

УДК 621.391

НИЗКОСКОРОСТНОЕ КОДИРОВАНИЕ РЕЧЕВОГО СИГНАЛА
ПРИ ПЕРЕМЕННОЙ ДЛИТЕЛЬНОСТИ СЕГМЕНТА АНАЛИЗА

А. А. Афанасьев, Р. С. Власов

Академия Федеральной службы охраны Российской Федерации (г. Орёл)

Поступила в редакцию 28.06.2019 г.

Аннотация. В статье изложены материалы по исследованию таких свойств речевого сигнала (РС), как однородность (подобность) распределения мгновенных значений отсчётов. Данное свойство используется при выделении сегментов обработки речи на основе последовательного статистического анализа. Параметры линейного предсказания (ЛП), рассчитываемые на каждом шаге расширения, претерпевают незначительные (в пределах шага квантования) изменения по отношению к начальному сегменту. Это обстоятельство позволяет использовать результаты анализа начального участка для восстановления участка, расширенного вышеуказанным способом, что позволяет существенно снизить как вычислительную сложность, так и скорость кодирования РС.

Ключевые слова: речевой сигнал, сегмент анализа, линейные спектральные частоты, шаг квантования, расширение сегмента.

ВВЕДЕНИЕ

В настоящее время, наблюдаются устойчивые тенденции роста числа потребителей речевых услуг и расширения речевого сервиса сетей и систем связи. Реализация этих тенденций требует увеличение числа и повышения эффективности использования имеющихся каналов связи. Особую актуальность эта задача приобрела в условиях интенсивного развития и все более широкого применения систем компьютерной телефонии (Computer-Telephony Integration, СТИ).

Дальнейшее развитие технологий, ориентированных на передачу неравномерно-распределённого во времени трафика позволило экономически эффективно реализовывать транспортировку речи в сетях, изначально ориентированных на передачу данных.

Таким образом, можно утверждать, что основными тенденциями развития сетей и систем связи в настоящее время является следующее:

- повышение эффективности использования полосы пропускания каналов связи;
- динамическая адаптация к изменениям скорости информационных потоков;
- передача речевой информации;
- расширение речевого сервиса.

То есть, развитие современных телекоммуникационных технологий происходит при параллельном совершенствовании, с одной стороны, сетевых механизмов, а с другой – алгоритмов преобразования речи [1].

Высокое качество РС, передаваемого с помощью алгоритмов, в основе которых лежит метод линейного предсказания (ЛП), при относительно невысоких скоростях передачи обеспечило их широкое распространение в различных системах и сетях связи. Однако при обеспечении низкоскоростного кодирования со скоростями до 4 кбит/с, использование вышеуказанных способов приводит к снижению качества передачи.

Одним из направлений дальнейшего снижения скорости передачи речи с сохранением высокого качества – это переход к переменной скорости кодирования, учитывающие

информационные и статистические избыточности сегментов РС.

При ЛП, в качестве параметров, подвергаемых кодированию и передаче, используются значения передаточной функции голосового тракта и сигнала возбуждения.

Авторегрессионный анализ последовательностей значений отсчетов, является основой модели, представляющий РС в виде отклика линейной системы с переменными параметрами (голосового тракта) на соответствующий сигнал возбуждения (порождающий сигнал).

Рассчитываемые на основе последовательностей случайных величин (значении отсчетов) коэффициенты линейного предсказания (КЛП) a_m , характеризуют состояние линейной системы и сигнала возбуждения, позволяющие синтезатору восстановить исходный сигнал с требуемой степенью верности.

При этом полином передаточной функции синтезирующего фильтра имеет следующий вид:

$$A_M(Z) = 1 + \sum_{k=1}^M a_k z^{-k}. \quad (1)$$

Исследования в предметной области речевого кодирования показали невозможность прямого интерполирования и квантования КЛП (a_k) при переходе от одного субсегмента к другому в рамках единого однородного сегмента данных, так как это приводит к высокой вероятности получения неустойчивого фильтра синтеза модели предсказания [2, 3, 4].

