритма Витерби // 11-я Всероссийская открытая конференция «Современные проблемы дистанционного зондирования Земли из космоса», М., 2013.

44. Золотарёв В.В., Овечкин П.В. Характеристики декодирования блочных кодов по алгоритму Витерби для систем ДЗЗ // 13-я Всероссийская открытая конференция «Современные проблемы дистанционного зондирования Земли из космоса». М., 2015.

45. Arikan E. Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels // IEEE Transactions on Information

Theory. 2009, vol.55, no.7, pp.3051-3073.

46. Овечкин Г.В., Чикин А.В. Помехоустойчивость приемника спутниковых сигналов DVBS2 // 11-я межд. конф. и выст. «Цифровая обработка сигналов и ее применение – DSPA 2009». М., 2009, с. 578-580.

47. Зубарев Ю.Б., Золотарёв В.В., Овечкин Г.В., Овечкин П.В. Оптимизационная теория кодирования: итоги 25 лет развития // 18-я Международная конференция «Цифровая обработка сигналов и её применение – DSPA 2016», Пленарный доклад. М., 2016, Т.1, с.6-12.

48. Кудряшов Б.Д. Основы теории кодирования. СПб.: БХВ-Санкт-Петербург, 2016, 393с.

49. Золотарёв В.В. Применение дивергентного кодирования в каналах спутниковой связи и ДЗЗ // 13-я Всероссийская открытая конференция «Современные проблемы дистанционного зондирования Земли из космоса». М., 2015.

50. Зубарев Ю.Б., Золотарёв В.В., Овечкин Г.В., Овечкин П.В. Итоги 25-летнего развития оптимизационной теории кодирования // Научноёмкие технологии. М., 2016, Т.17, с. 26-32.

51. Золотарёв В.В. Многопороговое декодирование // Проблемы передачи информации. М., 1986, Т. XX11, Вып. 1, с. 104-109.

52. Золотарёв В.В., Овечкин Г.В., Овечкин П.В. Блоковая модификация алгоритма Витерби // Вестник РГРТУ. Рязань, 2017, №59, с.30-35.

53. Золотарёв В.В., Овечкин Г.В., Баринов И.В. Применение самоортогональных кодов в каскадных схемах кодирования для каналов связи со стираниями // 19 Межд. конф. «Цифровая обработка сигналов и её применение – DSPA 2017». М., 2017, Т.1, с. 75-79.

54. Золотарёв В.В. Способ обнаружения и исправления стираний при приёме дискретной информации. Патент на изобретение РФ №2611235 от 21.02.2017.

ВЫЧИСЛИТЕЛЬНЫЕ АЛГОРИТМЫ ЦИФРОВОЙ ФИЛЬТРАЦИИ И АНАЛИЗ ИХ ЭФФЕКТИВНОСТИ

Кошелев В.И., заведующий кафедрой РТС РГПУ, д.т.н., профессор, e-mail: koshelev.v.i@rsreu.ru.

COMPUTATIONAL ALGORITHMS OF DIGITAL FILTRATION AND ANALYSIS OF ITS EFFECTIVENESS

Koshelev V.I.

The problem of organizing computations in digital signal processing systems taking into account the accumulation of sampling errors is solved, and it makes it possible to significantly improve the accuracy of calculations both in real-time systems and in analyzing the efficiency of filtering processes whose correlation matrix is poorly conditioned. Two calculation schemes are proposed, using methods for sorting data arrays before they are summed. The aim of the paper is to analyze the accuracy of calculations, to develop sorting algorithms and to estimate their computational costs.

Key words: accumulation of discretization errors, the eigenvalues and the conditionality of the correlation matrix of the process, summable methods for sorting arrays.

Ключевые слова: накопление ошибок дискретизации, число обусловленности корреляционной матрицы процесса, методы сортировки суммируемых массивов.

Введение

Во многих задачах цифровой фильтрации больших массивов данных ввиду накопления ошибок дискретизации данных по амплитуде важное значение имеет последовательность выполнения вычислений. Вначале рассмотрим анализ эффективности систем первичной межпериодной обработки радиолокационных сигналов с точки зрения порядка выполнения арифметических операций и его влияния на точность получаемого решения [1].

Проанализируем точность расчета эффективности алгоритмов межпериодной обработки сигналов на фоне некоррелированных помех. Покажем, что при относительно большой размерности решаемой задачи не всегда целесообразно проводить анализ эффективности по известным из [2, 3] расчетным соотношениям, т.к., при представлении чисел в ЭВМ ограниченным количеством разрядов, точность решения оказывается в ряде случаев недостаточной для решения поставленной задачи. В рассматриваемом классе задач причиной этого является плохая обусловленность матриц корреляций, используемых в расчетах в качестве ядра квадратичной формы.

В связи с этим в статье ставится актуальная задача построения вычислительных алгоритмов, отличающихся от прямых методов расчета более высокой точностью решения при приемлемых с практической точки зрения затратах машинного времени. В целом рациональное распределение вычислительных ресурсов может обеспечивать в одних ситуациях сокращение времени вычислений и объема занимаемой памяти ЭВМ, а в других – повышение точности типовых вычислительных процессов, характерных для анализа систем фильтрации радиолокационных сигналов [3]. Кроме того, модификация типовых вычислительных алгоритмов филь-

Решается задача организации вычислений в цифровых системах обработки сигналов, учитывающей накопление ошибок дискретизации и позволяющая значительно повысить точность вычислений как в системах реального времени, так и при анализе эффективности фильтрации процессов корреляционная матрица которых плохо обусловлена. Предложены две схемы вычислений, использующие методы сортировки массивов данных перед их суммированием.

трации радиолокационных и навигационных сигналов [4-6] может обеспечить повышение точности при обработке сигнала в реальном масштабе времени.

Решение задачи

Как известно [7], эффективность системы оптимальной обработки при гауссовской статистике полезного сигнала и коррелированной помехи наиболее полно характеризуется вероятностями правильного D и ложного F обнаружения. При флуктуирующем сигнале вероятностные характеристики однозначно связаны с коэффициентом улучшения сигнал-(помеха+шум) γ формулой $D = F^{1/(1+\gamma Q)}$, где Q – пороговое отношение сигнал-(помеха+шум). В свою очередь величина

$$\gamma = \left| \mathbf{s}^H \mathbf{w} \right|^2, \quad (1)$$

где $\mathbf{w} = (\mathbf{R}_c + \lambda \mathbf{E})^{-1} \mathbf{s}$ – вектор обработки с элементами

$$w_j = \sum_{i=1}^n r_{ji} e^{-i\varphi_i}, \quad r_{ji} \text{ – элементы матрицы } (\mathbf{R}_c + \lambda \mathbf{E})^{-1}.$$

Используем другую форму записи выражения (1):

$$\gamma = \left| \mathbf{s}^H \mathbf{w} \right|^2 / \mathbf{w}^H (\mathbf{R}_c + \lambda \mathbf{E}) \mathbf{w},$$

которая в случае произвольной корреляционной матрицы сигнала имеет вид

$$\gamma = \mathbf{w}^H \mathbf{R}_s \mathbf{w} / \mathbf{w}^H (\mathbf{R}_c + \lambda \mathbf{E}) \mathbf{w}. \quad (2)$$

Вычисление γ при заданных параметрах сводится к расчету отношения двух квадратичных форм. Для определения потенциальной эффективности обработки необходимо найти экстремум (максимум) выражения (2).

Пусть минимальное собственное значение матрицы $\mathbf{R}_c$ равно μ . Тогда с увеличением размерности задачи при неизменном характере корреляции и $\mu \leq \lambda$ имеется возможность рассчитать выражение (2) по упрощенной формуле, сократив непроизводительные затраты машинного времени. При $\lambda < \mu$ происходит частичная или полная потеря точности решения [8]. Причем для заданной величины μ величина n , при которой справедливы данные рассуждения, оказывается связанной с длиной разрядной сетки операндов ЭВМ. Аналогичная связь имеет место для параметра μ при заданном n . Как показано в [1], соотношение точности представления операндов в ЭВМ и число обусловленности играет определяющую роль в числе правильно вычисленных десятичных знаков результата вычислений линейных систем уравнений.

Рассмотрим упрощенный расчет эффективности систем фильтрации при $\lambda \geq \mu$. Пусть при минимальном собственном значении матрицы $\mathbf{R}_c$, равном $\mu = \alpha$, максимальное собственное значение матрицы $\mathbf{R}_s$ $\beta \leq \mu$. Тогда на основании результатов [8]

$$\mu \leq m / (\alpha + \lambda), \quad (3)$$

а если кроме того $\alpha \ll \lambda$, то $\mu \leq n / \lambda$, где величина n / λ определяет максимально достижимый коэффициент улучшения отношения сигнал-(помеха+шум), а вероятность правильного обнаружения для флуктуирующих сигналов определяется как

$$D = F^{1+nQ/\lambda}.$$

Наиболее близкое приближение к границе параметра μ (или D) наблюдается в задаче радиолокации при оптимальной скорости цели, чему соответствуют значения доплеровского сдвига фазы сигнала за период повторения импульсов $\varphi_s = j2\pi \pm \pi$, при $j = 0, 1, 2, \dots$

На рис. 1 приведены предельные зависимости $\chi(n)$ при различных λ , позволяющие без громоздких расчетов ориентировочно оценить необходимый объем обрабатываемой выборки, при которой достигается заданная эффективность. Пунктиром обозначены зависимости предельной эффективности от объема обрабатываемой пачки, а сплошными кривыми – аналогичные зависимости при различных значениях $dF_c T$ относительной ширины гауссовского спектра флуктуаций помехи. При увеличении уровня некоррелированного шума расширяется диапазон изменения n , в пределах которого сравниваемые характеристики сливаются друг с другом. Так, при $\lambda = 10^{-6}$ они сближаются при $n > 6$, а при $\lambda = 10^{-4}$ – уже при $n \geq 5$, что позволяет использовать упрощенный метод расчета для ориентировочных расчетов.

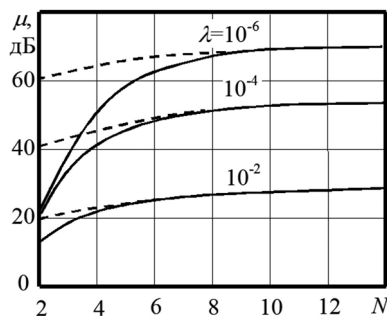


Рис. 1 Предельные зависимости $\mu(N)$ при различных λ

Более сложным и практически важным является вопрос обеспечения необходимой точности расчета качественных показателей фильтрации при $\lambda < \mu$.

Как известно [9], операции умножения и сложения величин с ограниченным числом разрядов на ЭВМ не являются коммутативными, ассоциативными и не связаны между собой законами дистрибутивности.

При расчете эффективности обычно пользуются формулой (2). При вычислении на ЭВМ числителя $\mathbf{w}^H \mathbf{R}_s \mathbf{w}$, имеющего порядок величины n , обычно не возникает проблем обеспечения точности решения, т.к. при этом вычисляется величина, близкая максимальному собственному числу матрицы $\mathbf{R}_s$. Напротив, при вычислении знаменателя формулы (2), ввиду плохой обусловленности матрицы (значение μ мало), возможна потеря катастрофическая точности решения.

Для ее предотвращения разделим при вычислениях знаменатель выражения (2), представляющий собой сумму вида $\beta_j = \omega_j \omega_i (r_{ji}^c + \lambda \delta_{ji})$, на сумму положительных β^+ и сумму отрицательных β^- слагаемых. Тогда

$$\beta^+ + \beta^- = \sum_{j=1}^n \sum_{i=1}^n \beta_{ji}^+ + \sum_{j=1}^n \sum_{i=1}^n \beta_{ji}^-, \quad \text{где } \beta_{ji}^+ = (\beta_{ji} + |\beta_{ji}|) / 2, \quad \beta_{ji}^- = (\beta_{ji} - |\beta_{ji}|) / 2.$$

$$\beta_{ji}^- = (\beta_{ji} - |\beta_{ji}|) / 2.$$

Рассмотрим более сложный случай, соответствующий вычислениям с плавающей запятой. При вычислении каждого из слагаемых β^+ и β^- получаем после округления мантиссы числа β до t -го разряда после запятой конечную дробь $fl(\beta)$:

$$fl(\beta) = (\dots((\beta_{11}^+ + \beta_{12}^+)(1 + \varepsilon_1^+) + \beta_{13}^+)(1 + \varepsilon_2^+) + \dots \beta_{nn}^+)(1 + \varepsilon_{n^2-1}^+) =$$

$$= (\beta_{11}^+ + \beta_{12}^+) \prod_{i=1}^{n-1} (1 + \varepsilon_i^+) + \beta_{13}^+ \prod_{i=1}^{n-1} (1 + \varepsilon_i^+) + \dots + \beta_{nn}^+ (1 +$$

$$\varepsilon_{n^2-1}^+) +$$

$$+ (\dots(\beta_{11}^- + \beta_{12}^-)(1 + \varepsilon_1^-) + \beta_{13}^-)(1 + \varepsilon_2^-) + \dots \beta_{nn}^-)(1 + \varepsilon_{n^2-1}^-).$$

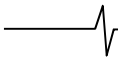
Здесь величина ε_j – погрешность округления числа, $|\varepsilon_j| \leq 0,5 \cdot 10^{-t+1}$. С некоторым приближением можно считать, что относительное возмущение решения имеет вид:

$$|E_1| \leq \frac{n^2 - 1}{2} 10^{-t+1},$$

$$|E_{ji}| \leq \frac{n^2 + 1 - j(n-1) + i}{2} 10^{-t+1}.$$

Тогда, как показано в [9], суммирование в ЭВМ чисел, имеющих одинаковый знак в режиме с плавающей запятой, эквивалентно точному суммированию возмущенных чисел с относительным возмущением E_{ji} . Величина возмущения максимальна в первых слагаемых (j, i – малы) и минимальна в последних (j, k приближается к n).

Очевидно, для уменьшения суммарной ошибки вычисления положительные и отрицательные суммы целесообразно образовывать в порядке возрастания их абсолютных значений и произвести вычитание на последнем шаге. Однако для реализации такого алгоритма, являющегося в процессорах обработки данных составной частью стандартной операции MAC – суммирования



с накоплением, необходимо вначале вычислить и запомнить все слагаемые суммы, что требует дополнительно n^2 ячеек памяти. После формирования массива слагаемых необходима его сортировка. Отметим, каждое слагаемое вносит свой вклад в суммарную ошибку столько раз, сколько раз оно участвует в суммировании.

Для устранения неравноправности слагаемых, следуя [9], произведем m -кратное попарное суммирование слагаемых, где $n = 2^m$ – число суммируемых слагаемых. При этом достигается повышение точности в $n/2 \log_2 n$ раз. Так, при $n = 8$ точность увеличивается примерно на порядок, а при $n = 32$ – на два порядка. В последнем случае занимаемая память сокращается вдвое и отпадает необходимость в сортировке массива.

В развитие этого подхода в [8] предложен рассматриваемый ниже метод, позволяющий в еще большей степени сократить необходимый объем памяти. Такое сокращение достигается при организации вспомогательного количества $2l$ ячеек (по числу десятичных разрядов мантиссы) для положительных и отрицательных слагаемых, в которые слагаемые вначале записываются, а затем полученные частичные суммы переписываются в ячейку, номер которой соответствует порядку суммируемого числа. Данный алгоритм не требует полной сортировки и связанного с этим выделения большого числа ячеек памяти. Это обеспечивает, как показали расчеты, погрешность вычисления квадратичной формы, не превышающей $\Delta \approx 10^{-8}$ при длине слова в 4 байта и 10^{-16} при длине слова в 8 байт, что вполне приемлемо для большинства практических задач [10-13].

Повышение точности при вычислении и обработке сигналов методом сортировки мультипликации

Рассмотрим подробнее вопрос повышения точности при выполнении алгоритмов цифровой фильтрации процессов в реальном времени при плохо обусловленных матрицах случайного процесса. Очевидно, погрешность вычислений растет с увеличением количества обрабатываемых отсчетов сигнала и разброса их абсолютных величин. Типичным примером таких вычислений являются рассмотренные выше операции над данными, характеризующимися плохо обусловленными корреляционными матрицами. Им соответствуют математические и эмпирические модели реальных узкополосных физических процессов. Ниже при решении практической задачи цифровой обработки узкополосных сигналов проиллюстрировано использование методов частичной и полной сортировки данных промежуточных вычислений и показано, что их применение приводит к существенному уменьшению погрешности вычислений или снижению требований к числу разрядов представления операндов.

Данные в вычислительных устройствах могут быть представлены в одном из двух форматов: с фиксированной точкой или с плавающей точкой. В случае представления значений в формате с фиксированной точкой при реальной длине разрядной сетки ($l \leq 16$) динамический диапазон $D = D_{\max} - D_{\min} = 2^l - 2$, а следовательно и разброс абсолютных величин значений получается небольшим (до 5 десятичных порядков). Все числа представляются с одинаковой абсолютной погрешностью

$\varepsilon_{\text{fix}} = 2^{l-1}$. При этом потеря точности при ЦОС может происходить уже на этапе представления чисел в двоичном коде из-за ошибок ограничения.

Формат с плавающей точкой при p -разрядном порядке и m -разрядной мантиссе позволяет представлять числа с одинаковой относительной погрешностью (при этом абсолютная погрешность $\varepsilon_{\text{fl}} = 2^{m-l} = 2^{p-l+2}$) в значительно более широком динамическом диапазоне $D = 3 \cdot 2^{2-l}$. Однако в этом случае практически полное исключение ошибок ограничений сопровождается потерями в точности вычислений, т.к. абсолютная погрешность представления чисел возрастает в $\varepsilon_{\text{fl}}/\varepsilon_{\text{fix}} = 2^{p-l+2}/2^{1-l} = 2^{p-l}$ раз в результате резервирования части разрядов для представления порядка операндов. Независимо от способа представления чисел источники возникновения погрешностей вычислительных операций следующие: квантование результатов перемножения (мультипликации) и отбрасывание части младших разрядов меньшего слагаемого при сложении чисел, абсолютные величины которых различаются на число порядков, соизмеримое с динамическим диапазоном их представления в процессоре обработки сигналов. В дальнейших рассуждениях будем пренебрегать погрешностью квантования мультипликации, которая относительно невелика и, как правило, не приводит к катастрофической потере точности вычислений. Рассмотрим основной источник погрешности, который связан с суммированием большого числа мультипликаций. Попадаем также, что значения хранятся и обрабатываются в формате с плавающей точкой.

Очевидный аппаратный метод уменьшения погрешности вычислений, связанный с увеличением разрядности запоминающих устройств и АЛУ, не всегда оправдан, т.к. ведет к усложнению и удорожанию аппаратных средств. Более целесообразным является программный метод, связанный с разработкой специальных алгоритмов с повышенной точностью и вычислительной эффективностью. При этом следует иметь в виду, что такие алгоритмы могут резко снизить скорость вычислений, что также не всегда оправдано практически. Поэтому необходимо находить компромисс между точностью вычислений и быстродействием [14].

Проанализируем повышение точности вычислений путем организации специального алгоритма формирования суммы мультипликаций, при котором исключается суммирование операндов, абсолютные величины которых различаются на несколько порядков. Одной из мер на пути к этому является раздельное накопление положительных и отрицательных слагаемых, начиная от наименьших значений к большим с последующим объединением результата. При формировании однополярных сумм S^+ и S^- и расчете выражения

$$\sum_{j=0}^{n-1} \sum_{i=0}^{n-1} w_j w_k w_{ji} = \sum_{j=0}^{n-1} \sum_{i=0}^{n-1} d_{ji} = \sum_{j=0}^{n-1} \sum_{i=0}^{n-1} d_{ji}^+ + \sum_{j=0}^{n-1} \sum_{i=0}^{n-1} d_{ji}^- = S^+ + S^-$$

применим методы полной или частичной сортировки, а также метод частичных сумм.

Вычислительный эксперимент показывает, что применение метода полной сортировки позволяет значительно увеличить точность вычислений. На рис 2 пред-

ставлен график зависимости числа λ обусловленности корреляционной матрицы от относительной ширины спектра узкополосного процесса dFT при различном числе обрабатываемых отсчетов.

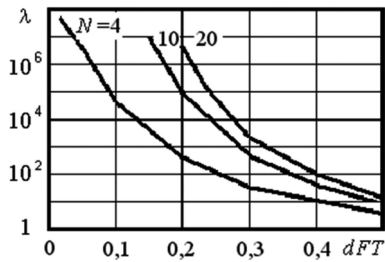


Рис. 2. Зависимость числа обусловленности от относительной ширины полосы процесса ($\lambda(dFT)$) при различном числе обрабатываемых отсчетов N

При хорошо обусловленных матрицах (числе обусловленности более 10^4) применение алгоритма сортировки нецелесообразно, т.к. точность вычислений достаточна без применения специальных алгоритмов. Его следует применять при числе обусловленности меньшем, чем 10^7 .

