ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«БЕЛГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ»
(Н И У « Б е л Г У »)

ИНСТИТУТ ИНЖЕНЕРНЫХ И ЦИФРОВЫХ ТЕХНОЛОГИЙ

КАФЕДРА ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМ И ТЕХНОЛОГИЙ

РАЗРАБОТКА И ИССЛЕДОВАНИЕ МЕТОДА СКРЫТНОГО ПОДТВЕРЖДЕНИЯ ЦЕЛОСТНОСТИ ЗВУКОЗАПИСИ

Выпускная квалификационная работа
обучающегося по направлению подготовки 1
1.04.02 Инфокоммуникационные технологии и системы связи
очной формы обучения, группы 12001736
Укуахамба Ядмилде Авелину

Научный руководитель
кандидат технических наук,
доцент кафедры
Информационно-
телекоммуникационных
систем и технологий
НИУ «БелГУ» Прохоренко Е.И.

Рецензент
к.т.н., доцент
начальник отдела программного
обеспечения информационных
средств
ООО «НПП “ЭИТ” БелГУ»
Соловьев В.И.

БЕЛГОРОД 2019

Оглавление

ВВЕДЕНИЕ

Речевая коммуникация между людьми, остается одним из самых распространенных способов передачи информации. Существующие технологии позволяют без труда фальсифицировать такой способ коммуникации как устная речь. Естественно, такой способ коммуникации нуждается в защите. К направлениям обеспечивающих цифровую защиту сообщения в виде звукозаписи устной речи, относят методы стеганографии и криптографии.

Стеганография, как и криптография, являются направлениям защиты информации, но их идеология различна. В криптографии сообщение шифруется для защиты его содержимого. Уточним, что зашифрованное сообщение не может быть расшифровано без ключа. После расшифровки защита снимается, и конфиденциальность больше не сохраняется. В стеганографии о существовании сообщения никому неизвестно, кроме получателя и отправителя. Именно из-за скрытности обеспечивается защита.

Известно «Криптографическое сообщение может быть перехвачено злоумышленникам, однако злоумышленник может даже не знать, что существует стеганографическое сообщения» [1]. Таким образом, стеганография не только защищает конфиденциальное сообщение, так и криптография, но также обеспечивает безопасность отправителя и получателя. Тем временем, стеганография и криптография могут быть объединены, чтобы обеспечить повышение безопасности. То есть шифрование сообщения с

использованием криптографии, а затем скрытие шифрованного сообщения внутри цифрового объекта.

Стоит отметить, что некоторые исследователи приводят термин, связанный с объединением стеганографии и криптографии – цифровые отпечатки.

Цифровые отпечатки – это идентификатор в виде секретного шифрованного сообщения внутри объекта, который позволяет выделить его уникальность относительно подобных объектов. Иными словами «цифровой отпечаток» относится к процессу идентификации и скрытной записи шифрованного сообщения в код который уже присущ объекту. Это часто рассматривается как метод, используемый для отслеживания авторизованных пользователей.

Отличие цифровых отпечатков от цифровых водяных знаков – это происхождение скрытого шифрованного сообщения (уникального кода) и способа его кодирования. В цифровых водяных знаках, код представляет собой произвольное сообщение, содержащее информацию о собственности. Цифровой отпечаток представляет собой последовательность формируемый передаваемыми данными и идентификатором авторизованного пользователя. Далее эта последовательность преобразуется в уникальный, но значительно более короткий номер или строку.

Стоит отметить, цифровой отпечаток состоит из трех основных частей, а именно: хэш-кодера (использующего идентификатор авторизованного пользователя и анализатор передаваемых данных), устройства для внедрения и детектора.

Теоретическая часть работы содержит математические модели формирования хэш-последовательности использующего идентификатор авторизованного пользователя и структуру передаваемых данных, являющихся звукозаписью. Описываются математические основы методов скрытного внедрения информации.

В практической части приведены:

– схема скрытного внедрения цифрового отпечатка;

– схема идентификации цифрового отпечатка;

– алгоритм формирования хэш-последовательности цифрового отпечатка;

– алгоритм скрытного внедрения цифрового отпечатка;

– алгоритм извлечения цифрового отпечатка;

– алгоритм идентификации цифрового отпечатка.

Приводятся планы вычислительных экспериментов и их результаты, оценивающие возможность идентификации звукозаписи на основе цифрового отпечатка.

В заключении приведены рекомендациями по дальнейшему применению разработанных алгоритмов.

Методы исследования. В магистерской работе использованы методы цифровой обработки сигналов, статистического анализа и вычислительных экспериментов.

Объект: кодовое представление звукозаписи устной речи.

Предмет: метод формирования цифрового отпечатка.

Целью работы является разработка схемы внедрения цифрового отпечатка в звукозапись устной речи, с обеспечением скрытности.

Глава 1. Кодирование звукового сигнала

Для повышения эффективности освоения, передачи и обработки знаний, часто прибегают к использованию информационного обмена. Для человека представляется естественным осуществлять информационный обмен, используя устную речь и визуальное отображение предметов, явлений или процессов. К основными этапам информационного обмена можно отнести:

– кодирование информации;

– декодирование информации.

Кодирование представляет собой отождествление предметов, явлений или процессов с некоторыми конструкциями, которые составляют некоторое конечное множество, формируемыми в процессе развития обмена информации.

Предполагается, что фрагмент звукозаписи, нуждающийся в защите, представлен декретированной эквидистантной последовательностью значений амплитуд размерности и обозначается вектором размерности N, вида:

$$\vec{x} = (f_1, f_2, .., f_i, .., f_N)^T, \quad (1.1)$$

где f_i – дискретное значение амплитуды в момент времени t_i.

1.1 Методы анализа и синтеза звукового сигнала

Подраздел посвящен описанию свойств звукозаписей устной речи, для случая цифровой регистрации устной речи – в дискретные моменты времени с преобразованием её в кодовое представление – речевые данные.

«Человек для информационного обмена достаточно часто использует устную речь» [2]. Индустрия создания информационного, образовательного и развлекательного контента применяет устную речь в виде сигнала,

разделенного на фрагменты, для звукового сопровождения информационных справок, фильмов и музыкальных композиций. Это приводит к росту потоков информации, содержащей речевые сообщения. Важным моментом использования устной речи в индустрии является субъективное качество, с которым воспроизводится устная речь [19, 33]. Наличие тех или иных частотных компонент в речи формирует её тональную окраску и помогает улучшить эмоциональное восприятие с одной стороны, но приводит к необходимости хранения избыточности (таблица 1.1), с другой стороны. При этом стоит учитывать, что устная речь представлена совокупностью последовательности звуков. В свою очередь звуки, имеющие идею или смысл, образуют речевое сообщение. Для дальнейшей обработки, хранения и передачи речевое сообщение регистрируют в виде речевых сигналов, которые представляют собой результаты регистрации звуков в течении заданного промежутка времени. Речевые сигналы, регистрируют как колебания электрического тока на выходе микрофона. В случае цифровой регистрации устной речи – в дискретные моменты времени с преобразованием её в кодовое представление – речевые данные. С математической точки зрения устная речь – это нестационарный, сложно-модулированный сигнал (в случае цифрового представления – отрезок речевых данных), порождаемый последовательностью звуков языка или их отсутствием. Свойство не стационарности является основным свойством речевых сигналов, среди учитываемых в задачах анализа и синтеза.

Таблица 1.1 – Свойства речевых сигналов доступные для модификации

Во временной области	**В частотной области**
Неравномерное амплитудное распределение	Неравномерный спектр
Корреляция между отчетами	Сосредоточенность спектра в узкой области частот
Корреляция между периодами основного тона	
Паузы в сигнале	

Избыточность устной речи, выявленную в процессе определения закономерностей речевых данных, можно использовать в ряде стеганографических задач, когда необходимо: подтверждение идентичности полученной информации; определение её целостности; хранение, при котором контрольную информацию невозможно обнаружить, если не знать о её существовании; также иногда необходима защита информации от несанкционированного доступа; помещение в информацию дополнительных сведений, без увеличения объема таким образом, чтобы сведения о ней отображались и воспринимались органами чувств человека тогда, когда в них возникнет необходимость. Поэтому речевые данные необходимо рассматривать с двух позиций:

- как объект, в котором осуществляют скрытое кодирование, обеспечивающее хранение и передачу контрольной информации;

- как объект, который сам представляет собой контрольную информацию.

Большинство методов скрытого кодирования предполагают предварительный анализ данных и адаптацию под каждый отрезок данных. Для речевых данных это так же необходимо, т.к. устная речь стационарна на небольшом промежутке времени 20-30 мс [3]. При этом, с течением времени, у речевых данных не только изменяются временное представление, но и частотное. Численные значения данных и распределение энергии по частотным компонентам зависят:

– от принадлежности звука к буквам русской речи;

– от местоположения (начало, середина, конец) звука в речевом сообщении (в русском языке выделяют 120 фонем [3]);

– от голосового аппарата диктора, воспроизводящего речевое сообщение;

– от параметров анализа отрезка речевых данных.

Это говорит о том, что для большей однозначности определения характеристик и закономерностей в речевых данных, следует использовать разбиение на отрезки, которые полностью порождены одним и тем же звуком речи. При этом математические основы должны адекватно отражать свойства речевых сигналов и позволять создавать алгоритмы их обработки.

Стоит отметить, что для дискретных сигналов с ограниченной энергией, т.е. сигнала $f(t)$ с ограниченным спектром рассматриваемого только в течении конечного интервала времени T точное разложение [38]:

$$\int_{-\infty}^{\infty} f^2(t)dt = \frac{1}{\vartheta_\partial} \sum_{l=-\infty}^{\infty} f_l^2 , \qquad (1.1)$$

заменяется приближенным:

$$\int_{0}^{T} f^2(t)dt \approx \frac{1}{\vartheta_\partial} \sum_{l=1}^{L} f_l^2 , \qquad (1.2)$$

Следовательно, для отрезка данных прямое преобразование Фурье (1.28) имеет вид [21, 36, 37, 39]

$$F_\upsilon = \sum_{l=1}^{L} f_l e^{-j2\pi\upsilon l\Delta} , \ \upsilon \in \left[-\frac{\vartheta_\partial}{2}, \frac{\vartheta_\partial}{2} \right] \qquad (1.3)$$

где F_υ – трансформанта Фурье сигнала,

Для частотных компонент (трансформант Фурье) обратное преобразование Фурье (ф. 1.13), определяется следующим образом [34]:

$$f_l = \frac{\Delta}{2\pi} \sum_{\upsilon=-\frac{\vartheta_\partial}{2}}^{\frac{\vartheta_\partial}{2}} F_\upsilon e^{j2\pi u l\Delta} , \ l = 1,2,\ldots,L . \qquad (1.4)$$

В соответствии со следствием из равенства Парсеваля [21, 34, 35, 36, 37]:

$$\int_{-\infty}^{\infty} f^2(t)dt = \int_{-\pi}^{\pi} F^2(\upsilon)d\upsilon/2\pi , \qquad (1.5)$$

важной характеристикой сигнала является его энергия, которая определяется на основе квадрата евклидовой нормы:

$$\|f\|^2 = \int_0^T f^2(t)dt\,, \tag{1.6}$$

где $\|\ \|$ - евклидова норма сигнала:

$$\|f\| = \sqrt{\langle f, f \rangle} = \sqrt{\int_G f^2(t)dt}\,. \tag{1.7}$$

Учитывая соотношения 1.6-1.7 с некоторой погрешностью δ можно определить численное значение энергии отрезка данных $E(f)$, получив из соотношения:

$$E(f) = \|f\|^2 = \sum_{l=1}^{L} f_l^2\,. \tag{1.8}$$

Так же примем, что энергия отрезка данных в частотной и временной областях равны, правило Парсеваля имеет вид:

$$\sum_{l=1}^{L} f_l^2 = \sum_{\upsilon=1}^{\vartheta_\partial} F_\upsilon^2\,. \tag{1.9}$$

Ниже приведен анализ речевых данных для фразы «Министерство распределяет научные кадры». Речевые данные взяты с характеристиками: частота дискретизации ϑ_∂ =8кГц, разрядность хранимых данных 16 бит. Для анализа распределения энергии использовалось разбиение на отрезки данных содержащих 64 значения (рис.1.1 б), 128 значений (рис.1.1 в), 256 значений (рис.13.г). Результаты исследований показывают, что с изменением времени меняется энергия отрезка данных, в паузах энергия, в случае отсутствия шумов, близка к нулю. Также некоторые звуки содержат малую долю энергии. Более наглядно нестационарность отрезков речевых данных отражена на рисунках 1.1–1.5, которые содержат огибающие сигналов для звуков, входящих в фразу и построены огибающие их спектров.

На рисунках 1.1–1.4 представлены огибающие отрезков речевых данных, порожденных звуком «а», взятого из слова «кадры» для двух дикторов.

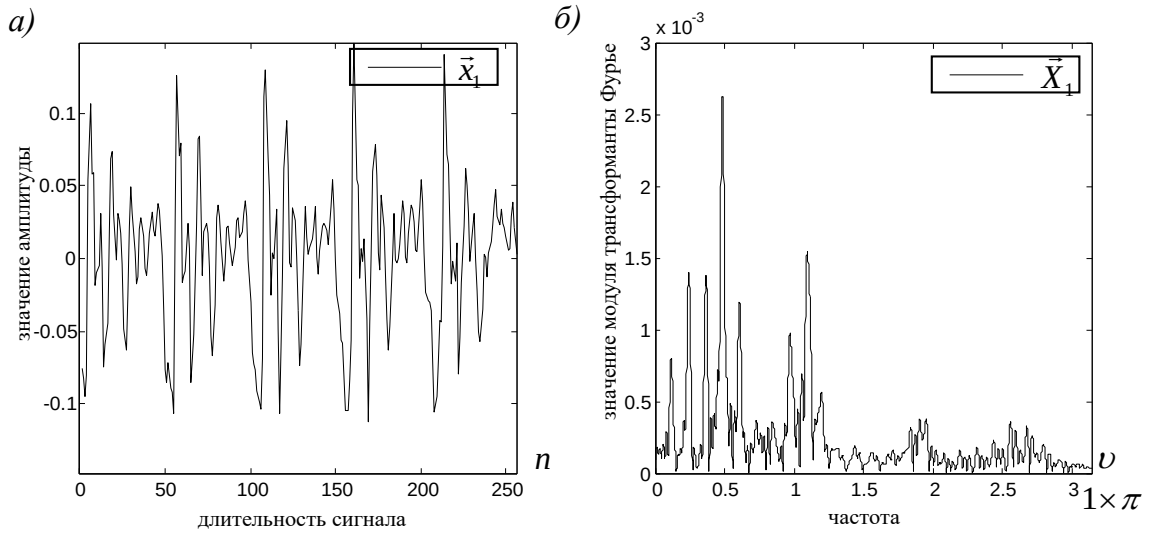


Рисунок 1.1 – Речевые данные, порожденного звуком «а», слова «кадры» произносимого первым диктором: а) огибающая данных во временной области $\vec{x}_1$; б) огибающая нормированного амплитудного спектра $\vec{X}_1$

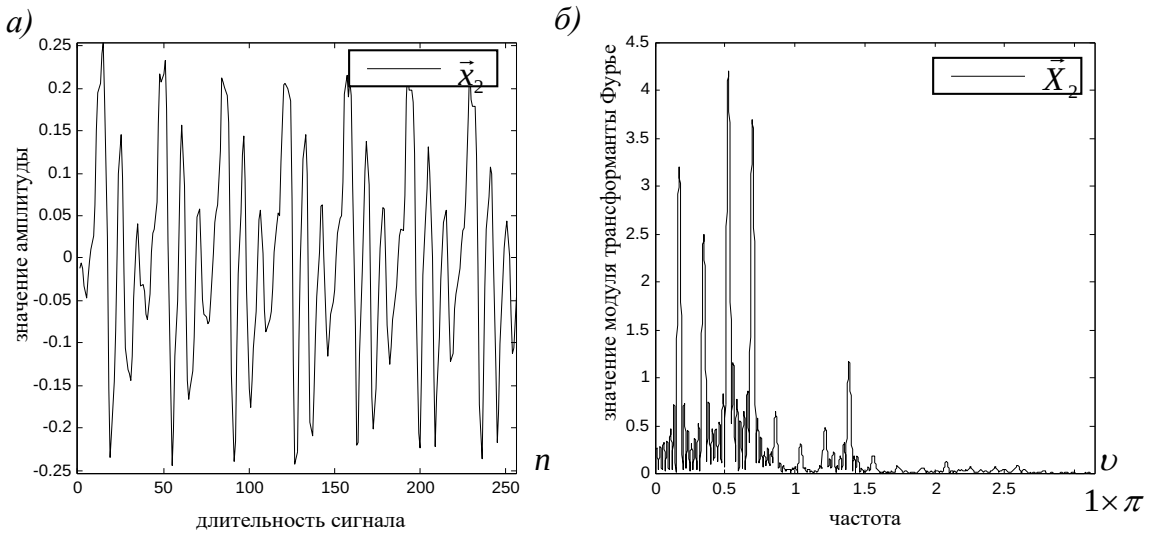


Рисунок 1.2 – Речевые данные, порожденного звуком «а», слова «кадры» произносимого вторым диктором: а) огибающая данных во временной области $\vec{x}_2$; б) огибающая нормированного амплитудного спектра $\vec{X}_2$ У

огибающей речевых данных во временной области рисунок 1.2 можно выделить периодичность, также данные содержат мало шумовых компонент.

У речевых данных, принадлежащих первому диктору (рис. 1.2) также имеется периодичность во временной области, но данные содержат больше

шумовых компонент. Энергии частотных компонент спектров также отличаются, но имеют одинаковые закономерности подавляющая доля частотных компонент расположена в спектре до значения $\pi/2$, максимальное значение находится у компонент близких к $\pi/4$.

На рисунках 1.3-1.4 представлены огибающие отрезков речевых данных, для одного диктора, порожденных звуками «а», для слова «распределяет» (рисунок 1.3) и для слова «научные» (рисунок 1.4).

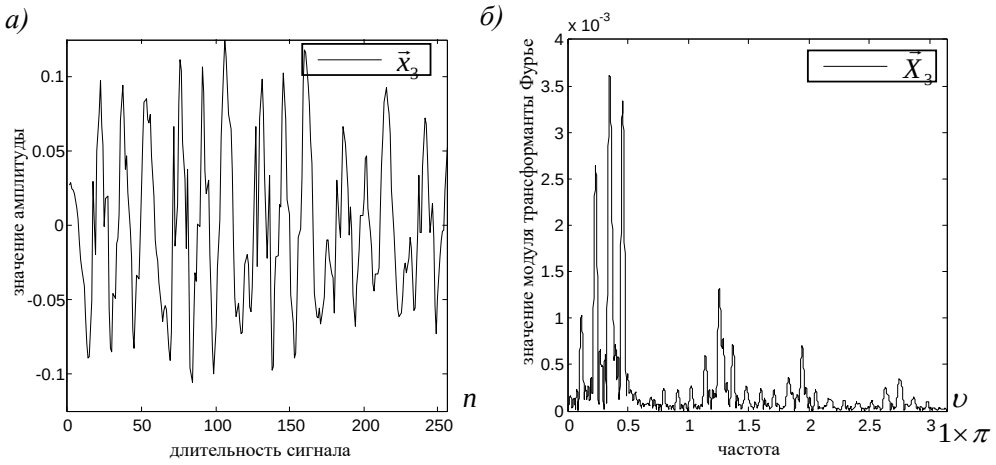


Рисунок 1.3 – Речевые данные, порожденного звуком «а», слова «распределяет» произносимого первым диктором: а) огибающая данных во временной области $\vec{x}_3$; б) огибающая нормированного амплитудного спектра $\vec{X}_3$

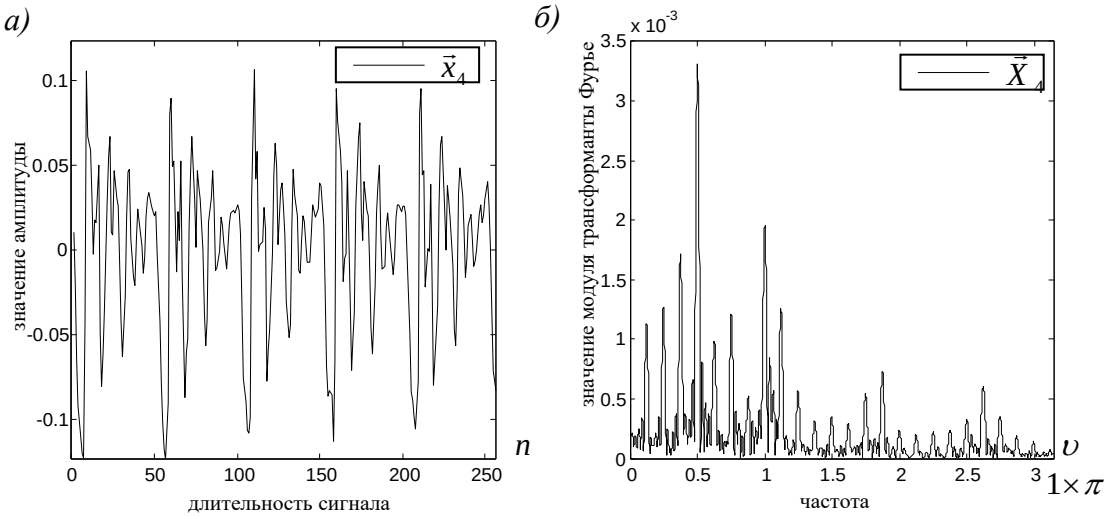


Рисунок 1.5 – Речевые данные, порожденного звуком «а», слова «научные» произносимого первым диктором: а) огибающая данных во временной области $\vec{x}_4$; б) огибающая нормированного амплитудного спектра $\vec{X}_4$

Из рисунков 1.2, 1.3-1.5, видно, что отрезки речевых данных для звука соответствующих одной букве «а», но взятые из разных слов речевого сигнала полученного из одной фразы отличаются, подтверждением этому можно привести качественную оценку.

Таблица 1.2 – Характеристики отрезков речевых данных

Отрезок речевых данных	$\vec{x}_1$	$\vec{x}_3$	$\vec{x}_4$	$\vec{x}_2$
Диктор	Первый диктор			Второй диктор
Математическое ожидание, m_x	-4.45e-04	5.0294e-04	-5.57e-04	-0.0017
Дисперсия, σ_x	0.0029	8.4123e-04	0.0022	0.0146
Энергия, E_x	0.7386	0.2154	0.5687	3.7409

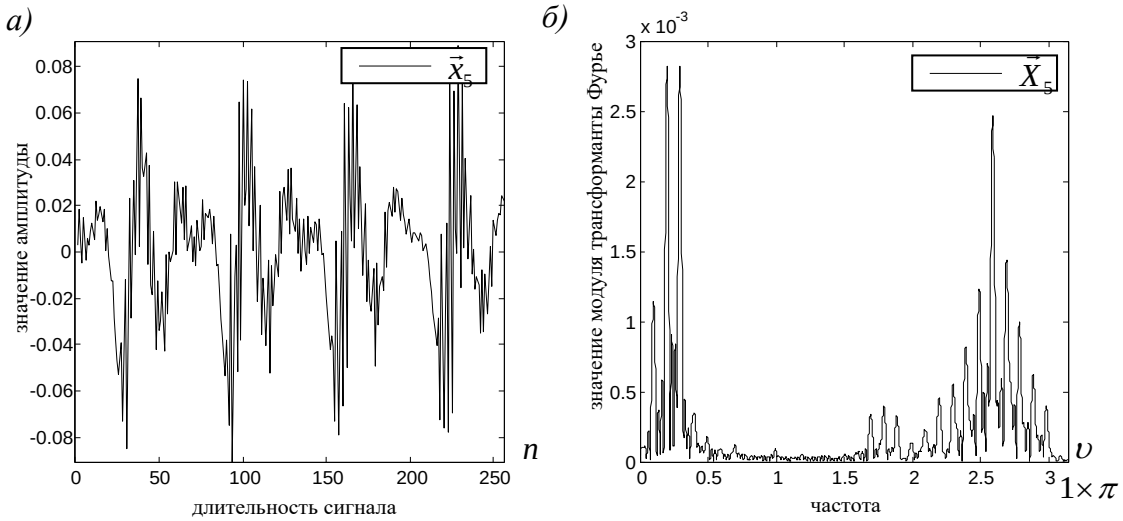


Рисунок1.6 – Речевые данные, порожденного началом первого звука «и», слова «министерство»: а) огибающая данных во временной области $\vec{x}_5$; б) огибающая нормированного амплитудного спектра $\vec{X}_5$

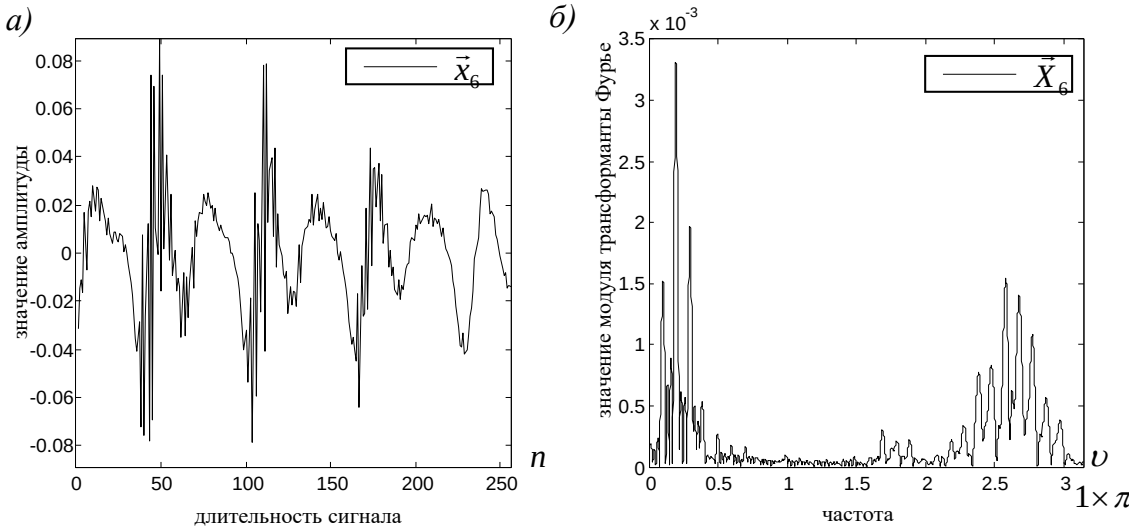


Рисунок 1.7 – Речевые данные, порожденного концом первого звука «и», слова «министерство»: а) огибающая данных во временной области $\vec{x}_6$; б) огибающая нормированного амплитудного спектра $\vec{X}_6$

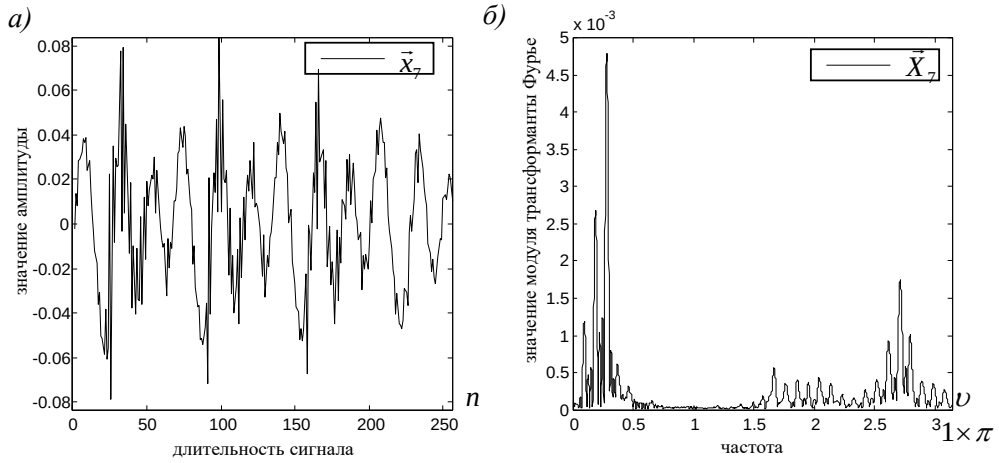


Рисунок 1.8 – Речевые данные, порожденного началом второго звука «и», слова «министерство»: а) огибающая данных во временной области $\vec{x}_7$; б) огибающая нормированного амплитудного спектра $\vec{X}_7$

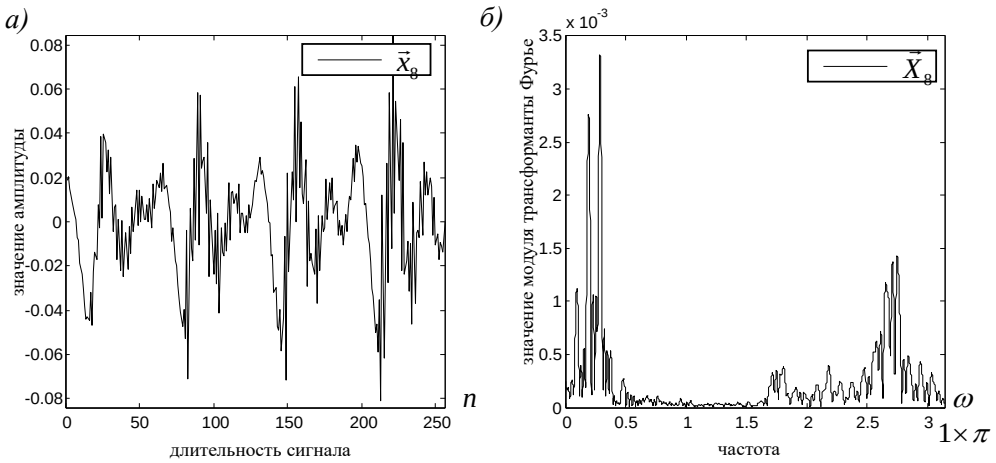


Рисунок 1.9 – Речевые данные, порожденного концом второго звука «и», слова «министерство»: а) огибающая данных во временной области $\vec{x}_8$; б) огибающая нормированного амплитудного спектра $\vec{X}_8$

Таблица 1.3 – Характеристики отрезков речевых данных

Расположение отрезка данных	Первая буква, начало звука	Первая буква, конец звука	Вторая буква, начало звука	Вторая буква, конец звука
	$\vec{x}_5$	$\vec{x}_6$	$\vec{x}_7$	$\vec{x}_8$
Энергия, E_x	0.1960	0.1966	0.1630	0.2197

Из табл. 1.3, что с течением времени наблюдается изменение характеристик сигнала, показывающих, что речь нестационарна.