Для описания передаточной функции наиболее предпочтительными являются параметры, предложенные Итакурой [3, 4] и названные линейные спектральные частоты (ЛСЧ) или линейные спектральные корни.

Алгоритм определения ЛСЧ заключается в поиске нулей тригонометрических вспомогательных полиномов, получаемых из передаточной функции (1). Данный способ определения ЛСЧ широко известен и описан в [5, 6]. В свою очередь имеются подходы к вычислению ЛСЧ, основанные на использовании статистических взаимосвязей в РС [11, 12].

Без потери общности в дальнейшем описании под параметрами ЛП будем понимать ЛСЧ.

Известные решения в области обработки РС используют фиксированные сегменты, на основании которых рассчитывают параметры состояния линейной системы и сигнала возбуждения, позволяющие синтезатору восстановить исходный сигнал с требуемой степенью верности, при этом входные отсчеты РС разделяются на сегменты фиксированной длины и для каждого из сегментов рассчитывают ЛСЧ. Выражение описывающее данный процесс выглядит следующим образом:

$$\vec{X} = \left\{ \begin{array}{l} \{\bar{X}_1\} \cup \{\bar{X}_2\} \cup \dots \cup \{\bar{X}_{m-1}\} \cup \{\bar{X}_m\} \cup \\ \cup \{\bar{X}_{m+1}\} \cup \dots \cup \{\bar{X}_{M-1}\} \cup \{\bar{X}_M\} \end{array} \right\}, \quad (2)$$

где $|\{\bar{X}_m\}| = \text{const}$, M – количество сегментов анализа, при фиксированной длине.

Для каждого из сегментов $\{\bar{X}_k\}$ рассчитывают коэффициенты формирующей модели, а затем обрабатывают с использованием алгоритма ЛП, так как декомпозиция РС на локальных временных интервалах достаточно хорошо соотносится с квазистационарной моделью РС на временных интервалах, соответствующих режиму установившихся звуков. При этом происходит разделение передаточной функции голосового тракта и возбуждающего сигнала. Недостатком такого подхода является использование фиксированной длительности сегмента анализа ($|\{\bar{X}_k\}| = \text{const}$), что не соотносится с природой возникновения и существования РС.

Недостатками существующих подходов к определению ЛСЧ являются следующее:

1. Отсутствие адаптивного изменения границ анализируемого сегмента РС на участках одинаковой природы (однородности) образования присутствующих в речи, и использование свойств однородности при определении ЛСЧ.

2. Значительный вычислительный ресурс, затрачиваемый на определение ЛСЧ на однородных сегментах.

По своей природе участки РС в зависимости от произносимых звуков можно разделить на три основных вида: вокализованные, невокализованные (шумовые) и участки со смешанным возбуждением (сонанты) [8]. На рис. 1–3 представлены участки РС, получен-

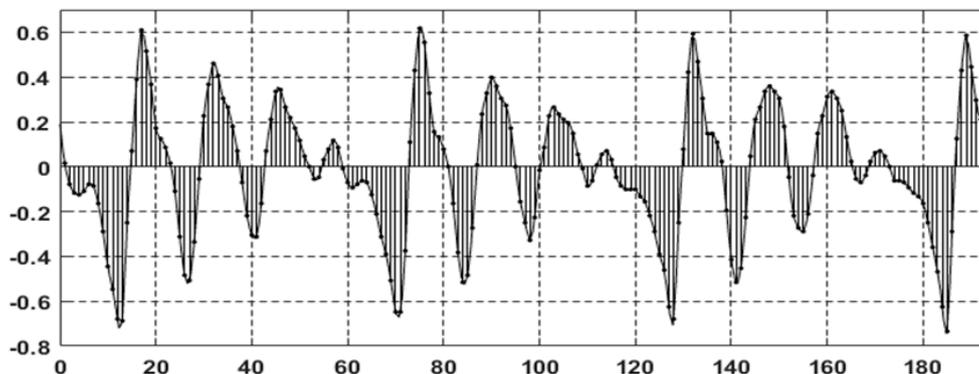


Рис. 1. Временное представление первых 24,5 мс звука «и» во фразе «Циифровая обработка сигналов» – вокализованный звук

