

УДК 621.391.13

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ПРОЕКТИРОВАНИЯ СИСТЕМ СВЯЗИ С ВЫСОКОЙ ЭНЕРГЕТИЧЕСКОЙ ЭФФЕКТИВНОСТЬЮ

Быховский М.А., д.т.н., профессор Московского технического университета связи и информатики, e-mail: bykhmark@gmail.com

THEORETICAL FUNDAMENTALS OF DESIGN OF TELECOMMUNICATION SYSTEMS WITH HIGH ENERGY EFFICIENCY

Bykhovskiy M.A.

The author determines the borders of telecommunication systems' energy efficiency, which allow for comparison of values of the main parameters of specific telecommunication systems (marginal speed of message transmission and signal-to-noise per bit ratio) with potentially possible values of these parameters. It is shown that when designing a telecommunication system with high energy efficiency, it is advisable to choose optimal modulation methods for message transmission – multidimensional volume-spherical or surface-spherical signal ensembles. The article discusses the issues of parameter selection of such signal ensembles that allow to approach Shannon's border the spectral efficiency of a system being designed. In addition, the work discusses the issues of such noise immunity code selection for a designed system, that allow simplification of the communication system through the usage of the simplest, in the technical sense, modulator/demodulator, as well as coder and decoder.

Key words: energy efficiency of telecommunication systems, Shannon's border, multidimensional signal ensemble, noise immunity code, telecommunication systems design.

Ключевые слова: энергетическая эффективность систем связи, границы Шеннона, многомерный ансамбль сигналов, помехоустойчивый код, проектирование систем связи.

Введение

Основные параметры ансамблей сигналов для телекоммуникационных систем должны выбираться так, чтобы проектируемая система связи обеспечивала требуемую надежность приема сообщений, определяемую вероятностью ошибки $P_{ser}(N, R_f, \rho_s)$ при демодуляции сигналов, передаваемых по каналу связи, и необходимую удельную скорость (R_f) передачи сообщений при заданной полосе частот (F) канала связи. Система связи является энергетически эффективной, если в ней требуемые характеристики (надежность связи и удельная скорость передачи сообщений) достигаются при отношении сигнал/шум на входе демодулятора (ρ_s), близком к минимально возможному, который определен для «идеальной» по Шеннону [1] системы связи. Она должна быть построена таким образом, чтобы эти требования удовлетворялись бы наиболее экономичным образом, т.е. используемые в системе передатчики должны иметь минимально возможную мощность (такие передатчики дешевле и создают меньше помех другим системам), а также, чтобы основные устройства (кодеры, модуляторы, демодуляторы и декодеры) были бы по возможности наиболее простыми в технической реализации. При проектировании спектрально-эффективной системы связи особое значение имеет выбор для нее метода модуляции (ансамбля сигналов, используемого для передачи сообщений) и помехоустойчивого кода.

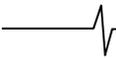
Определены границы энергетической эффективности систем связи, пригодные для сравнения значений основных параметров конкретных систем связи (удельной скорости передачи сообщений и отношения сигнал/шум, приходящегося на один принятый бит) с потенциально возможными значениями этих параметров. Показано, что при проектировании системы связи, обладающих высокой эффективностью, целесообразно для передачи сообщений выбирать оптимальные методы модуляции – многомерные объемно-сферические или поверхностно-сферические ансамбли сигналов. Обсуждаются вопросы выбора параметров таких ансамблей сигналов, которые позволяют приблизить спектральную эффективность проектируемой системы связи к границе Шеннона. Кроме того, обсуждаются вопросы выбора помехоустойчивого кода для проектируемой системы, позволяющего упростить систему связи за счет применения в ней наиболее простых в техническом отношении модулятора и демодулятора, а также кодера и декодера.

В фундаментальных работах Шеннона [1, 2] были установлены предельные возможности снижения энергетики линии связи, зависящей от ρ_{s0} – отношения сигнал/шум на входе демодулятора ($\rho_{s0} = \rho_{b0} R_f$, где ρ_{b0} – отношение сигнал/шум на входе демодулятора, приходящееся на один принятый бит), при котором по каналу связи можно передавать сообщения со скоростью, равной R_f . Шеннон установил знаменитую формулу

$$R_f = \log_2(1 + \rho_{b0} R_f), \text{ из которой следует}$$

$$\rho_{b0} = (2^{R_f} - 1) / R_f. \quad (1)$$

В (1) R_f – максимально возможная удельная скорость передачи сообщений по каналу связи, равная его пропускной способности. Следует отметить что формула (1) справедлива для предельного случая, когда для пе-



редачи сообщений используются N -мерные ансамбли сигналов (АС) большой размерности ($N = 2 \text{int}(FT) \gg 1$, $\text{int}(x)$ – целая часть числа x , T – длительность передаваемых сигналов, причем нормированная длительность сигналов (FT) должна быть очень большой). Шеннон показал, что в предельном случае можно обеспечить абсолютную надежность передачи сообщений, в которой вероятность ошибки ($P_{ser}(N, R_f, \rho_s)$) при демодуляции сигналов равна $P_{ser}(N, R_f, \rho_s) \rightarrow 0$ (при $N \rightarrow \infty$). Шеннон назвал «идеальной» [1] систему связи, которая позволяет, в принципе, передавать сообщения без ошибок.

При проектировании конкретной системы связи рациональный выбор ее параметров может быть сделан на основе сравнения выбранных параметров с потенциально возможными. Важность такого сравнения подчеркивал [3] В.А. Котельников.

Корректное сравнение разных систем связи подразумевает, что исследуемые системы обеспечивают одинаковые скорости передачи сообщений и одинаковую надежность их приема. В конкретных системах связи всегда длительность сигналов реализуемого ансамбля и надежность приема сообщений имеют определенные значения $P_{ser}(N, R_f, \rho_s) \neq 0$ и $(FT) \neq \infty$. Поэтому сравнение их данных с параметрами «идеальной» системы связи, для которой $P_{ser}(N, R_f, \rho_s) = 0$ и $(FT) = \infty$, не может привести к содержательному заключению об эффективности параметров конкретной системы, сделанному при ее проектировании.

Тем не менее, во многих работах (например, [4-7]) предлагалось проводить сравнение необходимого отношения сигнал/шум ρ_{s0} и скорости передачи сообщений R_f , которые должны быть обеспечены в конкретной разрабатываемой системе связи, с теми, которые имеет «идеальная» система Шеннона. При таком сравнении не учитывались длительность сигналов, входящих в выбранный для передачи сообщений АС, и длина использованного помехоустойчивого кода, от которых существенно зависит сложность демодулятора. Надежность приема переданных сообщений в принципе не могла быть учтена, так как в «идеальной» системе связи обеспечивается абсолютная надежность приема сообщений. Поэтому такое сравнение нельзя признать корректным.

В [8-10] были получены формулы, позволяющие определить значение $P_{ser}(N, R_f, \rho_s)$ для системы связи, в которой для передачи сообщений используются оптимальные по Шеннону объемно- и поверхностно-сферические N -мерные ансамбли сигналов(АС), имеющих конечную длительность. Для объемно-сферических АС конечной длительности в [10] была получена формула, обобщающая формулу Шеннона для пропускной способности канала связи. В обобщенной формуле Шеннона [10]: $R_f = \log_2(1 + \rho_s / \beta^2) = \log_2(1 + \rho_s 10^{-0,1\Delta\rho})$, параметр $\Delta\rho$ (дБ) зависит только от значения вероятности ошибки при демодуляции сигналов P_{ser} , и их длительности (FT).

В данной работе определены границы энергетиче-

ской эффективности систем связи и выполнено сравнение характеристик нескольких конкретных систем, в которых применяются двумерные сигналы с квадратурно-амплитудной модуляцией (КАМ), с оптимальными. Это сравнение основано на обобщенной формуле Шеннона, справедливой для объемно и поверхностно-сферических N мерных ансамблей сигналов.

Границы энергетической и спектральной эффективности систем связи

В обобщенной формуле Шеннона [10] параметр $\beta^2 = \rho_s / (2^{R_f} - 1)$ связан с $P_{ser}(FT, \beta^2)$ – вероятностью ошибки при демодуляции объемно-сферического АС и нормированной длительности передаваемых сигналов

(FT) соотношением $P_{ser}(FT, \beta^2) = \int_{(FT)\beta^2}^{\infty} p_n(z) dz$, в кото-

ром $p_n(z) = \frac{z^{(FT)-1} e^{-z/2}}{2^{(FT)} \Gamma(FT)}$ – плотность распределения

вероятностей случайной величины по закону χ^2 . Если обозначить $\Psi_{1n}(x)$ – функцию обратную к функции $P_{ser}(FT, \beta^2)$, то из обобщенной формулы следует соотношение, определяющее величину $\beta^2(P_{ser}, N) = \Psi_{1n}(P_{ser}) / (FT)$.

Для вычисления значений $P_{ser}(FT, \beta^2)$ и $\beta^2(P_{ser}, N)$ можно воспользоваться пакетом прикладных программ MathCard, в котором имеются встроенные функции $pgamma(x, n)$ и $qgamma(x, n)$, позволяющие рассчитать зависимость $P_{ser}(FT, \beta^2)$, используя соотношение $F_n(x) = 1 - pgamma(x, n)$, и зависимость $\beta^2(P_{ser}, N)$, применяя соотношение для обратной к $F_n(x)$ функции $x = qgamma(1 - F_n(x), n)$, где $n = FT$, $x = (FT)\beta^2$. Для вычисления функции $\Psi_{1n}(P_{ser})$ можно также применить формулы из [11].

Обобщенная формула Шеннона определяет для реальной оптимальной системы связи с объемно-сферическим АС значение удельной скорости передачи сообщений (R_f) при применении в ней сигналов с ограниченным значением N и при заданном значении вероятности ошибки при приеме сообщений. Из нее следует, что по сравнению с «идеальной» по Шеннону системой связи, в реальной для достижения той же удельной скорости R_f , которая обеспечивается в «идеальной» системе, необходимо увеличить параметр ρ_b в $\beta^2(P_{ser}, N)$ раз. Таким образом, для реальной оптимальной системы связи граница спектральной эффективности системы связи имеет следующий вид:

$$\rho_b = [(2^{R_f} - 1) / R_f] \Psi_{1n}(P_{ser}) / (FT). \quad (2)$$

Отметим, что в [8] была получена оценка для $P_{ser}(FT, \beta^2)$, из которой можно получить простую формулу для расчета ρ_b . Эта формула, вывод которой представлен в Приложении, имеет вид:

$$\rho_b = \left(\frac{2^{R_f} - 1}{R_f} \right) (1 + \gamma) e^w, \quad (3)$$

где $\gamma = \ln(1/P_{ser})/(FT)$ и $w = \sqrt{\left(\frac{\gamma}{1+\gamma}\right)^2 + \frac{2\ln(1+\gamma)}{1+\gamma}} - \frac{\gamma}{1+\gamma}$.

Для иллюстрации точности определения величины ρ_b с помощью приближенной формулы в табл. 1 приведены результаты ее расчетов по точной (2) и приближенной (3) формулам для разных значений вероятности P_{ser} . Сравнение данных табл. 1 показывает, что при достаточно больших значениях FT и $P_{ser} \geq 10^{-5}$ точность определения параметра ρ_b по формуле (3) весьма высока – погрешность не превышает долей дБ. При небольших значениях FT и высоких требований к надежности связи, определяемых величиной P_{ser} , точность определения параметра ρ_b снижается, но остается в пределах 1,5 дБ.

Таблица 1. Сравнение расчетов ρ_b по формулам (2) и (3)

R_f бит/с/Гц	FT	Номер расчетной формулы	P_{ser}			
			10^{-10}	10^{-5}	10^{-3}	10^{-2}
			отношение сигнал/шум ρ_b дБ			
4	2	(ф-ла 2)	16,9	14,2	12,3	10,9
		(ф-ла 3)	17,6	15,3	13,7	12,6
	32	(ф-ла 2)	9,8	8,6	7,9	7,3
		(ф-ла 3)	10,1	9,06	8,4	7,9
8	2	(ф-ла 2)	26,5	23,5	21,6	20,2
		(ф-ла 3)	26,8	24,5	23,05	21,9
	32	(ф-ла 2)	17,9	18	17,1	16,7
		(ф-ла 3)	19,4	18,3	17,7	17,2

Обычно сравнение эффективности систем связи, в которых применяются конкретные ансамбли сигналов, выполняют, считая, что в сравниваемых системах используются ансамбли сигналов одинаковой размерности N_s , в них обеспечивается одинаковая помехоустойчивость приема сигналов, определяемая значением p_b , и одинаковые удельные скорости передачи сообщений R_f . При этом часто полагают, что в сравниваемых системах обеспечиваются одинаковые вероятности ошибок, приходящиеся на один бит (ρ_b). Связь между вероятностями P_{ser} и ρ_b можно определить, используя формулу Л.М. Финка [12]

$$P_{ser} = 1 - (1 - p_b)^{(FT)} \quad (4)$$

Формулы (2) и (3) позволяют корректно сравнивать эффективность конкретной системы связи с пределом Шеннона, учитывая нормированную длительность (FT) сигналов оптимального ансамбля сигналов.

Следует отметить, что вероятности ошибки p_b в дискретном канале связи, вход которого является входом кодера, а выход – выходом демодулятора [4], определяет его пропускную способность, равную $C(\rho_b) = 1 + p_b \log_2(p_b) + (1 - p_b) \log_2(1 - p_b)$. Значение $C(\rho_b)$ при

$\rho_b = 10^{-3}$ весьма близко к 1 ($C(\rho_b) = 1 - 10^{-2}$). При этом максимально возможная удельная скорость передачи сообщений в системе связи, включающей как непрерывный, так и дискретный каналы (в системе связи, в которой используются как демодуляторы принимаемых сигналов, так и декодеры помехоустойчивых кодов), равна $R_{fc} = R_f C(\rho_b)$. Используя достаточно длинные помехоустойчивые коды, можно обеспечить сколь угодно высокую надежность декодирования последовательности символов, поступающих с ошибками с выхода демодулятора, если они имеют кодовую скорость R_c , удовлетворяющую условию $R_c < C(\rho_b)$.

На рис. 1 представлены графики зависимостей $\rho_b(R_f)$ для объемно-сферических АС при разных значениях FT и при $\rho_b = 10^{-10}$ (сплошные линии) и 10^{-5} (пунктирные линии) для $FT = 1, 8, 32$ и 256 .

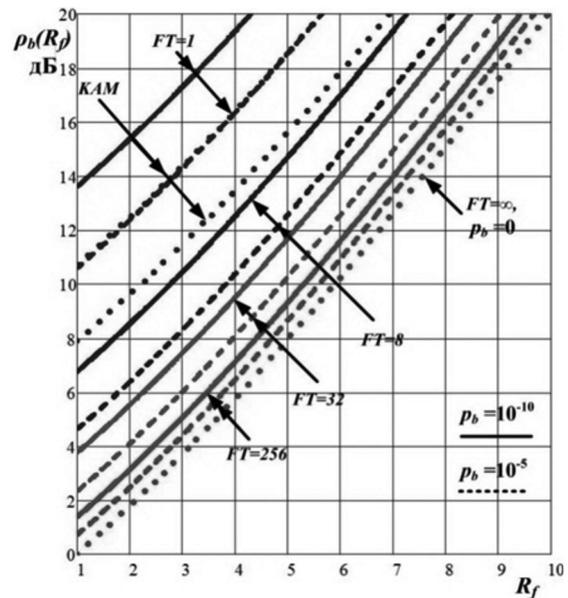


Рис. 1. Границы эффективности систем связи при разных значениях FT и ρ_b

Точками на рис. 1 нанесена граница спектральной и энергетической эффективности для «идеальной» системы связи, построенная по формуле Шеннона (1). Кроме того, на рис. 1 представлены зависимости $\rho_b(R_f)$ для систем связи, в которых применяются АС с КАМ. Зависимость для АС с КАМ при $\rho_b = 10^{-10}$ совпадает с зависимостью для объемно-сферического АС при $FT = 1$ и $\rho_b = 10^{-5}$.

Из рис. 1 видно, что с увеличением размерности объемно-сферического АС граница спектральной и энергетической эффективности системы связи для определенного значения P_{ser} приближается к границе Шеннона. Так, например, если $\rho_b = 10^{-10}$, то при $FT = 1$ эта граница выше границы Шеннона почти на 14 дБ, а при $FT = 256$ – всего на 1,2 дБ.

Система связи, в которой применяется объемно-сферический АС, при $\rho_b = 10^{-5}$ и $FT = 256$ имеет практически предельную энергетическую эффективность – она

уступает «идеальной» системе связи всего 0,3 дБ. Следует отметить, что, если к качеству приема в проектируемой системе связи предъявляется требование обеспечения вероятности ошибки при приеме сообщений, равной $\rho_b = 10^{-5}$, то надежность связи обеспечивается в системе с объемно-сферическим АС при оптимальной демодуляции принимаемых сигналов, причем при $FT = 256$ в системе нет необходимости в применении помехоустойчивого кодирования сообщений.

Из рис. 1 видно, что при снижении надежности приема сообщений на выходе демодулятора спектральная и энергетическая эффективность системы повышается. Например, при $FT = 8$ при демодуляции сообщений с малой вероятностью ошибочного приема $\rho_b = 10^{-10}$, когда применение помехоустойчивого кода в системе связи не требуется, на входе демодулятора должно быть обеспечено отношение сигнал/шум в $\rho_b = 12,5$ дБ при $R_f = 4$ бит/с/Гц. Если же вероятность ошибочного приема равна $\rho_b = 10^{-5}$, то при том же значении R_f необходимое отношение сигнал/шум на входе демодулятора становится равным $\rho_b = 10,2$ дБ. Так как значение $\rho_b = 10^{-5}$ достаточно мало, для повышения надежности связи может быть использован короткий помехоустойчивый код с кодовой скоростью $R_c \approx 1$ и простым алгоритмом декодирования. Такой код позволит дополнительно уменьшить значение ρ_b – вероятность ошибок при приеме сообщений на выходе системы связи (например, до $\rho_b = 10^{-10}$).

Следует иметь в виду, что при любом значении вероятности ошибки в последовательности принятых символов на выходе демодулятора, надежность связи можно повысить до любой требуемой величины, применяя при передаче сообщений помехоустойчивый код, исправляющий ошибки в этой последовательности. Поэтому приблизить энергетическую эффективность создаваемой системы связи к пределу Шеннона, т.е. уменьшить энергетику линии связи (уменьшить ρ_b), необходимую для обеспечения требуемой надежности приема сообщений, переданных с определенной скоростью, можно, используя следующие возможности: 1) снизить требования к надежности приема сообщений на выходе демодулятора (увеличив ρ_b) и применить для достижения требуемой надежности приема в системе связи помехоустойчивые коды, и 2) использовать для передачи сообщений оптимальные ансамбли сигналов большой размерности (N -мерные АС, где $N \gg 1$).

Уже отмечалось, что целью проектирования системы связи является обеспечение требований к удельной скорости (R_{fc}) и надежности (ρ_0) передачи в ней сообщений. Для любого значения вероятности ошибки ρ_b при демодуляции сигнала, используя достаточно длинные помехоустойчивые коды с избыточностью, согласно теореме Шеннона для дискретных каналов связи, возможно обеспечить любую, сколь угодно высокую надежность связи (любое сколь угодно малое значение вероятности ошибочного приема (ρ_0) символов на вы-

ходе декодера). Причем реализовать это можно, применяя коды, у которых кодовая скорость R_c – отношение количества (k) избыточных символов к длине кодовой комбинации (n) весьма близка к 1, т.е. $R_c \approx 1$. Следует иметь в виду: чем ближе R_c к пропускной способности канала связи, тем более длинным должен быть используемый код и более сложным декодер. Применяя для передачи сообщений сигналы, у которых удельная скорость передачи сообщений равна R_f , и коды с кодовой скоростью R_c , спектральная эффективность системы связи оказывается равной $R_{fc} = R_f R_c$.

При проектировании системы связи необходим поиск компромисса между сложностью устройств модуляции и демодуляции сигналов (связанных с параметрами и размерностью используемого АС и значением вероятности ρ_b ошибки при их демодуляции) и сложностью декодеров и декодеров (связанных с длиной используемого кода, его кодовой скоростью и значением вероятности ρ_0 ошибки символов в последовательности, сформированной на выходе декодера). Естественно, при решении о выборе АС и помехоустойчивого кода для системы связи следует учитывать не только показатель ее энергетической эффективности, но и экономические факторы, связанные со сложностью реализации основных узлов проектируемой системы: передатчика, модулятора и демодулятора, кодера и декодера.

На рис. 2 приведены вычисленные по формуле (2) границы эффективности систем связи при $\rho_b = 10^{-3}$ и разных значениях FT , принимающих значения от 1 до 128. Там же приведена граница Шеннона, для которой $\rho_b = 0$ и $FT = \infty$.

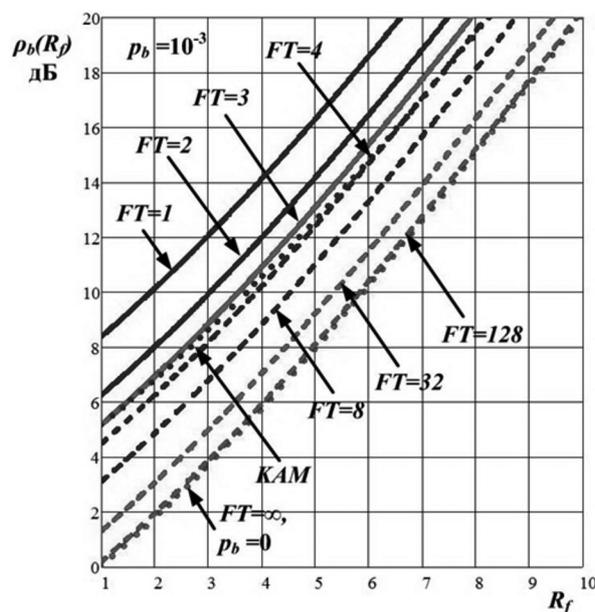


Рис. 2. Границы эффективности систем связи при разных значениях FT и $\rho_b = 10^{-3}$

Из рис. 2 видно, что в случае $\rho_b = 10^{-3}$ граница спектральной и энергетической эффективности Шеннона практически достигается при применении АС, размер-

ность которого равна $FT=128$. Ансамбль сигналов размерности 16 ($FT=8$), уступает в энергетической эффективности границе Шеннона всего 3 дБ. Точками на рис. 2 показана зависимость $\rho_b(R_f)$ для системы КАМ, при которой $FT=1$. В случае $\rho_b=10^{-3}$ эта зависимость лежит выше границы Шеннона на 4,5 дБ.

Отметим, что, как следует из приведенных в [8, 9] результатов, N -мерные объемно-сферические и поверхностно-сферические АС обеспечивают практически одинаковую надежность связи. Поэтому полученные для объемно-сферического АС данные могут быть полностью отнесены и к поверхностно-сферическим АС.

Сравнение энергетической эффективности конкретных систем связи и объемно-сферического АС

С представленными на рис. 1 и 2 границами энергетической эффективности оптимальных АС целесообразно сравнивать параметры конкретных систем модуляции. Это позволяет делать заключения о возможных путях их совершенствования с целью уменьшения энергетического потенциала линии связи. В [5] приведены параметры трех видов модуляции, применяемых в современных системах связи: 1) решетчатой кодированной модуляции (РКМ), в которой используются сверточные коды; 2) кодовой модуляции с битовым перемежением (МКБП) с турбокодом (ТК) и 3) многоуровневой кодовой модуляции (МКМ), в которой также применяется ТК.

В табл. 2 в строках 1...4 приведены данные видов модуляции.

Таблица 2. Сравнение параметров систем РКМ, МКБП, МКМ и объемно-сферического АС

№	Параметры системы	РКМ	МКБП с ТК	МКМ с ТК
1	Видмодуляции	КАМ-64	КАМ-64	КАМ-64
2	вид кода	сверточный код		
3	параметры кода	$K = 7;$ $R_c = 5/6$	$K = 4;$ $R_c = 5/6$	$K = 7;$ $R_c = 5/6$
4	длина переданного сообщения N_b , бит	56	576	20000
5	$\rho_b, \Delta\rho_b$ дБ	13, $\Delta\rho_b = 5$	11,8, $\Delta\rho_b = 3,8$	10,4, $\Delta\rho_b = 2,4$
6	объемно-сфер. АС, FT	8		
7	длина переданного сообщения N_b , бит	40		
8	$\rho_b, \Delta\rho_b$ дБ	12,5, $\Delta\rho_b = 4,5$		

Во всех этих системах применяются двумерные ($FT = 1$) сигналы 64-КАМ, у которых удельная скорость передачи сообщений составляет $R_f = 6$ бит/с/Гц. В них применяются сверточные коды (СВК) с кодовой скоростью $R_c = 5/6$, и поэтому результирующая удельная скорость передачи сообщений по каналу связи составляет $R_{fc} = R_f R_c = 5$ бит/с/Гц. В таблице указан параметр K – длина кодового ограничения СВК, которая у каждого вида модуляции разная, а также длина кодируемой по-

следовательности символов, причем для РКМ указана длина последовательности символов (N_b), которая обрабатывается в демодуляторе при принятии решения. Как пояснялось ранее, для СВК она составляет $N_b = 6 \div 8 K$. В пятой строке таблицы указаны значения отношения сигнал/шум ρ_b , приходящегося на один принятый бит сообщения, при котором вероятность ошибки в последовательности символов на выходе декодера равна $\rho_0 = 10^{-5}$.

Используя границы эффективности систем связи, согласно рис. 1, в строках 6...8 табл. 2 приведены данные о параметрах объемно-сферического АС (FT, N_b, ρ_b), при которых можно достичь такие же скорость передачи сообщений и надежность связи, которые обеспечиваются в системах РКМ, МКБП и МКМ. В строке 6 этой таблицы указаны значения $\Delta\rho_b$ дБ, определяющие значение ρ_b , необходимое для обеспечения надежности приема сообщений, равное $\rho_0 = 10^{-5}$, превышающее для этих систем границу Шеннона. Отметим, что в системе с использованием объемно-сферического АС, нужная надежность приема сообщений обеспечивается при демодуляции принятого сигнала, и поэтому применение в ней помехоустойчивого кода не требуется.

Как видно из табл. 2, в системах с РКМ, МКБП и МКМ при демодуляции и декодировании принимаемых сигналов длина обрабатываемых последовательностей символов составляет $N_b = 56$ для РКМ, $N_b = 576$ для МКБП и $N_b = 20000$ для МКМ. Сложность реализации системы связи тесно связана с величиной N_b . Отметим, что граница Шеннона указывает, что при $R_f = 5$ возможно осуществлять передачу сообщений без ошибок, если $\rho_b = 8$ дБ. В строке 5 таблицы указаны значения $\Delta\rho_b$ дБ, показывающие насколько значение ρ_b для этих систем превышает границу Шеннона. В системе с использованием сигналов оптимального объемно-сферического АС, длина последовательности двоичных символов, передаваемых каждым сигналом, составляет $N_b = (FT)R_f$.

В строках 6, 7 и 8 табл. 2 приведены параметры системы с объемно-сферического АС при $\rho_b = 10^{-5}$ и $R_f = 5$ бит/с/Гц (такую же надежность приема сообщений и скорость передачи обеспечивают системы РКМ, МКБП и МКМ) размерность оптимального АС должна составлять $N = 2FT = 16$. В такой системе в демодуляторе осуществляется обработка сигналов, при которой на выходе формируется двоичная последовательность символов с длиной $N_b = 40$. Эта длина существенно меньше данной для систем РКМ, МКБП и МКМ. Энергетическая эффективность системы с объемно-сферического АС примерно такая же, как и у систем РКМ, МКБП, но уступает 2 дБ системе МКМ с ТК. Следует, однако, обратить внимание на то, что в системе МКМ применяется длинный ($N_b = 20000$) помехоустойчивый код и, следовательно, декодер этого кода является весьма сложным устройством.

По этой причине реализация системы, использующей для передачи сообщений оптимальные объемно-сферические АС, может быть существенно проще, чем у сравниваемых с ней систем. Приблизить эффективность этой системы к пределу Шеннона возможно, снизив требования к надежности демодуляции принимаемых сигналов и определить, например, вероятность ошибок при демодуляции сигналов равной $\rho_b = 10^{-3}$ или $\rho_b = 10^{-2}$. Применяя в системах связи совместно оптимальные многомерные АС и методы помехоустойчивого кодирования, можно приблизиться к границе Шеннона, используя АС большей размерности, чем указанная в приведенном примере (например, $N = 64$), и помехоустойчивые коды (например, каскадные коды). При этом демодулятор и декодер будут достаточно простыми устройствами.

Приложение:

Определение отношения сигнал/шум, приходящегося на один принятый бит, для оптимального ансамбля сигналов

Используя в [9] формулу:

$$P_{ser} \cong \exp \left\{ - (FT) \left[\left(\frac{2^{C_f} - 2^{R_f}}{2^{R_f} - 1} \right) - \ln \left(\frac{2^{C_f} - 1}{2^{R_f} - 1} \right) \right] \right\},$$

найдем очень простую формулу для расчета ρ_b – отношения сигнал/шум, приходящегося на один принятый бит, для оптимального ансамбля сигналов.

Если обозначить $\gamma = \ln \left(\frac{1}{P_{ser}} \right) / (FT)$ и $V = (2^{C_f} - 1) \div (2^{R_f} - 1)$, а также учесть, что $2^{C_f} - 1 = \rho_b R_f$, то величины V и γ связаны соотношением $V - \ln(V) = (1 + \gamma)$.

Обозначив $V = (1 + \gamma)e^w$, для определения величины w получим следующее уравнение $e^w - w / (1 + \gamma) = 1 + \ln(1 + \gamma) / (1 + \gamma)$, из которого следует $w \leq 1$. Используя приближенную формулу $e^w \cong 1 + w + 0,5w^2$, получим уравнение для определения w : $w^2 + \frac{2w\gamma}{1 + \gamma} - \frac{2 \ln(1 + \gamma)}{1 + \gamma} =$

$$= 0. \text{ Отсюда } w = \sqrt{\left(\frac{\gamma}{1 + \gamma} \right)^2 + \frac{2 \ln(1 + \gamma)}{1 + \gamma}} - \frac{\gamma}{1 + \gamma} \text{ и формула}$$

для определения отношения сигнал/шум, приходящегося на один принятый бит:

$$\rho_b = \left(\frac{2^{R_f} - 1}{R_f} \right) (1 + \gamma) e^w.$$

Заключение

Получены границы энергетической эффективности систем связи, пригодные для сравнения значений основных параметров конкретных систем (удельной скорости передачи сообщений и отношения сигнал/шум, приходящегося на один принятый бит) с потенциально

возможными значениями этих параметров. Показано, что при проектировании системы целесообразно для передачи сообщений выбирать оптимальные методы модуляции – многомерные объемно-сферические или поверхностно-сферические ансамбли сигналов.

Обсуждены вопросы выбора параметров таких ансамблей сигналов, которые позволяют приблизить энергетическую эффективность проектируемой системы связи к границе Шеннона. Анализируются вопросы выбора помехоустойчивого кода, позволяющего упростить систему связи за счет применения в ней наиболее простых в техническом отношении модулятора и демодулятора, а также кодера и декодера.

Литература

1. Shannon C. Probability of error for optimal codes in Gaussian channel. BellSystemTechn. J., May, 1959. (Перевод на русский язык статьи «Вероятность ошибки для оптимальных кодов в гауссовском канале», опубликованной в книге Шеннон К. Работы по теории информации и кибернетике. М.: Из-во иностранной литературы. // Под ред. Р.Л. Добрушина и О.Б. Лупанова).
2. Shannon C. Communication in the presence of noise, Proc. IRE, № 1, 1949. (Перевод на русский язык статьи «Связь при наличии шума», опубликованной в книге Шеннон К. Работы по теории информации и кибернетике. М.: Из-во иностранной литературы/под ред. Р.Л. Добрушина и О.Б. Лупанова).
3. Котельников В.А. Теория потенциальной помехоустойчивости. М.: Госэнергоиздат, 1956.
4. Прокис Дж. Цифровая связь // Перевод с английского под ред. Д.Д. Кловского // М.: Советское радио, 2000.
5. Варгуаузин В.А., Цикин И.А. Методы повышения энергетической и спектральной эффективности цифровой радиосвязи. СПб.: БХБ-Петербург, 2013.
6. Зюко А.Г., Фалько И.П., Панфилов И.П., Банкет В.Л., Иващенко П.В. Помехоустойчивость и эффективность систем передачи информации // Под. ред. А.Г. Зюко. М.: Радио и связь, 1985.
7. Зяблов В.В., Коробков Д.Л., Портой С.Л. Высокоскоростная передача сообщений в реальных каналах связи. М.: Радио и связь, 1991.
8. Быховский М.А. Вероятность ошибки для оптимальных многомерных кодов в гауссовом канале связи и их основные характеристики // Электросвязь № 2, 2016.
9. Быховский М.А. Помехоустойчивость приема оптимальных сигналов, расположенных на поверхности N-мерного шара // Электросвязь № 3, 2016.
10. Быховский М.А. Пропускная способность канала связи при передаче сигналов ограниченной длительности // Электросвязь № 8, 2016.
11. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Основы моделирования и первичная обработка данных. Справочное изд. М.: Финансы и статистика, 1983.
12. Финк Л.М. Теория передачи дискретных сообщений. М.: Советское радио, 1970.

УДК 621.391.13

ЭНЕРГЕТИЧЕСКАЯ ЭФФЕКТИВНОСТЬ СИСТЕМ СВЯЗИ, ИСПОЛЬЗУЮЩИХ МНОГОМЕРНЫЕ АНСАМБЛИ СИГНАЛОВ И КАСКАДНОЕ КОДИРОВАНИЕ

Быховский М.А., д.т.н., профессор Московского технического университета связи и информатики, e-mail: bykhmark@gmail.com.

ENERGY EFFICIENCY OF TELECOMMUNICATION SYSTEMS THAT USE MULTIDIMENSIONAL SIGNAL ENSEMBLES AND CASCADE SOURCE CODING

Bykhovsky M.

The author investigates a telecommunication system that uses cascade source message coding and where the signals from multi-dimensional surface-spherical (SS) signal ensemble; such signals create the inner code without contributing excess symbols) are used for message transmission via continuous telecommunication channel; Reed Solomon (RS) code is used as an outer code. The article contains formulas allowing to determine the probability of error reception of the signals in the investigated telecommunication system; such formulas determine the impact of the SS signals' length and Reed-Solomon codes' length on the noise immunity of the signal reception.

It is shown that in such a system it is possible to provide energy efficiency of message transmission that is fairly close to energy efficiency of an «ideal», according to Shannon, telecommunication system while using rather simple demodulators of the received signals and decoders of the RS codes.

Key words: cascade source coding, multidimensional signal ensemble, noise immunity of signal reception, Reed-Solomon codes, complexity realization of demodulators and decoders.

Ключевые слова: каскадное кодирование, многомерный ансамбль сигналов, код Рида-Соломона, помехоустойчивость приема сообщений, сложность реализации демодуляторов и декодеров.

Введение

В работе К. Шеннона [1] показано, что в канале связи с полосой частот F при белой гауссовской помехе оптимальным методом передачи сообщений является использование N -мерных поверхностно-сферических ансамблей сигналов ($N = 2FT$, T – длительность сигналов). Применяя такие ансамбли сигналов большой длительности ($N \rightarrow \infty$), можно обеспечить прием сигналов с абсолютной надежностью в том случае, если удельная скорость передачи сообщений R_f (в бит/с/Гц) меньше максимально возможной $R_{f\max}$, равной пропускной способности

канала связи $C = \ln(1 + \rho_s)$, где ρ_s – отношение сигнал/шум на входе демодулятора, связанное с отношением сигнал/шум на один принятый бит $\rho_b : \rho_s = \rho_b R_{f\max}$. Надежность связи определяется вероятностью ошибки при демодуляции сообщений $P_{ser}(N, R_f, \rho_s)$. Систему связи, в которой $P_{ser}(N, R_f, \rho_s) \rightarrow 0$ при $N \rightarrow \infty$, Шеннон определил как «идеальную» [1]. В «идеальной» системе связи минимальное значение ρ_{b0} связано с максимальной удельной скоростью передачи сообщений $R_{f\max}$ соотношением

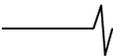
Проведено исследование системы связи с каскадным кодированием сообщений. Для передачи сообщений по непрерывному каналу связи используются сигналы, принадлежащие многомерному поверхностно-сферическому ансамблю сигналов - ПСАС (эти сигналы образуют внутренний код без внесения избыточных символов), а в качестве внешнего кода применяется код Рида Соломона (РС).

Приведены формулы, позволяющие определить вероятность ошибочного приема принимаемых сигналов в исследуемой системе связи, с помощью которых определено влияние длины сигналов, принадлежащих ПСАС, и длины кодов Рида-Соломона на помехоустойчивость приема сообщений.

Показано, что в такой системе возможно, используя достаточно простые демодуляторы принимаемых сигналов и декодеры кодов РС, обеспечить энергетическую эффективность передачи сообщений весьма близкую к энергетической эффективности «идеальной» по Шеннону системы связи. Отмечено, что по своей энергетической эффективности рассмотренные в статье каскадные коды превосходят современные турбокоды.

$$\rho_{b0} = (2^{R_{f\max}} - 1) / R_{f\max}. \quad (1)$$

Для передачи сообщений в такой системе применяется N -мерный поверхностно-сферический ансамбль сигналов (ПСАС), являющийся оптимальным. Преобразование информационной последовательности символов в определенный сигнал ПСАС представляет собой процесс модуляции, аналогичный квадратурно-амплитудной, амплитудно-фазовой и т.п. Поскольку при формировании сигналов ПСАС используются N фазовых модуляторов, в данном случае уместно назвать процесс передачи информационных сигналов с помощью ПСАС гиперфазовой модуляцией (ГПФМ).



В системе связи с каскадным кодированием в качестве внутреннего кода используются сигналы с ГПФМ, а в качестве внешнего – коды Рида-Соломона (РС), соответствующие классу кодов с максимально достижимым расстоянием [2]. При заданной кодовой скорости R_c кодовое расстояние $d_H = (2t_c + 1)$ между кодовыми комбинациями имеет наибольшее из возможных значений. Для кодов РС длина кодовой комбинации $n = (2^m - 1)$, $k = (2^m - 1 - 2t_c)$ – число информационных позиций в ней, t_c – количество ошибок, которое может быть исправлено кодом РС, $R_c(\hat{t}_c) = 1 - 2\hat{t}_c - 1 / (n + 1)$ – кодовая скорость $\hat{t}_c = (t_c / n)$. Такие коды при достаточно высокой кодовой скорости R_c могут обеспечить высокую надежность связи при сравнительно небольшой длине n и имеют достаточно простой алгоритм декодирования.

Сигнальная матрица каскадного кода показана на рис. 1.

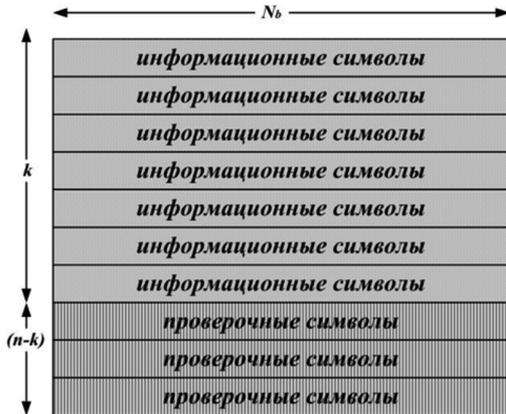


Рис. 1. Сигнальная матрица каскадного кода

Символы, расположенные в этой матрице, построчно передаются с помощью сигналов с ГПФМ, причем последовательности передаваемых символов каждой строки длиной N_b рассматриваются при модуляции как номер передаваемого сообщения. Если при демодуляции сигналов с вероятностью P_{ser} возникли ошибки, то ошибочно принятые символы распределяются между столбцами сигнальной матрицы. Информационные символы, расположенные в столбцах матрицы, кодируются с помощью кода P_c , имеющего длину n кодовой комбинации, в которой используются k информационных и $(n - k)$ проверочных символов. После демодуляции принятого сигнала с ГПФМ, можно считать, что сообщения передаются в дискретном канале связи, в котором вероятность ошибки одного принятого символа равна $\rho_b = P_{ser}$.

На рис. 2 представлены зависимости $P_{ser}(FT)$ ошибки при демодуляции сигналов с ГПФМ, рассчитанные с использованием [3] при значениях удельной скорости передачи сообщений $R_f = 4, 6, 8$ и 10 бит/с/Гц. Каждая группа практически совпадающих кривых 1, 2 и 3 на рис. 2 относится к случаю, когда для каждого конкретного значения R_f величина ρ_b – отношение сигнал/шум,

приходящееся на один принятый бит, превышает минимально возможное значение ρ_{b0} , согласно формулы Шеннона [1], примерно на $\Delta\rho_b = 0,7, 1,3$ и $1,8$ дБ.

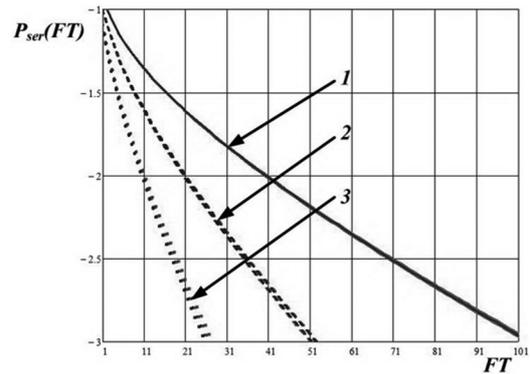


Рис. 2. Зависимости $P_{ser}(FT)$ ошибки при демодуляции сигналов с ГПФМ

В табл. 1 приведены точные значения $\Delta\rho_b$ для зависимостей 1, 2 и 3.