Результаты исследований показывают, что сигналы, соответствующие различным звукам русской речи, имеют разное распределение трансформант Фурье. Можно выделить ряд особенностей, от которых зависит распределение энергии на частотной оси:

- от особенностей диктора (рис. 1.4-1.8);
- от местоположения звука в фразе;
- от выбранного участка звука: начало, середина, конец;
- от длины отрезка анализа;
- от параметров данных.

Математические основы должны адекватно отражать свойства речевых сигналов и при этом позволять создавать алгоритмы их обработки.

Известно, что устная речь – это нестационарный, сложно-модулированный сигнал, порождаемый последовательностью звуков языка или их отсутствием [2]. Свойство не стационарности является основным свойством речевых сигналов, среди учитываемых в задачах анализа и синтеза. Для синтеза и реализации алгоритмов стеганографии звукозапись устной речи преобразуется к виду:

$$\vec{x} = (x_1, x_2, .., x_i, .., x_N)^T,\ x_i = f_i / 2^{(B-1)},\ x_i \in [-1, 1] \qquad (1.10)$$

где x_i – дискретное значение амплитуды в момент времени t_i; B – разрядность звукозаписи.

При реализации цифровых методов обработки сигналов $f(t)$ используются результаты регистрации колебаний, в виде численного значения амплитуд $f(t_l)$, взятых через промежутки времени Δ с дальнейшим их преобразованием в кодовые комбинации – данные:

$$\vec{f} = (f_1, f_2, \ldots, f_l, \ldots, f_L)^T, \qquad (1.11)$$

$$f_l = f(t_l),\ t_l \in [0, T], \qquad (1.12)$$

где ^{T} - символ транспонирования, подразумевающий формирование данных в виде столбца; f_l - отчет под номером l, т.е. мгновенное значение амплитуды сигнала измеренное в точке дискретного времени t_l:

$$t_l \in [0, T],\ l = 0, 1, ..., L, \qquad (1.13)$$

где L – длительность фрагмента сигнала в отсчетах.

Необходимо отметить, что далее будет использоваться равномерная дискретизация, как наиболее распространенная [21, 22, 23, 24, 25]. То есть далее речь будет вестись об обработке последовательностей отсчетов, регистрируемых через равные промежутки времени:

$$\Delta = t_{l-1} - t_l,\ \Delta = const, \tag{1.14}$$

где Δ - период дискретизации, равный величине обратной частоте дискретизации:

$$\Delta = 1/\vartheta_\partial, \tag{1.15}$$

$$L = \frac{T}{\Delta} = \vartheta_\partial \cdot T, \tag{1.16}$$

где L – количество цифровых значений.

Значения амплитуд отчетов f_l дискретного сигнала хранят в виде чисел с ограниченной разрядностью:

$$f_l = \sum_{b=0}^{(B-1)} c_b \cdot 2^b, \tag{1.17}$$

где c_k - значение разряда $c_k \in \{0,1\}$; B - разрядность численных значений амплитуды сигнала (соответствуют разрядности аналого-цифровых преобразователей $B \in \{8,12,16,24,32\}$); 2 - основание системы счисления.

Разрядность данных определяет шаг квантования и разрешающую способность μ, как расстояние между соседними уровнями квантования:

$$\mu = 2^{-B}. \tag{1.18}$$

1.2 Методы оценки синтезированных сигналов

Описаны основные меры оценивающие изменения в цифровом представлении звукозаписи устной речи.

Мера, отражающая абсолютное изменение энергии отрезков во временной области (*mean square error*), [Ozer H., 2000; Hicsonmez S., 2013; 17]:

$$MSE = \sum_{i=1}^{N} (x_i - y_i)^2 . \tag{1.19}$$

Оценка, определяющая порядок изменения энергии по отношению к общей энергии исходного сигнала (*signal-to-noise ratio* – SNR) [Ozer H., 2000; Hicsonmez S., 2013; 17]:

$$SNR = 10 \cdot \log_{10} \frac{\sum_{i=1}^{N} x_i{}^2}{\sum_{i=1}^{N} (x_i - y_i)^2} . \tag{1.20}$$

Степень структурной схожести синтезированного и исходного отрезка оценим с использованием корреляции ρ [Ozer H., 2000; 17]:

$$\rho = \frac{\left(\sum_{i=1}^{N} (x_i - \bar{x}) \cdot \sum_{i=1}^{N} (y_i - \bar{y}) \right)}{\sqrt{\sum_{i=1}^{N} (x_i - \bar{x})^2 \cdot \sum_{i=1}^{N} (y_i - \bar{y})^2}} . \tag{1.21}$$

Нормированное среднеквадратическое отклонение [19]:

$$\sigma = \sqrt{\sum_{i=1}^{N} (x_i - k \cdot y_i)^2 \Big/ \sum_{i=1}^{N} x_i^2} \; ; \; k = \sum_{i=1}^{N} x_i \cdot y_i \Big/ \sum_{i=1}^{N} x_i^2 . \tag{1.22}$$

1.3 Методы скрытного внедрения информации

Посвящен описанию методов стеганографии. Как было отмечено выше, цифровое представление фрагмента сигнала обладает избыточностью.

Цифровое представление фрагмента сигнала речевого сообщения в виде отрезка данных зачастую обладает избыточностью кодового представления [38-48]. Избыточность может возникать как при хранении численных значений амплитуд, так и в их частотных представлениях. Избыточность часто используют для фильтрации, сжатия и цифровой обработки данных, но при этом остаются частотные компоненты, во временной или частотной областях, обработка которых не проводится и при этом изменения, в которых неощутимы органами чувств человека. А значит, к ним можно применить алгоритмы, реализующие скрывающее кодирование.

Достоинством метода наименее значащего бита является достижение высокой емкости данных, с сохранением высокой скрытности. Если используется скрытое кодирование в четвертом разряде при шестнадцати битовом представлении данных, емкость составляет 64 закодированных бита на один хранимый килобайт. Если используется скрытое кодирование в четырех младших разрядах при шестнадцатиразрядном представлении данных, емкость составляет 256 закодированных бит на один хранимый килобайт. Но при этом метод обладает рядом недостатков:

- изменения происходят во временной и частотной областях. Математические основы метода не позволяют использовать предварительный анализ для адаптивного внедрения с целью минимизации изменений в частотной области данных;

- метод НЗБ можно адаптировать для внедрения во временной области, путем внедрения не в каждое значение данных, а выборочно согласно заранее определенных критериев скрытности;

- артефакты, вызванные кодированием информации, существеннее проявляются в той части частотной полосы, где частотные компоненты минимальны. На слух артефакты воспринимаются как белый шум;

- контрольная информация может быть разрушена при любом изменении в данных.

В работах [1, 2, 30-31] предложен метод расширения спектра, позволяющий осуществить крытое кодирование бита контрольной информации в отрезке данных. Суть метода заключается в добавлении, к отрезку данных – псевдослучайной последовательности, которая является отображением бита контрольной информации:

$$\vec{y} = \alpha_m \cdot e_m \cdot \vec{u} + \vec{x}, \tag{1.23}$$

где $\vec{x}$ - отрезок данных; $\vec{u}$ - отрезок данных, соответствующий псевдослучайной последовательности; e_m - бит контрольной информации $e_m \in \{-1, 1\}$; m - порядковый номер бита контрольной информации; α_m - весовой коэффициент, определяющий скрытность системы, который в работе [56, 57] предлагается выбирать равным:

$$\alpha_m = \frac{\langle \vec{y}, \vec{u} \rangle}{\|\vec{u}\|^2}, \tag{1.24}$$

Декодирование бита контрольной информации из данных происходит путем определения знака скалярного произведения отрезка данных и псевдослучайной последовательности:

$$\hat{e}_m = sign\left(\langle \vec{y}, \vec{u} \rangle\right), \tag{1.25}$$

где $sign(\ )$ - операция выделения знака.

Основой метода служит использование ПСП [54, 63], которая имеет такое же количество значений, как и отрезок данных (рис. 1.11). Достижение скрытности ПСП достигается за счет изменения коэффициента α_m.

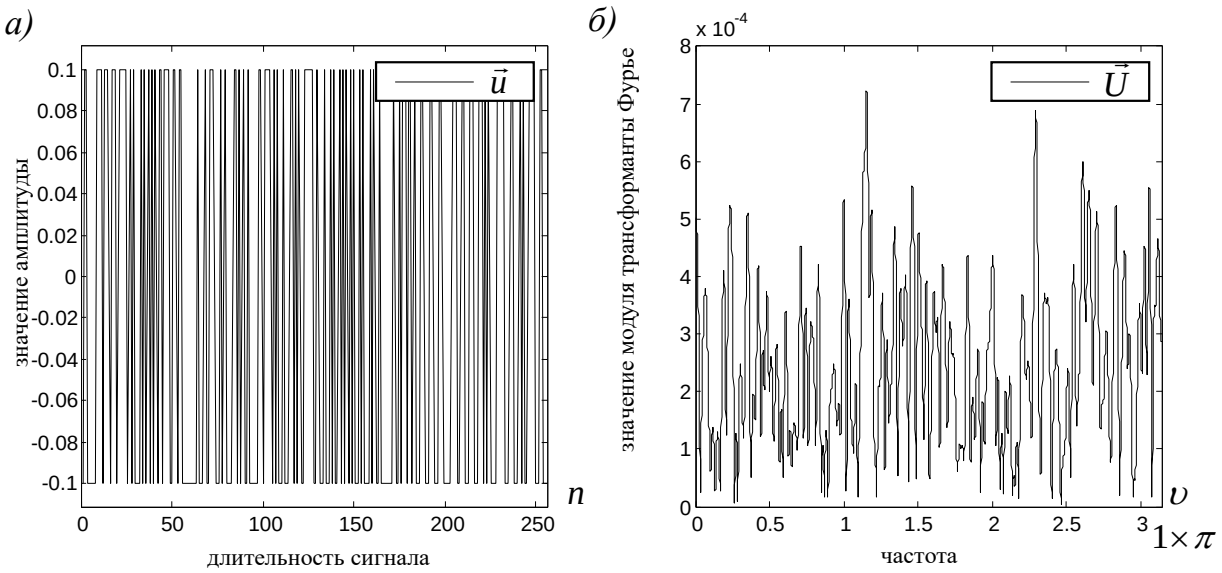


Рисунок 1.10 – Отрезок, соответствующий ПСП:
а) огибающая значения амплитуд $\vec{u}$; б) огибающая нормированного амплитудного спектра $\vec{U}$

Стоит отметить, что метод предполагает скрыто кодировать контрольную информации без перехода в частотную область. Метод расширения спектра позволяет учитывать энергетические свойства отрезка данных в целом, но не может учитывать распределение энергии по частотным компонентам (рис. 1.12). То есть не в полной мере использует закономерности в данных.

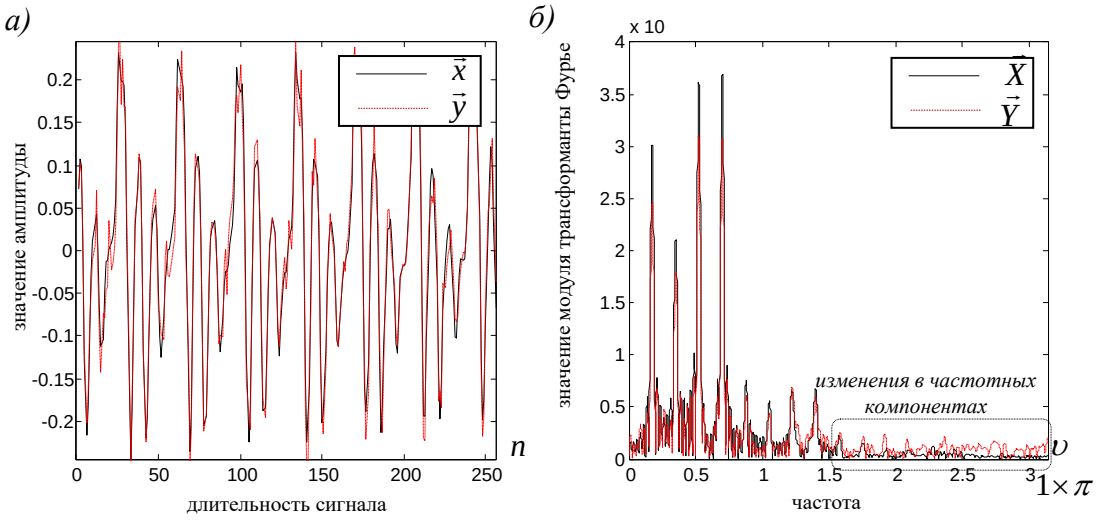


Рисунок 1.11 – Результат скрытого кодирования контрольной информации методом расширения спектра: а) огибающая данных во временной области; б) огибающая модуля трансформанты Фурье

Для учета частотных свойств данных предлагается модификация метода расширения спектра [2, 56, 57], путем увеличения части энергии частотных компонент ПСП. Для этого ПСП модулируют гармоническим сигналом, осуществляя перенос энергии в частотные интервалы, которые были определены при предварительном анализе:

$$\vec{y} = \alpha_m \cdot e_m \cdot \vec{g} + \vec{x}, m = 1,\ldots,M, \quad (1.26)$$

где M – количество бит, содержащихся в контрольной информации; $\vec{g}$ - отрезок данных, соответствующий гармоническому сигналу:

$$g_n = \sin(\upsilon_0 n / N) \cdot u_n, \; n = 1,\ldots,N, \quad (1.27)$$

где υ_0 - частота опорного сигнала; N – количество значений.

Данная модификация метода расширения спектра позволяет не только осуществить перенос на заданную частоту υ_0 (рис. 1.13), но и задать полосу занимаемую ПСП.

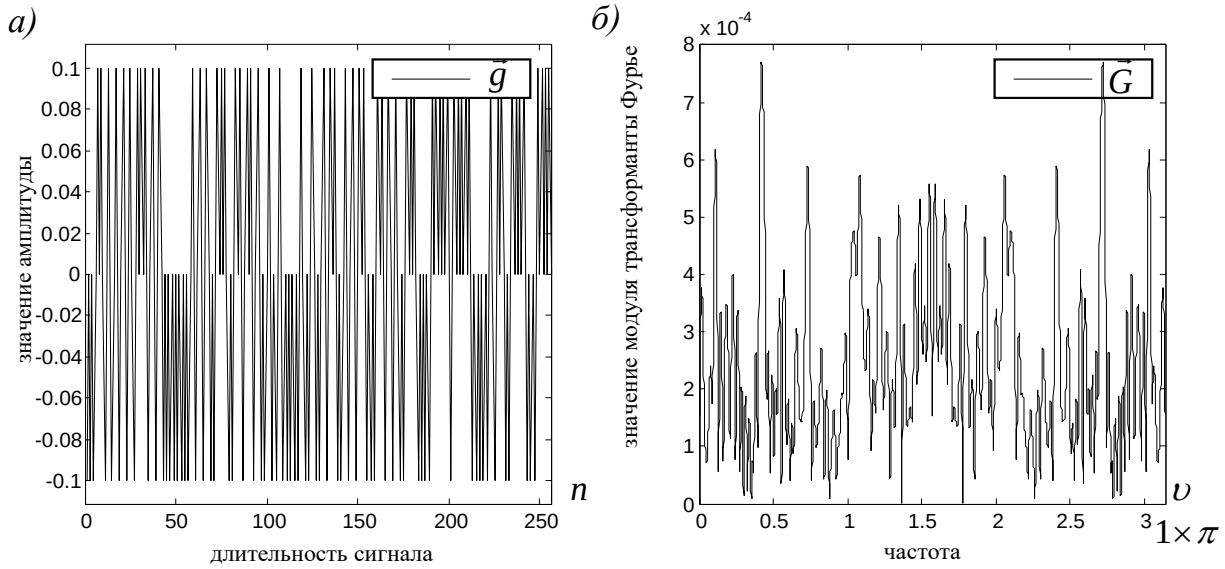


Рисунок 1.12 – Отрезок соответствующий ПСП:
а) огибающая значения амплитуд $\vec{g}$; б) огибающая нормированного амплитудного спектра $\vec{G}$

Экспериментальные исследования показали, что основным достоинством метода расширения спектра является помехоустойчивость при

высокой скрытности. В качестве параметров внедрения для отрезков речевых данных выбиралась частота дискретизации $\vartheta_д = 8$кГц, исследования проводились для коэффициентов α_m меньших одной десятой энергии отрезка речевых данных. Количество перекодированных отрезков составляла 23350000, количество отрезков шума - 10^4.

Таблица 1.4 – Значение вероятности $P_{ош}$ появления ошибочного бита при декодировании двоичных символов e_m

коэффициент, α_m / шум/сигнал, h_0^2	0.1		0.01		0.001	
	N = 128	N = 256	N = 128	N = 256	N = 128	N = 256
0.001	< 1×10-4	< 1×10-4	0.1290	0.0521	0.3614	0.3170
0.01	0.2×10-3	< 1×10-4	0.1285	0.0508	0.3497	0.3013
0.1	0.4×10-3	7.2×10-3	0.1439	0.0720	0.3661	0.3203
1	5.5×10-3	0.1395	0.2133	0.1297	0.4011	0.3602

В качестве параметров скрытого кодирования модифицированным методом расширения спектра для отрезков речевых данных выбиралась частота дискретизации $\vartheta_д = 8$кГц, коэффициент внедрения $\alpha_m = 0.01$, количество значений входящих в отрезок данных $N = 128$. Количество перекодированных отрезков 58150000, количество отрезков шума 10^4.

Таблица 1.5 – Значение вероятности $P_{ош}$ появления ошибочного бита при декодировании двоичных символов e_m, при $\alpha_m = 0.01$

частота, υ_0 / шум/сигнал, h_0^2	$\frac{\pi}{4}$		$\frac{\pi}{2}$		$\frac{3\pi}{4}$	
	N = 128	N = 256	N = 128	N = 256	N = 128	N = 256
0.001	0.1350	0.0523	0.1300	0.0624	0.1256	0.0517
0.01	0.1328	0.0569	0.1296	0.0504	0.1315	0.0529
0.1	0.1392	0.0578	0.1507	0.0661	0.1413	0.0576
1	0.2105	0.1255	0.2127	0.1330	0.2083	0.1253

Из экспериментально полученных данных (таблица 1.2 и таблица 1.3), видно, что для обеспечения стойкости контрольной информации к шуму необходимо использовать некоррелированную с данными ПСП, имеющую не

меньше 256 значений, при этом стойкость данных не зависит от частоты, которой была промоделирована ПСП.

Однако метод обладает рядом недостатков:

- низкий объем передаваемых данных, т.к. в один отрезок данных можно поместить один бит информации;

- для верного декодирования контрольной информации необходимо хранить ПСП или параметры, позволяющие её восстановить;

- для декодирования необходимо выполнение условия синхронизации;

- при внедрении контрольной информации в отрезок данных необходимо учитывать степень корреляции между ПСП и данным, т.к. в случае высокой корреляции и малой энергии ПСП имеется возможность неверного декодирования контрольной информации;

- метод изменяет энергию всего отрезка путем добавления энергии ПСП последовательности;

- метод не позволяет в адаптивном режиме скрыто закодировать информацию в строго определенных частотных подобластях отрезка данных. Энергию ПСП сложно сосредоточить в заданном частотном интервале, так как происходит просачивание её в другие частотные интервалы.

Стоит отметить, что метод расширения спектра осуществляет кодирование без преобразования значений отрезка данных в другие формы представления.

Рассмотрим один из распространённых методов стеганографического кодирования использующий прямое разложение отрезка аудио-сигнала $\vec{x}$, на DCT-коэффициенты вида [Malvar H. S., 1992]:

$$\alpha_1 = \frac{\sqrt{2}}{N}\sum_{i=1}^{N} x_i \,. \tag{1.28}$$

$$\alpha_k = \frac{2}{N}\sum_{i=0}^{N-1} x_i \cdot \cos\left(\frac{(2i+1)\cdot k\cdot \pi}{2N}\right);\ k = 2,3,\ldots, N\,. \tag{1.29}$$

$$\alpha = (\alpha_1, \alpha_2, .., \alpha_k, .., \alpha_N)^T, \quad (1.30)$$

где x_i – значение амплитуды сигнала $x_i \in \bar{x}$; m – номер DCT-коэффициента; α_k – DCT-коэффициент $k = 0,1,\ldots,(N-1)$.

Кодирование осуществляется путем замены знака коэффициента разложения:

$$\begin{gathered} \tilde{\alpha}_k = sign(e) \cdot |\alpha_k|; \; k \in \{1,2,\ldots,(N-1)\}; \\ \bar{\alpha} = (\alpha_1, \alpha_2, .., e \cdot |\alpha_k|, .., \alpha_N)^T = (\alpha_1, \alpha_2, .., \tilde{\alpha}_k, .., \alpha_N)^T \end{gathered} \quad (1.31)$$

так как гармонический сигнал можно представить в виде вектора (функция и спектр гармонического сигнала приведены на рисунке 6.):

$$\begin{gathered} \vec{g} = (g_1, g_2, \ldots, g_i, .., g_N)^T, \; g_i = \cos\left(\frac{2 \cdot i \cdot k \cdot \pi}{2 \cdot N}\right); \; i = 0,1,.., (N-1); \\ k \in \{1,2,.., (N-1)\}. \end{gathered} \quad (1.32)$$

то кодирование можно осуществить путем использования соотношения:

$$y = \bar{x} + (e - sign(\alpha)) \cdot |\alpha| \cdot \vec{g}, \; \alpha = \langle \bar{x}, \vec{g} \rangle. \quad (1.50)$$

в котором в качестве базисной функции используют гармонический сигнал, представленный на рисунке 6.

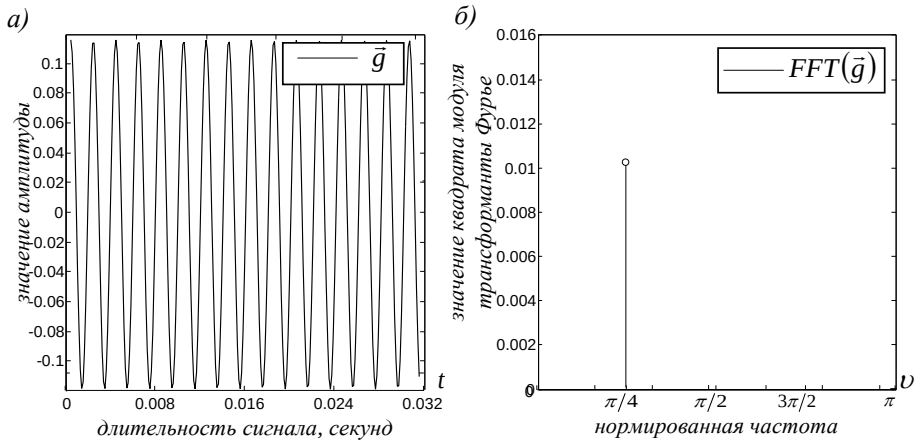


Рисунок 1.13 – Отрезок, соответствующий гармоническому сигналу

Декодирование дополнительной информации, при условии, что при кодировании использовалось дискретно-косинусная трансформация для коэффициента $k \in \{1,2,..,(N-1)\}$ осуществляется следующим образом:

$$\tilde{e} = sign\left(\frac{2}{N}\sum_{i=2}^{N} x_i \cdot \cos\left(\frac{i \cdot k \cdot \pi}{N}\right)\right). \quad (1.33)$$

Особенности использования стеганографического метода кодирования дополнительной информации, основанного на дискретно-косинусной трансформации: изменение происходит на строго определённых частотах; емкость зависит от свойств сигнала и выбранной психоакустической модели.

1.4 Постановка задач исследования

Рассмотрен способ цифрового представления звукозаписи. В звукозаписи устной речи присутствуют паузы, в связи с чем в схеме внедрения цифрового отпечатка предлагается использовать детектор пауз между словами.

Исходя из приведенного выше описания, были сформулированы следующие задачи исследования:

1. Проанализировать методы формирования хэш-функции.

2. Проанализировать методы, обеспечивающие скрытность при кодировании информации в звукозапись.

3. Разработать схему и его программный прототип, реализующий кодирование/декодирование цифрового отпечатка в звукозапись с обеспечением его скрытности.

4. На основе вычислительных экспериментов оценить работоспособность разработанного метода защиты звукозаписи устной речи.

Глава 2. Метод скрытного подтверждения целостности звукозаписи

В главе рассматриваются схемы формирования и внедрения символов цифровых отпечатков.

2.1 Анализ частотно-временных свойств звукозаписей

В подразделе рассматриваются свойства компонент звуков устной речи, с позиции их использования в процессе формирования цифровых отпечатков.

Важным моментом использования устной речи в индустрии является субъективное качество, с которым воспроизводится звукозапись устная речь. Наличие тех или иных частотных компонент в речи формирует её тональную окраску и помогает улучшить эмоциональное восприятие с одной стороны, но приводит к необходимости хранения избыточности, с другой стороны. При этом стоит учитывать, что устная речь представлена совокупностью последовательности звуков. В свою очередь звуки, имеющие идею или смысл, образуют речевое сообщение.

Для анализа свойств фрагменты сигнала, приводят к цифровому виду и разбивают на отрезки данных, представляющие собой одномерные массивы чисел с количеством элементов N. Стоит отметить, что числа имеют при хранении в цифровом виде ограниченную точность, при заданной разрядности данных K (рис 1.3):

$$\vec{x} = \left(x_1, x_2, \ldots, x_n, \ldots, x_N\right)^T,\ n = 1,..., N, \tag{2.1}$$

$$\vec{x} = (f_{(z-1)N+1}, f_{(z-1)N+2}, ..., f_{z \cdot N})^T .,\ z = 1, 2, ..., Z \tag{2.2}$$

где z - порядковый номер отрезка данных; Z - количество отрезков данных принадлежащих, фрагменту речевого сигнала.

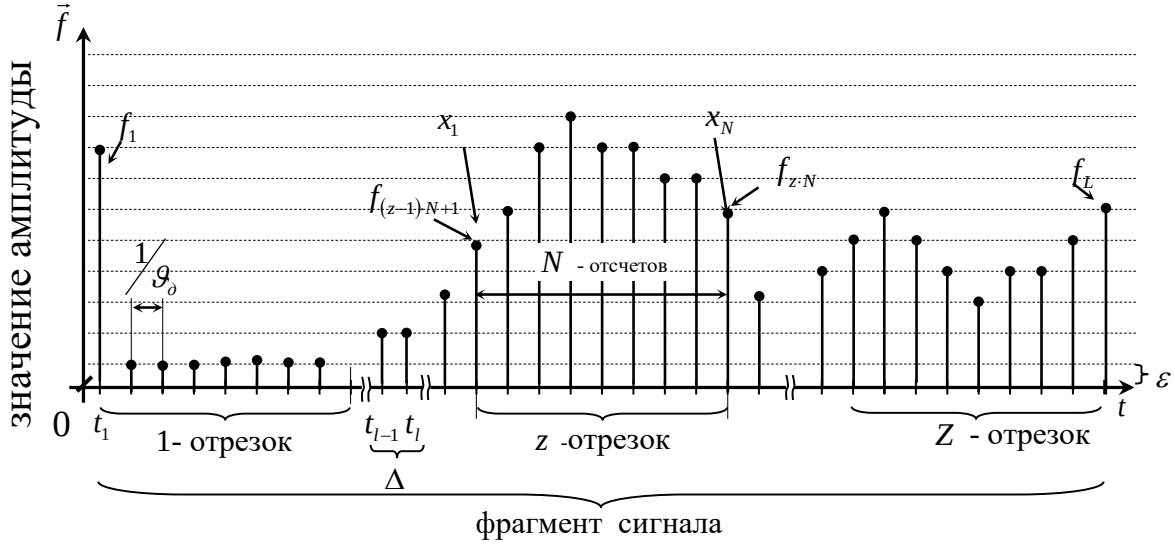


Рисунок 2.1 – Разбиение фрагмента сигнала на отрезки данных

В основе многих методов обработки и анализа сигналов $f(t)$ лежат разложения среди разнообразных систем ортогональных функций $g(t)$ [16, 18, 20, 22, 26-33]. В рамках данной работы используются множество $\mathbf{L}$ функций, в замкнутой области $\mathbf{G}$, для которых в общем случае выполняется соотношение [21]:

$$\langle f(t), g(t) \rangle = \int_{\mathbf{G}} f(t) g(t) dt, \; g(t) \in \mathbf{L}, \; f(t) \in \mathbf{L}. \qquad (2.3)$$

где $\langle , \rangle$ – проекция, результат скалярного произведение двух функций $g(t)$ и $f(t)$.