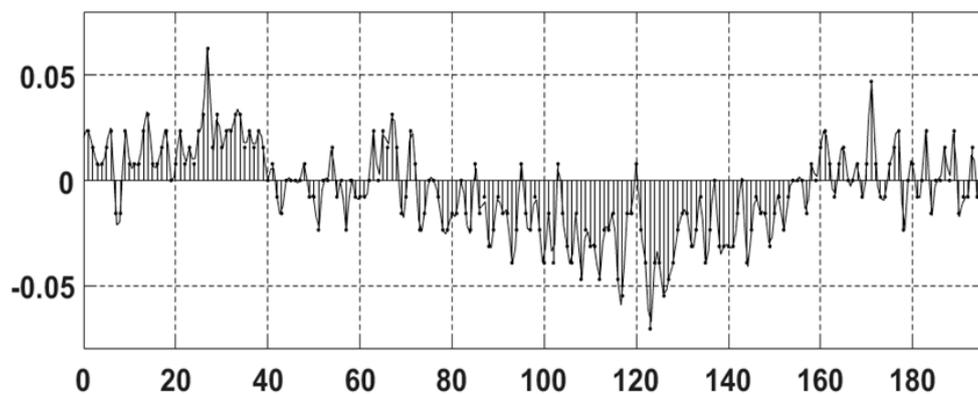


Рис. 2. Временное представление первых 24,5 мс звука «ф» во фразе «Цифровая обработка сигналов» – шумовой

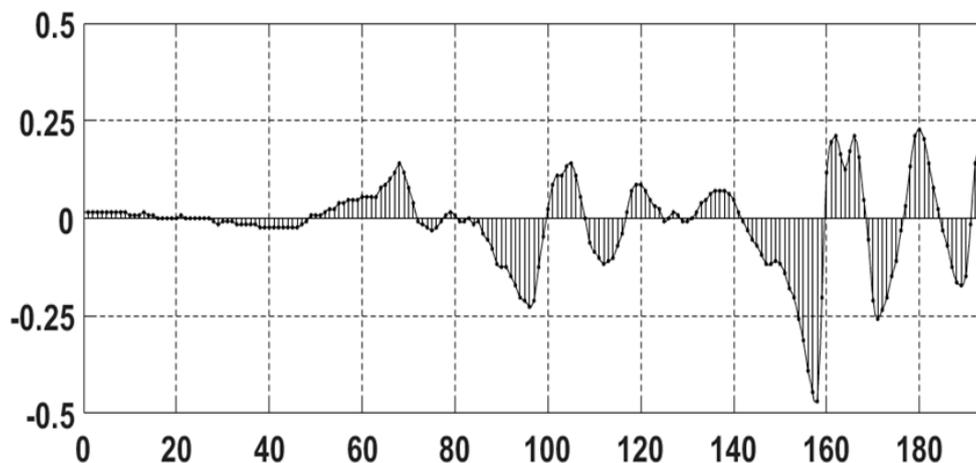


Рис. 3. Временное представление первых 24,5 мс звука «и» во фразе «Цифровая обработка сигналов» – сонанта

ные при произношении трёх разных звуков во фразе «цифровая обработка сигналов», полученные при равномерной дискретизации с частотой 8 кГц.

Как видно из представленных рисунков вокализованные звуки обладают свойством «квазипериодичности», сохраняющим его на значительном (по сравнению с длительностью

фиксированных сегментов анализа) протяжении. Особое внимание заслуживают и сонантные звуки т. к., несмотря на отсутствие явно выраженной периодичности, налицо наличие определённых статистических взаимосвязей между мгновенными значениями отсчётов.