Таблица 1. Таблица некоторых характеристик сигналов с ГПФМ

R_f бит/с/Гц	4	6	8	10	FT			
					10^{-2}	10^{-3}	10^{-5}	10^{-10}
P_{sb} дБ	11,76	18	24	30,1	$P_{ser}=0; FT=\infty$			
$\Delta\rho_{s1}$ дБ	0,82	0,7	0,9	1	40	102	227	648
$\Delta\rho_{s2}$ дБ	1,24	1,2	1,3	1,5	20	48	104	325
$\Delta\rho_{s3}$ дБ	1,74	1,66	1,8	2	11	25	68	178

Из этого рисунка следует, что для вероятности ошибки при демодуляции сигналов, равной $\rho_b = 10^{-2}$, необходимо использовать АС с ГПФМ с $FT = 40, 20$ и 11 . Если же $\rho_b = 10^{-3}$, то $FT = 10^2, 48$ и 25 . Таким образом, повышение требований к надежности приема сообщений при демодуляции сигналов с ГПФМ приводит к увеличению почти в 2,5 раза нормированной длительности этих сигналов (и размерности АС). Это, естественно, усложняет модулятор и демодулятор данного АС.

В табл. 1 указаны также достаточно большие значения FT для $\rho_b = 10^{-5}$ и $\rho_b = 10^{-10}$. Если при демодуляции сигналов обеспечивается столь высокая надежность приема сообщений, то применение помехоустойчивого кодирования может оказаться ненужным. Однако реализация модуляторов и демодуляторов сигналов с ГПФМ со столь большими значениями FT является весьма сложной технической задачей.

Для расчета вероятности ошибочного декодирования кодовой комбинации кода РС $P_{dec}(n)$, положим – вероятность ошибки одного символа этой комбинации, равна $\rho_b \cong P_{ser}$. При этом распределение вероятностей ошибочно принятых символов в кодовой комбинации подчиняется закону Бернулли, и для расчета $P_{dec}(n)$ можно использовать формулу:

$$P_{dec}(n) = \sum_{i=(t_c+1)}^n C_n^i \rho_b^i (1 - \rho_b)^{(n-i)} \quad (2)$$

Применение каскадного кодирования с внешним кодом РС позволяет, используя достаточно простые модуляторы и демодуляторы в системе связи с ГПФМ, по-

высить надежность приема сообщений до нужного значения, исправляя ошибки в информационных символах, возникших при демодуляции сигналов.

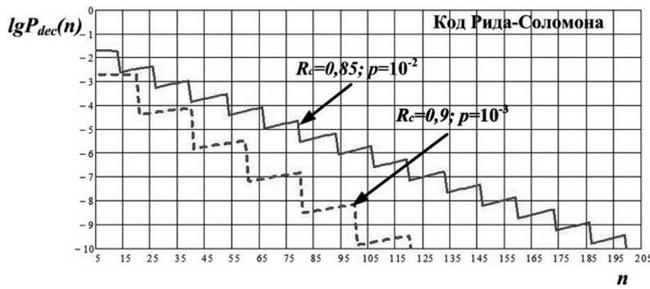


Рис. 3. Зависимости вероятности ошибки $P_{dec}(n)$ декодирования от длины кода

На рис. 3 показаны зависимости вероятности ошибки $P_{dec}(n)$ декодирования от длины кодовой комбинации кода РС для двух случаев:

- 1) при вероятности ошибки при демодуляции $\rho_b = 10^{-2}$ и кодовой скорости $R_c = 0,85$ и
- 2) при $\rho_b = 10^{-3}$ и $R_c = 0,9$.

Удельная скорость передачи сообщений в такой системе связи для $R_f = 4, 6, 8$ и 10 в первом случае составляет $R_{fc} = 3,4; 5,1; 6,8$ и $8,5$ бит/с/Гц, а во втором – $R_{fc} = 3,6; 5,4; 7,2$ и 9 бит/с/Гц.

Из рис. 3 следует: если требуемая надежность приема сообщений определяется условием $P_{dec}(n) = 10^{-5}$, то в канале при $\rho_b = 10^{-2}$ длина кода РС равна $n = 90$, а если в канале $\rho_b = 10^{-3}$, то $n = 50$. Если же $P_{dec}(n) = 10^{-10}$ длина кода РС равна $n = 200$, $\rho_b = 10^{-2}$ и $n = 120$ для канала при $\rho_b = 10^{-3}$.

В табл. 2 указаны длины $N_b = (FT)R_f$ последовательности двоичных символов, передаваемых каждым сигналом с ГПФМ для разных значений R_f и ρ_b .

Таблица 2. Длина внешнего и внутреннего кода в рассматриваемой системе связи

$R_f \rho_b$	N_{b3}				
	10^{-2}	10^{-3}	10^{-5}	10^{-10}	
4	44	100	272	712	
6	66	150	408	1068	
8	88	200	544	1424	
10	110	250	680	1788	
$P_{dec}(n)=10^{-5}$	n	90	50	-	-
$P_{dec}(n)=10^{-10}$	n	200	120	-	-

Значение N_{b3} соответствует случаю, когда энергетика линии связи выбрана так, что $\Delta\rho_b \approx 1,8$ дБ (этот случай соответствует группе кривых 3 на рис. 2, а также последней строке в табл. 3). В нижних двух строках табл. 2 указаны длины внешних кодов РС, которые обеспечивают значения ошибки декодирования $P_{dec} = 10^{-5}$ и $P_{dec} = 10^{-10}$, когда вероятность ошибки при демодуляции принимаемых сигналов равна $\rho_b = 10^{-2}$ или $\rho_b = 10^{-3}$.

Следует отметить, что при выборе энергетике линии связи, близкой к пределу Шеннона (при уменьшении

значения $\Delta\rho_b$), значения N_{b1} и N_{b2} увеличиваются и требуется применять более сложные структуры модулятора и демодулятора.

Из таблицы 2 видно, что при любом заданном значении R_f длина внутреннего кода существенно зависит от величины $\rho_b \cong P_{ser}$ – чем меньшее значение имеет ρ_b , тем более длинным должен быть внутренний код и, следовательно, тем более сложными в технической реализации будет модулятор и оптимальный демодулятор сигналов с ГПФМ.

Два последних столбца табл. 2 относятся к случаям, когда достигается высокая надежность приема сообщений ($\rho_b = 10^{-5}$ или $\rho_b = 10^{-10}$), и нет необходимости в применении помехоустойчивых кодов. При этом, как видно из приведенных данных, количество двоичных символов, передаваемых с помощью одного сигнала из ансамбля с ГПФМ, оказывается весьма значительным, особенно когда удельная скорость передачи сообщений должна быть достаточно большой ($R_f = 8$ или 10 бит/с/Гц). Из табл. 2 также следует, что при выборе размерности ансамбля сигналов с ГПФМ количество двоичных символов, которые передаются с помощью одного сигнала, принадлежащего ансамблю с ГПФМ, сравнительно невелико, а надежность принятых сообщений на выходе демодулятора оказывается невысокой ($\rho_b = 10^{-2}$ или $\rho_b = 10^{-3}$). В этом случае повысить надежность сообщений возможно, применив помехоустойчивое каскадное кодирование.

В последних двух строках табл. 2 приведены значения длины кодов РС ($n = 50 \dots 200$), с помощью которых можно, исправив ошибки, возникшие при демодуляции сигналов с ГПФМ, обеспечить требуемую надежность приема сообщений на выходе декодера. Техническая реализация декодера для кода РС при указанных сравнительно небольших значениях длины кодовых комбинаций практических затруднений не вызывает. Длина рассматриваемого каскадного кода, равная $N_c = N_{b3} \cdot n$, в $K_c(p_b, P_{dec}) = N_c(p_b) / N_{b3}(P_{dec})$ раз больше длины сигналов, используемых в ансамбле ГПФМ) в качестве внутреннего кода.

В табл. 3 даны значения коэффициента $K_c(p_b, P_{dec})$, определенного на основе данных табл. 2 при разных значениях ρ_b и P_{dec} . Очевидно, длина каскадного кода в $10 \dots 18$ раз может превосходить длину сигналов внутреннего кода.

Таблица 3. Значения коэффициента $K_c(p_b, P_{dec})$

P_{dec}	R_f	$\rho_b=10^{-2}$	$\rho_b=10^{-3}$
10^{-5}	6	14,6	18,4
	10	14,6	18,4
10^{-10}	6	12,4	16,8
	10	12,3	10

Однако применение каскадного кода с внешним кодом РС дает возможность существенно уменьшить длину внутреннего кода, снизив требования к надежности демодуляции сигналов с ГПФМ (к значению P_{dec}). Это позволяет упростить техническую реализацию модулятора и демодулятора сигналов, используемых в системе связи.



Поскольку в каскадном коде применен (в качестве внутреннего кода) оптимальный по Шеннону ансамбль сигналов с ГПФМ, то он по помехоустойчивости не уступает наиболее мощным на сегодня турбокодам [4], кодовые комбинации которого передаются по каналу связи с использованием многопозиционных сигналов при квадратурно-амплитудной или амплитудно-фазовой модуляцией [5]. При одинаковой помехоустойчивости приема сообщений кодовые комбинации турбокодов по сравнению с рассмотренным выше каскадным кодом должны иметь существенно большую длину. Кроме того, при декодировании турбокода должен применяться декодер, алгоритм работы которого основан на использовании сложной итерационной процедуры «мягкого» декодирования принятой кодовой комбинации [4].

Заключение

Применение в системах связи гиперфазовой модуляции – многомерных ансамблей поверхностно-сферических сигналов, а также каскадного кодирования позволяет обеспечить их энергетическую эффективность весьма близкую к энергетической эффективности «идеальной» по Шеннону системы связи, используя для этого сигналы, имеющие сравнительно небольшую длительность и коды Рида-Соломона небольшой длины.

При высокой удельной скорости передачи сообщений по своей энергетической эффективности использу-

емые каскадные коды превосходят современные турбокоды, в которых для передачи кодовых комбинаций по каналам связи применяются многопозиционные сигналы с квадратурно-амплитудной или амплитудно-фазовой модуляцией.

Автор выражает признательность профессору В.П. Дворковичу за внимание к работе и полезные замечания.

Литература

1. Shannon C. Probability of error for optimal codes in Gaussian channel. BellSystemTechn. J., May, 1959. (Перевод статьи «Вероятность ошибки для оптимальных кодов в гауссовском канале» из книги Шеннона К. Работы по теории информации и кибернетике. М.: Из-во иностранной литературы // Под ред. Р.Л. Добрушина и О.Б. Лупанова).
2. Питерсон У., Уэлдон Э. Коды, исправляющие ошибки. М.: Мир, 1976.
3. Быховский М.А. Помехоустойчивость приема оптимальных сигналов, расположенных на поверхности N-мерного шара // Электросвязь, № 3, 2016.
4. Варгуаузин В.А., Цикин И.А. Методы повышения энергетической и спектральной эффективности цифровой радиосвязи. СПб.: БХБ-Петербург, 2013.
5. Прокис Дж. Цифровая связь / Перевод с английского под ред. Д.Д. Кловского, М.: Советское радио, 2000.

УДК 621.396.42

ПРИМЕНЕНИЕ ЗАЩИТНЫХ ИНТЕРВАЛОВ В ОДНОЧАСТОТНЫХ КВ МОДЕМАХ ПЕРЕДАЧИ ДАННЫХ

*Маслаков М.Л., аспирант, научный сотрудник ПАО «Российский институт мощного радиостроения»,
e-mail: maslakovml@gmail.com.*

APPLICATION OF GUARD INTERVAL IN SINGLE-TONE HF DATA MODEMS

Maslakov M.L.

Application features of guard intervals in single-tone high frequency data modems is considered in this article. The influence of different types of guard intervals on the accuracy of calculation of the impulse response of the communication channel is proved and illustrated. A new method of guard intervals compensation is designed. The results of simulation of bit error rate of a single-tone HF data modem using guard interval are presented.

Key words: intersymbol interference, adaptive correction, impulse response of the channel, integral convolution type equation, ill-posed problem, guard interval, bit errorrate.

Ключевые слова: межсимвольная интерференция, адаптивная коррекция, импульсная характеристика канала, интегральное уравнение типа свертки, некорректная задача, защитный интервал, вероятность ошибки на бит.

Введение

При передаче сигналов по коротковолновому (КВ) радиоканалу возникают эффекты многолучевого распространения, связанные с переотражением излучаемой волны и ее распространением по нескольким путям, что приводит к межсимвольной интерференции (МСИ). В результате МСИ происходит наложение соседних символов друг на друга, что в конечном итоге приводит к ошибкам демодуляции.

В многочастотных КВ модемах передачи данных, использующих сигналы OFDM (Orthogonal Frequency Division Multiplexing – мультиплексирование с ортогональным частотным разделением каналов) для борьбы с МСИ вводят защитные интервалы (ЗИ) между символами. Длительность ЗИ выбирается исходя из потенциальной длительности интервала многолучевости, и для КВ канала составляет величину порядка нескольких миллисекунд [1].

Однако в одночастотных или последовательных КВ модемах передачи данных длительность символа обычно составляет менее одной миллисекунды, поэтому введение ЗИ аналогично многочастотным модемам существенно снизит информационную скорость. В известных стандартах последовательных КВ модемов [2, 3] ЗИ не используют, а для компенсации МСИ применяют методы адаптивной коррекции сигналов, заключающейся в построении корректирующего фильтра (КФ). Для этого в передаваемый сигнал осуществляют периодические вставки тестовых сигналов, не несущих информации, а служащих лишь для расчета импульсной характеристики (ИХ) канала и определения соответствующей ИХ КФ на приемной стороне.

В статье рассматриваются особенности применения защитных интервалов в одночастотных коротковолновых модемах передачи данных. Показано влияние различных видов защитных интервалов на точность расчета импульсной характеристики канала. Предложен способ компенсации защитного интервала. Приведены результаты моделирования помехоустойчивости одночастотного коротковолнового модема передачи данных с применением защитных интервалов.

Целью данной работы является повышение помехоустойчивости одночастотного КВ модема передачи данных с помощью введения ЗИ между тестовым и информационным сигналами при условии сохранения информационной скорости. Для этого в работе проведен анализ влияния различных ЗИ на точность расчета ИХ канала и определен оптимальный интервал, требуемый при расчете ИХ канала.

Адаптивная коррекция

Процедура адаптивной коррекции состоит из следующих задач: расчет ИХ канала связи; расчет ИХ КФ; коррекция (фильтрация) принимаемого информационного сигнала с помощью КФ.

Расчет ИХ канала связи сводится к решению уравнения вида:

$$s_0(t) * h_k(t) = \int_{T_1}^{T_2} s_0(t-\tau) \cdot h_k(\tau) d\tau = u_0(t), \quad t \in [T_3; T_4], \quad (1)$$

где $s_0(t)$ – передаваемый тестовый сигнал, $h_k(t)$ – ИХ канала связи, $u_0(t)$ – принимаемый искаженный тестовый сигнал.

Приведенное уравнение (1) относится к линейному интегральному уравнению типа свертки первого рода или уравнению Фредгольма первого рода [4], в котором одна из подынтегральных функций, в данном случае $h_k(t)$, является неизвестной. Известно множество методов для решения уравнения (1), используемых в зада-

чах адаптивной коррекции, которые можно разделить на две группы: методы расчета во временной области и методы расчета в частотной области. Разнообразные методы, относящиеся к первой группе, приведены, например, в [5, 6]. Методы второй группы предполагают использование дискретного преобразования Фурье (ДПФ) и обладают рядом преимуществ: во-первых большое разнообразие алгоритмов быстрого преобразования Фурье (БПФ) позволяет значительно уменьшить число вычислительных операций; во-вторых отсутствуют проблемы сходимости решения.

В результате выбран метод регуляризации Тихонова [7], применив который получим решение уравнения (1) в следующей форме:

$$h_k(t, \alpha) = \frac{1}{2\pi} \int_{-\Omega}^{\Omega} G(\omega, \alpha) \cdot \frac{U_0(\omega)}{S_0(\omega)} d\omega, \quad (2)$$

где $S_0(\omega)$, $U_0(\omega)$ – спектры передаваемого и принимаемого тестового сигнала, $G(\omega, \alpha)$ – стабилизирующая функция, зависящая от параметра регуляризации α . При этом существует некоторое значение $\alpha = \alpha_{opt}$, обеспечивающее наименьшее отклонение регуляризованного решения $h_k(t, \alpha)$ от точного в заданной метрике (для данной задачи в метрике C [8]).

Коэффициенты ИХ КФ $h_{кор}(t)$ можно получить, решив уравнение

$$h_k(t) * h_{кор}(t) = \begin{cases} 1, & t = 0, \\ 0, & t \neq 0, \end{cases} \quad (3)$$

подставив вместо $h_k(t)$, полученную из (2) $h_k(t, \alpha)$, после чего осуществить коррекцию принимаемого информационного сигнала, либо непосредственно осуществить коррекцию, решив уравнение

$$\int_{T_1}^{T_2} s_{инф}(t - \tau) \cdot h_k(\tau, \alpha) d\tau = u_{инф}(t), t \in [T_3; T_4], \quad (4)$$

где $s_{инф}(t)$ – передаваемый информационный сигнал, $u_{инф}(t)$ – принимаемый информационный сигнал.

Решение уравнения (4) осуществляется также методом регуляризации Тихонова с той особенностью, что теперь одна из подынтегральных функций $h_k(t, \alpha)$ известна с погрешностью, т.к. она является регуляризованным решением уравнения (1). Таким образом, решение уравнения (4) принимает вид:

$$s_{инф}(t, \alpha) = \frac{1}{2\pi} \int_{-\Omega}^{\Omega} G_1(\omega, \alpha) \cdot \frac{U_{инф}(\omega)}{H_k(\omega, \alpha)} d\omega, \quad (5)$$

где $G_1(\omega, \alpha)$ – стабилизирующая функция, аналогичная стабилизирующей функции в выражении (2).

Отметим, что параметр регуляризации в выражениях (2) и (5) одинаков. В [8] автором показано, что выбор одного общего оптимального значения α_{opt} обеспечивает меньшее отклонение результата коррекции и, следовательно, большую помехоустойчивость.

Анализ влияния различных ЗИ на точность расчета ИХ канала

В монографии Б.И. Николаева [9] приводится вариант реализации последовательного КВ модема с использованием ЗИ. В описываемом модеме предполагалось использование испытательного импульса (или испытательной комбинации) отделенного во времени с обеих сторон пассивными нулевыми ЗИ, достаточными для того, чтобы в точке приема иметь импульсный отклик без межсимвольных наложений. Там же показано, что вероятность ошибки на бит, по сравнению с системой без ЗИ немного меньше (примерно на $1 \div 2 \cdot 10^{-3}$).

Однако такой вид сигнала с балансной амплитудной модуляцией в конечном итоге не нашел широкого применения, так как использование только одного испытательного импульса (символа) или испытательной комбинации, состоящей из одного испытательного импульса и пассивного ЗИ, не позволяет адекватно оценить ИХ канала, а также способствовать периодической подстройке тактовой синхронизации и компенсации доплеровского смещения частоты.

В современных одночастотных модемах [2, 3] в качестве теста, как правило, используется фазоманипулированный сигнал на основе ПСП длительностью несколько миллисекунд (значительно больше интервала многолучевости) без ЗИ и занимающий всю выделенную полосу частот. Вставка ЗИ между тестовым и информационным сигналами приведет к структуре сигнала показанной на рис. 1.



Рис. 1. Структура сигнала с ЗИ

Между тестовым и информационным сигналами длительностью $L_{тест}$ и $L_{инф}$, соответственно, присутствует ЗИ длительностью $L_{ЗИ}$. Длительности $L_{тест}$ и $L_{инф}$ составляют несколько символов, например, в авиационном стандарте ARINC 635 [2] $L_{тест} = 15$, $L_{инф} = 30$. Длительность ЗИ $L_{ЗИ}$ удобно взять кратной длительности нескольких символов, при этом $L_{ЗИ}$ должна быть больше задержки между лучами ΔL .

Рассмотрим возможные варианты ЗИ между тестовым и информационным сигналами. Первым вариантом ЗИ является упомянутая выше пассивная пауза, т.е. интервал длительностью $L_{ЗИ}$, на котором отсутствует передача полезного сигнала. При приеме на этом интервале будем иметь «хвосты» тестового и информационного сигналов и аддитивный шум. Расчет ИХ канала может проводиться как на интервале тестового сигнала $L_{тест}$, так и на интервале $L_{м+ЗИ} = L_{тест} + 2L_{ЗИ}$. Удвоенное значение $L_{ЗИ}$ используется потому, что ЗИ расположен перед и после тестового сигнала.

Еще одним из возможных видов ЗИ может быть отрезок гармонического колебания. При этом несущая частота этого колебания может быть как равной частоте несущей, так и расположена вблизи крайней частоты выделенной полосы. В этом случае расчет ИХ канала предполагается проводить на длительности тестового сигнала $L_{месм}$. При этом, как в случае пассивной паузы, так и в случае гармонического колебания, принимаемый тестовый сигнал на длительности $L_{месм}$ или $L_{м+ЗИ}$ может быть дополнен нулями до $N \geq 2L_{месм}$.

Кроме того, для повышения точности расчета ИХ канала и соответствующей ИХ КФ автором предлагается использовать структуру ЗИ, учитывающую метод расчета ИХ. Будем считать свертку (1) циклической и решать уравнение в замкнутой области $\{T_1 \leq \tau \leq T_2; T_1 \leq t \leq T_1\}$. Тогда, перейдя к дискретной форме и решая это уравнение методом, основанным на переходе в частотную область, автором предлагается использование ЗИ, являющегося циклическим продолжением теста, что позволит перейти к циклической свертке в частотной области без дополнения нулями. В этом случае задача расчета ИХ канала сводится к решению уравнения:

$$\sum_{l=0}^{L_{месм}-1} s(mod((n-l), L_{месм})) \cdot h(l) = u(n), \quad (6)$$

$$n = 0 \dots L_{месм} - 1,$$

где $mod((n-l), L_{месм})$ – означает операцию взятия по модулю $L_{месм}$, а функции, входящие в это уравнение, имеют вид:

$$\begin{aligned} s(n) &= s_o(n), \quad n = 0 \dots L_{месм} - 1, \\ h(n) &= h_o(n), \quad n = 0 \dots L_{месм} - 1, \\ u(n) &= u_o(n), \quad n = 0 \dots L_{месм} - 1. \end{aligned} \quad (7)$$

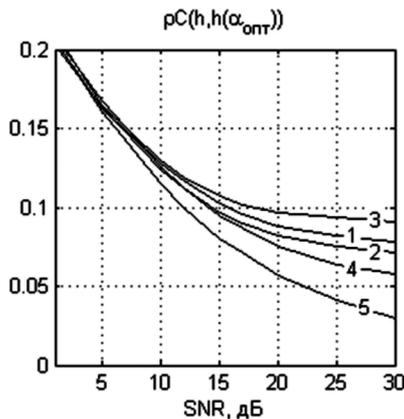


Рис. 2. Зависимости $\rho C(h, h(\alpha_{опт}))$ от ОСШ

при $\Delta L > L_{ЗИ} = 2L_{сумв}$, ДПФ на длительности $L_{месм}$:
 1 – без ЗИ; 2 – пассивная пауза; 3 – гармоническое колебание на частоте несущей; 4 – гармоническое колебание на частоте $f_c = 3,2$ кГц; 5 – циклическое повторение

Проведем сравнительный анализ влияния вида ЗИ на точность расчета ИХ канала. На рис. 2 и 3 показаны зависимости уклонения рассчитанной ИХ канала связи относительно заданной в метрике C от отношения сигнал/шум (ОСШ, SNR) при оптимальном параметре регуляризации $\alpha_{опт}$ для тестового сигнала без ЗИ и различ-

ных вариантов вида ЗИ. Зависимости $\rho C(h, h(\alpha_{опт}))$ усреднены по выборке из 1000 экспериментов, при этом значение параметра регуляризации $\alpha_{опт}$, обеспечивающее минимальное уклонение, в каждом эксперименте выбирается заново. В качестве модели канала выбрана двухлучевая модель с релейскими замираниями, полоса частот 0,3 – 3,4 кГц. Тестовый сигнал взят на длительности $L_{месм}$, ДПФ выполнено на той же длительности. Длительность ЗИ равна двум длительностям символа: $L_{ЗИ} = 2L_{сумв}$, при этом задержка между лучами ΔL : ЗИ $\Delta L < L_{ЗИ}$ (рис. 2); $\Delta L > L_{ЗИ}$ (рис. 3).

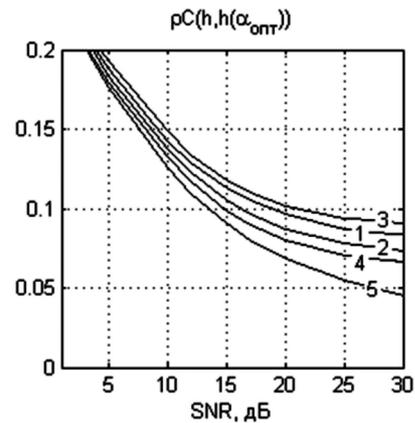


Рис. 3. Зависимости $\rho C(h, h(\alpha_{опт}))$ от ОСШ

при $\Delta L > L_{ЗИ} = 2L_{сумв}$, ДПФ на длительности $L_{месм}$:
 1 – без ЗИ; 2 – пассивная пауза; 3 – гармоническое колебание на частоте несущей; 4 – гармоническое колебание на частоте $f_c = 3,2$ кГц; 5 – циклическое повторение

Как видно, использование ЗИ позволило повысить точность расчета ИХ канала, причем даже при длительности ЗИ меньше интервала многолучевости, т.е. при $\Delta L > L_{ЗИ}$ (см. рис. 3). При этом минимальное уклонение получилось при ЗИ, представляющем собой циклическое повторение, так как в этом случае, взяв сигнал на длительности $L_{месм}$ и выполнив ДПФ на той же длительности, «хвосты» тестового сигнала, появившиеся в результате многолучевости, сохраняются. Кроме того длительность (число отсчетов) получаемой ИХ канала в этом случае будет достаточна и равна $L_{месм}$. Под «достаточностью» имеется ввиду, что число коэффициентов ИХ не должно превышать $L_{месм}$. Более того особенность решения состоит в том, что вторая половина вектора $h(\alpha_{опт})$ соответствует «отрицательному времени», т.е. времени до начала синхронизации, подобно тому, как вторая половина ДПФ соответствует отрицательным частотам. Таким образом, вектор отсчетов ИХ канала $h_k(\alpha_{опт})$ правильной представить в виде:

$$h_k(n, \alpha_{опт}) = \begin{cases} h\left(n + \frac{L_{месм}}{2}, \alpha_{опт}\right), & n = 0 \dots \frac{L_{месм}}{2} - 1, \\ h\left(n - \frac{L_{месм}}{2}, \alpha_{опт}\right), & n = \frac{L_{месм}}{2} \dots L_{месм} - 1. \end{cases} \quad (8)$$

Однако в дальнейшем для расчета ИХ корректора понадобится только $h(n, \alpha_{опт})$ при $n = 0 \dots \frac{L_{месм}}{2} - 1$. Во-

общее говоря, независимо от наличия и вида ЗИ, а также длительности, на которой выполняется ДПФ, для расчета ИХ корректора потребуется не более первой половины отсчетов. Так, дополнив принятый тест нулями до $N > 2L_{мест}$, для расчета ИХ КФ потребуется $N/2$ максимум отсчетов.

На рис. 4 показаны зависимости уклонения рассчитанной ИХ канала связи относительно заданной в метрике S от отношения сигнал/шум (ОСШ, SNR) при оптимальном параметре регуляризации $\alpha_{опт}$ для тестового сигнала без ЗИ и различных вариантов вида ЗИ. Тестовый сигнал взят на длительности $L_{т+ЗИ}$ и дополнен нулями до $N > 2L_{т+ЗИ}$, ДПФ выполнено на длительности N . Задержка между лучами $\Delta L < L_{ЗИ} = 2L_{симв}$.

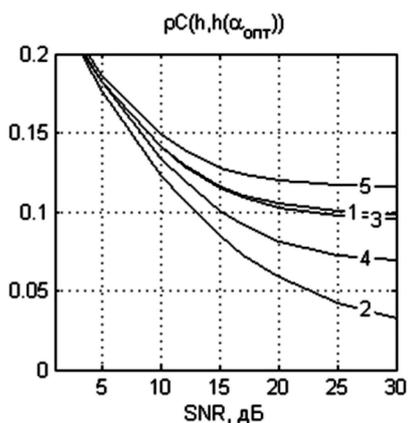


Рис. 4. Зависимости $\rho_C(h, h(\alpha_{опт}))$ от $\sigma_{ш}$

при $\Delta L < L_{ЗИ} = 2L_{симв}$, ДПФ на длительности N : 1 – без ЗИ; 2 – пассивная пауза; 3 – гармоническое колебание на частоте несущей; 4 – гармоническое колебание на частоте $f_c = 3,2$ кГц; 5 – циклическое повторение

В этом случае наименьшее уклонение получилось при использовании пассивной паузы, так как в принимаемом тестовом сигнале на длительности $L_{т+ЗИ}$ сохранены «хвосты» теста, появившиеся в результате многолучевости, и отсутствует наложение «хвостов» информационного сигнала.

Таким образом, применение ЗИ позволяет повысить точность расчета ИХ канала и, соответственно, ИХ КФ, а следовательно, увеличить помехоустойчивость, что будет показано далее. При этом должны быть учтены вид ЗИ и интервал, требуемый для ДПФ, принимаемого тестового сигнала ($L_{мест}$ или $L_{т+ЗИ}$) с точки зрения сохранения «хвостов» тестового сигнала.

Компенсация ЗИ

Рассмотренные выше варианты ЗИ обеспечивают повышение точности расчета ИХ канала, при этом на длительности $L_{мест}$ принимаемого тестового сигнала отсутствуют «хвосты» неизвестного информационного сигнала при $\Delta L < L_{ЗИ}$, либо их влияние не столь значительно при $\Delta L > L_{ЗИ}$. Однако, в процессе коррекции, т.е. при обработке принимаемого информационного сигнала, «хвосты» тестового сигнала, и «хвосты» ЗИ будут накладываться на информационный сигнал. Исключением является лишь применение пассивной паузы в

качестве ЗИ. В этом случае при достаточной длительности ЗИ, т.е. при $\Delta L < L_{ЗИ}$ отсутствует мешающее влияние «хвостов» тестового и информационного сигналов друг на друга. Иными словами пассивная пауза является ЗИ как для тестового сигнала, так и для информационного.

Одним из недостатков такого вида ЗИ является небольшое увеличение значения пик-фактора. Вторым, более существенным недостатком, является влияние автоматической регулировки усиления (АРУ) приемника. Так, при отсутствии многолучевости на месте пассивных пауз будет присутствовать только шум, уровень которого будет увеличен АРУ приемника до уровня сигнала, в результате чего уменьшится ОСШ и точность расчета ИХ канала и ИХ КФ.

Для использования преимуществ пассивной паузы и устранения указанных недостатков разработан способ адаптивной коррекции с компенсацией ЗИ, на который автором получен патент РФ на изобретение № 2573270 [10]. Способ предполагает вставку ЗИ между тестовыми и информационными сигналами таким образом, чтобы ЗИ до и после i -го теста были противоположны по фазе ЗИ до и после $(i+1)$ -го теста. На рис. 5 показана структура передаваемого сигнала с противоположными ЗИ.

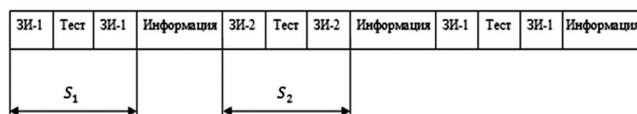


Рис. 5. Структура сигнала с противоположными ЗИ

В качестве ЗИ в таком случае могут быть использованы символы «0» и «1» с двухпозиционной фазовой модуляцией (ФМ-1, BPSK). Тогда при сложении сигналов S_1 и S_2 ЗИ будут скомпенсированы, а амплитуда тестового сигнала удвоится. При этом мощность тестового сигнала увеличится в 4 раза, а дисперсия шума на всей длительности $L_{т+ЗИ}$ вырастет лишь в два раза, в результате чего ОСШ увеличится в 2 раза, что подробно продемонстрировано автором в работе [11] при реализации квазикогерентного сложения. Отметим, что данный подход предполагает выполнение условия:

$$h_j(t) \cong h_{j\pm 1}(t), \quad (9)$$

что, впрочем, не является серьезным ограничением.

Для расчета ИХ канала и ИХ корректора необходимо получить суммарный тестовый сигнал на длительности $L_{т+ЗИ}$, дополнив его нулями до $N > 2L_{т+ЗИ}$. Для коррекции необходимо выделить информационный сигнал на длительности $L_{инф}$, дополнив его нулями до того же значения N .

Результаты моделирования помехоустойчивости

Проведем моделирование помехоустойчивости одночастотного КВ модема передачи данных с применением ЗИ в виде циклического повторения, пассивной паузы, а также при использовании способа компенсации ЗИ.

Определим длительность информационной последовательности для обеспечения линейки скоростей 1200, 2400 и 3600 бит/с. Относительную информационную скорость можно вычислить по формуле:

$$R = \frac{L_{инф}}{L_{инф} + L_{тест} + 2\Delta L_{ЗИ}}, \quad (10)$$

а соответствующую информационную скорость можно определить как:

$$V_{инф} = R \cdot V \cdot m, \quad (11)$$

где V – частота следования символов, m – кратность модуляции.

Так, в авиационном стандарте ARINC 635 [2] соотношение тест/информация составляет $L_{тест} / L_{инф} = 15 / 30$ символов, частота следования символов $V = 1800$ симв/с, следовательно, информационная скорость при двухпозиционной фазовой манипуляции (ФМ-1, BPSK) без ЗИ: $m = 1$, $L_{ЗИ} = 0$ составляет $R = 2 / 3 \approx 0,667$ и $V_{инф} = 1200$ бит/с. Аналогично, для обеспечения скорости 1200 бит/с при частоте следования символов $V = 1600$ симв/с, потребуется информационная последовательность длительностью $L_{инф} = 45$ символов.

Для сохранения информационной скорости при применении ЗИ потребуется увеличить длительность информационной последовательности в соответствии с табл. 1.

Таблица 1. Требуемые длительности информационной последовательности

$L_{ЗИ}$, симв	$L_{инф}$, симв ($V = 1800$ симв/с)	$L_{инф}$, симв ($V = 1600$ симв/с)
0	30	45
1	34	51
2	38	57
3	42	63
4	46	69

При моделировании выбраны следующие параметры: частота дискретизации: $F_s = 16$ кГц; полоса частот канала связи – 3,1 кГц; частота следования символов – $V = 1600$ симв/с; вид передаваемых сигналов: фазоманипулированные сигналы различной кратности ФМ-1 (BPSK), ФМ-2 (QPSK), ФМ-3 (PSK-8); тестовый сигнал: псевдослучайная последовательность длины 15. В качестве модели канала связи выбрана двухлучевая модель с релейскими замираниями и задержкой между лучами 2 мс. В этом случае потребуется длительность ЗИ порядка 3-4 символов.

На рис. 6 приведены кривые вероятности ошибки на бит в зависимости ОСШ при отсутствии ЗИ и применении ЗИ в виде циклического повторения теста и пассивной паузы для различных кратностей модуляции.

Отметим, что наибольшая помехоустойчивость обеспечивается при применении ЗИ в виде пассивной паузы, т.к. в этом случае пассивная пауза является ЗИ как для тестового сигнала, так и для информационного. Использование ЗИ в виде циклического повторения теста обеспечивает меньшую вероятность ошибки на бит по сравнению со случаем отсутствия ЗИ за счет более качественного расчета ИХ КФ. Однако при коррекции информационного сигнала «хвосты» от ЗИ являются помехой, не позволяющей обеспечить наилучший результат коррекции. Для повышения помехоустойчивости в этом случае потребуется ввести дополнительные ЗИ в

виде циклического повторения информационного сигнала, что позволит достигнуть помехоустойчивости, обеспечиваемой при применении пассивной паузы. К сожалению, в этом случае потребуется дополнительно увеличить длительность информационной последовательности, что может быть допустимо лишь в канале с медленными замираниями.

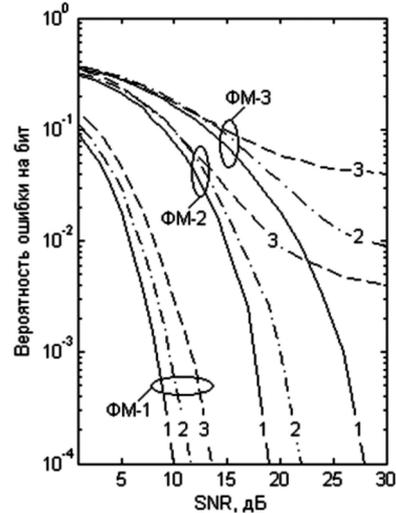


Рис. 6. Зависимости вероятности ошибки на бит от ОСШ в двухлучевом канале для различных кратностей фазовой манипуляции при: 1 – ЗИ пассивная пауза; 2 – ЗИ циклическое повторение теста; 3 – без ЗИ

На рис. 7 приведены кривые вероятности ошибки на бит в зависимости от ОСШ при отсутствии ЗИ и применении способа компенсации ЗИ для различных кратностей модуляции.

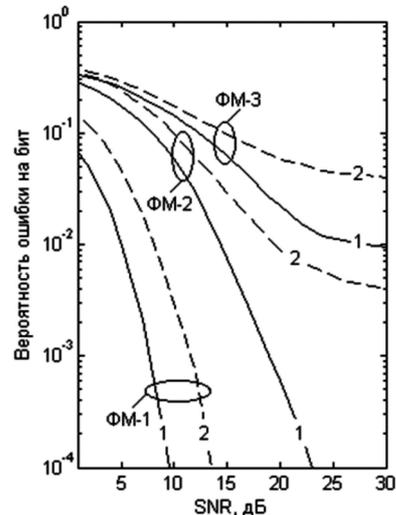


Рис. 7. Зависимости вероятности ошибки на бит от ОСШ в двухлучевом канале для различных кратностей фазовой манипуляции при: 1 – компенсации ЗИ; 2 – без ЗИ

Применение способа компенсации ЗИ, предполагающего использование противоположных по фазе символов, также обеспечивает снижение вероятности ошибки на бит по сравнению со случаем отсутствия ЗИ. Примечательно, что при низких значениях ОСШ вероятность ошибки на бит получилась немного меньше, чем при использовании пассивной паузы. Это объясняется эффектом квазиогерентного накопления теста [11] и увеличением ОСШ при компенсации ЗИ.

В работе был проведен анализ влияния ЗИ различного вида на уклонение рассчитанной ИХ канала связи относительно заданной при моделировании. Автором предложены использование ЗИ в виде циклического продолжения теста, а также новый способ компенсации ЗИ. Показано, что применение ЗИ позволяет повысить точность расчета ИХ канала и, соответственно, ИХ КФ; при этом должен быть учтен вид ЗИ и соответствующий интервал, на котором осуществляется ДПФ. Учет этих особенностей позволяет в значительной степени повысить точность расчета ИХ и, как следствие, помехоустойчивости модема.

Проведенное моделирование показало, что применение ЗИ в одночастотных КВ модемах позволяет повысить их помехоустойчивость в замирающем двухлучевом канале, при сохранении информационной скорости. Кроме того, при определенных условиях появляется возможность повысить информационную скорость передачи данных, перейдя к использованию сигналов повышенной кратности. Так, например, допуская вероятность ошибки на бит порядка $1 \div 3 \cdot 10^{-2}$, т.е. считая, что используемый в модеме помехоустойчивый код обеспечит в этом случае требуемую достоверность, при ОСШ в канале более 20 дБ появляется возможность перейти от ФМ-2 без ЗИ к ФМ-3 с применением ЗИ, тем самым увеличив скорость с 2400 до 3600 бит/с.

Таким образом, применение ЗИ позволяет повысить помехоустойчивость, а в ряде случаев даже повысить информационную скорость.

Литература

1. Бакунин М.Г., Крейнделин В.Б., Шлома А.М., Шумов А.П. Технология OFDM. Учебное пособие для вузов. – М.: Горячая линия – Телеком, 2016. – 352 с.
2. ARINCCharacteristic 635-2. HF Data Link Protocol. – Feb. 27, 1998.
3. MIL-STD-188-110C. Interoperability and Performance Standards for Data Modems. – Sept. 23, 2011.
4. Манжиров А.В., Полянин А.Д. Справочник по интегральным уравнениям: методы решения. – М.: Факториал Пресс, 2000. – 384 с.
5. Джиган В.И. Адаптивная фильтрация сигналов: теория и алгоритмы. М.: Техносфера, 2013. – 528 с.
6. Sayed A.H. Adaptive filters. – New Jersey: Hoboken: John Wiley & Sons, Inc., 2008. – 786 с.
7. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач / Учебное пособие для вузов. – Изд. 3-е испр. – М.: Наука, 1986. – 288 с.
8. Маслаков М.Л., Егоров В.В. Влияние выбора параметра регуляризации на помехоустойчивость в задачах адаптивной коррекции сигналов. Материалы IX Всероссийской конференции «Радиолокация и радиосвязь». Москва, 2015. – С. 182-187.
9. Николаев Б.И. Последовательная передача дискретных сообщений по непрерывным каналам с памятью. – М.: Радио и связь, 1988. – 264 с.
10. Маслаков М.Л. Пат. РФ №2573270. МПК H04L 1/20. Способ адаптивной коррекции с компенсацией защитных интервалов / Егоров В.В., Катанович А.А., Лобов С.А., Маслаков М.Л., Мингалев А.Н., Смаль М.С., Тимофеев А.Е. 2016. Бюл. №2.
11. Маслаков М.Л. Адаптивная коррекция с компенсацией шума. Материалы 16-й Международной конференции «Цифровая обработка сигналов и ее применение – DSPA2014». Москва, 2014. – С. 220-223.

УДК 621.396.96

МИНИМИЗАЦИЯ ИНТЕГРАЛЬНОГО УРОВНЯ БОКОВЫХ ЛЕПЕСТКОВ ФАЗОКОДО-МАНИПУЛИРОВАННЫХ СИГНАЛОВ НА ОТРЕЗКЕ ВРЕМЕННОЙ ОСИ

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: litvnov_dmitry@inbox.ru.

MISMATCHED FILTER SYNTHESIS, MINIMIZING INTEGRATED SIDELOBE LEVEL AT THE SEGMENT OF TIME AXIS FOR PHASE-CODED PULSE

Zaytsev G. V., Kondranina N. S., Litvinov D. M.