Оценим время вычислений, затрачиваемое на сортировку массива умножений, для чего также требуется дополнительный объем памяти для хранения всех промежуточных результатов. График, отражающий число операций при сортировке, пропорциональное увеличению времени вычислений t в зависимости от длины сортируемого массива, приведен на рис. 3. На графике по оси ординат отложен аппаратно независимый параметр, имеющий смысл количества операций перестановки, необходимое для сортировки массива. Известные быстрые алгоритмы сортировки обеспечивают сокращение времени сортировки в среднем и непригодны в синхронной обработке данных в условиях лимита времени [10].

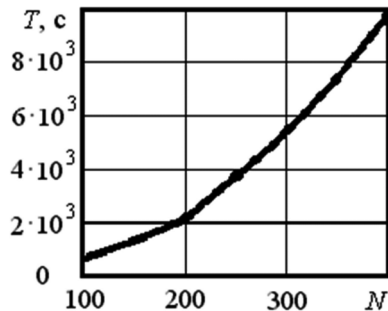


Рис. 3. Зависимость числа операций при сортировке от длины сортируемого массива

Временные затраты можно снизить, применив вместо полной сортировки частичную, при которой сортировке подвергаются умножения при учете только их порядков, т.е. без учета мантиссы. Это позволяет в l/p раз уменьшить вычислительные (временные) затраты. Структурная схема алгоритма вычислений приведена на рис. 4.

Другим предлагаемым алгоритмом, более экономичным по временным затратам, является алгоритм формирования частичных сумм. Каждая частичная сумма складывается из значений одного порядка. Как только порядок частичной суммы превышает порядок составляющих ее значений, она переносится в старшую час-

тичную сумму и обнуляется, а затем накапливается снова. Таким образом, осуществляется один «проход» по всему массиву значений, в результате которого большой массив исходных значений превращается в значительно меньший отсортированный массив частичных сумм. На последнем этапе частичные суммы складываются от младших (меньших) к старшим (большим), т.е. в порядке возрастания. При этом любые два суммируемых значения различаются по абсолютной величине не более, чем на один порядок, и их суммирование не приводит к ощутимой потере точности.

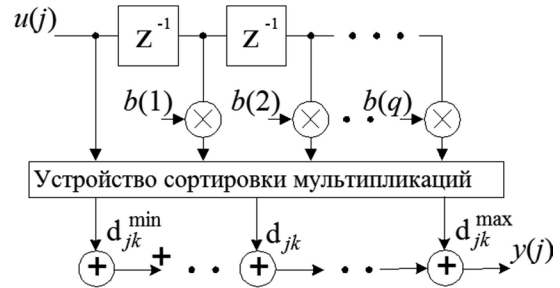


Рис. 4. Схема алгоритма сортировки массивов при обработке сигналов

Использование алгоритма формирования частичных сумм позволяет увеличить точность вычисления минимального собственного числа на 10^{-5} дБ, что несколько меньше, чем при использовании алгоритма с полной сортировкой данных. Однако при этом временные затраты растут пропорционально n , а не n^2 , как при полной сортировке. Таким образом, в отличие от известных методов безусловной сортировки предлагается условная сортировка в зависимости от промежуточных результатов вычислений. Кроме того, отличие данного метода от предложенного Кнудом 1973 г. метода M -рядной сортировки состоит в том, что в данном случае производится динамическая сортировка умножений, которую возможно встроить в стандартный алгоритм MAC.

Сформулируем критерий относительной вычислительной эффективности E , отражающий приоритет временных и аппаратных затрат. Для этого введем обозначения: T – относительные временные затраты, V – относительные аппаратные затраты. Тогда $E = (1/T)\alpha + V(1-\alpha)$, где α – коэффициент, изменяющийся в диапазоне $[0...1]$ и определяющий приоритет между временными и аппаратными затратами. В свою очередь временные затраты можно оценить следующим образом:

$$T = it_0 + (n^2 - 1)(lt_0) + n^2(l^2t_0), \quad (4)$$

где i – количество пересылок, $(n^2 - 1)$ – количество сложений, n^2 – количество умножений, t_0 – время выполнения элементарной операции.

В таблице для рассматриваемых методов вычислений приведены числовые данные по E и T , которые позволяют выбрать компромиссный вариант схемы организации вычислений.

Таблица

Относительные временные затраты	Без сортировки	С полной сортировкой	С частичной сортировкой
T	$184t_0$	$2744t_0$	$2488t_0$

Таким образом, при числе обусловленности менее 10^4 применение специальных алгоритмов нецелесообразно. Если число обусловленности превышает 10^7 , выбор алгоритма производится исходя из требований к точности и быстродействию.

Заключение

Выполнен анализ влияния точности представления операндов в задачах цифровой обработки сигналов на погрешность анализа эффективности систем первичной обработки радиолокационных и навигационных сигналов. Проведено сопоставление выигрыша в точности с «платой» за выигрыш в виде увеличения сложности вычислений. Сформулирован критерий относительной вычислительной эффективности E , позволяющий разработчику перераспределять требования между такими ресурсами процессора, как быстродействие и объем аппаратной части.

Определена последовательности операций при вычислениях, позволяющая повысить точность, и исследованы особенности вычислительных алгоритмов расчета накопления мультипликаций при выполнении стандартной операции МАС.

Разработаны специализированные алгоритмы расчета эффективности систем фильтрации для хорошо и плохо обусловленных корреляционных матриц. Первый обеспечивает упрощенный инженерный метод приближенной оценки эффективности, а второй – повышение точности вычислений и расширение диапазона изменения параметров корреляции преобразуемых процессов, для которых возможно получение численных результатов с приемлемой погрешностью. Точность вычислений повышена путем организации специального алгоритма частичной сортировки при формировании суммы мультипликаций, при котором исключается суммирование операндов, абсолютные величины которых различаются на несколько порядков. При этом сортировке подвергаются порядки мультипликаций без учета мантиссы, что позволяет в l/p раз уменьшить непроизводительные временные затраты. Предложенный метод по сравнению с известным методом позволяет значительно сократить необходимый при вычислениях объем памяти и обеспечивает точность вычисления квадратичной формы $\Delta \approx 10^{-8}$ при длине слова в 4 байта и 10^{-16} при длине слова в 8 байт, что позволяет решить большинство практических задач. Важно, что в разработанном методе временные затраты растут пропорционально n , а не n^2 , как при полной сортировке.

Литература

1. Дж. Голуб, Ч. Ван Лоун. Матричные вычисления. М.: Мир, 1999. – 548 с.
2. Кошелев В.И., Кирдяшкин В.В., Сычев М.И., Ясенцев Д.А. Актуальные вопросы радиолокации // Под ред. П.А. Бакулева.- 2016, 216 с.
3. Кошелев В.И., Андреев В.Г., Белокуров В.А. Современные методы повышения эффективности обнаружения радиолокационных сигналов.- 2016, 154 с.
4. Кошелев В.И. Многоканальная доплеровская фильтрация радиолокационных сигналов // Радиотехника. – 2012. – № 4. – С. 30-35.
5. Кошелев В.И., Белокуров В.А. Синтез алгоритма обнаружения цели, совмещенного с устранением неоднозначности по дальности // Изв. вузов. Радиоэлектроника. – 2012. – № 3. – С. 36-41.
6. Кошелев В.И., Белокуров В.А. Использование фильтра Калмана с перекрестными связями в системе ориентации высокоманевренного объекта // Вестник Рязанского государственного радиотехнического университета. – Рязань: РГРТУ, 2011. – № 1. – вып. 35. – С. 3-7.
7. Адаптивная обработка радиолокационных сигналов на базе процессора БПФ Кошелев В.И. Цифровая обработка сигналов. 2001. № 4. с. 12.
8. Кошелев В.И. АРСС модели случайных процессов. Прикладные задачи синтеза и оптимизации.– М: Радио и связь, 2002.– 116 с.
9. Воеводин В.В. Вычислительные основы линейной алгебры.– М.: Наука, 1977.– 304 с.
10. Блейхут Р. Быстрые алгоритмы цифровой обработки сигналов.– М.: Мир, 1989.– 448 с.
11. Кошелев В.И., Козлов Д.Н. Оптимизация каналов ускорения многоканального обнаружителя при действии коррелированной помехи // Радиотехнические и телекоммуникационные системы. – 2014. – № 3. – С. 5-11.
12. Оценка вычислительных затрат при реализации алгоритма защиты радиолокационных систем от активных шумовых помех. Кошелев В.И., Штрунова Е.С. // Радиотехнические и телекоммуникационные системы. 2011. № 2. С. 39-41.
13. Кошелев В.И. Параметры многоканального обнаружителя доплеровских сигналов Вестник Рязанского государственного радиотехнического университета. 2001. № 8. С. 18.
14. Витязев В.В. Многоскоростная обработка сигналов. – М.: Горячая линия -телеком, 2017. – 336 с.

УДК 537.86+621.396.96

СТАТИСТИЧЕСКИЙ ПОДХОД К РАЗРАБОТКЕ И АНАЛИЗУ АЛГОРИТМОВ ОБНАРУЖЕНИЯ ФРАКТАЛЬНЫХ СИГНАЛОВ И ИЗОБРАЖЕНИЙ

Паршин А.Ю., к.т.н., доцент кафедры РТУ Рязанского государственного радиотехнического университета, email: parshin.a.y@rsreu.ru;

Паршин Ю.Н., д.т.н., профессор, заведующий кафедрой РТУ Рязанского государственного радиотехнического университета, email: parshin.y.n@rsreu.ru.

STATISTICAL APPROACH TO DESIGN AND ANALYSIS OF DETECTION ALGORITHMS OF FRACTAL SIGNALS AND FRACTAL IMAGES

Parshin A. Yu., Parshin Yu. N.

Description and analysis of the parameters of fractal phenomena are considered in context of the theory of optimal statistical solutions. The algorithms for object and edge detection on the image are obtained. The fractal image is given in the form of a two-dimensional fractal Brownian motion - a fractal Brownian surface, as well as a set of distances between vectors in a pseudo-phase space. A refinement of the description of the distances between vectors in a pseudophase space with considering of their given topological dependence, and their probabilistic description is proposed. Expressions for the likelihood ratio logarithm are obtained, the error probabilities are calculated for the object and edge detection in case of Gaussian and non-Gaussian approximations. It is shown that the maximum likelihood estimate of the correlation dimension is sufficient statistics for the detection problem. It is proposed to determine the quality of the edge detection by a system of error probabilities - the probability of false edge fields and the probability of skipping the edge.

Key words: correlation dimension, optimal estimation, maximum likelihood estimation, fractal Brownian motion, error probability of edge detection.

Ключевые слова: корреляционная размерность, оптимальное оценивание, максимально правдоподобные оценки, фрактальное броуновское движение, вероятности ошибок выделения границы.

Введение

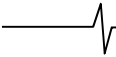
Область применения обработки сигналов и изображений постоянно расширяется, а разнообразие решаемых задач часто выходит за рамки классических методов. Сигналы и изображения наблюдаются в условиях действия помех различной природы, априорной неопределенности, что дополнительно затрудняет применение традиционных моделей и методов.

Исследования последних десятилетий позволили установить свойства самоподобия и дробной меры у сигналов и изображений, получаемых при приеме отраженных от различных объектов сигналов [1, 2]. Исследуемые процессы рассматриваются не как простая совокупность отдельных элементов с определенными характеристиками, а как некоторая структура, обладающая внутренними топологическими связями между элементами и характеризующая сложный объект в целом. Отличительным свойством таких процессов, представленных в некотором топологическом пространстве, является нецелый характер их размерности. Несмотря на существование различных определений и величины размерности для заданного сигнала или изображения [3], каждая из них характеризует общее свойство самоподобия. Это позволяет использовать величину размерности как показатель при решении задач обнаружения, классификации, оценивания параметров

Описание и анализ параметров фрактальных процессов рассмотрены с точки зрения теории оптимальных статистических решений. Получены алгоритмы обнаружения, выделения границы объектов на изображении. Фрактальное изображение задается в виде двумерного фрактального броуновского движения – фрактальной броуновской поверхности, а также совокупностью расстояний между векторами в псевдофазовом пространстве. Проведено уточнение описания расстояний между векторами в псевдофазовом пространстве с учетом их топологической зависимости, дано их вероятностное описание. Получены выражения для логарифма отношения правдоподобия, рассчитаны вероятности ошибок при обнаружении и выделении границ в гауссовском и негауссовском приближении. Показано, что достаточной статистикой для задачи обнаружения является максимально правдоподобная оценка корреляционной размерности. Предложено определять качество выделения границы системой вероятностей ошибок – вероятностью ложных полей границы и вероятностью пропуска границы.

[1, 2]. Вместе с тем теория оптимальной обработки сигналов и изображений на основе фрактальных представлений разработана недостаточно.

Широко известны и применяются методы и алгоритмы оптимальной обработки сигналов и изображений на основе вероятностных моделей и теории оптимальных статистических решений. Наиболее общая постановка задачи и модели сигналов и помех реализованы в оценочно-корреляционно-компенсационном подходе [4-6]. Статистический подход используется также и при обработке сигналов и изображений с фрактальными свойствами, например, фрактального броуновского движения [7]. Другим примером эффективного применения статистических методов является интерпретация корреляци-



онного интеграла как вероятности неперевышения расстояния между векторами заданного значения [8-10].

Обработка фрактальных сигналов является предметом исследований различных ученых [1-3, 8-15]. В работе [1] исследуются различные подходы к использованию фрактальных свойств сигналов, в частности, фрактальной сигнатуры для повышения качества обнаружения малозаметных объектов. Авторы работы [15] подробно исследуют точность оценивания различных фрактальных характеристик, а также ее влияние на качество обнаружения. В работах [11,12] предложено для повышения качества обнаружения комплексировать фрактальный и энергетический обнаружители, что позволяет расширить класс сигнально-помеховых ситуаций для задачи обнаружения.

Целью исследования является разработка и совершенствование статистического подхода в задачах синтеза и анализа эффективности оптимальных алгоритмов обнаружения фрактальных сигналов методом максимального правдоподобия, а также алгоритмов выделения границ на изображении с помощью текстурного анализа путем использования фрактальных свойств анализируемого фрагмента изображения.

Максимально правдоподобное оценивание корреляционной размерности

Известны несколько определений для понятия размерности объектов [3, 16]: емкость, информационная и корреляционная размерности, размерность покрытия, размерность Хаусдорфа, показатель Херста. Данные определения размерности используются для описания одного и того же свойства сигнала или изображения, однако приводят к различным по величине значениям. Используемая далее корреляционная размерность имеет вычислительные преимущества при обработке упорядоченных во времени или пространстве отсчетов. В работах [2, 16] для математического описания используется понятие корреляционного интеграла, который определяет вероятность того, что два независимых наблюдаемых вектора находятся на расстоянии меньше r : $C_w(r) = P(\| \mathbf{x} - \mathbf{y} \|_E \leq r)$, где $\mathbf{x}, \mathbf{y}$ – независимые одинаково распределенные E -мерные векторы, w – вероятностная мера. Корреляционная размерность также описывается вероятностным способом на основе энтропии Колмогорова для хаотического аттрактора [26].

При наблюдении выборки E -мерных векторов $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ корреляционная размерность определяется [17] как двойной предел

$$D = \lim_{r \rightarrow 0} \lim_{n \rightarrow \infty} \frac{\log C_n(r)}{\log r}, \quad (1)$$

где $C_n(r)$ – корреляционный интеграл:

$$C_n(r) = \frac{2}{n(n-1)} \sum_{i=1}^n \sum_{j<i}^n H(r - \|\mathbf{x}_i - \mathbf{x}_j\|_E), \quad (2)$$

где $\|\bullet\|_E$ – означает норму в E -мерном пространстве

вложения, $H(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$ – функция Хэвисайда. Для

измерения расстояния между векторами используются

различные метрики, например евклидова метрика. В работах [18–20] описан способ реконструкции динамической системы путем использования задержанных во времени значений наблюдаемой последовательности отсчетов в качестве значений остальных компонент динамической системы $\mathbf{x}_k = \{x(t_k), x(t_{k+1}), \dots, x(t_{k+E-1})\}$.

Максимально правдоподобная оценка корреляционной размерности предложена в работах [8, 9] и основана на предположении, что корреляционный интеграл вычисляется для независимых случайных расстояний

$$r_k = \|\mathbf{x}_i - \mathbf{x}_j\|_E, \quad i = 1, \dots, n, \quad j = i, \dots, n,$$

$$k = 1, \dots, M = n(n-1)/2,$$

подчиняющихся степенному закону распределения вероятностей.

Для заданного значения корреляционной размерности корреляционный интеграл (2) имеет вид $C(r) \approx r^D$, что позволяет представить расстояния между векторами как случайную величину со степенным законом распределения. При условии нормировки расстояний $r_k / r_{\max}$, закон распределения вероятностей для расстояний имеет вид $F(r) = r^D$, а функция плотности распределения вероятности равна [9]:

$$w(r) = \frac{dF(r)}{dr} = D \times r^{D-1}, \quad 0 < r < 1.$$

Полагая измеренные значения расстояний $\mathbf{r} = \{r_1, r_2, \dots, r_M\}$ между векторами в фазовом пространстве статистически независимыми, запишем многомерную плотность распределения вероятностей

$$w(\mathbf{r}, D) = \prod_{m=1}^M w(r_m) = \prod_{m=1}^M D \times r_m^{D-1}. \quad (3)$$

Так как полученная многомерная плотность (3) является также функцией неизвестной размерности D , то она может рассматриваться как функция правдоподобия. Оценка максимального правдоподобия корреляционной размерности получается в результате решения экстремальной задачи:

$$\hat{D} = \arg \max_D w(\mathbf{r}, D).$$

Переходя к логарифму функции правдоподобия и используя необходимое условие экстремума

$\frac{d}{dD} \ln w(\mathbf{r}, D) = 0$, получим максимально правдоподобную оценку [10]:

$$\hat{D} = -\frac{M}{\sum_{i=1}^M \ln r_i}. \quad (4)$$

Приведенная оценка является эффективной и асимптотически несмещенной. Анализ смещения и дисперсии ошибки оценивания проведен в работе [27].

Фрактальное броуновское движение как модель фрактальных объектов

В качестве модели фрактального сигнала широко используется фрактальное броуновское движение (ФБД), характеризуемое интенсивностью σ^2 и показа-

телем Херста H [7]. Размерность одномерного ФБД определяется показателем Херста $D = 2 - H$, а двумерного ФБД – $D = 3 - H$. Если сигналы вида ФБД наблюдаются в смеси с гауссовским шумом, то это дополнительно затрудняет их обнаружение и идентификацию. Поэтому актуальной задачей является синтез алгоритмов оптимального обнаружения и различения сигналов вида ФБД на фоне гауссовского шума.