В табл. 1 и 2 представлены значения средней длительности гласных звуков (фонем),

Таблица 1

Средняя длительность гласных звуков

Звук	Длительность гласных звуков, с						Среднее значение длительности, с	
	в словах			во фразах				
	в середине слова		в начале и в конце слова		произвольное положение		ударные	неударные
	ударные	неударные	ударные	неударные	ударные	неударные		
"у"	0,19	0,11	0,25	–	0,17	0,1	0,2	0,11
"о"	0,22	–	0,25	–	0,17	0,12	0,2	0,14
"а"	0,23	0,1	0,26	0,18	0,18	0,08	0,21	0,08
"э"	0,2	0,08	0,25	0,19	0,14	0,11	0,18	0,12
"и"	0,18	0,1	0,26	0,16	0,16	0,09	0,19	0,11
"ы"	0,19	0,1	–	0,17	0,18	0,08	0,19	0,11

Таблица 2

Средняя длительность согласных звуков

Типы согласных звуков		Длительность согласных звуков, с						
		твердые				мягкие		
Смычковые	Звонкие	"б"	"д"	"г"	"бь"	"дь"	"гь"	
		0,10	0,10	0,07	0,09	0,09	0,10	
	Глухие	"л"	"т"	"к"	"ль"	"ть"	"гь"	
		0,02	0,02	0,03	0,03	0,10	0,04	
Щелевые	Звонкие	"в"	"ж"	"з"	"вь"	"зь"		
		0,10	0,11	0,11	0,08	0,13		
	Глухие	"ф"	"ш"	"с"	"х"	"фь"	"шь"	"сь"
		0,15	0,14	0,14	0,12	0,12	0,24	0,15
Сонанты	Носовые	"м"	"н"	"мь"	"нь"			
		0,09	0,09	0,09	0,10			
	Щелевые	"л"	"ль"	"й"				
		0,08	0,08	0,08				
	Дрожащие	"р"	"рь"					
		0,06	0,04					
Аффрикаты	"ц"	"ч"						
	0,14	0,09						

полученные в результате анализа звукового состава русской речи.

Анализ данных, содержащихся в таблицах, показывает, что приведенные значения значительно превосходят принятую фиксированную длительность сегмента анализа речи в существующих системах обработки, это дает предпосылки для возможности снижения информативной избыточности пара-

метров РС, используемых при реализации алгоритмов низкоскоростного кодирования.

МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

Наличие статистических взаимосвязей в РС, которые сохраняются на протяжении большем, чем фиксированная длительность

сегментов анализа, используемая в существующих системах кодирования, создает предпосылки перехода от фиксированной длины сегмента анализа, к использованию сегментов анализа, границы которых адаптивно формируются в зависимости от природы образования РС, и, следовательно, являются переменными по длине. Очевидно, что в этом случае выражение, описывающее декомпозицию РС в системах обработки на основе метода ЛП примет с учётом адаптации к природе образования, примет следующий вид:

$$\bar{X} = \left\{ \left\{ \bar{X}_1 \right\} \cup \left\{ \bar{X}_2 \right\} \cup \dots \cup \left\{ \bar{X}_{k-1} \right\} \cup \left\{ \bar{X}_k \right\} \cup \left\{ \bar{X}_{k+1} \right\} \cup \dots \cup \left\{ \bar{X}_{K-1} \right\} \cup \left\{ \bar{X}_K \right\} \right\}, \quad (3)$$

где K – количество сегментов анализа при адаптивном формировании границ $\left\{ \bar{X}_k \right\} = \text{var}$.

Существуют технические решения использующие переменную длину сегмента анализа, которые основаны на увеличении длины анализируемого сегмента. Они основаны на использовании корреляционных связей между отсчётами РС, которые выявляются с помощью анализа автокорреляционной функции (АКФ). [9].

Однако применение данных лишь о корреляционных связях между отсчётами, во-первых: предполагает нормальность распределения результатов наблюдений, а во вторых не полностью устраняет статистическую избыточность РС при кодировании связанную

с выбором длительности сегмента анализа. В задачах прикладного характера распределение значений РС отличается от нормального, однако сам вид распределения остаётся неизвестным. Таким образом, увеличение границ анализируемого сегмента может быть основано на использовании более полной информации о статистических взаимосвязях между мгновенными значениями РС, получаемой с помощью непараметрических методов, т.е. при отсутствии данных о характере распределения.