Стоит отметить, что общем случае результат скалярного произведения есть вещественное число [34].

Сигнал $f(t)$ конечной длительности является функцией времени и, следовательно, можно найти его спектральное разложение среди разнообразных систем ортогональных функций [16, 20, 21, 22, 35, 36]. Наиболее часто используются частотные представления, базисом которых являются гармонические сигналы (например, прямое и обратное преобразование Фурье) [16, 20, 21, 22, 35, 36]:

$$g(t) = e^{-j2\pi\nu t}, \tag{2.4}$$

$$F(\upsilon) = \int_0^T f(t) e^{-j2\pi\nu t} dt, \tag{2.5}$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\nu) e^{j2\pi\nu t} d\nu, \tag{2.6}$$

где $F(\nu)$ - трансформанта Фурье.

Трансформанта Фурье – «коэффициент пропорциональности между длиной малого интервала частот и отвечающей ему комплексной амплитудой гармонического сигнала с частотой υ» [].

В рамках работы символами латинского алфавита обозначаются частотные представления сигналов $F(\nu)$, $X(\nu)$, $Y(\nu)$, $U(\nu)$, $S(\nu)$, $C(\nu)$, $G(\nu)$, $Q(\nu)$. Использование частотных представлений $F(\nu)$ оправдано (адекватно) как с точки зрения воспроизведения, так и ее восприятия [40, 41]. Для унификации описания нормированной частотной области используют круговую частоту, отражающую периодичность изменений сигнала $f(t)$ с изменением времени t (ν – величина обратная периоду):

$$\nu = 1/T, \ \upsilon = 2\pi\nu. \tag{2.7}$$

Во многих алгоритмах исследующих устную речь, применяют методы спектрального анализа в основе которых лежит преобразование Фурье. Согласно разложению в бесконечный ряд Фурье, фрагмент речевого сигнала, можно представить в синус-косинусной форме:

$$x_n = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left(a_k \cdot \cos\left(n \cdot \Delta t \cdot k \cdot \Delta\omega_0\right) + b_k \cdot \sin\left(n \cdot \Delta t \cdot k \cdot \Delta\omega_0\right) \right), \ n = 1, 2, \ldots, N \tag{2.8}$$

где Δt – шаг дискретизации по времени; N – длительность сигнала в отчетах; $\Delta\omega_0$ – шаг дискретизации по частоте; $n \cdot \Delta t$ – абсолютное значение дискретного

времени; $k \cdot \Delta\omega_0$ – абсолютное значение дискретной частоты; a_k , b_k – коэффициенты определяемые соотношениями:

$$a_k = \frac{2}{N}\sum_{n=1}^{N} x_n \cdot \cos\left(n \cdot \Delta t \cdot k \cdot \Delta\omega_0\right),\ n = 1, 2, \ldots, N, \tag{2.9}$$

$$b_k = \frac{2}{N}\sum_{n=1}^{N} x_n \cdot \sin\left(n \cdot \Delta t \cdot k \cdot \Delta\omega_0\right),\ n = 1, 2, \ldots, N, \tag{2.10}$$

где N – длительность сигнала в отчетах:

$$N = \frac{T}{\Delta t}, \tag{2.11}$$

где T – длительность сигнала в секундах.

Дискретность сигнала приводит к ограничению верхней частоты, называемой частотой Найквеста [23, 25, 28]:

$$\vartheta_в = \frac{1}{2 \cdot \Delta t} = \frac{\vartheta_д}{2}. \tag{2.12}$$

Тогда с учетом (1.18-1.19), выражение (1.1) может быть переписано в виде:

$$x_n = \frac{a_0}{2} + \sum_{k=-\vartheta_д/2}^{\vartheta_д/2}\left(a_k \cdot \cos\left(n \cdot \Delta t \cdot k \cdot \Delta\omega_0\right) + b_k \cdot \sin\left(n \cdot \Delta t \cdot k \cdot \Delta\omega_0\right)\right),\ n = 1, 2, \ldots, N. \tag{2.13}$$

В выражение (2.6) входят коэффициенты, определяемые соотношениями (2.2) и (2.3), точность определения которых зависит от количества суммируемых значений амплитуд сигнала, т.е. от длины отрезка сигнала.

Хранимые фрагменты сигнала содержат устную речь конечной длины, поэтому при анализе спектра с использованием математического аппарата опирающегося на соотношения (2.2), (2.3), (2.6) и (2.7) стоит учитывать то, что операции проводимые над сигналом малой длительности могут давать погрешности при определении спектра, т.е. нахождение значений модуля трансформант Фурье X_k зависит от длительности сигнала:

$$\mathrm{X}_k = \sqrt{a_k^2 + b_k^2},\ k = \left(-\vartheta_д/2\right), \ldots, -1, 0, 1, \ldots, \left(\vartheta_д/2\right), \tag{2.14}$$

Так же от длительности сигнала зависит первое слагаемое в разложении Фурье (2.1), которое содержит среднее значение сигнала:

$$\frac{a_0}{2} = \frac{1}{N}\sum_{n=1}^{N} x_n \,. \tag{2.15}$$

Коэффициенты действительной части в приведенном ранее разложении Фурье (2.6), можно найти используя скалярное произведение:

$$a_k = \langle \vec{x}, \vec{c}_k \rangle, \tag{2.16}$$

где $\vec{c}$ - функция косинуса:

$$\vec{c} = \cos(n \cdot \Delta t \cdot k \cdot \Delta\omega_0), \; n = 1, 2, \ldots, N \,. \tag{2.17}$$

Коэффициенты мнимой части в приведенном ранее разложении Фурье, можно найти используя скалярное произведение:

$$b_k = \langle \vec{x}, \vec{s}_k \rangle \tag{1.26}$$

где $\vec{c}$ - функция синуса:

$$\vec{s} = \sin(n \cdot \Delta t \cdot k \cdot \Delta\omega_0), \; n = 1, 2, \ldots, N \,. \tag{1.27}$$

Учитывая (2.9) и (2.11), выражение (2.7), можно представить в виде:

$$\mathrm{X}_k = \sqrt{\left(\langle \vec{x}, \vec{c}_k \rangle\right)^2 + \left(\langle \vec{x}, \vec{s}_k \rangle\right)^2} \,, \; k = \left(-\vartheta_\partial/2\right), \ldots, -1, 0, 1, \ldots, \left(\vartheta_\partial/2\right) \tag{2.18}$$

При анализе формулы (2.13) с учетом процедуры формирования векторов (2.10) и (2.12), можно заметить, что если частоты гармонического сигнала точно не совпадут с частотной компонентой в исследуемом сигнале, то разложение (2.18) будет проведено с некоторой погрешностью. Эта погрешность повлияет на вычисление спектра (2.13), т.к. от точности выбора

дискретной частоты зависит результат скалярного произведения. Выше описанное приводит к эффекту растекания спектра (рисунок 2.1 б).

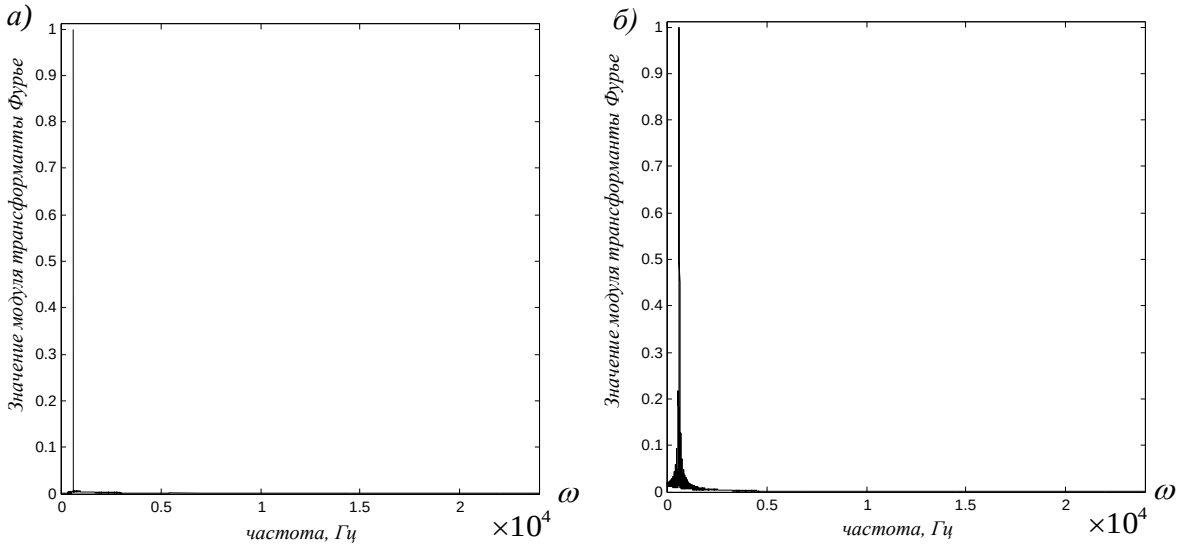


Рисунок 2.2 – Спектр функции косинуса на центральной частоте f_0=1кГц: а) в случае совпадения узлов сетки дискретных частот; б) в случае не совпадения узлов сетки дискретных частот

Ниже представлены блок-схема (рисунок 1.12) и словесное описание алгоритма определения модулей трансформант Фурье. Алгоритм предусматривает определение модулей трансформанты Фурье для дискретных значений частоты в положительной части спектра. Шаг дискретизации по частоте равен 1 Гц, для того чтобы уменьшить эффект растекания.

Словесное описание алгоритма расчета модулей трансформант Фурье.

Ввод: отрезок данных $x=\left(x_1,x_2,\ldots,x_n,\ldots,x_N\right)^{\mathrm{T}}$.

Вывод: результаты расчета модулей трансформант Фурье: $\vec{X}=\left(X_1,X_2,\ldots,X_k,\ldots,X_{\mathrm{K}}\right)^{\mathrm{T}}$.

1. Ввести отрезок данных $x=\left(x_1,x_2,\ldots,x_n,\ldots,x_N\right)^{\mathrm{T}}$.

2. Определить параметры цифрового представления:

– частота дискретизации: ϑ_0 .

– разрядность данных: B.

3. Вычислить шаг дискретизации по частоте $\Delta\omega = 2\pi/\vartheta_д$.

4. Задать начальное значение нормированной частоты: $k = 1$.

5. Найти абсолютное значение частоты: $\omega_k = \Delta\omega \cdot (k\text{-}1)$.

6. Задать начальное значение нормированного дискретного времени: $i = 1$.

7. Вычислить значение функции косинуса в точке дискретного времени i, для абсолютное значение частоты ω_k: $c_k(i) = \cos(\omega_k \cdot (i\text{-}1))$.

8. Вычислить значение функции синуса в точке дискретного времени i, для абсолютное значение частоты ω_k: $s_k(i) = \sin(\omega_k \cdot (i\text{-}1))$.

9. Увеличить дискретное время: $i = i + 1$.

10. Проанализировать превышает ли дискретное время i длительность исследуемого сигнала N:

– если не превышено, т.е. $i < N$, то перейти к п. 9.

– если превышено, т.е. $i > N$, то перейти к п. 13.

11. Увеличить значение нормированной частоты: $k = k + 1$.

12. Проанализировать превышает ли значение нормированной частоты k верхнее значение частоты $\vartheta_д/2$:

– если не превышено, т.е. $\omega_k < \vartheta_д/2$, то перейти к п. 7.

– если превышено, т.е. $\omega_k > \vartheta_д/2$, то перейти к п. 15.

13. Сохранить результаты расчета гармонических сигналов: $\vec{c}_k$, $\vec{s}_k$.

14. Задать начальное значение нормированной частоты: $k = 1$.

15. Определить проекцию для четной функции частоты: $a_k = \langle \vec{x}, \vec{c}_k \rangle$.

16. Определить проекцию для нечетная функции частоты: $b_k = \langle \vec{x}, \vec{s}_k \rangle$.

17. Определить модуль трансформанты Фурье: $X_k = \sqrt{a_k^2 + b_k^2}$.

18. Увеличить значение нормированной частоты: $k = k + 1$.

19. Проанализировать превышает ли значение частоты ω_k верхнее значение частоты $\vartheta_д/2$:

– если не превышено, т.е. $\omega_k < \vartheta_д/2$, то перейти к п. 15.

– если превышено, т.е. $\omega_k > \vartheta_д/2$, то перейти к п. 20.

20. Сохранить результаты расчета модулей трансформант Фурье: X_k.

21. Конец.

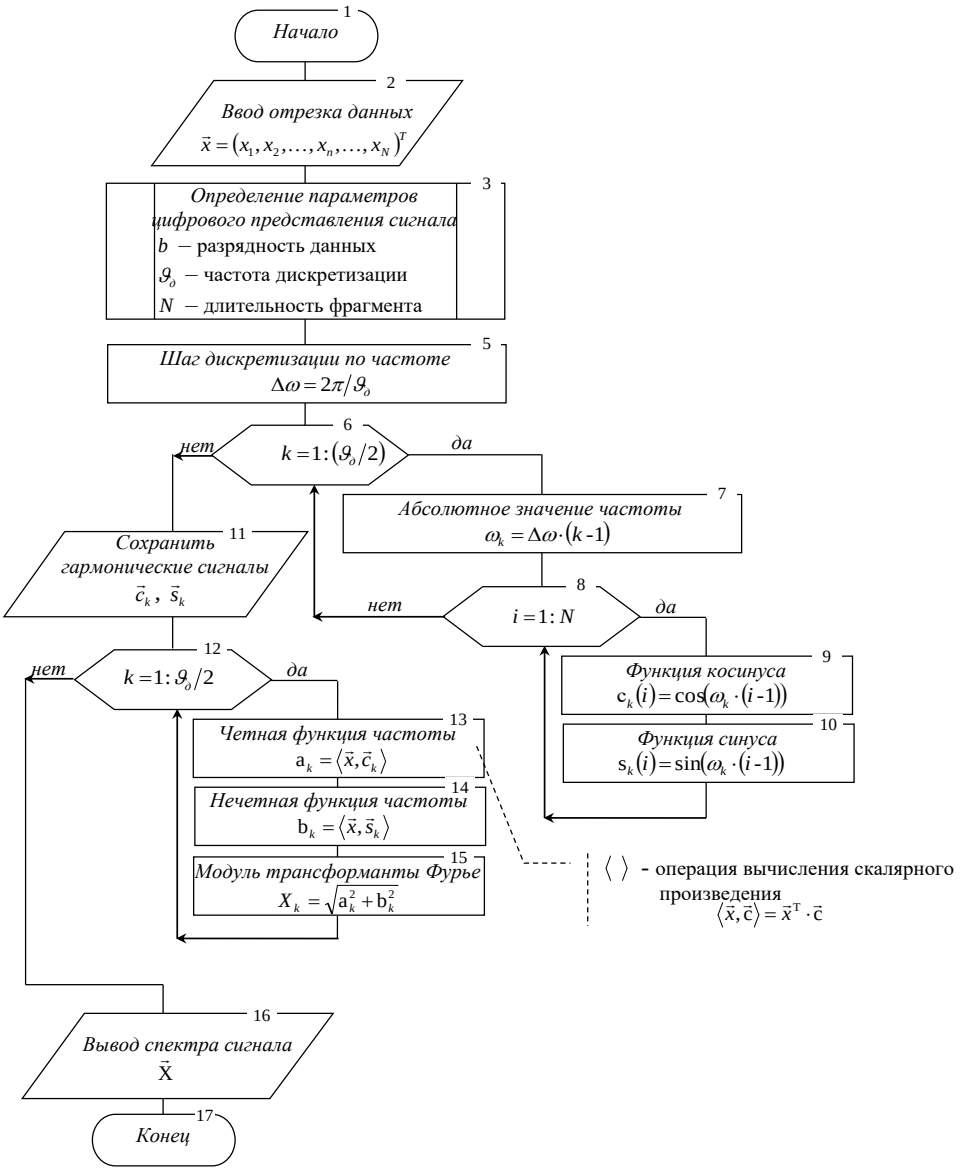


Рисунок 2.3 – Блок-схема алгоритма оценивающего спектр используя расчет модулей трансформант Фурье

Ниже приведены результаты оценки спектра для фрагментов речевых сигналов, порожденных различными звуками русской речи одного диктора (диктор мужчина, частота дискретизации $\vartheta_д$=48кГц, разрядность 16 бит, длительность 32мкс).

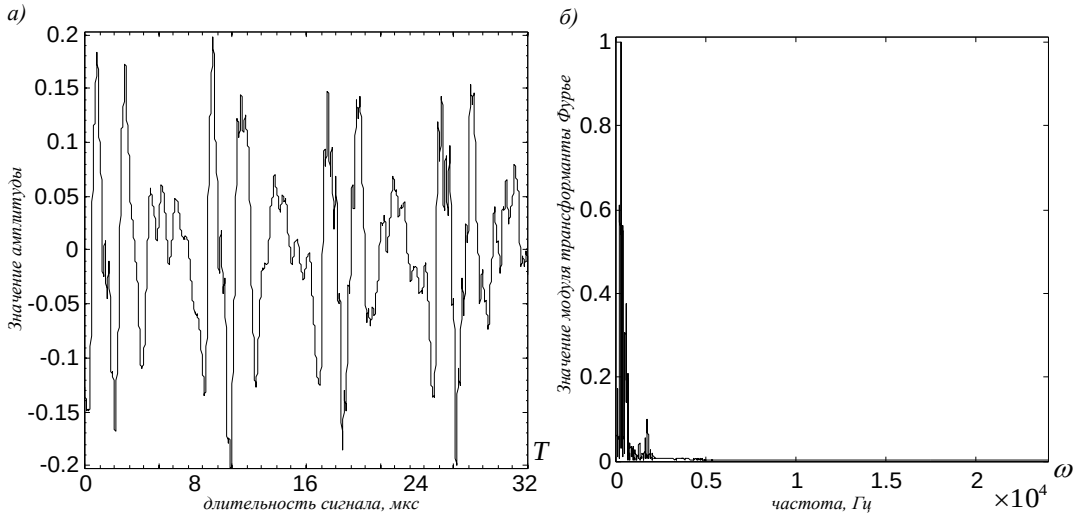


Рисунок 2.4 – Фрагмент речевого сигнала (а) порожденного звуком «а» и его спектр (б), (звук взят из слова «Артист», частота дискретизации $\vartheta_д$ =48 кГц, разрядность B=32 бита, длительность фрагмента сигнала T=32мс)

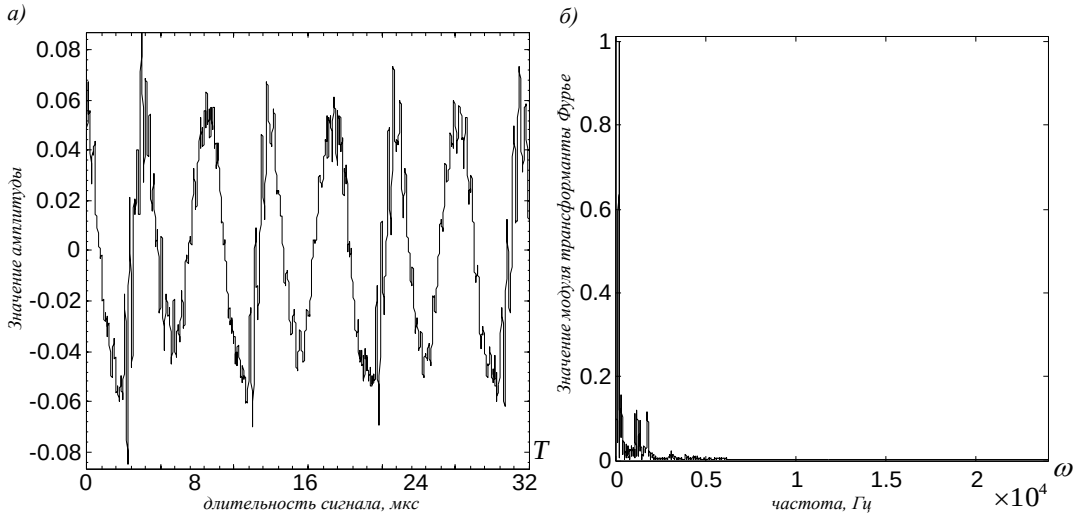


Рисунок 2.5 – Фрагмент речевого сигнала (а) порожденного звуком «и» и его спектр (б), (звук взят из слова «Идут», частота дискретизации $\vartheta_д$=48 кГц, разрядность B=32 бита, длительность фрагмента сигнала T=32мс)

Как видно из спектров (рисунки 2.3-2.4 часть б, приложение А), для частотных компонент звуков, соответствующих буквам «а», «б», «в», «г», «д», «е», «и», «к», «л», «м», «н», «у», «э», визуально можно выделить период.

Речевой сигнал звуков, соответствующих буквам: «ж», «з», «с», «ф», «ц», «ч», похож на шум (рисунки 2.5-2.6 часть б, приложение А).

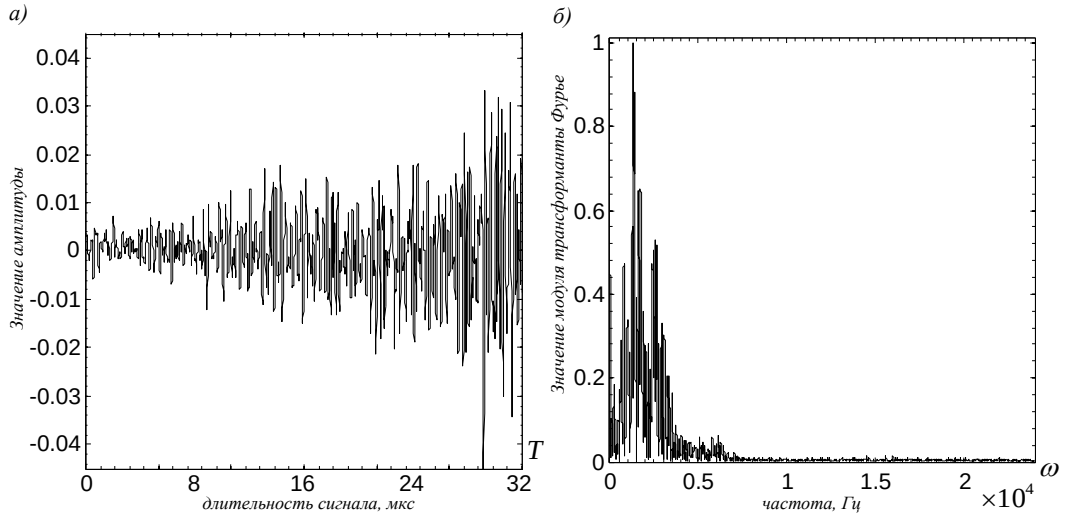


Рисунок 2.6 – Фрагмент речевого сигнала (а) и его спектр (б), порожденного звуком «ш», слово «Штурман»

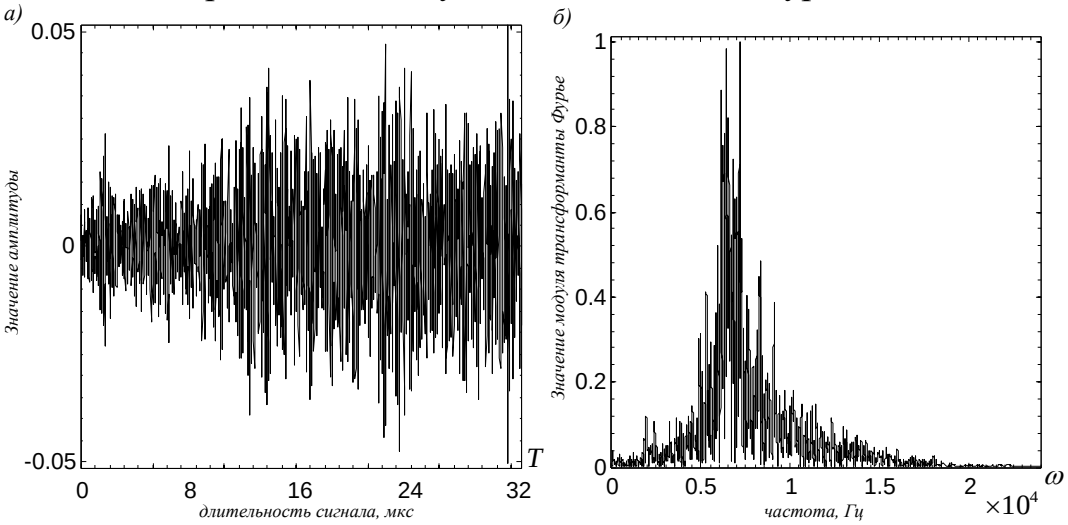


Рисунок 2.7 – Фрагмент речевого сигнала (а) и его спектр (б), порожденного звуком «с», слово «Срочно»

Речевой сигнал звуков, соответствующих буквам: «к», «н», «р», «т», имеют сложную структуру в которой кроме периода наблюдается модуляция по амплитуде (рисунки 2.7-2.8 часть б).

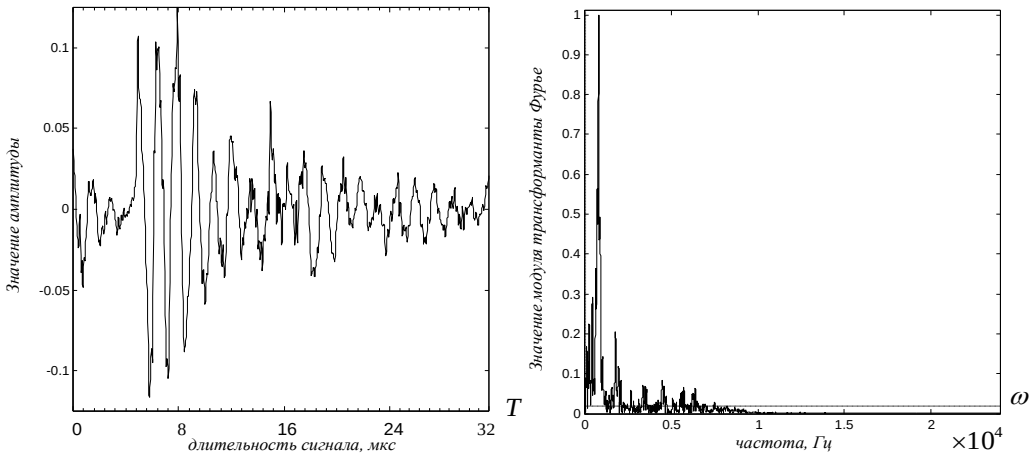


Рисунок 2.8 – Фрагмент речевого сигнала (а) порожденного звуком «К» и его спектр (б)

(звук взят из слова «кОмандир», частота дискретизации $\vartheta_д$=48 кГц, разрядность B=32 бита, длительность фрагмента сигнала T=32мс)

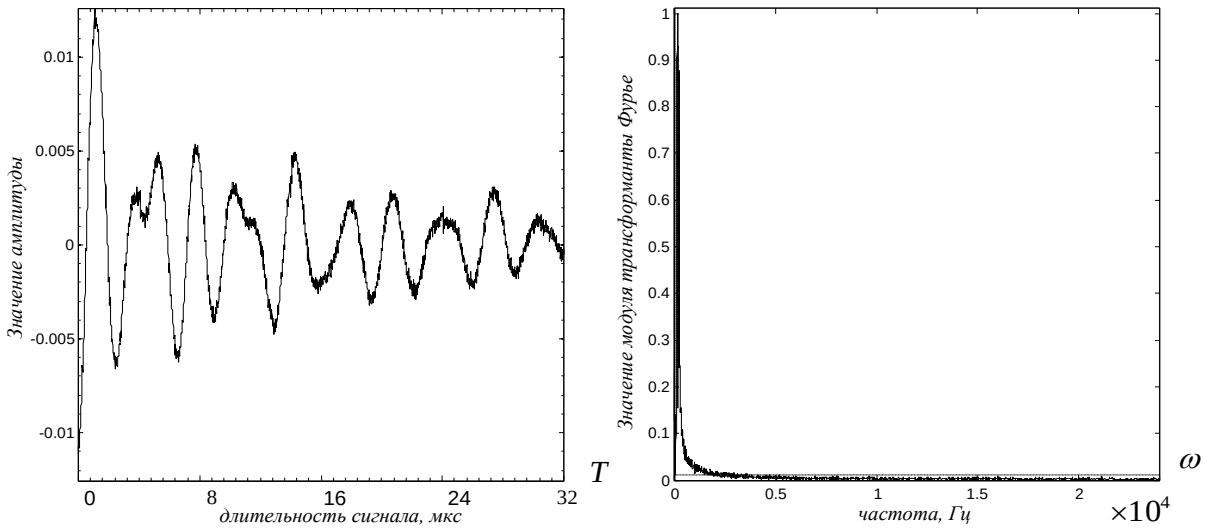


Рисунок 2.9 – Фрагмент речевого сигнала (а) порожденного звуком «Т» и его спектр (б)

(звук взят из слова «разворОт», частота дискретизации $\vartheta_д$=48 кГц, разрядность B=32 бита, длительность фрагмента сигнала T=36мс)

Частотные компоненты, энергия которых больше средней (2.14), речевого сигнал соответствующие большинству звуков русской речи занимают узкую полосу в области низких частот, в частности:

– частотные компоненты для звуков соответствующие буквам: «в», «г», «н», «у», расположены до частоты 500 Гц;

– частотные компоненты для звуков соответствующие буквам: «б», «т», «р», расположены до частоты 1 кГц;

– частотные компоненты для звуков соответствующие буквам: «а», «е», «и», «о», «э», «ю», «я», «д», «л», «м», «п», расположены до частоты 2 кГц;

– частотные компоненты для звуков соответствующие буквам: «ж», «з», «ц», «ш», расположены не ниже частоты 6 кГц;

– частотные компоненты для звуков соответствующие буквам: «с», «ф», расположены не ниже частоты 10 кГц иногда до частоты 16 кГц.