Пусть сигнал вида ФБД x_n принимается на фоне гауссовского шума v_n :

$$y_n = x_n + v_n, \quad n = 1, \dots, N,$$

где N – число отсчетов во времени наблюдаемого процесса, v_n – независимые отсчеты гауссовского шума с дисперсией D_v . Отсчеты ФБД формируются одним из известных методов [7] во временной или спектральной областях. ФБД являются гауссовскими случайными процессами, поэтому их свойства полностью определяются корреляционными матрицами для одномерного сигнала [7]:

$$\begin{aligned} \mathbf{M}\{(x(t_2) - x(t_1))(x(t_4) - x(t_3))\} = \\ = \frac{1}{2}\sigma^2 \left[-(t_2 - t_4)^{2H} + (t_2 - t_3)^{2H} + (t_1 - t_4)^{2H} - \right. \\ \left. - (t_1 - t_3)^{2H} \right], \end{aligned} \quad (5)$$

а также для двумерного изображения [28]:

$$\begin{aligned} R_{ij} = \mathbf{M}\{x_i x_j\} = \\ = \frac{1}{2}\sigma^2 [(i - j - 1)^{2H} - 2(i - j)^{2H} + (i - j + 1)^{2H}] \end{aligned}$$

Таким образом, по заданным N отсчетам формируется $M \times M$ матрица $\mathbf{R}$ корреляций всех $M = \frac{1}{2}N(N - 1)$ возможных приращений

$$\Delta x_m = x(t_i) - x(t_j), \quad i = 1, \dots, N, \quad j, i, m = 1, \dots, M.$$

Заданный таким образом вектор приращений $\Delta \mathbf{x} = \{\Delta x_1, \dots, \Delta x_M\}^T$ имеет многомерную плотность распределения вероятностей

$$w(\Delta \mathbf{x}) = \frac{1}{(2\pi)^{M/2} \sqrt{\det \mathbf{R}}} \exp \left[-\frac{1}{2} \Delta \mathbf{x}^T \mathbf{R}^{-1} \Delta \mathbf{x} \right],$$

где корреляционная матрица $\mathbf{R}$ зависит от показателя Херста H , а плотность распределения вероятностей может быть рассмотрена как функция правдоподобия

$$\begin{aligned} w(\Delta \mathbf{x} / H) = \frac{1}{(2\pi)^{M/2} \sqrt{\det \mathbf{R}(H)}} \times \\ \times \exp \left[-\frac{1}{2} \Delta \mathbf{x}^T \mathbf{R}^{-1}(H) \Delta \mathbf{x} \right]. \end{aligned} \quad (6)$$

Отношение правдоподобия при наблюдении приращений сигнала в виде ФБД на фоне гауссовского шума имеет вид

$$\begin{aligned} \Lambda(H) = \sqrt{\frac{\det \mathbf{R}_v}{\det(\mathbf{R}(H) + \mathbf{R}_v)}} \times \\ \times \exp \left[-\frac{1}{2} \Delta \mathbf{x}^T (\mathbf{R}(H) + \mathbf{R}_v)^{-1} \Delta \mathbf{x} + \frac{1}{2D_v} \Delta \mathbf{x}^T \Delta \mathbf{x} \right]. \end{aligned} \quad (7)$$

В общем виде произвольно заданных моментов вре-

мени вычисление определителя и обращение корреляционной матрицы приращений $\mathbf{R}$ представляет собой сложную вычислительную задачу. Поэтому в некоторых случаях целесообразно принимать в рассмотрение только некоррелированные приращения (5) на непересекающихся интервалах времени, для которых корреляция равна:

$$\mathbf{M}\{\Delta x_i \Delta x_j\} = \delta_{ij} \sigma^2 \Delta t_i^{2H},$$

где δ_{ij} – символ Кронекера, $i = 1, \dots, N - 1$. В этом случае

матрица $\mathbf{R}$ является диагональной, определитель

$$\det \mathbf{R} = \sigma^{2N} \prod_{n=1}^{N-1} \Delta t_n^{2H}.$$

Многомерная плотность распределения вероятностей для этого случая имеет вид

$$\begin{aligned} w(\Delta \mathbf{x}) = \frac{1}{(2\pi)^{(N-1)/2} \sqrt{\sigma^{2N} \prod_{n=1}^{N-1} \Delta t_n^{2H}}} \times \\ \times \exp \left[-\frac{1}{2\sigma^2} \sum_{n=1}^{N-1} \frac{\Delta x_n^2}{\Delta t_n^{2H}} \right], \end{aligned}$$

откуда получаем отношение правдоподобия для случая наблюдения сигнала вида приращений ФБД на фоне приращений гауссовского шума:

$$\begin{aligned} \Lambda = \frac{(2D_v)^{(N-1)/2}}{\sqrt{\sigma^{2N} \prod_{n=1}^{N-1} \Delta t_n^{2H}}} \times \\ \times \exp \left[-\frac{1}{2} \sum_{n=1}^{N-1} \left(\frac{\Delta x_n^2}{\sigma^2 \Delta t_n^{2H}} - \frac{1}{2D_v} \right) \Delta x_n^2 \right]. \end{aligned}$$

Данный алгоритм является квазиоптимальным, так как не учитывает корреляцию приращений на перекрывающихся интервалах. Вместе с тем он обладает существенными вычислительными преимуществами, так как не содержит затратных в вычислительном отношении операций обращения матриц.

Получение некоррелированных наблюдений ФБД возможно также в спектральной области. Известно [16], что спектральная плотность мощности ФБД равна

$$G(f) \propto \frac{1}{f^{2H+1}}.$$

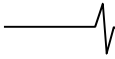
Модель ФБД может быть также задана в спектральной области в виде совокупности гармоник $\underline{\mathbf{S}} = \{\underline{\mathbf{S}}_1, \dots, \underline{\mathbf{S}}_M\}$, полученных дискретным преобразованием Фурье от значений ФБД $\{x_1, \dots, x_N\}$, где все гармоники являются независимыми комплексными гауссовскими

числами, дисперсии которых равны $|\underline{\mathbf{S}}_m|^2 = \frac{G_0}{m^{2H+1}}$,

$m = 1, \dots, M = N/2$ – номер гармоники.

Многомерная плотность распределения вероятностей спектральных составляющих ФБД имеет вид

$$w(\underline{\mathbf{S}}) = \prod_{m=1}^M \frac{m^{2H+1}}{\pi^M G_0^M} \exp \left[-\frac{1}{G_0} \sum_{m=1}^M |\underline{\mathbf{S}}_m|^2 m^{2H+1} \right].$$



Для спектрального представления наблюдаемых данных отношение правдоподобия имеет вид

$$\Lambda = \frac{1}{\prod_{m=1}^M \left(\frac{G_0}{G_{v0} m^{2H+1}} + 1 \right)} \times \exp \left[- \sum_{m=1}^M \left(\frac{1}{\frac{G_0}{m^{2H+1}} + G_{v0}} - \frac{1}{G_{v0}} \right) |S_m|^2 \right]. \quad (8)$$

Таким образом, в спектральной области алгоритм вычисления отношения правдоподобия получается более простым, так как исключается операция обращения матрицы. Обнаружение ФБД на фоне гауссовского шума в спектральной области производится в результате вычисления статистики (8) и сравнения ее с порогом.

Анализ характеристик обнаружения фрактального сигнала

Полученные соотношения используются для оценки качества обнаружения, а также для расчета характеристик обнаружения в гауссовском приближении при $N \gg 1$. Вероятность пропуска сигнала и вероятность ложной тревоги соответственно равны:

$$P = \int_{-\infty}^h w_1(z) dz, \quad F = \int_h^{\infty} w_0(z) dz, \quad (9)$$

где h – значение порога принятия решения, $z = \ln \Lambda$ – статистика обнаружения. Используя выражение для

интеграла вероятности $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$, запишем

вероятности (9) ошибочных решений

$$P = \frac{1}{2} + \frac{1}{2} \text{erf} \left(\frac{h - m_{z/1}}{\sqrt{2D_{z/1}}} \right), \quad F = \frac{1}{2} - \frac{1}{2} \text{erf} \left(\frac{h - m_{z/0}}{\sqrt{2D_{z/0}}} \right).$$

Более точные результаты при $N > 1$ можно получить путем вычисления плотности распределения вероятностей суммы нескольких независимых χ^2 случайных величин с различными дисперсиями методом характеристических функций. Так как плотность распределения вероятностей суммы случайных величин представляет собой свертку, то получим

$$w_\theta(z) = F^{-1} \left\{ \prod_{m=1}^M F \{ w_{m\theta}(S) \} \right\},$$

где $F\{\dots\}$, $F^{-1}\{\dots\}$ – обозначает прямое и обратное преобразования Фурье соответственно, $w_{m\theta}(S) =$

$$= \frac{1}{D_{m\theta}} \exp \left(- \frac{S}{D_{m\theta}} \right) \quad - \quad \text{экспоненциальное}$$

распределение при наличии $\theta = 1$ и отсутствии $\theta = 0$ ФБД в наблюдаемых данных,

$$D_{m\theta} = \left(\frac{1}{G_{v0}} - \frac{1}{\frac{G_0}{m^{2H+1}} + G_{v0}} \right) \times \left(\frac{\theta G_0}{m^{2H+1}} + G_{v0} \right).$$

Анализ обнаружителя ФБД, проведенный методом статистического моделирования, показал хорошее согласование с теоретическими результатами при различных отношениях сигнал-шум, показателя Херста, числе гармоник спектра ФБД.

Полученные результаты обобщаются на случай, когда сигнал представлен как сумма нескольких ФБД с различными параметрами, а также когда шумы наблюдения также являются ФБД, например, если на входе действует сумма K независимых ФБД с различными показателями Херста H_k , $k = 1, \dots, K$. Получаемые при этом аналитические выражения для логарифма отношения правдоподобия являются достаточной статистикой при решении как задачи обнаружения, так и многих задач оценивания параметров сигналов, моделируемых ФБД.

Расчеты характеристик проводятся при заданном отношении сигнал-шум, определяемом следующим

образом: $q = \frac{D_X}{D_V}$, где $D_X = \sum_{m=1}^M \overline{|S_m|^2} = G_X \sum_{m=1}^M \frac{1}{m^{2H+1}}$ –

дисперсия фрактального броуновского движения,

$D_V = \sum_{m=1}^M \overline{|V_m|^2} = M G_V$ – дисперсия гауссовского шума.

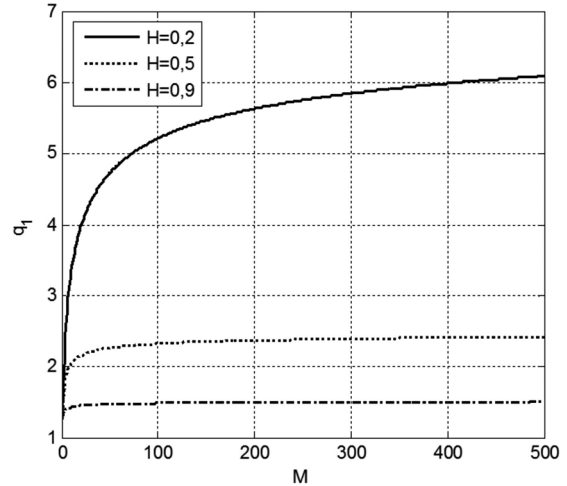


Рис. 1. Зависимость отношения сигнал-шум от количества гармоник, $q = 10$

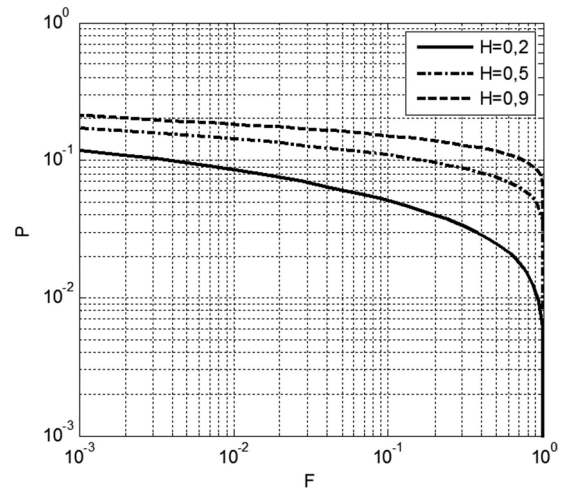


Рис. 2. Характеристики обнаружения в гауссовском приближении, $q = 1$, $M = 100$

Результаты расчетов (рис. 1) показывают, что отношение сигнал-шум увеличивается при уменьшении показателя Херста. Характеристики обнаружения, полученные при различных значениях отношения сигнал-шум q и показателя Херста, представлены на рис. 3, 4.

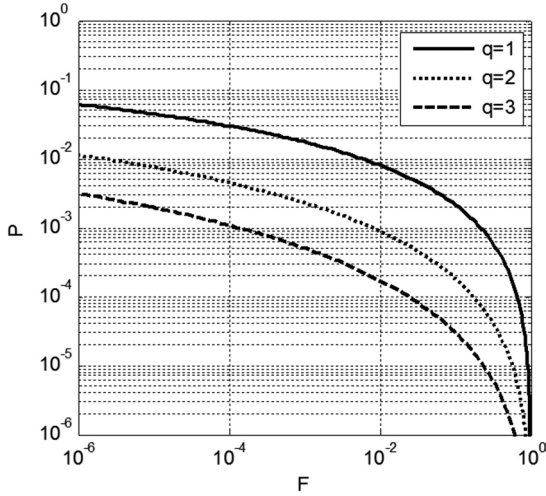


Рис. 3. Характеристики обнаружения, $M = 100$, $H = 0,9$

Алгоритм выделения границы по фрактальным признакам фона и объекта

Метод проверки статистических гипотез на основе достаточной статистики в виде отношения правдоподобия используется также для решения задачи выделения границы между фоном и объектом:

$$\Lambda = \frac{w_1(\mathbf{r})}{w_2(\mathbf{r})} = \frac{\prod_{m=1}^M D_1 r_m^{D_1-1}}{\prod_{m=1}^M D_2 r_m^{D_2-1}}, \quad (10)$$

где D_1 , D_2 – корреляционные размерности объекта и фона соответственно.

Для выделения границы объекта изображение разбивается на кадры квадратной формы, состоящие из полей анализа. При формировании границы кадра из 2-х полей граница может располагаться в одном из четырех априорно заданных направлений – горизонтальном, вертикальном и по двум направлениям диагонали. Оценка границы представляется как одно из полей. Для определенности граница отображается верхним, правым и нижним диагональным полями.

При формировании границы квадратного кадра из 9 полей граница также может располагаться в одном из четырех априорно заданных направлений: горизонтально или вертикально посередине кадра, а также по двум направлениям диагонали. Оценка границы представляется как набор из 3 полей, расположенных вертикально или горизонтально в середине кадра, а также по диагонали кадра.

Граница разделяет кадр на две части, в каждой из которых формируется N векторов [14]. Достаточная статистика для задачи выделения границы равна отношению правдоподобия для границы различного вида в предположении, что в частях кадра по обе

стороны границы располагаются объекты с различным значением корреляционной размерности. При условии равновероятных расположений объекта и фона относительно границы отношение правдоподобия (10) принимает вид

$$\Lambda_k = \frac{\frac{1}{2} \prod_{m=1}^{M_1} D_1 r_{1km}^{D_1-1} \times \prod_{m=1}^{M_2} D_2 r_{2km}^{D_2-1} + \frac{1}{2} \prod_{m=1}^{M_2} D_1 r_m^{D_1-1} + \frac{1}{2} \prod_{m=1}^{M_1} D_2 r_m^{D_2-1}}{\frac{1}{2} \prod_{m=1}^{M_1} D_2 r_{1km}^{D_2-1} \times \prod_{m=1}^{M_2} D_1 r_{2km}^{D_1-1} + \frac{1}{2} \prod_{m=1}^{M_2} D_1 r_m^{D_1-1} + \frac{1}{2} \prod_{m=1}^{M_1} D_2 r_m^{D_2-1}}, \quad (11)$$

где $k = 1, \dots, 4$ – номер разновидности границы, r_{1km}, r_{2km} – расстояния в полях анализа 1 и 2 при k -й разновидности границы, r_m – расстояния между векторами в кадре.

Определение наличия границы производится путем сравнения с порогом максимального из отношений правдоподобия (11), вычисленного для каждой из проверяемых гипотез, состоящей в наличии объектов с различными значениями корреляционной размерности в соседних полях одного и того же кадра при наличии границы определенного вида. Принимается решение о наличии границы, соответствующей кадру с номером $\hat{k} = \arg \max_k \Lambda_k$,

если значение отношения правдоподобия $\Lambda_{\hat{k}}$ превышает некоторое пороговое значение.

Анализ эффективности алгоритма выделения границы

Анализ эффективности предложенных алгоритмов выделения границ изображения проведен на примере объекта с границей в виде восьмиугольника, что позволяет более полно отразить влияние нелинейности границы объекта.

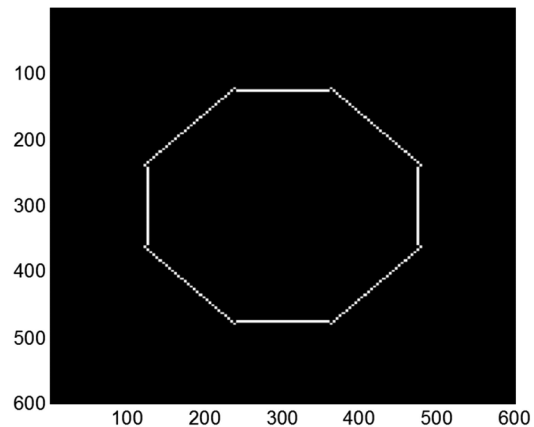
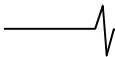


Рис. 4. Идеальная граница объекта

для исследуемого метода выделения границы

Выделение границы проводилось методом статистического моделирования отношения правдоподобия (11). На рис. 4 представлена оценка границы, полученная при идеальных условиях измерений границы объекта: ошиб-



ки, вызванные конечным объемом выборки при статистическом оценивании исключены. Некоторая неровность границы на рис. 4 объясняется особенностями формирования кадра, принятыми видами границы и конечным размером исследуемого кадра. На рис. 5 приведены результаты выделения границы объекта с использованием алгоритма максимального правдоподобия (19). На изображении заметны ошибки выделения границы в виде ложных отметок и пропуска границы.

Визуальные методы оценки качества обработки изображения не позволяют объективно произвести сравнение анализируемых алгоритмов и определить их эффективность. Предложено оценивать эффективность алгоритмов выделения границ по вероятностным характеристикам, в качестве которых приняты вероятность ложных полей границы на всем изображении $P_{\text{ЛГ}} \approx \frac{N_{\text{ЛГ}}}{N_{\text{НГР}}}$ и вероятность пропуска границы $P_{\text{ПГ}} \approx \frac{N_{\text{ПГ}}}{N_{\text{ГР}}}$, где $N_{\text{ЛГ}}$ – число полей на изображении, не совпадающих с отметками эталонной границы, $N_{\text{НГР}}$ – число полей эталонного изображения, не относящихся к эталонной границе, $N_{\text{ПГ}}$ – число пропущенных полей на эталонной границе, $N_{\text{ГР}}$ – общее число полей на эталонной границе.

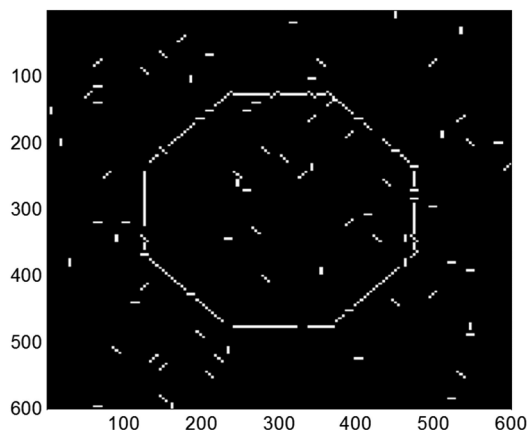


Рис. 5. Результат обработки алгоритмом максимального правдоподобия, $D_1=1,3$, $D_2=1,3$, граница из трех полей

Выбирая априорную величину вероятности появления ложных полей, можно устанавливать величину порога, с которым производится сравнение статистики (11).

При этом обработка осуществляется с помощью алгоритма максимального правдоподобия (11), полученного в предположении независимых расстояний. Для уменьшения геометрической зависимости расстояний предлагается аналогично работам [10, 13] применить усечение, при котором учитываются расстояния, не превышающие некоторого заданного значения. При этом оценка корреляционной размерности получается асимптотически эффективной, что позволяет использовать алгоритм максимального правдоподобия, синтезированный для независимой выборки. Однако число усеченных расстояний невелико и приближенно соответствует числу введенных в данной статье независимых расстояний.

Сравнение результатов анализа, проведенного для независимых расстояний, с результатами для зависи-

мых расстояний показывает заметное увеличение вероятностей ошибок при использовании зависимых расстояний. Полученные результаты согласуются с результатами работы [27], в которой также установлено сильное влияние зависимости расстояний на точность оценивания корреляционной размерности.

Зависимость расстояний при оценивании корреляционной размерности

Для оценивания корреляционной размерности методом максимального правдоподобия необходимо определить расстояния между всеми парами N векторов в выборке. Если для проведения расчета корреляционной размерности используются все возможные расстояния, то такой набор называется V -статистикой [21], а число расстояний равно d_E . Для данной статистики характерна повторяемость расстояний, что приводит к неинформативности более половины элементов статистики. Кроме того, часть неповторяющихся расстояний является взаимозависимой, что не согласуется с моделью, принятой при синтезе максимально правдоподобного алгоритма оценивания корреляционной размерности. Вследствие этого V -статистика практически не используется для получения оптимальных оценок.

Для получения более информативной U -статистики [22] выделяют неповторяющиеся расстояния между векторами:

$$\|\mathbf{x}_i - \mathbf{x}_j\|, i = 2, \dots, N, j = 1, \dots, i-1.$$

Количество расстояний в статистике равно $M_U = N(N-1)/2$. В этом случае также имеется геометрическая взаимозависимость части расстояний, что не позволяет строго применять модель независимых расстояний для синтеза оптимального алгоритма.

Для уменьшения взаимозависимости расстояний используют T -статистику [22] с применением окна Тейлера [23]: статистика содержит расстояния

$$\|\mathbf{x}_i - \mathbf{x}_j\|, i = T+1, \dots, N, j = 1, \dots, i-T,$$

где T – размер окна Тейлера. При этом разность индексов устанавливается равной T , что влияет на независимость расстояний между векторами. Число расстояний в статистике равно $N - d_E - 1$. В частном случае $T=1$ получаем U -статистику. В некоторых работах [25] предлагается выбирать размер окна Тейлера равным $T = E - 1$.

Более точно независимость расстояний может быть установлена путем рассмотрения геометрических свойств множества векторов. Рассмотрим E -мерное пространство векторов общего положения, в котором задана метрика и определены расстояния между всеми парами векторов. Координаты векторов могут быть не зависимыми, например, статистически независимыми случайными числами.

Для выявления свойства взаимозависимости расстояний между N векторами общего положения в евклидовом пространстве рассмотрим сначала более простой случай векторов на плоскости: $E=2$, анализируя расстояния последовательно для каждого вектора.