The paper presents method of mismatched filter synthesis, minimizing integrated sidelobe level at the segment of time axis for binary phase-coded pulse. Main characteristics of the filtering output are analyzed (integrated sidelobe level, loss in signal-to-noise ratio) and their dependences on the synthesis parameters are investigated.

Key words: phase-coded pulse, mismatched filtering, minimization of integrated sidelobe level, low correlation zone.

Ключевые слова: фазокодоманипулированные сигналы, несогласованная фильтрация, минимум интегрального уровня боковых лепестков, зона с малой корреляцией.

Введение

фазокодоманипулированных (ФКМ) сигналов [1, 2], снижающий боковые лепестки сигнала на заданном отрезке временной оси. Свойства ФКМ сигнала определяются кодовой последовательностью $\mathbf{s}_0 = (s_{0 \text{ code}}, s_{1 \text{ code}}, \dots, s_{n-1 \text{ code}})$, задающей начальные фазы сегментов сигнала. В данной работе рассматриваются бинарные последовательности, в которых элементы последовательности s_0 могут принимать только два значения ± 1 и которые наиболее часто используются на практике.

Помимо радиолокации, ФКМ сигналы используются также в физике, системах передачи информации, биомедицинских исследованиях и других областях науки и техники. Задача их оптимальной обработки весьма актуальна во многих приложениях.

Сигнальные характеристики радиолокатора при использовании ФКМ сигнала определяются взаимной корреляционной функцией (ВКФ) принимаемого сигнала и некоторой опорной последовательности $\mathbf{x} = (x_0, x_1, \dots, x_{N-1})$, имеющейся в приемном устройстве. Длина N последовательности \mathbf{x} , вообще говоря, может быть больше n : $N \geq n$, а составляющие ее элементы – действительные числа.

В данной работе доплеровская частота принимаемого сигнала считается нулевой. При этом исследование сигнальных характеристик сводится к анализу взаимной корреляционной функции последовательностей \mathbf{x} и \mathbf{s}_0 [1, 2].

В радиолокации наиболее часто используется согласованная фильтрация, позволяющая максимизировать отношение сигнал/шум на выходе системы обработки. При этом в случае аддитивного гауссовского шума опор-

В работе рассматривается метод синтеза оптимальной опорной последовательности для несогласованной фильтрации фазокодоманипулированных сигналов, обеспечивающий минимизацию интегрального уровня боковых лепестков на заданном отрезке временной оси. Анализируются получаемые характеристики (интегральный уровень боковых лепестков, потери в отношении сигнал/шум) и зависимости характеристик от основных параметров.

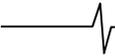
ная последовательность определяется соотношением $\mathbf{x} = \mathbf{s}_0$, $N = n$, а взаимная корреляционная функция становится автокорреляционной (АКФ) [1, 2]. АКФ имеет максимальный лепесток, называемый главным, и ряд боковых лепестков меньшего уровня. Боковые лепестки являются паразитными, а их величина часто превосходит допустимый уровень.

Отступая от согласованности и увеличивая длину опорной последовательности \mathbf{x} , можно добиться снижения уровня боковых лепестков за счет некоторого уменьшения отношения сигнал/шум.

Выбор последовательности \mathbf{x} при заданной последовательности \mathbf{s} зависит от решаемой задачи. Обзор методов синтеза опорной последовательности может быть найден, например, в [3]. В настоящей работе рассматривается задача минимизации интегрального уровня лепестков (строгое определение – в следующем разделе) на заданном отрезке временной оси, или в заданной зоне корреляционной функции. Подобная задача снижения уровня лепестков ВКФ в заданной зоне для группы из нескольких ФКМ сигналов широко обсуждается в приложении к связанным задачам [4]. Однако применительно к одиночным радиолокационным сигналам этот вопрос, насколько известно авторам, в литературе не рассматривался. В данной работе предлагается простой метод решения указанной задачи.

Метод синтеза опорной последовательности

Введем необходимые обозначения. Дополним последовательность \mathbf{s}_0 нулями до длины N . В данной работе будем использовать симметричное расположе-



ние дополнительных нулей относительно середины последовательности, так как асимметрия приводит к ухудшению исследуемых характеристик [5]. Пусть начальный отрезок такой последовательности содержит m нулей. Тогда в конце последовательности также имеется m нулей, и общая длина равна $N = n + 2m$. Обозначим такую последовательность той же буквой, что и исходную, но без нулевого индекса:

$$\mathbf{s}(m) = (0, 0, \dots, 0, s_{0 \text{ code}}, s_{1 \text{ code}}, \dots, s_{n-1 \text{ code}}, 0, 0, \dots, 0) = \{s_i\}, i = 0, 1, \dots, N-1. \quad (1)$$

Аргумент m этой последовательности для краткости будем опускать, если приводимые выражения справедливы при любом m или если значение индекса очевидно из контекста.

Взаимная корреляционная функция $\mathbf{c} = \{c_k\}$, последовательностей \mathbf{x} и $\mathbf{s}(m)$ может быть вычислена по формуле

$$c_k = \sum_{i=0}^{N-1} x_i s_{i+k}^*$$

где считается, что $s_r = 0$ при $r < 0$ и $r \geq N$, а символ $*$ означает операцию комплексного сопряжения для чисел и транспонирования и комплексного сопряжения элементов для векторов и матриц.

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 :

$$\mathbf{D}_k = \begin{pmatrix} \overbrace{0 \ 0 \ \dots \ 0}^{k+1} 1 \ 0 \ 0 \ \dots \ 0 \\ 0 \ 0 \ \dots \ 0 \ 0 \ 1 \ 0 \ \dots \ 0 \\ 0 \ 0 \ \dots \ 0 \ 0 \ 0 \ 1 \ \dots \ 0 \\ \dots \dots \dots \dots \dots \dots \dots \\ 0 \ 0 \ 0 \ \dots \ \dots \ 0 \end{pmatrix}, k \geq 0,$$

и на k элементов вниз при отрицательном k : $\mathbf{D}_{-k} = \mathbf{D}_k^T, k > 0$.

Критерий качества ВКФ, используемый в данной работе, характеризует взвешенный интегральный уровень боковых лепестков относительно главного

$$Q = \frac{|c_M|^2}{\sum_{k \in \Omega, k \neq M} w_k |c_k|^2} \quad (3)$$

где M – индекс главного лепестка, который исключается из суммы знаменателя, а $w_k > 0$ – положительный коэффициент веса k -го лепестка, $k \in \Omega$. Для параметра (3) сохраним название коэффициента качества ВКФ,

общепринятое для случая одинаковых весов [3, 5].

Использование неодинаковых весов позволяет получать разнообразные формы ВКФ. В настоящей работе рассматривается задача снижения уровня боковых лепестков (2) на отрезке временной оси, или в некоторой зоне ВКФ. Эту зону будем задавать множеством $\Psi \in \Omega$ последовательных индексов $k \in \Psi = \{a, a+1, \dots, b\}$,

$M \notin \Psi$, где индекс главного лепестка M исключается из множества Ψ в случае $a \leq M \leq b$. При этом в данной работе веса боковых лепестков в заданной зоне принимаются равными единице, а вне зоны некоторому малому числу $\nu > 0$. При $\nu \rightarrow 0$ лепестки вне зоны перестают учитываться в критерии качества.

Рассмотрим метод, позволяющий при заданной последовательности \mathbf{s} синтезировать последовательность \mathbf{x} , которая максимизирует коэффициент качества (3). Излагаемый метод является обобщением известного метода для случая максимизации стандартного коэффициента качества [3, 5].

Подставляя (2) в (3), получаем формулу для коэффициента качества ВКФ:

$$Q = \frac{|c_M|^2}{\sum_{k \in \Omega, k \neq M} w_k |\mathbf{x} \mathbf{D}_k \mathbf{s}^*|^2} \quad (4)$$

Будем искать максимум величины (4) при условии, что индекс максимального лепестка равен нулю: $M = 0$. Это условие не уменьшает общности, так его всегда можно выполнить путем перераспределения нулей в последовательности \mathbf{s} (1). Однако, как показывают численные расчеты [5], несимметричное расположение нулей менее эффективно и поэтому в данной работе не рассматривается. Тогда из (2) следует, что выражение в числителе (4) равно $c_M = \mathbf{x} \mathbf{s}^*$, и формула (4) приобретает вид

$$Q = \frac{|\mathbf{x} \mathbf{s}^*|^2}{\mathbf{x} \mathbf{R} \mathbf{x}^*}, \quad (5)$$

где матрица \mathbf{R} определяется выражением

$$\mathbf{R} = \sum_{k \in \Omega, k \neq M} w_k \mathbf{D}_k \mathbf{s}^* \mathbf{s} \mathbf{D}_k^* \quad (6)$$

Величина $\mathbf{x} \mathbf{R} \mathbf{x}^*$ включает сумму квадратов всех боковых лепестков с положительными весами. Так как все боковые лепестки не могут быть равны нулю одновременно, то $\mathbf{x} \mathbf{R} \mathbf{x}^* > 0$ для всех векторов \mathbf{x} , и, следовательно, матрица \mathbf{R} является положительно определенной. Поэтому существует единственный положительно определенный квадратный корень $\mathbf{R}^{1/2}$ из матрицы \mathbf{R} ; матрицы \mathbf{R}^{-1} и $\mathbf{R}^{-1/2}$ также являются положительно определенными и удовлетворяют соотношению $\mathbf{U} = \mathbf{U}^*$ [6].

Представим числитель (5) в виде

$$|\mathbf{x} \mathbf{s}^*|^2 = |(\mathbf{x} \mathbf{R}^{1/2})(\mathbf{R}^{-1/2} \mathbf{s}^*)|^2 \quad (7)$$

Применяя к правой части (7) неравенство Коши-Буняковского, получаем

$$|\mathbf{x} \mathbf{s}^*|^2 \leq (\mathbf{x} \mathbf{R} \mathbf{x}^*)(\mathbf{s} \mathbf{R}^{-1} \mathbf{s}^*) \quad (8)$$

Подставляя это выражение в (5), имеем

$$Q \leq \mathbf{s} \mathbf{R}^{-1} \mathbf{s}^* . \quad (9)$$

Максимальное значение коэффициента качества (9) достигается при условии коллинеарности входящих в (7) векторов:

$$\mathbf{x} \mathbf{R}^{1/2} = a \mathbf{s} \mathbf{R}^{-1/2} , \quad (10)$$

где a – некоторый коэффициент. Так как масштабный коэффициент a в данной работе несущественен, то, полагая его равным единице, из последнего соотношения получаем формулу для вычисления оптимальной последовательности \mathbf{x} :

$$\mathbf{x} = \mathbf{s} \mathbf{R}^{-1} . \quad (11)$$

Для действительных последовательностей \mathbf{s} получаемые опорные последовательности \mathbf{x} (11) также действительны. В дальнейших разделах рассматриваются свойства оптимального решения (11).

Основные характеристики оптимальных ВКФ при $N = n$

В приводимых ниже результатах используются следующие параметры ВКФ. Первый из них – это коэффициент качества (3). Так как коэффициент качества изменяется в широких пределах, то ниже он вычисляется в децибелах: $10 \log Q$ дБ.

Второй параметр, весьма важный для радиолокационных применений, – отношение сигнал/шум на выходе алгоритма обработки. В случае аддитивного гауссовского шума со среднеквадратичным значением σ это отношение может быть вычислено по формуле

$$SNR(\mathbf{s}, \mathbf{x}) = \frac{|c_M|^2}{\sigma^2 \sum_{k \in \Omega} |x_k|^2} . \quad (12)$$

Величина уменьшения этого отношения по сравнению со случаем согласованной фильтрации называется энергетическими потерями или просто потерями, оценивается в децибелах и вычисляется по формуле

$$L = 10 \log[SNR(\mathbf{s}, \mathbf{s}) / SNR(\mathbf{s}, \mathbf{x})] . \quad (13)$$

Нетрудно видеть, что энергетические потери (13) не зависят от величины σ .

В литературе часто используется третий параметр – относительный уровень максимального бокового лепестка (в данном случае в исследуемой зоне подавления), определяемый как отношение уровня главного лепестка к уровню максимального бокового и вычисляемый в децибелах. В данной работе этот параметр практически не используется, так как исследуемые ниже зависимости этого параметра и коэффициента качества аналогичны, при этом относительный уровень максимального бокового лепестка больше коэффициента качества на величину порядка 10 дБ.

Описываемые ниже характеристики ВКФ для оптимального решения (11) не удается получить аналитическим путем, что связано с разнообразием свойств исходных последовательностей. Поэтому приводимые ниже закономерности получены путем анализа результатов расчета ВКФ на ПЭВМ для многочисленных последовательностей. Эти закономерности не являются строгими, а носят эмпирический характер.

Перейдем к анализу ВКФ. Проанализируем сначала

случай, когда длины сигнала и опоры одинаковы: $m = 0, N = n$. Эти значения являются, при прочих равных условиях, наиболее целесообразными. Действительно, при $N < n$ в главном лепестке не накапливается весь сигнал, и, следовательно, возникают потери. С другой стороны, увеличение длины опоры сверх длины сигнала приводит к ряду нежелательных последствий: возрастанию длины ВКФ, увеличению энергетических потерь, повышению вычислительной сложности алгоритма фильтрации и т. д. В этом разделе положим также, что зона подавления расположена симметрично относительно главного лепестка ВКФ.

Типичный вид ВКФ, получаемой при синтезе опорной последовательности по формуле (11), приведен на рис. 1. Величина главного лепестка на рисунке нормирована к единице (0 дБ). В качестве исходной последовательности для этого и последующих рисунков использована последовательность длины $n = 211$ с максимальным известным коэффициентом качества АКФ [7], имеющая в шестнадцатеричном представлении вид

70FE7E3FDC794CA6E31035FA43BB47A835136E7CC1A48AB6A6AD2.

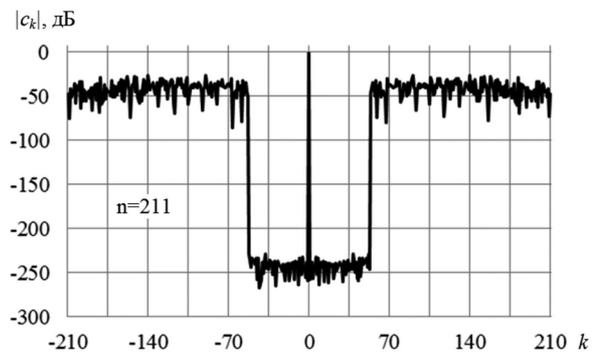


Рис. 1. Типичный вид ВКФ с зоной подавления лепестков

Модулирующая последовательность \mathbf{s}_0 получается из представленной последовательности путем преобразования ее в двоичный вид и замены нулей на -1. Ширина z зоны подавления на рис. 1 $z = 111$, весовые коэффициенты по принятому в данной работе соглашению в зоне подавления равны 1, вне зоны – $\nu = 10^{-10}$.

Как видно из рисунка, опорная функция, синтезированная по формуле (11), обеспечивает практически полное подавление лепестков в заданной зоне (уровень максимального бокового лепестка в зоне – 230 дБ, коэффициент качества $Q = 219$ дБ). Чем меньше значение весовых коэффициентов вне зоны подавления, тем больше уровень подавления в зоне. С точки зрения решаемой задачи целесообразно выбирать как можно меньшее значение веса ν вне зоны подавления. Однако при слишком малом значении ν и конечной разрядности вычислений матрица (6) может стать вырожденной. Опыт расчетов показывает, что для исключения вырожденности матрицы при использовании 64-разрядной арифметики с плавающей запятой значение весового коэффициента ν целесообразно выбирать не менее 10^{-12} .

Количество переменных в синтезируемом векторе \mathbf{x} равно N . Поэтому из общих соображений ясно, что значительно подавить боковые лепестки можно в зоне

шириной $z < N$. Расчеты подтверждают этот факт. На рис. 2 показана типичная зависимость коэффициента качества ВКФ (3) от ширины зоны подавления. На этом и последующих рисунках ширина зоны z нормирована к длине N -опорной последовательности x . В качестве исходной использована последовательность предыдущего примера. Зависимости рис. 2 приведены для трех значений весового коэффициента ν вне зоны подавления – 10^{-4} , 10^{-7} , 10^{-10} .

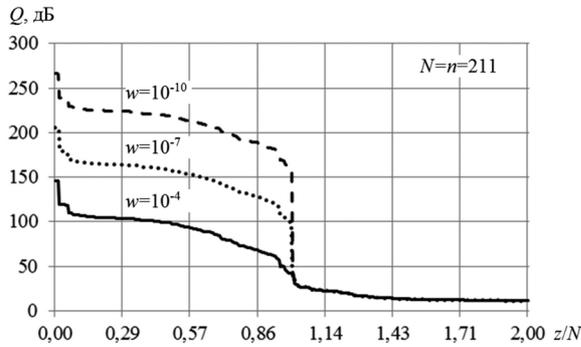


Рис. 2. Зависимость коэффициента качества от ширины зоны подавления

Как и следовало ожидать, коэффициент качества весьма значителен, пока ширина зоны меньше длины сигнала: $z < N$. В этой области величина весового коэффициента ν значительно влияет на коэффициент качества. При уменьшении ν на порядок, коэффициент качества растет примерно на 20 децибел. При дальнейшем расширении зоны подавления ($z > N$) коэффициент качества быстро стремится к значению, которое можно получить, не выделяя зоны [5]. Поэтому в дальнейшем рассматриваются характеристики только для зон шириной $z < N$. «Релейный» вид зависимостей, подобных приведенным на рис. 2, имеют графики для всех последовательностей как при $m = 0$, так и при $m > 0$.

На рис. 3 показана зависимость энергетических потерь от ширины зоны для того же примера при двух значениях весового коэффициента – 10^{-3} и 10^{-7} . Потери велики, если ширина зоны подавления близка к длине сигнала (величина $|z - N|$ невелика). Для других z потери существенно меньше. Подобное поведение функции энергетических потерь вызвано следующими причинами. Когда ширина зоны подавления больше числа переменных N в синтезируемом векторе x , их количество недостаточно для минимизации уровня всех лепестков, стоящих в знаменателе коэффициента качества (3). Максимизация коэффициента качества происходит за счет увеличения числителя (3), т.е. фильтр приближается к согласованному, и потери невелики. При малой ширине зоны подавления количество переменных достаточно как для уменьшения знаменателя (3), так и для максимизации числителя. При этом фильтр снова приближается к согласованному. При $z \approx N$ отличие от согласованности существенное, и потери могут быть значительны. Необходимо также отметить, что в этой области потери возрастают при уменьшении весового коэффициента ν вне зоны подавления. Для снижения потерь целесообразно выбирать максимальное значение этого коэффициента, ко-

торое еще обеспечивает необходимый уровень подавления лепестков в зоне.

В связи с важностью для радиолокационных применений параметра энергетических потерь рассмотрим их величину для следующих семейств исходных последовательностей s_0 : M -последовательности [8] длины $2^k - 1$ в диапазоне $k = 6 - 10$, последовательности Лежандра [9] до длины 300 и последовательности с наилучшим известным коэффициентом качества АКФ, табулированные до длины 300 [7]. Первые два семейства имеют асимптотический коэффициент качества АКФ (без зоны, т.е. $w \equiv 1$), соответственно, 3 и 6 в линейном масштабе, для последнего семейства средний коэффициент качества равен 7,6 [7, 9]. Краткое описание этих последовательностей может быть также найдено в [5].

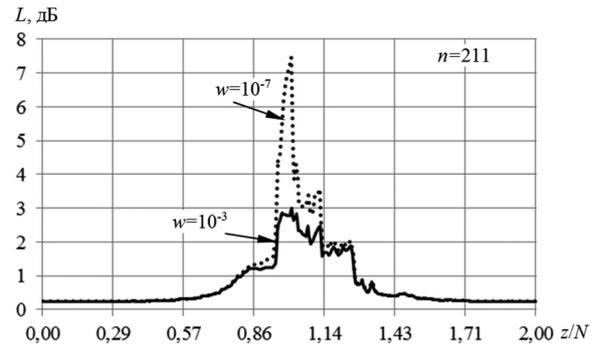


Рис. 3. Зависимость потерь от ширины зоны подавления

На рис. 4 приведены зависимости средних энергетических потерь от ширины z зоны подавления для трех указанных семейств. Весовые коэффициенты вне зоны подавления приняты равными $\nu = 10^{-10}$. Зависимость для каждого семейства рассчитана как усредненное значение потерь по последовательностям данного семейства. Дисперсия такой оценки, как показывают расчеты, невелика при $z < 0,9n$. При увеличении ширины зоны потери растут. Среднее значение потерь не превосходит 1 дБ при ширине зоны $z < (0,7 - 0,9)n$ в зависимости от семейства. Для последовательностей Лежандра и последовательностей с наилучшим коэффициентом качества АКФ средние потери не превосходят 0,5 дБ при $z < 0,5n$.

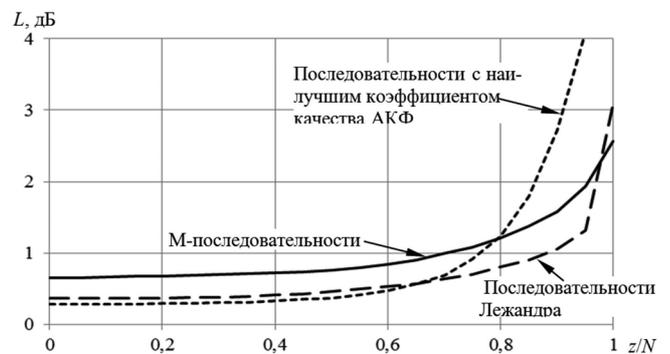
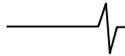


Рис. 4. Зависимость средних потерь от ширины зоны подавления

При небольшом значении потерь их величина, как правило, тем меньше, чем более высокий коэффициент качества АКФ имеет исходная последовательность. Та-



ким образом, для подавления боковых лепестков в зоне целесообразно в качестве исходных использовать последовательности с максимально возможным коэффициентом качества АКФ. Аналогичный вывод справедлив и для максимизации коэффициента качества без выделения зоны [5].

Влияние других параметров на характеристики ВКФ

В предыдущем пункте рассмотрены основные характеристики ВКФ для случая $N = n$. Однако поведение рассмотренных функций не изменяется, если $N > n$, $m > 0$. Поэтому на представленных выше графиках использован параметр N вместо n .

В частности, при увеличении N возрастает количество переменных, которые можно использовать для максимизации коэффициента качества, что приводит к возможности пропорционального расширения зоны подавления. Это положение иллюстрирует рис. 5, на котором для приведенной выше последовательность длины $n = 211$ показаны зависимости потерь от ширины зоны подавления при $z < N$ для трех величин параметра m : $m = 0, n, 2n$.

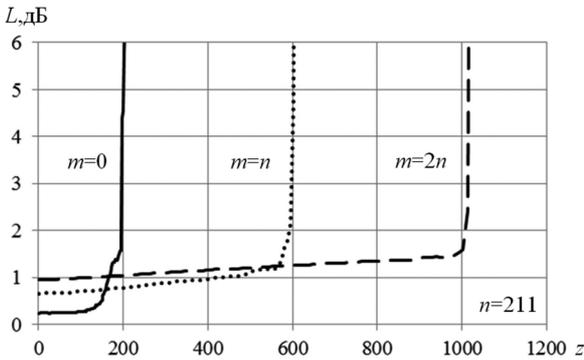


Рис. 5. Увеличение максимальной ширины зоны подавления при возрастании m

Рисунок наглядно демонстрирует возможность расширения зоны при увеличении длины опорной последовательности. Вместе с тем, с увеличением m происходит также некоторое возрастание потерь.

Приведенные выше результаты относятся к симметричному расположению зоны подавления боковых лепестков относительно главного лепестка. Рассмотрим, как изменяются результаты при отказе от симметрии. На рис. 6 показана зависимость коэффициента качества Q от положения центра зоны подавления на временной оси для использованной в предыдущих примерах исходной последовательности. Длина опорной последовательности для рис. 6а равна $N = n$, для рис. 6б – $N = 3n = 633$ ($m = n$). Ширина зоны подавления в обоих случаях равна $z = 41$, весовые коэффициенты вне зоны подавления $v = 10^{-10}$. По оси абсцисс отложена координата g центра зоны подавления относительно середины ВКФ.

Как следует из рис. 6, наилучшее подавление боковых лепестков происходит при расположении центра зоны в районе середины ВКФ. При сдвигах зоны коэффициент качества падает, при этом при больших значениях m (рис. 6б) существенно меньше, чем при малых (рис. 6а). Близкий к линейному в логарифмическом

масштабе спад характеристик имеет место на интервалах длины n у краев ВКФ. Тем не менее, коэффициент качества остается достаточно большим для практических применений.

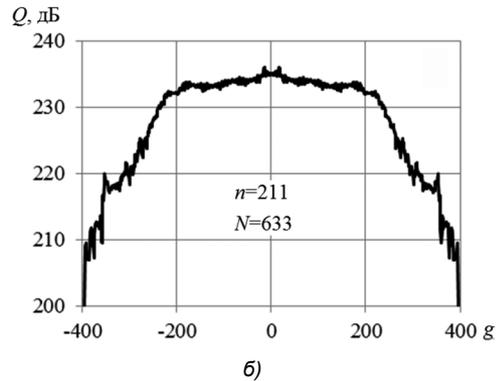
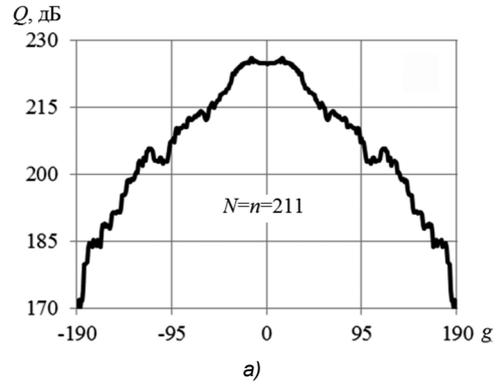


Рис. 6. Зависимость коэффициента качества от положения центра зоны подавления

На рис.7 показаны зависимости энергетических потерь от сдвига зоны подавления для тех же двух случаев. В случае $m = 0$ (рис. 7а) потери возрастают в несколько раз в диапазоне возможных сдвигов, оставаясь при этом приемлемыми. В случае $m = n$ (рис. 7б) изменение величины потерь составляет всего около 10 % во всем диапазоне сдвигов, хотя сама величина больше, чем для случая $m = 0$ при центральном расположении зоны. Таким образом, при сдвиге зоны подавления от центра наблюдается небольшая деградация как коэффициента качества, так и энергетических потерь, причем при значительной величине параметра m она существенно меньше.

Еще одно заслуживающее внимание свойство оптимальной ВКФ состоит в том, что подавление боковых лепестков в зоне не сильно сказывается на уровне лепестков вне зоны, если ширина зоны подавления не слишком велика. Для иллюстрации этого свойства на рис. 8 сравниваются две оптимальные ВКФ для вышеприведенной последовательности длины 211 при параметре $m = 0$. Одна из них получена минимизацией интегрального уровня лепестков при $w_k \equiv 1$ (кривая из точек). Другая – с выделением зоны подавления шириной $z = 100$, расположенной симметрично относительно главного лепестка при весовых коэффициентах вне зоны $v = 10^{-10}$ (непрерывная кривая). Для второй функции уровень максимального лепестка для всей ВКФ больше всего на 1,5 дБ, а энергетические потери возросли на 0,5 дБ. Иными словами, цена за создание зоны подавления весьма невелика.

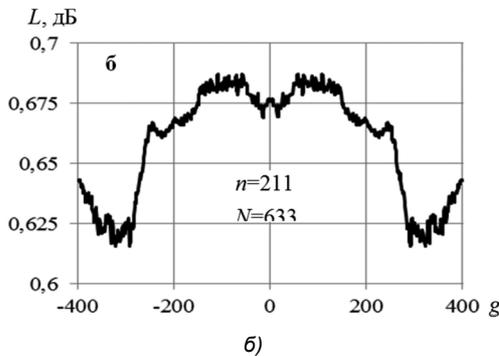
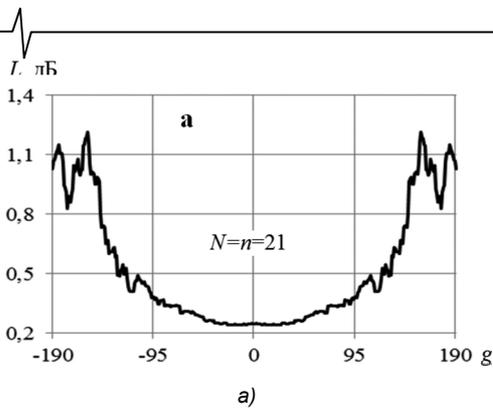


Рис. 7. Зависимость энергетических потерь от положения центра зоны подавления

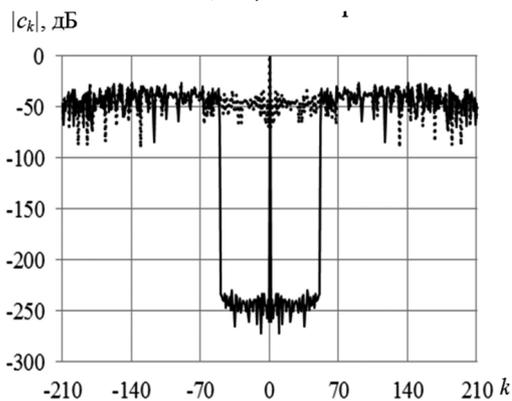


Рис. 8. Сравнение ВКФ с зоной и без зоны подавления давления

Описанный эффект носит систематический характер. Например, для исходных последовательностей Лежандра создание зоны подавления шириной $z < N/2$ ухудшает уровень максимального лепестка вне зоны подавления не более чем на 2 дБ.

Заключение

В работе предложен метод синтеза оптимальной опорной последовательности для несогласованной фильтрации фазокодированных сигналов, обеспечивающий минимизацию интегрального уровня боковых лепестков на заданном отрезке временной оси.

Показано, что результирующая ВКФ обладает следующими свойствами:

- в заданной зоне обеспечивается практически полное подавление боковых лепестков при ширине зоны $z < N$;
- для опорной функции, имеющей длину сигнала, энергетические потери лежат в диапазоне 0,5-1 дБ при ширине зоны $z < (0,5 - 0,8)n$ и выборе исходной последовательности с высоким коэффициентом качества АКФ;

– за счет увеличения длины опорной последовательности ширина зоны подавления может быть пропорционально увеличена;

– при несимметричном расположении зоны подавления относительно главного максимума ВКФ величина подавления несколько снижается, оставаясь, тем не менее, достаточно высокой для практических применений;

– при небольшой ширине зоны подавления уровень боковых лепестков ВКФ вне зоны изменяется незначительно по сравнению с ВКФ, синтезированной без зоны подавления, так что «стоимость» введения зоны невелика.

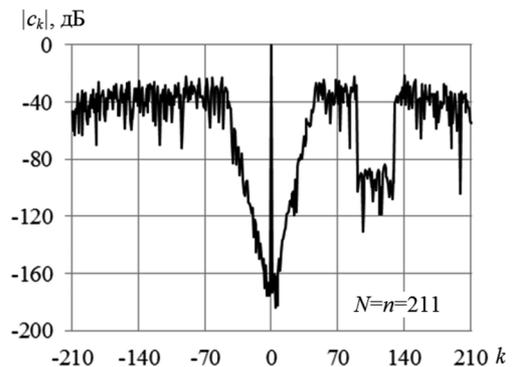


Рис. 9. Пример ВКФ сложной формы

Отметим также, что рассмотренный метод синтеза опорной последовательности позволяет получать ВКФ произвольно заданной формы путем соответствующего изменения весовых коэффициентов при синтезе. В качестве примера на рис. 9 приведена ВКФ с двумя зонами подавления, в одной из которых веса выбраны неодинаковыми для получения треугольной характеристики в логарифмическом масштабе.

Литература

1. Levanon N., Mozeson E. Radar Signals. John Wiley&Sons Inc. New Jersey. 2004. 412 p.
2. Трухачев А.А. Радиолокационные сигналы и их применения. М. Воениздат. 2005. 320 с.
3. Stoica P., Li J., and Xue M. Transmit codes and receive filters for radar. IEEE Signal Processing Magazine. Vol. 25. pp. 94–109. November 2008.
4. Jung-Soo Chung and Jong-Seon No. Low Correlation Zone Sequences. Sequences and Their Applications – SETA 2010: 6th International Conference. Paris, France. September 13-17, 2010. Proceedings. Springer-Verlag, Berlin, Heidelberg. pp. 1-29.
5. Зайцев Г.В., Кондранина Н.С., Литвинов Д.М. Оценка характеристик метода несогласованной фильтрации, минимизирующего интегральный уровень боковых лепестков фазокодированных сигналов // Цифровая обработка сигналов. № 1. 2017. с. 61-69.
6. Horn R.A., Johnson C.R. Matrix Analysis. Cambridge University Press. 1985. 652 p.
7. Knauer J. Merit Factor Records. Nov. 2004. Available in Oct. 2016 at URL: http://labraj.feri.um.si/en/Low-Autocorrelation_Binary_Sequence_Problem.
8. Питерсон У., Уэлдон Э. Коды, исправляющие ошибки. М. Мир. 1976. 594 с.
9. Jedwab J. A Survey of the Merit Factor Problem for Binary Sequences. Sequences and Their Applications. Proceedings of SETA 2004. ed. T. Helleseth et al. Lecture Notes in Computer Science. Vol. 3486. pp. 30–55. Springer-Verlag, Berlin, Heidelberg. 2005.

УДК 621.396.96

ВЛИЯНИЕ СКОРОСТИ ОБЪЕКТА НА ЕГО РАДИОЛОКАЦИОННОЕ ИЗОБРАЖЕНИЕ, ФОРМИРУЕМОЕ В РАДИОЛОКАТОРЕ БОКОВОГО ОБЗОРА С СИНТЕЗИРОВАННОЙ АПЕРТУРОЙ

Брызгалов А.П., начальник лаборатории, к.т.н., доцент ФГУП «Государственный научно-исследовательский институт авиационных систем», e-mail: a_bryzgalov@gosniias.ru.

EFFECT OF THE OBJECT VELOCITY ON THE RADAR IMAGE FORMED BY A SYNTHETIC-ARRAY SIDE-LOOKING RADAR

Bryzgalov A.P.

We estimate the offset and distortion of radar images formed by a side-looking synthetic-aperture radar while locating moving objects. We study the effect of radial and tangential velocity of objects. The results are proved on the model of radar images formation coordinated with the area location (fixed objects).

Key words: radar, synthetic aperture, side looking, radar image, shift, velocity effect, tangential velocity, radial component.

Ключевые слова: радиолокатор, синтезированная апертура, боковой обзор, радиолокационное изображение, смещение, влияние скорости, тангенциальная составляющая, радиальная составляющая.

Введение

В радиолокаторах бокового обзора с синтезированной апертурой (РСА), осуществляющих дистанционное зондирование земной поверхности и формирование ее радиолокационных изображений (РЛИ), при лоцировании движущихся объектов происходит смещение изображений этих объектов на местности, а также их искажение. Это создает определенные трудности как при решении задач обнаружения и распознавания движущихся объектов, оценки их положения и скорости, так и при летных испытаниях и математическом моделировании систем с синтезированной антенной решеткой (САР). Частично смещение РЛИ движущихся объектов уже рассматривалось, в частности в [1], но многие вопросы требуют дальнейших исследований.

Влияние радиальной составляющей скорости лоцируемого объекта на его РЛИ

Влияние радиальной составляющей скорости объекта, т.е. скорости $V_{об_x}$, направленной вдоль оси X , при том, что траектория носителя РСА совпадает с осью Y , может быть оценено, исходя из рис. 1. Для упрощения анализа полагаем, что движение объекта происходит только вдоль оси X . На рис. 1 показано взаимное расположение объекта и носителя при i -м зондировании: объект находится в точке A , совпадающей с m -м элементом поверхности с координатами x_{mi} , y_{mi} , а носитель РСА находится в точке с координатами 0 , y_i .

В РСА при формировании РЛИ m -го элемента поверхности при каждом зондировании (на интервале синтеза его РЛИ) осуществляется согласованная внутриимпульсная (частотно-временная), а затем и

Оцениваются смещение и искажения радиолокационных изображений, формируемых в радиолокаторе бокового обзора местности с синтезированной апертурой при лоцировании движущихся объектов. Рассматривается влияние радиальных и тангенциальных составляющих скорости объектов. Результаты подтверждаются на модели формирования радиолокационных изображений, согласованной с лоцированием местности (неподвижных объектов).

межпериодная (за интервал синтеза) обработка входного сигнала. При этом из параметров, характеризующих взаимное расположение и движение носителя РСА и лоцируемого объекта, при обработке используются 2 параметра (см., например, [2]): задержка эхосигнала, определяемая дальностью $R_{лоц_mi}$ до m -го элемента – расстоянием между фазовым центром антенны РСА и этим элементом, и радиальная составляющая скорости их взаимного перемещения $V_{лоц_r_mi}$. Необходимо совпадение этих параметров входного сигнала и обработки, чтобы получить максимальный сигнал РЛИ, соответствующий согласованной обработке для m -го элемента поверхности.

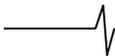
Для локатора бокового обзора при лоцировании неподвижных объектов скорость $V_{лоц_r_mi}$ на этапе синтеза, исходя из рис. 1, меняется в соответствии с выражением

$$V_{лоц_r_mi} = V_{нос_r_mi} = V_{нос} \sin \alpha_{mi}, \quad (1)$$

где $V_{нос}$ – скорость носителя РСА на интервале синтеза, $V_{нос_r_mi}$ – радиальная составляющая этой скорости, соответствующая проекции скорости носителя на линию визирования m -го элемента носителем, α_{mi} – угол визирования m -го элемента на i -ом зондировании относительно нормали к траектории носителя, который определяется выражением

$$\operatorname{tg} \alpha_{mi} = \Delta y_{mi} / R_{mi}, \quad (2)$$

где R_{mi} – расстояние от m -го элемента до траектории полета носителя; при совпадении положения объекта с положением m -го элемента R_{mi} равно расстоянию $R_{об\dot{i}}$ от объекта до траектории (см. рис. 1),



$$\Delta y_{mi} = y_{mi} - y_i.$$

При этом согласованная обработка для каждого m -го элемента рассчитана на N зондирований за интервал синтезирования $D_{сум}$ при изменении Δy_{mi} в пределах $\pm D_{сум}/2$.

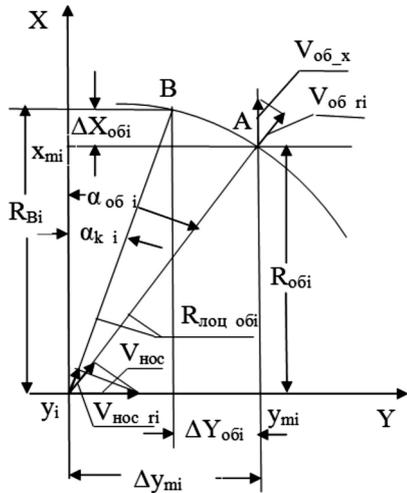


Рис. 1. Взаимное расположение носителя, объекта и его РЛИ при i -м зондировании ($V_{об_x} > 0$)

При нахождении носителя на траверзе m -го элемента ($\Delta y_{mi} = 0$), что для лоатора бокового обзора при лоцировании неподвижного объекта соответствует середине интервала синтезирования, $\alpha_{mi} = 0$, и поэтому входной сигнал и согласованный с ним опорный сигнал имеют параметры: $V_{лоц_r_mi} = 0$ и $R_{лоц_mi} = R_{mi}$.

При лоцировании движущегося объекта параметры меняются. В общем случае, в соответствии с рис. 1, $V_{лоц_r_mi} = V_{нос} \sin \alpha_{mi} - V_{об_x} \cos \alpha_{mi}$, (3) и, в частности, на траверзе m -го элемента, когда $\alpha_{mi} = 0$, $V_{лоц_r_mi} = V_{об_x}$. То есть, в этом случае применение для m -го элемента обработки, согласованной с неподвижным объектом, приводит к несогласованной обработке сигнала.

Однако на каждом зондировании, в целях формирования РЛИ участка местности, проводят обработку входного сигнала, согласованную с другими точками, и при этом угол визирования α_{mi} для этих точек меняется. При этом может оказаться, что для некоторого направления при угле визирования α_{ki} скорость $V_{лоц_r_ki}$, получаемая из (1) при лоцировании неподвижного объекта, будет равна радиальной составляющей (3) для движущегося объекта при угле визирования α_{mi} , т.е. выполняется равенство:

$$V_{нос} \sin \alpha_{ki} = V_{нос} \sin \alpha_{mi} - V_{об_x} \cos \alpha_{mi} \quad (4)$$

Это означает, что сигнал, отраженный от движущегося объекта, совпадающего с m -м элементом РЛИ, пройдет согласованную обработку для k -го элемента, для которого угол визирования равен α_{ki} , а второй параметр обработки $R_{лоц_ki} = R_{лоц_mi}$. При этом сигнал, прошедший согласованную обработку, окажется не в точке A – в точке нахождения объекта, а переместится в точку B (см. рис. 1) – в точку пересечения окружности (с радиусом $R_{лоц_mi}$ и центром в точке расположения носителя РСА y_i) с линией визирования под углом α_{ki} .

Исходя непосредственно из рис. 1, выражение (4) можно записать в виде:

$$V_{нос} \frac{\Delta y_{ki}}{R_{лоц_mi}} = V_{нос} \frac{\Delta y_{mi}}{R_{лоц_mi}} - V_{об_x} \frac{R_{mi}}{R_{лоц_mi}}, \quad (5)$$

а так как $\Delta y_{ki} = \Delta y_{mi} - \Delta Y_{об_i}$, то из этого следует, что смещение по оси Y изображения объекта, движущегося вдоль оси X ,

$$\Delta Y_{об_i} = R_{об_i} \frac{V_{об_x}}{V_{нос}}, \quad (6)$$

что при $R_{об_i} \gg D_{сум} V_{об_x} / V_{нос}$ можно записать как

$$\Delta Y_{об} \approx R_{об} \frac{V_{об_x}}{V_{нос}} \quad (7)$$

и что соответствует [1].