При этом стоит отметить, что звуки, соответствующие буквам русского языка, имеют разную длительность. Длительность одного и того же звука может меняться не только от слова к слову или его позиции в слове, но и зависит от темпа речи диктора, от его интонации.

Из проведенных экспериментов при изменении длительности отрезков речевого сигнала меняется спектр исследуемых отрезков, что связано с перераспределением энергии между частотными компонентами (растеканием спектра). Растекание спектра, приводит к уменьшению количества частотных компонент, имеющих энергию меньше средней. Среднее значение модуля трансформанты Фурье.

На рисунке 1.9 б, г приведены спектры, рассчитанные соответственно для отрезков речевых сигналов рисунке 2.9 а, в. Уменьшение длительности сигнала в 4раза привело к изменению среднего значение модуля трансформанты Фурье в два раза (на рисунке среднее значение отмечено пунктиром). При этом можно отметить уменьшение частотных полос, содержащих частотные компоненты, энергия которых меньше средней.

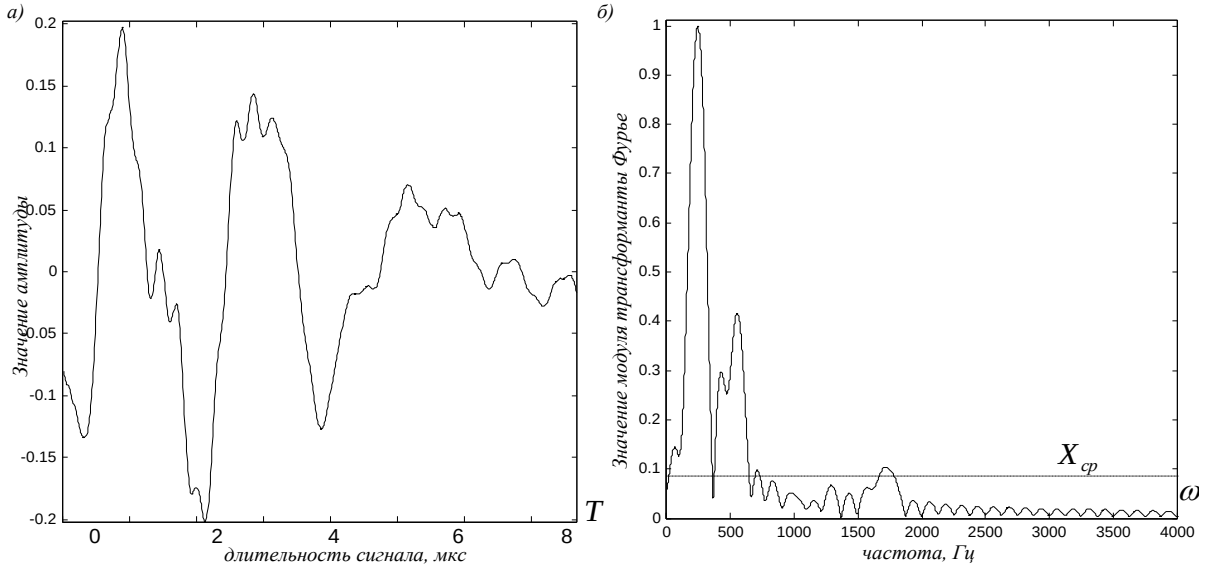


Рисунок 2.10 – Фрагмент речевого сигнала (а) порожденного звуком «а» и его спектр (б), (звук взят из слова «Артист», частота дискретизации $\vartheta_д$ =48 кГц, разрядность B =32 бита)

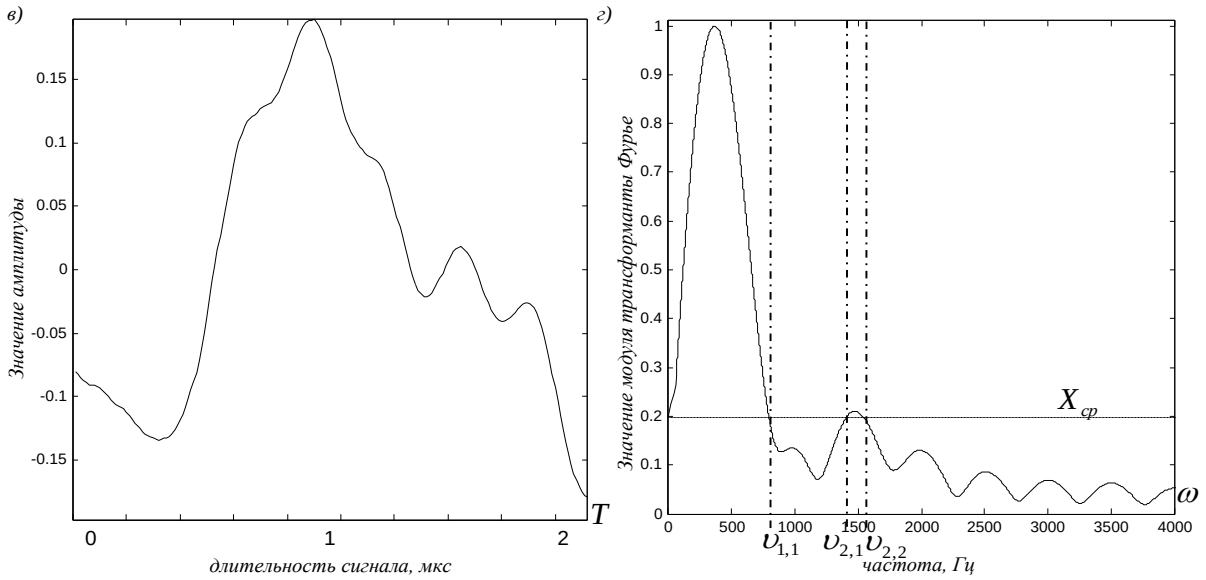


Рисунок 2.11 – Фрагмент речевого сигнала (а) порожденного звуком «а» и его спектр (б), (звук взят из слова «Артист», частота дискретизации $\vartheta_д$ =48 кГц, разрядность B =32 бита)

Наличие частотных областей с частотными компонентами имеющих большую энергию и малую энергию (т.е. компоненты больше и меньше средней), имеет практическую значимость для скрытого кодирования информации, так как частотные компоненты с большой энергией могу

маскировать частотные компоненты с малой энергией. То есть скрытое кодирование дополнительной информации предполагает наличие в анализируемом отрезке частотных компонент с малой и большой долей энергией. Следовательно, для осуществления скрытого кодирования, в спектре необходимо наличие не менее двух частотных полос с разной энергией. На рисунке 1.9 г приведен пример в котором частотная полоса от 0 Гц до $\upsilon_{1,1}$= 850 Гц и частотная полоса $\upsilon_{2,1}$= 1400 Гц до $\upsilon_{2,2}$= 1550 Гц, содержит частотные компоненты больше среднего, а частотные полосы с частотами от $\upsilon_{1,1}$= 850 Гц до $\upsilon_{2,1}$= 1400 Гц и выше $\upsilon_{2,2}$= 1550 Гц, содержат частотные компоненты меньше среднего.

В связи с тем, что точность определения энергии частотных компонент в спектре сигнала, зависит от его длительности, то необходимо провести исследование определяющее минимальную длительность, при которой возможно разделение на полосы. Ниже представлены блок-схема (рисунок 1.10) и словесное описание алгоритма оценивания величины частотной концентрации отрезка сигнала.

Вычислительный эксперимент

Цель: определение минимальной длительности отрезка речевого сигнала, при котором возможно разделение спектра на частотные полосы.

План:

1) Загрузить речевой сигнал, соответствующий букве русского языка.

2) Задать параметры цифрового представления речевого сигнала.

3) Задать начальную длительность анализируемого речевого сигнала.

4) Найти спектр речевого сигнала, состоящий из модулей трансформант Фурье.

5) Найти среднее значение спектральной компоненты.

6) Определить границы частотных полос, содержащих частотные компоненты, которые меньше среднего значения.

7) Определить ширину частотных полос.

8) Определить долю энергии, содержащуюся в частотной полосе.

9) Усреднить исследуемые характеристики для каждой буквы и сравнить

полученные результаты при различных значениях длительности речевого сигнала.

Исходные данные:

1) Речевые сигналы, порожденные различными звуками русской речи, записанные с частотой дискретизации $\vartheta_д$ =48кГц, и разрядностью кодовой последовательности 32 бита.

Словесное описание алгоритма оценки частотного распределения энергии отрезка сигнала.

1. Загрузить речевой сигнал длительностью не менее $T_1 = 2/\vartheta_д$ (шаг изменения времени $\Delta t_1 = T_1/2$, номер шага итерации $i=1$).

2. Задать параметры цифрового представления речевого сигнала:

- частота дискретизации $\vartheta_д$ =48кГц;

- разрядность кодовой последовательности В =32 бита.

3. Получить цифровое представление речевого сигнала, для заданной длительности T_i: $\vec{x} = (x_1, \ldots, x_k, \ldots, x_N)^T$, где $N = T_i \cdot \vartheta_д$.

4. Найти спектр речевого сигнала, состоящий из модулей трансформант Фурье: $\vec{X} = (X_1, \ldots, X_k, \ldots, X_{\vartheta_в})^T$, где $\vartheta_в = \vartheta_д/2$.

5. Найти среднее значение спектральной компоненты: $X_{ср} = \frac{1}{\vartheta_в} \sum_{k=1}^{\vartheta_в} X_k$.

6. Задать номер частотной полосы с модулями трансформант Фурье энергия которых меньше средней: $r=1$.

7. Задать начальную ширину частотной полосы с модулями трансформант Фурье энергия которых меньше средней: $\Delta\Omega_r = 0$.

8. Задать начальную энергию для частотной полосы с модулями трансформант Фурье энергия которых меньше средней: $\mathrm{P}_r = 0$.

9. Задать верхнюю частоту частотной полосы с модулями трансформант Фурье энергия которых меньше средней: $\upsilon_2(r) = 0$.

10. Задать начальное значение частоты: $\omega_k = 1$ ($k = 1$).

11. Проверить превышает ли модуль трансформанты Фурье на текущей частоте X_k среднее значение X_{cp}

– если не превышено, т.е. $X_k < X_{cp}$, то перейти к п. 12.

– если превышено, т.е. $X_k > X_{cp}$, то перейти к п. 19.

12. Проверить верность условия: «текущую частоту ω_k можно причислить к интервалу r, т.к. она является следующим значением дискретной сетки частот, т.е. $\omega_k = (\upsilon_2(r)+1)$»:

– условие выполняется, т.е. $\omega_k = (\upsilon_2(r)+1)$ и модуль трансформанты Фурье принадлежит текущей частотной полосе, то перейти к п. 16.

– условие не выполняется, т.е. $\omega_k \neq (\upsilon_2(r)+1)$ и модуль трансформанты Фурье принадлежит следующей частотной полосе, то перейти к п. 13.

13. Увеличить номер частотной полосы с модулями трансформант Фурье энергия которых меньше средней: $r = r + 1$.

14. Задать ширину частотной полосы r с модулями трансформант Фурье энергия которых меньше средней: $\Delta\Omega_r = 0$.

15. Задать начальную энергию частотной полосы r с модулями трансформант Фурье энергия которых меньше средней: $\mathrm{P}_r = 0$.

16. Задать верхнюю частоту частотной полосы r с модулями трансформант Фурье энергия которых меньше средней: $\upsilon_2(r) = \omega_k$.

17. Увеличить ширину частотной полосы $\Delta\Omega_r$ с модулями трансформант Фурье, энергия которых меньше средней: $\Delta\Omega_r = \Delta\Omega_r + \Delta\omega = \Delta\Omega_r + 1$ (в связи с

тем, что модули трансформант Фурье рассчитывался с использованием дискретной сетки частот имеющий шаг 1 Гц, то шаг $\Delta\omega = 1$).

18. Увеличить энергию частотной полосы r с модулями трансформант Фурье энергия которых меньше средней: $\mathrm{P}_r = \mathrm{P}_r + \left(\mathrm{X}_k\right)^2$.

19. Увеличить значение частоты: $\omega_k = \omega_k + \Delta\omega$ (т.е. $k = k + 1$).

20. Проанализировать превышает ли значение частоты ω_k верхнее значение частоты $\vartheta_o/2$:

– если не превышено, т.е. $\omega_k < \vartheta_o/2$, то перейти к п. 11.

– если превышено, т.е. $\omega_k > \vartheta_o/2$, то перейти к п. 21.

21. Сохранить количество интервалов: $R = r$.

22. Записать значения: P_r, $r = 1, 2, \ldots, R$; $\Delta\Omega_r$ $r = 1, 2, \ldots, R$; $\upsilon_2(r)$ $r = R$.

23. Рассчитать шаг изменения времени: $\Delta t_{i+1} = \Delta t_i / 2$.

24. Увеличить шаг итерации: $i = i + 1$.

25. Проанализировать количество интервалов содержащих модули трансформант Фурье энергия которых меньше среднего:

– если количество интервалов $R > 2$, то уменьшить длительность $T_i = T_{i-1} - \Delta t_i$, выполнить п.3-п.25;

– если количество интервалов $R = 2$, то перейти к п. 26.

– если количество интервалов $R = 1$, то увеличить длительность $T_i = T_{i-1} + \Delta t_i$, выполнить п.3-п.25

26. Записать минимальное значение длительности речевого сигнала T_{i}, при котором можно осуществить скрытое кодирование.

27. Конец.

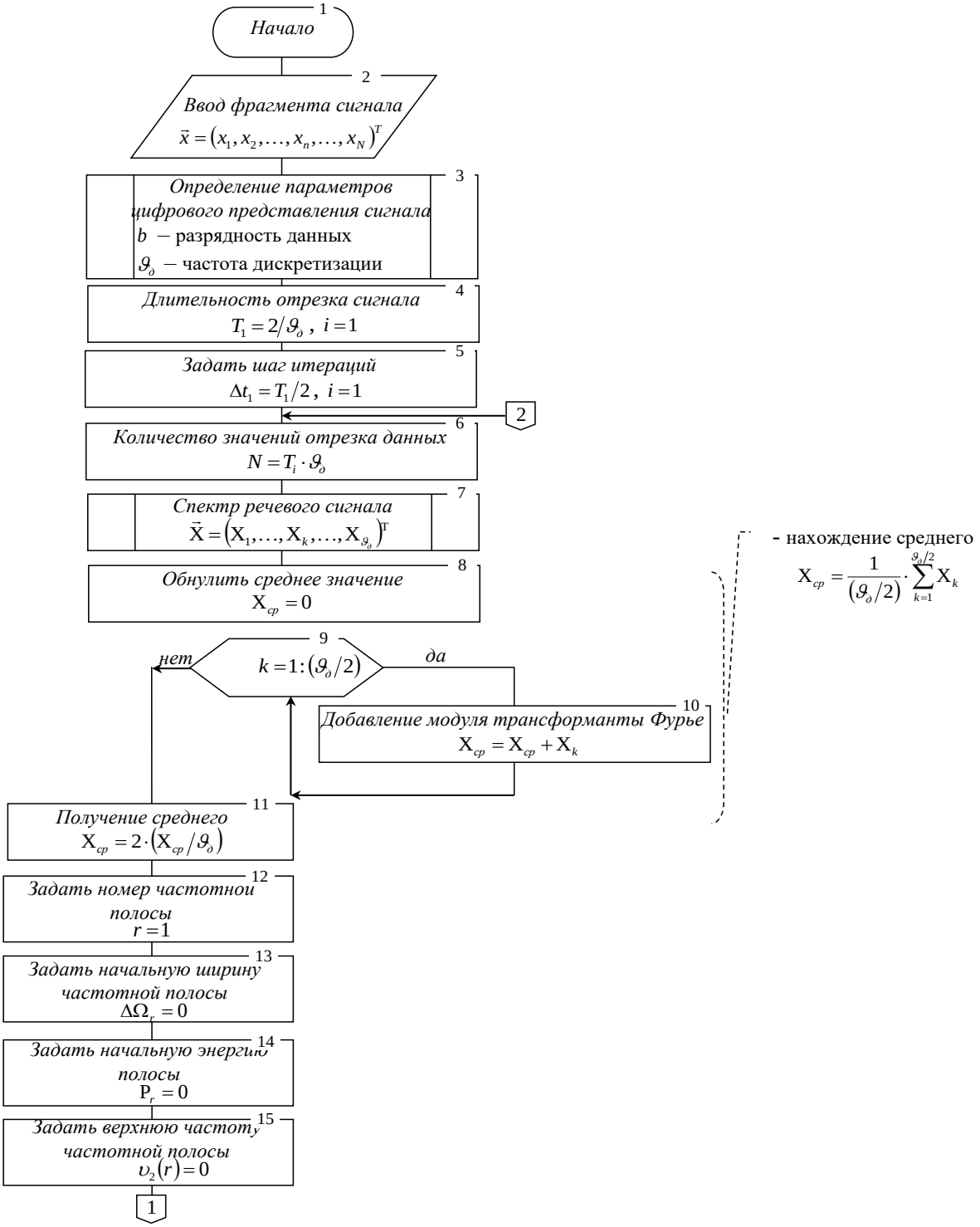


Рисунок 2.12 – Блок-схема алгоритма поиска минимальной длительности

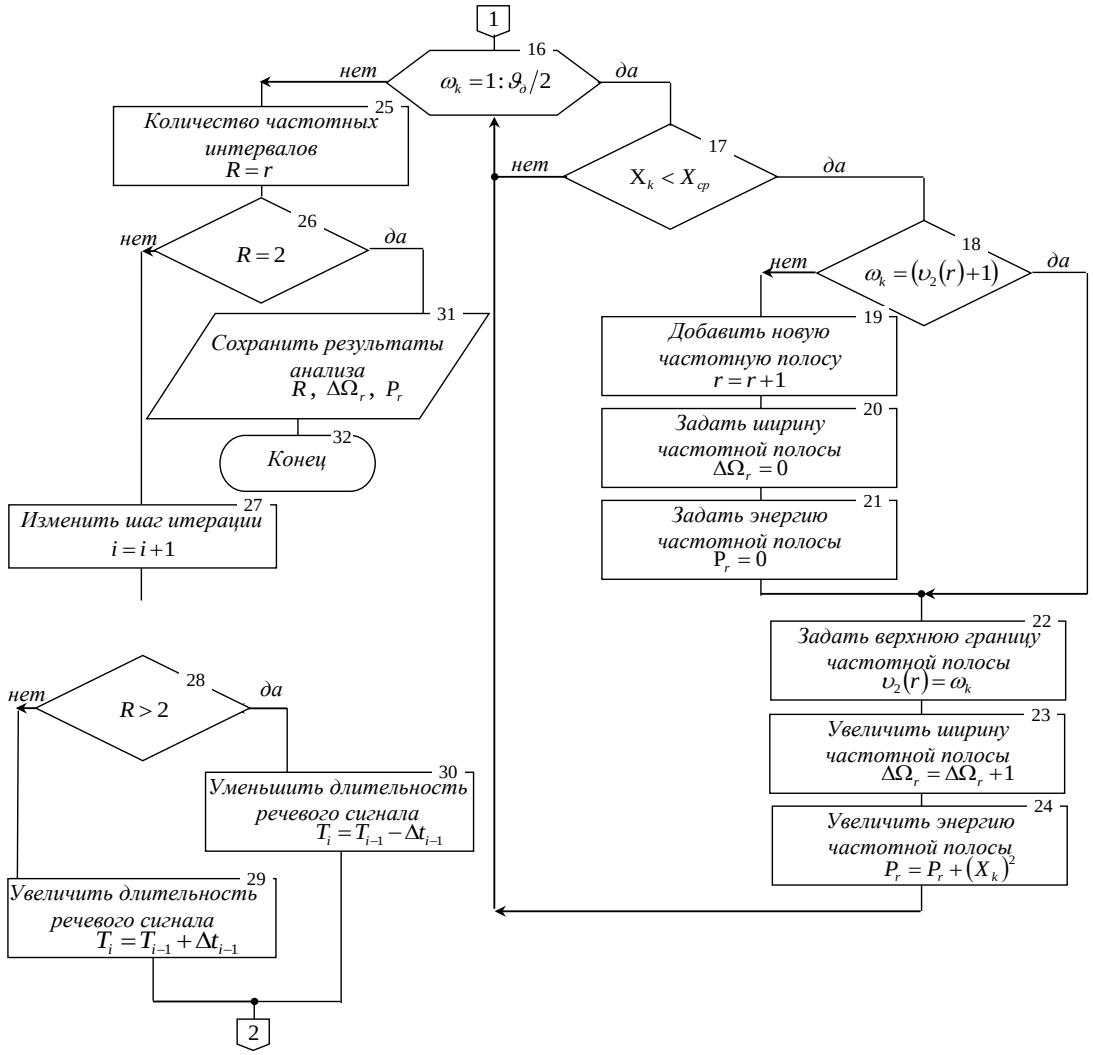


Рисунок 2.13 – Блок-схема алгоритма
поиска минимальной длительности (окончание)

2.2 Алгоритм формирования контрольной информации

приведено описание алгоритма формирования хеш-последовательности из отрезка звукозаписи устной речи. Отмечено, что не стационарность сигнала и соответственно слабая корреляция получаемых из сигнала данных, являющееся звукозаписью устной речи позволяет избежать коллизий. При этом для формирования хеш-последовательности, предлагается использовать не все разряды амплитуды сигнала, а только 50% старших. Предложение позволяет уменьшить количество ошибок при идентификации.

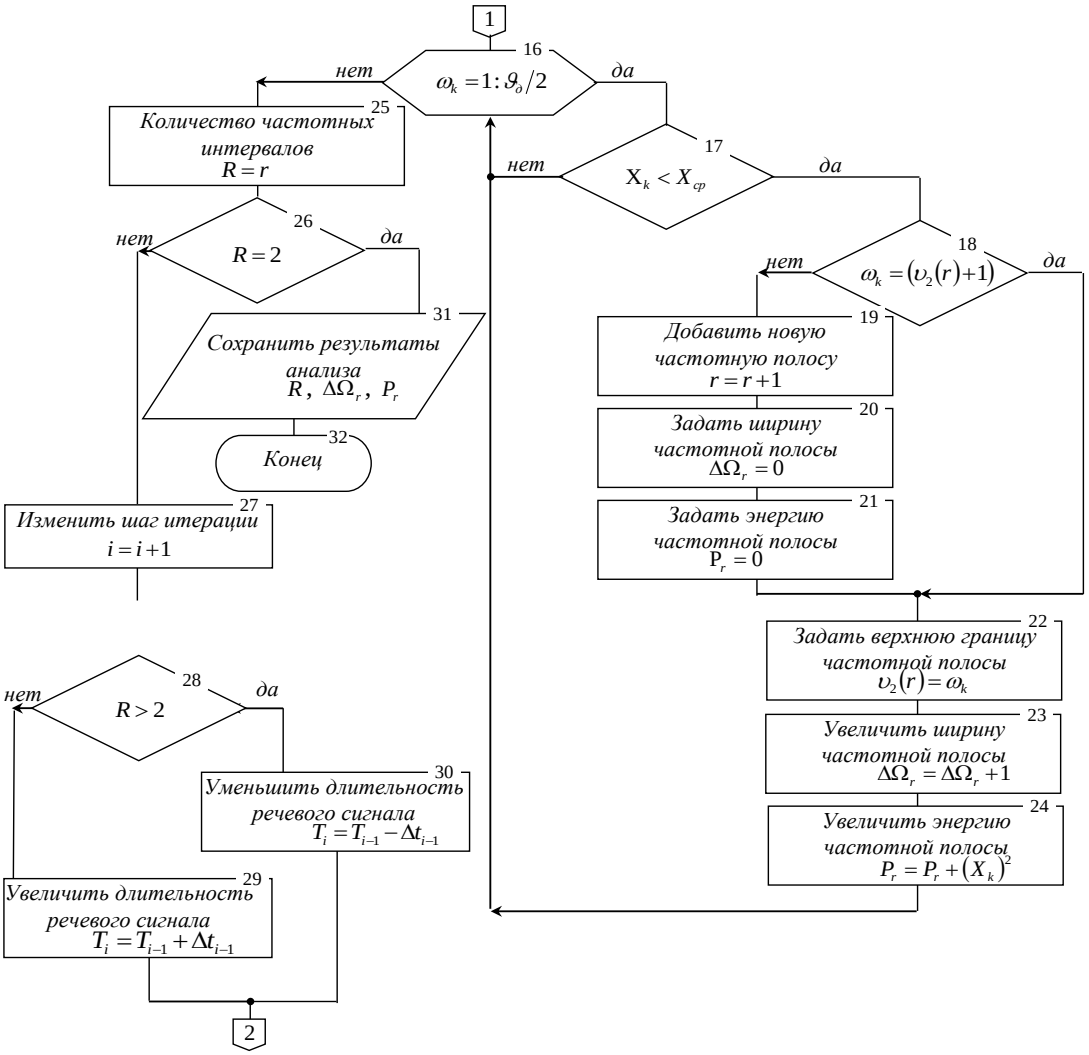


Рисунок 2.13 – Блок-схема алгоритма
поиска минимальной длительности (окончание)

2.2 Алгоритм формирования контрольной информации

Опишем схему формирования хеш-последовательности из отрезка звукозаписи устной речи. Отмечено, что не стационарность сигнала и соответственно слабая корреляция получаемых из сигнала данных, являющееся звукозаписью устной речи позволяет избежать коллизий. При этом для формирования хеш-последовательности, предлагается использовать не все разряды амплитуды сигнала, а только 50% старших. Предложение позволяет уменьшить количество ошибок при идентификации.

2.2.2 Помехоустойчивое кодирование контрольной информации

Коды БЧХ предназначены для исправления независимых ошибок кратности не менее. Их описывают с помощью корней порождающего многочлена G(x). В качествекорней G(x) выбирают последовательных элементов из поля Галуа GF(pm), где 1≤j≤pm, часто принимают j=1, тогда в качестве корней выбирают элементы поля ε, ε2,... Для примитивных двоичных кодов БЧХ длина комбинации равна 32. В этом случае все не нулевые элементы поля Галуа GF(2m) εi будут корнями двучленного уравнения. Но, как следует из теории полей Галуа, каждый из элементов поля εi будет корнем некоторого неприводимого многочлена mi(x) меньшей степени, т.em(εi)=0. Многочлены mi(x) называют минимальными функциями, все они входят в разложение двучлена на неприводимые сомножители. Порождающий многочленG(x) должен быть наименьшим общим кратным всех минимальных функций, т. е. G(x) = НОК[ml(x), m2(x),..., m2t(x)]. Таким образом, вектор, представленный многочленом f(x),будет кодовым тогда и только тогда, когда он делится без остатка как на каждую из минимальных функций m2(x), ..., так и на их наименьшее общее кратное. Тогда для любого из корнейε, ε2, ..., ε2tсправедливо уравнение, которое можно записать в виде произведения двух матриц. Но так как корнями f(x) должны быть все элементы ε, ε2, ..., то можно сделать вывод, что вектор [c0c1...cn–1] будет кодовым тогда и только тогда, когда он принадлежит нулевому пространству матрицы. Из свойств полей следует, что если εi корень какой-либо минимальной неприводимой по mod2 функции mi(x) степени k, то остальными корнями будут. Тогда при определении порождающего многочлена G(x) и проверочной матрицы кода. Должны быть учтены только нечетные степени, т.е. G(x)= m1(x)m3(x) ... m2.

Схема алгоритма формирования кода БЧХ приведена на рисунке 2.14.