Необходимо определить возможность задать положение каждого из векторов в пространстве расстояний с точностью до сдвига и поворота. Расстояние r_{12}, r_{13} между 1-м и 2-м векторами может быть произвольным и задается только положением этих векторов в пространстве. Расстояние r_{13} между 1-м и 3-м векторами также может быть произвольным, а расстояние r_{23} между 2-м и 3-м векторами зависит от значений выше рассмотренных расстояний в соответствии с правилом треугольника: $0 \leq r_{23} < r_{12} + r_{13}$. Вместе с тем минимальное значение расстояния зависит от заданной метрики. Например, в пространстве с евклидовой метрикой нижняя граница диапазона расстояния r_{23} равна $|r_{12} - r_{13}| \leq r_{23} \leq r_{12} + r_{13}$. Таким образом, появляется расстояние, которое зависит от других расстояний, а положение 3-го вектора определяется расстояниями r_{13}, r_{23} с точностью до отражения. Однозначность положения 3-го вектора гарантируется, если добавляется 4-й вектор и фиксируется еще одно расстояние r_{34} .

Обобщая проведенные рассуждения на произвольное число векторов, можно заключить:

1) полностью независимыми являются только $N-1$ расстояний;

2) расстояние между 2-м и 3-м векторами является условно независимым от r_{12}, r_{13} и ограничивается правилом треугольника $|r_{12} - r_{13}| < r_{23} < r_{12} + r_{13}$;

3) расстояние между 3-м и 4-м векторами принимает одно из двух predetermined значений; величина этих расстояний зависит от выбранной метрики, остальные расстояния до 3-го вектора полностью определяются предыдущими расстояниями;

4) добавление каждого последующего вектора общего положения однозначно фиксирует все предыдущие векторы, что делает все условно независимые расстояния фиксированными, и дает одно независимое расстояние и несколько условно независимых расстояний.

Учитывая сложный характер взаимозависимости расстояний между векторами, затруднительно установить общее выражение для распределения вероятности этих расстояний. Более простой задачей является определение максимального числа M_1 независимых расстояний между N векторами в E -мерном пространстве.

Пространство конфигураций, задаваемых координатами всех векторов, имеет размерность $D_{conf} = NE - E(E+1)/2$, где вычитаемый член соответствует количеству движений E -мерного пространства, преобразующих конфигурации векторов в тождественные, то есть E сдвиги и $E(E-1)/2$ вращений [8, 24]. При этом количество фиксирующих расстояний между векторами не может быть меньше размерности пространства конфигураций, т.е. если количество расстояний меньше D_{conf} , то они конфигурацию однозначно задать не могут. И наоборот, если количество расстояний больше D_{conf} , то они образуют избыточную систему координат и являются

зависимыми.

Вместе с тем система, содержащая E выбранных точек и расстояния от каждой точки до них всех, имеет как раз такой размер вращений [24]. Однако пространство конфигураций отображается в соответствующее евклидово пространство неоднозначно: каждому набору расстояний соответствует дискретный набор конфигураций, расположенных по разные стороны от E выбранных точек. Для устранения неоднозначности необходимо добавить расстояния до ещё одной точки от всех, т.е. $N-E-1$ расстояний. Таким образом, получается верхняя и нижняя граница для необходимого количества расстояний:

– нижняя граница $M_{UL} = D_{conf}$,

– верхняя граница $M_{UH} = D_{conf} + N - E - 1$.

Например, для двумерного изображения $E=2$, а число расстояний равно $M_{UL} = 2N - 3$, $M_{UH} = 3N - 6$. Можно предположить, что в приложениях более существенной является размерность системы, т.е. нижняя граница необходимого числа расстояний.

При условии, что все M расстояний $\mathbf{r} = \{r_m, m = 1, \dots, M\}$ между векторами независимы, а корреляционная размерность равна D совместная плотность распределения вероятностей расстояний имеет вид (2).

Зависимые расстояния формируются в соответствии со следующим алгоритмом.

1. Генерируются $N-1$ независимых случайных чисел со степенным распределением вероятностей в диапазоне значений $(0;1)$; данные числа задают расстояния от первого вектора до всех $N-1$ остальных векторов. Совместная плотность распределения вероятностей этих $N-1$ независимых расстояний между векторами имеет вид

$$w_1(\mathbf{r}_1 / D) = \begin{cases} \prod_{m=1}^{N-1} D r_{1m}^{D-1}, & \mathbf{r}_1 \in [0,1) \\ 0, & \mathbf{r}_1 \notin [0,1) \end{cases}.$$

2. Генерируется $N-2$ независимых случайных чисел со степенным распределением вероятностей в диапазоне значений $(r_{\min k}; r_{\max k})$, $k = 3, \dots, N$; данные числа задают условно независимые расстояния от второго вектора до всех $N-2$ остальных векторов кроме 1-го вектора. Минимальное и максимальное значения определяются правилом треугольника: $r_{\min k} = |r_{12} - r_{1k}|$,

$r_{\max k} = r_{12} + r_{1k}$, $k = 3, \dots, N$. Совместная плотность распределения вероятностей этих $N-2$ независимых расстояний между векторами имеет вид

$$w_2(\mathbf{r}_2 / \mathbf{r}_1, D) = \begin{cases} \prod_{m=N}^{2N-3} \frac{D}{r_{\max k} - r_{\min k}} (r_m - r_{\min k})^{D-1}, & r_m \in [r_{\min k}, r_{\max k}), k = 3, \dots, N \\ 0, & r_m \notin [r_{\min k}, r_{\max k}), k = 3, \dots, N, \\ & m = N + k - 3. \end{cases}$$

Так как минимальные и максимальные значения $r_{\min k}, r_{\max k}$ зависят от расстояний с номерами

1,...,N-1, то расстояния r_1, r_2 также статистически зависимы, а их совместная плотность распределения вероятностей равна

$$w_{12}(r_1, r_2 / D) = w_1(r_1 / D) w_2(r_2 / r_1, D).$$

На основе данной плотности распределения вероятностей рассчитывается отношение правдоподобия – универсальной решающей статистики.

3. Формируются координаты 1-го и 2-го векторов в пространстве вложения $D_E = 2$: $x_1 = 0, y_1 = 0, x_2 = r_{12}, y_2 = 0$. Координаты остальных $i = 3, \dots, N$ векторов определяются из геометрии их положения с использованием теоремы косинусов и расстояний от i -го вектора до 1-го и 2-го векторов:

$$x_i = \frac{r_{12}^2 + r_{1i}^2 - r_{2i}^2}{2r_{12}}, \quad y_i = \pm \sqrt{r_{1i}^2 - x_i^2}.$$

Для определенности знак координат y_i выбирается положительным.

4. По найденным координатам всех векторов вычисляются оставшиеся

$$\frac{N^2}{2} - \frac{5N}{2} - 3$$

зависимых

$$\text{расстояний } r_3 = \left\{ r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, \right. \\ \left. i = 3, \dots, N, j = i+1, \dots, N \right\}.$$

Так как расстояния r_3 определены полностью расстояниями r_1, r_2 , то они не содержат дополнительной информации для оценивания корреляционной размерности.

Заключение

Показано, что методы теории оптимальных статистических решений могут с успехом применяться также и для обработки фрактальных сигналов и изображений. Основой эффективности статистических методов является нерегулярный характер, а также сравнительно большой объем наблюдаемых данных. В этих условиях статистическое описание фрактальных сигналов и изображений производится различными методами: использование модели одномерного и двумерного фрактального броуновского движения, статистическое описание расстояний между векторами в псевдофазовом пространстве. Такой подход позволяет на основе теории оптимальных статистических решений получать алгоритмы обработки для решения различных задач: обнаружения, различения, выделения границ, оценивания параметров, а также проводить анализ эффективности обработки. Вместе с тем, статистическое описание получено не для всех фрактальных сигналов и изображений и их характеристик, что делает актуальным продолжение исследований в этом направлении.

Литература

1. Потапов А.А. Фракталы в радиофизике и радиолокации: топология выборки. М.: Университетская книга, 2005. 847 с.
2. Culter C.D. A review of the theory and estimation of fractal dimension // Nonlinear time series and chaos v.1. Dimension estimation and models, 1993. P. 1-107.
3. Паркер Т.С., Чжуа Л.О. Введение в теорию

хаотических систем для инженеров // ТИИЭР, 1987, Т. 75, № 8. С. 6-40.

4. Сосулин Ю.Г. Теория обнаружения и оценивания стохастических сигналов. М.: Советское радио, 1978. 320 с.

5. Сосулин Ю.Г., Паршин Ю.Н. Оценочно-корреляционно-компенсационная обработка многомерных сигналов // Радиотехника и электроника. 1981. Т. 26, № 8. – С. 1635-1643.

6. Сосулин Ю. Г., Костров В. В., Паршин Ю. Н. Оценочно-корреляционная обработка сигналов и компенсация помех. М.: Радиотехника, 2014. 632 с.

7. Кроновер Р.М. Фракталы и хаос в динамических системах. Основы теории. М.: Постмаркет, 2000. 352 с.

8. Pisarenko D. V., Pisarenko V.F. Statistical estimation of the correlation dimension // Physics Letters, A (197), 1995. P. 31-39.

9. Luca L., Lasocki S., Luzio D., Vitale M. Fractal dimension confidence interval estimation of epicentral distribution // Annali di Geofisica, 42(5), 1999. P. 911-925

10. Luciana De Luca, Dario Luzio, Massimo Vitale. A ML Estimator of the Correlation Dimension for Left-Hand Truncated Data Samples // Pure and applied geophysics, V.159, № 11-12, 2002. P. 2789-2803.

11. Паршин А.Ю., Паршин Ю.Н. Комплексирование энергетического и фрактального обнаружителей // Проблемы передачи и обработки информации в сетях и системах телекоммуникаций. 17-я Международная науч. техн. конф. Часть 2. Рязань: РГПУ, 2012. С. 37-39.

12. Parshin A., Parshin Yu. Usage of non-Gaussian statistics for RF signals detection by complex energy and fractal detector // International radar symposium – IRS 2013, Proceeding, volume I, II, Drezden, Germany. – German /institute of navigation, 2013, pp. 779-784.

13. Паршин А.Ю., Паршин Ю.Н. Использование максимально правдоподобных оценок фрактальной размерности в негауссовских статистиках для обнаружения радиосигналов// Цифровая обработка сигналов. 2013. №1. С. 42-46.

14. Parshin A.Yu., Parshin Yu.N. Synthesis and analysis of efficiency of algorithm for objects edges detection by maximum likelihood method // 2nd Mediterranean Conference on Embedded Computing (MECO). Budva. Montenegro. June 15th-21st 2013. 282 p.

15. Сосулин Ю.Г., Русскин А.Б. Фрактальное обнаружение протяженных мало контрастных объектов на изображениях // Радиотехника. 2009. № 12. С. 48-57.

16. Pesin Ya. B. On rigorous mathematical definitions of correlation dimension and generalized spectrum for dimensions // Journal of Statistical Physics, V.71, 1993. P. 529-547.

17. Grassberger P., Procaccia I. Characterization of strange attractors // Physical Review Letters, V. 50, 1983. P. 346-349.

18. Packard N.H., Crutchfield J.P., Farmer J.D., Shaw R.S. Geometry from a time series // Physical Review Letters, V. 45, 1980. P. 712-716.

19. Takens F. Detecting strange attractors in turbulence // Lecture Notes in mathematics, V. 898, 1981. P. 366-381.

20. Ying-Cheng Lai, David Lerner. Effective scaling regime for computing the correlation dimension from chaotic time series // Physica D, V.115, 1998. P. 1-18.

21. Антипов О.И., Неганов В.А., Потапов А.А. Детер-

минированный хаос и фракталы в дискретно-нелинейных системах. М.: Радиотехника, 2009. 235 с.

22. Cees Diks. Nonlinear Time Series Analysis: Methods And Applications (Nonlinear Time Series and Chaos: v. 4) World Scientific. 1999. 209 p.

23. Theiler J. Spurious dimensions from correlation algorithms applied to limited time-series data // Phys. Rev. A 34. 1986. pp. 2427-2432.

24. Fried E., Absztrakt Algebra, Elemi Uton. Budapest, MuszakiKonyvkiado, 1972. 260 p.

Provenzale A., Smith J.A., Vio R., Murate G. Distinguishing between low-dimensional dynamics and randomness in

measured time series. Physica D 58. 31-49. 1992.

25. Шустер Г. Детерминированный хаос: Введение. М.: Мир. 1988. 240 с.

26. Паршин Ю.Н., Паршин А.Ю. Максимально правдоподобное оценивание корреляционной размерности с учетом влияния смещения оценки и усечения диапазона масштабов // Вестник Рязанского государственного радиотехнического университета. 2013. №4-1 (46.). С. 18-25.

27. Hong Qian, Gary M Raymond, James B Bassingthwaite. On two-dimensional fractional Brownian motion and fractional Brownian random field // J. Phys. A: Math. Gen. 31 (1998) L527-L535.

РАСЧЕТ ЛОГАРИФМА ОТНОШЕНИЯ ПРАВДОПОДОБИЯ OFDM-СИГНАЛА ПРИ ИСПОЛЬЗОВАНИИ ТЕХНИКИ ПОВОРОТА СОЗВЕЗДИЯ

Левченко А.С., аспирант ФГБОУ ВПО «Национальный исследовательский университет «Московский энергетический институт, e-mail: levch.andrew@gmail.com.

LLR CALCULATION FOR OFDM SIGNAL USING CONSTELLATION ROTATION TECHNIQUE

Levchenko A.S.

Method that allows to specify log-likelihood ratio for OFDM signal using constellation rotation technique has been proposed. Method allows to specify log-likelihood ratio by taking into account multiplicative noise component of signal caused by multipath propagation. Results of competitive modeling using terrestrial broadcasting system RAVIS model are listed for cases when constellation rotation technique is used and without it, SNR advantage is determined. Modeling shows that for QPSK constellation SNR advantage is more than 1 dB with code rate $\frac{3}{4}$, in range from 0.2 to 0.5 with code rate $\frac{1}{2}$. For QAM16 constellation SNR advantage is more than 0.3 dB with code rate $\frac{3}{4}$, with code rate $\frac{1}{2}$ there is no SNR advantage.

Key words: COFDM, soft demodulator, log-likelihood ratio, VHF band, RAVIS, rotated, constellation rotation.

Ключевые слова: COFDM, демодулятор с «мягким» выходом, логарифм отношения правдоподобия, ОБЧ-диапазон частот, RAVIS, поворот созвездия.

Введение

В стандарте DVB-T2 [1] предлагается для повышения эффективности эфирной цифровой системы передачи данных, основанной на модуляции COFDM, использовать технику поворота сигнального созвездия. Суть этой техники приведена ниже. Далее будут использоваться термины «ячейка OFDM», «символ OFDM» в соответствии с [2].

На передающей стороне последовательность перемеженных по времени бит $b_0, b_1, \dots, b_N$, поступая на вход квадратурного амплитудного модулятора, отображается на координаты Re и Im исходной ячейки OFDM (т.е. каждой битовой последовательности ставится в соответствие комплексное число $\dot{X} = Re + i \cdot Im$, где i – мнимая единица). Все возможные 2^N отображений образуют сигнальное созвездие. При этом, поскольку созвездие повернуто, каждой исходной последовательности бит соответствует уникальное значение как по координате Re , так и по координате Im . В этом первое принципиальное различие рассматриваемой схемы модуляции от классической.

Таблица 1. Точки повернутого созвездия QPSK

Последовательность бит	Номер точки созвездия	Координата Re	Координата Im
00	1	+0,3173	+0,9483
01	2	-0,9483	+0,3173
10	3	+0,9483	-0,3173
11	4	-0,3173	-0,9483

Рассмотрим на примере созвездия QPSK. В табл. 1 приведены точки повернутого на 26,5 градуса созвездия QPSK (см. рис. 1, слева).

Предложен способ расчета логарифма отношения правдоподобия сигнала COFDM для случая, когда применяется техника поворота созвездия. Способ позволяет уточнить оценки логарифмов отношения правдоподобия, учитывая мультипликативную составляющую ошибки принимаемого сигнала, возникающую при многолучевом распространении. Приведены результаты сравнительного моделирования для системы эфирного вещания RAVIS с использованием техники поворота созвездия и без ее использования, определен энергетический выигрыш от ее использования. Для созвездия QPSK выигрыш составляет более 1дБ при скорости кода $\frac{3}{4}$ и от 0,2 до 0,5 дБ при скорости кода $\frac{1}{2}$. Для созвездия QAM16 выигрыш составляет более 0,3 дБ при скорости кода $\frac{3}{4}$, при скорости кода $\frac{1}{2}$ имеет место проигрыш в отношении сигнал/шум.

После указанного отображения точки созвездия формируются в блок из M ячеек и осуществляется циклический сдвиг координаты Im , как указано в табл. 2. В этом второе принципиальное различие рассматриваемой схемы модуляции от классической.

Таблица 2. Координаты Re и Im ячеек до и после сдвига

Номер ячейки	1	2	...	M-1	M
Координата Re до сдвига	Re_1	Re_2	...	Re_{M-1}	Re_M
Координата Im до сдвига	Im_1	Im_2	...	Im_{M-1}	Im_M
Координата Re после сдвига	Re_1	Re_2	...	Re_{M-1}	Re_M
Координата Im после сдвига	Im_M	Im_1	...	Im_{M-2}	Im_{M-1}

В результате циклического сдвига координаты Im все возможные сочетания уникальных координат Re_i и Im_j образуют новое, расширенное сигнальное созвездие (см. рис. 1, справа). Точки расширенного созвездия QPSK приведены в табл. 3.

За счет последующего перемежения ячеек координаты R_i и I_{i-1} -ой ячейки OFDM передаются в канале на удаленных друг от друга частотах (обозначим их как f_R и f_I) в ячейках OFDM $\dot{X}_1$ и $\dot{X}_2$, принадлежащих расширенному созвездию. Благодаря этому, если одна из координат (R_i или I_i) будет утеряна (например, ячейка $\dot{X}_1$ будет подавлена глубоким замиранием), исходную битовую последовательность возможно восстановить по ячейке $\dot{X}_2$.

На приёмной стороне демодулятор формирует логарифмы отношения правдоподобия (далее LLR) для бит на основании принятых ячеек OFDM. В документе [3] приведено выражение для логарифма отношения правдоподобия, который рассчитывается демодулятором:

$$LLR(b_i) = \quad (1)$$

$$= \log \left[\frac{\sum_{x \in C_i^0} \exp \left\{ \frac{(I - \rho_I I_x)^2 + (Q - \rho_Q Q_x)^2}{2\sigma^2} \right\}}{\sum_{x \in C_i^1} \exp \left\{ \frac{(I - \rho_I I_x)^2 + (Q - \rho_Q Q_x)^2}{2\sigma^2} \right\}} \right],$$

где I – принятая на частоте f_R координата Re -ой ячейки OFDM, Q – принятая на частоте f_I координата Im -ой ячейки OFDM, ρ_I и ρ_Q – амплитудные множители, описывающие замирание на частотах f_R и f_I , I_x и Q_x – проекции точки x повернутого созвездия на оси Re и Im .

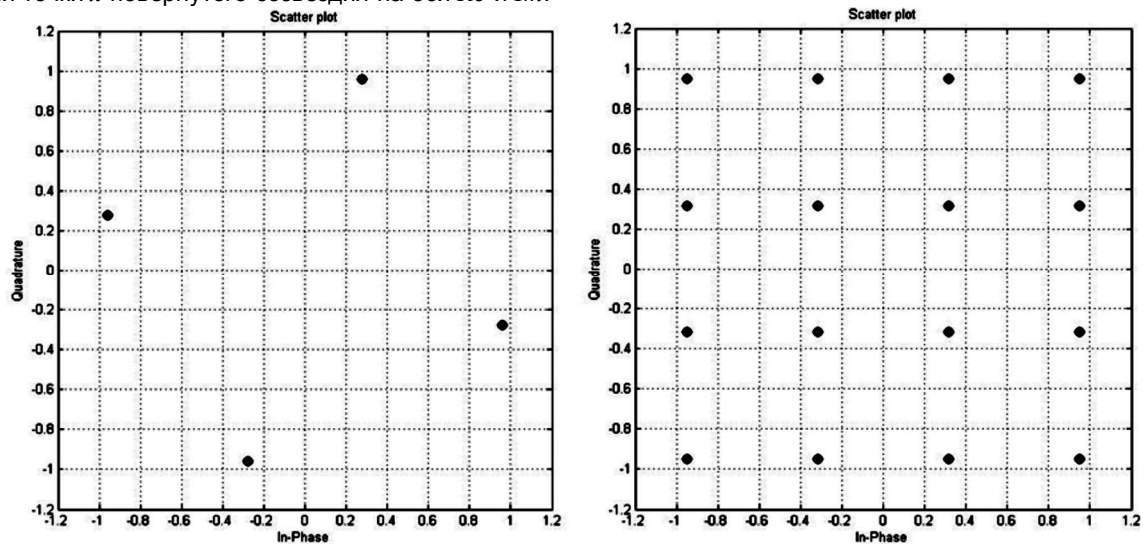


Рис. 1. Сигнальное созвездия до (слева) и после (права) циклического сдвига координаты Im

Таблица 3. Точки расширенного созвездия QPSK

1	2	3	4
+0.3173	+0.9483	+0.3173	+0.3173
00	00	00	01
5	6	7	8
-0.9483	+0.9483	-0.9483	-0.3173
01	00	01	10
9	10	11	12
+0.9483	+0.9483	+0.9483	-0.3173
10	00	10	10
13	14	15	16
-0.3173	+0.9483	-0.3173	-0.3173
11	00	11	10

1^1
-0.9483^2
01^4
-0.3173^3
10^5

Примечания:

1. Номер точки созвездия
2. Координата Re
3. Координата Im
4. Принятая битовая последовательность, если точка созвездия принята в канале Re
5. Принятая битовая последовательность, если точка созвездия принята в канале Im

Постановка задачи

В данной работе была поставлена задача получить аналитическое выражение LLR с учетом мультипликативной ошибки для использующей поворот сигнального созвездия системы эфирного цифрового вещания, основанной на модуляции COFDM.