Из рис. 1 следует так же, что при i -ом зондировании смещение на РЛИ положения движущегося объекта по оси X

$$\Delta X_{mi} = R_{mi} - R_{Bi}. \quad (8)$$

Учитывая в (8), что $R_{Bi} = R_{лоц_mi} \cos \alpha_{ki}$, а $R_{лоц_mi} = R_{mi} / \cos \alpha_{mi}$, можно записать

$$\Delta X_{mi} = R_{mi} \left(1 - \frac{1}{\cos \alpha_{mi}} \cos \alpha_{ki} \right). \quad (9)$$

Вместе с тем, из (4) при малых значениях α_{ki} и α_{mi} следует равенство:

$$\alpha_{ki} \approx \frac{V_{нос} \alpha_{mi} - V_{об_x}}{V_{нос}}. \quad (10)$$

В этом случае можно показать, что

$$\Delta X_{mi} \approx -\frac{R_{mi}}{2} \left(2\alpha_{mi} \frac{V_{об_x}}{V_{нос}} - \frac{V_{об_x}^2}{V_{нос}^2} \right). \quad (11)$$

На интервале синтезирования α_{mi} меняется от $D_{сум}/2R_{mi-нач}$ в начале интервала до 0 в центре интервала и до $-D_{сум}/2R_{mi-кон}$ на последнем зондировании, и в соответствии с (11) меняется смещение ΔX_{mi} . Но за счет движения самого объекта происходит также реальное смещение этого объекта относительно его положения на начало интервала синтезирования на величину

$$\Delta X_{об_i} \approx \left[\frac{D_{сум}}{2} - \Delta y_{mi} \right] \frac{V_{об_x}}{V_{нос}}. \quad (12)$$

В соответствии с изменением Δy_{mi} значение $\Delta X_{об_i}$ за интервал синтезирования меняется от 0 до

$$\Delta X_{об_i} \approx D_{сум} \frac{V_{об_x}}{V_{нос}}.$$

В результате при каждом зондировании суммарное смещение изображения за счет первого слагаемого в (11) и смещения (12) сводит все сигналы, получаемые от объекта за интервал синтезирования, в одну точку по дальности, соответствующую положению центральной точки интервала синтезирования. Соответственно, не происходит и расширение ГЛ РТИ по дальности (по оси X). Но максимум РЛИ объекта смещен относительно указанной точки по оси X на величину ΔX_{m_pez} , определяемую вторым слагаемым в (11).

Значения величин ΔY_{mi} и ΔX_{m_pez} для РЛИ можно проиллюстрировать численными примерами. Так, при $R_{об_m} = 10000$ м, $V_{нос} = 100$ м/с, $V_{об_x} = -0,1$ м/с в соответствии с (7) и вторым слагаемым в (11) $\Delta Y_{mi} = -10$ м, а $\Delta X_{m_pez} = 0,005$ м. При возрастании скорости объекта до 1 м/с – $\Delta Y_{mi} = -100$ м и $\Delta X_{m_pez} = 0,5$ м.

Полученные результаты были проверены путем математического моделирования РСА. Использовалась корреляционно-фильтровая согласованная обработка, подробно изложенная в [2] и [3]. На рис. 2-5 приведены результаты моделирования РСА при согласованной обработке когерентной последовательности импульсов на несущей частоте 1 ГГц с линейной частотной модуляцией (ЛЧМ) в импульсе с девиацией 200 МГц. Длительность каждого импульса составляла 0,2 мс, период повторения – 10 мс. Использовалась обработка пачки импульсов за 10 с, что при скорости носителя $V_{нос} = 100$ м/с обеспечивало интервал синтезирования $D_{синт} = 1000$ м.

На рис. 2 приведено яркостное РЛИ (в координатах X, Y) при наличии единственного отражателя с координатами на момент прохождения центра интервала синтезирования $x = 10000$ м, $y = 0$ (т.е. $R_{об} = 10$ км) при $V_{нос} = 100$ м/с и $V_{об,x} = -0,1$ м/с. На рис. 3 приведено то же РЛИ в формате 3D. Из рис. 2 следует, что смещение максимума отклика на одиночный точечный движущийся объект по оси Y составляет -10 м, а по оси X – незначительно, что соответствует ранее полученным численным результатам.

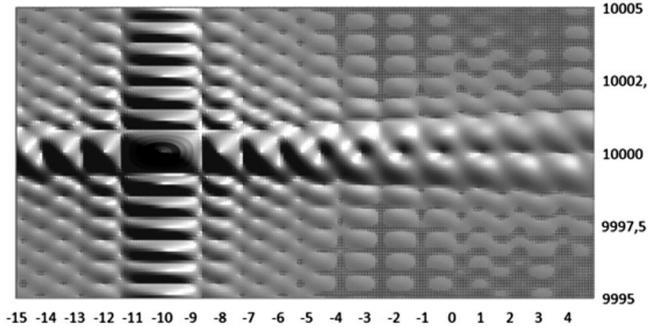


Рис. 2. РЛИ отражателя (в координатах X, Y) при $x = 10000$ м, $y = 0$ и $V_{об,x} = -0,1$ м/с

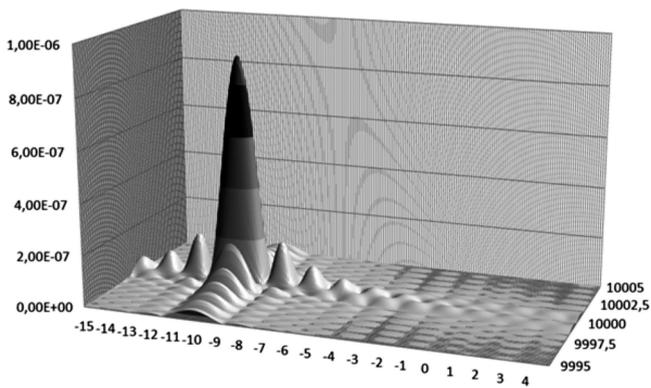


Рис. 3. РЛИ (в формате 3D) отражателя при $x = 10000$ м, $y = 0$ и $V_{об,x} = -0,1$ м/с

На рис. 4 приведено двумерное (яркостное) РЛИ при тех же исходных данных, что и для рис. 2, но при $V_{об,x} = -1$ м/с. Из рис. 4 и численных данных, полученных при его вычислении, следует, что $\Delta Y_m = -100$ м и $\Delta X_{m,рез} = 0,6$ м, что также соответствует ранее полученным аналитическим данным.

Представляет интерес анализ влияния радиального движения объекта на его РЛИ в азимутальном направлении. Косвенно изменения формы главного лепестка РЛИ объекта можно оценить по изменению параметров

обработки сигнала на интервале синтезирования.

Для неподвижного объекта радиальная составляющая скорости взаимного перемещения носителя и объекта определяется в соответствии с выражением (1):

$$V_{лоц,r,mi} = V_{нос} \sin \alpha_{mi},$$

где угол α_{mi} на интервале синтезирования меняется, учитывая его малость, в пределах $\pm D_{синт}/2R_{об}$. На эти скорости настроена согласованная обработка для неподвижных объектов. Для центральной точки интервала синтезирования, как уже отмечалось, и скорость $V_{лоц,r,mi} = 0$, и угол визирования $\alpha_{mi} = 0$.

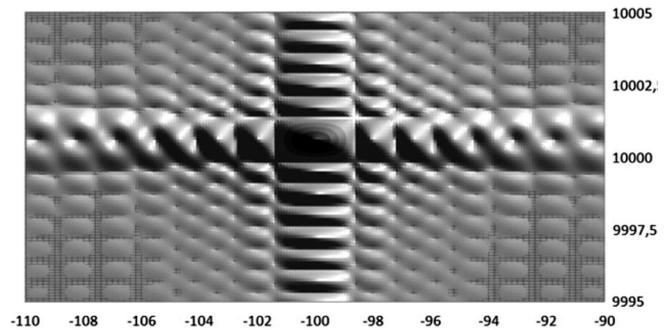


Рис. 4. РЛИ (в координатах X, Y) отражателя при $x = 10000$ м, $y = 0$ и $V_{об,x} = -1$ м/с

При лоцировании движущегося объекта равенство нулю $V_{лоц,r,mi}$ реализуется при некотором смещенном угле визирования $\alpha_{см}$. В соответствии с (3) с учетом, что $\alpha_{см} \ll 1$, $\alpha_{см} \approx V_{об}/V_{нос}$, (13) и это есть угол визирования объекта из центра интервала синтезирования.

Для одной из крайних точек интервала синтезирования при смещении угла визирования на некоторый угол $\Delta\alpha$ должно выполняться:

$$V_{нос} \sin(\alpha_{см} + \Delta\alpha) - V_{об,x} \cos(\alpha_{см} + \Delta\alpha) = V_{нос} \sin(D_{синт}/2 R_{об}), \quad (14)$$

или приближенно

$$V_{нос} (\alpha_{см} + \Delta\alpha) - V_{об,x} \approx V_{нос} D_{синт}/2 R_{об}. \quad (15)$$

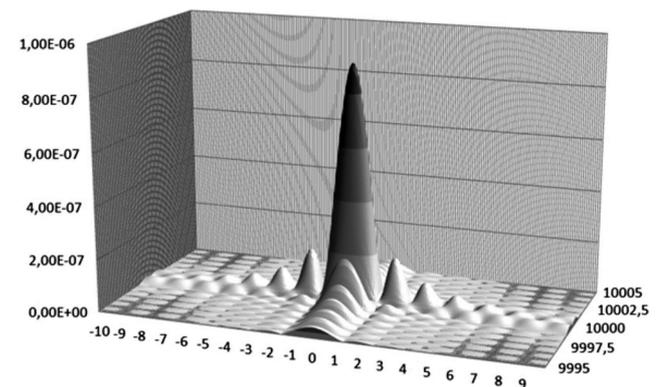


Рис. 5. РЛИ неподвижного объекта с координатами $x = 10000$ м, $y = 0$

С учетом (13) получаем, что $\Delta\alpha \approx D_{синт}/2R_{об}$. Аналогично для другой крайней точки. А это означает, что интервал синтезирования при переходе к лоцированию движущегося вдоль оси X объекта несколько сместился, но величина его не изменилась. Следовательно, ГЛ РЛИ в азимутальном сечении (вдоль оси Y) для движущегося вдоль оси X объекта также практически не изменился.

Это подтверждают результаты моделирования. На рис. 5 приведено РЛИ для неподвижного объекта при тех же исходных данных, что и для рис. 2, полученного для движущегося объекта. По форме ГЛ на обоих рисунках хорошо совпадают.

Влияние тангенциальной составляющей скорости лоцируемого объекта на его РЛИ

Для оценки влияния на РЛИ объекта тангенциальной составляющей его скорости $V_{об_y}$ (см. рис.6) следует рассмотреть это влияние в трех точках интервала синтезирования этого РЛИ: в центральной и в двух крайних. Очевидно, что в центральной точке носитель находится на траверзе объекта, и в этой точке радиальные составляющие скоростей объекта и носителя равны нулю и не зависят от тангенциальной составляющей скорости объекта. При зондировании объекта в этой точке производится согласованная обработка входного сигнала, не зависящая от значения $V_{об_y}$.

Влияние $V_{об_y}$ при зондировании в двух крайних точках интервала синтезирования можно оценить, исходя из рис. 6 и 7.

На рис. 6 приведено расположение лоцируемого объекта (точка A , $V_{об_y} < 0$) и носителя (точка 1, $V_{нос} > 0$) при некотором i -ом зондировании, соответствующем первому зондированию объекта на интервале синтезирования его РЛИ, при котором суммарная радиальная скорость сближения носителя и объекта в точке A , определяемая как

$$V_{лоц_r_дв_обі} = V_{нос} \sin \alpha_{об_i} - V_{об_y} \sin \alpha_{об_i} \quad (16)$$

сравнилась с максимальной расчетной скоростью сближения

$$V_{лоц_r_макс} = V_{нос} \sin \alpha_{синт} \quad (17)$$

из диапазона скоростей, используемых при согласованной обработке для неподвижного объекта в некоторой точке B , для которой $\alpha_{k_i} = \alpha_{синт}$ при том, что весь диапазон изменения α_{k_i} определяется двумя крайними значениями

$$\pm \alpha_{синт} \approx \pm \arctg \frac{D_{синт}}{2R_B}.$$

В рассматриваемом случае при отрицательной скорости объекта $\alpha_{синт} \geq \alpha_{k_i} \geq \alpha_{об_i}$. При этом при первом и последующих зондированиях существует точка, в которой выполняется равенство

$$V_{нос} \sin \alpha_{k_i} = V_{нос} \sin \alpha_{об_i} - V_{об_y} \sin \alpha_{об_i} \quad (18)$$

или

$$\sin \alpha_{об_i} = \frac{V_{нос}}{V_{нос} - V_{об_y}} \sin \alpha_{k_i}. \quad (19)$$

Таким образом, при движении носителя отклик согласованной обработки при первом зондировании объекта появится в точке B , в которой выполняется равенство (19) при $\alpha_{k_i} = \alpha_{синт}$. В этом случае первый параметр обработки – расчетная радиальная скорость в точке B – соответствует реальной радиальной скорости при нахождении объекту в точке A . Вторым параметром – расстояние $R_{лоц_обі}$ – определяется выражением

$$R_{лоц_обі} = D_{синт} / 2 \sin \alpha_{k_i}. \quad (20)$$

Это означает, что на этом такте зондирования сигнал, отраженный от движущегося объекта в точке A ,

пройдет согласованную обработку при формировании РЛИ в некоторой точке B , расположенной на пересечении линии визирования $\alpha_{k_i} = \alpha_{синт}$ и окружности с радиусом $R_{лоц_обі}$ с центром на оси Y в точке 1.

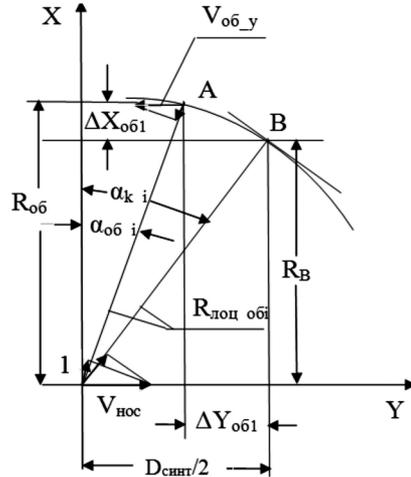


Рис. 6. Взаимное расположение носителя, объекта и его РЛИ при первом зондировании на интервале синтезирования антенны и формирования РЛИ этого объекта

При этом смещение этого сигнала по оси Y в соответствии с рис. 6 составит:

$$\Delta Y_{об1} = D_{синт}/2 - R_{лоц_обі} \sin \alpha_{об_i} \quad (21)$$

или, с учетом (20) и (19),

$$\Delta Y_{об1} = \frac{D_{синт}}{2} \left[1 - \frac{V_{нос}}{V_{нос} - V_{об_y}} \right] = - \frac{D_{синт}}{2} \frac{V_{об_y}}{V_{нос} - V_{об_y}}. \quad (22)$$

Вместе с тем, как уже отмечалось ранее, при прохождении носителем траверза лоцируемого объекта (точки 2) радиальные составляющие скоростей носителя и объекта равны нулю. Поэтому сигнал даже после обработки, согласованной с неподвижным объектом, не смещается и находится в точке, определяемой положением объекта (и носителя) по оси Y на этот момент.

Это положение определяется взаимным перемещением объекта и носителя относительно точки 1 до траверза, т.е., до их совмещения по оси Y . В соответствии с рис. 6 и выражением (22) расстояние по оси Y между объектом и носителем в точке 1 составляет:

$$\Delta Y_{нос_об1} = \frac{D_{синт}}{2} - \Delta Y_{об1} = \frac{D_{синт}}{2} \left(1 + \frac{V_{об_y}}{V_{нос} - V_{об_y}} \right) = \frac{D_{синт}}{2} \frac{V_{нос}}{V_{нос} - V_{об_y}}. \quad (23)$$

Соответственно, к моменту достижения точки траверза (точки 2) объект (и его изображение) за счет своего движения сместится относительно точки 1 на величину:

$$\Delta Y_{об1_2} = \Delta Y_{нос_об1} \frac{V_{об_y}}{V_{нос} - V_{об_y}} = \frac{D_{синт}}{2} \frac{V_{нос} V_{об_y}}{(V_{нос} - V_{об_y})^2}. \quad (24)$$

При $V_{нос} \gg V_{об_y}$

$$\Delta Y_{об1_2} \approx \frac{D_{синт}}{2} \frac{V_{об_y}}{V_{нос} - 2V_{об_y}} \approx \frac{D_{синт}}{2} \frac{V_{об_y}}{V_{нос}}, \quad (25)$$

что примерно соответствует (22), но противоположно по знаку.

На такую же величину объект сместится при переходе из точки 2 в точку 3, соответствующую последнему зондированию рассматриваемого объекта на интервале синтезирования его РЛИ. На рис. 7 приведено расположение лоцируемого объекта и носителя при *i*-ом зондировании, когда носитель расположен в точке 3, а объект – в точке *A*. При этом необходимо учесть, что точка *A* по сравнению с точкой *A* на рис. 6 за счет движения самого объекта сместилась на величину

$$\Delta Y_{об1_3} \approx D_{снтт} \frac{V_{об_y}}{V_{нос} - 2V_{об_y}}. \quad (26)$$

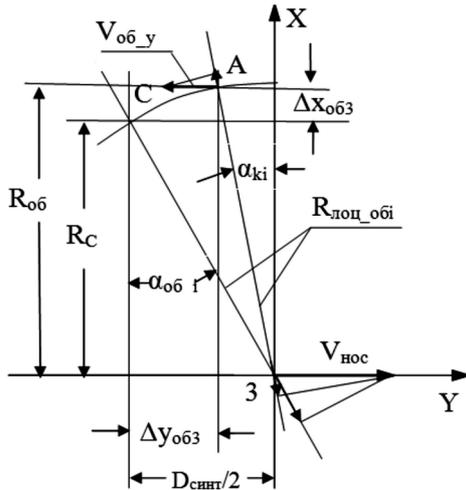


Рис. 7. Взаимное расположение носителя, объекта и его РЛИ при последнем зондировании на интервале синтезирования

Вместе с тем, как и в случае, приведенном на рис. 6, согласованная обработка входного сигнала, отраженного от движущегося объекта в точке *A*, произойдет в точке *C*, в которой выполняется условие (18). И по аналогии с (21), (22), можно показать, что

$$\Delta Y_{об3} = \frac{D_{снтт}}{2} \frac{V_{об_y}}{V_{нос} - V_{об_y}}. \quad (27)$$

Это означает, что суммарное результирующее смещение максимума сигнала объекта за интервал синтезирования его РЛИ, обусловленное движением самого объекта и смещением максимума сигнала при его обработке (из-за не учета движения объекта), составляет

$$\Delta Y_{рез} \approx 2D_{снтт} \frac{V_{об_y}}{V_{нос}}, \quad (28)$$

и это определяет расширение ГЛ изображения объекта по азимуту. При этом, так как нарушается когерентность сигналов, принятых за интервал синтезирования, то ГЛ оказывается модулированным по амплитуде и может иметь несколько максимумов. По этой же причине нельзя однозначно определить ширину ГЛ. В качестве оценки можно использовать оценку по уровню 0,5 от максимуму амплитуды ГЛ. Изложенное иллюстрируется рис. 8, на котором приведены сечения РЛИ по оси *Y* и который соответствует лоцированию движущегося вдоль оси *Y* объекта при $V_{об_y} = 1$ м/с. Остальные параметры моделирования соответствуют рис. 2. Рис. 8 подтверждает

полученное ранее выражение (28) для определения расширения ГЛ при не учете тангенциальной скорости объекта, в соответствии с которым для рассматриваемого случая $\Delta Y_{рез} = 20$ м.

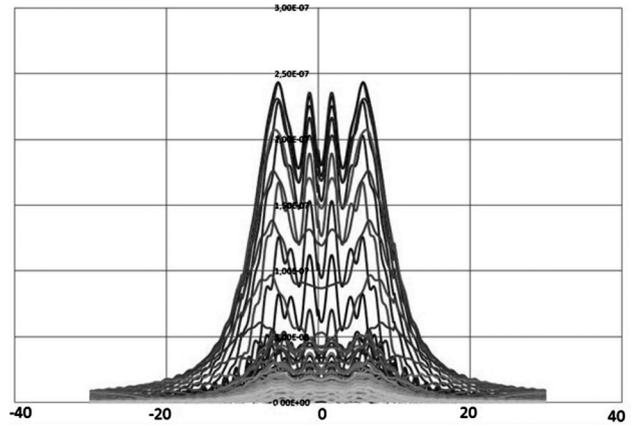


Рис. 8. Сечения РЛИ по оси *Y*, соответствующие лоцированию движущегося вдоль оси *Y* объекта при $V_{об_y} = 1$ м/с, $V_{нос_y} = 100$ м/с и $D_{снтт} = 1000$ м

На рис. 9 приведены результаты моделирования РСА в виде сечений РЛИ в азимутальной плоскости при наличии единственного отражателя с координатами на начало моделирования $x_{об0} = 10000$ м (т.е. $R_{об} = 10$ км), $y_{об0} = 0$ при $V_{нос_y} = 100$ м/с, $V_{об_y} = 2$ м/с и $D_{снтт} = 400$ м. Координаты носителя на начало моделирования составляли $x_{нос0} = 0$, $y_{нос0} = -210$ м. При этих условиях в соответствии с выражением (28) расширение ГЛ равно 16 м, что соответствует ширине ГЛ примерно по уровню 0,5 на рис. 9.

При этом, как уже отмечалось ранее, положение центра ГЛ РЛИ объекта по азимуту соответствует положению объекта на момент прохождения носителем траверза объекта, а это есть момент $t_{тр}$, когда $y_{об} = y_{нос}$. Исходя из уравнений движения объекта и носителя по оси *Y*, нетрудно получить, что относительно начала моделирования $t_{тр}$ определяется выражением:

$$t_{тр} = (y_{об0} - y_{нос0}) / (V_{об_y} - V_{нос_y}),$$

и для рассматриваемого случая $t_{тр} = 2,14$ с, а положение объекта (и носителя) на этот момент по оси *Y* равно 4,28 м, что хорошо соответствует рис. 9.

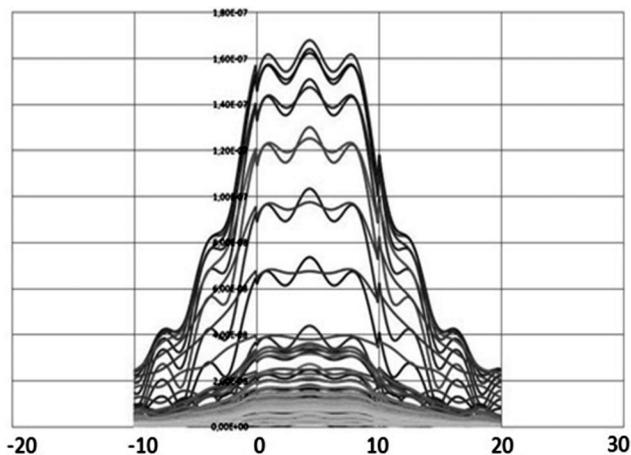
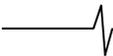


Рис. 9. Сечения в азимутальной плоскости РЛИ отражателя с координатами на начало моделирования $x_{об0} = 10000$ м, $y_{об0} = 0$ при $V_{об_y} = 2$ м/с, $V_{нос_y} = 100$ м/с и $D_{снтт} = 400$ м



Непосредственно из рис. 6 и 7 следует, что движение объекта с тангенциальной скоростью, неучтенной при обработке, не только расширяет ГЛ отклика (РЛИ объекта) по азимуту, но и на этапе синтезирования меняется дальность до него. Вследствие этого при отрицательной скорости объекта отклик изгибается вниз, как показано на рис. 10, или вверх при положительной скорости.

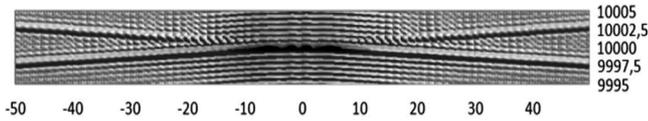


Рис. 10. РЛИ (в координатах X, Y) отражателя при $x = 10000$ м, $y = 0$ и $V_{об,y} = -1$ м/с

Порядок величины изменения дальности можно оценить на примере рис. 6 (первая точка лоцирования объекта на интервале формирования его РЛИ). Значение этого изменения равно

$$\Delta X_{об1} = R_{об} - R_B. \quad (29)$$

Исходя из рис. 6, а также полагая, что углы $\alpha_{об,i}$ и $\alpha_{k,i}$ значительно меньше 1, можно записать: $\Delta X_{об1} = R_{лоц,об1} \times$

$\times (\cos \alpha_{об,i} - \cos \alpha_{k,i}) \approx R_{лоц,об1} (\alpha_{k,i}^2 - \alpha_{об,i}^2)$ Для 1-й точки $\alpha_{k,i} \approx D_{синт}/2 R_{лоц,об1}$, а, исходя из выражения (19), следует:

$$\Delta X_{об1} \approx \frac{D_{синт}^2 V_{об,y}}{2 R_{об} V_{нос,y}}. \quad (30)$$

Для примера, при $V_{нос,y} = 100$ м/с, $V_{об,y} = 1$ м/с, $R_{об} = 10000$ м, $D_{синт} = 1000$ м расчетное смещение в 1-й точке $\Delta X_{об1} = 0,5$ м.

Заключение

В радиолокаторах бокового обзора, предназначенных для мониторинга земной и водной поверхности и осуществляющих обработку, согласованную с неподвижной местностью, лоцирование движущихся объектов в общем случае приводит к искажению и смещению РЛИ этих объектов. Наличие радиальной составляющей скорости лоцируемого объекта практически не искажает форму РЛИ этого объектов, но приводит к его смещению

на местности. Это смещение в продольном и поперечном относительно траектории носителя РСА направлениях пропорционально дальности до лоцируемого объекта. Одновременно в продольном направлении смещение пропорционально отношению радиальной скорости объекта к скорости носителя и может достигать существенных значений, которые необходимо учитывать как при реальных работах, так и при моделировании. Поперечные смещения незначительны.

При наличии у скорости объекта тангенциальной составляющей основные изменения его РЛИ связаны со значительным расширением этого РЛИ в продольном (азимутальном) направлении. Расширение пропорционально интервалу синтезирования и отношению продольной скорости объекта и скорости носителя. При этом центр РЛИ объекта привязан по своему положению к положению объекта при прохождении носителем его траверза (с учетом тангенциальной составляющей скорости объекта). При указанном расширении РЛИ в азимутальном направлении появляется некоторое смещение краев этого РЛИ по дальности. Поэтому РЛИ объекта имеет вид дужки.

Полученные соотношения позволяют оценивать смещения и расширения РЛИ движущихся объектов.

Литература

1. Кондратенков Г.С., Фролов А.Ю. Радиовидение. – М.: Радиотехника. 2005.
2. Брызгалов А.П., Караулова Е.В., Хныкин А.В. Аналого-цифровая обработка информации в радиолокаторах с синтезированной апертурой, использующих сверхширокополосные сигналы с линейной частотной модуляцией. // Цифровая обработка сигналов. №4, 2004 г.
3. Брызгалов А.П., Ковальчук И.В., Хныкин А. В., Шелва И. А., Юсупов Р. Г. Моделирование радиолокатора с синтезированной апертурой при решении задач его внутреннего и внешнего проектирования. Электронный журнал «Труды МАИ», Московский авиационный институт (национальный исследовательский университет), №43, март 2011.

УДК 621.391

МНОГОСКОРОСТНАЯ ОБРАБОТКА СИГНАЛОВ В ЗАДАЧАХ РЕЖЕКЦИИ УЗКОПОЛОСНЫХ ПОМЕХ

*Витязев В.В., д.т.н., профессор, зав. кафедрой Рязанского государственного радиотехнического университета,
e-mail: vityazev.v.v@rsreu.ru;*

Никишкин П.Б., аспирант, м.н.с. Рязанского государственного радиотехнического университета.

MULTIRATE SIGNAL PROCESSING FOR NARROWBAND NOISE REJECTION

Vityazev V.V., Nikishkin P.B.

Narrowband filters based methods for high selectivity frequency selection systems construction are considered. These methods are implemented in two-step structure using multirate signal processing. Modeling of the proposed approach and research of its efficiency are conducted.

Key word: multirate, processing, rejection, noise, narrowband, frequency selection, modeling.

Ключевые слова: многоскоростная, обработка, режекция, помеха, узкополосная, частотная селекция, моделирование.

Цифровые режекторные фильтры узкополосных помех: на основе децимации и интерполяции импульсной характеристики
Постановка задачи

Одной из классических задач цифровой частотной селекции сигналов является режекция (подавление) узкополосных помех и, в частности, низкочастотной составляющей. Если при этом ширина полосы режекции составляет незначительную часть (менее одной десятой) от общей рабочей полосы частот, которая определяет частоту дискретизации цифрового сигнала, то порядок N проектируемого КИХ-фильтра будет расти пропорционально отношению частоты дискретизации к ширине полосы режекции и составит сотни, тысячи и десятки тысяч единиц.

С целью реализации потенциальной возможности многократного уменьшения приведенных вычислительных затрат при построении узкополосного режекторного фильтра в работе [1] за основу был принят метод компенсации помехи по структуре на рис. 1,а, и предложена двухступенчатая реализация, представленная на рис. 1,б, отличающаяся от известных использованием эффекта децимации и интерполяции импульсной характеристики при реализации узкополосного фильтра с передаточной функцией $H(z)$ в цепи компенсации помехи.

Предполагается, что из входного сигнала $x(nT)$, задержанного на половину длины импульсной характеристики узкополосного фильтра, вычитается сигнал на его выходе $w_2(nT)$, что и компенсирует в значительной степени влияние узкополосной помехи в заданной полосе частот относительно центральной частоты ω_0 . Заметим, что степень режекции помехи в данной схеме определяется точностью аппроксимации желаемой частотной характеристики узкополосного фильтра в поло-

Рассматриваются методы построения системы частотной селекции сигналов с высокой избирательностью на основе узкополосного фильтра в цепи компенсации помех, реализуемого по двухступенчатой структуре с использованием многоскоростной обработки сигналов. Проводится моделирование и исследование эффективности данного подхода с позиции достижимости заданных характеристик и требуемых вычислительных затрат.

се пропускания. Это накладывает особые требования на синтез передаточной функции узкополосного фильтра в полосе пропускания, одновременно ослабляя требования к точности аппроксимации в полосе непропускания.

Возможны два подхода к эффективной реализации узкополосного фильтра с передаточной функцией $H(z) = H_1(z)H_2(z)$ и произвольной центральной частотой ω_0 полосы пропускания. Первый предполагает использование структуры с квадратурной модуляцией, позволяющей свести задачу построения цифрового полосового фильтра к задаче построения НЧ фильтра с заданными свойствами частотной избирательности. Второй исходит из предпосылки, что как гребенчатый (с передаточной функцией $H_1(z)$), так и сглаживающий (с передаточной функцией $H_2(z)$) фильтры реализуются с использованием комплексных импульсных характеристик и соответственно произвольным положением центральной частоты полосы пропускания [2].

Необходимость использования квадратурной модуляции или комплексных импульсных характеристик как минимум удваивает вычислительные затраты. Вместе с тем, в частном случае, когда значение центральной частоты ω_0 и коэффициент децимации ν импульсной характеристики гребенчатого фильтра отвечают условию $\omega_0 = 2\pi k / \nu$, где k – целое число, импульсные характеристики гребенчатого и сглаживающего фильтров принимают действительные значения, а следовательно, вычислительные затраты становятся соизмеримыми с затратами на режекторный фильтр с нулевой центральной частотой полосы режекции.

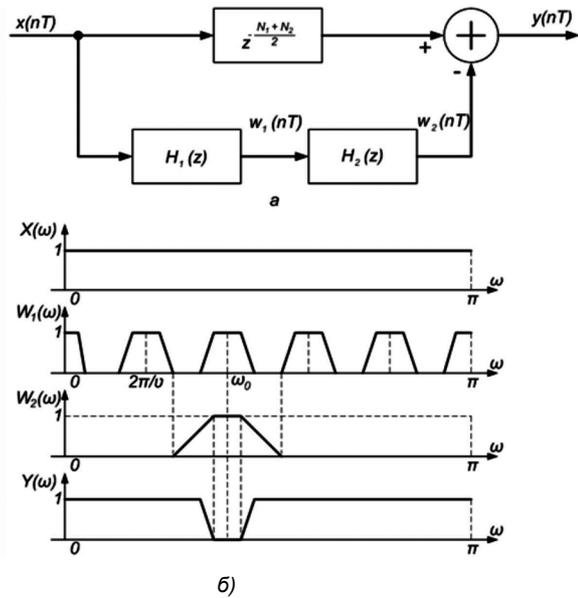
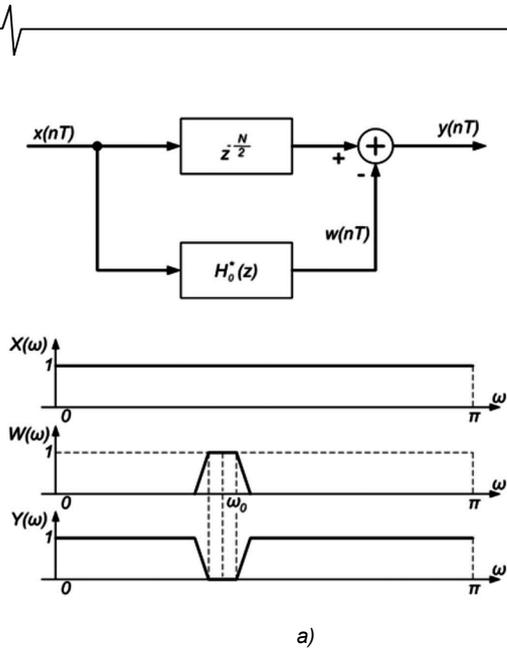


Рис. 1. Структурные схемы узкополосного режекторного фильтра: а – без децимации; б – с децимацией и интерполяцией импульсной характеристики

Общая вычислительная эффективность рассматриваемой структуры узкополосного режекторного фильтра полностью определяется степенью уменьшения вычислительных затрат на реализацию узкополосного фильтра в цепи компенсации помехи.

Решение задачи оптимизации

Пусть показатели частотной избирательности проектируемого фильтра с желаемой АЧХ $H(\omega)$: коэффициент прямоугольности АЧХ α , показатель узкополосности β , логарифмический показатель частотной избирательности $L(\varepsilon_{1\text{дон}}, \varepsilon_{2\text{дон}})$ принимают следующие значения [3]:

$$\alpha = \omega_{c1} / (\omega_{c2} - \omega_{c1}); \quad \beta = 2\pi / \omega_{c1};$$

$$L(\varepsilon_{1\text{дон}}, \varepsilon_{2\text{дон}}) = -\frac{2}{3} \lg(10 \cdot \varepsilon_{1\text{дон}} \cdot \varepsilon_{2\text{дон}}).$$

При этом порядок КИХ-фильтра, реализуемого по обычной трансверсальной структуре, $N = \alpha\beta L(\varepsilon_{1\text{дон}}, \varepsilon_{2\text{дон}})$, однозначно определяет требуемый объем вычислительных затрат в единицу времени $R_T = Nf_{кв}$.

Для рассматриваемой двухкаскадной структуры приведенные вычислительные затраты

$$R_T = \left[\frac{\alpha}{v} + \frac{v}{\beta - \gamma v} \right] \beta L \left(\frac{\varepsilon_1}{2}, \varepsilon_2 \right) f_{кв}. \quad (1)$$

определяются как показателями частотной избирательности режекторного фильтра, так и принятым значением коэффициента прореживания импульсной характеристики гребенчатого фильтра v !

Взяв производную правой части выражения (1) по v и решив уравнение $\partial R_T / \partial v = 0$, найдем оптимальное значение коэффициента прореживания:

$$v_{\text{opt}} = (\beta \sqrt{\alpha\beta} - \alpha\beta\gamma) / (\beta - \alpha\gamma^2), \quad (2)$$

где $\gamma = (1 + 2\alpha) / \alpha$.

Минимальный объем вычислительных затрат на реализацию двухкаскадной структуры узкополосного фильтра

при оптимальном значении параметра v (2) составит

$$R_T(v_{\text{opt}}) = \min_v R_T = \frac{\alpha\alpha(a - b\gamma) - b^2\beta}{b(a - b\gamma)} L \left(\frac{\varepsilon_1}{2}, \varepsilon_2 \right) f_{кв}, \quad (3)$$

где $a = \beta - \gamma\alpha$, $b = \sqrt{\alpha\beta} - \gamma\alpha$.

Для узкополосных фильтров $\beta \gg \gamma\alpha$, и оценка вычислительных затрат (3) принимает приближенное значение

$$R_T^* = 2\sqrt{\alpha\beta} L(\varepsilon_1/2, \varepsilon_2) f_{кв}.$$

Вывод: при оптимальном значении коэффициента прореживания v_{opt} приведенные вычислительные затраты уменьшаются пропорционально корню квадратному от порядка фильтра N .

Моделирование

На рис. 2 представлены результаты моделирования в среде MATLAB+Simulink узкополосного режекторного фильтра без децимации и интерполяции, а на рис. 3 – с децимацией и интерполяцией импульсной характеристики.

Предполагалось, что проектируемый узкополосный режекторный фильтр, работающий на частоте дискретизации 3000 Гц, должен обладать следующими свойствами частотной избирательности относительно центральной частоты полосы режекции, равной 1 кГц: коэффициент прямоугольности АЧХ $\alpha = 2,5$; показатель узкополосности $\beta = 150$; неравномерность АЧХ в полосе пропускания $\varepsilon_{1\text{дон}} = 10^{-3}$; уровень затухания в зоне непрозрачности $\varepsilon_{2\text{дон}} = 10^{-4}$.

Подставив заданные параметры частотной избирательности в (2) и (3), получим: $v_{\text{opt}} = 14$ и $R_T^*(v_{\text{opt}}) = 778469$ операций умножения / секунду. Таким образом, по отношению к однокаскадной реализации узкополосного режекторного КИХ-фильтра, требующей $N = \alpha\beta L \times (\varepsilon_{1\text{дон}}, \varepsilon_{2\text{дон}}) = 7537500$ операций умножения / секунду, приведенные вычислительные затраты уменьшаются на порядок.

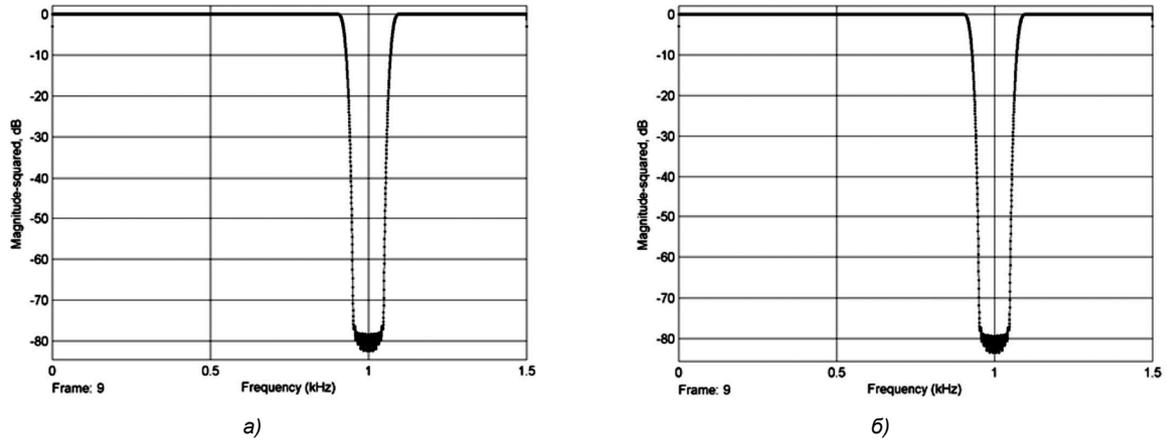


Рис. 2. Результаты моделирования: а – АЧХ режекторного фильтра, полученная с помощью программы FDATAoo; б – АЧХ режекторного фильтра без децимации и интерполяции, полученная моделированием структуры, представленной на рис. 1,а

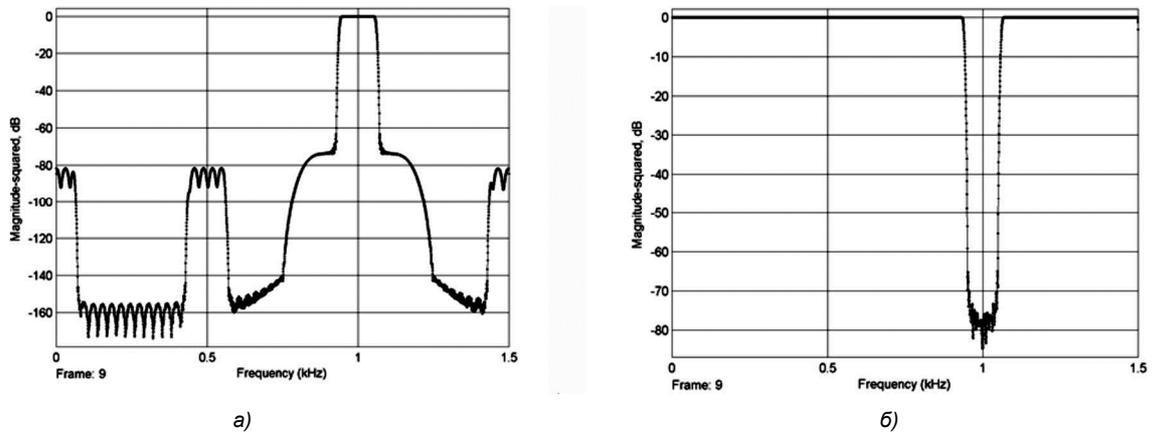


Рис. 3. Результаты моделирования: а – АЧХ узкополосного фильтра; б – АЧХ режекторного фильтра с децимацией и интерполяцией импульсной характеристики

Альтернативным способом эффективной реализации узкополосного режекторного фильтра является использование децимации и интерполяции преобразуемого сигнала в цепи компенсации [3, 4].