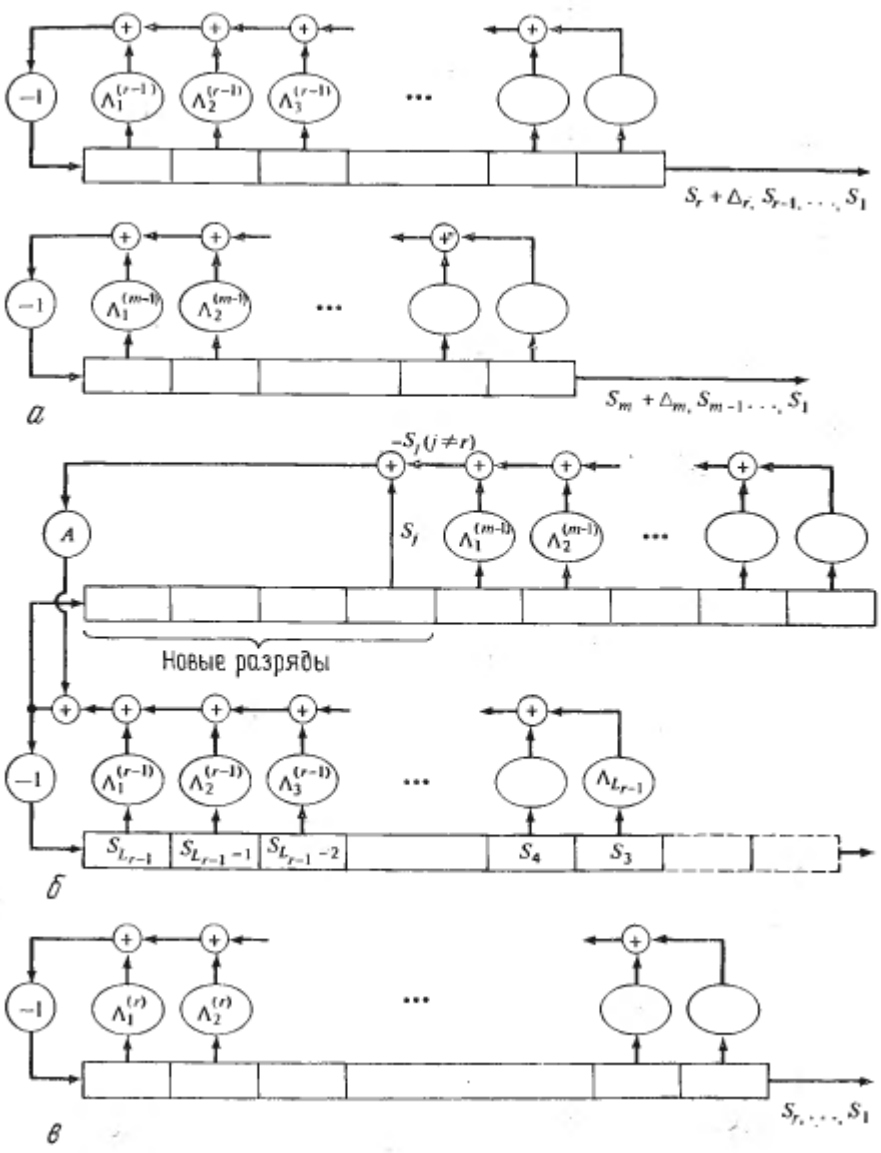


Рисунок 2.14 – Схема алгоритма формирования кода БЧХ

2.3 Алгоритм скрытного подтверждения целостности звукозаписи

2.3.1 Синтез защищенной звукозаписи

Обобщающее решение схемы формирования (рис. 2.15) и внедрения цифровых отпечатков

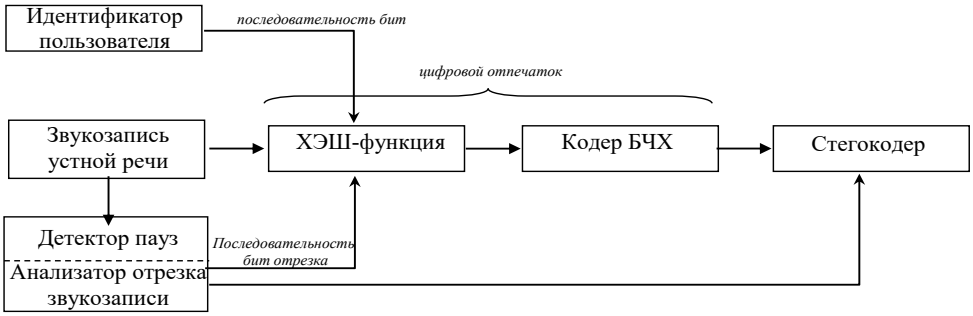


Рисунок 2.15 – Структурная схема внедрения цифрового отпечатка

Приведена схема восстановления внедрённых цифровых отпечатков. Производится описание разработанной схемы и проведены основные алгоритмы и схемы формирующие цифровой отпечаток.

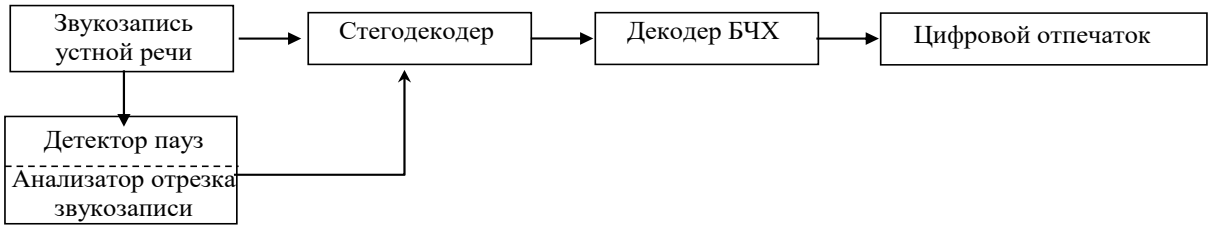


Рисунок 2.16 – Структурная схема извлечения цифрового отпечатка

Приведенные выше схемы показывают взаимосвязь всех используемых подходов, при построении схемы формирования и внедрения символов цифровых отпечатков.

2.4 Основные результаты и выводы главы

Разработана структурная схема внедрения цифрового отпечатка и структурная схема извлечения цифрового отпечатка из звукозаписи устной речи.

Проведен анализ звукозаписей устной речи с позиции возможности их защиты на основе предложенного метода.

Глава 3. Исследование работоспособности метода подтверждения целостности звукозаписи

3.1 Оценка скрытности контрольной информации

Рассмотрим схему внедрения и параметров, влияющих на скрытность внедряемой информации.

Исходные данные: В качестве речевого материала, была использована база речевого корпуса [45], в которой были взяты размеченные звукозаписи устной речи (т.е. определённые автором звукозаписи паузы и звуки, а также параметры звукозаписи: с частота дискретизации 8кГц и разрядность амплитуды 16 бит). Общая длительность речевого корпуса включая паузы составляет два с половиной часа.

Методы анализа: в работе использованы методы цифровой обработки сигналов, статистического анализа и вычислительных экспериментов.

Цель эксперимента: оценить секретность цифрового отпечатка, при прохождении всех процедур внедрения.

Выходные данные: усредненные по всему корпусу звукозаписей оценки скрытности.

Результаты вычислительных экспериментов по внедрению цифрового отпечатка, с обеспечением скрытности, приведены в табл. 1.

Таблица 3.1 - Оценки скрытности цифрового отпечатка

№	Метод	*MSE*	*SNR*	ρ
1	Расширения спектра (SSp)	0.142	19.9	0.873
2	Метод собственных векторов (QR)	0.029	48.9	0.987
3	Метод, основанный на дискретно-косинусном преобразовании (DCT)	0.032	46.7	0.993

Из результата анализа секретности цифрового отпечатка, приведённого в табл. 1 несложно заметить, что наибольшей скрытностью обладают методы,

основанные на дискретно-косинусном преобразовании, с совместно используемым методом наименее значащего бита.

3.2 Оценка достоверности интерпретации контрольной информации

Рассмотрим оценку достоверности с которой может быть восстановлен цифровой отпечаток из звукозаписи устной речи. В разделе приводится вычислительный эксперимент, в котором:

Проверяемая основная гипотеза H0: целостность звукозаписи подтверждается.

Проверяемая основная альтернативна гипотеза H1: звукозапись сфальсифицирована.

Методы анализа: в работе использованы методы цифровой обработки сигналов, статистического анализа и вычислительных экспериментов.

Цель эксперимента: оценить вероятность возникновения ошибки, приходящаяся на единицу звукозаписи (на единицу времени секунду) при прохождении всех процедур внедрения.

В качестве исходного речевого материала использовался такой же корпус как и в предыдущем эксперименте.

Выходные данные: вероятность появления ошибки, приходящаяся на единицу звукозаписи.

План эксперимента

Часть внедрения

1. Считать звукозапись устной речи.
2. Провести анализ звукозаписи устной речи, сегментированной на отрезки с целью определения отрезков, относимых к паузе.
3. Осуществить хеширование отрезков звукозаписи.
4. Осуществить помехоустойчивое кодирование цифрового отпечатка.
5. Внедрить цифровые отпечатки в отрезки звукозаписи устной речи.

Часть идентификации

6. Провести анализ звукозаписи устной речи, сегментированной на отрезки с целью определения отрезков, относимых к паузе.

7. Извлечь цифровой отпечаток из отрезков звукозаписи.

8. Сопоставить цифровой отпечаток с хранимым в базе.

9. Повторить эксперимент 10^4 раз.

Результаты эксперимента сведены в табл. 3.

Таблица 3.2 Вероятность ошибки

№	Метод	Вероятность ошибки
1	Расширения спектра (SSp)	0,13
2	Метод собственных векторов (QR)	0,16
3	Метод, основанный на дискретно-косинусном преобразовании (DCT)	0,15

3.3 Прототип технологии подтверждения целостности звукозаписи

Процедуру кодирования можно разделить на ряд этапов:

– преобразование контрольной информации, заданной символами к виду необходимому для скрывающего кодирования;

– ввод сигнала содержащего устную речь;

– определение параметров цифрового представления речевого сигнала;

– анализ параметров цифрового представления и выбор оптимальных параметров кодированная, оптимальных в смысле обеспечения заданной скрытности при максимально возможной достоверности декодирования контрольной информации;

– формирования банка фильтров;

– формирование ортогонального базиса;

– преобразование фрагмента сигнала в отрезки данных;

– анализ отрезков;

– адаптивное кодирование.

Этап I. Задание исходных данных

1. Сформировать кодируемую информацию состоящую из M - двоичных символов (вектор кодируемых данных):

$$\vec{e} = (e_1, e_2, \ldots, e_m, \ldots, e_M)^T,\ m = 1,2,\ldots,M \qquad (3.1)$$

где e_m - кодируемый символ, принимающий значение $e_m \in \{-1,1\}$, ($e_m = 1$ - осуществляется кодирование соответствующего значению бита «1», $e_m = -1$ - осуществляется кодирование соответствующего значению бита «0»); ^{T} - символ транспонирования, подразумевающий формирование данных в виде столбца; M - период повторения символов информации.

Кодируемая информация сформирована генератором двоичных последовательностей произвольной длины, на основе рекуррентного соотношения [1, с.54]:

$$\Phi(X) = X^{24} \oplus X^4 \oplus X^3 \oplus X \oplus 1, \qquad (3.2)$$

где X - триггер генератора; $\oplus$ - знак суммы по модулю два.

Вероятности появления «1» или «0» приближаются к ½, за начальное состояние принята последовательность вида «10100001110011010001011 00». После получения по (2) псевдослучайная последовательность (ПСП) была преобразована к виду $e_m \in \{-1,1\}$, для этого число «0» заменялось числом «-1». Период повторения последовательности составил $M = \left(2^{24} - 1\right) = 16777215$.

2. Формируется отрезки данных со значениями $\vec{\xi} = \left(\xi_1, \xi_2, \ldots, \xi_N\right)^T$, каждый из которых соответствующий флуктационному шуму описываемого нормальным законом распределения.

Для этого последовательность построенная на основе линейного конгруэнтного генератора [1, с.31]:

$$\zeta_{n+1} = \left(\zeta_n \cdot j + b\right) \bmod M,\ n = 1,\ldots,M, \qquad (3.3)$$

где m - модуль; j - множитель; b - приращение; x_0 - начальное заполнение.

$$\xi_{n+1} = \sin\left(2\pi \cdot \widehat{\zeta}_n\right)\sqrt{-2\ln\left(\widehat{\zeta}_{n+1}\right)},\ n = 1,\ldots, M\ , \tag{3.4}$$

где ξ_n значения случайной величины соответствующей флотационному шуму описываемой нормальным законом распределения.

3. Задать соотношение шум/сигнал h_0^2, $h_0^2 = \left[10^{-4},\ldots, 10\right]$.

4. Вести длину отрезка аудиосигнала $N = 128$ ($N = 256$) в отчетах.

Этап III. Анализ отрезка данных

6. Из файла, содержащего аудио-сигнал $\vec{f} = (f_1,..., f_L)^T$, выделяется вектор, соответствующий отрезку данных $\vec{x}$ в количестве N отчетов, ($\vec{x}$ - данные хранимые в виде файла; x_n - отчет отрезка данных $\vec{x}_z = \left(x_1,..., x_n,..., x_N\right)^T$, $x_n \in \left[-1,\ \ 1\right]$. Отрезки нумеруются с первого по Z (где $\vec{x}_z$ - отрезок с порядковым номером z файла-контейнера $\vec{X} = \left(\vec{x}_1,..., \vec{x}_z,..., \vec{x}_Z\right)$, $z = 1, 2,\ldots, Z$ (рис. 2)).

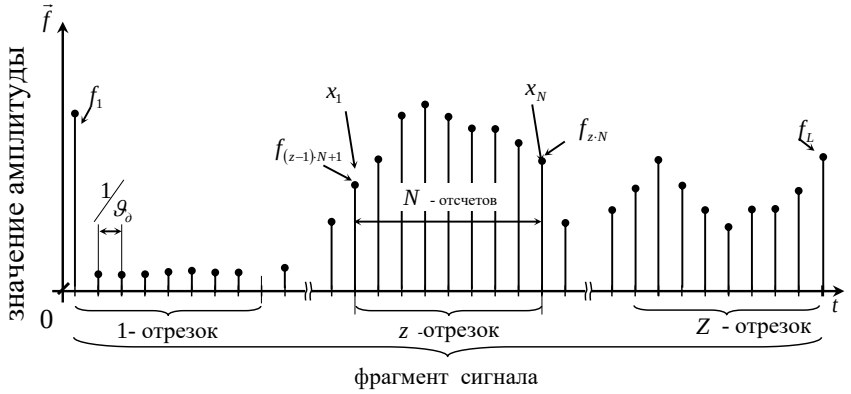


Рисунок 2 – Разбиение данных на отрезки

7. Для отрезка данных рассчитывают часть энергии P_r, приходящейся на частотный интервал r:

$$P_r = \vec{x}^T \cdot A_r \cdot \vec{x},\ r = 1, 2,\ldots, R \tag{3.5}$$

8. Для отрезка данных определяют энергию:

$$E_x = \left\|\vec{x}\right\|^2 = \int_0^T x^2 dt \cong \sum_{n=1}^{N} x_n^2\ , \tag{3.6}$$

9. Оценивается среднее значение энергии приходящееся на частотный интервал:

10. Определяется порядковый номер частотного интервала, часть энергии которого максимальна среди частотных интервалов энергия которых меньше средней:

$$\min_{r\in\mathbf{R}}\left(\overline{E}_x - P_r\right)\geq 0 \rightarrow \bar{r} = r \qquad (3.7)$$

12. Порядковый номер частотного интервала $D \in R$, в который будет осуществляется кодирование сохраняем.

Этап IV. Кодирование информации

14. Определяется значения скалярного произведения $\vec{\alpha}_i^r$ собственного вектора субполосной матрицы A_D с вектором данных $\vec{x}$:

$$\alpha_i^r = \left\langle \vec{x}, \vec{q}_i^{\,r} \right\rangle,\ i \in \mathbf{J}, \qquad (3.8)$$

где i - индекс обозначающий порядковый номер собственного вектора $\vec{q}_i$ субполосной матрицы A_r, с упорядоченными собственными числами и соответствующим им собственными векторами; $\mathbf{J}$ - множество собственных чисел отличных от нуля $\forall\lambda_{(j)}^r > 0$, в случае по убыванию упорядочивания $j = 1, 2, \ldots, J_r$, где J_r - величина определяющая количества собственных чисел больших нуля:

$$J_r \cong trec(A_r) \cong \left\lfloor \sum_{i=1}^{N} \lambda_i^r + 2 \right\rfloor. \qquad (3.9)$$

15. Осуществляем фильтрацию частотной компоненты содержащейся в векторе данных $\vec{x}$:

$$\vec{c} = \vec{x} - \alpha_i^{\bar{r}} \cdot \vec{q}_i^{\,\bar{r}},\ i \in \mathbf{J}, \qquad (3.10)$$

где $\vec{c}$ - результат фильтрации (контейнер).

16. Осуществляем кодирование e_m в отрезке данных:

$$\vec{s} = \vec{c} + sign(e_m) \cdot |\alpha_i| \cdot \vec{q}_i,\ i \in J, \qquad (3.11)$$

где $\vec{s}$ - отрезок данных с закодированной информацией; $|\ |$ - абсолютное значение взятое с положительным знаком; $sign(\ )$ - операция выделения знака.

Этап V. Воздействие шума

17. Формируется отрезок данных состоящий из N значений и соответствующий флуктационной помехи, т.е. шум $\vec{\xi} = (\xi_1, \xi_2, \ldots, \xi_N)^T$ описываемый нормальным законом распределения.

18. К отрезок данных с закодированной информацией добавляем сгенерированный шум, с соотношением шум/сигнал h_0^2:

$$\vec{y} = \vec{s} + h_0 \cdot \frac{\|\vec{s}\|}{\|\vec{\xi}\|} \cdot \vec{\xi}\ , \tag{3.12}$$

где $\vec{y} = (y_1, y_2, \ldots, y_n, \ldots, y_N)^T$ - отрезок данных с закодированной информацией после добавления шума; $\|\ \|$ - норма данных:

$$\|\vec{s}\| = \sqrt{\sum_{n=1}^{N} s_n^2}\ , \tag{3.13}$$

$$\|\vec{\xi}\| = \sqrt{\sum_{n=1}^{N} \xi_n^2}\ , \tag{3.14}$$

Стоит отметить, что значение $|y_n|$, может превышать единицу, но в использовании масштабирующего коэффициента нет необходимости, т.к. использование вычислительных возможностей MatLab с типом данных double позволяет оперировать с значениями превышающими динамический диапазон цифровой системы $\delta = \frac{1}{2^{bit}} < \frac{1}{10^{32}}$.

Этап VI. Декодирование информации

19. Для частотного интервала с сохраненным номером $r \in \mathbf{R}$, осуществляется декодирование:

$$\hat{\alpha}_i^r = \left\langle \vec{y}, \vec{q}_i^r \right\rangle,\ i \in \mathbf{J}\ , \tag{3.15}$$

$$\hat{e}_b = sign\left(\hat{\alpha}_i^r\right),\ i \in \mathbf{J}\ , \tag{3.16}$$

где $\hat{e}_b$ - декодированный бит внедренной информации; b - номер извлекаемого бита;

20. Оценка возникновения ошибки:

$$\varepsilon_b = \begin{cases} 1, & e_m \neq \hat{e}_b \\ 0, & e_m = \hat{e}_b \end{cases}, \quad (3.17)$$

где ε_b - ошибка принятия бита («1» - бит принят неверно, «0» - бит принят верно); b - порядковый номер наблюдения.

21. Воздействие новым отрезком шума $\vec{\xi}$, увеличение счетчика декодированных бит:

$$b = b + 1. \quad (3.18)$$

Оценивается количество воздействий на один отрезок, если $z > 10^4$

- нет, то переход к п. 17.

- да, то переход к п. 22.

22. Кодирование нового символа в новом отрезке данных:

$$m = m + 1. \quad (3.19)$$

Оценивается закончились ли отрезки не содержащие паузы:

- нет, то переход к п. 6.

- да, то переход к п. 23.

Алгоритм алгоритма кодирования информации состоит из четырех этапов: этапа анализа, этапа синтеза, оценки и непосредственного этапа кодирования.

1. Задается ряд параметров определяющих работу алгоритма:
- длина отрезка данных, определяемая количеством отчетов - N;
- устанавливается граница нулевого значения $\varepsilon = \frac{1}{\left(2^{bit-1} - 1\right)}$.

2. Скрываемая информация ЦВЗ (метка - $\vec{W} = \left(w_1, w_2, \ldots, w_i, \ldots, w_M\right)^T$), кодируется в виде последовательности знаков $\vec{e}$. Где e_i - компоненты внедряемого вектора $\vec{e} = \{e_1, e_2, \ldots, e_M\}$, $e_i \in \{-1, 0, 1\}$ ($e_i = 0$ - внедрение не осуществляется, $e_i = 1$ - осуществляется внедрение компоненты соответствующей значению бита «1», $e_i = -1$ - осуществляется внедрение компоненты соответствующей значению бита «0»).

Таблица 3.1 – Кодирование внедряемой информации в виде знака*

Символ кириллицы	$ASCII_{10}$	$ASCII_2$	Биполярное представление	знак $sign(\ )$
1	2	3	4	5
Н	238	11101110	1 1 1 -1 1 1 1 -1	+ + + – + + + –
И	233	11101001	1 1 1 -1 1 -1 -1 1	+ + + – + – – +
У	245	11110101	1 1 1 1 -1 1 -1 1	+ + + + – + – +
Б	226	11100010	1 1 1 -1 -1 -1 1 -1	+ + + – – – + –
е	197	11000101	1 1 -1 -1 -1 1 -1 1	+ + – – – + – +
л	204	11001100	1 1 -1 -1 1 1 -1 -1	+ + – – + + – –
Г	231	11100111	1 1 1 -1 -1 1 1 1	+ + + – – + + +
У	245	11110101	1 1 1 1 -1 1 -1 1	+ + + + – + – +

* 1 столбец – внедряемая информация; 2 столбец – внедряемая информация в $ASCII_{10}$ – коде десятичной системы счисления; 3 столбец – внедряемая информация в $ASCII_2$ – код двоичной системы счисления; 4 столбец – внедряемая информация в биполярном представлении; 5 столбец – внедряемая информация в виде значение знака компоненты внедряемого вектора $\vec{e} = \{e_1, e_2, \ldots, e_M\}$ (e_1 ="+", e_2 ="+", e_3 ="–", …, e_M ="+", для примера табл. 1 количество внедряемых символов равно $M = 64$).

5. Из файла содержащего цифровые данные соответствующие фрагменту речевого сигнала, выделяют отрезки данных $\vec{x}$ содержащие N отчетов, (x_n - отчет сигнала $\vec{x}_z = (x_1, ..., x_n, ..., x_N)^T$ (рис. 3). Отрезки нумеруют с первого по Z (где $\vec{x}_z$ - отрезок с порядковым номером z файла-контейнера $\vec{x} \in (\vec{x}_1, ..., \vec{x}_z, ..., \vec{x}_Z)$, $z = 1, 2, \ldots, Z$).

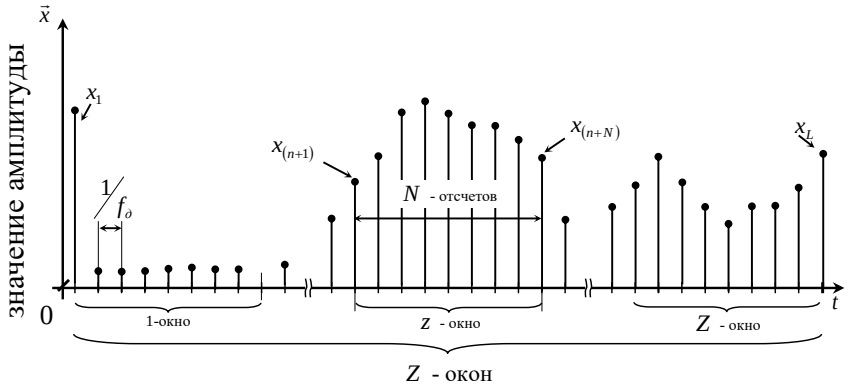


Рисунок 3 – Разбиение данных на отрезки

6. Для отрезка данных определяют энергию:

$$E_x = \|\vec{x}\|^2 = \int_0^T x^2 dt \cong \sum_{n=1}^{N} x_n^2 , \quad (12$$

7. Оценивается среднее значение энергии приходящееся на частотный интервал:

$$\overline{E}_x = \frac{\Delta\Omega_r \cdot \|\vec{x}\|^2}{\pi} , \ r = 2, \ldots, R \quad (13$$

где $\bar{}$ - символ обозначающий принадлежность к среднему значению.

$\overline{E}_x$ - пороговое значение части энергии, соответствующее среднему значению энергии.

8. Осуществляем фильтрацию частотной компоненты содержащейся в векторе данных $\vec{x}$:

$$\vec{c} = \vec{x} - \sum_{i \in J} \alpha_i^r \cdot \vec{q}_i^r \tag{19}$$

где $\vec{c}$ - результат фильтрации (контейнер).

9. Осуществляем кодирование $e(m)$ в отрезке данных, в случае если проекция больше ε:

$$\vec{s} = \vec{c} + \sum_{r \in D} \sum_{i \in J} k_i^r \cdot sign(e_m) \cdot \left|\alpha_i^r\right| \cdot \vec{q}_i^r \ , \ i \in J \ , \tag{20}$$

где $\vec{s}$ - отрезок данных с закодированной информацией; $|\ \ |$ - абсолютное значение взятое с положительным знаком; $sign(\ )$ - операция выделения знака; k_i^r.

Этап 1. Преобразование контрольной информации.

Этап 2. Ввод сигнала содержащего устную речь.

Этап 3. Выбор параметров кодированная сигнала содержащего устную речь.

Этап 4. Формирования банка фильтров

Этап 5. Формирование ортогонального базиса.

Этап 6. Преобразование фрагмента сигнала в отрезки данных.

Этап 7. Анализ отрезков.

Этап 8. Формование оптимальной сигнально-кодовой конструкции

Этап 9. Адаптивное кодирование

Стоит отметить, что поиск границ (временных меток) содержащих речь в сигнале не представляется целесообразным, т.к. кодирование может быть осуществлено в паузу содержащую шум. Это можно пояснить тем, что за стойкость отвечает энергия частотной компоненты, а точнее знак её проекции на собственной вектор, являющийся отображением контрольной информации. Если же значение энергии частной компоненты (проекции), будут малы на

столько, что не смогут обеспечить заданную достоверность извлечения, то кодирование в этой компоненте (в этом отрезке) не целесообразно.

Для работы алгоритма необходимо:

Ввод: $y(t)$, $\vec{e}$, $P_{ош}\left(h_0^2\right)$, σ .

Вывод:.

1. Ввод звукового сигнала $y(t)$.
2. Ввод контрольной информации: $\vec{e} = \left(e_1, e_2, \ldots e_m, \ldots, e_M\right)^T$.
3. Ввод достоверности: $P_{ош}\left(h_0^2\right)$.
4. Ввод скрытности: σ .
5. Задать полосу: $\Omega = \pi$.
6. Определить длительность сигнала: T .
7. Определить параметры цифрового представления (частоты дискретизации $\vartheta_д$, разрядности B), исходя из средней ёмкости данных, длительности фрагмента речевого сигнала ($y(t)$), объема контрольной информации ($\vec{e}$), а также из заданных параметров скрытности (σ) и достоверности сигнала ($P_{ош}$).
8. Определить параметры разбиения (разбиения полосы K_Ω, длительности отрезка ΔT), исходя из средней ёмкости данных, длительности фрагмента речевого сигнала ($y(t)$), объема контрольной информации ($\vec{e}$), а также из заданных параметров скрытности (σ) и достоверности сигнала ($P_{ош}$).
9. Осуществить цифровое преобразование $[\vec{y}] = \mathrm{F}_\Delta(y(t), \vartheta_д)$, $y_l \in [-1, 1]$, $l = 1, \ldots, L$.
10. Определить параметр стойкости $[\delta_n] = \mathrm{F}_\delta(y(t), \vec{e}, P_{ош}, \sigma)$.

11. Определение количество значений амплитуды входящих в отрезок $N = \Delta T \cdot \vartheta_o$.

12. Определить количество отрезков $Z = \lfloor T / \Delta T \rfloor$.

13. Определить ширину частотного интервала $\Delta\omega = (K_\Omega \cdot \Omega)/(\Delta T \cdot \vartheta_o)$.

14. Определение количества частотных интервалов: $R = \lfloor \Omega / \Delta\omega \rfloor + 1$

15. Задать шаг итерации по столбцам субполосной матрицы для первого частотного интервала: $k = 1$.

16. Задать шаг итерации по строкам субполосной матрицы для первого частотного интервала: $i = 1$.

17. Оценить является ли элемент диагональным: $i = k$,

– если элемент субполосной матрицы для первого частотного интервала расположен на диагонали $i = k$, то его значение $a_1(i,k) = 2 \cdot \Delta\omega_r / \pi$;

– если элемент субполосной матрицы для первого частотного интервала расположен на диагонали $i \neq k$, то его значение $a_1(i,k) = \sin(\Delta\omega_r \cdot (i\text{-}k))/\pi$.

18. Увеличить шаг итерации по строкам: $i = i + 1$.

19. Оценить превышает ли значение счетчика строкам $i > N$

– если значение счетчика по строкам превышает длительность отрезка $i > N$, то перейти к п. 20

– если значение счетчика по строкам не превышает длительность отрезка $i \le N$, то перейти к п. 17.

20. Увеличить шаг итерации по столбцам: $k = k + 1$.

21. Оценить превышает ли значение столбцов размерности $k > N$

– если значение счетчика по столбцам превышает размерность $k > N$, то перейти к п. 22

– если значение счетчика по столбцам превышает размерность $k > N$, то $k \le N$, то перейти к п. 16.

22. Задать номер первоначального частотного интервала: $r = 2$.

23. Рассчитать центральную частоту интервала $\omega_r = \Delta\omega_r \cdot (r - 1)$.

24. Задать шаг итерации по столбцам субполосной матрицы для r частотного интервала: $k = 1$.

25. Задать шаг итерации по строкам субполосной матрицы для r частотного интервала: $i = 1$.

26. Оценить является ли элемент диагональным, т.е. $i = k$:

– если элемент субполосной матрицы для первого частотного интервала расположен на диагонали $i = k$, то его значение $a_1(i,k) = \Delta\omega_r / \pi$;

– если элемент субполосной матрицы для первого частотного интервала расположен на диагонали $i \neq k$, то его значение $a_r(i,k) = a_1(i,k) \cdot \cos(\omega_r \cdot (i\text{-}k))$

27. Увеличить шаг итерации по строкам субполосной матрицы для r частотного интервала: $i = i + 1$.

28. Оценить превышает ли значение счетчика строкам $i > N$

– если значение счетчика по строкам превышает длительность отрезка $i > N$, то перейти к п. 29

– если значение счетчика по строкам не превышает длительность отрезка $i \leq N$, то перейти к п. 26.