Синтез логарифма отношения правдоподобия

В общем случае логарифм отношения правдоподобия выражается как

$$LLR(b_i) = \ln \left[\frac{P\{b_i = 0\}}{P\{b_i = 1\}} \right] \quad (2)$$

Прохождение ячеек на частотах f_R и f_I через канал можно описать как

$$\begin{cases} \dot{Y}_1 = \dot{X}_1 \dot{H}_1 + \dot{N}_1, \\ \dot{Y}_2 = \dot{X}_2 \dot{H}_2 + \dot{N}_2, \end{cases}$$

где $\dot{Y}_1$ и $\dot{Y}_2$ – принятые на частотах f_R и f_I ячейки OFDM, $\dot{N}_1$ и $\dot{N}_2$ – аддитивные шумы с нулевым матожиданием и дисперсией $\left(\frac{\sigma_{R1}^2}{\sigma_{I1}^2}\right)$ и $\left(\frac{\sigma_{R2}^2}{\sigma_{I2}^2}\right)$, $\dot{H}_1$ и $\dot{H}_2$ – коэффициенты передачи канала (здесь и далее символ с надстрочной точкой обозначает комплексную величину).

На основании принятых на частотах f_R и f_I символов $\dot{Y}_1$ и $\dot{Y}_2$ необходимо сделать вывод о том, какая последовательность бит была передана. Поскольку одной и той же последовательности бит соответствуют разные сочетания значений $\dot{X}_1$ и $\dot{X}_2$, выражение (2) можно уточнить:

$$LLR(b_i) = \ln \left[\frac{\sum_j \sum_k P\{X_1 = x_j, X_2 = x_k | Y_1 = y_1, Y_2 = y_2\}}{\sum_m \sum_n P\{X_1 = x_m, X_2 = x_n | Y_1 = y_1, Y_2 = y_2\}} \right], \quad (4)$$

где j, k принимают значения такие, что $\dot{X}_1 = x_j$ и $\dot{X}_2 = x_k$ соответствуют нулю в i -ом бите, m и n принимают значения такие, что $\dot{X}_1 = x_m$ и $\dot{X}_2 = x_n$ соответствуют единице в i -ом бите. Так, исходя из табл. 3, при передаче информации в виде точек сигнального созвездия QPSK, для старшего бита индексы принимают следующие значения: $j = 1, 2, 3, 4, 5, 6, 7, 8$; $k = 1, 2, 5, 6, 9, 10, 13, 14$; $m = 9, 10, 11, 12, 13, 14, 15, 16$; $n = 3, 4, 7, 8, 11, 12, 15, 16$.

Набор событий $\dot{X}_1 = x_j, \dot{X}_2 = x_k$ и $\dot{X}_1 = x_m, \dot{X}_2 = x_n$ образует полную группу событий. В процессе передачи обязательно должно произойти одно из таких событий, при этом все эти события попарно несовместимы. Тогда можно воспользоваться формулой Байеса:

$$\begin{aligned} P(X_1 = x_j, X_2 = x_k | Y_1 = y_1, Y_2 = y_2) = \\ = \frac{P(X_1 = x_j, X_2 = x_k) P(Y_1 = y_1, Y_2 = y_2 | X_1 = x_j, X_2 = x_k)}{\sum_{s,l} P(X_1 = x_s, X_2 = x_l) P(Y_1 = y_1, Y_2 = y_2 | X_1 = x_s, X_2 = x_l)} = \end{aligned} \quad (5)$$

$$= \frac{P(Y_1 = y_1, Y_2 = y_2 | X_1 = x_j, X_2 = x_k)}{\sum_{s,l} P(Y_1 = y_1, Y_2 = y_2 | X_1 = x_s, X_2 = x_l)},$$

где индексы s, l принимают все допустимые значения. Применим (5) к числителю и знаменателю (4), знаменатель (5) сократится, получим

$$LLR(b_i) = \ln \left[\frac{\sum_j \sum_k P(Y_1 = y_1, Y_2 = y_2 | X_1 = x_j, X_2 = x_k)}{\sum_m \sum_n P(Y_1 = y_1, Y_2 = y_2 | X_1 = x_m, X_2 = x_n)} \right] \quad (6)$$

Поскольку при зафиксированных X_1 и X_2 случайные величины Y_1 и Y_2 не зависят,

$$\begin{aligned} P(Y_1 = y_1, Y_2 = y_2 | X_1 = x_j, X_2 = x_k) = \\ = P(Y_1 = y_1 | X_1 = x_j) P(Y_2 = y_2 | X_2 = x_k), \end{aligned}$$

тогда $LLR(b_i) =$ (7)

$$= \ln \left[\frac{\sum_j \sum_k P(Y_1 = y_1 | X_1 = x_j) P(Y_2 = y_2 | X_2 = x_k)}{\sum_m \sum_n P(Y_1 = y_1 | X_1 = x_m) P(Y_2 = y_2 | X_2 = x_n)} \right]$$

Выражение (7) имеет скрытый недостаток: имея векторы вероятностей $P(\dot{Y}_1 = \dot{y}_1 | \dot{X}_1 = \dot{x}_j), j = 1 \dots \frac{N^2}{2}$ и $P(\dot{Y}_2 = \dot{y}_2 | \dot{X}_2 = \dot{x}_k), k = 1 \dots \frac{N^2}{2}$ (где N – размер исходного созвездия), нам понадобится рассчитать $2 * \left(\frac{N^4}{4} - \frac{N^2}{2}\right)$ попарных произведений этих вероятностей. Так, для повернутого созвездия QPSK ($N=4$) потребуется рассчитать 112 умножений, для QAM16 ($N=16$) 32512 умножений, для QAM64 – более 8 миллионов умножений. Это делает выражение (7) мало-применимым в реальной аппаратуре. Однако, с учетом того, что j и k, m и n друг от друга не зависят, его можно упростить:

$$LLR(b_i) = \quad (8)$$

$$= \ln \left[\frac{\sum_j P(Y_1 = y_1 | X_1 = x_j) \sum_k P(Y_2 = y_2 | X_2 = x_k)}{\sum_m P(Y_1 = y_1 | X_1 = x_m) \sum_n P(Y_2 = y_2 | X_2 = x_n)} \right]$$

Теперь вместо 8 миллионов умножений для QAM64 будет достаточно двух (!).

С учетом того, что оценка коэффициента передачи канала в приемнике возможна лишь с некоторой точностью, можно переписать (3) как

$$\begin{cases} Y_1 = X_1 \left(H_1 + N_{H1} \right) + N_1, \\ Y_2 = X_2 \left(H_2 + N_{H2} \right) + N_2, \end{cases}$$

где $\dot{N}_{H1}$ и $\dot{N}_{H2}$ – нормально распределенные комплексные случайные величины с нулевым математическим ожиданием и дисперсией $\left(\sigma_{H1R}^2\right)$ и $\left(\sigma_{H2R}^2\right)$. В [4] было показано, что

$$p_{\bar{y}}\left(\bar{y} \mid \bar{X} = x_i\right) = \frac{1}{\pi\left(\sigma_N^2 + \sigma_H^2\left|x_i\right|^2\right)} \exp\left(\frac{-\left|\bar{y} - x_i H\right|^2}{\sigma_N^2 + \sigma_H^2\left|x_i\right|^2}\right) \quad (9)$$

Тогда

$$LLR(b_i) = \ln \frac{\sum_j \frac{1}{\left(\sigma_{N_1}^2 + \sigma_{H_1}^2\left|x_j\right|^2\right)} \exp\left(\frac{-\left|y_1 - x_j H_1\right|^2}{\sigma_{N_1}^2 + \sigma_{H_1}^2\left|x_j\right|^2}\right)}{\sum_m \frac{1}{\left(\sigma_{N_1}^2 + \sigma_{H_1}^2\left|x_m\right|^2\right)} \exp\left(\frac{-\left|y_1 - x_m H_1\right|^2}{\sigma_{N_1}^2 + \sigma_{H_1}^2\left|x_m\right|^2}\right)} \times \frac{\sum_j \frac{1}{\left(\sigma_{N_1}^2 + \sigma_{H_1}^2\left|x_j\right|^2\right)} \exp\left(\frac{-\left|y_2 - x_j H_1\right|^2}{\sigma_{N_1}^2 + \sigma_{H_1}^2\left|x_j\right|^2}\right)}{\sum_m \frac{1}{\left(\sigma_{N_1}^2 + \sigma_{H_1}^2\left|x_m\right|^2\right)} \exp\left(\frac{-\left|y_2 - x_m H_1\right|^2}{\sigma_{N_1}^2 + \sigma_{H_1}^2\left|x_m\right|^2}\right)} \quad (10)$$

где j, k принимают значения такие, что $\dot{X}_1 = x_j$ и $\dot{X}_2 = x_k$ соответствуют нулю в i -ом бите, m и n принимают значения такие, что $\dot{X}_1 = x_m$ и $\dot{X}_2 = x_n$ соответствуют единице в i -ом бите, $\bar{y}_1$ и $\bar{y}_2$ – значения, которые приняли случайные величины $\dot{Y}_1$ и $\dot{Y}_2$ (т.е. ячейки, которые были приняты на частотах f_R и f_I), $\sigma_{N_1}^2 = \sigma_{R1}^2 + \sigma_{I1}^2$, $\sigma_{N_2}^2 = \sigma_{R2}^2 + \sigma_{I2}^2$, $\sigma_{H_1}^2 = \sigma_{H1R}^2 + \sigma_{H1I}^2$, $\sigma_{H_2}^2 = \sigma_{H2R}^2 + \sigma_{H2I}^2$.

Результаты моделирования

Для оценки эффективности техники поворота сигнального созвездия было проведено моделирование его работы в составе системы цифрового эфирного вещания РАВИС [8] с использованием функциональной модели, разработанной в среде Simulink. Данная модель позволяет имитировать функционирование системы «передатчик-канал-приёмник» для различных видов многолучевого канала в соответствии с [5]. В качестве критерия оценки эффективности демодулятора выбрана вероятность битовой ошибки в приёмнике.

На рис. 2-8 приведены зависимости для различных каналов и режимов передачи данных. Результаты моделирования при модуляции QAM64 не приведены, поскольку поворот такого созвездия не дает выигрыша, а наоборот, приводит к потерям.

На рис. 2-9 приведены зависимости для различных каналов и режимов передачи данных. Результаты моделирования при модуляции QAM64 не приведены, поскольку поворот такого созвездия не дает выигрыша, а наоборот, приводит к потерям.

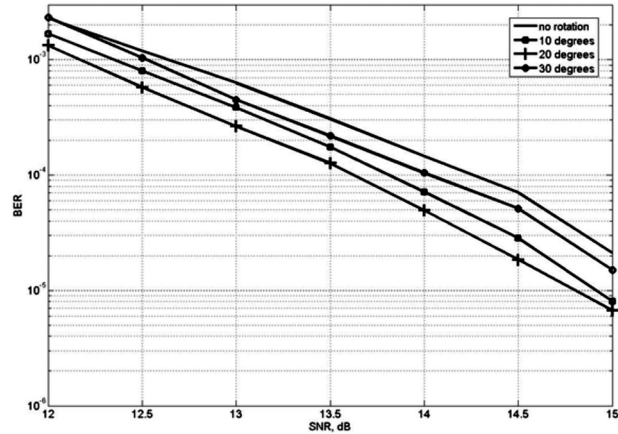


Рис. 2. Канал Rural, скорость кода $\frac{3}{4}$, QPSK, без коррекции шума

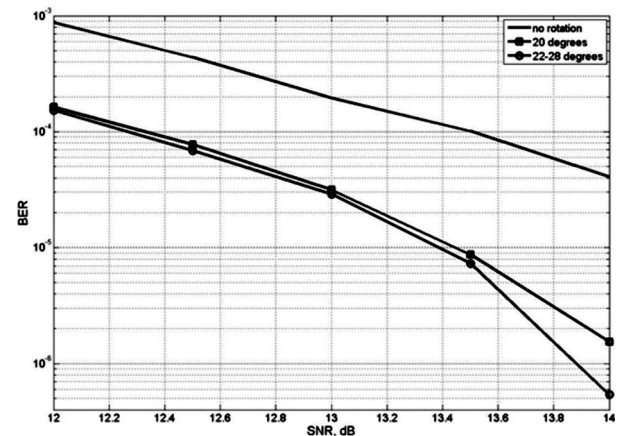


Рис. 3. Канал Rural, скорость кода $\frac{3}{4}$, QPSK, коррекция шума фильтром Винера

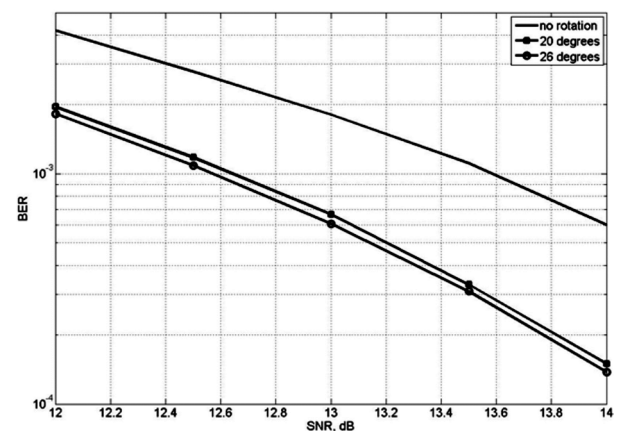


Рис. 4. Канал Urban, скорость кода $\frac{3}{4}$, QPSK, коррекция шума фильтром Винера

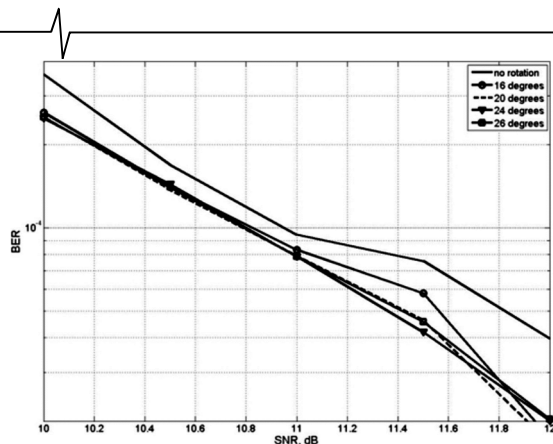


Рис. 5. Канал Urban, скорость кода $\frac{1}{2}$, QPSK, коррекция шума фильтром Винера

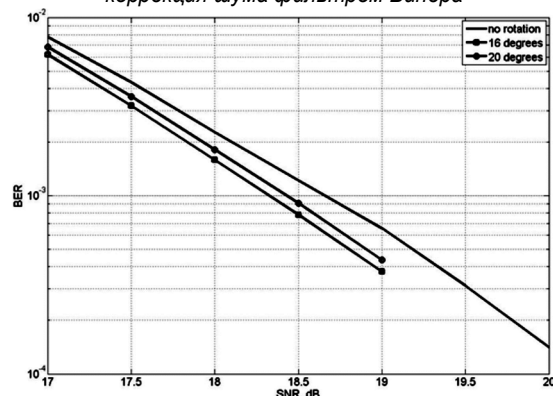


Рис. 6. Канал Rural, скорость кода $\frac{3}{4}$, QAM16, без коррекции шума

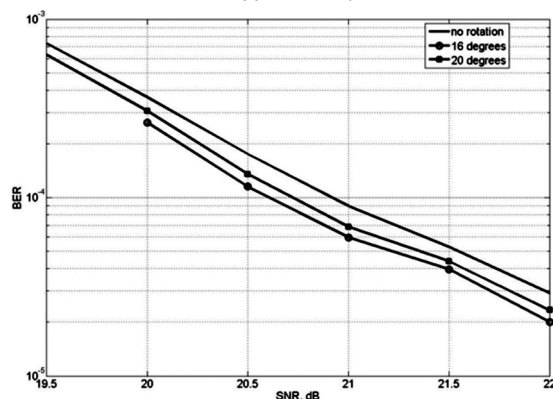


Рис. 7. Канал Urban, скорость кода $\frac{3}{4}$, QAM16, коррекция шума фильтром Винера

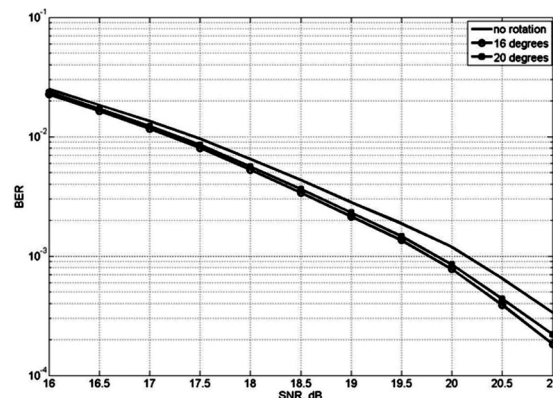


Рис. 8. Канал Urban, скорость кода $\frac{3}{4}$, QAM16, без коррекции шума

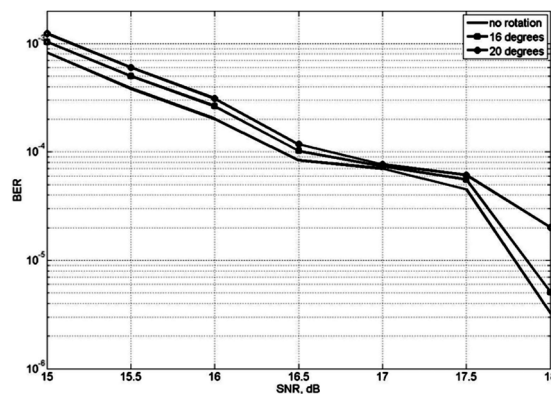


Рис. 9. Канал Urban, скорость кода $\frac{1}{2}$, QAM16, без коррекции шума

Закключение

Получено выражение для расчета логарифма отношения правдоподобия с учетом мультипликативной ошибки для использующей поворот сигнального созвездия системы эфирного цифрового вещания, основанной на модуляции COFDM.

Результаты моделирования показывают, что техника поворота созвездия более эффективна при высоких скоростях кода. Так, для созвездия QPSK выигрыш составляет более 1 дБ при скорости кода $\frac{3}{4}$ и от 0,2 до 0,5 дБ при скорости кода $\frac{1}{2}$. Для созвездия QAM16 выигрыш составляет более 0,3 дБ при скорости кода $\frac{3}{4}$; при скорости кода $\frac{1}{2}$ имеет место проигрыш в отношении сигнал/шум (см. рис. 9).

Технику поворота созвездия нецелесообразно использовать при сигнальных созвездиях высоких порядков, ввиду малого выигрыша в отношении сигнал/шум или даже его ухудшения. Разумно ограничиться его её использованием только для созвездия QPSK.

Оптимальный угол поворота созвездия зависит от порядка созвездия и от свойств канала передачи данных. Для исследуемых каналов наилучшие результаты получены при углах поворота 26° для QPSK и 16° для QAM16. Разброс угла поворота $\pm 1^\circ$ относительно оптимального не приводит к ощутимому снижению эффективности техники поворота созвездия.

Литература

1. ETSI EN 302 755: «DigitalVideoBroadcasting (DVB); Framestructurechannelcodingandmodulationforasecondgeneration digital terrestrial television broadcasting system (DVB-T2)», 2014.
2. ГОСТ Р 54309-2011. Аудиовизуальная информационная система реального времени (РАВИС). Процессы формирования кадровой структуры, канального кодирования и модуляции для системы цифрового наземного узкополосного радиовещания в ОВЧ диапазоне. Технические условия.
3. ETSI TS 102831:«Digital Video Broadcasting (DVB); Implementation guidelines for a second generation digital terrestrial television broadcasting system (DVB-T2)», 2012.
4. Левченко А.С. Демодулятор COFDMсигнала с мультипликативными шумами // Радиотехнические и телекоммуникационные системы. №1, 2016.
5. ETSI ES 201 980 V3.1.1 (2009-08). Digital Radio Mondiale (DRM); System Specification.