Цифровые режекторные фильтры узкополосных помех: на основе децимации и интерполяции преобразуемого сигнала в цепи компенсации
Постановка задачи

На рис. 4 представлена структурная схема режекторного фильтра, реализуемого на основе децимации и интерполяции преобразуемого сигнала в цепи компенсации. Предполагается, как и ранее, что из входного сигнала $x(nT)$, задержанного на половину длины импульсной характеристики узкополосного фильтра, вычитается сигнал на его выходе $w(nT)$, что и компенсирует в значительной степени влияние узкополосной помехи в заданной полосе частот относительно центральной частоты ω_0 . Вычислительная эффективность рассматриваемой структуры режекторного фильтра полностью определяется степенью уменьшения вычислительных затрат на реализацию узкополосного фильтра в цепи компенсации помехи, что достигается оптимизацией его параметров.

Для рассматриваемой структуры узкополосного фильтра приведенные вычислительные затраты

$$R_T = 2\alpha\beta L \left(\frac{\varepsilon_1}{2}, \varepsilon_2 \right) \frac{1}{V} f_{KB} \tag{4}$$

определяются как показатели частотной избирательности режекторного фильтра, так и принятым значением коэффициента прореживания ν сигнала $w^*(n\nu T)$ на выходе фильтра-дециматора с передаточной функцией $H_0^*(z)$.

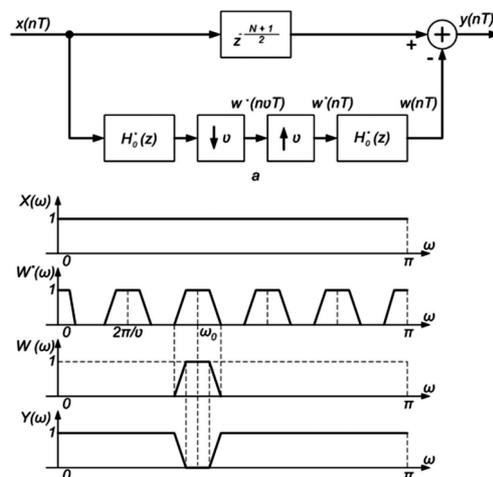


Рис. 4. Структурная схема режекторного фильтра, реализуемого на основе децимации и интерполяции преобразуемого сигнала в цепи компенсации

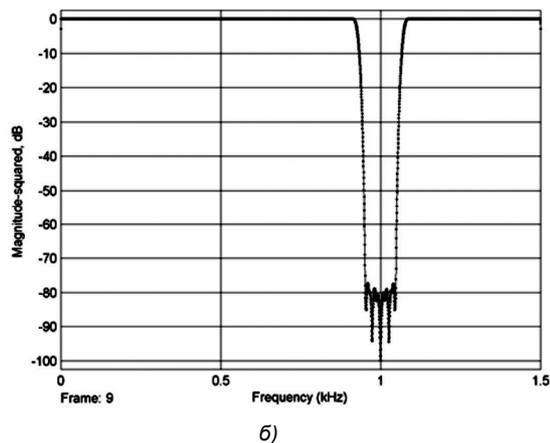
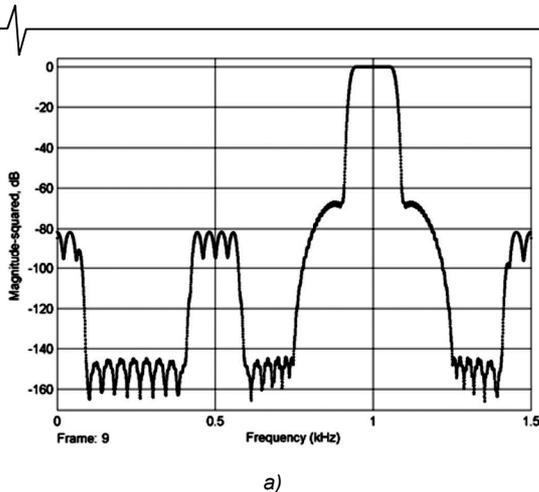


Рис. 5. Результаты моделирования: а – АЧХ узкополосного фильтра;

б – АЧХ режекторного фильтра с децимацией и интерполяцией преобразуемого сигнала в цепи компенсации

Выражение (4) достигает своего минимума при максимально допустимом значении коэффициента децимации [3, 4]

$$v_{\max} = \alpha\beta / (2\alpha + 1).$$

Моделирование

На рис. 5 представлены результаты моделирования узкополосного режекторного фильтра с децимацией и интерполяцией преобразуемого сигнала в цепи компенсации для рассматриваемого выше примера.

Сравнительный анализ характеристик, представленных на рис. 2, 3 и 5, говорит о достижимости заданных свойств частотной избирательности узкополосного режекторного фильтра (рис. 2), но при существенно меньших вычислительных затратах. Для рассматриваемого примера $v_{\max} = \alpha\beta / (2\alpha + 1) = 62$ и $R_{r\min} = 243145$ операций умножения/секунду, что в 31 раз меньше по отношению к прямой форме реализации режекторного фильтра и в 3 раза меньше по отношению к методу реализации на основе децимации и интерполяции импульсной характеристики узкополосного фильтра.

При переходе к многоступенчатой многоскоростной реализации узкополосного фильтра в цепи компенсации возможно дальнейшее уменьшение вычислительных затрат, но при этом надо учитывать «паразитное» влияние эффекта наложения (элайзинга) при уменьшении частоты дискретизации и задавать более жесткие требования к уровню затухания в зоне его непрозрачности (полосе пропускания режекторного фильтра) [4].

Цифровой полосовой фильтр с высокой прямоугольностью АЧХ

Постановка задачи

Рассматривается задача синтеза структуры НЧ фильтра с высокой прямоугольностью АЧХ ($\alpha \gg 1$) и относительно небольшим значением показателя узкополосности ($\beta < \alpha$). Предполагается, что задача построения полосового фильтра может быть сведена к задаче построения НЧ фильтра с использованием структуры с квадратурной модуляцией [3, 4]. Относительно малое значение показателя узкополосности β не позволяет эффективно использовать методы децимации и интерполяции преобразуемого сигнала или импульсной характеристики фильтра для существенно-

го уменьшения приведенных вычислительных затрат.

В работе [1] впервые был предложен метод построения структуры НЧ фильтра с высокой прямоугольностью АЧХ с применением узкополосного режекторного фильтра с передаточной функцией $H_1(z)$, формирующего переходную зону проектируемого широкополосного фильтра по представленной выше (рис. 1,б) структуре, и сглаживающего фильтра с передаточной функцией $H_2(z)$, подавляющего все высокочастотные составляющие за полосой режекции входного фильтра (рис. 6).

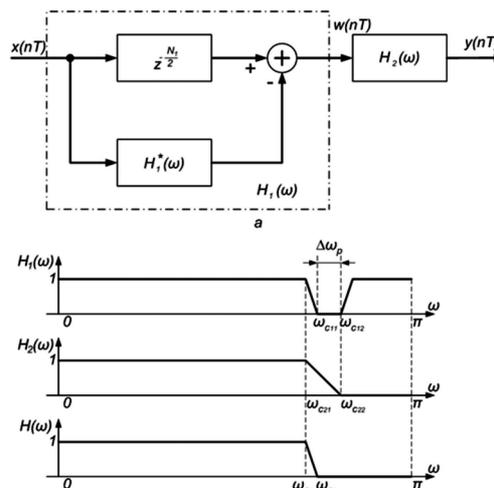


Рис. 6. Структурная схема НЧ фильтра с высокой прямоугольностью АЧХ

Порядок N_1 узкополосного режекторного фильтра определяется порядком N проектируемого широкополосного фильтра, так как оба фильтра имеют одну и ту же относительную ширину переходной зоны АЧХ, но вычислительные затраты на его реализацию могут уменьшаться пропорционально коэффициенту децимации v импульсной характеристики гребенчатого фильтра с передаточной функцией $H_{1r}(z)$, используемого в цепи компенсации с общей передаточной функцией $H_1^*(z)$.

Увеличение коэффициента децимации v возможно только путем уменьшения ширины полосы режекции и, как следствие, пропорциональным увеличением показателя узкополосности фильтра в цепи компенсации. С другой стороны, уменьшение ширины полосы режекции приводит к увеличению показателя прямоугольности АЧХ α_2

выходного сглаживающего фильтра с функцией передачи $H_2(z)$ и соответственно пропорциональному увеличению вычислительных затрат на его реализацию.

Решение задачи оптимизации

Поиск оптимального значения коэффициента децимации ν импульсной характеристики гребенчатого фильтра приводит к решению следующей задачи оптимизации:

$$R_T(\nu) = \left(\frac{2N_{1Г}}{\nu} + 2N_{1С} + N_2 \right) f_{кв} \rightarrow \min_{\nu} \tag{5}$$

где $N_{1Г} \cong N$ и $N_{1С}$ – порядки цифрового гребенчатого и сглаживающего фильтров в цепи компенсации; N_2 – порядок выходного сглаживающего фильтра. Увеличение затрат в 2 раза на реализацию цепи компенсации связано с использованием квадратурной модуляции и появлением дополнительного квадратурного канала.

Порядок гребенчатого фильтра

$$N_{1Г} = \alpha\beta L \left(\frac{\varepsilon_{1доп}}{2}, \varepsilon_{2доп} \right), \tag{6}$$

несколько превышает порядок N проектируемого широкополосного фильтра, что связано с повышением требований точности аппроксимации его частотной характеристики в зоне режекции.

Порядки $N_{1С}$ и N_2 зависят от заданной ширины полосы режекции $\Delta\omega_p$ и выбранного значения коэффициента децимации ν импульсной характеристики гребенчатого фильтра:

$$N_{1С} \cong \frac{2\pi\nu}{2\pi - \Delta\omega_p\nu} L \left(\frac{\varepsilon_{1доп}}{2}, \varepsilon_{2доп} \right);$$

$$N_2 \cong \frac{2\pi}{\Delta\omega_p} L \left(\frac{\varepsilon_{1доп}}{2}, \varepsilon_{2доп} \right). \tag{7}$$

Предполагается, что при $\alpha \gg 1$ ширина переходной зоны АЧХ режекторного фильтра многократно меньше полосы режекции $\Delta\omega_p$ и ее влиянием при оценке порядков $N_{1С}$ и N_2 сглаживающих фильтров можно пренебречь. Как и следовало ожидать, при расширении полосы режекции $\Delta\omega_p$ пропорционально уменьшается порядок N_2 выходного сглаживающего фильтра, но увеличивается порядок $N_{1С}$ сглаживающего фильтра в цепи компенсации. Заметим, что при этом для выбранного значения ширины полосы режекции $\Delta\omega_p$ на максимально допустимое значение коэффициента децимации ν накладывается ограничение:

$$\nu \leq \frac{\pi}{\Delta\omega_p} = \frac{f_{кв}}{2\Delta f_p}. \tag{8}$$

Подставив (6) и (7) в (5), получим

$$R_T(\nu, \Delta\omega_p) = \left(\frac{2\alpha\beta}{\nu} + \frac{4\pi\nu}{2\pi - \Delta\omega_p\nu} + \frac{2\pi}{\Delta\omega_p} \right) \times \tag{9}$$

$$\times L \left(\frac{\varepsilon_{1доп}}{2}, \varepsilon_{2доп} \right) f_{кв} \rightarrow \min_{\nu, \Delta\omega_p}.$$

Задача нелинейного программирования (9) решается путем простого перебора целочисленных значений коэффициента децимации ν на ограниченном множестве значений (8). Заметим, что при каждом фиксированном целочисленном значении коэффициента децимации ν можно найти оптимальное значение ширины полосы режекции $\Delta\omega_{p,опт}$, при котором достигается минимум целевой функции (9):

$$\left(\frac{4\pi\nu}{2\pi - \Delta\omega_p\nu} + \frac{2\pi}{\Delta\omega_p} \right) \rightarrow \min_{\nu, \Delta\omega_p},$$

если $\Delta\omega_{p,опт} = \frac{2(\sqrt{2}-1)\pi}{\nu}.$ (10)

В частном случае, когда центральная частота ω_0 полосы режекции кратна периоду $2\pi/\nu$, цепь компенсации не требует введения квадратурной модуляции, а следовательно вычислительные затраты на ее реализацию уменьшаются в 2 раза, и задача оптимизации (9) принимает вид:

$$R_T(\nu, \Delta\omega_p) = \left(\frac{\alpha\beta}{\nu} + \frac{2\pi\nu}{2\pi - \Delta\omega_p\nu} + \frac{2\pi}{\Delta\omega_p} \right) \times$$

$$\times L \left(\frac{\varepsilon_{1доп}}{2}, \varepsilon_{2доп} \right) f_{кв} \rightarrow \min_{\nu, \Delta\omega_p}.$$
 (11)

При этом для каждого фиксированного целочисленного значения коэффициента децимации ν на ограниченном множестве значений (8) оптимальное значение $\Delta\omega_{p,опт}$ ширины полосы режекции достигается при

$$\Delta\omega_{p,опт} = \frac{\pi}{\nu}, \tag{12}$$

т.е. фактически при максимально допустимом его значении при заданной ширине полосы режекции.

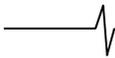
Заметим, что вычислительные затраты в соответствии с (11) могут быть существенно уменьшены только при условии многократного уменьшения затрат на реализацию гребенчатого фильтра (первое слагаемое), что достигается при условии, что коэффициент децимации $\nu \geq 4$.

Дополнительное условие – центральная частота ω_0 полосы режекции кратна периоду $2\pi/\nu$ при заданном ее значении в еще большей степени ограничивает множество допустимых значений ν и тем самым ускоряет поиск оптимального сочетания параметров ν и $\Delta\omega_{p,опт}$.

Как показали расчеты, ожидаемый выигрыш в минимизации вычислительных затрат по отношению к обычной некаскадной структуре лежит в диапазоне от 1,5 до 5 раз, в зависимости от показателя прямоугольности АЧХ проектируемого широкополосного фильтра.

Моделирование

С целью иллюстрации достижимости заданных частотных характеристик широкополосного НЧ фильтра с относительно высокой прямоугольностью АЧХ в рамках



рассматриваемой структуры было произведено моделирование в среде MATLAB+Simulink при использовании в качестве узкополосного режекторного фильтра его реализации на основе децимации и интерполяции импульсной характеристики (рис. 1,а). Результаты моделирования представлены на рис. 7.

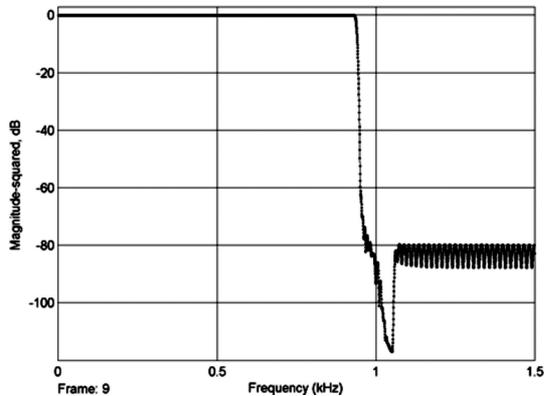


Рис. 7. Результат моделирования НЧ фильтра с высокой прямоугольностью АЧХ

Заключение

С целью реализации потенциальной возможности многократного уменьшения приведенных вычислительных затрат при построении узкополосного режекторного фильтра за основу был принят метод компенсации помехи по двухступенчатой реализации, отличающийся от известных использованием эффекта децимации и интерполяции импульсной характеристики при реализации узкополосного фильтра в цепи компенсации помехи. Показано, что при оптимальном значении коэффициента прореживания v_{opt} приведенные вычислительные затраты уменьшаются пропорционально корню квадратному от порядка фильтра N .

Альтернативным способом эффективной реализации узкополосного режекторного фильтра является использование децимации и интерполяции преобразуемого сигнала в цепи компенсации. Сравнительный анализ

представленных характеристик говорит о достижимости заданных свойств частотной избирательности узкополосного режекторного фильтра при существенно меньших вычислительных затратах. При переходе к многоступенчатой многоскоростной реализации узкополосного фильтра в цепи компенсации возможно дальнейшее уменьшение вычислительных затрат, но при этом надо учитывать «паразитное» влияние эффекта наложения (элайзинга) при уменьшении частоты дискретизации и задавать более жесткие требования к уровню затухания в зоне его непрозрачности (полосе пропускания режекторного фильтра).

Рассматривается задача синтеза структуры НЧ фильтра с высокой прямоугольностью АЧХ ($\alpha \gg 1$) и относительно небольшим значением показателя узкополосности ($\beta < \alpha$). Предложен метод построения структуры НЧ фильтра с высокой прямоугольностью АЧХ с применением узкополосного режекторного фильтра, формирующего переходную зону проектируемого широкополосного фильтра, и сглаживающего фильтра, подавляющего все высокочастотные составляющие за полосой режекции входного фильтра. Как показали расчеты, ожидаемый выигрыш в минимизации вычислительных затрат по отношению к обычной некаскадной структуре лежит в диапазоне от 1,5 до 5 раз, в зависимости от показателя прямоугольности АЧХ проектируемого широкополосного фильтра.

Литература

1. Витязев В.В., Горюнов Ю.Н. Оптимальное проектирование цифровых полосовых фильтров с высокой прямоугольностью // *Электросвязь*. – 1995. № 4. – С. 30-32.
2. Витязев В.В., Морозов Е.А. Оптимальное проектирование цифровых полосовых фильтров на процессорах обработки сигналов // *Электросвязь*. – 1995. № 12. – С. 29-31.
3. Витязев В.В. Цифровая частотная селекция сигналов. М.: Радио и связь, 1993. 240 с.
4. Витязев В.В. Многоскоростная обработка сигналов. – М.: Горячая линия – Телеком, 2017. – 336 с.

УДК 621.372

СИНТЕЗ ЦЕЛОЧИСЛЕННЫХ БИХ-ФИЛЬТРОВ МИНИМАЛЬНОЙ ДЛИНЫ СЛОВА КОЭФФИЦИЕНТОВ

*Бугров В.Н., к.т.н., доцент Нижегородского государственного университета им. Н.И. Лобачевского,
e-mail: bug@rf.unn.ru.*

SYNTHESIS OF THE INTEGER IIR FILTERS WITH SHORT COEFFICIENT WORD LENGTH

Bugrov V.N.

The integer simulation and development infinite impulse response (IIR) filters taking into account the possibilities of their realization on digital integer platforms are considered. The problem statement and solution of multifunctional synthesis of digital filters such a problem on the basis of the numerical methods of integer nonlinear mathematical programming are given. As an several examples, the problem solution of synthesis of high-pass IIR-filters with short coefficient word length has been given. The analysis of their characteristics is resulted.

Key words: the integer IIR-filter, integer nonlinear programming, multifunctional synthesis, criterion function.

Ключевые слова: целочисленный рекурсивный фильтр, целочисленное нелинейное программирование, многофункциональный синтез, целевая функция.

Введение

Наряду с селективной способностью совокупное качество цифрового фильтра определяется и его быстродействием, то есть минимальным временем расчёта отклика при реализации на заданной цифровой платформе. При этом под селективностью обычно понимают возможность удовлетворения совокупности требуемых характеристик фильтра в частотной области. К основным факторам, определяющих быстродействие, относятся, прежде всего, арифметика вычислений и длина слова коэффициентов (разрядность) цифрового БИХ-фильтра. Обзор публикаций по методам расчёта и синтеза рекурсивных фильтров показывает, что в настоящее время преобладает косвенное проектирование рекурсивных фильтров по аналоговому прототипу с применением метода билинейного преобразования [1 - 3] либо различных подходов, на нём базирующихся [4 - 6], при использовании в алгоритмах цифровой фильтрации вещественной арифметики вычислений обычно в формате с плавающей точкой. Вещественный формат представления данных вынуждает квантовать их значения, что приводит к весьма нежелательным для работы такого БИХ-фильтра последствиям, таким как искажение частотных характеристик, появление шумов квантования, необходимость масштабирования вещественных коэффициентов фильтра и возможность появления малых предельных циклов при квантовании результатов внутренних вычислений. К тому же проектное решение в вещественных числах может быть реализовано только на специализированных сигнальных процессорах. Вопросам квантования при аналитическом расчёте БИХ-фильтра посвящено много работ [5, 6]. В некоторых из

Рассматриваются вопросы моделирования и синтеза рекурсивных (IIR) цифровых фильтров с учетом возможностей их реализации на цифровых платформах с целочисленной арифметикой вычислений. Приводится постановка задачи многофункционального синтеза IIR-фильтров методами целочисленного нелинейного математического программирования. Иллюстрируется эффективность применения данного подхода на примерах решения задач синтеза рекурсивных целочисленных фильтров верхних частот с минимальной длиной слова коэффициентов. Проводится анализ характеристик синтезированных фильтров.

них рассматривалась также возможность ограничения радиусов полюсов передаточной функции фильтра. В целом можно отметить, что при высоком порядке фильтра ($N > 20$) весьма существенные трудности возникают уже при квантовании вещественных данных до 8 - 10 бит.

Однако возможность прямого синтеза цифровых БИХ-фильтров непосредственно в целочисленном пространстве состояний может быть предоставлена методологией целочисленного нелинейного программирования (ЦНП), общее описание которой приведено в работах [7 - 11]. Под пространством состояний в данном случае понимается, прежде всего, многомерное пространство целочисленных параметров (коэффициентов фильтра), входных и выходных сигналов – целочисленных временных последовательностей, а также базовые целочисленные операции над данными в алгоритме цифровой фильтрации. Идеология ЦНП позволяет эффективно проектировать целочисленные рекурсивные фильтры с заданной разрядностью представления данных и максимальном выполнении требований к совокупности частотных характеристик фильтра при произвольной форме их задания. В настоящее время в общей номенклатуре коммерческих цифровых платформ существенную долю занимают 8-разрядные цифровые платформы с целочисленной арифметикой вычислений. При реализации высокоскоростных целочисленных цифровых фильтров (ЦЦФ) на специализированных платфор-

мах или на кристалле разрядность представления данных может быть ещё ниже (до 4 или даже до 3 бит).

В данной статье рассматривается возможность решения задач синтеза каскадных рекурсивных ЦЦФ минимальной разрядности представления данных компьютерным пакетом ЦНП (версии 3.1), примеров, иллюстрирующих принципиальные возможности данного подхода к многофункциональному проектированию цифровых систем.

Моделирование и синтез рекурсивных ЦЦФ малой разрядности

Рекурсивные фильтры, являясь дискретными линейными системами с обратной связью, обладают значительно большими селективными возможностями по сравнению с КИХ-фильтрами и позволяют реализовать требуемые частотные характеристики значительно меньшей разрядностью представления данных и меньшим порядком N фильтра. Что касается структуры построения рекурсивных фильтров, то в [5, 12] приведено сравнение каскадных, параллельных, прямых и волновых структур построения, и показано, что каскадная структура является наилучшей. Поэтому в настоящее время построение рекурсивных ЦЦФ в форме каскадного соединения звеньев первого или второго порядка на практике используется наиболее часто.

Передаточная функция для рекурсивного ЦЦФ, состоящего из каскадного соединения m -звеньев второго порядка ($m = N/2$), имеет следующий вид [7, 8]:

$$H(z) = \prod_{i=1}^m \frac{b_{0i} + b_{1i}z^{-1} + b_{2i}z^{-2}}{a_{0i} + a_{1i}z^{-1} + a_{2i}z^{-2}}, \quad (1)$$

где комплексная переменная z при переходе к описанию комплексной частотной характеристики $K(e^{j\omega}) = |K(e^{j\omega})| \times e^{j\varphi(\omega)}$ принимает значение $z = e^{j\omega}$, а $\omega = 2\pi f / fs$ – приведенная круговая частота.

Все коэффициенты системной функции (1) являются целочисленными, а их интервал изменения (вариации) определяется заданной длиной битового слова (разрядностью) коэффициентов фильтра. Из соотношения (1) легко получается разностное уравнение для одного звена целочисленного фильтра:

$$y_n = (b_0x_n + b_1x_{n-1} + b_2x_{n-2} - a_1y_{n-1} - a_2y_{n-2}) / a_0, \quad (2)$$

где x_n, y_n – входная и выходная целочисленные временные последовательности, a_0 – масштабирующий множитель.

Как видно из (2), при вычислении отклика фильтра должна выполняться операция деления на целочисленный коэффициент a_0 , которая может быть реализована операцией сдвига при условии принадлежности каждого i -го коэффициента биномиальному целочисленному ряду:

$$a_{0i} \in \{2^q\}, \quad q = \overline{0, W_k} \quad i = \overline{1, m}, \quad (3)$$

где W_k – длина (разрядность) битового слова целочисленных коэффициентов фильтра.

На рис. 1 приведена типичная структура звеньев ре-

курсивного целочисленного фильтра, соответствующая разностному уравнению (2). Как видно, при вычислении отклика фильтра кроме традиционных операций сложения, умножения и задержки на такт присутствует операция сдвига на $B = \log_2 a_0$ бит, с помощью которой реализуется целочисленное деление на биномиальный коэффициент a_0 .

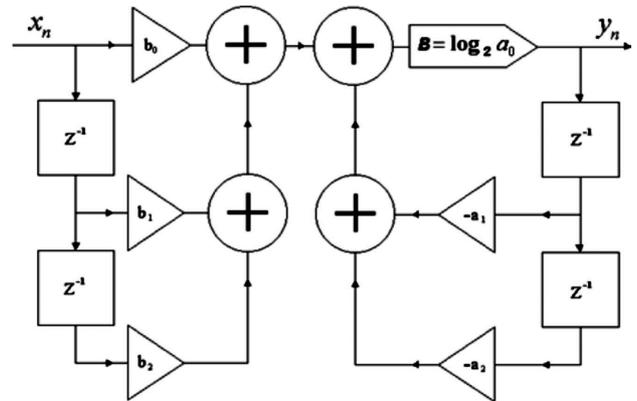


Рис. 1. Структура звена рекурсивного ЦЦФ

Можно отметить, что именно наличие итоговой операции сдвига в алгоритме расчёта отклика рекурсивного целочисленного звена нарушает коммутативность расчётной процедуры реализации нулей и полюсов его передаточной функции (1), что не позволяет, в свою очередь, реализовать целочисленное рекурсивное звено в канонической форме.

Как известно, рекурсивный фильтр будет устойчив, если все полюсы p_i передаточной функции лежат внутри единичного круга в z -плоскости:

$$|Zp_i| < 1. \quad (4)$$

Заметим, однако, что это стандартное условие является необходимым, но далеко не достаточным, так как устойчивость работы БИХ-фильтра, как системы с обратной связью, может быть нарушена появлением предельных циклов (limit cycle), когда вроде бы устойчивый по соотношению (4) фильтр начинает демонстрировать неустойчивое поведение, приводящее к возникновению на выходе периодических колебаний. При этом предельные циклы типа «granular» [3] возникают, когда при отсутствии сигнала на входе амплитуда выходного сигнала затухает, но из-за вычислительных погрешностей не доходит до нуля. На рис. 2, в качестве примера проявления таких циклов, представлены осциллограммы сигналов на выходе рекурсивных фильтров различного порядка при нулевом входе.

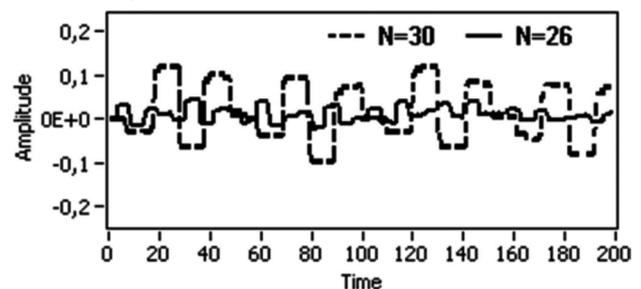


Рис. 2. Амплитуда предельных циклов при нулевом входе «Переполняющие» (overflow) предельные циклы имеют место, когда амплитуда выходного сигнала не затухает

хает, а возрастает, вызывая в итоге переполнение. Как известно, возможность возникновения предельных циклов возрастает с увеличением порядка рекурсивного фильтра и во многом определяется добротностью полюсов его передаточной функции, т.е. коэффициентами знаменателя системной функции (1). В настоящее время провести теоретический анализ, расчёт предельных циклов весьма непросто даже для БИХ-фильтров малого порядка. Поэтому действенной, а порой и единственной альтернативой является практическая реализация синтезированного ЦЦФ на конкретной цифровой платформе (MCU, DSP или FPGA) и экспериментальный анализ амплитуды выходного сигнала фильтра при нулевом входе. Именно поэтому проектирование цифровых БИХ-фильтров высокого порядка ($N > 20$) не должно ограничиваться только теоретическим нахождением коэффициентов (как это обычно принято), но обязательно этапом реализации фильтра и практического исследования решения на возможное наличие предельных циклов. В том же случае, когда предельный цикл наблюдается, действенной мерой его устранения является проведение повторного синтеза фильтра под меньшую допустимую добротность его полюсов. А так как добротность полюсов пропорциональна их радиусу в z -плоскости, то достаточно задать меньшее значение допустимого радиуса полюсов передаточной функции и вновь синтезировать фильтр. Попытки задания различных радиусов полюсов [6] на стадии аналитического проектирования БИХ-фильтров малоэффективны из-за низкой адекватности и чрезмерных вычислительных затрат, тогда как поисковый синтез численными методами ЦНП позволяет это сделать без затруднений. Таким образом, совместное условие устойчивости проектируемого БИХ-фильтра с учётом отсутствия предельных циклов можно представить как

$$|Z_{pi}| < r_{max} ,$$

где r_{max} – допустимый максимальный радиус полюсов передаточной функции фильтра в z -плоскости, при котором предельные циклы в системе отсутствуют.

Так в пакете ЦНП 3.1 возможно задание следующей линейки значений максимальных радиусов полюсов при решении конкретной проектной задачи: $\Gamma_{max} = \{1,0, 0,98, 0,9, 0,82 \text{ и } 0,75\}$. Как показала практика, при синтезе с меньшим значением максимального радиуса, практически всегда удаётся получить проектное решение без предельных циклов того или иного рода, хотя селективная способность фильтра при этом, естественно, снижается.

Что касается результатов необходимых для расчёта отклика ЦЦФ промежуточных вычислений, то все они также являются целочисленными, и при заданной битовой разрядности квантования входного сигнала W_x достаточно выделить внутренний аккумуляторный регистр с разрядностью

$$W_{ak} = W_x + W_k + 2 \text{ [бит]}$$

для хранения результата целочисленного умножения с накоплением, осуществляемого по алгоритму (2). Колебаний переполнения, то есть возникновения больших предельных циклов, вызванных переполнением разрядной сетки регистра-аккумулятора, при таком расчёте его

разрядности практически никогда не возникает, особенно если учесть, что накопление результата целочисленного умножения в (2) осуществляется алгебраически, с учётом знака слагаемых, что существенно понижает разрядность результата.

Как известно, в каскадных формах построения цифровых фильтров необходима процедура масштабирования сигнала, что позволяет каскадному фильтру работать в широком динамическом диапазоне входных сигналов. Однако в каскадных ЦЦФ расчёт такого масштабирования для целочисленного звена гораздо легче осуществлять не применением L_p -нормы, а прямым введением требования обеспечения малого разброса коэффициентов передачи отдельных звеньев [9, 10]. Формально требование масштабирования усиления записывается двусторонними функциональными ограничениями (9) экстремальной задачи ЦНП-синтеза.

В общем виде задачу целочисленного нелинейного программирования при машинном синтезе рекурсивного ЦЦФ с заданной разрядностью представления данных можно записать так [10]:

$$F^o(\mathbf{IX}^o) = \min F(\mathbf{IX}) \quad \mathbf{IX} \in I^{6m} \quad (5)$$

$$-2^{W_k} - 1 \leq a_i \leq 2^{W_k} - 1 \quad i = \overline{1, m}, \quad (6)$$

$$-2^{W_k} - 1 \leq b_i \leq 2^{W_k} - 1 \quad i = \overline{1, m},$$

$$a_{0i} \in \{2^q\}, \quad q = 0, W_k \quad i = \overline{1, m}, \quad (7)$$

$$|Z_{pi}| < r_{max} \quad i = \overline{1, m}, \quad (8)$$

$$K_i^{\min} \leq |K_i(e^{j\omega})| \leq K_i^{\max} \quad i = \overline{1, m}, \quad (9)$$

где m – число звеньев второго порядка, \mathbf{IX} – вектор многомерного целочисленного пространства параметров (коэффициентов), $F(\mathbf{IX})$ – целевая функция, K_i^{\min} , K_i^{\max} – допустимые границы изменения коэффициента усиления i -го звена.

Экстремальная задача синтеза (5) записана относительно целочисленного пространства I^{6m} параметров (коэффициентов фильтра), размерностью $6m$. Ограничения (6) задают границы изменения этих целочисленных коэффициентов, а соотношение (7) определяет принадлежность коэффициентов a_{0i} биномиальному ряду. Функциональные ограничения (8) контролируют в процессе синтеза условие устойчивости рекурсивного фильтра по всем полюсам коэффициента передачи с радиусами, не выше Γ_{max} , а ограничения (9) масштабируют коэффициенты передачи звеньев в заданный интервал.

Многофункциональное задание целевой функции наиболее часто формируется в виде взвешенной суммы (10) частных целевых функций $f_i(\mathbf{IX})$, которые определяют выполнение функциональных требований по той или иной частотной характеристике фильтра [10, 13]:

$$F(\mathbf{IX}) = \sum_i \beta_i \cdot f_i(\mathbf{IX}) . \quad (10)$$

Коэффициент β_i задает значимость (вес) характеристики (i -го частотного окна). Сами частные целевые функции $f_i(\mathbf{IX})$ формирует функциональный редактор пакета синтеза обычно по критерию минимума среднеквадратичного отклонения

$$f_i(\mathbf{IX}) = \sqrt{\frac{1}{p} \cdot \sum_{n=1}^p [Y_n(\mathbf{IX}) - Y_n^T]^2}, \quad (11)$$

где $Y_n(\mathbf{IX})$ – текущее значение характеристики фильтра на n -ой дискретной частоте диапазона определения, а Y_n^T – требуемое значение частотной характеристики.

Критерий максимальной ошибки

$$f_i(\mathbf{IX}) = \max_n |Y_n(\mathbf{IX}) - Y_n^T| \quad (12)$$

используется значительно реже. Очень часто ради экономии времени расчёта в соотношении (11) корень не извлекают и работают по ненормированному квадратичному критерию

$$f_i(\mathbf{IX}) = \sum_{n=1}^p [Y_n(\mathbf{IX}) - Y_n^T]^2. \quad (13)$$

Поисковое итеративное решение экстремальной задачи ЦНП (5) в заданном пространстве параметров осуществляет программный алгоритмический комплекс [7, 13], обращаясь к модельному блоку программы для расчёта текущих функциональных характеристик фильтра. Вектор \mathbf{IX}^0 , минимизирующий скалярную целевую функцию $F(\mathbf{IX})$ на множестве допустимых целочисленных решений (6), является эффективным решением задачи параметрического синтеза рекурсивного ЦЦФ.

Следует отметить, что, в отличие от классического аналитического расчёта, поисковое проектирование является, безусловно, интеллектуальным процессом. Множество сценариев решения сложной проектной задачи может быть предложено, много специфических приёмов и навыков может быть применено опытным проектировщиком-поисковиком для успешного решения сложной проектной задачи.

Типовым сценарием поискового проектирования каскадных ЦЦФ является сценарий в стиле динамического программирования, как последовательность поисковых задач с поэтапным повышением порядка проектируемого фильтра. На первом, стартовом этапе используют структуру низкого порядка (4-го или 6-го, не выше). Естественно, выполнение совокупных требований таким фильтром будет низким. На втором этапе уже данное решение используется в качестве исходного решения. Порядок фильтра при этом повышают путём дублирования найденных ранее коэффициентов звена (что в пакете синтеза может делаться автоматически). После нескольких подобных итераций и определяется итоговый порядок проектируемого фильтра, при котором погрешность выполнения сложных совокупных требований лежит в пределах заданного допуска.

Ниже показана возможность синтеза методом ЦНП целочисленных рекурсивных фильтров верхних частот минимальной разрядности представления данных с учётом заданного радиуса полюсов его передаточной функции в z -плоскости. Многофункциональный синтез при этом осуществляется как по критерию требуемой частотной селективности, так и по критерию возможного обеспечения линейности фазы в полосе пропускания фильтра. Таким образом, целевая функция в приведённых ниже проектных задачах формировалась в виде

взвешенной суммы двух частных целевых функций $f_{AЧХ}(\mathbf{IX})$ и $f_{ФЧХ}(\mathbf{IX})$, обеспечивающих соответственно выполнение требований как амплитудной селекции фильтра, так и линейности его фазы

$$F(\mathbf{IX}) = \beta_1 f_{AЧХ}(\mathbf{IX}) + \beta_2 f_{ФЧХ}(\mathbf{IX}), \quad (14)$$

где сами частные целевые функции формировались по аддитивному квадратичному критерию (13). Весовые коэффициенты целевых функций определялись практически в ходе поэтапного решения экстремальных задач целочисленного программирования.

Техническое задание для синтеза ФВЧ с требуемой фазовой линейностью в целочисленном пространстве 3-битовых параметров было следующим:

1. Полоса пропускания фильтра 0,24 – 0,5 в шкале относительных частот ff/s , где fs – частота дискретизации;
2. Коэффициент передачи в полосе пропускания 0 дБ с допуском $\pm 1,2$ дБ;
3. Неравномерность фазы в полосе не выше 10° ;
4. Переходная полоса 0,07 – 0,24;
5. Уровень подавления на частотах ниже 0,07 не менее 40 дБ;
6. Длина слова коэффициентов W_x – 3 бита;
7. Порядок рекурсивного фильтра – 20;
8. Максимально-допустимый радиус полюсов – 0,75;
9. Масштабирование коэффициентов передачи звеньев в интервал $\{0,5 - 2,5\}$.

Требуемая АЧХ фильтра приведена на рис. 3а (выделена пунктиром). При синтезе рекурсивного ФВЧ с линейной фазой целевая функция формировалась в виде взвешенной суммы частных целевых функций $f_{AЧХ}(\mathbf{IX})$ и $f_{ФЧХ}(\mathbf{IX})$ по соотношению (14). Сами частные функций формировались по аддитивному критерию (13).

Задача целочисленного программирования для многофункционального синтеза ФВЧ в форме каскадного соединения 10 звеньев второго порядка записывалась так

$$F^0(\mathbf{IX}^0) = \min F(\mathbf{IX}) \quad \mathbf{IX} \in I^{50} \quad (15)$$

$$-3 \leq a_i, b_i \leq 3 \quad i = \overline{1, 10}, \quad (16)$$

$$a_{0i} = 4 \quad i = \overline{1, 10}, \quad (17)$$

$$|Z_{pj}| < 0,75, \quad (18)$$

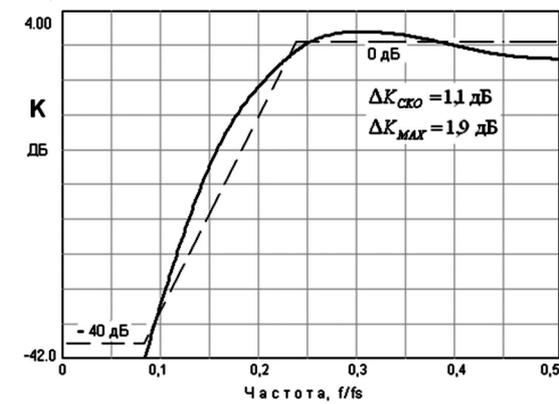
$$0,5 \leq K_i(e^{j\omega}) \leq 2,5 \quad i = \overline{1, 10}. \quad (19)$$

Таким образом, итоговая минимизация целевого функционала осуществлялась на 50-мерном целочисленном пространстве 3-битовых параметров в допустимой области (16) при выполнении функциональных ограничений устойчивости фильтра (18) по всем полюсам передаточной функции с радиусами, не превышающими $r_{\max} = 0,75$ в z -плоскости, и масштабировании (19) коэффициентов усиления звеньев в заданный интервал. Нормирующие коэффициенты (17) всех звеньев фильтра были одинаковыми. Для ввода структуры ППФ в программу использовался встроенный топологический редактор ЦНП-пакета, позволяющий сформировать файл исходных данных к решению конкретной задачи синтеза с указанием числа варьируемых коэффициентов, их начальных значений и границ изменения, а также возможного их дублирования в случае необходимости.

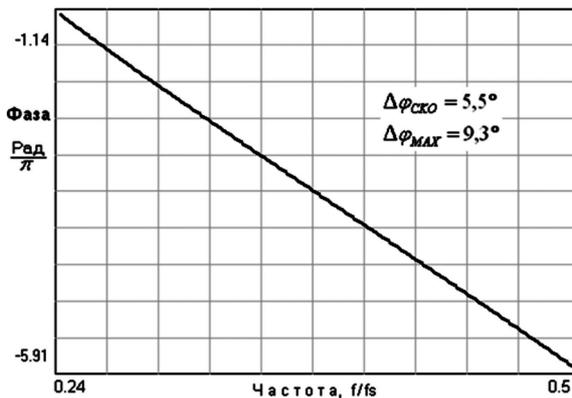
Таблица 1. Оптимальные коэффициенты трёхбитового ФВЧ

Звено фильтра	Коэффициенты передаточной функции фильтра					Усиление звена K_{max}
	a_1	a_2	b_0	b_1	b_2	
1	2	0	2	-3	-1	2,0
2	0	0	-2	3	0	1,3
3	0	2	0	3	1	1,6
4	2	1	0	-3	0	1,2
5	-3	2	1	3	-3	1,5
6	-2	0	1	3	-2	0,9
7	0	0	-3	3	1	1,3
8	0	0	0	3	-3	1,5
9	0	0	1	3	-2	1,1
10	1	0	-3	0	-1	1,3

В табл. 1 приведены оптимальные значения целочисленных 3-битовых коэффициентов передаточной функции по синтезу ФВЧ с одинаковой значимостью ($\beta_1 = \beta_2$) требований к АЧХ и ФЧХ, а также достигнутые значения коэффициентов усиления его звеньев. Графики его частотных характеристик представлены на рис. 3. Время решения данной задачи на стандартном персональном компьютере по сценарию динамического программирования и целевой функции (14) не превышало 10 минут.