Наиболее предпочтительным способом, учитывающим статистические взаимосвязи и «свободным» от допущения о нормальном законе распределения мгновенных значений отсчётов РС является способ выделения сегментов обработки речи на основе последовательного статистического анализа. Этот способ использует методы последовательной проверки статистических гипотез в сочетании с непараметрическими подходами к определению параметров распределения значений отсчётов, составляющих анализируемый сегмент [10].

Исследованиями доказано, что ЛСЧ на сегменте анализа, границы которого адаптивно расширены вышеуказанным способом, не претерпевают значительных изменений по отношению к начальному в 20 мс. Данное свойство однородности может быть использовано для снижения вычислительного ре-

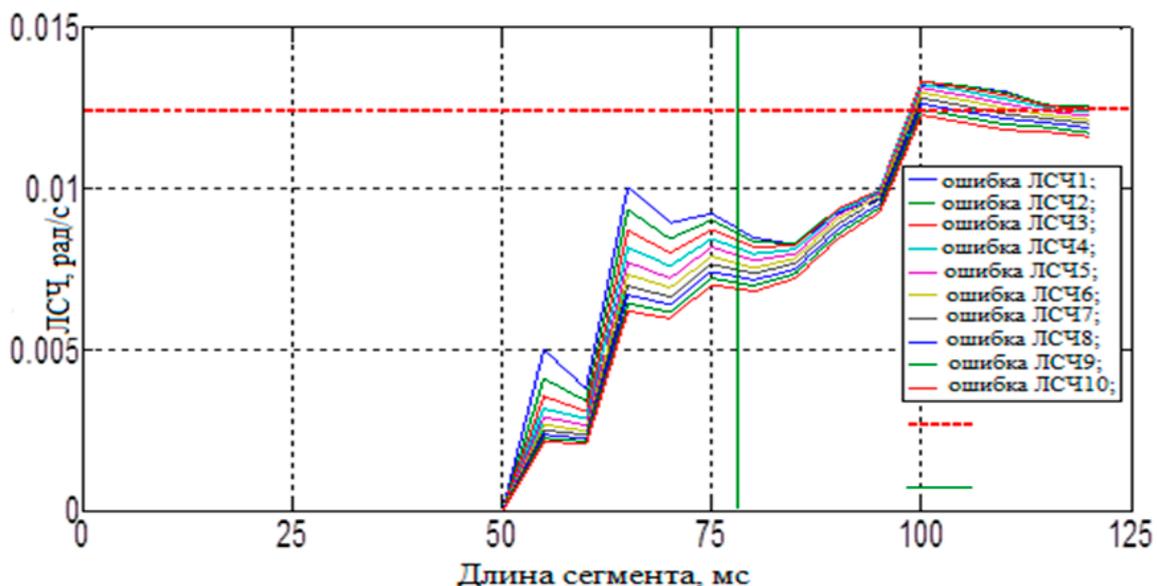


Рис. 4. Отклонения (ошибки) ЛСЧ на однородных сегментах по отношению к начальному

сурса затрачиваемого для определения параметров предсказания, а также количества бит выделяемых для их кодирования.

Анализ результатов исследований, приведённых на рис. 1, показал, что отклонение ЛСЧ от значений, вычисленных на начальных сегментах, не превышает величину шага квантования (красная пунктирная линия). Это свойство сохраняется на протяжении большем, чем максимально возможный размер анализируемого сегмента (80 мс, Рекомендации G.114 Международного союза электросвязи).

Данное обстоятельство позволяет вырабатывать подходы к формированию данных для кодирования сегмента РС, с сохранением параметров предсказания на всём протяжении участка РС одинаковой природы образования.