29. Увеличить шаг итерации по столбцам субполосной матрицы для r частотного интервала: $k = k + 1$.

30. Оценить превышает ли значение столбцов размерности субполосной матрицы $k > N$:

– если значение счетчика по столбцам превышает размерность субполосной матрицы $k > N$, то перейти к п. 31.

– если значение счетчика по столбцам не превышает размерность субполосной матрицы $k \leq N$, то перейти к п. 26.

31. Увеличить номер частотного интервала: $r = r + 1$.

32. Определить является ли анализируемый частотный интервал последним, т.е. выполняется условие: $r = R$:

– если частотный интервал не последний $r < R$, то перейдём к п. 23.

– если частотный интервал последний $r = R$, то перейдём к п. 33.

33. Задание размерности $K_r = 0$ $(r = 1, 2, \ldots, R)$ пространства собственных векторов $\mathbf{R}_\mathbf{A}$.

34. Задать номер первоначального частотного интервала: r=1.

35. Поиск собственных векторов $(\mathbf{Q}_r = (\vec{q}_1(r), \vec{q}_2(r), \ldots, \vec{q}_k(r), \ldots, \vec{q}_N(r)))$ и соответствующих им собственных чисел ($\mathbf{\Lambda}_r = diag\left[\lambda_{1,1}(r), \ldots, \lambda_{k,}(r), \ldots, \lambda_{N,N}(r)\right]$): $\mathbf{A}_r\mathbf{Q}_r = \mathbf{\Lambda}_r\mathbf{Q}_r$.

36. Задать шаг итерации по номерам собственных векторов: $k = 1$.

37. Оценить превышает ли доля энергии собственного вектора $\vec{q}_k(r)$ субполосной матрицы в частотном интервале r порог определяемый скрытностью, т.е. $\lambda_k \geq \sigma$.

– если превышает, то перейдём к п. 38;

– если неравенство ложно $i > N$, то перейдём к п. 40.

38. Увеличение размерности пространства $\mathbf{R}_\mathbf{A}$, $K_r = K_r + 1$.

39. Добавление собственного вектора $\vec{q}_k(r)$ к пространству $\mathbf{R}_\mathbf{A}$, т.е. $\vec{q}_k(r) \to \mathbf{R}_\mathbf{A}$.

40. Увеличить шаг итерации по номерам собственных векторов субполосной матрицы для r частотного интервала: $k = k + 1$.

41. Оценить превышает ли значение номера размерности субполосной матрицы $k > N$.

– если значение счетчика по номерам превышает размерность субполосной матрицы $k > N$, то перейти к п. 42.

– если значение счетчика по номерам не превышает размерность $k \leq N$, то перейти к п. 37.

42. Увеличить номер частотного интервала: $r = r + 1$.

43. Определить является ли анализируемый частотный интервал последним, т.е. выполняется условие: $r = R$:

– если частотный интервал не последний $r < R$, то перейдём к п. 35.

– если частотный интервал последний $r = R$, то перейдём к п. 44.

44. Задать номер кодируемого символа: $m = 0$.

45. Задать номер анализируемого отрезка данных: $z = 1$.

46. Задать номер первого цифрового значения амплитуды отрезка данных: $n = N \cdot (z - 1) + 1$.

47. Выделить отрезок данных: $\vec{x} = (y_n, y_{n+1}, \ldots, y_{n+N})^T$.

48. Задать размер подпространства сигнально-кодовой конструкции (СКК): $j = 0$.

49. Задать значение решающей функции: $H_x = \left(2 \cdot \|\vec{x}\|^2 \cdot \Delta\omega\right) / \Omega$.

50. Задать номер первоначального частотного интервала: $r = 1$.

51. Найти значение энергии в частотном интервале: $P_r = \vec{x} \cdot \mathbf{A}_r \cdot \vec{x}^T$.

52. Увеличить номер частотного интервала: $r = r + 1$.

53. Определить является ли анализируемый частотный интервал последним, т.е. выполняется условие: $r = R$:

– если частотный интервал не последний $r < R$, то перейдём к п. 51.

– если частотный интервал последний $r = R$, то перейдём к п. 54.

54. Задать номер частотного интервала, содержащего максимальную энергию: $d = 1$.

55. Задать часть энергии $\hat{\mathbf{P}}_d = 0$ (крышечка тут означает отсортированные значения от максимального к минимальному), соответствующего d частотному интервалу.

56. Зафиксировать номер $k = 0$ частотного интервала с максимальной энергией среди оставшихся в массиве $\mathbf{P}_r$.

57. Задать номер первоначального частотного интервала: $r = 1$.

58. Оценить превышает ли значение энергии в r-м интервале, значение энергии в d-м интервале отсортированных значений, т.е. $\mathbf{P}_r > \hat{\mathbf{P}}_d$:

– если не превышает $\mathbf{P}_r < \hat{\mathbf{P}}_d$, то перейдём к п. 62.

– если превышает $\mathbf{P}_r > \hat{\mathbf{P}}_d$, то перейдём к п. 59.

59. $\hat{\mathbf{P}}_d = \mathbf{P}_r$.

60. $D_d = r$.

61. $k = r$.

62. Увеличить номер частотного интервала: $r = r + 1$.

63. Определить является ли анализируемый частотный интервал последним, т.е. выполняется условие: $r = R$:

– если частотный интервал не последний $r < R$, то перейдём к п. 58.

– если частотный интервал последний $r = R$, то перейдём к п. 64.

64. $\mathbf{P}_k = 0$.

65. Увеличить номер частотного интервала: $d = d + 1$.

66. Определить является ли анализируемый частотный интервал последним, т.е. выполняется условие: $d = R$:

– если частотный интервал не последний $d < R$, то перейдём к п. 55.

– если частотный интервал последний $d = R$, то перейдём к п. 67.

67. Вычислить значение адаптивного порога: $H_x = \left(\|\vec{x}\|^2 \cdot \Delta\omega \right) / \Omega$.

68. Задать номер частотного интервала: $d = 1$.

69. Оценить принадлежит ли частотный интервал частотному потенциалу, т.е. превышает ли часть энергии в частотном интервале значение адаптивного порога: $H_x > \hat{\mathbf{P}}_d$:

– если частотный интервал принадлежит частотному потенциалу $H_x \geq \hat{\mathbf{P}}_d$, то перейдём к п. 70.

– если частотный интервал не принадлежит частотному потенциалу $H_x < \hat{\mathbf{P}}_d$, то перейдём к п. 92.

70. Определить номер частотного интервала: $r = D_d$.

71. Задать номер собственного вектора субполосной матрицы входящего в пространство $\mathbf{R}_{\mathbf{A}}$: $k=1$.

72. Выбрать собственный вектор $\vec{q}_k(r)$ субполосной матрицы из пространства собственных векторов $\mathbf{R}_{\mathbf{A}}$.

73. Определить проекцию: $\alpha=\langle\vec{x},\vec{q}_k(r)\rangle$.

74. Оценить превышает ли значение энергии проекции значение порога определяющего стойкость: $|\alpha|>\delta_n$:

– если значение энергии проекции превышает значение порога $|\alpha|>\delta_n$, то перейдём к п. 75.

– если значение энергии проекции не превышает значение порога $|\alpha|\le\delta_n$, то перейдём к п. 90.

75. Установить флаг принадлежности собственного вектора к подпространству сигнально-кодовой конструкции: $flag=true$.

76. Задать номер функции сигнально кодовой конструкции: $j=1$.

77. Извлечение функции сигнально кодовой конструкции из подпространства: $\vec{\varphi}_j\leftarrow\mathbf{R}_J(z)$.

78. Определить проекцию: $\beta=\langle\vec{\varphi}_j,(\alpha\cdot\vec{q}_k(r))\rangle$.

79. Оценить повлияет ли просачивание энергии проекции $\alpha\cdot\|\vec{q}_k(r)\|^2$, на вектор $\vec{\varphi}_j$ входящий в сигнально-кодовую конструкцию, т.е. будет ли превышено значение порога определяющего стойкость: $|\beta|>\delta_n$:

– если значение энергии проекции превышает значение порога $|\beta|>\delta_n$, то перейдём к п. 80.

– если значение энергии проекции не превышает значение порога $|\beta|\le\delta_n$, то перейдём к п. 81.

80. Так как собственный вектор влияет на сигнально-кодовую конструкцию, то установить флаг принадлежности собственного

вектора к подпространству сигнально-кодовой конструкции: $flag = false$.

81. Увеличить номер сигнально кодовой конструкции: $j = j + 1$.

– если номер больше размерности пространства сигнально-кодовой конструкции $j > J_z$, то перейдём к п. 82.

– если номер меньше размерности пространства сигнально-кодовой конструкции $j \le J_z$, то перейдём к п. 77.

82. Проанализировать собственный вектор $\alpha \cdot \|\vec{q}_k(r)\|^2$ влияет на сигнально-кодовую конструкцию $\mathbf{R}_J(z)$:

– если не влияет, т.е. $flag = true$, то перейдём к п. 83.

– если влияет, т.е. $flag = false$, то перейдём к п. 90.

83. Увеличить размер подпространства сигнально-кодовой конструкции $j > J_z$.

84. Перейти к следующему номеру кодируемого символа: $m = m + 1$.

85. Определить существует ли символ, т.е. $m \le M$:

– если номер кодируемого символа не последний $m \le M$, то перейдём к п. 86.

– если символа с таким номером не существует $m > M$, то перейдём к п. 88.

86. Формирование вектора сигнально-кодовой конструкции содержащей e_m символ кодируемой информации: $\vec{\varphi}_j = \left(e_m \cdot |\alpha| - \alpha\right) \cdot \vec{q}_k(r)$.

87. Перейдём к п. 89.

88. Формирование вектора сигнально-кодовой конструкции фиксирующий конец внедрения: $\vec{\varphi}_j = -\alpha \cdot \vec{q}_k(r)$.

89. Добавим вектор к сигнально-кодовой конструкции: $\vec{\varphi}_j \to \mathbf{R}_J(z)$.

90. Увеличить номер собственного вектора субполосной матрицы входящего в пространство $\mathbf{R}_A$: $k = k + 1$.

91. Определить принадлежит ли номер собственного вектора субполосной матрицы размерности пространства $\mathbf{R}_{\mathbf{A}}$, т.е. выполняется условие: $k > K_r$:

– если номер больше размерности пространства $k > K_r$, то перейдём к п. 92.

– если номер меньше размерности пространства $k \le K_r$, то перейдём к п. 72.

92. Увеличить номер частотного интервала: $d = d + 1$.

93. Определить является ли анализируемый частотный интервал последним, т.е. выполняется условие: $d = R$:

– если частотный интервал не последний $d < R$, то перейдём к п. 55.

– если частотный интервал последний $d = R$, то перейдём к п. 67.

94. Задать номер функции сигнально кодовой конструкции: $j = 1$.

95. Извлечение функции сигнально кодовой конструкции из подпространства: $\vec{\varphi}_j \leftarrow \mathbf{R}_J(z)$.

96. $\vec{x} = \vec{x} + \vec{\varphi}_j$

97. Увеличить номер сигнально кодовой конструкции: $j = j + 1$.

– если номер больше размерности пространства сигнально-кодовой конструкции $j > J_z$, то перейдём к п. 82.

– если номер меньше размерности пространства сигнально-кодовой конструкции $j \le J_z$, то перейдём к п. 95.

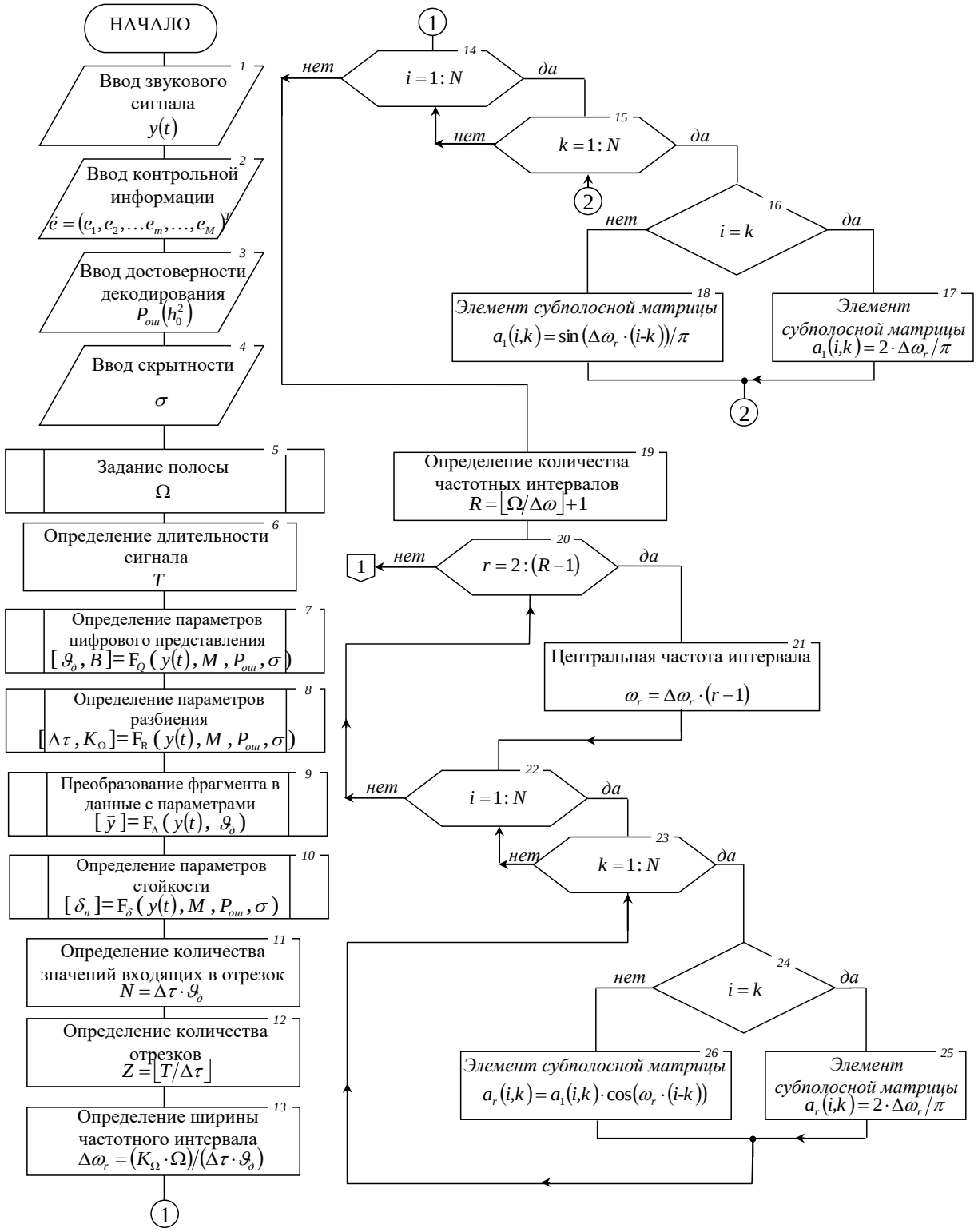


Рисунок 3.2 – Блок-схема алгоритма скрытного кодирования контрольной информации в звуковом фрагменте (начало)

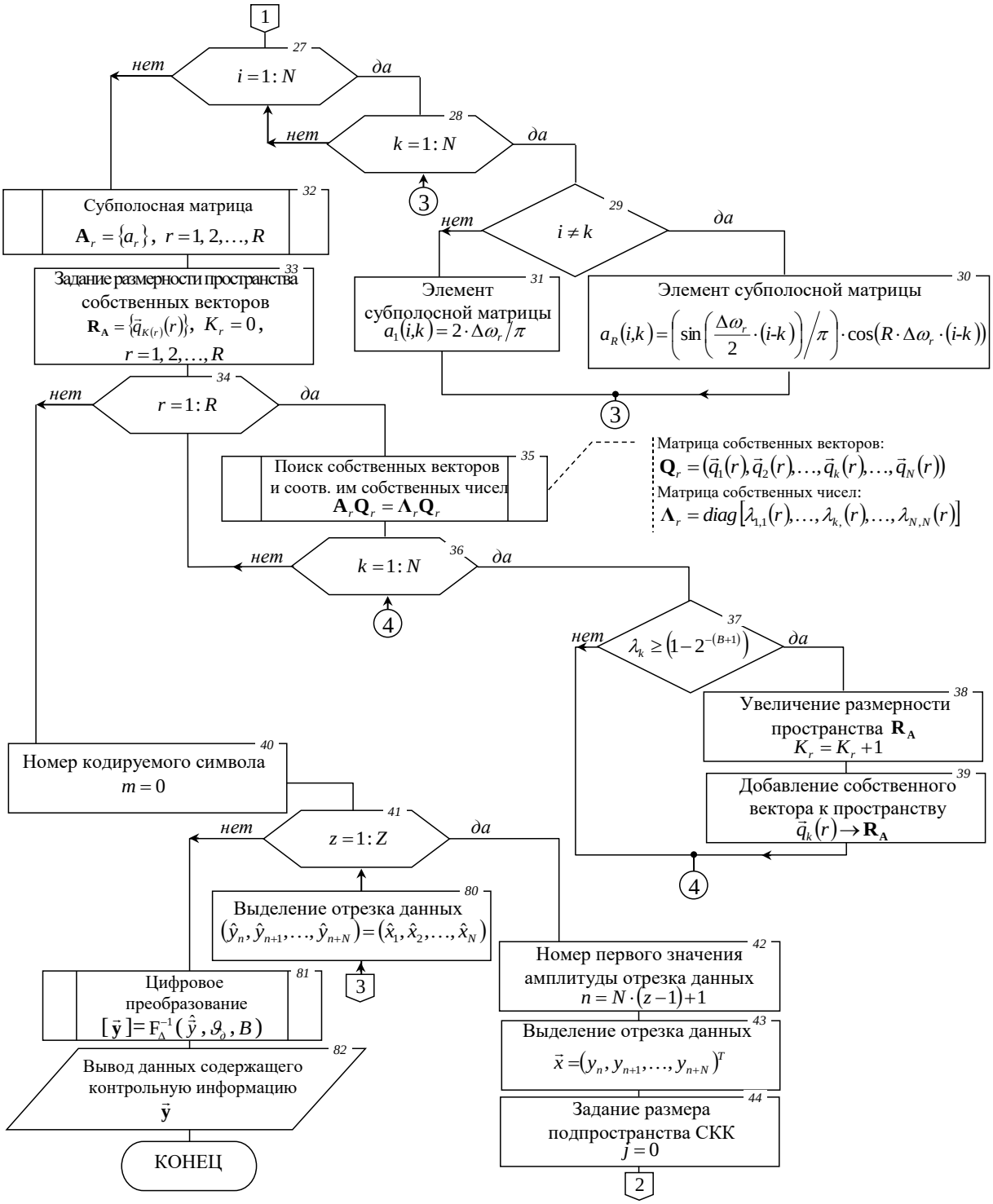


Рисунок 3.3 – Блок-схема алгоритма скрытного кодирования контрольной информации в звуковом фрагменте (продолжение)

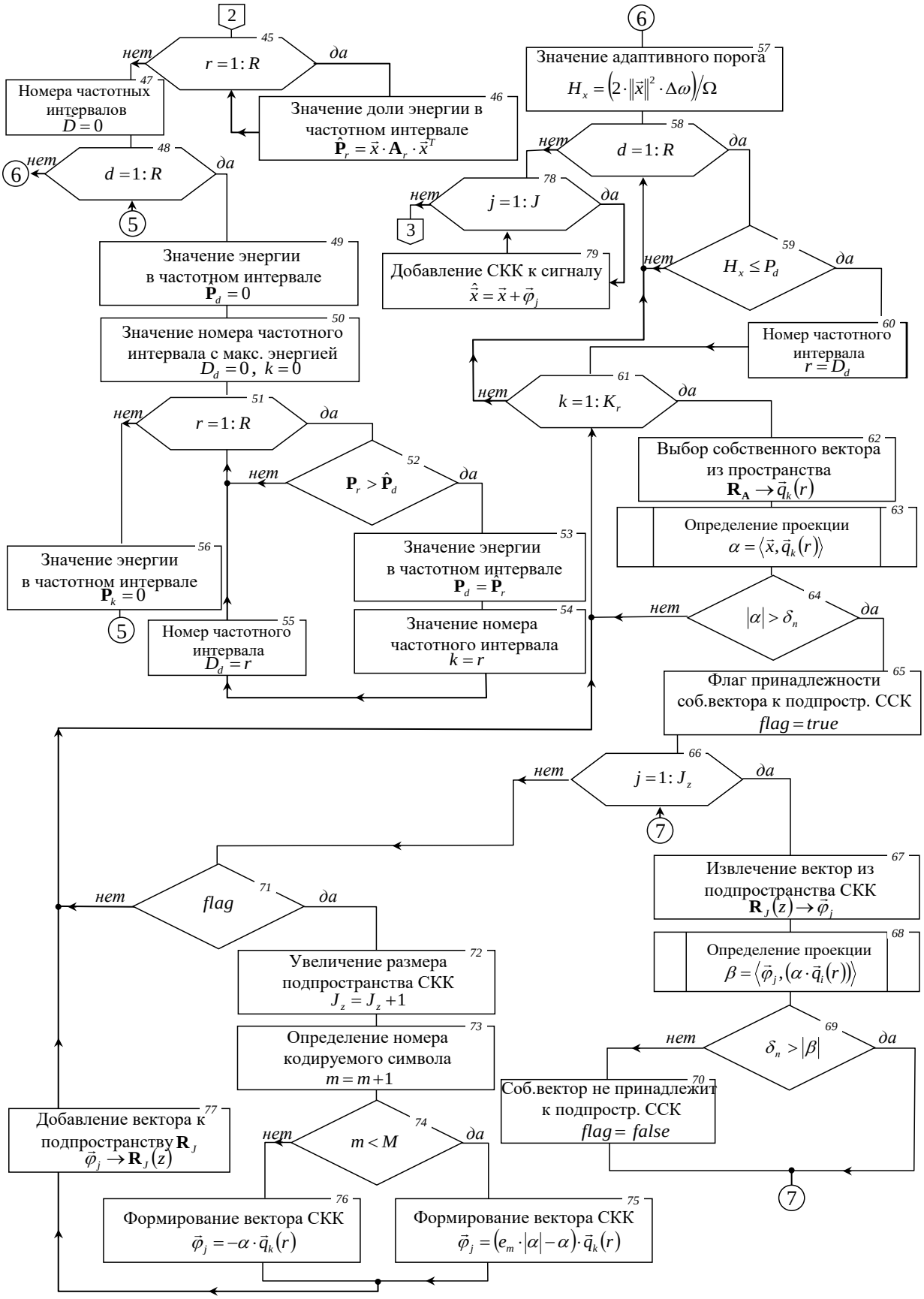


Рисунок 3.4 – Блок-схема алгоритма скрытного кодирования контрольной информации в звуковом фрагменте (окончание)

Блоки один и два вводят фрагмент сигнала и соответственно контрольную информацию.

Блоки с третьего по пятый определяют, параметры скрытности и стойкости контрольной информации, которые предполагается достичь в результате кодирования.

В блоках с седьмого по десятый согласно рекомендаций полученных в второй и третьей главах, согласно выявленным закономерностям, определяют параметры разбиения в временной и частотных областях.

Блок девять преобразует данные (сигнал) в цифровой вид с нормированием значений амплитуд к единице.

Блоки с одиннадцатого по тринадцатый включительно, определяют разбиения фрагмента сигнала в цифровом виде во временной и частотной областях.

Блоки с четырнадцатого по тридцать первый включительно, позволяют рассчитать субполосные матрицы согласно выбранного разбиения. Стоит отметить, что для первого и последнего частотного интервалов используются отдельные циклы. Это вызвано тем, что частотные интервалы для которых рассчитываются эти суполосные матрицы имеют ширину в два раза меньше. Субполосные матрицы рассчитанные для частотных интервалов (не включая первый и последний) используют для расчета субполосную матрицу полученную для первого частотного интервала. Использование суплосной матрицы найденной для первого частотного интервала увеличивает скорость работы алгоритма, за счет уменьшения количества операций.

Тридцать второй блок объединяет субполосне матрицы в одну базу с параметрами разбиения.

Блоки с тридцать третьего по тридцать девятый формирую векторное пространство состоящее из собственных векторов субполосной матрицы отобранных исходя из разрядности данных (или необходимой величины просачивания энергии за пределы полосы).

В блоках с сорокового по восьмидесятый осуществляется кодирование контрольной информации в сигнал.

Цикл включающий блоки сорок пять и сорок шесть определяет значение долей энергии содержащихся в частотных интервалах.

Цикл включающий блоки с сорок седьмого по пятьдесят шестой, осуществляет сортировку частей энергии с максимальным значение на первой позиции и далее убыванием. Сортировка значений энергии с фиксацией номеров частотных интервалов необходима для формирования устойчивой сигнально-кодовой конструкции. Устойчивость осуществляется за счет того, что кодирование будет осуществляется вначале в частотные интервалы с большей энергией (с большими значениями субполосных проекций), а потом по убыванию энергии.

Блоки с пятьдесят седьмого по семьдесят седьмой формируют на сигнально-кодовую конструкцию исходя из параметров скрытности и стойкости.

В блоке пятьдесят семь определяется значение решающей функции, необходимой для адаптивного выбора частотных интервалов.

В блоках с шестьдесят четвертого по семидесятый определяется степень влияния субполосной проекции (найденной в блоке 63) на сигнально-кодовую конструкцию.

Если изменение проекции не повлияет на уже закодированную информацию блок семьдесят один, то проекцию отображающую элемент контрольной информации (блоки с семьдесят четыре по семьдесят шесть) добавляют к сигналу (блоки семьдесят восемь и семьдесят девять) и вносят в сигнально-кодовую конструкцию (блоки с семьдесят второго по семьдесят седьмой).

Разработанный алгоритм позволяет скрытно закодировать контрольную информацию в данные являющиеся цифровым представлением устной речи зафиксированной на выходе микрофона. В основу работы алгоритма положено

использование в качестве базисных функций собственных векторов субполосной матрицы. Используемый базис позволяет осуществить формирование сигнально-кодовой конструкции с долей энергии в задаваемой полосе. Стоит отметить, что для обеспечения скрытности, энергию сигнально-кодовой конструкции концентрируют в адаптивно выбранной частотной полосе (наборе частотных интервалах), а её величину выбирают равной той которой она была изначально. Адаптация к отрезку речевых данных так же осуществляется за счет выбора в качестве пригодных для стеганографического кодирования, частотных интервалов, часть энергии в которых меньше среднего значение энергии, приходящееся на частотную полосу. Формирование подпространства СКК из собственных векторов субполосной матрицы, удовлетворяющих соотношению ( ), позволяет минимизировать изменения в частотных интервалах, в которых не содержится подавляющая энергия СКК. Иными словами, алгоритм позволяет избирательно использовать часть частотной полосы, не допуская искажения остального спектра речевого сигнала. Использование в совокупности описанных выше подходов позволяет получить высокую емкость данных.