а) АЧХ



б) ФЧХ

Рис. 3. Характеристики ФВЧ по синтезу

Что касается оценки фазовой нелинейности в полосе пропускания синтезированного фильтра, то по критерию влияния на форму выходного сигнала корректно применение в данном случае аддитивной среднеквадратичной оценки (СКО) фазовых искажений

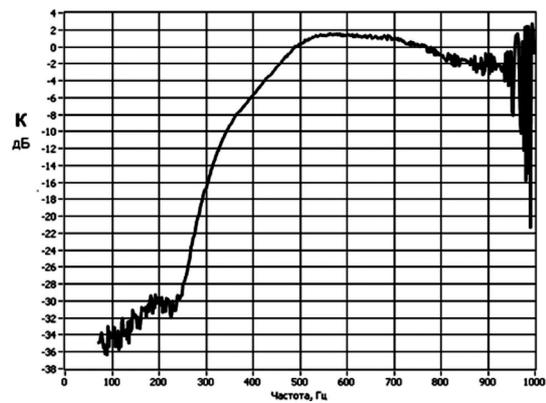
$$\Delta\varphi_{ско} = \sqrt{\frac{1}{P} \cdot \sum_{n=1}^P [\varphi_n(\mathbf{IX}) - \varphi_n^L]^2}, \quad (20)$$

а не оценки по критерию максимальной ошибки

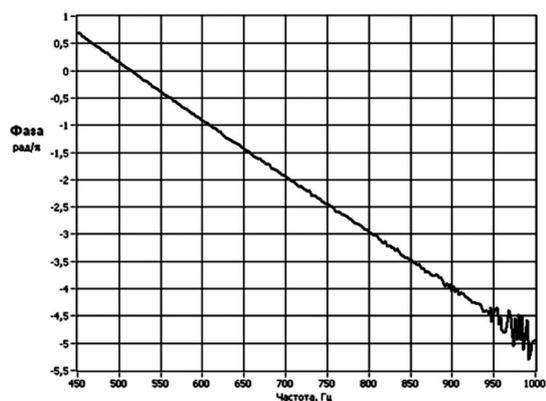
$$\Delta\varphi_{\max} = \max_n |\varphi_n(\mathbf{IX}) - \varphi_n^L|,$$

где $\varphi_n(\mathbf{IX})$ – текущее значение ФЧХ на n -ой дискретной частоте интервала оценки, а φ_n^L – требуемое значение линейной ФЧХ. Это следует непосредственно из преобразования Фурье, которое определяет именно аддитивное влияние фазового и амплитудного спектра всех гармоник обрабатываемого сигнала на искажение его формы. Поэтому в дальнейшем для оценки фазовой нелинейности, а также неравномерности АЧХ в полосе пропускания фильтра, будет применяться, в основном, оценка по критерию СКО, хотя и максимальная ошибка будет также приводиться. В данном случае среднеквадратичная неравномерность АЧХ и ФЧХ в полосе пропускания 3-битового ФВЧ были равны соответственно $\Delta K_{ско} = 1,1$ дБ и $\Delta\varphi_{ско} = 5,5^\circ$. Максимальный радиус полюсов передаточной функции при этом был равен 0,71.

Практическая реализация фильтра осуществлялась на многофункциональном микроконтроллере MSP430F1611 с целочисленным RISC-ядром [14]. Результаты автоматизированного измерения частотных характеристик фильтра на всём интервале Найквиста для частоты дискретизации $f_s = 2$ кГц. приведены на рис. 4. Так как выходной сигнал снимался непосредственно с ЦАП, то также наблюдается возрастание шумов измерения вблизи частоты Найквиста, где эффекты квантования сигнала наиболее выражены.



а) АЧХ



б) ФЧХ

Рис. 4. Экспериментальные измерения 3-битового ФВЧ

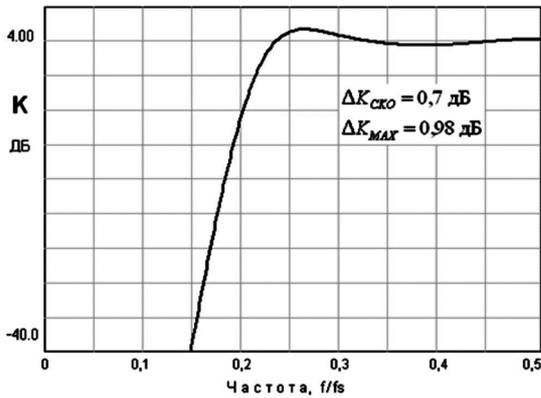
Анализ выходного сигнала при нулевом входе показал, что предельные циклы того или иного рода при максимальном радиусе полюсов 0,71 в синтезированном

фильтре отсутствуют. Таким образом, функциональные требования при ЦНП-синтезе трёхрядного рекурсивного ФВЧ с одинаковой значимостью ($\beta_1 = \beta_2$) требований к АЧХ и ФЧХ фильтра были выполнены.

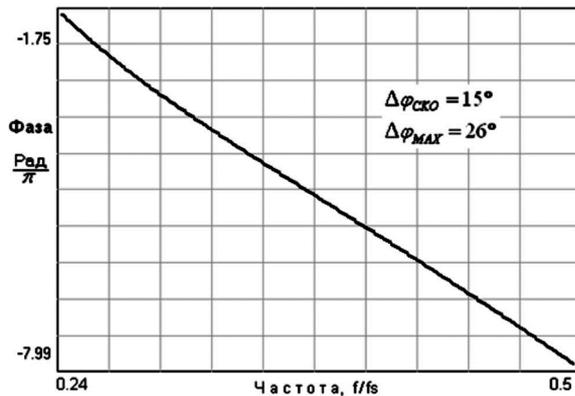
В табл. 2 приведены оптимальные значения 3-битовых коэффициентов передаточной функции уже по синтезу ФВЧ с различной значимостью ($\beta_1 = 5\beta_2$) требований к АЧХ и ФЧХ. Графики его частотных характеристик представлены на рис. 5.

Таблица 2. Оптимальные коэффициенты трёхбитового ФВЧ

Звено фильтра	Коэффициенты передаточной функции фильтра					Усиление звена K_{max}
	a_1	a_2	b_0	b_1	b_2	
1	-2	2	-1	-3	-2	1,8
2	1	0	-2	2	0	1,3
3	-1	2	0	2	-2	1,4
4	-1	1	2	-3	0	1,1
5	2	0	1	2	-3	2,0
6	-1	2	0	2	-3	1,7
7	0	0	-3	2	1	1,2
8	1	0	0	2	-3	1,7
9	0	0	0	2	-2	1,0
10	1	0	0	2	-2	1,3



а) АЧХ



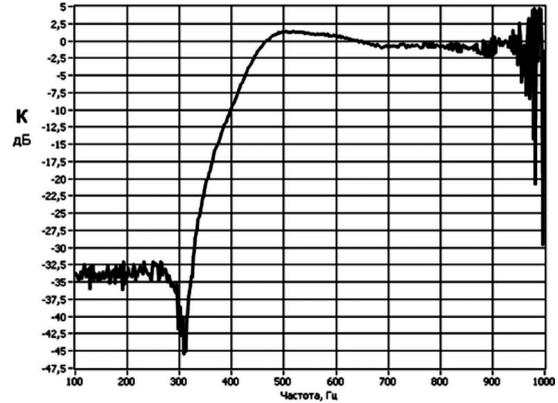
б) ФЧХ

Рис. 5. Характеристики ФВЧ по синтезу

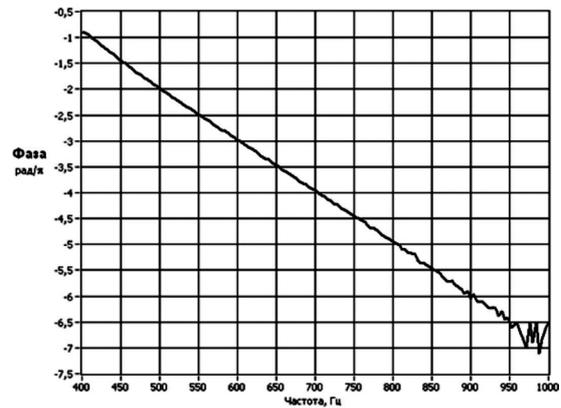
Как видно, в данном случае ширина переходной зоны существенно уменьшилась, а среднеквадратичная неравномерность АЧХ в полосе пропускания снизилась до значения $\Delta K_{CKO} = 0,7$ дБ, в то время как фазовые искажения возросли до $\Delta \varphi_{CKO} = 15^\circ$. Этого и следовало ожидать, так как в последовательной форме построения рекурсивный фильтр является квазимиимальнофазовой структурой, в которой модуль и аргумент передаточной функции связаны преобразованием Гильберта.

Строго минимальнофазовой структурой синтезированный фильтр не является, поскольку имеет нули коэффициента передачи вне пределов единичной окружности в z -плоскости.

Результаты измерения частотных характеристик фильтра (рис. 6) на всём интервале Найквиста для частоты дискретизации $f_s = 2$ кГц соответствуют данным синтеза.



а) АЧХ



б) ФЧХ

Рис. 6. Экспериментальные измерения 3-битового ФВЧ

Анализ выходного сигнала при нулевом входе показал, что предельные циклы того или иного рода при максимальном радиусе полюсов 0,71 в синтезированном фильтре также отсутствуют.

Заключение

Методы целочисленного нелинейного программирования в приложении к задачам проектирования линейных цифровых фильтров являются современной альтернативой традиционным методам проектирования цифровых фильтров [2 – 5]. Классическое проектирование базируется, как известно, на аналитическом представлении функциональных зависимостей, начиная с аналитической аппроксимации требуемой частотной характеристики БИХ-фильтра, что, естественно, приводит к чрезвычайной сложности аналитических вычислений даже в относительно простых задачах. Так аналитический расчёт цифровых фильтров или фазовых корректоров даже с учётом возможности обеспечения только требуемой фазовой линейности, как справедливо указано в [2, 3, 15], уже вызывает значительные математические трудности. Если же необходимо реализовать некий сложный закон изменения характеристики, например,

фазы коэффициента передачи либо её производных (ГВЗ и дисперсии), то аналитический расчёт фильтра становится невозможен. Принципиальное отличие ЦНП-синтеза состоит в применении современных численных методов машинного проектирования, позволяющих работать не с аналитическим, а с дискретным представлением характеристик проектируемого фильтра, тогда как исходные требуемые, так и текущие характеристики табулированы с заданной дискретностью их представления в частотной области и в вычислительной системе представлены двумерными массивами (векторами). Это даёт возможность, с одной стороны, без труда, применением численных методов, рассчитывать с заданной точностью все требуемые характеристики фильтра (включая ГВЗ и дисперсионные характеристики). С другой стороны – применять для синтеза технического решения весьма эффективные поисковые методы дискретного программирования, позволяющие осуществлять проектирование цифровых БИХ-фильтров непосредственно в целочисленном пространстве состояний. Критерием поиска при этом является соответствие совокупного текущего функционирования синтезируемого фильтра его требуемому функционированию. Современные алгоритмические комплексы целочисленной минимизации позволяют решать такие проектные задачи надёжно и эффективно при выполнении всех внешних требований и ограничений к работе цифрового фильтра, что даёт возможность существенно повысить качество проектируемых фильтров и сократить время их разработки.

Из материалов, приведённых в статье видно, что в сравнении с традиционными подходами, синтез цифровых фильтров методом целочисленного нелинейного программирования позволяет:

1. Осуществлять синтез фильтра по совокупности требуемых его частотных характеристик при произвольной форме их задания и заданной частотной шкале (линейной, логарифмической и др.).

2. Гарантировать устойчивость решения для целочисленных БИХ-фильтров приоритетным выполнением функциональных условий устойчивости в процессе ЦНП-синтеза фильтра. При этом возможно задание требуемого максимального радиуса полюсов передаточной функции при синтезе, что позволяет эффективно управлять добротностью проектируемого фильтра в случае возникновения предельных циклов того или иного рода.

3. Проектировать фильтры с любой, наперёд заданной разрядностью представления данных (вплоть до 3-х бит) непосредственно в целочисленном пространстве параметров (коэффициентов) фильтра.

Глобальная модельная идеология поиска определяет высокую надёжность отыскания эффективного решения экстремальной задачи ЦНП. Хорошего начального приближения (прототипа) здесь не требуется. Как правило, в проектных задачах со сложными селективными требованиями оптимальное решение определяется не из начальной точки, заданной пользователем, а из точки, сгенерированной самим поисковым алгоритмом решения задачи.

Необходимое масштабирование сигнала в каскад-

ных структурах может быть обеспечено непосредственно в ходе ЦНП-синтеза целочисленного фильтра. Здесь нет необходимости использования косвенных приёмов масштабирования усиления применением, например, L_p -нормы.

Целочисленная дискретизация пространства параметров (коэффициентов) фильтра позволяет получать проектные решения в целых числах, что снимает все ограничения по арифметике вычислений при реализации ЦНП-фильтров на любых цифровых платформах (сигнальных процессорах, контроллерах, FPGA) с заданной разрядностью представления данных, а также на кристаллах заказных или полужаказных СБИС.

Литература

1. Айфичер Э., Джервис Б. Цифровая обработка сигналов: практический подход. М.: Издательский дом «Вильямс», 2004.
2. Рабинер Л., Гоулд Б. Теория и применение цифровой обработки сигналов. - М.: Мир, 1978.-848 с.
3. Сергиенко А.Б. Цифровая обработка сигналов. СПб.: Питер, 2002. 608 с
4. EzIIR filter design package. <http://www.ti.com/tool/sprc072>.
5. Мингазин А.Т. Синтез передаточных функций цифровых фильтров в области дискретных значений коэффициентов (обзор). // Электронная техника. Сер. 10. 1993. № 1,2. С. 3-35.
6. Mathias Lang. Algorithms for the Constrained Design of Digital Filters with Arbitrary Magnitude and Phase Responses. // Vienna, June 1999.
7. Бугров В.Н. Проектирование цифровых фильтров методами целочисленного нелинейного программирования. // Вестник ННГУ, 2009, № 6. с. 61 – 70.
8. Шкелев Е.И., Бугров В.Н., Пройдаков В.И., Артемьев В.В. Целочисленные цифровые фильтры – эффективное решение для 8-битовых цифровых платформ. М., Компоненты и технологии, № 10, 2013, с. 104 – 110.
9. Бугров В.Н., Пройдаков В.И., Артемьев В.В. Синтез цифровых фильтров методами целочисленного нелинейного программирования. 17-ая международная конференция «Цифровая обработка сигналов и её применение – DSPA-2915», Тезисы докладов. М.: НТО РЭС им. А.С. Попова, 2015 г, с. 200 – 204.
10. Бугров В.Н. Сложноселективные задачи целочисленной цифровой фильтрации. М., Компоненты и технологии, № 10, 2016, с. 100-120.
11. Бугров В.Н, Морозов Н.С. Синтез целочисленных цифровых КИХ-фильтров с линейной фазой. // Цифровая обработка сигналов, №1, 2016, с. 14-19
12. Dempster A.G., Macleod M.D. IIR digital filter design using minimum adder multiplier blocks.//IEEE Trans.on Circuits and Systems-II, 1998, v. 45, N 6.
13. Воинов Б.С., Бугров В.Н., Воинов Б.Б. Информационные технологии и системы: поиск оптимальных, оригинальных и рациональных решений. М.: Наука, 2007, 730 с.
14. Семенов Б.Ю. Микроконтроллеры MSP430. Первое знакомство, М.: Изд-во «Солон-пресс», 2006.
15. Кисель В.А. Аналоговые и цифровые корректоры: Справочник. - М.: Радио и связь, 1986, 184 с.

МЕТОД ИЗМЕРЕНИЯ ЧАСТОТЫ ОСНОВНОГО ТОНА С МЕЖПЕРИОДНЫМ НАКОПЛЕНИЕМ РЕЧЕВОГО СИГНАЛА

Савченко В.В., д.т.н., профессор, профессор кафедры «Математика и информатика» Нижегородского государственного лингвистического университета; e-mail: svv@lunn.ru.

MEASUREMENT METHOD OF PITCH WITH INTER PERIOD ACCUMULATION OF SPEECH SIGNAL

Savchenko V.V.

A new method for measuring the pitch frequency of a speech signal in conditions of increased noise is proposed. The effectiveness of the method is achieved due to the effect of interperiodic accumulation. This effect is realized in a multichannel frequency measurement system using several parallel connected signal recirculator with adjustable delay periods in the feedback circuits. The effectiveness of the method has been studied theoretically and experimentally. The accuracy and sensitivity of the method are estimated as a function of the interference intensity at the input of the measurer. It is shown that for a signal-to-interference ratio of 20 dB or more, the error of the new method does not exceed (1...2) % of the nominal value of the frequency. It is agrees well with the potentially achievable accuracy under the conditions in question. In conditions of high noise the gain the threshold value of the signal-to-interference ratio in comparison with the world analogs is 4-5 dB or more.

Key words: speech, speech signal, pitch frequency, automatic speech processing, speech technology.

Ключевые слова: речь, речевой сигнал, частота основного тона, автоматическая обработка речи, речевые технологии.

Введение

Частота основного тона (ЧОТ) относится к наиболее информативным акустическим характеристикам речевого сигнала [1-3] и в этом качестве играет важную роль при автоматической обработке речи в системах самого разного назначения [4, 5], в том числе на выходе телефонного канала связи [6]. Исследования в данной области непрерывно продолжаются [7, 8], в том числе, по актуальному направлению [9, 10] повышения помехоустойчивости измерителя ЧОТ. Однако из-за известной сложности такого рода задач в теории до сих пор отсутствует ее эффективное решение. Так, речевой сигнал принципиально вариативен по своей тонкой структуре [11], модулирован по амплитуде [12] и не стабилен по динамике [13, 14]. При этом действие акустического (фоновое) шума [15, 16] только усиливает все перечисленные факторы. Как следствие, большинство известных методов измерения ЧОТ хорошо зарекомендовало себя в «тепличных», бесшумных условиях и, вместе с тем, сильно теряет по эффективности при действии случайных помех средней мощностью минус 10 дБ и выше относительно мощности сигнала [17]. А это, между прочим, распространенные условия [3] производства и восприятия речи в процессе коммуникаций. Проблема обостряется условиями малых выборок наблюдений в пределах интервалов вокализации речевого сигнала конечной длительности T [18]. Разработке и исследованию метода измерения ЧОТ повышенной помехоустойчивости посвящена настоящая статья. В ней используется математический аппарат теории сигналов [19, 20].

Предложен новый метод измерения частоты основного тона речевого сигнала в условиях повышенного зашумления. Результативность метода достигается за счет эффекта межпериодного накопления. Указанный эффект реализуется в системе многоканального измерения частоты с использованием нескольких параллельно включенных накопителей-рециркуляторов с регулируемым периодом задержки сигнала в цепях обратной связи. Эффективность метода исследована теоретически и экспериментально. По результатам проведенного исследования даны оценки точности и чувствительности метода в зависимости от интенсивности помехи на входе измерителя. Показано, что при отношении сигнал-помеха 20 дБ и более погрешность нового метода не превышает (1...2) % от номинала частоты основного тона, что хорошо согласуется с потенциально достижимой точностью в рассматриваемых условиях. В условиях повышенного зашумления выигрыш в пороговой величине отношения сигнал-помеха по сравнению с мировыми аналогами составляет 4-5 дБ и более.

Постановка задачи

Согласно акустической теории речеобразования [2], сигнал $x(t)$ на выходе речевого тракта диктора на интервалах действия гласных фонем имеет линейчатый частотный спектр. Его минимальная (нижняя) частота F_0 и определяет текущее значение ЧОТ. В диапазоне ее значений (80...160) Гц, характерном для разговорной речи мужчин, при $T=150...200$ мс [11] будем иметь порядка $M = TF_0 = 15...30$ периодов основного тона в пределах вокализованных отрезков речевого сигнала. А это серьезный стимул для применения межпериодного накопления [21] в качестве радикального средства повышения помехоустойчивости обработки.

Известно [19-21], что указанное накопление в общем случае реализуется по схеме гребенчатого фильтра

накопления (ГФН), настроенного на частоту F_0 . В условиях априорной неопределенности теория [20] рекомендует многоканальную систему обработки сигнала с использованием набора из N параллельно включенных ГФН, перекрывающих своими частотными характеристиками весь анализируемый диапазон. Ее структурная схема изображена на рис. 1, где приняты следующие обозначения: АД – амплитудный (квадратичный) детектор, ФНЧ – фильтр нижних частот. Система АД–ФНЧ служит для измерения средней мощности или дисперсии речевого сигнала $y_i(t)$ на выходе ГФН в составе i -го канала. Решение в отношении оптимальной оценки частоты \hat{F}_0 здесь принимается по принципу максимума средней мощности $P(y_i)$ накопленного сигнала $y_i(t)$. Вопросы практического осуществления и анализа многоканальной системы (рис. 1) в задаче измерения ЧОТ подробно рассматриваются далее.

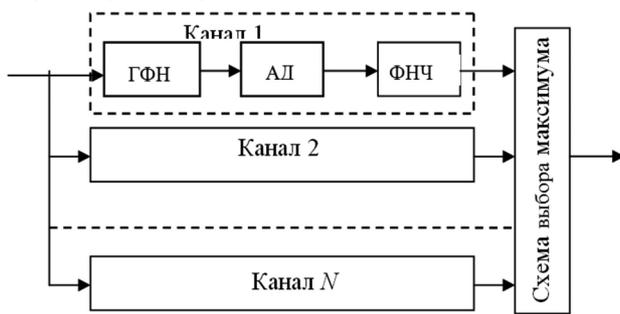


Рис. 1. Система оптимального измерения частоты

Синтез алгоритма

При большом числе каналов $N \gg 1$ (рис. 1) разработчики отдают предпочтение линейным рекурсивным фильтрам простейшего типа – по схеме рециркулятора [21]. Его динамика описывается рекуррентным уравнением первого порядка

$$y_i(t) = x(t) + by(t - T_i), \tag{1}$$

где $T_i = 1/F_i$ – период накопления сигнала $x(t)$, $b < 1$ – коэффициент его усиления в цепи обратной связи. Чем ближе значение b к единице, тем более выражен в (1) эффект накопления. Однако на практике следует учитывать конечную длительность речевого сигнала на интервалах его вокализации. Поэтому существует оптимальное значение b_0 параметра b из условия максимизации достигаемого выигрыша в отношении сигнал–помеха (ОСП) за счет накопления сигнала. В первом приближении можно записать (см., напр., монографию [21], ф. 5.4.9 на с. 202) $b_0 \approx 1 - 1,27/M$, или примерно 0,95 в рассматриваемой нами задаче. При этом период накопления сигнала в каждом канале (рис. 1) соответствует определенному варианту ЧОТ $F_i = 1/T_i$ в пределах диапазона ее ожидаемых значений $[F_{\min}; F_{\max}]$. В режиме дискретного времени $t = 0, \Delta T, 2\Delta T, \dots$ с частотой дискретизации речевого сигнала $F = const$ будем иметь $T_i = T_{i-1} + \Delta T$, где $\Delta T = 1/F$ – период взятия отсчетов. При равенстве $T_0 = 1/F_{\max} - \Delta T$ получаем

$$T_i = i\Delta T, i = \overline{1, N}, \tag{2}$$

где число каналов $N = (T_N - T_1) / \Delta T = F(F_{\min}^{-1} - F_{\max}^{-1})$. Например, при частоте дискретизации сигнала $F = 8$ кГц, согласованной с полосой пропускания стандартного телефонного канала связи [3], в расчете на мужские голоса будем иметь $N = 8000(80^{-1} - 160^{-1}) = 8000 / 160 = 50$.

Дополняя (1) оценкой средней мощности накопленного сигнала по формуле эмпирической дисперсии [19]

$$P(y_i) = (\tau F)^{-1} \sum_{k=1}^{\tau F} y_i^2(t + k\Delta T), \tag{3}$$

где τ – длительность речевого фрейма или интервала квазистационарности речевого сигнала, предполагаемого центрированным, получим квазиоптимальный алгоритм измерения ЧОТ на основе межпериодного накопления:

$$\hat{F}_0 = 1/T_m, m = \text{Arg max}_{i \leq N} P(y_i). \tag{4}$$

Его эффективность может быть охарактеризована как точностью результирующей оценки ЧОТ, так и помехоустойчивостью или надежностью обработки речевого сигнала в условиях действия случайных помех [20].

Анализ точности

Точность измерения ЧОТ согласно алгоритму (1)-(4) определяется, главным образом, его инструментальной погрешностью. При учете достаточно высокой разрешающей способности ГФН по частоте [21] $\Delta f \approx \pi^{-1} F_i \times (1 - b_0) / (1 + b_0)$, которая при $F_i \leq 160$ Гц и $b_0 = 0,95$ не превышает 1,3 Гц, определим инструментальную погрешность алгоритма через величину частотной расстройки

$$\begin{aligned} \Delta F_i &= 0,5(1/T_i - 1/T_{i+1}) = \\ &= 0,5\Delta T / (T_i T_{i+1}) = 0,5 / (F T_i T_{i+1}). \end{aligned}$$

двух соседних каналов измерителя ЧОТ (рис. 1). Нетрудно понять, что она ограничена сверху предельным уровнем $0,5 / (F T_1^2) = 0,5 F_{\max}^2 / F$. Например, в рассматриваемом диапазоне частот будем иметь $\Delta F_i \leq 0,5 \times 160^2 / 8000 = 1,6$ Гц. Отметим, что полученный результат отвечает требованиям госстандарта [22] к частотным измерителям резонансного типа и при этом превышает на (30...40) % показатели точности современных методов измерения ЧОТ из обзорной работы [17]. Аналогичный вывод можно сделать и в отношении помехоустойчивости предложенного метода.

Со ссылкой на монографию [21] при действии гауссовских помех помехоустойчивость метода может быть охарактеризована выигрышем ОСП по мощности q^2 величиной порядка $W \approx 0,8M$. В нашей задаче она составляет примерно 12 дБ. Сказанное в равной мере относится как к помехам типа белого шума, так и к подобным речевому сигналу помехам с «окрашенными» спектрами мощности [15]. Второй тип помех имеет неотличимую от полезного сигнала внутрпериодную тонкую (формантную) структуру [18] и поэтому представляет неразрешимую на данный момент проблему [17] для большинства известных методов измерения ЧОТ. Одна-

ко, с точки зрения межпериодной обработки (1) подобные помехи неотличимы от белого шума, т.к. не содержат в своем спектре достаточно мощных периодических компонент в пределах рабочего диапазона частот. Поэтому вне зависимости от типа действующих помех эффективность предложенного метода в статистическом смысле близка к потенциально достижимой эффективности измерения частоты. В пересчете к среднеквадратичной величине ошибки измерений [20] через эффективную длительность речевого сигнала, примерно равную WT_i , при $q^2 = 20$ дБ получим $\sigma_0^* \approx 1 \div \sqrt{q^2 (WT_i)^2} \leq F_{\max} / (W \sqrt{q^2}) = 160 / (0,8 \times 20 \times 10) = 1$ Гц. Подчеркнем, это справедливо при почти идеальных условиях. Но, например, уже при $q^2 = 10$ дБ будем иметь как минимум $\sigma_0^* = 3,16$ Гц, а при $q^2 = 0$ дБ – в три раза больше: $\sigma_0^* = 10$ Гц. Исследование эффективности предложенного в статье метода в условиях существенного зашумления речевого сигнала проводится далее экспериментальным путем. Об актуальности этой задачи свидетельствует, например, следующий факт: заявленные в работе [9] в качестве среднеквадратичной величины ошибки измерения ЧОТ (0,039...0,045) Гц при применении так называемого «метода активного восприятия»¹ для случая действия белого гауссовского шума при двух рассмотренных выше значениях ОСП более чем на два порядка (!) превышают потенциальную точность оптимального измерителя частоты.

Программа и методика эксперимента

Исследования проводились в два этапа. На первом этапе условный диктор (автор статьи) несколько раз прочитал через микрофон художественный текст из первой главы поэмы А.С. Пушкина «Евгений Онегин». И каждый раз записал прочитанный текст на персональный компьютер в виде звукового файла. Затем автоматически из каждой записи были выделены в отдельные файлы все ее вокализованные отрезки. При этом использовалась известная [23] методика. При длительности каждой записи 3 мин. и более в результате было получено не менее 200 файлов речевого сигнала $x(t)$ длительностью 150-200 мс каждый. Согласно рекомендациям работы [24] этого вполне достаточно для получения в дальнейшем надежных статистических оценок. Тем самым на первом этапе экспериментальных исследований был сформирован необходимый речевой материал.

На втором этапе полученные файлы были подвергнуты компьютерной обработке согласно разработанному выше алгоритму (1)-(4). Для этого была создана и использована специальная авторская программа

«Speech Transform», включающая в себя рециркулятор с регулируемым периодом задержки T_i сигнала $x(t)$ в цепи обратной связи. Главное окно программы показано на рис. 2. Здесь в правой части указаны параметры алгоритма и введено обозначение $n = T_i F$ для относительной величины периода T_i . Оптимальная оценка ЧОТ определяется в данной программе через частоту дискретизации речевого сигнала простым соотношением вида $\hat{F}_0 = 1 / T_m = F / n^*$, где n^* – значение параметра для канала (рис. 1) с максимальным сигналом на выходе. В нашем случае мы получили $n^* = 81$ и, следовательно, $\hat{F}_0 = 8000 / 81 = 98,8$ Гц.

В верхней части окна показана временная диаграмма одного из файлов $x(t)$. В данном случае это был сигнал ударной гласной «А» из слова «прАвил». В нижней части окна показана временная диаграмма сигнала на выходе ГФН в составе определенного канала измерителя

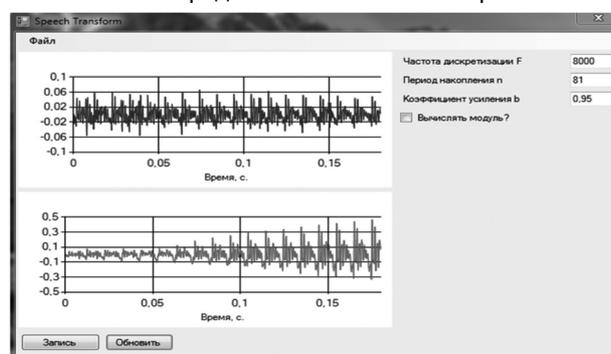


Рис. 2. Скриншот главного окна компьютерной программы

ЧОТ (рис. 1). В нашем случае это был оптимальный, m -й канал с сигналом $y_m(t)$ максимальной мощности. Из сопоставления двух диаграмм на рис. 2 хорошо виден эффект межпериодного накопления речевого сигнала в ГФН. Его исследованию и был посвящен, главным образом, весь дальнейший эксперимент. При этом использовалась аддитивная модель смеси сигнала с помехой $\tilde{x}(t) = x(t) + \eta(t)$, в рамках которой помеха формировалась программным способом с использованием стандартного генератора белого гауссовского шума. Его дисперсия, а вслед за ней и ОСП варьировались в эксперименте в широких пределах. И каждый раз вычислялась средняя мощность $P(y_m)$ накопленного сигнала $y_m(t)$ на интервале в один речевой фрейм длительностью $\tau = 10$ мс или на интервале в $\tau F = 80$ отсчетов данных. Полученные результаты представлены на рисунках ниже.

Основные результаты и выводы

Для оценки чувствительности алгоритма (1)-(4) в качестве важнейшей технической характеристики измерителей частоты [22] на рис. 3 представлена зависимость избирательной способности метода по частоте $\gamma_m(i) = P(y_{m+i}) / P(y_m)$ через относительную величину средней мощности накопленного сигнала $P(y_{m+i})$ в каждом отдельном канале системы обработки (рис. 1), начиная с оптимального, m -го канала. При ее вычисле-

¹ А может ли быть восприятие не активным? На этот риторический вопрос в свое время четко ответил авторитетный ученый-психолог Дж. Гибсон: «Восприятие – это активный процесс извлечения информации об окружающем мире, включающий в себя реальные действия по исследованию того, что воспринимается» (см. его монографию «Экологический подход к зрительному восприятию» - М.: Прогресс, 1988).

ниях использовалась выборка из 200 независимых реализаций вокализованных отрезков речевого сигнала $x(t)$, а ОСП составляло 20 дБ.

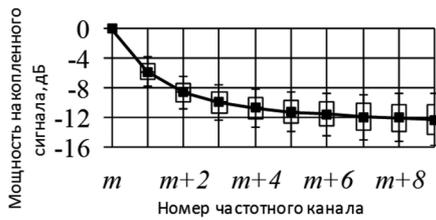


Рис. 3. Избирательная способность по частоте

Как видим, зависимость в данном случае носит ярко выраженный спадающий характер, причем, особенно сильно – в начале координат. Ее первое значение $\gamma_m(1)$ составляет порядка минус 6 дБ. В этих условиях чувствительность предложенного метода может быть охарактеризована половиной частотного интервала между двумя соседними каналами измерителя ЧОТ (рис. 1). Путем несложных вычислений получаем $\Delta F_m = 0,5 \div \div (FT_m T_{m+1}) \approx 0,6$ Гц. Это с запасом удовлетворяет требованиям действующего стандарта [22].

Для оценки помехоустойчивости метода далее была исследована зависимость его избирательной способности $\gamma^* = \gamma_m(1)$ от ОСП. Для этого было сформировано достаточно представительное множество независимых реализаций белого гауссовского шума $\eta(t)$ разной дисперсии в расчете на фиксированную среднюю мощность речевого сигнала $x(t)$. Полученная зависимость $\gamma^*(q^2)$ представлена в виде диаграммы на рис. 4.

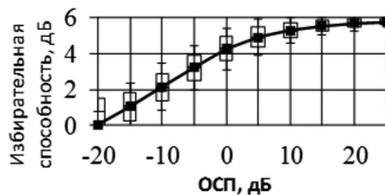


Рис. 4. Зависимость избирательной способности от ОСП

Пороговая величина ОСП здесь составила (-4...-10) дБ. Это на 4-5 дБ лучше, чем в методе, предложенном в работе [14], который был разработан автором специально для использования в условиях повышенного зашумления.

В развитие полученных результатов далее была исследована зависимость $\gamma^*(q^2)$ при действии случайной помехи с окрашенным спектром мощности. В эксперименте она была сформирована пропусканием нормального белого шума через ФНЧ с эффективной полосой пропускания, равной $0,1F = 800$ Гц. При этом для каждой точки q^2 использовалась выборка из $L = 260$ независимых файлов окрашенной помехи $\eta(t)$. Полученные оценки, как и ожидалось, по форме повторили зависимость, представленную на рис. 4. Потери в ОСП при этом не вышли за пределы (1...2) дБ – на уровне статистической погрешности измерений. В самом деле, следуя гауссовской аппроксимации усредненной выборочной величины γ^* , воспользуемся в качестве ее вероят-

ностной характеристики классическим выражением для половины длины доверительного интервала в его относительном выражении $\delta = z_p / \sqrt{L}$ [24]. Здесь z_p – коэффициент надежности или «доверия» на уровне значимости α , определяемый корнем уравнения $\Phi(z_p) = p$ с интегралом вероятности в левой части. При равенстве $\alpha = 0,1$ (соответствующая доверительная вероятность равна $p = 1 - \alpha = 0,9$) и объеме выборки $L = 260$ по таблицам нормального распределения [25] находим $z_{0,9} \approx 1,65$. И, следовательно, получаем $\delta \approx 10\%$. При учете номинального значения характеристики $\gamma^* \approx 6$ дБ (см. рис.4) данный результат представляется вполне приемлемым с точки зрения точности проведенного исследования.

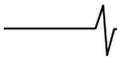
Заключение

В статье предложен новый метод измерения ЧОТ на основе межпериодного накопления речевого сигнала, исследована его эффективность. Для реализации эффекта накопления используется ГФН простейшей структуры – рециркулятор с регулируемым периодом задержки сигнала в цепи обратной связи. Благодаря накоплению предложенный метод обеспечивает выигрыш по сравнению с его известными аналогами как в точности измерения ЧОТ, так и в его чувствительности.

Полученные результаты и сделанные по ним выводы позволяют рекомендовать данный метод для практического применения в системах автоматической обработки речи при их работе в условиях повышенного зашумления.

Литература

1. Christensen M, Jakobsson A. Multi-pitch Estimation. – Morgan and Claypool, 2009. – 432 p.
2. Фант Г. Акустическая теория речеобразования. – М.: Наука, 1964. – 304 с.
3. Савченко В.В. Оценка качества цифровой передачи речи по конечной выборке речевого сигнала // Электросвязь. 2017. № 3. С. 52-57.
4. Лебедева Н.Н., Каримова Е.Д., Казимирова Е.А. Анализ речевого сигнала в исследованиях функционального состояния человека // Биомедицинская радиоэлектроника. 2015. № 2. С. 3-12.
5. Андреева Н.Г., Смирнова Т.А. Восприятие синтезированных моделей одноформантных гласных с разной частотой основного тона // Сенсорные системы. 2014. Т. 28. № 4. С. 13-21.
6. Чернобельский С.И. Сравнение результатов акустического анализа голоса при различных способах его записи // Вестник оториноларингологии. 2014. № 1. С. 41-43.
7. Алимуратов А.К. Исследование частотно-избирательных свойств методов декомпозиции на эмпирические моды для оценки частоты основного тона речевых сигналов // Труды Московского физико-технического института. 2015. Т. 7. № 3 (27). С. 56-68.
8. Вишнякова О.А., Лавров Д.Н. Гибридный алгоритм выделения частоты основного тона // Математические структуры и моделирование. 2016. № 1 (37). С. 59-65.



9. Гай В.Е. Метод оценки частоты основного тона в условиях помех // Цифровая обработка сигналов. 2013. № 4. С. 65-71.
10. Архипов И.О., Гиниятуллин Р.Р. Автокорреляционный выделитель основного тона с предварительным оцениванием частоты колебаний голосовых связок // В сборнике: «Молодые ученые – ускорению научно-технического прогресса в XXI веке». – Ижевск: Изд-во: ИННОВА, 2016. С. 421-428.
11. Savchenko V.V., Savchenko A.V. Information Theoretic Analysis of Efficiency of the Phonetic Encoding–Decoding Method in Automatic Speech Recognition // Journal of Communications Technology and Electronics. 2016. Vol. 61. No. 4. P. 430-435.
12. Азаров И.С., Вашкевич М.И., Петровский А. Алгоритм оценки мгновенной частоты основного тона речевого сигнала // Цифровая обработка сигналов. 2012. № 4. С. 49-57.
13. Вольф Д.А., Мещеряков Р.В. Модель и программная реализация сингулярного оценивания частоты основного тона речевого сигнала // Труды СПИИРАН. 2015. № 6. С. 191-209.
14. Савченко В.В. Тестирование случайных временных рядов на стационарность на основе принципа минимума информационного рассогласования // Известия вузов. Радиофизика. 2017. Т. 60. № 1. С. 89-96.
15. Savchenko V.V. Enhancement of the Noise Immunity of a Voice-Activated Robotics Control System Based on Phonetic Word Decoding Method // Journal of Communications Technology and Electronics. 2016. Vol. 61. No. 12. P. 1374 -1379.
16. Савченко В.В., Акатьев Д.Ю., Афонин М.В. Автоматическое распознавание речи на фоне шума // Со-временные тенденции развития науки и технологий. 2015. № 6-2. С. 99-102.
17. Hasan M.A. Pitch Detection Algorithm Based on Windowless Autocorrelation Function and Modified Cepstrum Method in Noisy Environments // IJCSNS International Journal of Computer Science and Network Security. 2017. Vol.17. No.2. P. 106-112.
18. Savchenko V.V. The Principle of the Information-Divergence Minimum in the Problem of Spectral Analysis of the Random Time Series Under the Condition of Small Observation Samples // Radiophysics and Quantum Electronics. 2015. Vol. 58. No. 5. P. 373-379.
19. Радиотехнические системы / Под ред. Ю.М. Казаринова. – М.: Издательский центр «Академия», 2008. – 592 с.
20. Лезин Ю.С. Введение в теорию и технику радиотехнических систем. – М.: Радио и связь, 1986. – 140 с.
21. Лезин Ю.С. Оптимальные фильтры и накопители импульсных сигналов. - М.: Советское радио, 1969. - 448 с.
22. ГОСТ 12692-67. Измерители частоты резонансные. Методы и средства проверок. - М.: Изд-во стандартов, 1967. - 7 с.
23. Savchenko V.V. Ponomarev D.A. Automatic Segmentation of Stochastic Time Series Using a Whitening Filter // Optoelectronics, Instrumentation and Data Processing. 2009. Vol. 45. No. 1. P. 37-42.
24. Савченко В.В. Определение объема контрольной выборки в условиях априорной неопределенности по принципу гарантированного результата // Научные ведомости БелГУ. Серия: Экономика. Информатика. 2015. № 1 (198). Вып. 33/1. С.74-78.
25. Большаков В.Д. Теория ошибок наблюдений. - М.: Недра, 1983. - 223 с.

УДК 621.396.962.2

МЕТОДИКА ОБОСНОВАНИЯ ТРЕБОВАНИЙ К АНАЛОГО-ЦИФРОВОМУ ПРЕОБРАЗОВАНИЮ В РЛС С СИНТЕЗИРОВАННОЙ АПЕРТУРОЙ НЕПРЕРЫВНОГО ИЗЛУЧЕНИЯ

Рязанцев Л.Б., к.т.н., докторант «Военного учебно-научного центра Военно-воздушных сил «Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж), e-mail: kernel386@mail.ru;

Купряшкин И.Ф., к.т.н., доцент, доцент кафедры «Военного учебно-научного центра Военно-воздушных сил «Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж), e-mail: ifk78@mail.ru;

Лихачев В.П., д.т.н., профессор кафедры «Военного учебно-научного центра Военно-воздушных сил «Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж), e-mail: lvp_home@mail.ru.

METHODOLOGY OF JUSTIFICATION OF REQUIREMENT TO THE ANALOG TO DIGITAL CONVERTER IN THE LFM CW SYNTHETIC APERTURE RADAR

Ryazantsev L.B., Kupryashkin I.F., Likhachev V.P.