Разработанные решения могут быть использованы в речепреобразующих устройствах на основе ЛП, функционирующих в составе систем ориентированных на современные системы и сети связи при переходе и к протоколам с переменными скоростями, применением высокоскоростных мультиплексов потоков с переменной скоростью.

Наиболее удобным, с точки зрения минимизации искажений на стыках анализируемых сегментов является последовательное расширение начального сегмента длиной в 20 мс (160 отсчётов при частоте дискретизации 8 кГц) на величину периода основного тона с принятием границы вблизи нуля. То есть, имеет место следующее выражение:

$$\bar{X}_m = \left\{ \begin{array}{l} \left\{ \bar{X}_m^{120} \right\} \cup \left\{ \bar{X}_1^{OT_m} \right\} \cup \dots \\ \cup \left\{ \bar{X}_n^{OT_m} \right\} \cup \dots \cup \left\{ \bar{X}_N^{OT_m} \right\} \end{array} \right\}, \quad (4)$$

где $\left\{ \bar{X}_m^{120} \right\}$ – начальный сегмент анализа длиной 20 мс; $\left\{ \bar{X}_1^{OT_m} \right\}$ – множество отсчётов РС, на которое увеличивается сегмент анализа на 1-м этапе расширения; $\left\{ \bar{X}_n^{OT_m} \right\}_{n \neq 1}$ – множество отсчётов РС на протяжении периода основного тона m -го сегмента анализа; N – число периодов основного тона на которое был расширен сегмент.

Следует отметить что $\left| \left\{ \bar{X}_1^{OT_m} \right\} \right| \neq \left| \left\{ \bar{X}_n^{OT_m} \right\}_{n \neq 1} \right|$. Данное обстоятельство объясняется тем, что конечная граница начального сегмента анализа, как правило, отличается от нуля. Корректировка границы сегмента (принятие вблизи нуля) после первого расширения приводит к тому что размерность $\left| \left\{ \bar{X}_1^{OT_m} \right\} \right|$ отличается от количества отсчётов, составляющих участок на длине периода основного тона. При этом все последующие расширения $\left| \left\{ \bar{X}_n^{OT_m} \right\}_{n \neq 1} \right|$ полностью соответствуют ему.

Алгоритм реализации предлагаемого решения представлен на рис. 5

Очевидно, что динамическое выделение сегмента анализа РС приводит к переменной скорости кодирования. Значение данного параметра будет снижаться на однородных сегментах, и достигать своего верхнего предела на шумовых участках. В связи с этим остаётся открытым вопрос об использовании данного подхода в системах кодирования речи, особенности функционирования которых существенно затрудняют переход к кодированию с переменной скоростью. К таким системам относятся кодеки, использующие разновидности алгоритма CELP. В таких системах сигнал возбуждения декодера формируется из т.н. «кодовых книг». Такие книги содержат различные варианты сигналов возбуждения и для формирования последнего необходимо лишь знать его расположение. Количество отсчётов, составляющих сигнал возбуждения, в таких системах фиксировано, а структура оптимально подобрана для синтеза сегмента речи на приемной стороне [13].

Ввиду того, что расширение сегмента анализа происходит на вокализованных (квазипериодических) участках, то сигнал возбуждения так же имеет квазипериодическую структуру. Следовательно, решением проблемы фиксированной длины участка возбуждения является расширение последнего на число периодов основного тона равное числу периодов, на которое был расширен сегмент анализа на приёмной стороне.