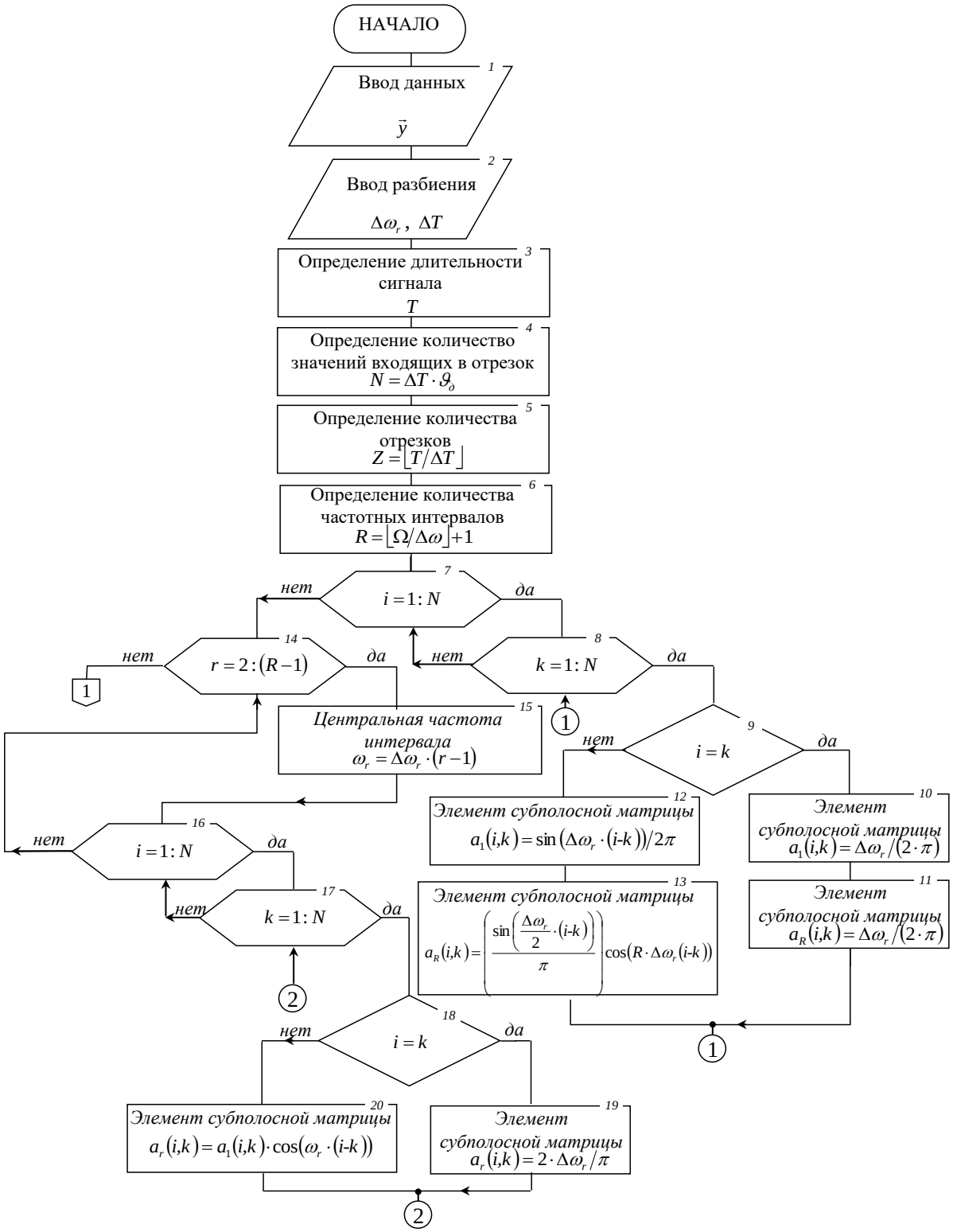


Рисунок 3.5 – Блок- схема алгоритма декодирования контрольной информации из звукового фрагмента (начало)

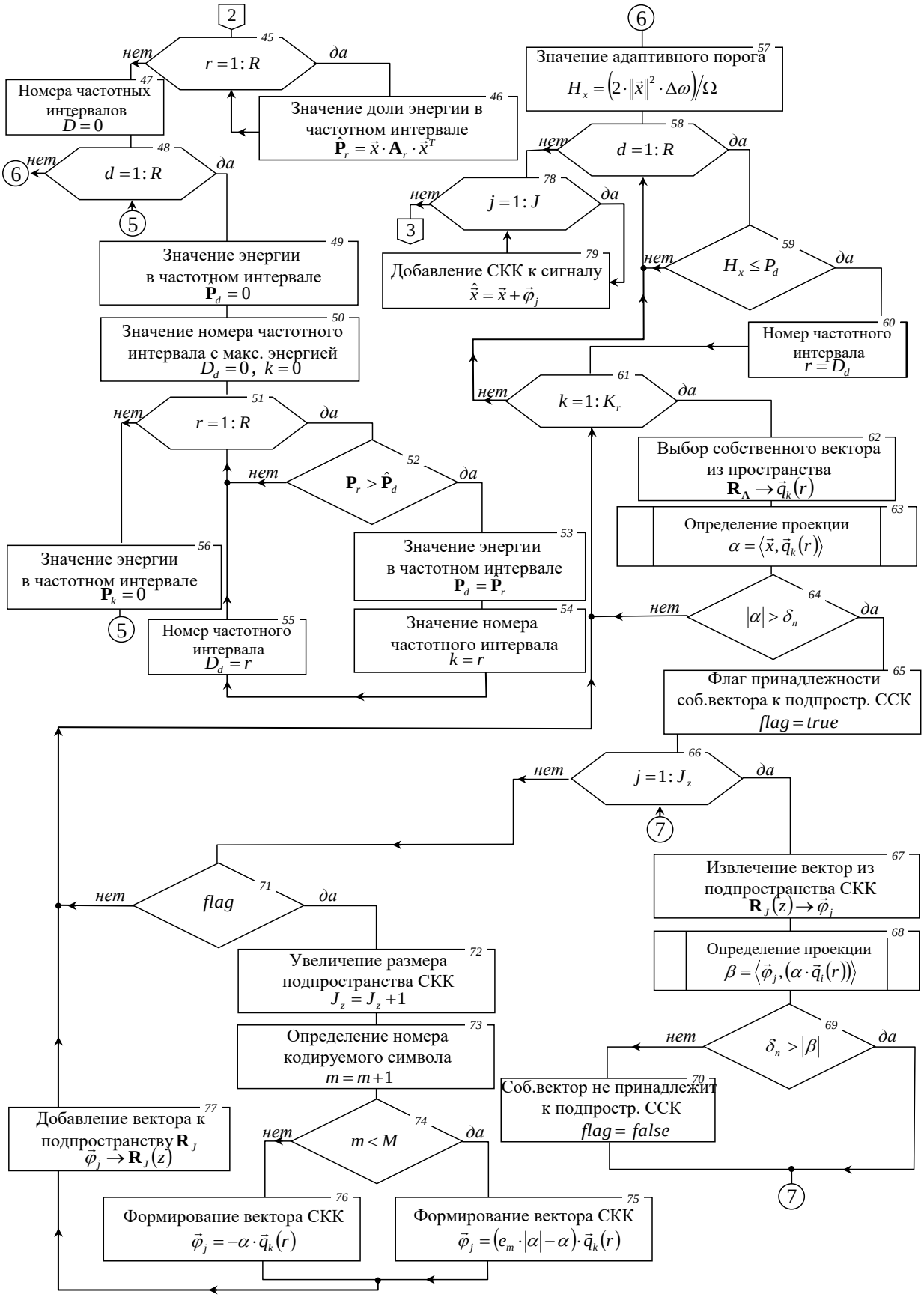


Рисунок 3.6 – Блок- схема алгоритма декодирования контрольной информации из звукового фрагмента (продолжение)

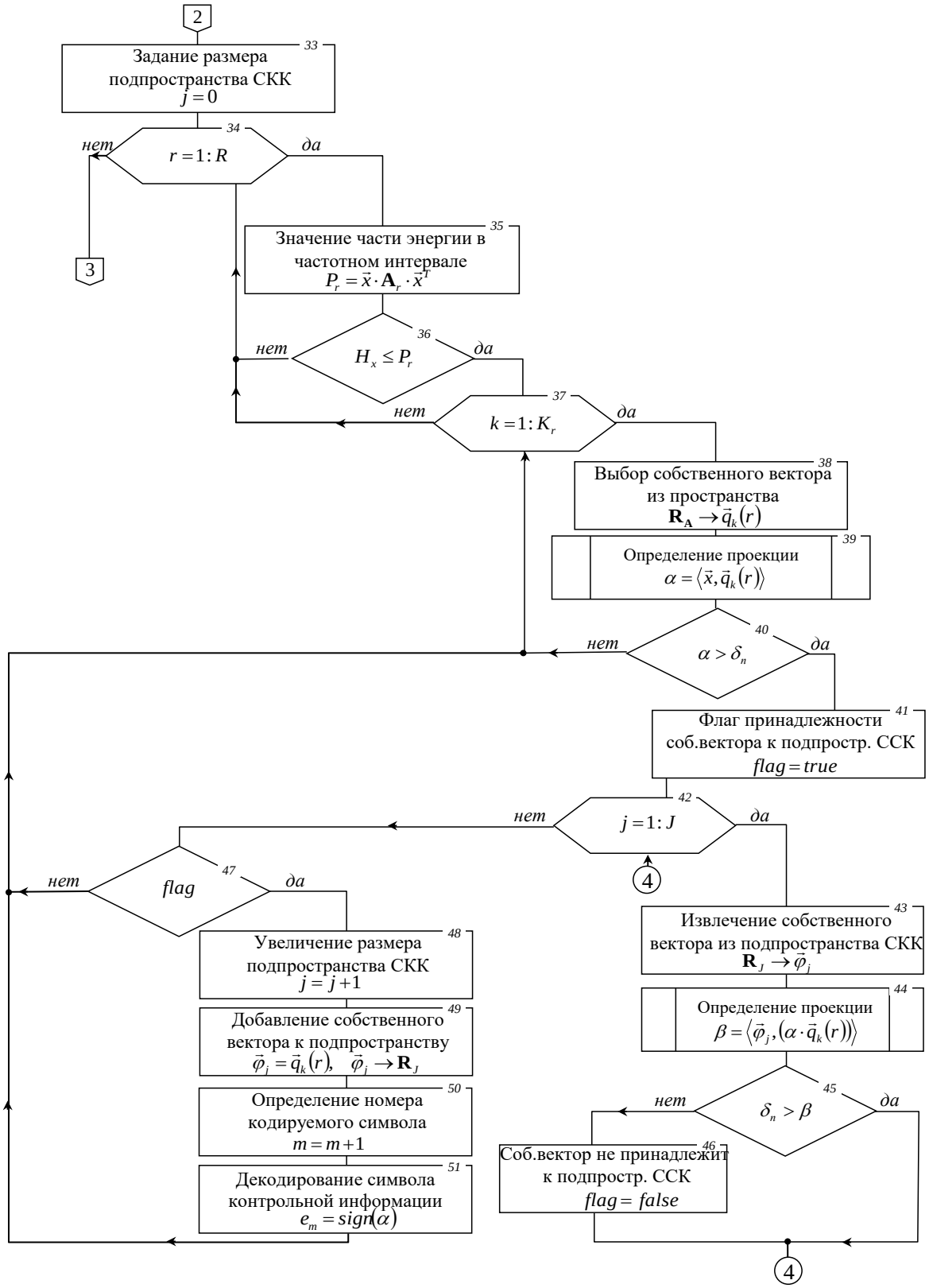


Рисунок 3.7 – Блок- схема алгоритма декодирования контрольной информации из звукового фрагмента (окончание)

Программная поддержка процедур внедрения и восстановления

Архитектура программно-аналитической поддержки разработанных методов и алгоритмов кодирования контрольной информации в фрагмент речевых данных состоит из следующих программных модулей:

– модуль ввода звукового сигнала;

– модуль преобразования контрольной информации в коды;

– модуль выбора параметров кодирования;

– модуль разбиения данных на отрезки;

– модуль формирования банка субполосных фильтров;

– модуль формирование ортогонального базиса;

– модуль анализа отрезков;

– модуль синтеза оптимальной сигнально-кодовой конструкции;

– модуль скрытного кодирования контрольной информации в отрезок;

– модуль записи сигнала в виде данных;

– модуль чтения речевых данных;

– модуль декодирования контрольной информации.

На рисунке **Ошибка! Источник ссылки не найден.** представлена с труктурная схема программной поддержки скрытного кодирования контрольной информации в звуковом фрагменте созданная на основе разработанных алгоритмов.

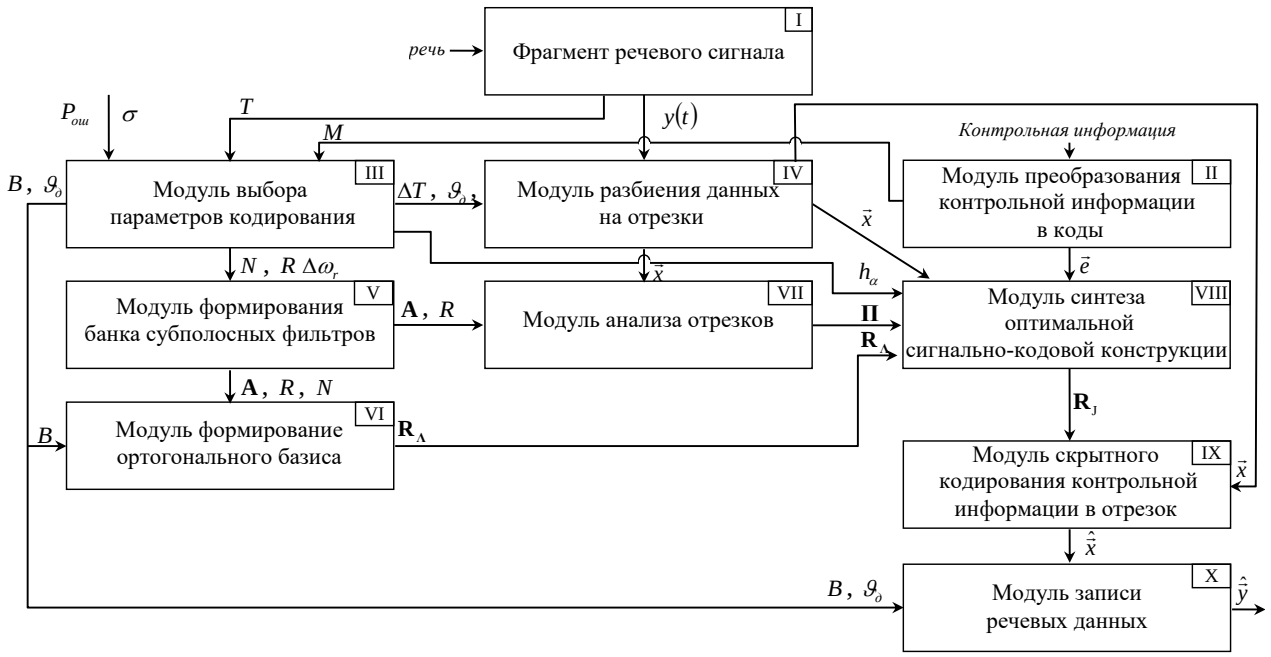


Рисунок 3.8 – Структурная схема программной поддержки скрытного кодирования контрольной информации в звуковом фрагменте

Модуль ввода фрагмента звукового сигнала, входной информацией которого, является аудио-сигнал поступивший с АЦП. Выходными данными данного модуля являются: фрагментом сигнала $y(t)$ и его длительность T.

Модуль преобразования контрольной информации в коды. Контрольная информация поступавшая модуль, преобразуется в символы, а символы согласно общепринятых таблиц кодирования (к примеру ASCII) преобразуется в двоичный вид. Далее контрольная информация, представленная двоичной системой счисления преобразуется в знаки, т.е. бит соответствующий единице кодируется положительным знаком а бит соответствующий нулю отрицательным знаком. На выход модуля поступает контрольная информация в цифровом виде $\vec{e}$ ($\vec{e} = (e_1, e_2, \ldots, e_m, \ldots, e_M)^T$, $e_m = \{sign(1), sign(-1)\}$) и объем контрольной информации M.

Модуль выбора параметров кодирования. Оценивает объем контрольной информации M, полученной от модуля преобразования контрольной информации; длительность сигнала T, полученной от модуль ввода

фрагмента звукового сигнала; допустимую вероятность ошибки $P_{ош}$ и рекомендуемый уровень качества σ, задаваемые пользователем. Затем модуль используя банк данных осуществляется выбор оптимального разбиения частотной полосы (т.е. ширину частотных интервалов $\Delta\omega_r$, $r \in R$, частоту дискретизации $\vartheta_д$) и параметров временного представления сигнала цифровой вид (разрядность хранимых данных B, длительность отрезка ΔT). Под оптимальностью тут понимаем обеспечение заданной скрытности с максимально возможной достоверностью декодирование контрольная информации.

Модуль разбиения данных на отрезки, выполняет хранение фрагмента сигнала $y(t)$ в цифровом виде и его преобразование в данные с частотой дискретизации $\vartheta_д$, при этом мгновенная амплитуда $y(t)$ нормируется к единице и на выходе представляется фрагментом данных $\vec{y}$ с плавающей точкой. Далее происходит разбиение фрагмента речевых данных $\vec{y}$ на отрезки $\vec{x}$ с длительностью отрезка ΔT и частотой дискретизации $\vartheta_д$. Выходными данными этого модуля являются: отрезки речевых данных $\vec{x}$.

Модуль формирования банка субполосных фильтров. Используя значение длительности отрезка анализа ΔT (т.е. с количеством значений N) и частоту дискретизации $\vartheta_д$, а так же параметры разбиения частотной оси, т.е. ширину частотных интервалов $\Delta\omega_r$, и количество частотных интервалов R. Модуль рассчитывает субполосные матрицы $\mathbf{A}_r$, $r = 1,2,\ldots,R$. Выходными данными этого модуля являются: субполосные матрицы $\mathbf{A} = \{\mathbf{A}_r\}$, и их количество R.

Модуль формирование ортогонального базиса $\mathbf{R}_\Lambda$. Осуществляет поиск ортогональных функций $\mathbf{R}_\Lambda = \{\vec{q}_i(r)\}$, $i = 1,\ldots,J$, $r = 1,2,\ldots,R$, то есть таких сигналов, которые имеют подавляющую долю энергии в частотном интервале r (т.е. собственных векторов $\vec{q}_i(r)$ субполосной матрицы $\mathbf{A}_r$ найденных для в

частотной полосы шириной $\Delta\omega_r$ и центральной частоты ω_r), и при этом энергия которых оказывает минимальное влияние на соседние частотные интервалы $\int_{\upsilon\notin\Delta\omega_r}\left\|\vec{Q}_i(r)\right\|^2 d\upsilon\approx 0$. Выходными данными этого модуля являются: набор собственных векторов субполосной матрицы $\mathbf{R}_{\Lambda}=\{\vec{q}_i(r)\}$, с заданной долей просачивания ($\int_{\upsilon\notin\Delta\omega_r}\left\|\vec{Q}_i(r)\right\|^2 d\upsilon\le\delta^{-B}/2$), найденные для частотных интервалов $r=1,2,\ldots,R$.

Модуль анализа отрезков. Согласно выбранному частотно/временному представлению сигнала $\{\Delta\omega_r,\ \vartheta_\partial,\ R,\ \Delta T\}$ и соответствующего отрезка данных $\vec{x}$, при использовании субполосных матриц $\mathbf{A}=\{\mathbf{A}_r\}$, $r=1,2,\ldots,R$ осуществляется анализ распределения энергии P_r, по частотным интервалам $r\in R$, с целью выбора оптимальной полосы $\mathbf{\Pi}=\{(\Delta\omega_r,\omega_r)\}$, $r\in R$, изменения в которой не ухудшат восприятие сигнала. Иными словами происходит определение частотного потенциала $\mathbf{\Pi}$ отрезка речевых данных $\vec{x}$, состоящего из адаптивно выбираемых частотных интервалов $\{(\Delta\omega_r,\omega_r)\}$. Адаптивность достигается за счет использования решающего правила опирающегося на ширину частотного интервала $\Delta\omega_r$ и энергию отрезка $\|\vec{x}\|^2$. Выходными данными этого модуля являются: номера частотных интервалов входящих в частотный потенциал $\mathbf{\Pi}$ для отрезка речевых данных $\vec{x}$.

Модуль синтеза оптимальной сигнально-кодовой конструкции. Используя знания о частотном потенциале $\mathbf{\Pi}=\{(\Delta\omega_r,\omega_r)\}$ $r\in R$, в котором хранятся информация о номерах частотных интервалов которые можно осуществить внедрение, также стоит отметить, что номера частотных интервалов представлены в порядке убывания энергии определенной для этого частотного интервала. В модуле начиная с первого частотного интервала входящего в частотный потенциал определяются проекции $\alpha_i(r)$ собственных

векторов $\vec{q}_i(r)$. Далее происходит анализ, состоящий из двух этапов. Первый этап определяет стойкость проекции $\alpha_i(r)$, проекция может быть проанализирована на втором этапе, если эта проекция больше порога h_α, определяемого ошибкой верного декодирования. Второй этап заключается проверке окажет ли влияние эта проекция на значения других проекций не принадлежащих этому частотному интервалу, но входящих в подпространство $\mathbf{R}_J$, $\beta_j \le \sigma_\alpha$, $\forall j \in J$, где $\beta_j = \langle(\alpha_j(r)\cdot\vec{\varphi}_j),(\alpha_i(r)\cdot\vec{q}_i(r))\rangle$, если влияние не указано то формируется вектор сигнально-кодовой конструкции $\vec{g}_j = sign(e_m)\cdot|\alpha_i(r)|\cdot\vec{q}_i(r)$ и увеличивается подпространство $\vec{\varphi}_j = \vec{q}_i(r)$. Выходными данными этого модуля являются: $\mathbf{R}_J$, т.е. набор сигналов $\vec{\varphi}_j$ ($j = 1,2,\ldots,J$), являющихся отображением контрольный информации e_m.

Модуль скрытного кодирования контрольной информации в отрезок данных осуществляет предварительную фильтрацию набором из $\mathbf{R}_J$ для соответствующего отрезка данных $\vec{x}$ и осуществляет кодирование $\vec{e}$, добавляя сигнально-кодовую конструкцию $\vec{g}_j$, $j = 1,2,\ldots,J$. Далее исходя из разрядности хранимых данных отрезок $\hat{\vec{x}}$ квантуется, далее все отрезки объединяются в фрагмент $\hat{\vec{y}}$. На выходе этого модуля получается фрагмент речевых данных $\hat{\vec{y}}$.

Разработанная программная поддержка метода скрывающего кодирования контрольной информации позволяет при декодировании использовать только параметры разбиения данных в частотной и временной областях, а все оставшиеся параметры могут быть определены из отрезка данных. На рисунке **Ошибка! Источник ссылки не найден.** представлена с труктурная схема программной поддержки декодирования контрольной информации из звукового фрагмента.

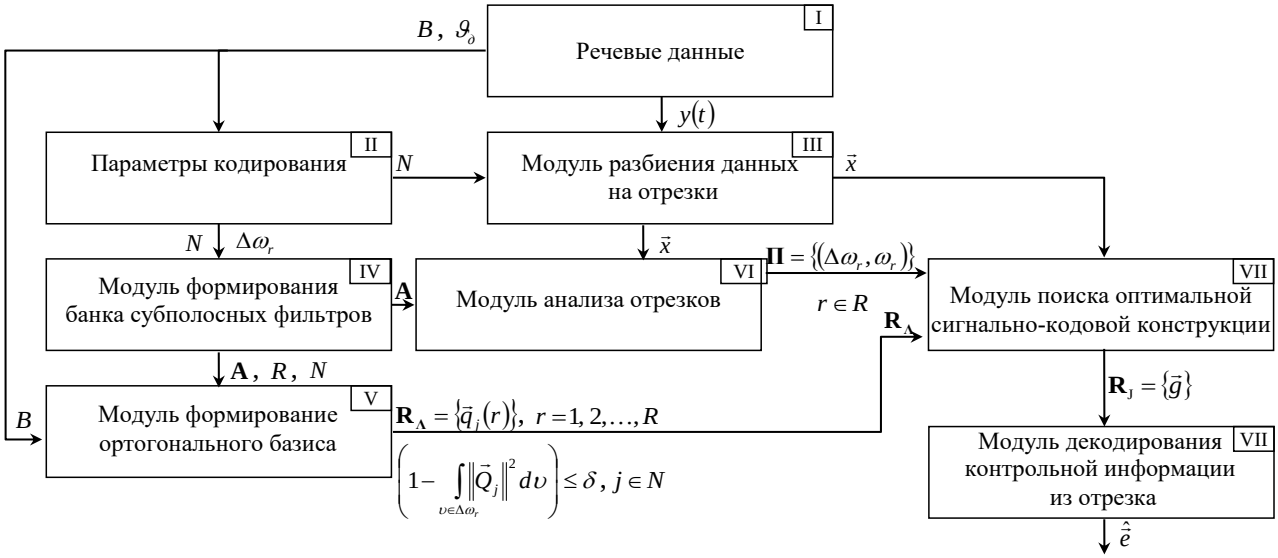


Рисунок 3.9 Структурная схема программной поддержки скрытного декодирования контрольной информации из данных

– модуль записи сигнала содержащего устную речь и контрольную информацию в виде речевых данных с разрядностью и частотой дискретизации;

– модуль чтения речевых данных содержащей фрагмент устной речи и контрольную информацию;

– модуль декодирования контрольной информации. Преобразующий сигнально-кодовую конструкцию отображающую контрольную информацию в символьный вид приемлемый для её восприятия.

3.4 Основные результаты и выводы главы

В качестве платформы для реализации программно-алгоритмической поддержки разработанных алгоритмов была выбрана ЭВМ класса IBM PC, с операционной системой на базе ядра Windows NT.

ЗАКЛЮЧЕНИЕ

В работе были представлены результаты анализа цифрового представления информации. В частности, подходы к кодированию идентификатора отправителя. Оценена емкость одной звукозаписи.

При этом отмечается, что в зависимости от разрядности АЦП на одно значение амплитуды f_i отводиться от 8 до 32 бит, но наибольшее распространение получило использование амплитуд с разрядностью 16 бит.

Из приведенного анализа длительности голосовых звукозаписей в мессенджерах (Viber, WhatsApp, Vk), можно заключить, что средняя длина звукозаписи составляет от четырех до шести секунд. Учитывая частоту дискретизации 8кГц, длительность защищаемой звукозаписи (без сжатия) в среднем составляет 40000 значений амплитуды. При этом не менее 80% составляют паузы между словами. Было показано целесообразность выбора размерности отрезка анализа в 128 значений амплитуды. Учитывая наложенные выше ограничения, в среднем из 312 отрезков звукозаписи для внедрения пригодны не более 94.

Рассмотренные алгоритмы внедрения информации на основе метода наименее значащего бита (LSB), метода расширения спектра (SSp), метода основанного на дискретно-косинусном преобразовании (DCT) и метода внедрения основанного на разложении по собственным векторам (QR).

Из описанных методов только последние два (DCT и QR), позволяют поместить и достоверно извлечь до 8 бит информации с обеспечением их скрытности. Иными словами, эти методы позволяют внедрить в одну секунду звукозаписи около 500 бит информации.

Так же отмечается, что в связи с сжатием звукозаписи внедренная информация может быть утеряна, а также существующие методы стеганографии не всегда позволяют однозначно восстановить внедренную информацию. В связи с этим предлагается использовать методы

помехоустойчивого кодирования, для повышения целостности внедренной информации.

Разработанный метод по формированию цифрового отпечатка включает: алгоритм хеширования SHA-2 (Secure Hash Algorithm) для случая с длиной дайджеста сообщения 256 (бит) и алгоритм помехоустойчивого кодирования БЧХ (15, 11, 5) с шестнадцатеричным кодом Рида - Соломона. Соответственно размерность цифрового отпечатка составит 512 бит.

В связи с размерностью цифрового отпечатка и доступным объемом внедряемой информации, внедрение предполагается осуществлять на с использованием методов, основанных на дискретно-косинусном преобразовании (DCT), замещая при этом наименее значащие биты.

В работе проведена проверка работоспособности разработанного метода внедрения цифровых отпечатков для подтверждения целостности звукозаписи, которая показала, что в 60 процентах случаев целостность будет подтверждена.

По полученным результатам можно сделать вывод о том, что необходимо уменьшить размерность цифрового отпечатка до 64 бит (в 8 меньше существующих), что позволит повысить помехоустойчивость отпечатка. А также целесообразно использовать модифицированные стеганографические алгоритмы позволяющие внедрить информацию в параметры компонент, не подвергающиеся изменению.

СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Advances in Audio and Speech Signal Processing: Technologies and Applications: Technologies and Applications. / Perez-Meana H.: Igi Global, 2007.
2. Ahmed N., Natarajan T., Rao K. R. Discrete cosine transform // IEEE transactions on Computers. - 1974. - Т. 100, № 1. - С. 90-93.
3. Al-Haj A., Mohammad A. A., Bata L. DWT-based audio watermarking // Int. Arab J. Inf. Technol. - 2011. - Т. 8, № 3. - С. 326-333.
4. Arnold M., Chen X. M., Baum P., Gries U., Doerr G. A Phase-Based Audio Watermarking System Robust to Acoustic Path Propagation // Ieee Transactions on Information Forensics and Security. - 2014. - Mar. - Т. 9, № 3. - С. 411-425.
5. Audio watermark: A comprehensive foundation using MATLAB. / Lin Y. a. L. W. H.: Springer International Publishing, 2015.
6. Bidelman G.M., Jennings S.G., Strickland E.A. PsyAcoustX: A flexible MATLAB® package for psychoacoustics research // Front. Psychol. 2015. Vol. 6, № OCT. P. 1–11.
7. Boll S. Suppression of acoustic noise in speech using spectral subtraction // IEEE Transactions on acoustics, speech, and signal processing. - 1979. - Т. 27, № 2. - С. 113-120.
8. Bos C.E. Masking and Discrimination // J. Acoust. Soc. Am. 1966. Vol. 39, № 4. P. 708.
9. Brungart D.S. et al. Informational and energetic masking effects in the perception of multiple simultaneous talkers. // J. Acoust. Soc. Am. 2001. Vol. 110, № 5 Pt 1. P. 2527–2538.
10. Carlile S. Psychoacoustics // The Sonification Handbook. 2011. P. 41–61.
11. Cheddad A. et al. Digital image steganography: Survey and analysis of current methods // Signal Processing. 2010. Vol. 90, № 3. P. 727–752.

12. Chen B., Wornell G.W. Quantization index modulation: A class of provably good methods for digital watermarking and information embedding // IEEE Trans. Inf. Theory. 2001. Vol. 47, № 4. P. 1423–1443.

13. Cohen A., Holmgren J., Nishimaki R., Vaikuntanathan V., Wichs D. Watermarking cryptographic capabilities // Proceedings of the forty-eighth annual ACM symposium on Theory of Computing. - Cambridge, MA, USA: ACM, 2016. - C. 1115-1127.

14. Cohen A., Holmgren J., Nishimaki R., Vaikuntanathan V., Wichs D. Watermarking cryptographic capabilities // Proceedings of the forty-eighth annual ACM symposium on Theory of Computing. - Cambridge, MA, USA: ACM, 2016. - C. 1115-1127.

15. Cox I. J. a. K. J. a. L. T. a. S. T. Secure, imperceptable yet perceptually salient, spread spectrum watermark for multimedia // IEEE. - 1996. - C. 192-197.

16. Cvejic N., Seppanen T. Increasing the capacity of LSB-based audio steganography // Proceedings of 2002 IEEE Workshop on Multimedia Signal Processing, MMSP 2002. 2002.