The article is devoted to the investigation of the requirements for the bit depth of the analog-to-digital converter of a digital processing system with a synthesized antenna aperture radar with LFM CW signal. It is shown that taking into account the effect of oversampling occurring during the synthesis of radar images provides a significant reduction in the requirements for the capacity of the analog-digital converter.

Key words: FMCW SAR, dynamic range, ADC resolution, unmanned aerial vehicle, oversampling.

Ключевые слова: РЛС с синтезированной апертурой антенны, динамический диапазон, непрерывное излучение, разрядность АЦП, беспилотный летательный аппарат, сверхдискретизация.

Введение

На сегодняшний день РЛС с синтезированной апертурой непрерывного излучения (РСА НИ) в силу их малой массы, габаритов и относительной простоты конструкции являются одним из наиболее перспективных типов радиолокационных систем для малых беспилотных летательных аппаратов (БЛА) [1]. Примером таких РЛС являются зарубежные РЛС MiniSAR [2], ImSAR [3], и отечественная система, разработанная авторами [4].

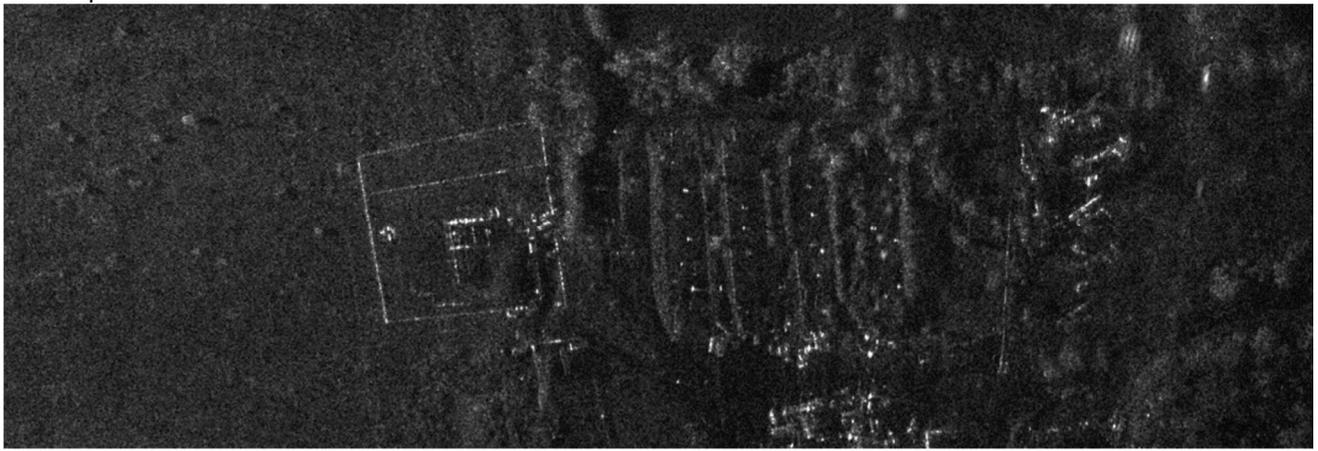
В ходе разработки РСАНИ выбор разрядности аналого-цифрового преобразователя (АЦП) осуществляется исходя из динамического диапазона (ДД) эхо-сигналов на выходе приемного устройства, который по сравнению с импульсными системами характеризуется большими значениями в силу того, что его верхняя граница определяется суммарной мощностью не отдельной разрешаемой полосы дальности, а суммарной мощностью эхо-сигнала всей картографируемой поверхности в пределах следа диаграммы направленности (ДН) антенны. При этом реализация в малогабаритных РСА НИ традиционных мер по расширению ДД, применяющихся в импульсных РЛС (временная регулировка усиления, косекансная форма ДН антенн), не представляется возможной, вследствие частотного принципа измерения дальности и жестких ограничений на массогабаритные характеристики антенных систем

Статья посвящена исследованию требований к разрядности аналого-цифрового преобразователя системы цифровой обработки РЛС с синтезированной апертурой антенны, использующей в качестве зондирующего непрерывный частотно-модулированный сигнал. Показано, что учет эффекта сверхдискретизации, имеющего место в процессе синтеза радиолокационных изображений в РЛС с непрерывным излучением, обеспечивает значительное снижение требований к разрядности аналого-цифрового преобразователя.

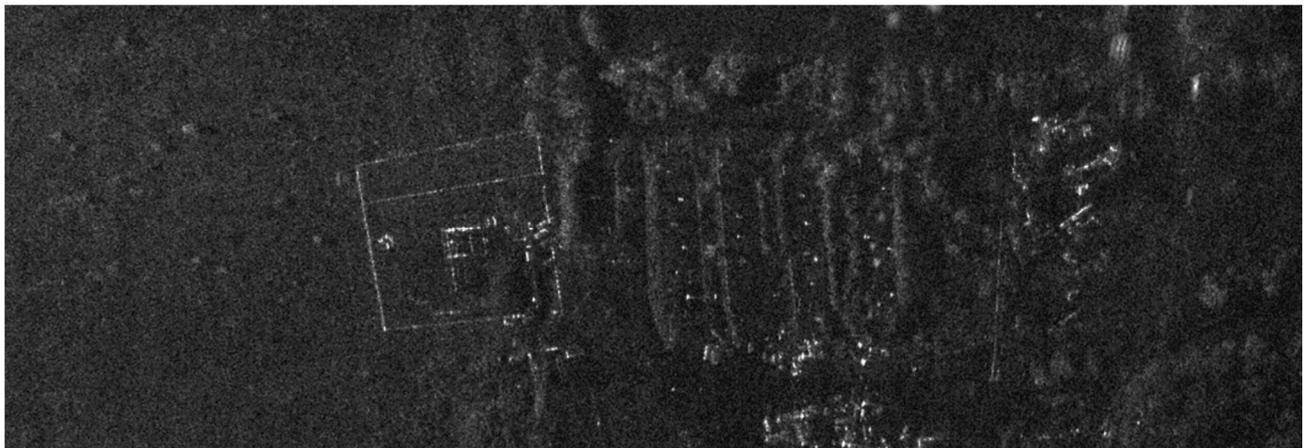
[5]. В результате ДД сигналов может достигать значений порядка 70...80 дБ, что предполагает использование АЦП с разрядностью 12-14 бит и частотой дискретизации до нескольких десятков МГц.

В ходе проведения исследований, направленных на уменьшение объема регистрируемой информации в интересах реализации возможности ее передачи по радиоканалу с борта БЛА на наземный пункт, было установлено, что последовательное уменьшение разрядности АЦП с 16 до 8, 4, 2 и 1 бит практически не приводит к заметному визуальному ухудшению качества формируемых радиолокационных изображений (РЛИ) и их дешифровочных свойств (рис. 1).

Это свидетельствует о том, что начальные требования к разрядности АЦП оказались избыточными, т.к. при оценке требуемой разрядности не учитывался эффект сверхдискретизации [6]. Такой эффект используется во временной области в цифровой обработке сигналов и состоит в том, что за счет увеличения частоты дискретизации и последующей низкочастотной фильтрации возможна оцифровка сигналов, амплитуда которых намного меньше единицы младшего разряда (ЕМР) АЦП [7]. При этом обязательным условием наступления эффекта является наличие в сигнале аддитивной шумовой состав-



а) 16 бит



б) 1 бит

Рис. 1. Фрагменты РЛИ земной поверхности (51°48'33"N, 39°08'25"E) в 3 см диапазоне при различных значениях $n_{\text{ацп}}$

ляющей со среднеквадратическим отклонением, превышающем значение ЕМР [8]. Аналогичный эффект может быть достигнут в частотной области при выполнении дискретного преобразования Фурье (ДПФ) над выборкой зашумленного сигнала, содержащей достаточно большое количество отсчетов.

Так как процесс формирования РЛИ в РСА НИ при условии качественной коррекции миграции дальности и фокусировке траекторного сигнала по сути сводится к вычислению его двумерного ДПФ [9], а суммарный эхо-сигнал следа ДН носит выраженный шумоподобный характер, то условия наступления эффекта сверхдискретизации можно считать выполняющимися. Отсюда следует, что при обосновании требований к разрядности АЦП РСА НИ необходимо принимать во внимание возможность ее снижения за счет учета описанного эффекта сверхдискретизации.

В то же время чрезмерное снижение разрядности АЦП приводит к появлению в спектре квантованного сигнала паразитных гармонических составляющих, проявляющихся на РЛИ в виде ложных отметок при наличии мощных отражателей, находящихся на малых дальностях. Этот факт также необходимо учитывать при обосновании требований к разрядности АЦП.

Таким образом целью статьи является разработка методики обоснования требований к разрядности АЦП, обеспечивающей приемлемое качество формируемых

радиолокационных изображений с учетом допустимого уровня искажений, обусловленных квантованием принятого сигнала.

Для обоснования методики необходимо оценить диапазон значений отношения уровня эхо-сигнала точечной цели к суммарному эхо-сигналу следа ДН на входе АЦП при типовых для РСА НИ условиях съемки земной поверхности. Далее рассчитывается минимальная разрядность АЦП, обеспечивающая оцифровку эхо-сигнала точечной цели в пределах всего диапазона уровней с учетом ограничений на максимальный уровень ложных отметок на РЛИ, появление которых обусловлено квантованием эхо-сигналов наиболее интенсивных отражателей в пределах зоны обзора.

Определение ДД сигналов на входе АЦП РСА НИ

Требуемая разрядность АЦП $n_{\text{ацп}}$ РСА НИ может быть выбрана исходя из отношения мощности сигнала точечной цели к мощности суммарного сигнала следа ДН

$$q = P_c / P_\phi, \quad (1)$$

где $P_c \approx \sigma_{\text{ц}} / R_{\text{ц}}^4$; $\sigma_{\text{ц}}$, $R_{\text{ц}}$ – эффективная площадь рассеяния (ЭПР) и наклонная дальность до точечной цели, соответственно:

$$P_\phi \approx \sigma_\phi \int_0^{r_{\text{max}}} \int_{-\pi}^{\pi} f^4(\epsilon, \beta) R^{-4}(\epsilon) \gamma(\epsilon) d\beta dr,$$

где σ_ϕ – средняя удельная ЭПР фона в пределах следа ДН; $\gamma(\varepsilon) = \cos(\varepsilon)$ – коэффициент рассеяния, обуславливающий влияние угла облучения на отражающие характеристики местности; $f^2(\varepsilon, \beta)$ – функция, описывающая вид диаграммы направленности антенны; ε , β , r – угол места, азимут и дальность, соответственно; $R(\varepsilon) = h/\cos(\varepsilon)$; h – высота полета носителя РСА. Учитывая, что $r = htg(\varepsilon)$ и, применяя метод замены переменной интегрирования, получим

$$P_\phi \approx \sigma_\phi \int_0^{\varepsilon_{\max}} \int_{-\pi}^{\pi} \frac{\sin(\varepsilon) \cos(\varepsilon)}{h^2} f^4(\varepsilon, \beta) \gamma(\varepsilon) d\beta d\varepsilon, \quad (2)$$

где $\varepsilon_{\max} = \arccos(h/R_{\max})$; R_{\max} – расстояние до дальней границы зоны обзора РСА.

Так, на рис. 2 представлены зависимости q от $R_{\text{ц}}$ для следующих условий съемки: $h = 500$ м, $R_{\max} = 10$ км; $\sigma_\phi = -15$ дБ (степь); $f^2(\varepsilon, \beta) = 1$ при $\beta \in [-\Theta_{\text{аз}}/2; \Theta_{\text{аз}}/2]$ и $f^2(\varepsilon, \beta) = 0$ при $\beta \notin [-\Theta_{\text{аз}}/2; \Theta_{\text{аз}}/2]$; $\Theta_{\text{аз}} = 20^\circ$.

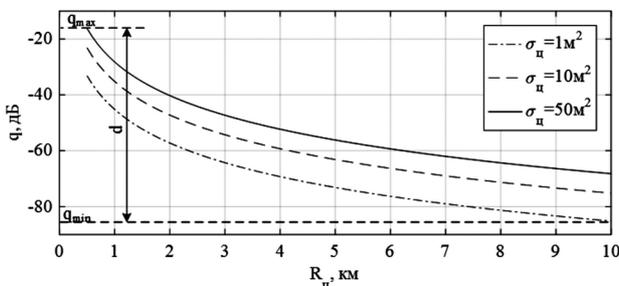


Рис.2. Зависимости q от $R_{\text{ц}}$ и $\sigma_{\text{ц}}$

Из рис. 2 видно, что при заданных условиях диапазон относительных уровней эхо-сигналов точечных целей (динамический диапазон) $d = |q_{\max} - q_{\min}|$ достигает 70 дБ. При других условиях съемки, а также в случае решения задач обнаружения распределенных целей и картографирования земной поверхности ДД сигналов может быть выше на 10...20 дБ.

Выбор разрядности АЦП, обеспечивающего оцифровку сигналов в заданном ДД

Аналогично тому, как во временной области увеличение частоты дискретизации с последующей низкочастотной фильтрацией обуславливает снижение мощности шумов квантования, за счет чего достигается выделение более слабого сигнала на их фоне [7, 10], так и большое количество отсчетов в выборке при выполнении ДПФ приводит к аналогичному эффекту за счет перераспределения их энергии в большем количестве

отсчетов спектральных составляющих. В этом случае спектральная плотность шумов квантования определяется как

$$N_0^2 = \frac{2n_0^2}{M}, \quad (3)$$

где $n_0^2 = \frac{1}{12} \left(\frac{A_0}{2^{n_{\text{ант}}}} \right)^2$ – средняя мощность шума квантования [7, 10]; M – общее количество отсчетов, по которым выполняется ДПФ; A_0 – амплитуда сигнала.

Из (3) следует, что при уменьшении $n_{\text{ант}}$ на некоторое количество бит w обеспечение постоянного значения N_0^2 может быть достигнуто за счет использования выборки из M отсчетов. То есть учитывая, что $M(2^{n_{\text{ант}}})^2 = 2^{n_{\text{ант}} + 0,5 \log_2(M)}$, использование выборки из M отсчетов при выполнении ДПФ эквивалентно увеличению разрядности АЦП на величину, определяемую выражением $w = 0,5 \log_2(M)$.

Тогда $n_{\text{ант}}$ можно определить с учетом полученной оценки d как

$$n_{\text{ант}} \geq \begin{cases} \log_2(d) - w, & \text{при } \log_2(d) > w; \\ 1, & \text{при } \log_2(d) \leq w. \end{cases} \quad (5)$$

Так, например, при частоте оцифровки $f_{\text{ант}} = 12$ МГц и времени синтезирования апертуры $T_c = 3$ с общее количество отсчетов $M = f_{\text{ант}} T_c = 36 \times 10^6$, что эквивалентно увеличению дополнительной разрядности АЦП w на 12 бит. Тогда из (5) следует, что для обеспечения оцифровки сигналов с ДД $d = 80$ дБ требуется $n_{\text{ант}} \geq \log_2(10^{80/20}) - 12 \geq 2$ бит.

Выбор разрядности АЦП, обеспечивающего заданный уровень искажений на РЛИ

Следствием процесса дискретизации сигнала является появление дополнительных ложных отметок на РЛИ, обусловленных конечной разрядностью АЦП (рис. 3 а, б). Однако при наличии шума в оцифровываемом сигнале даже при малой разрядности АЦП возможно значительное уменьшение уровня этих отметок (рис. 3 в, г) [6]. При этом уменьшение уровня отметок на РЛИ обусловлено не маскированием шумовой составляющей, а именно снижением их амплитуды.

Степень выраженности негативного эффекта дискретизации наглядно характеризует показатель $I_{\max} = I_{\text{н}}/I_{\text{л}}$, где $I_{\text{н}}$ и $I_{\text{л}}$ – значения амплитуды истинной отметки

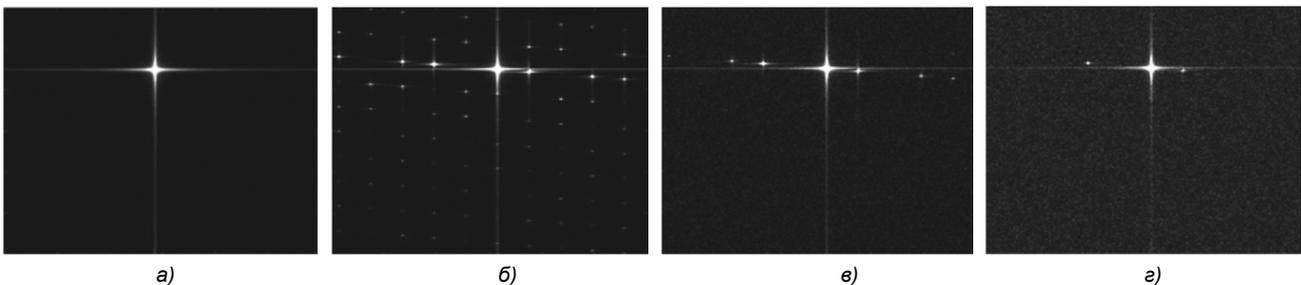
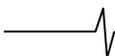


Рис.3. РЛИ точечной цели при различных значениях q :

а) неквантованное, без шума; б) $n_{\text{ант}} = 1$, $q = 40$ дБ; в) $n_{\text{ант}} = 1$, $q = 20$ дБ; г) $n_{\text{ант}} = 1$, $q = 10$ дБ



(основной гармоники) и максимальной ложной (паразитной гармоники), соответственно. Показатель I_{\max} является аналогом характеристики АЦП, известной как SFDR – Spurious-free dynamic range. В радиовидении аналогичный показатель носит название максимального уровня боковых лепестков [11].

Для определения влияния $n_{\text{ант}}$ и q на значение показателя I_{\max} проведено численное моделирование процесса формирования РЛИ по аддитивной смеси отсчетов траекторного сигнала точечной цели и шума, результаты которого иллюстрируют зависимости на рис. 4.

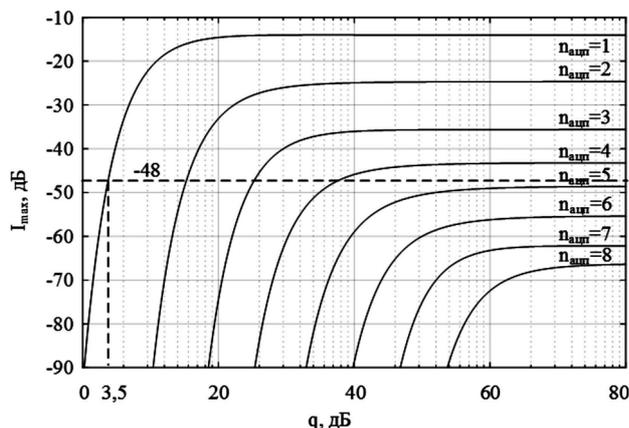


Рис.4. Зависимости уровня гармонических составляющих от q

Характер зависимостей на рис. 4 практически не меняется при изменении размера выборки M . Из их анализа также следует, что снижение относительного уровня эхо-сигнала точечной цели на входе АЦП до уровня $q < 6(2^{n_{\text{ант}}} - 1)$ и менее приводит к быстрому снижению уровня искажений I_{\max} , характер изменения которого становится близким к линейному (в логарифмическом масштабе) при $q < 3(2^{n_{\text{ант}}} - 1)$.

Из анализа зависимостей следует, что при $n_{\text{ант}} = 1$ и использовании для отображения сформированных РЛИ местности мониторов персональных компьютеров, ДД которых не превышает 48 дБ, появление ложных отметок будет наблюдаться при значениях $q > 3, 5$ дБ. В этом случае, при значениях q , не превышающих минус 15 дБ (рис. 2), возможно применение бинарного квантования сигнала без появления ложных отметок. Однако, например, при формировании РЛИ объектов при других условиях съемки и отображения требования к $n_{\text{ант}}$ могут быть выше.

Исходя из полученных результатов методика обоснования требований к АЦП в РСА НИ включает ряд последовательных действий: определение ДД сигналов d и максимального значения q_{\max} на входе АЦП для заданных условий съемки и характеристик РСА; вычисление на основе полученного значения d минимального значения $n_{\text{ант}}$ с использованием (5), обеспечивающего оцифровку сигналов в требуемом ДД; определение ми-

нимального значения $n_{\text{ант}}$, обеспечивающего заданный уровень искажений на РЛИ на основе полученного значения q_{\max} с помощью зависимостей, представленных на рис. 4; выбор разрядности АЦП, удовлетворяющей условиям оцифровки сигналов в требуемом ДД и заданного уровня искажений на РЛИ.

Заключение

Таким образом, приведенная методика позволяет обосновать требования к разрядности АЦП РСА НИ для конкретных условий съемки земной поверхности. Полученные результаты свидетельствуют о возможности существенного уменьшения разрядности АЦП (вплоть до бинарного квантования) практически без потери качества формируемых РЛИ. Кроме того, уменьшение объема регистрируемой информации в процессе работы РСА НИ снизит требования к вычислителю при синтезе РЛИ (в том числе на борту носителя) и к пропускной способности радиоканала при передаче зарегистрированного сигнала на наземный пункт управления в количестве раз, пропорциональное снижению разрядности АЦП.

Литература

1. Лихачев В.П., Рязанцев Л.Б., Чередников И.Ю. Применение беспилотных летательных аппаратов для ведения тактической радиолокационной разведки // Военная мысль. 2016. № 3. С. 24-28.
2. Sandia National Laboratories [Электронный ресурс] // URL: <http://www.sandia.gov> (дата обращения: 04.04.2017).
3. ImSAR LLC [Электронный ресурс] // URL: <http://www.imsar.com> (дата обращения: 04.04.2017).
4. Богомолов А.В., Купряшкин И.Ф., Лихачев В.П., Рязанцев Л.Б. Малогабаритная двухдиапазонная РСА для беспилотного авиационного комплекса. Труды XXIX Всероссийского симпозиума «Радиолокационное исследование природных сред» / под общ. ред. М.М. Пенькова. – СПб.: ВКА имени А.Ф. Можайского, 2015. Вып. 11. С. 237-242.
5. Рязанцев Л.Б. Обоснование конструкции антенной системы РЛС с синтезированной апертурой для БЛА малого класса // Антенны. 2016. № 5. С. 49-55.
6. Лайонс Р. Цифровая обработка сигналов. М.: ООО «Бином-Пресс», 2006 г.
7. Кестер У. Аналого-цифровое преобразование. М.: Техносфера, 2007.
8. J.C. Candy, G.C. Temes, «Oversampling Methods for A/D and D/A Conversion», pp. 1-29, IEEE Press, 1992.
9. Купряшкин И.Ф., Лихачев В.П. Космическая радиолокационная съемка земной поверхности в условиях помех: монография. Воронеж: Научная книга, 2014.
10. Improving ADC Resolution by Oversampling and Averaging // [Электронный ресурс] // URL: <http://www.silabs.com> (дата обращения: 04.04.2017).
11. Школьный Л.А. Радиолокационные системы воздушной разведки, дешифрирование радиолокационных изображений. М.: ВВИА им. Проф. Н.Е. Жуковского, 2008.

УДК 004.932

КОМБИНИРОВАННЫЙ МЕТОД СОВМЕЩЕНИЯ РАЗНОРОДНЫХ ИЗОБРАЖЕНИЙ В АВИАЦИОННЫХ СИСТЕМАХ ТЕХНИЧЕСКОГО ЗРЕНИЯ

Новиков А.И., к.т.н., доцент кафедры «Высшая математика» Рязанского государственного радиотехнического университета, e-mail: novikovanatoly@yandex.ru;

Логинов А.А., к.т.н., доцент кафедры «Электронные вычислительные машины» Рязанского государственного радиотехнического университета, e-mail: loginal@mail.ru;

Колчаев Д.А., магистрант кафедры «Системы автоматизированного проектирования» Рязанского государственного радиотехнического университета, e-mail: 3force3@mail.ru.

COMBINED METHOD OF HETEROGENEOUS IMAGES IMPOSITION IN AVIATION COMPUTER VISION SYSTEM

Novikov A.I., Loginov A.A., Kolchaev D.A.

A combined algorithm for imposition of a real image with an image synthesized from a digital terrain map, using a prediction of navigational data is presented. The synthesized image in it is formed on the basis of the predicted values of navigation parameters. The prediction of the navigation parameter values is carried out for each frame of the video sequence according to the linear models in the finite memory scheme.

Key words: real and synthesized images, correlation-extreme algorithm, linear prediction, combination of images, quality of combining.

Ключевые слова: реальное и синтезированное изображения, корреляционно-экстремальный алгоритм, линейный прогноз, совмещение изображений, качество совмещения.

Введение

Необходимость решения задачи совмещения изображений и использования полученных результатов возникает во многих областях науки и техники. Создание систем технического зрения для летательных аппаратов (ЛА) обусловлено, в первую очередь, требованиями безопасности полетов при плохой видимости (сложные метеословия, ночной полет). Совмещение реальных изображений и изображений, синтезированных по цифровой карте местности, позволяет создавать и отображать на экране комбинированное изображение и, как следствие, обеспечивать управление ЛА в сложных условиях.

Для совмещения реального и синтезированного изображений корреляционно-экстремальными методами приходится генерировать некоторое множество ракурсов виртуальной карты местности. Под ракурсом понимается конкретный набор значений шестимерного арифметического вектора $\mathbf{v} = (x, y, h, \psi, \theta, \gamma)$ – навигационных параметров позиционирования ЛА. Параметры позиционирования ЛА являются исходными данными для формирования синтезированного изображения по цифровой карте местности.

Требования к точности совмещения изображений и быстродействию вычислительных алгоритмов, обеспечивающих решение задачи совмещения изображений и других вспомогательных задач, определяются документом [1]. В соответствии с этим документом ошибка совмещения реального и синтезированного изображений

Приведен комбинированный метод совмещения реального изображения и изображения, синтезированного по цифровой карте местности. Синтезированное изображение в нем формируется на основе прогнозируемых значений навигационных параметров, скорректированных на величины поправок. Поправки вычисляются однократно с помощью корреляционно-экстремального алгоритма при выходе прогнозных значений за пределы заданной погрешности. Прогноз значений навигационных параметров осуществляется для каждого кадра видеопоследовательности по линейным моделям в схеме с конечной памятью.

не должна превышать 5 миллирадиан по горизонтали и по вертикали. Корреляционно-экстремальные алгоритмы при корректном выборе размеров сетки и расположения ее узлов в шестимерном пространстве параметров обеспечивают хорошее качество совмещения, но требуют недопустимо высоких затрат машинного времени на их реализацию [2]. Еще одно преимущество корреляционно-экстремальных методов совмещения заключается в том, что одновременно с решением задачи совмещения изображений решается навигационная задача. Точка экстремума целевой функции – это оценка $\hat{\mathbf{v}}_{opt}$ вектора $\mathbf{v}_{ист} = (x_{ист}, y_{ист}, h_{ист}, \psi_{ист}, \theta_{ист}, \gamma_{ист})$, истинных значений навигационных параметров.

Желание использовать преимущества корреляционно-экстремальных методов совмещения изображений, заключающееся в гарантированно высоком качестве совмещения, с одной стороны, и недопустимо высокие затраты времени на реализацию алгоритмов этой группы, с другой стороны, вынуждают искать компромиссные решения. Один вариант решения этой проблемы предложен в работе [3]. В основу подхода, предлагаемого авторами этой работы, положены три предложения:



- отказ от перебора углов курса, крена и тангажа;
- использование пирамиды изображений Гаусса;
- использование данных о динамике перемещения летательного аппарата в последовательности кадров.

В работе [4] было показано, что если ошибки сенсоров углов тангажа, крена и курса не превосходят $0,1^\circ$, то рассогласование между парой соответствующих точек, обусловленное предельными значениями только этих ошибок, не превосходит 5 мрад. Это соответствует приблизительно 5 пикселям для камеры 704×576 пикселей и с углами обзора $40^\circ \times 30^\circ$. Поэтому отказ от перебора углов курса, крена и тангажа, с одновременным расширением поля зрения виртуальной камеры как по вертикали, так и по горизонтали, вполне обоснован.

Как следует из приведенных в [3] результатов экспериментальной проверки алгоритма совмещения с использованием пирамиды изображений совместно с расширением поля зрения, его быстрое действие увеличивается более чем в 30 раз по сравнению с полным перебором навигационных параметров. Значит, поиск способов комбинирования различных методов совмещения разнородных изображений является оправданным.

Схема алгоритма совмещения

Предлагаемый алгоритм совмещения реального видеоизображения и отвечающего ему виртуального изображения (ВИ), синтезированного по цифровой карте местности состоит из четырех основных этапов:

- выделение контуров на реальном изображении (РИ);
- улучшение контурных препаратов на РИ и ВИ;
- подготовительный режим к совмещению изображений;
- основной режим совмещения изображений.

Первые два этапа – выделение контуров градиентными методами [5, 6] и улучшение контурных препаратов с учетом специфики решаемых задач – рассмотрены подробно в работах [7, 8]. Выделение контуров производилось с помощью алгоритма Canny [9], а улучшение контурных изображений – с помощью базовых морфологических операций дилатации и эрозии [6].

На подготовительном этапе происходит накопление кадров изображений и отвечающих им значений навигационных параметров в количестве, равном заданной памяти прогнозной модели (5-7 кадров). Пусть $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ – первые k значений шестимерного вектора навигационных параметров, полученные от соответствующих сенсоров. Для каждого из 6 параметров в составе вектора $\mathbf{v} = (x, y, h, \psi, \theta, \gamma)$ строится линейная модель по соответствующим k значениям этого параметра. В точке k одновременно формируется синтезированное изображение по ЦКМ, выполняется совмещение этого изображения с отвечающим ему реальным изображением с помощью корреляционно-экстремального алгоритма и находится оптимальная оценка вектора навигационных параметров $\hat{\mathbf{v}}_{opt}^{(k)}$.

На основном этапе для каждого i -го кадра ($i > k$) в скользящем режиме производится оценка параметров

a_0, a_1 прогнозных моделей по k точкам, предшествующим этому кадру, выполняется прогноз каждого параметра на одну точку вперед с корректировкой прогнозного значения на величину уклонения оптимальной оценки $\hat{u}_k^{КЭ}$ от фактического (см. выражение (2), верхний и нижний индексы?) значения данного параметра в этой точке:

$$\hat{u}_i = a_0^{(i)} + a_1^{(i)}(k+1) + \Delta u_k, \quad (1)$$

где

$$\Delta u_k = \hat{u}_k^{КЭ} - \hat{u}_k^{факт}, \quad (2)$$

Здесь u – один из 6 навигационных параметров.

По полученной прогнозной оценке $\hat{\mathbf{v}}_{прогн}^{(k+1)}$ вектора навигационных параметров по ЦКМ формируется синтезированное изображение и осуществляется совмещение (наложение) реального и виртуального изображений.

Вычислительный алгоритм

На подготовительном этапе для каждого из 6 навигационных параметров после получения первых k кадров в соответствии с методом наименьших квадратов вычисляются оценки $\hat{a}_0^{(k)}, \hat{a}_1^{(k)}$ прогнозных моделей (1):

$$\hat{a}_1^{(k)} = \frac{12}{k(k_2-1)} \left(B - \frac{k+1}{2} A \right) \quad (3)$$

$$\hat{a}_0^{(k)} = \frac{1}{k} \left(A - \frac{k(k+1)}{2} B \right),$$

$$\text{где } A = \sum_{j=1}^k u_j, \quad B = \sum_{j=1}^k j \cdot u_j.$$

Для конкретного значения параметра k – числа точек, по которым производится оценка коэффициентов линейной модели, – формулы (3) максимально упрощаются. Например, при $k=5$ будем иметь

$$\hat{a}_1^{(5)} = \frac{1}{10}(B-3A), \quad \hat{a}_0^{(5)} = \frac{1}{5}(A-15B), \quad (4)$$

Первое прогнозируемое значение параметра u в этом случае находится для 6-го кадра по формуле $\hat{u}_6 = \hat{a}_0^{(5)} + 6\hat{a}_1^{(5)} + \Delta u_5$, где $\Delta u_5 = \hat{u}_5^{КЭ} - \hat{u}_5^{факт}$. Для следующего 6-го кадра по формулам (4) находятся новые оценки параметров $\hat{a}_0^{(6)}, \hat{a}_1^{(6)}$ прогнозных моделей по фактическим значениям u_2, u_3, \dots, u_6 соответствующего навигационного параметра и осуществляется прогноз его значения к 7-у кадру: $\hat{u}_7 = \hat{a}_0^{(6)} + 6\hat{a}_1^{(6)} + \Delta u_5$ и т.д.

Оценка $\hat{u}_5^{КЭ}$ находится с помощью оптимизированного варианта корреляционно-экстремального алгоритма в схеме с расширенным ракурсом. Введение расширенного ракурса вместе с другими модификациями классического варианта корреляционно-экстремального алгоритма позволяет уменьшить вычислительную сложность алгоритма. Поле зрения расширяется, как показано на рис. 1 [10], по вертикали и по горизонтали с учетом интервалов неопределенности углов тангажа и курса соответственно. Такой подход позволяет отказаться от перебора углов курса и тангажа при поиске оптимального соответствия между РИ и ВИ.

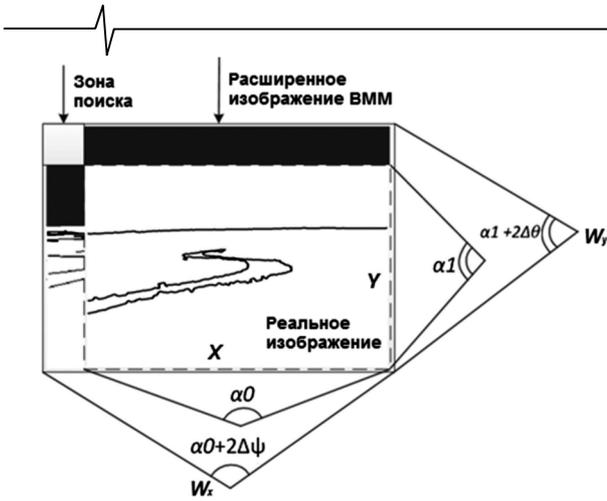


Рис. 1. Размеры расширенного ВИ

Алгоритм имеет два этапа. На первом этапе осуществляется сканирования координат по широте, долготе и высоте. На втором этапе обход координат осуществляется с уменьшенным в два раза шагом вокруг точки, найденной на первом этапе. На каждом этапе для уменьшения вычислительных затрат уточнение точки глобального экстремума осуществляется за счет последовательного трёхуровневого уточнения смещении левого верхнего угла обработанного реального изображения относительно левого верхнего угла расширенного виртуального изображения с учетом результатов предыдущего уровня. Значение критериальной функции вычисляется в каждом узле сетки.

На рис. 2 приведена схема корреляционно-экстремального алгоритма совмещения

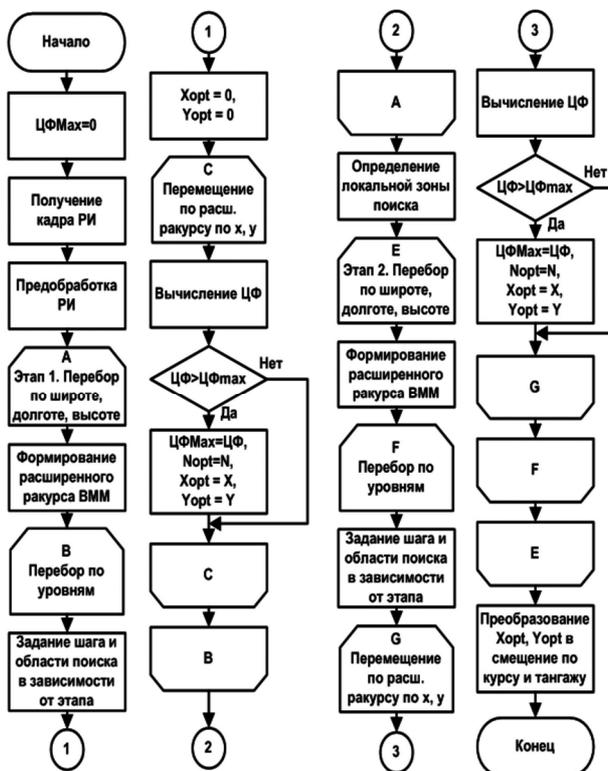


Рис. 2. Схема корреляционно-экстремального алгоритма
Входные данные алгоритма:

- реальное изображение в формате BMP 24 бит, полученное из AVI-файла;
- навигационные данные (информация о текущем мес-

тоположении летательного аппарата в пространстве);
– значения каждого из трех параметров (широты, долготы и высоты) в узлах сетки сканирования на 1-м этапе: $\{-100, -50, 0, 50, 100\}$, либо $\{-100, 0, 100\}$, т.е. $5^3 = 125$ точек в первом и $3^3 = 27$ – во втором случае.

В качестве целевой функции для оценки степени совпадения реального и виртуального изображений использовалась парная критериальная функция Рао [11]

$$K = \frac{a}{a + e + b + c}.$$

В этой формуле использованы обозначения:

$a = F_{11}(I_{PI}, I_{ВИ})$ – число совпадений черных пикселей на сравниваемых реальном (I_{PI}) и виртуальном ($I_{ВИ}$) изображениях;

$b = F_{10}(I_{PI}, I_{ВИ})$ – число совмещенных пикселей, имеющих различную яркость на сравниваемых изображениях: черную на реальном (I_{PI}) и белую на виртуальном ($I_{ВИ}$);

$c = F_{01}(I_{PI}, I_{ВИ})$ – число совмещенных пикселей, имеющих различную яркость на сравниваемых изображениях: белую на реальном (I_{PI}) и черную на виртуальном ($I_{ВИ}$);

$e = F_{00}(I_{PI}, I_{ВИ})$ – число совпадений белых пикселей на сравниваемых реальном (PI) и виртуальном (ВИ) изображениях.

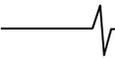
Рассмотрим схему работы корреляционного алгоритма.

1. На первом этапе последовательно выбираются узлы заданной сетки.

2. Вычисляются значения координат точки с наибольшим значением критериальной функции из окрестности возможного глобального экстремума. Шаг перемещения РИ относительно расширенного виртуального изображения задается равным 8-10 пикселям ($h_1 = 8 - 10$), как для перемещения по горизонтали, так и по вертикали. После того, как найдено значение максимума критериальной функции на этой сетке, запоминаются координаты ($X_{отн}$ и $Y_{отн}$) РИ относительно расширенного виртуального изображения, а так же значение максимума критериальной функции для этих координат.

3. На втором уровне формируется сетка с шагом $h_2 = h_1 / 2$, т.е. 4 или 5 пикселей в окрестности точки экстремума критериальной функции. Границы области поиска выбираются по определенной схеме в зависимости от соотношения величин $X_{отн}, Y_{отн}$ и шага h_1 смещений на предыдущем этапе. Для каждого смещения по координатам x, y вычисляются значения критериальной функции. Затем осуществляется поиск максимума критериальной функции в построенной области с вычислением новых значений $X_{отн}, Y_{отн}$ координат РИ относительно расширенного виртуального изображения.

4. На заключительном, третьем уровне шаг смещений по каждой координате выбирается равным 1 пикселю ($h_3 = 1$), а зона поиска формируется с учетом шага h_2 смещений предыдущего уровня.



В основном режиме совмещения изображений осуществляется поиск оптимальных значений параметров $\hat{a}_0^{(i)}, \hat{a}_1^{(i)}$ линейных моделей для каждого из 6 навигационных параметров. Оптимальные оценки параметров для каждого i -го кадра находятся по фактическим значениям каждого из этих параметров на k предыдущих кадрах. В результате формулы (3) видоизменяются следующим образом

$$\begin{cases} \hat{a}_1^{(i)} = \frac{12}{k(k_2-1)} \left(B_i - \frac{k+1}{2} A_i \right) \\ \hat{a}_0^{(i)} = \frac{1}{k} \left(A_i - \frac{k(k+1)}{2} B_i \right), \end{cases}$$

где $A_i = \sum_{j=1}^k u_{i+j-1}$, $B_i = \sum_{j=1}^k j \cdot u_{i+j-1}$, $i > k$.

Для каждого навигационного параметра осуществляется прогноз на один шаг (кадр) вперед по формуле (1) с корректировкой прогнозного значения на величину поправки $\Delta u_k = \hat{u}_k^{КЭ} - \hat{u}_k^{факт}$. В результате получаем вектор \hat{v}_{i+1} прогнозных значений навигационных параметров для следующего $i+1$ -го кадра. На основании этого вектора формируется ВИ и осуществляется наложение этого изображения на реальное изображение.

Прогнозное значение каждого параметра $k+1$ -му кадру сравнивается с его фактическим значением в этой точке. Если модуль уклонения $|\hat{u}_{i+1} - u_{i+1}^{факт}|$ окажется больше заданного порога ε , то производится запуск корреляционно-экстремального алгоритма и с его помощью формируется новое значение поправки $\Delta u_{i+k} = \hat{u}_{i+k}^{КЭ} - \hat{u}_{i+k}^{факт}$.

Экспериментальная часть и обсуждение

Апробация представленного алгоритма совмещения разнородных изображений проводилась на реальных видеоизображениях, синхронизированных с навигационными параметрами, полученными в ходе реальных полетов.

На рис. 3 приведены графики изменения параметров высоты и тангажа, полученные по навигационным данным в одном из таких полетов. По горизонтальной оси на рисунке отложены номера кадров, а по вертикальной оси – значения высоты в метрах и здесь же – угла тангажа в градусах. Четыре значения угла тангажа: $-11^\circ, 0^\circ, 2,5^\circ, 10^\circ$ – выделены на рисунке горизонтальными линиями с соответствующими надписями с правой стороны рисунка.

Можно отметить достаточно гладкий характер изменения высоты и значимые флуктуации угла тангажа. Кроме того, угол тангажа изменяется достаточно резко, но при этом согласованно с изменением высоты (5 раз за время полета в рассматриваемом фрагменте из 200000 кадров).

Количество прогнозируемых кадров до срыва зависит как от навигационных данных, так и от результата корреляционного совмещения. В среднем, это количество на рассматриваемых сюжетах равно 58.