ОЦЕНКА ХАРАКТЕРИСТИК МЕТОДА НЕСОГЛАСОВАННОЙ ФИЛЬТРАЦИИ, МИНИМИЗИРУЮЩЕГО ИНТЕГРАЛЬНЫЙ УРОВЕНЬ БОКОВЫХ ЛЕПЕСТКОВ ФАЗОКОДОМАНИПУЛИРОВАННЫХ СИГНАЛОВ

Зайцев Г.В., д.т.н., начальник СКБ ПАО «НПО «Алмаз», e-mail: gennady-zaytsev@yandex.ru;

Кондранина Н.С., ведущий инженер СКБ ПАО «НПО «Алмаз», e-mail: kondranina.nataliya@gmail.com;

Литвинов Д.М., начальник отдела СКБ ПАО «НПО «Алмаз», e-mail: litvinov_dmitry@inbox.ru.

CHARACTERIZATION OF MISMATCHED FILTERING, MINIMIZING INTEGRATED SIDELobe LEVEL FOR PHASE-CODED PULSE

Zaytsev G. V., Kondranina N. S., Litvinov D. M.

The paper investigates method of mismatched filtering, minimizing integrated sidelobe level for binary phase-coded pulse. Main characteristics of the output signal are analyzed (integrated sidelobe level, peak sidelobe level, loss in signal-to-noise ratio) for several classes of modulation sequence families: M-sequences, Legendre sequences, and sequences up to length 271 with best known integrated sidelobe level of autocorrelation function. Resulting parameters of mismatched filtering are described and tabulated for best found sequences with linear growth of quality criterion.

Key words: phase-coded pulse, mismatched filtering, minimization of integrated sidelobe level.

Ключевые слова: фазокодоманипулированные сигналы, несогласованная фильтрация, минимум интегрального уровня боковых лепестков.

Введение

В работе анализируются двоичные последовательности с низким уровнем боковых лепестков непериодической корреляционной функции. Такие последовательности используются в радиолокации, физике, системах передачи информации, биомедицинских исследованиях и других областях науки и техники. Для определенности ниже рассматриваются радиолокационные приложения, причем доплеровская частота в данной работе считается нулевой.

При использовании в радиолокации фазокодоманипулированный импульсный сигнал формируется следующим образом. Прямоугольный импульс гармонического колебания разбивается на n одинаковых по длительности сегментов, а начальная фаза колебания в каждом сегменте изменяется в соответствии с некоторым кодом. В данной работе рассматривается наиболее распространенный случай, при котором код является двоичным. Тогда операцию формирования фазокодоманипулированного импульса можно описать следующим образом. Сегмент гармонического колебания, имеющий номер i , умножается на i -ый коэффициент кода $s_{i \text{ code}}$, $i = 0, 1, \dots, n-1$, который может принимать значения ± 1 . Сигнальные характеристики радиолокатора при использовании описанного импульса полностью определяются свойствами последовательности $s_0 = (s_{0 \text{ code}}, s_{1 \text{ code}}, \dots, s_{n-1 \text{ code}})$, поэтому дальнейший анализ достаточно провести для дискретных последовательностей вместо непрерывного сигнала [1, 2]. Далее кодовую последовательность s_0 будем называть исходной.

Рассматривается метод несогласованной фильтрации фазокодоманипулированных сигналов, обеспечивающий минимизацию интегрального уровня боковых лепестков. Анализируются получаемые характеристики (уровень максимального бокового лепестка, интегральный уровень боковых лепестков, потери в отношении сигнал/шум и др.) для трех семейств двоичных модулирующих последовательностей: M-последовательности, последовательности Лежандра и наилучшие известные последовательности до длины 271. Приводится таблица найденных последовательностей с линейным и максимально быстрым ростом показателей качества от длины импульсной характеристики фильтра.

В приемной аппаратуре радиолокатора вычисляется взаимная корреляционная функция (ВКФ) принимаемого из эфира сигнала, содержащего полезную составляющую s_0 , и опорной последовательности $x = (x_0, x_1, \dots, x_{N-1})$, длина N которой, вообще говоря, может быть больше n : $N \geq n$, а составляющие элементы – действительные числа. При этом изучение боковых лепестков, обусловленных полезным сигналом, сводится к анализу ВКФ последовательностей x и s_0 [1, 2].

Для краткой записи этих вычислений дополним последовательность s_0 нулями до длины N . Пусть начальный отрезок такой последовательности содержит m нулей. Тогда в конце последовательности имеется $N - n - m$ нулей. Обозначим такую последовательность той же буквой, что и исходную, но без нулевого индекса:

$$s(m) = (0, 0, \dots, 0, s_{0 \text{ code}}, s_{1 \text{ code}}, \dots, s_{n-1 \text{ code}}, 0, 0, \dots, 0) = \{s_i\}, \quad i = 0, 1, \dots, N-1 \quad (1)$$

Аргумент m этой последовательности для краткости будем опускать, если приводимые выражения справедливы при любом m или если значение индекса очевидно.

Взаимная корреляционная функция $c = \{c_k\}$ последовательностей x и $s(m)$ может быть вычислена по формуле

$$c_k = \sum_{i=0}^{N-1} x_i s_{k+i}^*,$$

где считается, что $s_r = 0$ при $r < 0$ и $r \geq N$, а символ $*$ означает операцию сопряжения, т.е. транспонирования и комплексного сопряжения элементов вектора или матрицы.

В данной работе будем рассматривать ВКФ как конечную последовательность, содержащую только члены, не равные тождественно нулю. При введенных обозначениях индекс k таких элементов принадлежит множеству $\Omega = \{-(N-m-1), -(N-m-1)+1, \dots, -1, 0, 1, \dots, n+m-1\}$. Количество $N+n-1$ элементов этого множества будем называть длиной ВКФ. Элементы ВКФ могут быть представлены в виде

$$c_k = \mathbf{x} \mathbf{D}_k \mathbf{s}^*, k \in \Omega, \quad (2)$$

где $\mathbf{D}_k - N \times N$ – матрица сдвига последующего за ней вектор-столбца на k элементов вверх при неотрицательном k :

$$\mathbf{D}_k = \begin{pmatrix} \overbrace{0 \ 0 \ \dots 0}^{k+1} 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & & & & 0 \end{pmatrix}, \quad k \geq 0,$$

и сдвига на k элементов вниз при отрицательном k :

$$\mathbf{D}_k = \mathbf{D}_{-k}^T, \quad k < 0.$$

Заметим, что поскольку последовательность $\mathbf{s}$ является действительной, то опорную последовательность $\mathbf{x}$ не имеет смысла делать комплексной. Поэтому все рассматриваемые в работе последовательности являются действительными. Тем не менее, приводимые ниже формулы будут записываться в общем виде.

Выбор последовательности $\mathbf{x}$ при заданной последовательности $\mathbf{s}$ зависит от решаемой задачи. В радиолокации общепринято максимизировать отношение сигнал/шум на выходе приемного устройства. При этом в случае аддитивного гауссовского шума оптимальная опорная последовательность определяется соотношением $\mathbf{x} = \mathbf{s}_0$, $N = n$, обработка сигнала называется согласованной фильтрацией, а взаимная корреляционная функция становится автокорреляционной (АКФ) [1, 2]. На рис. 1 показан типичный вид полученной АКФ на выходе согласованного фильтра. Для краткой записи модулирующей последовательности из элементов ± 1 , использованной при построении рисунка, будем использовать следующее соглашение: сначала последовательность преобразуется в двоичный вид путем замены -1 на 0 , а затем представляется в шестнадцатеричном коде. Подобное соглашение используется также везде ниже. Для рис. 1 выбрана последовательность $\mathbf{s}_0 = \text{CE0549}$ длины 24. Значения АКФ c_k соединены на рис. 1 прямыми линиями для получения АКФ соответствующего аналогового сигнала [1, 2].

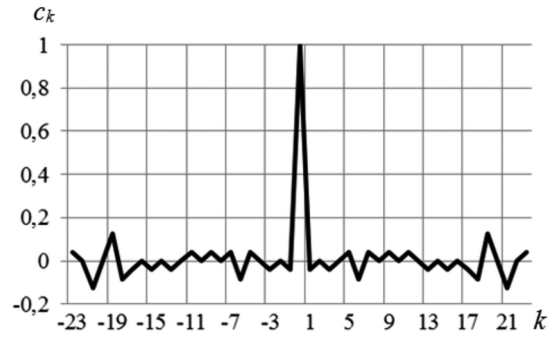


Рис. 1. Вид автокорреляционной функции

Как видно из рисунка, АКФ имеет максимальный лепесток, называемый главным, и ряд боковых лепестков меньшего уровня. Величина главного лепестка нормирована к единице.

Боковые лепестки АКФ являются паразитными, однако их величина не может быть сделана произвольно малой. Типовое значение имеет порядок 0,1 от величины главного максимума, что не удовлетворяет требованиям многих прикладных задач. Отступая от согласованности и увеличивая длину опорной последовательности $\mathbf{x}$, можно добиться снижения уровня боковых лепестков за счет уменьшения амплитуды главного лепестка и некоторого снижения выходного отношения сигнал/шум.

В случае аддитивного гауссовского шума со среднеквадратичным значением σ это отношение может быть вычислено по формуле

$$SNR(\mathbf{s}, \mathbf{x}) = \frac{|c_M|^2}{\sigma^2 \sum_{k \in \Omega} |x_k|^2}, \quad (3)$$

где M – индекс главного лепестка. Коэффициент уменьшения этого отношения по сравнению со случаем согласованной фильтрации называется энергетическими потерями, оценивается в децибелах и вычисляется по формуле

$$L = 10 \log[SNR(\mathbf{s}, \mathbf{s}) / SNR(\mathbf{s}, \mathbf{x})]. \quad (4)$$

Нетрудно видеть, что энергетические потери (4) не зависят от величины σ .

Для радиолокационных применений энергетические потери являются одним из главных параметров, и при выборе последовательности $\mathbf{x}$ необходимо следить, чтобы потери не превысили допустимую величину.

Существует несколько общепринятых критериев качества получаемой ВКФ. Одним из критериев является отношение модуля главного лепестка к модулю максимального бокового:

$$P = \min_{k \in \Omega, k \neq M} |c_M / c_k|. \quad (5)$$

В данной работе этот параметр для краткости будем называть уровнем подавления, или просто подавлением.

Другой критерий характеризует интегральный уровень боковых лепестков относительно главного:

$$Q = \frac{|c_M|^2}{\sum_{k \in \Omega, k \neq M} |c_k|^2}. \quad (6)$$

Этот параметр будем называть коэффициентом качества ВКФ по аналогии с общепринятым параметром АКФ, который в англоязычной литературе называется meritfactor.

Так как параметры (5) и (6) изменяются в широких пределах, то ниже они, как правило, вычисляются в децибелах: $20 \log P$ дБ, $10 \log Q$ дБ. В литературе часто используются параметры, обратные к (5) и (6), однако представленные формулы удобнее, позволяют оперировать с числами больше единицы, а при вычислении в децибелах дают положительные значения.

Задача совместного нахождения последовательностей s_0 и x , максимизирующих функции (5) и/или (6), исследуется на протяжении полувека, однако до сих пор не найдено регулярного способа ее решения. Основным методом, позволяющим решать подобные задачи, остается метод компьютерного перебора с использованием различных способов снижения вычислительной сложности [3, 4]. При этом объем вычислений, тем не менее, растет экспоненциально с увеличением длины последовательности, поэтому оптимальные решения найдены лишь для сравнительно небольших длин.

Задача существенно упрощается, если при нахождении последовательности x зафиксировать исходную последовательность s_0 . В литературе описано несколько методов решения подобных задач [5].

В данной работе анализируется один из таких методов, позволяющий вычислить оптимальную последовательность x , максимизирующую коэффициент качества (6) при заданной последовательности s_0 [5]. Несмотря на то, что сам метод известен, свойства получаемого решения исследованы слабо, что связано с отсутствием строгих закономерностей в получаемых результатах. В литературе представлены лишь примеры для нескольких исходных последовательностей s_0 [5, 6].

Целью данной работы является анализ характеристик результирующих ВКФ, полученных с помощью указанного метода, и выработка рекомендации по выбору исходной последовательности s_0 для получения максимального коэффициента качества.

Метод синтеза опорной последовательности

Рассмотрим метод, позволяющий при заданной последовательности s синтезировать последовательность x , которая максимизирует коэффициент качества (6) [5].

Подставляя (2) в (6), получаем формулу для коэффициента качества ВКФ:

$$Q = \frac{|c_M|^2}{\sum_{k \in \Omega, k \neq M} |x D_k s^*|^2} \quad (7)$$

Будем искать решение задачи при условии, что индекс максимального лепестка равен нулю: $M = 0$. Это условие не уменьшает общности, так как его всегда можно выполнить путем перераспределения нулей в последовательности s (1). Тогда из (2) следует, что выражение в числителе (7) равно $c_M = x s^*$. При этом формула (7) может быть выражена в виде

$$Q = \frac{|x s^*|^2}{x R x^*}, \quad (8)$$

где матрица R определяется выражением

$$R = \sum_{k \in \Omega, k \neq M} D_k s^* s D_k^* \quad (9)$$

Эта матрица является положительно определенной, т.е. $x R x^* > 0$ для всех комплексных векторов x .

Действительно, выражение $x R x^*$ равно сумме квадратов амплитуд боковых лепестков и, следовательно, не меньше нуля. Равным нулю оно также быть не может, так как это означало бы, что все лепестки нулевые, что невозможно. Из положительной определенности матрицы (9) следует, что существует единственный положительно определенный квадратный корень $R^{1/2}$ из матрицы R [7]. Кроме того, известно, что матрицы R^{-1} и $R^{-1/2}$ также являются положительно определенными, и что любая положительно определенная матрица U является эрмитовой, т.е. $U = U^*$ [7].

Представим числитель (8) в виде

$$|x s^*|^2 = |(x R^{1/2})(R^{-1/2} s^*)|^2 \quad (10)$$

Применяя к правой части (10) неравенство Коши-Буняковского и используя приведенные выше свойства положительно определенных матриц, получаем

$$|x s^*|^2 \leq (x R x^*)(s R^{-1} s^*) \quad (11)$$

Подставляя это выражение в (8), имеем

$$Q \leq s R^{-1} s^* \quad (12)$$

Максимальное значение (12) коэффициента качества достигается в случае, когда неравенство Коши-Буняковского (11) превращается в равенство. Известно, что это достигается при условии коллинеарности входящих в (10) векторов:

$$x R^{1/2} = a s R^{-1/2}, \quad (13)$$

где a – некоторый коэффициент. Так как масштабный коэффициент a в данной работе не существен, то, полагая его равным единице, из последнего соотношения получаем формулу для вычисления оптимальной последовательности x :

$$x = s R^{-1} \quad (14)$$

В дальнейших разделах рассматриваются свойства оптимального решения (14).

Выбор значения параметра m

При выводе формулы (14) индекс главного лепестка был принят равным нулю. С одной стороны от этого лепестка находятся $N - m - 1$ ненулевых членов ВКФ с отрицательными индексами, а с другой – $(n + m - 1)$ членов с положительными индексами. Поэтому, изменяя параметр m , можно передвигать положение главного лепестка по корреляционной функции.

Заметим, что величина параметра m в (1) может изменяться в пределах $m \in (0, 1, \dots, N - n)$ пока окно длины N содержит все члены последовательности s_0 .

В то же время, ВКФ (2) имеет большее число $N + n - 1$ ненулевых членов. Однако при индексах m вне указанного диапазона последовательность s_0 фактически укорачивается, и происходит анализ укороченной последовательности. Поэтому в данной работе такие значения параметра m не рассматриваются.

Проанализируем, как влияет параметр m на коэффициент качества. На рис. 2 представлен типичный график функции $Q(m)$ (6) в децибелах (сплошная линия) для последовательности $s_0 = 39C0A92$ с параметрами $n = 26$, $N - n = 52$. Пунктирной линией на рисунке показан график функции $P(m)$ (5).

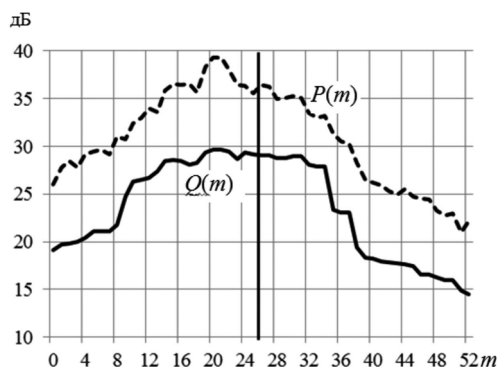


Рис. 2. Графики функций $Q(m)$ и $P(m)$

Отметим следующие характерные черты приведенных зависимостей. Обе кривые не являются монотонными и имеют волнообразный характер. Максимальные значения, как функции $Q(m)$, так и функции $P(m)$, достигаются в районе середины диапазона изменения параметра m , обозначенной на рис. 2 жирной вертикальной линией, однако эти максимумы несколько смещены относительно середины. Смещение максимумов можно объяснить разной структурой начала и конца последовательности s_0 .

Заметим также, что поведение функции $P(m)$ аналогично поведению функции $Q(m)$, при этом разность $P(m) - Q(m)$ составляет в среднем 7,5 дБ.

Еще одним параметром, важным для радиолокационных применений, являются потери L (4) в отношении сигнал/шум. Этот параметр слабо зависит от величины m . Для рассматриваемого примера потери находятся в диапазоне 0,4–0,57 дБ во всем анализируемом диапазоне изменения параметра m .

Чтобы уменьшить количество переменных параметров, от которых зависит оптимальное решение (14), в дальнейшем ограничимся его исследованием при параметре m в середине возможного диапазона. Получаемое при этом решение близко к наилучшему по параметру m . Тем не менее, при выборе последовательности для конкретных приложений рекомендуется исследовать полученное решение на оптимальность по параметру m , варьируя его в небольших пределах около середины.

Для дальнейших исследований положим, кроме того, что число нулей m в последовательности s (1) слева и справа от s_0 одинаково, т.е. величина $N - n$ четна. Это

условие не ограничивает получаемые результаты, так как коэффициент качества слабо зависит от увеличения параметра N на единицу. При указанных соглашениях параметры N , n , m связаны соотношением

$$N = n + 2m. \quad (15)$$

Выбор исходной последовательности s_0

Формула (14) дает способ вычисления оптимальной последовательности x при фиксированной исходной последовательности s_0 . Вопрос о выборе последней для получения наилучшего решения остается открытым. Поэтому в данной работе анализируются результаты, получаемые при использовании последовательностей из нескольких известных семейств в качестве исходных.

Сначала изложим несколько общих соображений. Известно, что при увеличении параметра m коэффициент качества монотонно увеличивается [5]. Начальное значение этого коэффициента определяется коэффициентом качества АКФ исходной последовательности s_0 . Поэтому логично предположить, что чем больше коэффициент качества АКФ последовательности s_0 , тем больше результирующий коэффициент качества ВКФ (12) для любого m .

Для проверки этого предположения проанализированы все возможные начальные последовательности для нескольких сравнительно коротких длин n в диапазоне 13–20. Для каждой длины вычислялись коэффициенты качества для всех возможных последовательностей как для АКФ, так и для оптимальных ВКФ, рассчитанных при использовании выражения (14), для нескольких значений параметра m . Результаты вычислений для всех указанных длин имеют одинаковый характер, рис. 3 иллюстрирует их для $n = 14$.

Графики этого рисунка построены следующим образом. Сначала проводилась сортировка всех возможных последовательностей данной длины по величине коэффициента качества АКФ. Для всех графиков рис. 3 по оси абсцисс отложен номер последовательности после такой сортировки. Левый верхний график показывает значения коэффициента качества АКФ этих последовательностей. Из него следует, что количество лучших, также как и худших, последовательностей весьма мало. Это подтверждает известный факт, состоящий в том, что поиск лучшей последовательности значительной длины — весьма трудоемкая в вычислительном плане задача.

На остальных графиках приведены аналогичные зависимости для оптимальных последовательностей x (14) при $m = 0,5n$; n ; $1,5n$. Белым цветом внутри залитых областей на этих рисунках показаны линии тренда.

Из сравнения графиков рис. 3 можно сделать следующие выводы. Во-первых, линии тренда примерно повторяют поведение функции, описывающей коэффициенты качества АКФ исходных последовательностей. Во-вторых, среднеквадратичное отклонение от линии тренда весьма значительно и растет с ростом параметра m . При одном и том же коэффициенте качества исходных последовательностей коэффициенты качества ВКФ могут отличаться на порядок и более.