17. Digital communication: Communication, multimedia, security. / Meinel C., Sack H.: Springer Science & Business Media, 2014.

18. Digital watermarking. / Cox I. J., Miller M. L., Bloom J. A., Honsinger C.: Springer, 2002.

19. Freyman R.L., Balakrishnan U., Helfer K.S. Effect of number of masking talkers and auditory priming on informational masking in speech recognition // J. Acoust. Soc. Am. 2004. Vol. 115, № 5. P. 2246–2256.

20. Glasberg B.R., Moore B.C.J. Derivation of auditory filter shapes from notched-noise data // Hear. Res. 1990. Vol. 47, № 1-2. P. 103–138.

21. Hanselman D., Littlefield B. Mastering MATLAB® 5: A Comprehensive Tutorial and Reference // The MATLAB® Curriculum Series. 1998. P. xviii + 638.

22. Hartung F., Kutter M. Multimedia watermarking techniques // Proc. IEEE. 1999. Vol. 87, № 7. P. 1079–1107.

23. Hu H.T., Hsu L.Y. A DWT-Based Rational Dither Modulation Scheme for Effective Blind Audio Watermarking // Circuits, Syst. Signal Process. Springer US, 2016. Vol. 35, № 2. P. 553–572.

24. Huang H. H. H., Rahardja S., Rongshan Y., Xiao L. X. L. A fast algorithm of integer MDCT for lossless audio coding // IEEE International Conference on Acoustics, Speech, and Signal Processing. - 2004.

25. Kirovski D., Malvar H. Robust spread-spectrum audio watermarking. P. 3–6.

26. Kirovski D., Malvar H. S. Spread-spectrum watermarking of audio signals // IEEE transactions on signal processing. - 2003. - Т. 51, № 4. - С. 1020-1033.

27. Kirovski D., Malvar H.S. Spread-spectrum watermarking of audio signals // IEEE Trans. Signal Process. 2003. Vol. 51, № 4. P. 1020–1033.

28. Kyr.R. Steganography-The Art of Hiding Data // Inf. Technol. J. 2004. Vol. 3, № 3. P. 245–269.

29. Lang A., Dittmann J., Spring R., Vielhauer C. Audio watermark attacks: from single to profile attacks // Proceedings of the 7th workshop on Multimedia and security. - New York, NY, USA: ACM, 2005. - С. 39-50.

30. Levitt H. Five guys named Moe: A tribute to Moe Bergman // Int. J. Audiol. 2014. Vol. 53, № 10. P. 774–776.

31. Levitt H. Transformed up- down methods in psychoacoustics // J. Acoust. Soc. Am. 1971. Vol. 49, № 2 pt 2. P. 467–477.

32. Lofberg J. YALMIP : a toolbox for modeling and optimization in MATLAB // 2004 IEEE Int. Conf. Comput. Aided Control Syst. Des. 2004. P. 284–289.

33. MacLean K. VoxForge // Ken MacLean.[Online]. Available: http://www.voxforge.org/home. [Acedido em 2012].

34. Petitcolas F. La cryptographie militaire, 1883.

35. Pinel J. et al. A high-capacity watermarking technique for audio signals based on MDCT-domain quantization // Int. Congr. Acoust. 2010. № August. P. 1–7.

36. Podilchuk C.I., Delp E.J. Digital watermarking: Algorithm and application // IEEE Signal Process. Mag. 2001. Vol. 18, № 4. P. 33–46.

37. Schyndel R.G. Van, Tirkel A.Z., Osborne C.F. A digital watermark // Proc. 1st Int. Conf. Image Process. 1994. Vol. 2. P. 86–90.

38. Seok J., Hong J., Kim J. A novel audio watermarking algorithm for copyright protection of digital audio // ETRI J. 2002. Vol. 24, № 3. P. 181–189.

39. Signal processing, perceptual coding and watermarking of digital audio: Advanced technologies and models. / He X.: IGI Global, 2011.

40. Soranzo A., Grassi M. Psychoacoustics: A comprehensive MATLAB toolbox for auditory testing // Front. Psychol. Frontiers Research Foundation, 2014. Vol. 5, № JUL.

41. Viemeister N.F. Temporal modulation transfer functions based upon modulation thresholds // J. Acoust. Soc. Am. 1979. Vol. 66, № 5. P. 1364–1380.

42. W.Bender et al. Techniques for data hiding // J. IBM Syst. 1996. Vol. 35, № 3-4. P. 313–336.

43. Zmudzinski S., Steinebach M. Psycho-acoustic model-based message authentication coding for audio data // Proceedings of the 10th ACM workshop on Multimedia and security. - Oxford, United Kingdom: ACM, 2008. - С. 75-84.

44. Алдошина И. Основы психоакустики // М.: Оборонгиз. - 2000.

45. Библев Д. И. ГОСТ 16600-72. Передача речи по трактам радиотелефонной связи. Требования к разборчивости речи и методы артикуляционных измерений 2003. Систем, требования: URL: https://www.researchgate.net/publication/312167036_Recording_Gost_16600-72

46. Вариационные методы анализа и построения функций по эмпирическим данным на основе частотных представлений. / Жиляков Е. Г. - Белгород: Изд-во БелГУ, 2007. - 160 с. с.

47. Верификация и валидация моделей для инженерных расчетов. - 2017. - URL: https://multiphysics.ru/stati/blog/verifikatciia-i-validatciia-modelei-dlia-inzhenernykh-raschetov.htm (дата обращения: 12.10.2017.2017).

48. Верификация и валидация моделей для инженерных расчетов. - 2017. - URL: https://multiphysics.ru/stati/blog/verifikatciia-i-validatciia-modelei-dlia-inzhenernykh-raschetov.htm (дата обращения: 12.10.2017.2017).

49. ГОСТ Р. Р 34.10-2001—Криптографическая защита информации // Процедуры выработки и проверки электронной цифровой подписи на базе ассиметричного криптографического алгоритма.

50. ГОСТ Р. Р 34.10-2001—Криптографическая защита информации // Процедуры выработки и проверки электронной цифровой подписи на базе ассиметричного криптографического алгоритма.

51. ГОСТ Р. Р 34.11–94. Информационная технология. Криптографическая защита информации. Функция хэширования // М.: Госстандарт России. - 1994.

52. ГОСТ Р. Р 34.11–94. Информационная технология. Криптографическая защита информации. Функция хэширования // М.: Госстандарт России. - 1994.

53. Жиляков Е. Г. Анализ речевых сигналов и изображений, 2016.

54. Жиляков Е. Г., Бабаринов С. Л. Феноменологическая математическая модель психоакустики слуха человека // Научные ведомости БелГУ. - 2017. - Т. 43, № 16 (265). - С. 8.

55. Жиляков Е. Г., Белов С. П., Медведева А. А., Курлов А. В., Лихолоб П. Г. Исследование чувствительности решающей функции при обнаружении частотных интервалов в условиях воздействия помех // Инфокоммуникационные технологии. - 2016. - Т. 14, № 2. - С. 122-129.

56. Жиляков Е. Г., Белов С. П., Медведева А. А., Курлов А. В., Лихолоб П. Г. Об одном алгоритме определения информационных частотных интервалов // Наука. Инновации. Технологии. - 2016. - С. 23-30.

57. Жиляков Е. Г., Лихолоб П. Г., Девицына С. Н. Определение возможного объёма внедряемой информации при скрытой передаче меток в речевых данных // Научные ведомости Белгородского государственного университета. Серия: Экономика. Информатика. - 2012. - Т. 23, № 13-1 (132). - С. 222-226.

58. Жиляков Е. Г., Лихолоб П. Г., Медведева А. А., Прохоренко Е. Н. Исследование чувствительности некоторых мер качества скрытия информации в речевых сигналах // Научные ведомости Белгородского государственного университета. Серия: Экономика. Информатика. - 2016. - Т. 9, № 230. - С. 174-179.

59. Жиляков Е. Г., Пашинцев В. П., Белов С. П., Лихолоб П. Г. О методе скрытного кодирования контрольной информации в речевые данные // Инфокоммуникационные технологии. - 2015. - Т. 13, № 3. - С. 325-333.

60. Жиляков Е., Лихолоб П., Цыбина Я., Лихогодина Е. О применении стеганографических методов для аутентификации сигналов, содержащих речевое сообщение // Научные ведомости Белгородского государственного университета. Серия: Экономика. Информатика. - 2017. - Т. 42, № 9 (258). - С. 10.

61. Захаров И. Д., Ожиганов А. А. Использование порождающих полиномов М-последовательностей при построении псевдослучайных кодовых шкал // Известия высших учебных заведений. Приборостроение. - 2011. - Т. 54, № 6.

62. И.А.Алдошина. Основы психоакустики. Санкт-Петербург: Оборонгиз, 2000. 154 р.

63. И.А.Вартанян. Слуховая оценка приближения и удаления источника звука: психоакустические корреляты и нейрофизиологические механизмы // Информационные технологии. 1999. Vol. 30, № 1. P. 1900.

64. Марков А. С., Цирлов В. Л., Барабанов А. В. Методы оценки несоответствия средств защиты информации //. - 2012.

65. Марков А. С., Цирлов В. Л., Барабанов А. В. Методы оценки несоответствия средств защиты информации //. - 2012.

66. Марпл С. Л. Цифровой спектральный анализ и его приложения (нет оглавл.) //. - 1990. - Т. 2. - С. 234.

67. Меркушева А. Фильтрация нестационарного сигнала (речи) в вейвлет-области с адаптацией к виду и динамике шума // Научное приборостроение. - 2003. - Т. 13, № 2. - С. 73-87.

68. Н.В.Пинчук. Н . В . Пинчук // Информационные технологии. 2004. Vol. 44, № 1. P. 2004.

69. Синицын С., Налютин Н. Верификация программного обеспечения // М.: БИНОМ. - 2008. - Т. 368. - С. 6.

70. Синицын С., Налютин Н. Верификация программного обеспечения // М.: БИНОМ. - 2008. - Т. 368. - С. 6.

71. Теория, применение и оценка качества генераторов псевдослучайных последовательностей. / Иванов М., Чугунков И.: Кудиц-Образ М., 2003.

72. Цифровая обработка сигналов. / Лайонс Р.: Бином, 2007.

73. Цифровая стеганография. / Грибунин В. Г., Оков И. Н., Туринцев И. В.: СОЛОН-Пресс, 2016.

74. Шелухин О. И., Канаев С. Д. Оценка качества неподвижного изображения при стеганографическом скрытии цифровых водяных знаков методом расширения спектра // Наукоемкие технологии в космических исследованиях земли. - 2016. - Т. 6, № 8. - С. 59-64.

75. Э.И.Вологдин. Аудиотехника. Москва: Горячая линия–Телеком, 2013. 742 p.

76. Электроакустика и звуковое вещание. Для высших учебных заведений. / Алдошина, Вологдин, Ефимов, Катунин, Кацнельсон - М.: Горячая линия-Телеком, 2007. Для высших учебных заведений. - 872 с. с.

77. Электроакустика и звуковое вещание. Для высших учебных заведений. / Алдошина И. А., Вологдин, Ефимов, Катунин, Кацнельсон - М.: Горячая линия-Телеком, 2007. Для высших учебных заведений. - 872 с. с.

78. Электроакустические измерения и оценка качества звучания. / Алдошина И. А. - СПб, 1998. - 65 с. с.

79. Ю.А.Ковалгин. Цифровое кодирование звуковых сигналов. Москва: КОРОНА принт, 2004. 160 p.

ПРИЛОЖЕНИЕ А

ЛИСТИНГ ПРОГРАММНОГО КОДА ПРЕДВАРИТЕЛЬНОЙ ОБРАБОТКИ КОНТРОЛЬНОЙ ИНФОРМАЦИИ

```
% Функция работы со спектром от RES
% Синтаксис:
%   p=fftR(r,sig,dots/nf)
% Параметры:
%   r - режим работы
%       0 - Найти спектр сигнала sig. К-во точек равно к-ву отсч. sig или nf
(третий пр-р)
%       1 - Обратное преобразование фурье sig
%       2 - Вычислить спектр в заданных точке(ах) dots
%       3 - Обратное Преобразование (точек)
%       4 - фазовую хар-ку
%   sig - сигнал (временной или частотный) для преобразования
% Необязательные параметры (для 0 и 1):
%   nf - кол-во точек для частотной области
%   dots - массив координат точек для вычисления
function p=fftR(varargin)
switch size(varargin,2)
    case 0
        disp('Где параметры???');
        r=-1;
    case 1
        r=-2;
        disp('Где сигнал???');
    case 2
        r=varargin{1};
        sig=varargin{2};
        d=length(sig);
    case 3
        r=varargin{1};
        sig=varargin{2};
        d=varargin{3};
end
if r==0 % Вычислим спектр использовав d точек
    p=zeros(1,d);
    N=length(sig);
    n=1:N;
    if length(sig)==d
        for k=1:d
            p(k)=sum(sig(n).*exp(-1i*2*pi*(k-1)*(n-1)/N));
        end
    else
        di=linspace(1,N,d);
        for k=1:d
            p(k)=sum(sig(n).*exp(-1i*2*pi*(di(k)-1)*(n-1)/N));
        end
    end
end
if r==1 % Вычислим спектр использовав d точек
    N=length(sig);
    p=zeros(1,d);
    k=1:N;
    if length(sig)==d
        for n=1:d
            p(n)=real((1/d)*sum(sig(k).*exp(1i*2*pi*(k-1)*(n-1)/N)));
```

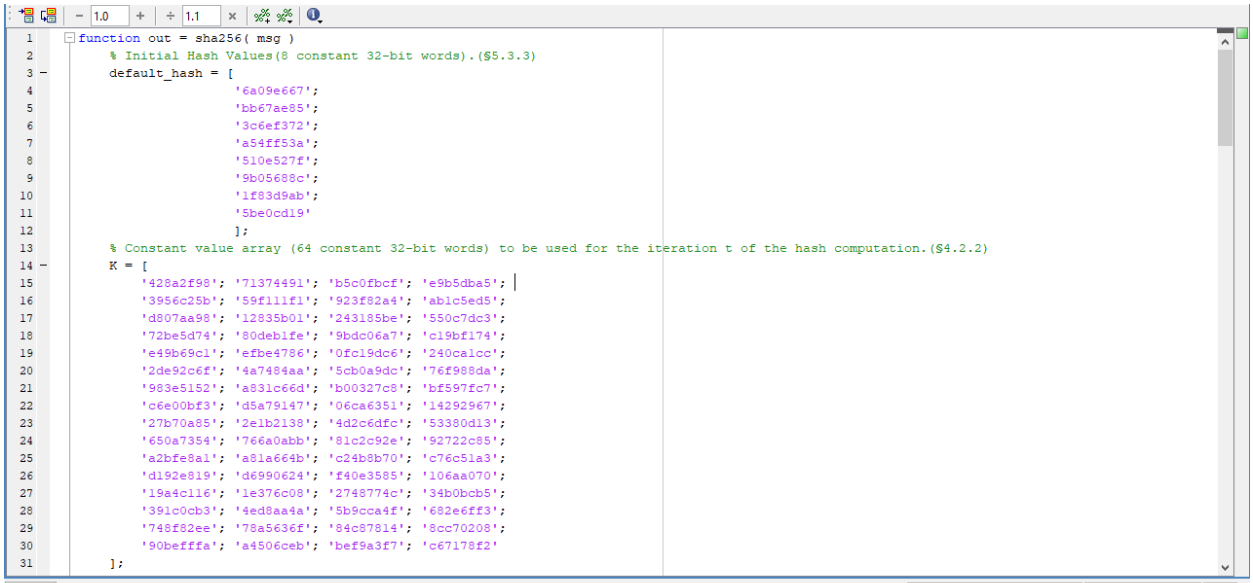
```
        end
    else
        di=linspace(1,N,d);
        for n=1:d
            p(n)=real((1/length(di))*sum(sig(k).*exp(1i*2*pi*(k-1)*(di(n)-
1)/N)));
        end
    end
end
if r==2 % Вычислим спектр использовав d
    p=zeros(1,length(d));
    N=length(sig);
    n=1:N;
    for k=1:length(d)
        p(k)=sum(sig(n).*exp(-1i*2*pi*(d(k)-1)*(n-1)/N));
    end
end
if r==3 % Вычислим спектр использовав d точек
    N=length(sig);
    p=zeros(1,length(d));
    k=1:N;
    for n=1:length(d)
        p(n)=real((1/N)*sum(sig(k).*exp(1i*2*pi*(k-1)*(d(n)-1)/N)));
    end
end
if r==4 % Вычислим фазу использовав d точек
    p=zeros(1,d);
    N=length(sig);
    n=1:N;
    if size(sig,2)==1
        sig=sig';
    end
    for k=1:d
        p(k)=sum(sig(n).*exp(-1i*2*pi*(k-1)*(n-1)/N));
    end
    p=atan(imag(p)./real(p));
end
```

ПРИЛОЖЕНИЕ Б

Хэширование контрольной информации

(GIT, sha256)

https://github.com/lostpfg/SHA-256-Matlab

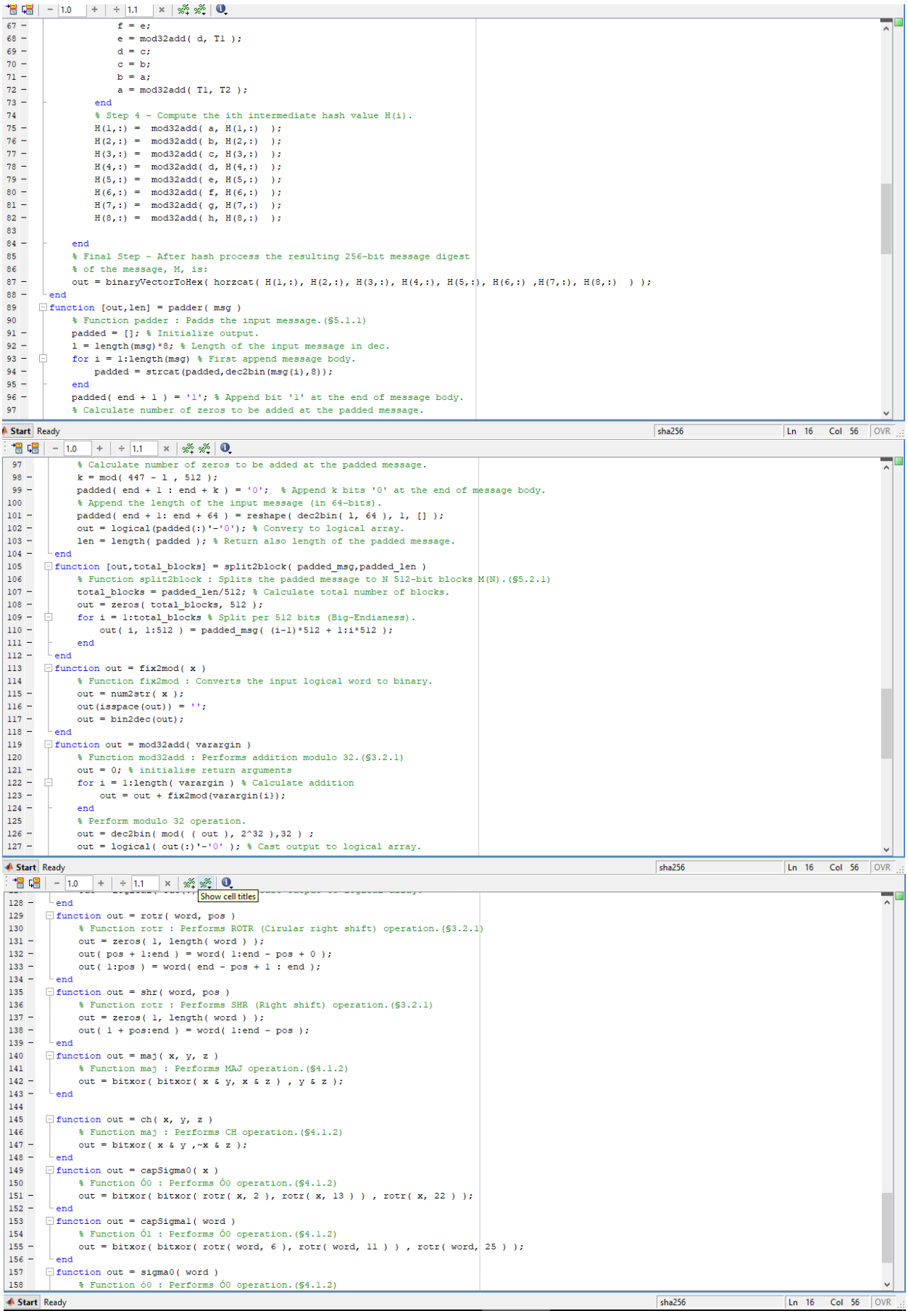


```
function out = sha256( msg )
    % Initial Hash Values(8 constant 32-bit words).(§5.3.3)
    default_hash = [
                    '6a09e667';
                    'bb67ae85';
                    '3c6ef372';
                    'a54ff53a';
                    '510e527f';
                    '9b05688c';
                    '1f83d9ab';
                    '5be0cd19'
                    ];
    % Constant value array (64 constant 32-bit words) to be used for the iteration t of the hash computation.(§4.2.2)
    K = [
        '428a2f98'; '71374491'; 'b5c0fbcf'; 'e9b5dba5';
        '3956c25b'; '59f111f1'; '923f82a4'; 'ab1c5ed5';
        'd807aa98'; '12835b01'; '243185be'; '550c7dc3';
        '72be5d74'; '80deb1fe'; '9bdc06a7'; 'c19bf174';
        'e49b69c1'; 'efbe4786'; '0fc19dc6'; '240calcc';
        '2de92c6f'; '4a7484aa'; '5cb0a9dc'; '76f988da';
        '983e5152'; 'a831c66d'; 'b00327c8'; 'bf597fc7';
        'c6e00bf3'; 'd5a79147'; '06ca6351'; '14292967';
        '27b70a85'; '2e1b2138'; '4d2c6dfc'; '53380d13';
        '650a7354'; '766a0abb'; '81c2c92e'; '92722c85';
        'a2bfe8a1'; 'a81a664b'; 'c24b8b70'; 'c76c51a3';
        'd192e819'; 'd6990624'; 'f40e3585'; '106aa070';
        '19a4c116'; '1e376c08'; '2748774c'; '34b0bcb5';
        '391c0cb3'; '4ed8aa4a'; '5b9cca4f'; '682e6ff3';
        '748f82ee'; '78a5636f'; '84c87814'; '8cc70208';
        '90befffa'; 'a4506ceb'; 'bef9a3f7'; 'c67178f2'
    ];
```

Start Ready sha256 Ln 15 Col 57 OVR

```
    % Split padded message to N (512-bit) blocks.(§6)
    [M,total_blocks] = split2block( padded_msg,padded_len );
    W = zeros( 64, 32 );
    H = zeros( 8, 32 );
    % Main SHA-256 compuation process.(§6.2.2)
    for j = 1:8 % Load initial hash values at first iteration.
        H(j,:) = hexToBinaryVector( default_hash( j , : ) , 32 );
    end
    for i = 1:total_blocks % For every block M(i).
        % Step 1 - Prepare the message schedule.
        for j = 1:64
            if j >= 1 && j <= 16
                W( j, 1:32 ) = M( i, 32*(j-1)+1:j*32 );
            else
                W( j, 1:32 ) = mod32add( sigma1( W( j-2, : ) ), W( j-7, : ) , sigma0( W( j-15, : ) ), W( j-16, : ) );
            end
        end
        % Step 2 - Initialize the eight working variables, a, b, c, d, e, f, g,
        % and h, with the (i-1)st hash value.
        a = H(1,:);
        b = H(2,:);
        c = H(3,:);
        d = H(4,:);
        e = H(5,:);
        f = H(6,:);
        g = H(7,:);
        h = H(8,:);
        % For t=0 to 63.
        for t = 1:64
            T1 = mod32add( h, capSigma1(e),ch( e, f, g ),hexToBinaryVector( K( t,: ) , 32 ), W( t,:  ) );
            T2 = mod32add( capSigma0(a), maj( a, b ,c )  );
```

Start Ready sha256 Ln 16 Col 56 OVR



```
            f = e;
            e = mod32add( d, T1 );
            d = c;
            c = b;
            b = a;
            a = mod32add( T1, T2 );
        end
        % Step 4 - Compute the ith intermediate hash value H(i).
        H(1,:) =  mod32add( a, H(1,:)  );
        H(2,:) =  mod32add( b, H(2,:)  );
        H(3,:) =  mod32add( c, H(3,:)  );
        H(4,:) =  mod32add( d, H(4,:)  );
        H(5,:) =  mod32add( e, H(5,:)  );
        H(6,:) =  mod32add( f, H(6,:)  );
        H(7,:) =  mod32add( g, H(7,:)  );
        H(8,:) =  mod32add( h, H(8,:)  );

    end
    % Final Step - After hash process the resulting 256-bit message digest
    % of the message, M, is:
    out = binaryVectorToHex( horzcat( H(1,:), H(2,:), H(3,:), H(4,:), H(5,:), H(6,:) ,H(7,:), H(8,:)  ) );
end
function [out,len] = padder( msg )
    % Function padder : Padds the input message.(§5.1.1)
    padded = []; % Initialize output.
    l = length(msg)*8; % Length of the input message in dec.
    for i = 1:length(msg) % First append message body.
        padded = strcat(padded,dec2bin(msg(i),8));
    end
    padded( end + 1 ) = '1'; % Append bit '1' at the end of message body.
    % Calculate number of zeros to be added at the padded message.
```

Start Ready sha256 Ln 16 Col 56 OVR

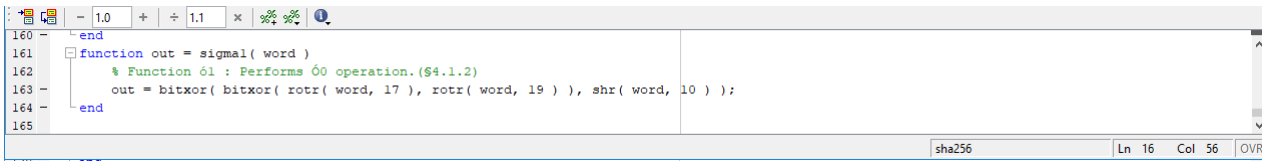
```
    % Calculate number of zeros to be added at the padded message.
    k = mod( 447 - l , 512 );
    padded( end + 1 : end + k ) = '0';  % Append k bits '0' at the end of message body.
    % Append the length of the input message (in 64-bits).
    padded( end + 1: end + 64 ) = reshape( dec2bin( l, 64 ), 1, [] );
    out = logical(padded(:)'-'0'); % Convery to logical array.
    len = length( padded ); % Return also length of the padded message.
end
function [out,total_blocks] = split2block( padded_msg,padded_len )
    % Function split2block : Splits the padded message to N 512-bit blocks M(N).(§5.2.1)
    total_blocks = padded_len/512; % Calculate total number of blocks.
    out = zeros( total_blocks, 512 );
    for i = 1:total_blocks % Split per 512 bits (Big-Endianess).
        out( i, 1:512 ) = padded_msg( (i-1)*512 + 1:i*512 );
    end
end
function out = fix2mod( x )
    % Function fix2mod : Converts the input logical word to binary.
    out = num2str( x );
    out(isspace(out)) = '';
    out = bin2dec(out);
end
function out = mod32add( varargin )
    % Function mod32add : Performs addition modulo 32.(§3.2.1)
    out = 0; % initialise return arguments
    for i = 1:length( varargin ) % Calculate addition
        out = out + fix2mod(varargin{i});
    end
    % Perform modulo 32 operation.
    out = dec2bin( mod( ( out ), 2^32 ),32 ) ;
    out = logical( out(:)'-'0' ); % Cast output to logical array.
```

Start Ready sha256 Ln 16 Col 56 OVR

```
end
function out = rotr( word, pos )
    % Function rotr : Performs ROTR (Cirular right shift) operation.(§3.2.1)
    out = zeros( 1, length( word ) );
    out( pos + 1:end ) = word( 1:end - pos + 0 );
    out( 1:pos ) = word( end - pos + 1 : end );
end
function out = shr( word, pos )
    % Function rotr : Performs SHR (Right shift) operation.(§3.2.1)
    out = zeros( 1, length( word ) );
    out( 1 + pos:end ) = word( 1:end - pos );
end
function out = maj( x, y, z )
    % Function maj : Performs MAJ operation.(§4.1.2)
    out = bitxor( bitxor( x & y, x & z ) , y & z );
end

function out = ch( x, y, z )
    % Function maj : Performs CH operation.(§4.1.2)
    out = bitxor( x & y ,~x & z );
end
function out = capSigma0( x )
    % Function Ó0 : Performs Ó0 operation.(§4.1.2)
    out = bitxor( bitxor( rotr( x, 2 ), rotr( x, 13 ) ) , rotr( x, 22 ) );
end
function out = capSigma1( word )
    % Function Ó1 : Performs Ó0 operation.(§4.1.2)
    out = bitxor( bitxor( rotr( word, 6 ), rotr( word, 11 ) ) , rotr( word, 25 ) );
end
function out = sigma0( word )
    % Function ó0 : Performs Ó0 operation.(§4.1.2)
```

Start Ready sha256 Ln 16 Col 56 OVR


1.0
1.1
end
function out = sigma1(word)
% Function ó1 : Performs Ó0 operation.(§4.1.2)
out = bitxor(bitxor(rotr(word, 17), rotr(word, 19)), shr(word, 10));
end
sha256
Ln 16
Col 56
OVR