На рис. 4 приведена иллюстрация совмещения РИ и ВИ в каждом из двух режимов работы комбинированного корреляционно-экстремального алгоритма, дополненного блоком прогнозирования навигационных параметров. Значение параметра k в блоке прогнозирования РИ навигационных параметров было взято равным 5. На рис. 4 а приведен результат простого наложения ВИ, сгенерированного по фактическим значениям навигационных параметров для 5-го кадра, содержащих ошибки позиционирования. Этому кадру соответствуют по 5 предшествующих значений каждого из 6 навигационных параметров, необходимых для построения прогнозных моделей. В этой точке выполняется одновременно и корреляционно-экстремальное совмещение ВИ и РИ. Результат такого совмещения ВИ с РИ приведен на рис. 4 б. На визуальном уровне качество совмещения можно оценить как достаточно высокое.

В основном режиме, начиная с 6-го кадра, совмещение ВИ с РИ производится только на основании прогнозных значений навигационных параметров. Пример такого совмещения для 51-го кадра приведен на рис. 4 в и 4 г. На рис. 4 в представлен результат простого наложения ВИ на РИ. Можно отметить существенное расхождение реального и виртуального изображений. На рис. 4 г приведен результат совмещения ВИ и РИ на основании спрогнозированных значений навигационных данных. Качество совмещения изображений можно оценить как вполне удовлетворительное.

Отметим, что преобразование найденного смещения изображения ВММ относительно расширенного изображения в изменения углов курса и тангажа производится по формулам:

$$\begin{cases} dKurs = \frac{(X_1 - X_0) cm_x}{W}, \\ dTang = \frac{(Y_1 - Y_0) cm_y}{H}, \end{cases}$$

где W, H – ширина и высота реального изображения, cm_x, cm_y – углы зрения камеры, X_0, Y_0 – координаты левого верхнего угла изображения ВММ относительно расширенного изображения при совмещении центров изображений, X_1, Y_1 – координаты левого верхнего угла изображения ВММ после совмещения.

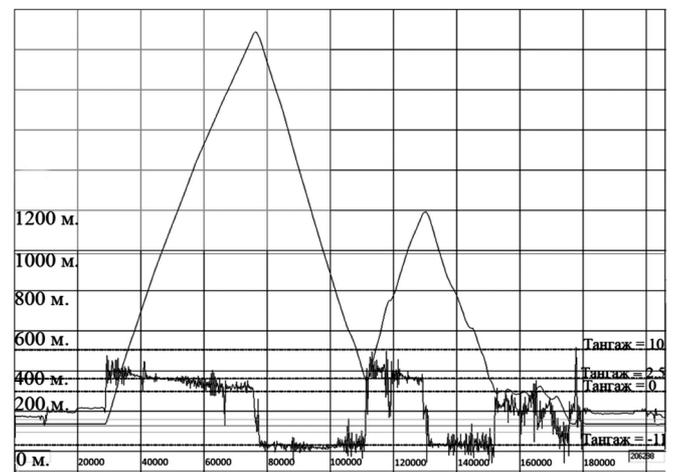


Рис. 3. Динамика изменения параметров высоты и тангажа

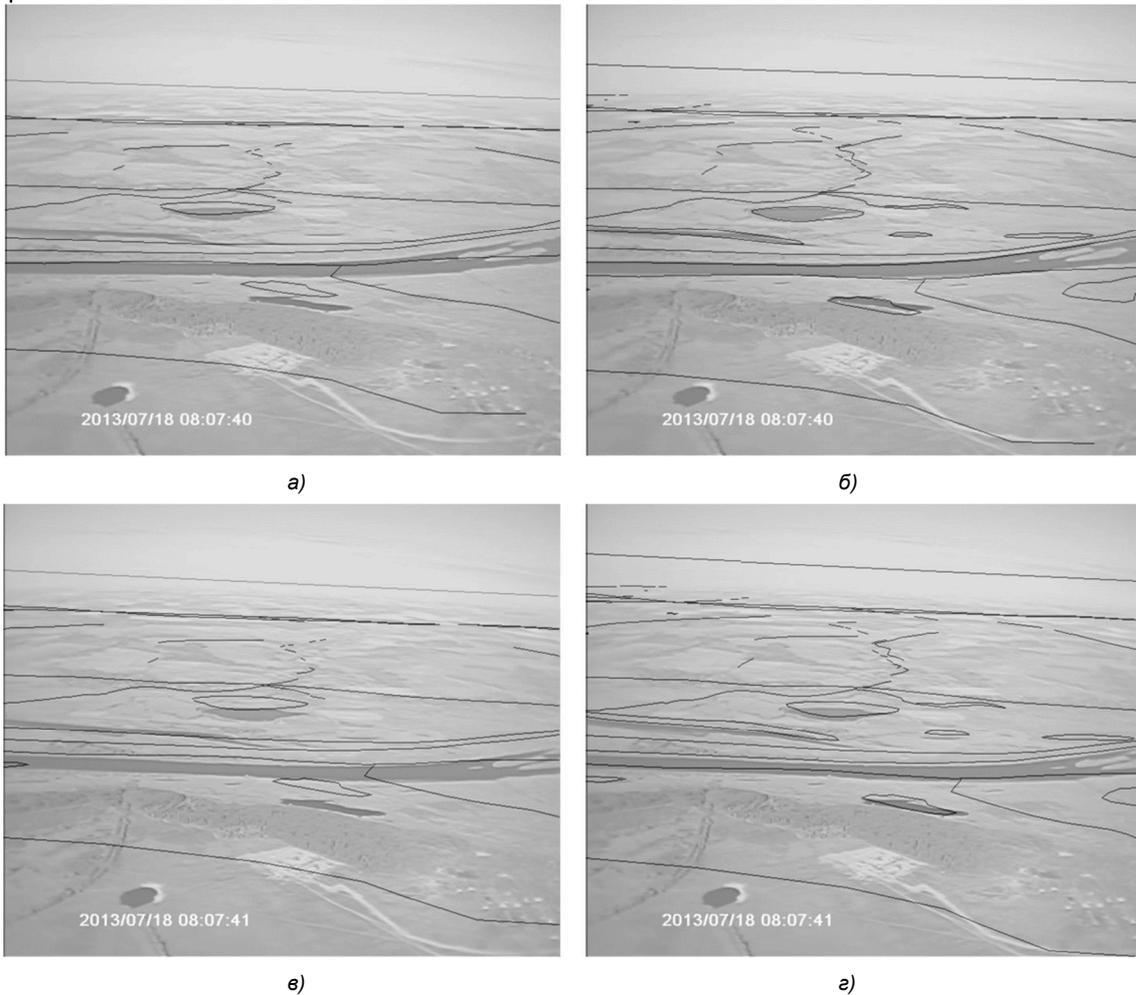


Рис. 4. Совмещение изображений: а – простое наложение ВИ на РИ на 5-м кадре;
 б – результат совмещения ВИ и РИ корреляционным алгоритмом на 5-м кадре;
 в – простое наложение ВИ на РИ на 51-м кадре;

г – результат совмещения ВИ и РИ на 51-м кадре на основании спрогнозированных значений навигационных данных

Вычисления проводились на компьютере с процессором Intel Core i3-550, ОЗУ 2Гб, по оригинальной программе, написанной на языке СИ++. Затраты машинного времени на выполнение совмещения корреляционно-экстремальным алгоритмом для одного кадра в среднем равны 4,8 сек. В основном режиме работы алгоритма расчет параметров прогнозных моделей для каждого кадра осуществляется практически мгновенно, но при этом 0,15 сек тратится на формирование ВИ по ЦКМ и наложение ВИ на РИ.

Апробация алгоритма выполнена на различных видах подстилающей поверхности по видеопоследовательностям длиной в 500-1000 кадров. Ожидаемо, наилучшие результаты совмещения получены на сюжетах, содержащих объекты постоянного присутствия на поверхности Земли: водные объекты, дороги. Частота перехода из основного, быстрого режима работы алгоритма совмещения изображений – прогнозирования значений навигационных параметров, – в режим корректировки поправок к значениям навигационных параметров корреляционно-экстремальным алгоритмом в значительной мере зависит от частоты и величин изменений компонент вектора навигационных параметров.

На рис. 5 представлен фрагмент графика показателя качества совмещения РИ и ВИ по предложенному алгоритму.

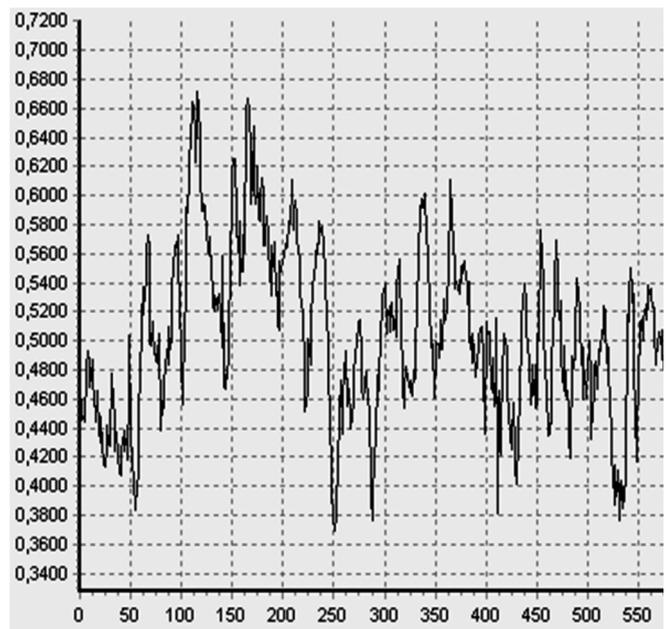


Рис. 5. Зависимость качества совмещения предложенного алгоритма от номера кадра

На графике по горизонтали отложены номера кадров, а по вертикали значения показателя α качества совмещения. Значение интегрального показателя качества

α для каждого кадра находилось как отношение числа совпавших черных пикселей при наложении ВИ на контурный образ реального изображения к общему числу черных пикселей на РИ [8]. Данный показатель является нормированным и его значения изменяются в пределах от 0 до 1. На реальных изображениях даже при хорошем (по мнению экспертов) качестве совмещения его значение находится на уровне 0,45...0,65. Среднее значение показателя качества в эксперименте, соответствующем рис. 5, оказалось равным 0,512. В других сериях сюжетов оно находилось в пределах 0,48...0,53. Минимальные значения показателя качества на рис. 5 в основном соответствуют моментам переключения алгоритма из режима совмещения ВИ с РИ на основании прогнозных значений навигационных параметров в режим корреляционно-экстремального совмещения.

Было выполнено сравнение предлагаемого алгоритма с алгоритмами «Трехуровневая пирамида изображений» и «Анализ последовательности изображений» [3]. Вычисления проводились в одинаковых условиях на программном стенде и на одинаковой видеопоследовательности изображений, состоящей из 799 кадров. Оценивались качество совмещения для каждого кадра и время работы алгоритма для одного кадра. Результаты этого сравнения представлены в таблице.

Таблица. Результаты сравнение алгоритмов совмещения

Название алгоритма	Качество совмещения	Время выполнения, с.
1. Простое наложение РИ на ВИ	0,311	0
2. Трехуровневая пирамида изображений	0,357	14,3 (44,6)
3. Анализ последовательности изображений	0,326	0,1 (7,6)
4. Корреляционное совмещение и экстраполяция	0,512	0,15 (4,8)

Для алгоритмов 2 и 3 в третьем столбце таблицы в скобках указано время работы алгоритмов, приведенное в статье [3], а без скобок – время работы, зафиксированное в ходе исследований алгоритмов на программном стенде. Для алгоритма 4 в скобках указано время работы предлагаемого алгоритма в режиме корреляционно-экстремального совмещения, а без скобок – в основном режиме прогнозирования.

Можно отметить, что предлагаемый алгоритм по показателю качества совмещения и времени, затрачиваемому на обработку одного кадра, не уступает в целом сравниваемым алгоритмам.

Заключение

Представлен алгоритм, основанный на комбинировании двух режимов совмещения: корреляционно-экстремального и совмещения по экстраполированным значениям навигационных параметров, скорректированных на величины соответствующих поправок. Выполнены сравнительные испытания предложенного ал-

горитма с другими известными алгоритмами на реальных видеопоследовательностях.

Сочетание преимуществ корреляционно-экстремального алгоритма при совмещении разнородных изображений с высоким быстродействием алгоритма совмещения, основанного на использовании прогнозных значений навигационных параметров, позволяет реализовать его в бортовом вычислительном комплексе.

Литература

1. Руководство Р-315 «По минимальным стандартам характеристик авиационных систем (MASPS) для систем улучшенного видения, систем искусственного видения, комбинированных систем искусственного видения и бортовых систем увеличения дальности видения» Авиационного Регистра Межгосударственного Авиационного Комитета (АР МАК). М. 2008. 86 с.
2. Баклицкий В.К. Корреляционно-экстремальные методы навигации и наведения. Тверь: Книжный клуб, 2009. 216 с.
3. Бабаян П.В., Ершов М.Д.. Алгоритмы устранения рассогласования разнородных изображений в бортовой системе видения // Вестник РГРТУ, № 54, Часть 2, Рязань, 2015. с. 15-20.
4. Герман Е.В., Муратов Е.Р., Новиков А.И. Математическая модель формирования зоны неопределенности в задаче совмещения изображений // Вестник РГРТУ, № 4, вып. 46, часть 2. Рязань, 2013. с. 10-16.
5. Форсайт Д.А., Понс Ж. Компьютерное зрение. Современный подход. М.: Вильямс, 2004. 928 с.
6. Гонзалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2005. – 1072 с.
7. Ефимов А.И., Новиков А.И. Алгоритмы совмещения изображений на основе преобразования в комплексной плоскости// Тезисы научно-технической конференции «Техническое зрение в системах управления», Москва, 14-16 марта 2017. с. 34-36.
8. Anatoly I. Novikov, Victoria A. Sablina, Aleksey I. Efimov, Michael B. Nikiforov. Contour Analysis in the tasks of real and virtual images superimposition// Journal Coupled Systems and Multiscale Dynamics, vol 4(4), 2016 pp. 251-259. (Doi: 10.1166/jcsmd. 2016.1112 J. Coupled Syst. Multiscale Dyn. vol 4(4)/2330-152X/2016/251/009).
9. J. Canny; A computational approach to edge detection; Proc. Of IEEE Transactions on Pattern and Machine Intelligence PAMI-8, 679 (1986).
10. S. Elesina, O. Lomteva. Increase of image combination performance in combined vision systems using genetic algorithm. Proceedings of the 3rd Mediterranean Conference on Embedded Computing (MECO). Montenegro, Budva, 2014, pp. 158-161.
11. Елесина С.И., Ефимов А.И. Отбор критериальных функций для систем улучшенного и комбинированного видения // Известия ТулГУ, технические науки, вып. 9, ч. 1. 2013. – С. 229-236.

УДК 004.8; 621.396.96

ОЦЕНКА НАПРАВЛЕНИЯ ПРИХОДА СИГНАЛА С ИСПОЛЬЗОВАНИЕМ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

Шевгунов Т.Я., к.т.н., доцент кафедры «Теоретическая радиотехника» Московского авиационного института (национального исследовательского университета) (МАИ), доцент кафедры бизнес-аналитики НИУ ВШЭ, e-mail: shevgunov@gmail.com;

Ефимов Е.Н., инженер НИО-4 Московского авиационного института (национального исследовательского университета) (МАИ), ведущий инженер ФНЦ «НИИ системных исследований» РАН, e-mail: omegatype@gmail.com;

Филимонова Д.В., студентка факультета «Радиоэлектроника летательных аппаратов» Московского авиационного института (национального исследовательского университета) (МАИ), e-mail: daria-fili@rambler.ru;

Воскресенский Д.И., д.т.н., профессор, заведующий кафедрой «Радиофизика, антенны и микроволновая техника» Московского авиационного института (национального исследовательского университета) (МАИ), e-mail: voskr@mai.ru.

ESTIMATION OF DIRECTION-OF-ARRIVAL BY MEANS OF MAXIMUM LIKELIHOOD ARTIFICIAL NEURAL NETWORKS

Shevgunov T.Ya., Efimov E.N., Filimonova D.V., Voskresenskiy D.I.

Feedforward artificial neural network based approach to implementing maximum likelihood direction-of-arrival estimator for passive radar systems is proposed. Comparison of the proposed estimator, optimal numerical solution and referenced Cramer-Rao lower bound was carried out; absence of significant dependency of the estimator accuracy in wide angle range is shown.

Key words: artificial neural networks, maximum likelihood, parametric estimation, direction-of-arrival method, circular antenna array.

Ключевые слова: искусственные нейронные сети, метод максимального правдоподобия, параметрическое оценивание, угломерный метод, кольцевые антенные решетки.

Введение

В настоящее время среди ученых и исследователей, занимающихся разработкой и совершенствованием алгоритмов цифровой обработки сигналов, предназначенных для исполнения в программных комплексах радиотехнических устройств, обозначились две тенденции: создание высоко оптимизированных алгоритмов, предназначенных для специализированных вычислителей, и построение адаптивных алгоритмов на основе универсальных вычислительных структур. В качестве последних наиболее популярными на протяжении нескольких десятилетий остаются системы, построенные на основе теории и методов искусственных нейронных сетей (ИНС).

Возможности ИНС как платформы для реализации универсальных алгоритмов открывают следующую стратегию их применения. Наиболее затратные и плохо прогнозируемые по объёму вычислений этапы обучения ИНС производится на высокопроизводительной рабочей станции, а в тех случаях, когда производительности последней недостаточно, для обучения ИНС может быть задействован вычислительный кластер. Затем на основе матриц внутренних параметров обученная ИНС воспроизводится на других платформах, входящих в состав вычислительных средств технических комплексов. Такими платформами могут выступать персональ-

Предложен способ формирования оценки пеленга в угломерной пассивной радиолокационной системе с использованием искусственных нейронных сетей прямого распространения сигнала, приближающих оценки по методу максимального правдоподобия. Проведено сравнение точности предложенного способа с оптимальным численным решением и границей Крамера-Рао, установлена независимость точности оценки пеленга в широком диапазоне углов.

ные компьютеры средней производительности, бортовые ЭВМ, критичные к продолжительной вычислительной нагрузке, и специализированные параллельные вычислители, предназначенные для аппаратной реализации ИНС.

В настоящей работе рассматривается задача создания алгоритма для пассивной радиолокационной системы угломерного типа, формирующей оценку пеленга по методу максимального правдоподобия (МП оценка) на основе ИНС прямого распространения сигнала. В научной литературе данный подход освещен слабо: в ведущих отечественной [1] и зарубежной [2] монографиях ИНС данного типа не описаны. В работе [2] была предложена идея создания ИНС, реализующих оценивание параметров по методу максимального правдоподобия, а не аппроксимацию по методу минимизации среднеквадратической ошибки; был предложен алгоритм обучения сети, реализующей только бинарную классификацию, формализованный в дальнейшем в работе [3]. Более строгое математическое обоснование процедуры синтеза и обучения ИНС, предназначенной для осуществления параметрического МП оценивания при известном типе распределения наблюдаемой величины, представ-

лено в работе [4]; там же предложены примеры ИНС, реализующих оценки параметров нормального, логнормального и Рэлеевского распределений. Ранее авторами настоящей работы был предложен подход к синтезу ИНС произвольной топологии на основе универсальных адаптивных элементов [6]. В дальнейших исследованиях [7] предложенная методология была обобщена и стала пригодной для организации процедур обучения ИНС не только на основе минимизации среднеквадратической ошибки, но и других критериев.

Несмотря на существующий скептицизм по отношению к ИНС среди исследователей, решение задачи оценки местоположения посредством применения алгоритмов ИНС в радиолокационных системах (РЛС) в настоящее время начинает набирать популярность в профессиональном сообществе [8]. Среди работ отечественных исследователей необходимо отметить работу [9], авторы которой провели поиск оптимальной конфигурации антенной решетки для оценки направления прихода сигнала с использованием классической ИНС типа «многослойный перцептрон».

Настоящая работа подготовлена по материалам доклада [10] и результатов его обсуждения на 19-й Международной конференции «Цифровая обработка сигналов и её применение».

Формирование МП оценки

Пусть требуется оценить значения элементов вектора параметров θ на основе наблюдений случайного вектора ξ . Его плотность вероятности $p_{\xi}(x; \theta)$ определяется в соответствии с некоторой априорно известной вероятностной моделью, связывающей вектор параметров θ с вектором наблюдений ξ . Оценка вектора $\theta_{ML}(x)$ по методу максимального правдоподобия представляет собой значение вектора θ , при котором достигается максимум функции $p_{\xi}(x; \theta)$ по параметру θ при наблюдаемых в конкретной реализации значениях x случайного вектора ξ :

$$\theta_{ML}(x) = \arg \max_{\theta} p_{\xi}(x; \theta). \quad (1)$$

В практическом случае задача оценки по методу максимального правдоподобия состоит в отыскании максимума целевой функции $T(x; \theta)$, которой может являться как сама плотность вероятности $p_{\xi}(x; \theta)$, так и функция, полученная её монотонной трансформацией, например, логарифмированием. Часто оказывается возможным пойти дальше: разложить выражение для плотности вероятности или её логарифма в сумму или произведение и затем исключить из рассмотрения слагаемые или сомножители, независимые от оцениваемого вектора параметров [5].

В некоторых редких, но важных случаях, например, таких как нормальное или равновероятное распределение, удаётся получить МП оценку в явной форме, представляющей собой функцию от произвольных значений вектора наблюдений x . В большинстве практически интересных случаев поиск максимума функции правдоподобия в пространстве вектора θ будет приводить к составлению системы нелинейных уравнений относительно его элементов. Поскольку в общем случае получить

решение этой системы в форме аналитической функции от x не удаётся, то его ищут с использованием численных методов поиска экстремума функции – безусловного или условного, если на вектор параметров наложены дополнительные ограничения.

При проведении серии наблюдений $\{x_n\}$ случайного вектора ξ возникает практическая потребность вычислить для каждого наблюдения оценку вектора θ_n , что потребует при использовании метода МП многократного проведения итеративной процедуры, что станет препятствием при реализации МП оценщика в системах реального времени. Существенную проблему создаёт тот факт, что число итераций численного алгоритма, используемого для поиска экстремума целевой функции, заранее неизвестно – оно зависит от истинного значения вектора θ , наблюдаемой реализации, выбранного начального приближения и параметров самого оптимизационного алгоритма. В дополнение к этому, каждое вычисление оценки посредством численной оптимизации целевой функции требует мониторинга: проверки условий сходимости итеративного процесса и выбора момента остановки алгоритма, а также контроля времени выполнения и потребляемых вычислительных ресурсов.

Рассмотрим подход к формированию МП оценки с ИНС, обладающей архитектурой прямого прохождения сигнала, вполне характеризуемой набором внутренних параметров w . На вход сети подаётся вектор наблюдений x , а на своём выходе сеть формирует вектор оценок θ_{NN} . Такая ИНС сможет заменить алгоритм минимизации целевой функции $T(x; \theta)$, если решения θ_{NN} , предлагаемые обученной сетью, будут хорошим приближением оценок максимального правдоподобия θ_{ML} , получаемых непосредственно с использованием алгоритма минимизации.

Так при использовании квадратичной нормы при вычислении ошибки скалярную меру близости $\varepsilon(x)$ двух оценок, вычисленных по наблюдению x , можно выразить как:

$$\varepsilon(x) = (\theta_{ML}(x) - \theta_{NN}(x))^H W (\theta_{ML}(x) - \theta_{NN}(x)), \quad (2)$$

где θ_{NN} и θ_{ML} – вектор-столбцы, «H» обозначает эрмитово сопряжение, а через W обозначена весовая матрица. Так, в случае выбора в качестве неё единичной матрицы $\varepsilon(x)$ будет измерять Евклидово расстояние между векторами, а в случае произвольной положительно определенной – расстояние Махаланобиса [12]. Выбор в качестве W диагональной матрицы с положительными элементами позволит учесть относительный вклад каждого компонента вектора оцениваемых параметров в общую ошибку. С практической точки зрения этот случай наиболее прост в реализации, если компоненты вектора оценки имеют разные физические единицы измерения или разные масштабы изменения, и может быть вполне строго обоснован, если между компонентами вектора оценки θ отсутствует корреляция, а также в том случае, когда не делается каких-либо предположений об их совместных распределениях.

На рис. 1 представлена параллельная схема обучения сети, в которой оптимизационный алгоритм, вычисляющий оценку максимального правдоподобия (ОМП), функционирует параллельно с ИНС.

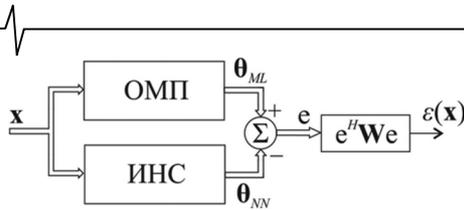


Рис. 1. Параллельная схема включения ИНС для её обучения

Задача обучения представленной сети состоит в том, чтобы за счёт подбора параметров w минимизировать на обучающем множестве примеров Λ суммарную среднеквадратическую ошибку J :

$$J = \sum_{n \in \Lambda} \varepsilon(x_n). \tag{3}$$

Недостатком данного решения является необходимость вычисления оценок по методу максимального правдоподобия для всех примеров тестового множества. Более того, нетрудно предположить, что в том случае, если примеры тестового множества Λ создаются на основе модели специально для обучения сети по известным значениям вектора параметров θ , лучший результат обучения будет получен, если вместо его оценки θ_{ML} использовать непосредственно истинные значения θ .

Для построения нейронной сети, формирующей на своём выходе оценку максимального правдоподобия, рассмотрим альтернативную каскадную схему включения сети в процесс формирования оценки, изображенную на рис. 2.

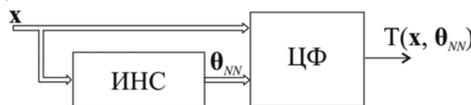


Рис. 2. Каскадная схема включения ИНС для её обучения

Входной вектор x подаётся на вход блока ИНС и один из входов блока вычисления целевой функции (ЦФ). Вектор параметров θ_{NN} , сформированный на выходе блока ИНС, поступает на другой вход блока вычисления целевой функции. При такой схеме включения задача обучения сети заключается в достижении такого набора её внутренних параметров (весов w), который обеспечит для каждого примера из обучающего множества формирование такой оценки вектора параметров θ_{NN} , которая доставит максимум целевой функции $T(x, \theta)$.

В том случае, если целевой функцией является логарифм функции максимального правдоподобия $p_{\xi}(x; \theta)$ интегральной оценкой качества на обучающем множестве примеров Λ будет сумма максимумов:

$$G = \sum_{n \in \Lambda} T(x_n, \theta_{NN}(x_n)). \tag{1}$$

Такой способ формирования интегральной оценки позволяет построить схему обучения ИНС на основе подходящего обобщения метода обратного распространения ошибки. Заметим при этом, что не требуется знать ни истинное значение вектора параметров θ , ни его МП оценки θ_{ML} , полученной численным методом. Для обучения сети достаточно, чтобы градиент целевой функции $\nabla_{\theta} T = grad\{T(x, \theta)\}$ по вектору параметров θ мог быть представлен в форме аналитического выражения или, по крайней мере, был вычисляемым по схе-

ме конечных разностей на основе значений самой функции T .

Представим градиент целевой функции по вектору весов w в нейронной сети в форме вектор-столбца:

$$\nabla_w T(x, \theta_{NN}(x)) = \left\{ \frac{\partial T(x, \theta_{NN}(x))}{\partial w_k} \right\}^T, \tag{5}$$

компоненты которого вычисляются по правилу производной сложной функции:

$$\frac{\partial T(x, \theta_{NN}(x))}{\partial w_k} = \sum_{n=1}^N \left\{ \frac{\partial T(x, \theta_{NN}(x))}{\partial \theta_m} \cdot \frac{\partial \theta_m}{\partial w_k} \right\}, \tag{6}$$

где $\partial T(x, \theta_{NN}) / \partial \theta_m$ – частная производная целевой функции по m -ой компоненте вектора параметров θ_{NN} , а $\partial \theta_m / \partial w_k$ – частная производная m -го выхода нейронной сети по k -му элементу вектора весов w .

Поскольку каждое слагаемое в (6) ограничено сверху, достижение максимума G обеспечивается максимизацией каждого из его слагаемых. Для этого изменение вектора весов ИНС w при обучении необходимо проводить в направлении роста целевой функции T , т.е. в направлении градиента. При использовании пакетного режима обучения и методов первого порядка, уравнение обновления вектора весов имеет вид:

$$w^{q+1} = w^q + L(q) \sum_{n \in \Lambda} \nabla_w T(x_n, \theta_{NN}(x_n)), \tag{7}$$

где $L(q)$ – квадратная матрица, определяющая скорость обучения в q -ю эпоху по каждому из весов w_k . Так при простейшей реализации метода градиентного спуска, матрица $L(q)$ выбирается диагональной, на главной диагонали которой размещаются не изменяющиеся во время обучения одинаковые элементы.

Для достижения лучшего приближения оценки обученной ИНС, а также для сокращения времени обучения следует использовать методы обучения второго порядка и смешанных порядков [13]. Выбор для обучения ИНС метода Левенберга-Марквардта [6] приводит к процедуре обновления:

$$w^{q+1} = w^q + \left\{ \sum_{n \in \Lambda} [\nabla_w T(x_n, \theta) \times \nabla_w^T T(x_n, \theta)] + \mu(q) I \right\}^{-1} \nabla_w T(x_n, \theta), \tag{8}$$

где через I обозначена единичная матрица размера N , $\mu(q)$ – комбинационный коэффициент, определяющий вклад компоненты наискорейшего спуска в приближение матрицы, обратной матрице Гессе.

Оценка направления прихода радиосигнала

Задача определения направления прихода сигнала состоит в оценке двух углов: пеленга α и угла места β – между направлением прихода фронта электромагнитной волны и опорным направлением. Модель рассматриваемой задачи включает в себя источник радиоизлучения (ИРИ) и узкобазовую пассивную РЛС [15], в которой реализуется радиопеленгационный (direction-of-arrival) метод [16] определения местоположения. В настоящем исследовании был рассмотрен двумерный случай, в котором оценивается только один параметр – пеленг. Для

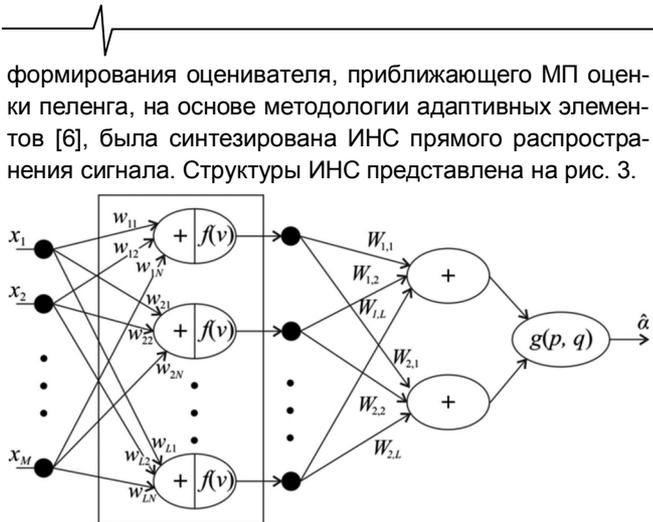


Рис. 3. Структурная схема ИНС. W и w – веса выходного и скрытого слоев соответственно

На вход ИНС подается вектор признаков X , элементы которого представляют собой разности времени прихода сигналов, принимаемых точками приёма антенной системы (АС), входящей в состав РЛС. На своём выходе ИНС формирует искомую оценку $\hat{\alpha}$ пеленга. В нейронах внутреннего слоя использованы сигмоидальные функции активации

$$f(v) = \frac{1}{1 + \exp(-v)}. \quad (9)$$

Функция активации нейрона выходного слоя g представляет собой функцию аргумента комплексного числа, вычисляющую угол вектора, заданного в декартовых координатах:

$$g(p, q) = \arg(p + jq). \quad (10)$$

Было проведено численное моделирование для кольцевой АС, состоящей из 7 точек приема: по три точки расположены на двух концентрических кольцах, одна точка – в центре системы. Для обучения сети было синтезировано множество обучающих примеров, охватывающих

диапазон пеленгов от -45 градусов до 225 градусов с шагом 1 градус. Каждый пример представляет собой вектор задержек, вычисленный на основе аналитической модели АС.

Каждая ИНС из семейства сетей, имеющих представленную на рис. 3 структуру и отличающихся размером скрытого слоя, была обучена методом Левенберга-Марквардта в соответствии с правилом (8). После завершения обучения выполнен анализ остаточной ошибки на обучающем множестве, по результатам которого была выбрана ИНС, содержащая в скрытом слое 12 нейронов.

Практический интерес представляют результаты сравнения точности оценок пеленга, получаемых с использованием ИНС, и оценок, вычисляемых путем применения численных методов для непосредственной максимизации функции правдоподобия. На рис. 4 представлены зависимости точности, измеряемой среднеквадратическим отклонением (СКО), от отношения сигнал-шум (ОСШ). Каждая оценка получена путем усреднения 10000 результатов опытов. Также на рис. 4 представлена нижняя граница Крамера-Рао, аналитическое выражение которой для концентрической многокольцевой кольцевой АС было получено в работе [17].

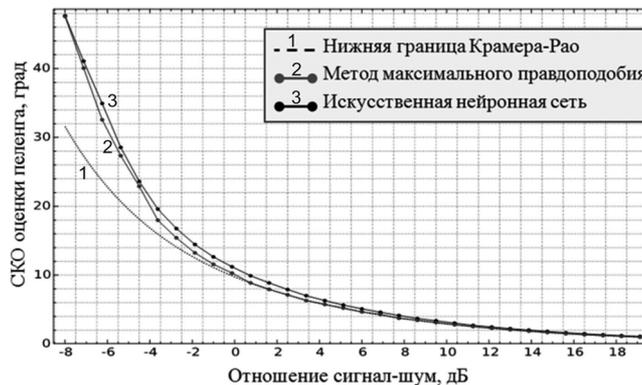


Рис. 4. Зависимость точности пеленга от отношения сигнал-шум

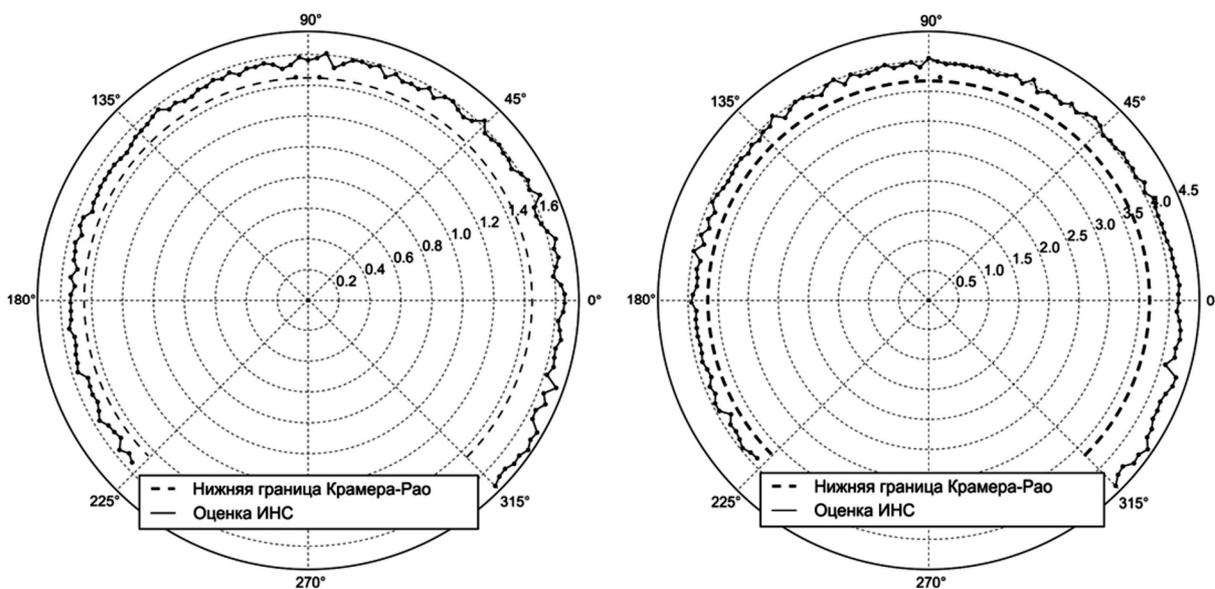


Рис. 5. Точность оценки пеленга в зависимости от его истинного значения для отношений сигнал-шум: 8 дБ (слева) и 10 дБ (справа)

Таблица 1. Сравнительный анализ вычислительных затрат оценщиков

Параметр		ИНС	Численный алгоритм оптимизации
Количество вычислительных операций		Определяется количеством нейронов в структуре ИНС	Определяется функцией правдоподобия и числом итераций алгоритма
Зависимость времени вычисления от выбора начального приближения		нет	да
Время единичной оценки пеленга		2,74 мс.	33,4 мс.
Время на обучение		1,32 мин.	нет
Увеличение СКО к нижней границе Крамера-Рао	-4 дБ	30 %	26 %
	0 дБ	13 %	4 %
	10 дБ	9 %	2 %

Анализ зависимостей на рис. 4 позволяет отметить, что с увеличением ОСШ, МП-оценки, одна из которых получена с использованием ИНС, а другая – с использованием численной оптимизации, асимптотически приближаются к нижней границе Крамер-Рао, минимально достижимой в рамках данной модели СКО. Оценка, получаемая с использованием ИНС, обладает меньшей точностью, чем оценка численного метода, однако СКО возрастает не более чем в 1,1 раза в диапазоне ОСШ от – 8 до 10 дБ.

В результате проведенного численного моделирования также было установлено, что значимая зависимость точности, получаемой с использованием ИНС оценки пеленга, от истинного его значения отсутствует. На рис. 5 представлены зависимости СКО оценки пеленга в секторе обзора шириной 270 градусов для двух значений ОСШ. Основные результаты сравнительного анализа вычислительных затрат алгоритмов представлены в табл. 1. Оценка временных затрат приведена при выполнении вычислений на персональном компьютере, оснащенный процессором Intel Core i 52.0 ГГц и 4 ГБ ОЗУ. Отметим, что время на обучение ИНС затрачивается однократно.

Заключение

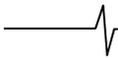
Применение ИНС для эффективного приближения МП оценок позволило создать алгоритм оценки пеленга в пассивной РЛС, реализация которого требует значительно меньше вычислительных ресурсов по сравнению с максимизацией функции правдоподобия посредством итеративного численного алгоритма. Результаты проведенного численного моделирования показывают, что вычисление единичной оценки с использованием ИНС требует в 12 раз меньше времени по сравнению с методами численной оптимизации. При этом для вычисляемых посредством ИНС оценок пеленга отсутствует зависимость от его истинного значения в широком диапазоне углов. Снижение точности оценки по сравнению с методом максимального правдоподобия составляет не более 10 % в диапазоне ОСШ от – 8 дБ до 18 дБ.

Исследование выполнено при финансовой поддержке

ке РФФИ в рамках научного проекта № 16-37-00395 «мола».

Литература

1. Галушкин А.И., Нейронные сети. Основы теории – М.: Горячая Линия-Телеком, 2012 – 496 с.
2. Хайкин С., Нейронные сети: полный курс, 2-е изд., испр.: Пер. сангл. – М.: ООО «И.Д. Вильямс», 2006. – 1104 с.
3. Baum E., Supervised Learning of Probability Distributions by Neural Networks // American Institute of Physics, 1988, pp. 52–61.
4. Setiono R., A neural network construction algorithm which maximizes the likelihood function // Connection Science, vol. 7, no. 2, 1995, pp. 147–166.
5. C. Cervellera, D. Maccio, M. Muselli, Deterministic learning for maximum-likelihood estimation through neural networks // IEEE Transactions on Neural Networks, vol. 19, no. 8, 2008, pp. 1456–1467.
6. Ефимов Е.Н., Шевгунов Т.Я., Построение нейронных сетей прямого распространения с использованием адаптивных элементов // Журнал радиоэлектроники, М.: ИРЭ РАН, №8, 2012.
7. Ефимов Е.Н., Филимонова Д.В., Шевгунов Т.Я. Применение нейронных сетей прямого распространения для формирования оценок параметров по методу максимального правдоподобия // Научно-технические технологии. М.: Радиотехника. 2015. № 8. С. 42–47.
8. Татузов А.Л., Нейронные сети в задачах радиолокации. – М.: Радиотехника. 2009. – 432 с.
9. Строцев А.А., Ломанцова Ю.А., Оценка влияния топологии антенной решетки пеленгатора на качество функционирования нейронной сети, формирующей оценку пеленга // 19-я Международная конференция «Цифровая обработка сигналов и её применение», М.: ИПУ РАН, 29-31 марта 2017, С. 818–822.
10. Шевгунов Т.Я., Ефимов Е.Н., Филимонова Д.В., Синтез искусственных нейронных сетей прямого распространения, приближающих оценки максимального правдоподобия // 19-я Международная конференция «Цифровая обработка сигналов и её применение», М.: ИПУ РАН, 29-31 марта 2017, С. 818–822.



11. Kay S.M. *Fundamentals of Statistical Signal Processing: Estimation Theory*. – Upper Saddle River. – Prentice Hall, 1993.

12. Большаков А.А., Каримов Р.Н., *Методы обработки многомерных данных и временных рядов*, 2-е изд., М.: Горячая линия – Телеком, 2014.

13. R. Battiti, *First- and Second-Order Methods for Learning: Between Steepest Descent and Newton's Method* // *Neural Computation*, 1992, Vol. 4, № 2, pp. 141–166.

14. D. W. Marquardt, *An algorithm for least-squares estimation of non-linear parameters* // *J. Soc. Ind. Appl. Math.*, vol. 11, 1963, pp. 431–441.

15. Дубровин А. В., Сосулин Ю. Г. Одноэтапное оценивание местоположения источника радиоизлучения пассивной системой, состоящей из узкобазовых подсистем // *Радиотехника и электроника*, 2004, т. 49, № 2, С. 156–170.

16. Zekavat R., Buehrer R. M., *Handbook of Position Location: Theory, Practice and Advances*. – Wiley-IEEE Press, 2011.

17. Дубровин А.В. Потенциальная точность пеленгования комплексами с антенными решетками, имеющими конфигурацию в виде набора произвольного числа колец // *Радиотехника и электроника*, 2006, т. 51, №13, С. 268–270.