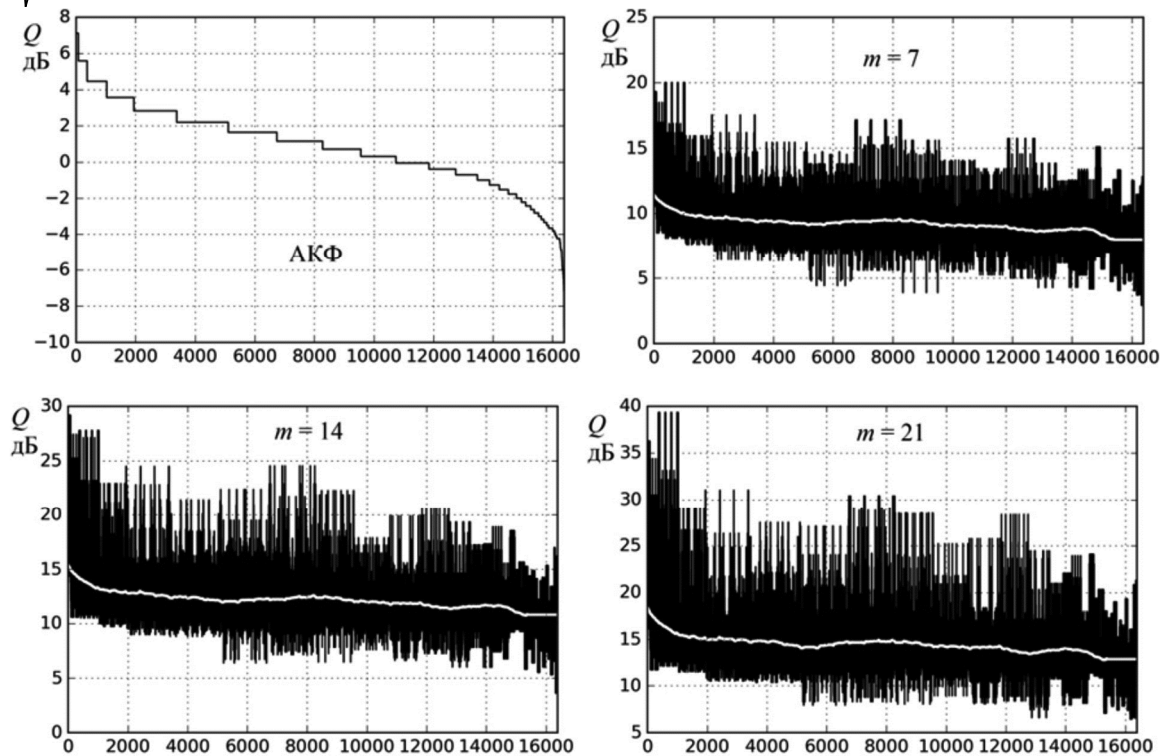


Рис. 3. Коэффициенты качества всех последовательностей длины 14:
слева сверху – для АКФ, далее – для ВКФ при $m = 7, 14$ и 21

Тем не менее, для всех значений параметра m максимальный коэффициент качества ВКФ реализуется для последовательностей с коэффициентом качества АКФ, близким к максимальному. Поэтому рассмотрим сначала параметры решения (14) для последовательностей с максимально возможным коэффициентом качества АКФ.

Параметры ВКФ для последовательностей с максимальным коэффициентом качества АКФ

Поиск последовательности с максимальным коэффициентом качества АКФ при заданной длине производится на протяжении полувека [4]. Результаты поиска можно найти, например, в [8], где перечислены лучшие известные последовательности для всех длин в диапазоне 4-271 за исключением длин 248, 260, 266. Последовательности до длины 60 найдены методом исчерпывающего перебора и, следовательно, являются наилучшими по коэффициенту качества АКФ. В данном разделе анализируются последовательности из этого перечня [8], причем для каждой длины оставлена одна последовательность, дающая наилучшие результаты при несогласованной фильтрации.

Один из основных вопросов состоит в том, насколько быстро растет коэффициент качества $Q(m)$ при увеличении параметра $m = 0,5(N - n)$. Для ответа на этот вопрос проанализированы параметры ВКФ, получаемых по изложенному методу, для всех последовательностей из указанного списка в диапазоне параметра $m = (0...4)n$. На рис. 4 представлены два типичных случая поведения функции $Q(m)$ (6): для последовательности 3FC08F1EEEDCAE2B13496 длины $n = 82$ (рис. 4 а) и последовательности 7FD556BC078D3B64B4CE66 дли-

ны $n = 87$ (рис. 4 б). На этом же рисунке приведены графики функции подавления $P(m)$ (5). По оси абсцисс на обоих графиках отложено отношение m/n .

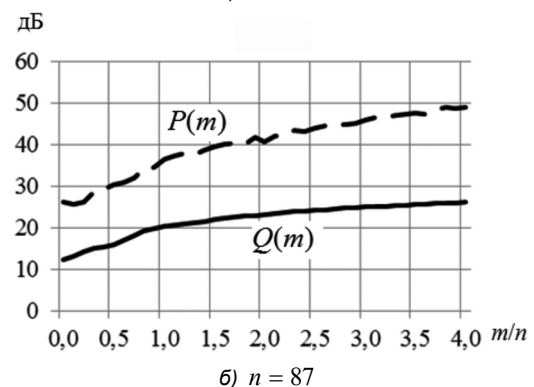
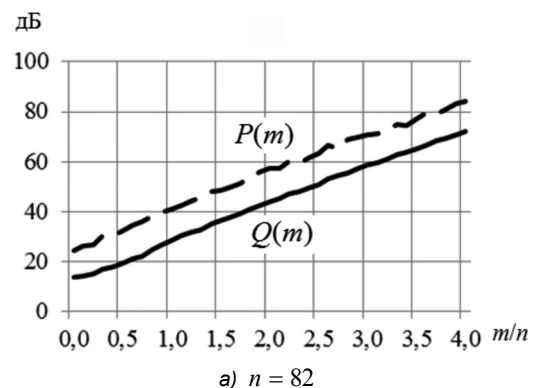
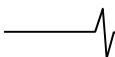


Рис. 4. Зависимости $Q(m)$ и $P(m)$

Зависимости рис. 4 а близки к линейным с небольшими колебаниями вокруг средней линии. Зависимости рис. 4б демонстрируют сильную нелинейность с постепенным уменьшением наклона. С точки зрения исследуемой задачи предпочтительными являются линейные



зависимости, так как они позволяют при меньшей длине N получить необходимый коэффициент качества. В анализируемом списке таких последовательностей примерно половина. Следовательно, для произвольно заданной длины в диапазоне $n = 4 - 271$ всегда можно найти близкую по длине последовательность с примерно линейной зависимостью $Q(m)$.

Для того, чтобы оценить скорость возрастания коэффициента качества $Q(m)$ при увеличении параметра m , для всех последовательностей из рассматриваемого списка рассчитаны коэффициенты аппроксимирующих прямых $a(m/n) + b$, оптимальных по критерию минимума среднеквадратичной ошибки. Хотя линейные функции хорошо аппроксимируют только зависимости вида рис. 4 а, коэффициент a может служить индикатором степени нелинейности функции.

Гистограмма распределения коэффициента a показана на рис. 5. Последовательности с коэффициентом a в диапазоне 2-5 демонстрируют, как правило, сильную нелинейность, в диапазоне 5-8 нелинейность слабее, при $a > 8$ – зависимости приближаются к линейным. Для функции $Q(m)$ рис. 4а коэффициенты равны $a = 15,1$, $b = 13,1$; а рис. 4 б – $a = 3,0$, $b = 16,0$. В первом случае это означает увеличение коэффициента качества ВКФ примерно на 15 дБ при увеличении отношения m/n на единицу. При $m/n = 4$ коэффициент качества достигает величины $Q = 72$ дБ, а уровень подавления – $P = 83$ дБ. Среднее значение наклона a аппроксимирующих прямых для последовательностей из рассматриваемого множества составляет $\bar{a} = 8,44$.

к-во

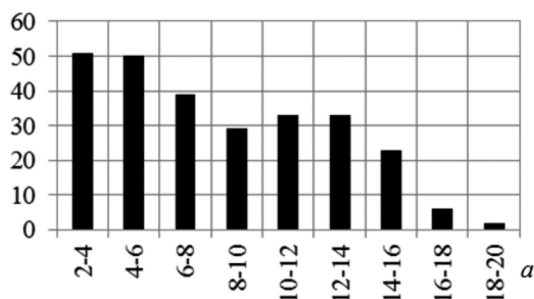


Рис. 5. Гистограмма наклона a аппроксимирующей прямой

Необходимо также отметить, что уже при $m = 0$ переход к несогласованной фильтрации улучшает коэффициент качества в среднем на 2,8 дБ, а для отдельных последовательностей – до 7 дБ.

Зависимость подавления $P(m)$ (5) по форме соответствует зависимости $Q(m)$. Если функция $Q(m)$ хорошо аппроксимируется прямой, то это же относится, как

правило, и к функции $P(m)$. При этом в среднем график функции $P(m)$ лежит на 10-15 дБ выше, а соответствующий наклон аппроксимирующей прямой на 1,2 дБ больше. При нелинейном характере $Q(m)$, функция $P(m)$ растет быстрее, чем $Q(m)$.

Следующий важный параметр для радиолокационных применений – это потери в отношении сигнал/шум (4). Так как критерий оптимизации (6) содержит в числителе уровень главного лепестка, то следует ожидать, что

оптимальная последовательность (14) не сильно отличается от согласованной, и поэтому потери будут невелики. Расчеты подтверждают эти ожидания. В качестве примера на рис. 6 показаны наложенные друг на друга исходная последовательность s при $s_0 = \text{CE0549}$ длины 24 (сплошная линия) и последовательность x , полученная по формуле (14) при $m = 12$ (линия из точек). По оси абсцисс отложены номера членов последовательностей, по оси ординат – их величина. Соседние точки на графике соединены прямыми линиями.

В области ненулевых членов s (номера точек 12-36) последовательности отличаются мало, а за пределами этой области элементы последовательности x невелики, что приводит к небольшим потерям.

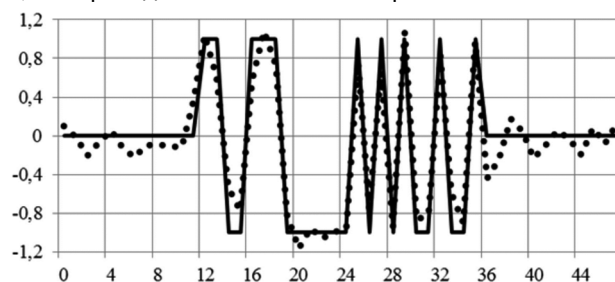


Рис. 6. Подобие последовательностей s и x

Для рассматриваемого семейства исходных последовательностей средняя величина потерь при $m = 0, n/2, n, 2n$ составляет соответственно 0,29; 0,55; 0,70; 0,85 дБ. Гистограмма распределения потерь показана на рис. 7 для $m = 0$ темными прямоугольниками и для $m = 2n$ светлыми прямоугольниками. При $m = 0 \dots 2n$ подавляющее большинство последовательностей имеют потери в диапазоне 0-1,25 дБ и лишь единичные – порядка 2-3 дБ. Указанные величины сравнительно невелики и являются адекватной платой за снижение уровня боковых лепестков. Заметим также, что приведенные величины потерь меньше, чем потери при введении весовых функций для подавления боковых лепестков при спектральном анализе [1, 2].

к-во

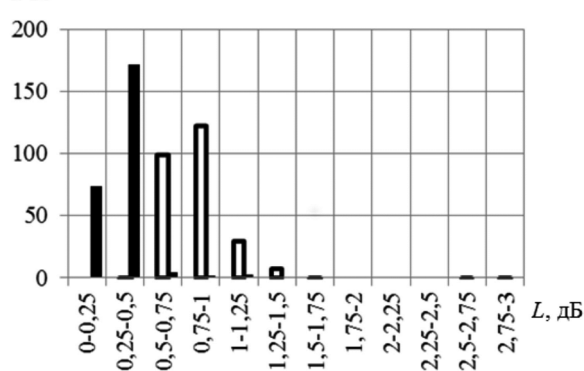


Рис. 7. Гистограмма распределения потерь

В табл. 1 представлены характеристики оптимального решения (14) для последовательностей из рассматриваемого перечня, имеющих длину $n \geq 13$ и коэффициент наклона аппроксимирующей прямой $a \geq 14$ для функции $Q(m)$. Для каждой последовательности приведены длина n , ее представление в шестнадцатеричном виде, коэффициент качества $Q(m)$ при согласованной

Таблица 1. Параметры ВКФ для лучших последовательностей с максимальным коэффициентом качества АКФ

n	s ₀	Q, дБ, при m=					P, дБ, при m=					L, дБ, m=		a	
		СФ*	0	n/2	n	2n	СФ*	0	n/2	n	2n	0	2n	Q	P
13	1f35	11,49	15,68	19,46	30,03	49,60	22,28	24,00	27,65	38,47	58,59	0,14	0,21	18,8	18,8
14	3f35	7,12	14,25	19,33	29,21	44,01	16,90	19,77	25,74	35,75	50,88	0,64	0,78	15,9	15,6
15	7cca	8,75	13,59	19,88	29,81	45,11	13,98	23,80	28,66	37,80	53,75	0,37	0,60	16,7	17,1
17	18329	6,55	10,85	16,25	24,90	38,58	15,07	17,65	24,32	34,32	47,14	0,82	1,14	14,6	15,1
26	39c0a92	8,76	13,22	17,80	29,06	46,46	18,76	21,99	26,51	36,43	56,77	0,33	0,57	17,5	17,3
29	1ce02a4d	8,31	11,19	16,32	28,16	44,33	19,71	19,96	26,70	37,80	55,72	0,29	0,69	15,8	16,5
38	3fc3cb5599	9,19	13,58	18,87	28,41	42,07	22,05	23,09	26,08	40,32	52,19	0,39	0,68	14,6	14,6
39	7f92d8f1aa	8,85	12,88	20,65	31,17	47,68	17,84	23,56	31,93	42,05	58,33	0,40	0,71	17,5	17,7
45	1fe4b6270e55	9,34	12,45	19,64	27,63	43,26	19,08	24,00	32,15	42,19	55,41	0,30	0,63	14,8	15,0
46	3fce1d896caa	9,07	12,45	20,08	28,91	42,72	21,21	23,51	30,93	38,78	54,04	0,29	0,57	14,3	14,2
47	7fce1d896caa	9,13	12,61	19,80	27,91	42,69	19,46	22,85	31,99	40,05	56,70	0,30	0,64	15,0	15,3
55	7fc9392f1b1caa	9,47	12,12	19,52	27,92	44,01	17,91	21,77	31,34	40,77	58,12	0,22	0,61	16,1	16,3
59	7f0184c34c5952a	9,29	13,29	19,23	28,56	42,78	21,44	26,52	30,71	42,69	55,40	0,52	0,95	14,9	15,2
67	7ff09ce1e96c9d2aa	9,69	13,28	20,32	28,52	44,44	19,62	22,82	33,83	41,66	57,32	0,39	0,85	15,2	15,5
77	1f6fe16cb3730c7a5475	9,18	11,49	17,60	24,75	39,36	16,90	21,19	28,01	37,34	52,52	0,23	0,76	14,0	14,9
82	3fc08f1eedcae2b13496	9,50	13,47	19,42	28,68	43,94	24,30	24,53	32,44	40,84	57,43	0,29	0,58	15,1	14,9
95	7319c1fe828fadf5ea949932	9,74	13,82	18,36	28,43	43,12	20,47	27,31	31,09	41,51	56,65	0,33	0,54	14,7	15,0
99	7cf8364ffe59e984aac635aca	9,29	12,93	19,16	29,24	43,10	20,83	25,88	33,86	43,22	58,45	0,30	0,64	14,0	14,4
103	61f3f099bafacfb99d52b296	9,80	13,58	19,93	28,86	43,45	23,35	26,83	34,44	44,36	57,32	0,27	0,52	15,0	15,2
104	e1f3f099bafacfb99d52b296	9,46	13,19	19,34	28,18	42,02	22,28	25,93	31,60	41,86	55,43	0,29	0,58	14,2	14,6
129	1c63e319ca1fdffadc155d5a0d9b2526d	9,56	12,93	19,18	28,78	44,20	21,38	25,84	32,70	43,57	58,45	0,29	0,62	15,3	15,7
130	3c63e319ca1fdffadc155d5a0d9b2526d	9,51	12,75	18,84	28,78	43,92	22,28	25,85	32,66	41,93	56,97	0,29	0,65	15,8	15,7
134	3f8ce603e0a09edbcb8e9d7d6b566cdab	9,27	13,55	20,17	30,63	45,25	18,46	27,77	33,13	45,64	58,09	0,41	0,65	15,4	15,6
135	6ad9b356b5f5cb8e99edbc8283e03398fe	9,48	13,86	19,73	30,74	45,74	19,08	26,71	33,23	46,24	60,70	0,39	0,60	16,2	16,5
140	ff01fdb484843167199261345c5c38a952a	9,20	12,38	20,74	29,92	44,30	20,00	26,13	35,51	46,03	59,54	0,46	0,99	14,0	14,5
141	1fe03fb69090862ce3324c268b8b87152a55	9,46	12,75	20,79	30,16	45,71	19,46	25,04	36,20	46,93	63,03	0,40	0,83	15,0	16,0
146	3fffc03784e58e4b49998f0e49e4e972ad555	9,23	12,34	18,54	27,82	42,11	23,29	24,94	32,62	41,23	56,60	0,32	0,79	14,9	14,6
158	3f0600f83071ff993cc57ecd39955b6b294aa6b5	9,30	13,59	19,35	28,96	43,32	23,97	27,31	29,80	43,81	59,40	0,34	0,61	14,4	15,2
169	1c1c7c623b8eb1fd05dafdd41debd5b0491226d6dad	9,69	12,63	17,96	26,69	41,67	23,73	28,29	34,44	42,12	57,61	0,24	0,57	14,8	15,0

*СФ – согласованная фильтрация.

(СФ) и несогласованной фильтрации с параметрами $m=0, n/2, n, 2n$, уровень подавления $P(m)$ для тех же параметров, потери L в отношении сигнал/шум, а также коэффициенты a наклона прямых, аппроксимирующих функции $Q(m)$ и $P(m)$.

Для $n \geq 170$ не имеется ни одной последовательности с коэффициентом $a \geq 14$. Это может быть связано либо с тем, что при $n > 60$ наилучшие последовательности не найдены, либо с тем, что максимальный коэффициент качества ВКФ ухудшается с ростом длины. Какой вариант имеет место в действительности, на сегодняшний день неизвестно.

Параметры ВКФ для М-последовательностей

Наиболее известным семейством псевдослучайных последовательностей, широко используемых в радиолокации для фазокодовой манипуляции, являются рекуррентные последовательности максимальной длины, или М-последовательности [1, 2]. М-последовательность

$a = (a_0, a_1, \dots, a_{n-1})$ из элементов 0, 1 длины $n = 2^k - 1$, $k = 2, 3, \dots$, генерируется по формуле

$$a_{i+k} = \left(\sum_{j=0}^{k-1} a_{i+j} p_j \right) \bmod 2, \quad i \geq k, \quad (16)$$

где $p_j, j = 0, 1, \dots, k-1$, – коэффициенты двоичного при-

митивного полинома $\sum_{j=0}^k p_j x^j$ степени k , а ненулевой

начальный отрезок последовательности $a_0, a_1, \dots, a_{k-1}$ задается произвольно. Таблица примитивных полиномов может быть найдена в [9]. Для модуляции радиолокационных сигналов используются как полные, так и

усеченные М-последовательности, при этом нулевые элементы заменяются на -1.

Асимптотический коэффициент качества АКФ М-последовательностей при $n \rightarrow \infty$ равен 4,77 дБ (3 в линейном масштабе) [4], что существенно меньше, чем у последовательностей предыдущего раздела.

Количество М-последовательностей, включая усеченные, для любой длины весьма велико. Поэтому для анализа получаемых ВКФ при использовании оптимальных решений (14) выбраны три подмножества таких последовательностей:

– М-последовательности длины 63 с минимальным уровнем максимального бокового лепестка; их количество равно 52, подробные данные могут быть найдены в [2];

– последовательности длины 70, получаемые усечением М-последовательностей длины 127, порожденных полиномом $x^7 + x + 1$ (127 штук);

– М-последовательности длины 2047, порожденные полиномами $x^{11} + x^2 + 1$ и $x^{11} + x^9 + x^2 + x + 1$ (по 25 штук для каждого полинома).

Средние значения основных параметров ВКФ для указанных наборов приведены в табл. 2 наряду с параметрами других рассматриваемых в данной работе семейств последовательностей. Обозначения параметров в таблице описаны в предыдущих разделах. Черта над параметром означает усреднение.

Анализ таблицы показывает следующее. По всем рассматриваемым параметрам М-последовательности уступают последовательностям предыдущего пункта (меньше параметры $Q(m)$ и $P(m)$, больше энергетические потери, меньше скорость роста a и т.д.). Причем различные виды полных и усеченных М-последовательностей имеют значительный разброс обсуждаемых пара-

метров. Характеристики усеченных последовательностей, как правило, существенно хуже, чем полных, а степень деградации тем больше, чем значительнее усечение. При всех длинах функции $Q(m)$ и $P(m)$ демонстрируют значительную степень нелинейности. Заметим также, что при $n=2047$ функция $\bar{P}(m)$ не является монотонной. Таким образом, использование M -последовательностей в рассматриваемом методе несогласованной фильтрации неэффективно.

Параметры ВКФ для последовательностей Лежандра

Последовательности Лежандра [4] представляют собой бесконечное семейство, имеющее один из наиболее высоких асимптотических коэффициентов качества АКФ, равный 7,78 дБ (6 в линейном масштабе). Для сравнения укажем, что рассмотренные выше последовательности с максимальным коэффициентом качества АКФ имеют этот параметр в среднем немногим более 9 дБ, а при длинах $4 \leq n \leq 60$, для которых приведенные последовательности являются наилучшими, – в среднем 8,64 дБ (7,31 в линейном масштабе). Таким образом, последовательности Лежандра имеют довольно высокий коэффициент качества.

Последовательность Лежандра $\mathbf{z} = (z_0, z_1, \dots, z_{n-1})$ определена для простых длин $n=p$ и может быть построена следующим образом [4]. Символ z_i равен +1, если сравнение $x^2 \equiv i \pmod{p}$ имеет решение, и равен -1 в противном случае, $i=1, 2, \dots, p-1$. Нулевой символ полагается равным +1. Последовательность s_0 с высоким коэффициентом качества АКФ, используемая в качестве исходной в (14), получается циклическим сдвигом последовательности $\mathbf{z}$ на $\lfloor rp \rfloor$ символов, где $r \in (0,25; 0,75)$, а выражение $\lfloor x \rfloor$ означает наибольшее целое, не превосходящее x . Таким образом

$$s_{i \bmod p} = z_{(i + \lfloor rp \rfloor) \bmod p} \quad (17)$$

Последовательности (17) также будем называть последовательностями Лежандра.

Для анализа получаемых ВКФ этих последовательностей и оптимальных решений (14) выбраны два подмножества последовательностей (17):

- последовательности с длинами в диапазоне 4-271 для сравнения с результатами предыдущих пунктов; количество таких последовательностей равно 110;
- последовательности с длинами в диапазоне 2000-2200 для оценки асимптотических характеристик; их количество равно 48.

Усредненные параметры этих подсемейств приведены в табл. 2. Последовательности Лежандра (17) по всем параметрам занимают промежуточное положение между последовательностями с наилучшей АКФ и M -последовательностями. Для больших длин последовательности Лежандра являются хорошим выбором. Аналогичные характеристики имеют последовательности Якоби [4] (асимптотический коэффициент качества АКФ равен 6). Можно также использовать расширенные последовательности Лежандра с наиболее высоким известным асимптотическим коэффициентом качества АКФ 8,02 дБ (6,34 в линейном масштабе) [10].

Заключение

Рассмотренный метод синтеза оптимальной опорной последовательности для фазокодоманипулированного сигнала, позволяющий минимизировать интегральный уровень боковых лепестков, является аналитически простым и без труда реализуется на персональном компьютере. Для большого количества сигналов при увеличении длины опорной последовательности в 3-5 раз

($m = n \dots 2n$) достигается увеличение коэффициента качества до 25-45 дБ и уровня подавления до 45-60 дБ при потерях в отношении сигнал/шум менее 1 дБ.

В качестве исходной последовательности для рас-

Таблица 2. Средние значения параметров ВКФ для различных семейств последовательностей

Параметр ВКФ	m	Пос-ти с лучшей АКФ, $n = 4 \dots 271$	М-пос-ти, $n = 63$, с максимальным P для АКФ	М-пос-ти, $n = 70$	М-пос-ти, $n = 2047$	Лежандра, $n = 4 \dots 271$	Лежандра $n = 2000 \dots 2200$
$\bar{Q}(m)$, дБ	СФ*	9,1	5,6	2,0	4,8	6,8	7,8
	$m = 0$	11,9	7,8	5,4	7,2	9,4	10,4
	$m = n/2$	16,5	10,9	9,3	10,5	14,6	16,2
	$m = n$	22,8	14,5	12,5	14,0	19,4	21,1
	$m = 2n$	31,3	18,4	16,4	17,6	26,2	28,1
$\bar{P}(m)$, дБ	СФ*	21,9	20,4	15,5	32,0	21,2	33,9
	$m = 0$	24,6	20,5	17,2	31,5	22,4	34,4
	$m = n/2$	29,6	22,1	22,1	21,9	28,2	40,1
	$m = n$	37,6	28,4	26,2	30,2	33,8	42,2
	$m = 2n$	47,1	33,3	31,7	37,3	41,7	53,1
$\bar{L}(m)$, дБ	$m = 0$	0,29	0,47	2,0	0,51	0,44	0,30
	$m = n/2$	0,55	0,93	3,0	0,93	1,00	0,75
	$m = n$	0,70	1,48	3,7	1,40	1,32	0,88
	$m = 2n$	0,85	2,17	4,6	2,03	1,84	1,11
$\bar{a}$ для Q	–	8,4	3,8	3,7	5,2	6,3	6,0
$\bar{a}$ для P	–	9,6	5,0	5,0	4,4	7,7	7,5

*СФ – согласованная фильтрация

смаатриваемого метода целесообразно выбирать последовательности с максимальным коэффициентом качества автокорреляционной функции. Для длин $n < 300$ такие последовательности приведены в литературе. Для больших длин целесообразно использовать последовательности Лежандра или Якоби.

Увеличение длины опорной последовательности приводит к некоторому возрастанию объема вычислений при фильтрации сигнала. Тем не менее, реализация рассмотренного метода в реальном масштабе времени на современных сигнальных процессорах не вызывает затруднений.

Таким образом, рассмотренный метод несогласованной фильтрации дает эффективный способ подавления боковых лепестков не только при малых длинах последовательности, но, вопреки распространенному мнению, и при больших длинах.

Литература

1. Levanon N., Mozeson E. Radar Signals. JohnWiley& SonsInc. NewJersey.2004.412 p.
2. Трухачев А.А. Радиолокационные сигналы и их применения. М. Воениздат. 2005. 320 с.
3. Nasrabadi M.A., Bastani M.H. A Survey on the Design

of Binary Pulse Compression Codes with Low Autocorrelation. Trends in Telecommunications Technologies. Christos J Bouras (Ed.). ISBN: 978-953-307-072-8. InTech.2010.

4. Jedwab J. A Survey of the Merit Factor Problem for Binary Sequences. Sequences and Their Applications. Proceedings of SETA 2004.ed. T. Hellesteth et al. Lecture Notes in Computer Science.vol. 3486. pp. 30–55. Springer-Verlag. Berlin Heidelberg. 2005.

5. P. Stoica, J. Li, and M. Xue. Transmit codes and receive filters for radar. IEEE Signal Processing Magazine. vol. 25. pp. 94–109. November 2008.

6. Levanon N. Cross-Correlation of Long Binary Signals with Longer Mismatched Filters. IEE Proceedings. Radar, Sonar, and Navigation.vol. 152. pp. 377–382. Dec. 2005.

7. Horn R.A. and Johnson C.R. Matrix Analysis. Cambridge University Press.1985. 652 p.

8. Knauer J. Merit Factor Records. Nov. 2004. Available in oct. 2016 at URL: http://labraj.feri.um.si/en/Low-Autocorrelation_Binary_Sequence_Problem.

9. Питерсон У., Уэлдон Э. Коды, исправляющие ошибки. М. Мир. 1976. 594 с.

10. Jedwab J., Katz D.J., Schmidt K.U. Advances in the Merit Factor Problem for Binary Sequences. Journal of Combinatorial Theory. Volume 120. No 4. pp. 882-906. May 2013.

ОСОБЕННОСТИ ВЫЧИСЛЕНИЯ И ОБРАБОТКИ СИГНАЛОВ ОСТАТКА В БЕЗКОНДЕНСАТОРНЫХ АЦП С СЕКЦИОНИРОВАННОЙ АРХИТЕКТУРОЙ

Кононов В.С., к.т.н., АО «Специализированное конструкторско-технологическое бюро электронных систем», e-mail: casandra1983@mail.ru.

ESTIMATION AND PROCESSING OF RESIDUE SIGNALS IN PIPELINED ANALOG-TO-DIGITAL CONVERTERS WITHOUT SWITCHED-CAPACITORS

Kononov V.S.

Technique for evaluation and processing of residue signals without classic sample and hold device is described. It is shown, that usage of RC-circuits with fixed delay for analogue and clock signals with compensation of integration effect, provides acceptable precision of evaluation of residue signals and removes classic limitation on frequency range.

Keywords: processing, architecture, capacitor, delay, residue, range.

Ключевые слова: обработка, архитектура, конденсатор, задержка, остаток, диапазон.

Введение

В конденсаторных АЦП с секционированной архитектурой [1] (рис. 1) сигнал остатка вычисляется как разность между сохраненным аналоговым сигналом и грубым цифровым приближением этого сигнала, полученным в результате последовательного преобразования сначала в секционном N -разрядном АЦП, а затем в секционном ЦАП аналоговой разрядности.

Сохранение аналогового сигнала в процессе грубого преобразования осуществляется в УВХ в виде заряда на конденсаторе C_{xp} , который заряжается при выборке через открытый МОП-транзистор (рис. 2).

Такой подход является достаточно эффективным и удобным при создании АЦП для приложений, не предназначенных для эксплуатации в «жестких» условиях. В последнем случае, как показал опыт эксплуатации АЦП с переключаемыми конденсаторами, наблюдается заметное ухудшение точностных характеристик, которые существенно деградируют с течением времени.

Одной из основных причин деградации точностных характеристик таких АЦП является утечка тока в стоковом рп-переходе ключевого МОП-транзистора (см. рис. 2), которая непосредственно влияет на зарядовое состояние конденсатора C_{xp} . Предотвратить возникновение и нарастание утечки тока не представляется возможным, так как этот процесс из-за влияния эффектов

Описан способ вычисления и обработки сигналов остатка без использования классического устройства выборки/хранения. Показано, что применение RC-линий фиксированной задержки аналоговых и тактовых сигналов с нивелированием влияния эффекта интегрирования на амплитуду последних обеспечивает приемлемую точность вычисления остатка и снимает классическое ограничение на ширину диапазона частот преобразования.

старения и воздействия условий эксплуатации является необратимым. В этом смысле УВХ можно считать наиболее уязвимыми устройствами в АЦП и поэтому их использование нежелательно, несмотря на очевидную значимость для конденсаторных АЦП.

Целью данной статьи является разработка способа вычисления и обработки сигналов остатка без использования УВХ для сохранения аналоговых сигналов.

В общем виде техника вычисления и обработки сигналов остатка уже рассматривалась в [2], где отмечалась необходимость использования двух линий задержки в цепях передачи аналогового и тактового сигналов. Забегая вперед, отметим, что при более детальном анализе этой техники потребовалось преодолеть определенные трудности, чтобы сохранить базовый принцип работы безконденсаторного АЦП. Согласно этому принципу такой АЦП должен работать в широком диапазоне частот преобразования – от нескольких килогерц до нескольких гигагерц (в нашем случае до 1 ГГц), что недоступно ни одному из известных конденсаторных АЦП.

Практическая реализация схемы вычисления и обработки сигналов остатка показана на рис. 3.

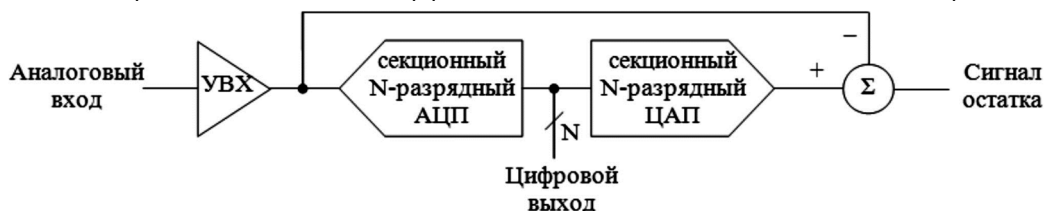


Рис. 1. N -разрядная секция конденсаторного АЦП: УВХ – устройство выборки/хранения

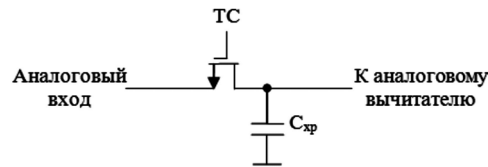


Рис. 2. Структура УВХ с переключаемым конденсатором: TC – тактовый сигнал

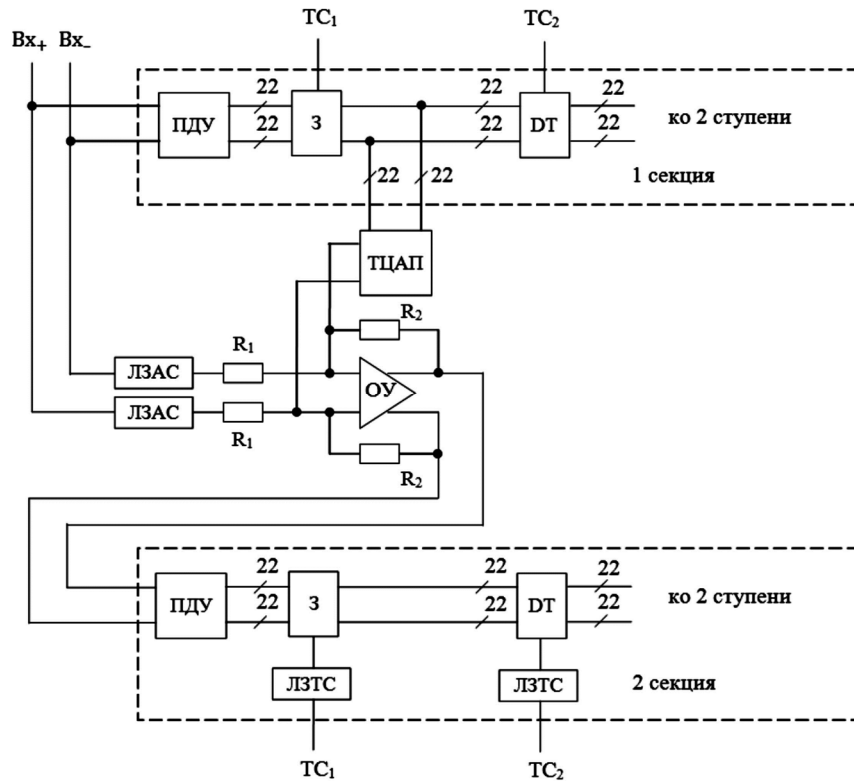


Рис. 3. Схема вычисления и обработки сигналов остатка:

ПДУ – предварительный усилитель; 3 – защелка; DT – D-триггер; ТЦАП – термометровый ЦАП; ОУ – операционный усилитель; ЛЗАС, ЛЗТС – линии задержки аналоговых (V_{x+} , V_{x-}) и тактовых (TC_1 , TC_2) сигналов

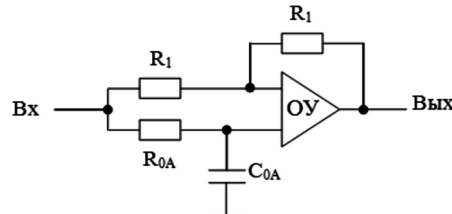


Рис. 4. Линия задержки аналогового сигнала

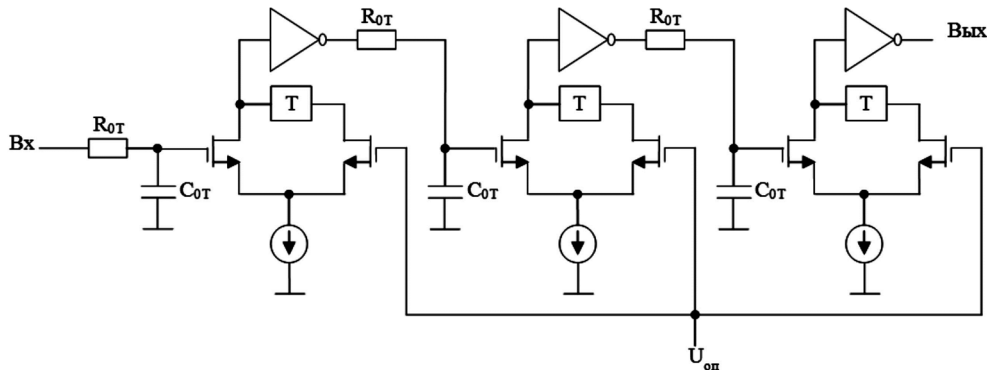
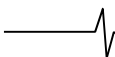


Рис. 5. Линия задержки тактовых сигналов TC_1 , TC_2 :

T – триггер; $U_{оп}$ – опорное напряжение; $R_{0T} \approx 0,3 \cdot R_{0A}$, $C_{0T} \approx C_{0A}$ (R_{0A} и C_{0A} – резистор и конденсатор на рис. 4)

Линии задержки ЛЗАС, ЛЗТС выполнены по различным схемам (рис. 4, 5), но имеют одинаковые времена задержки сигналов, сопоставимые с временем восстановления преобразованного сигнала на выходе ЦАП.

Строго говоря, время задержки ЛЗТС должно превышать время задержки ЛЗАС на величину, равную времени суммарной задержки ОУ и ПДУ на входе второй секции АЦП.



По результатам моделирования схемы (рис. 3) в КМОП-исполнении с проектными нормами 0,18 мкм время этой задержки составило около 3 нс. В результате при произвольной длительности тактовых сигналов TC_1 , TC_2 , которая может быть как больше, так и меньше полученного времени задержки, возникли определенные трудности при сохранении постоянной амплитуды задержанных тактовых сигналов из-за влияния эффекта интегрирования (рис. 6).

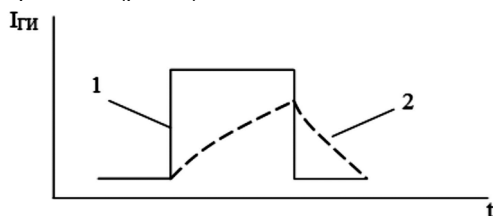


Рис. 6. Иллюстрация влияния эффекта интегрирования на амплитуду задержанного тактового сигнала при $\tau_d < \tau_{гн}$:

1 – исходный тактовый сигнал длительностью $\tau_{гн}$;

2 – задержанный тактовый сигнал;

τ_d – необходимое время задержки

Чтобы избежать влияния этого эффекта, фиксированная задержка набиралась с помощью трех одинаковых каскадов (см. рис. 5), каждый из которых обеспечивает время задержки около 1 нс, что соответствует минимальной длительности тактовых сигналов TC_1 , TC_2 .

При создании линии задержки для аналоговых сигналов (ЛЗАС) подобные трудности не возникали. Поэтому электрическая схема ЛЗАС (рис. 4) оказалась достаточно простой [3].

Завершая проведенный анализ техники вычисления и обработки сигналов остатка, необходимо отметить, что применение фиксированной задержки аналоговых и тактовых сигналов не приводит к ограничению длительности и периода тактовых сигналов. При этом точность установления задержанных аналоговых сигналов на входе ОУ оказалась по результатам моделирования соизмеримой с точностью установления аналоговых сигналов, сохраняемых на переключаемых конденсаторах в классических УВХ. Данное наблюдение позволило

отказаться от использования переключаемых конденсаторов, которые по сравнению с обычными RC-элементами несоизмеримо больше подвержены влиянию эффектов старения и воздействию «жестких» условий эксплуатации.

Выводы

Проведенный анализ позволяет сделать следующие выводы:

1. Применение конденсаторов хранения в классических УВХ является эффективным и удобным способом, обеспечивающим вычисление сигнала остатка, однако этот способ мало пригоден для специальных приложений, предназначенных для эксплуатации в «жестких» условиях.

2. Применение RC-линий задержки аналоговых и тактовых сигналов с нивелированием влияния эффекта интегрирования на амплитуду последних позволяет отказаться от использования классических УВХ с конденсаторами хранения аналоговых сигналов, которые из-за утечек тока ограничивают возможность работы АЦП в широком диапазоне частот преобразования.

3. Предложенный способ вычисления и обработки сигналов остатка обеспечивает приемлемую точность установления входного и восстановленного аналоговых сигналов на входах межсекционного операционного усилителя, что, благодаря использованию линий фиксированной задержки сигналов, не приводит к ограничению диапазона частот преобразования.

Литература

1. Кестер У. Аналого-цифровое преобразование [Текст]: монография / У. Кестер; пер. с англ. Е.Б. Володина. – М.: Техносфера, 2007. – 1016 с.
2. Кононов В.С. Архитектуры безконденсаторных КМОП-КНИ-АЦП для космического применения [Текст] / В.С. Кононов // Вестник Воронежского гос. технического университета. – 2016. – Т. 12. – №1, С. 70-73.
3. van Valkenburg M.E. Analog Filter Design [Text] / M.E. van Valkenburg – Oxford University Press, 1995. – 608 p.