В настоящее время системы передачи речи строятся с использованием сигнальных

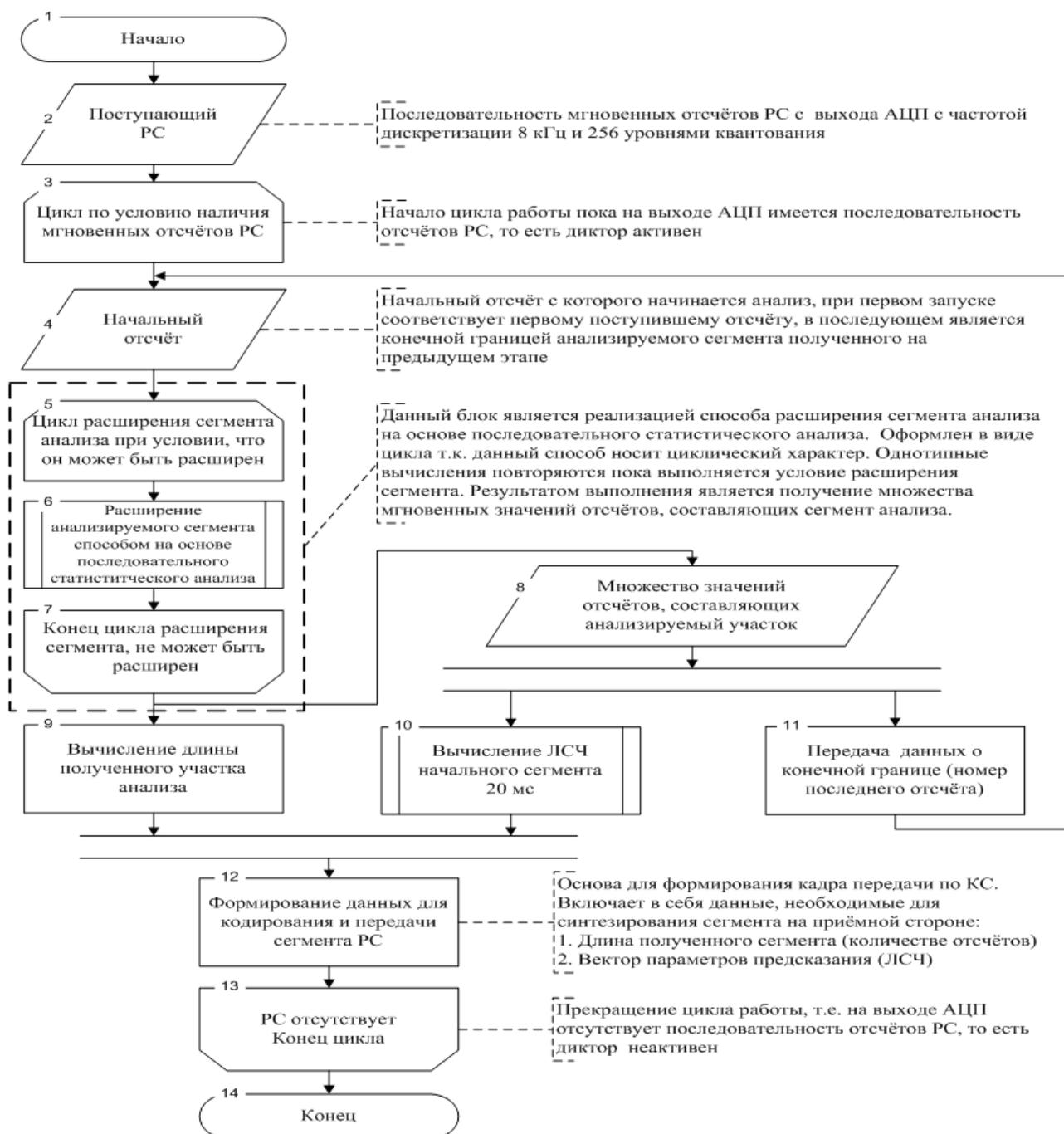


Рис. 5 Алгоритм формирования данных для кодирования и передачи сегмента РС в речепреобразующих устройствах с переменной скоростью передачи

процессоров. Следовательно, для применения вышеописанного подхода, необходимо выполнить следующие действия:

1. Видоизменить кадр передачи, добавив информацию о числе периодов основного тона, на которые был расширен сегмент анализа.
2. Включить в декодер процедуру расширения сигнала возбуждения, перед подачей его на вход синтезирующего фильтра.

В этом случае структурная схема системы кодирования речи при переменной длительности сегмента анализа выглядит следующим образом (рис. 6).

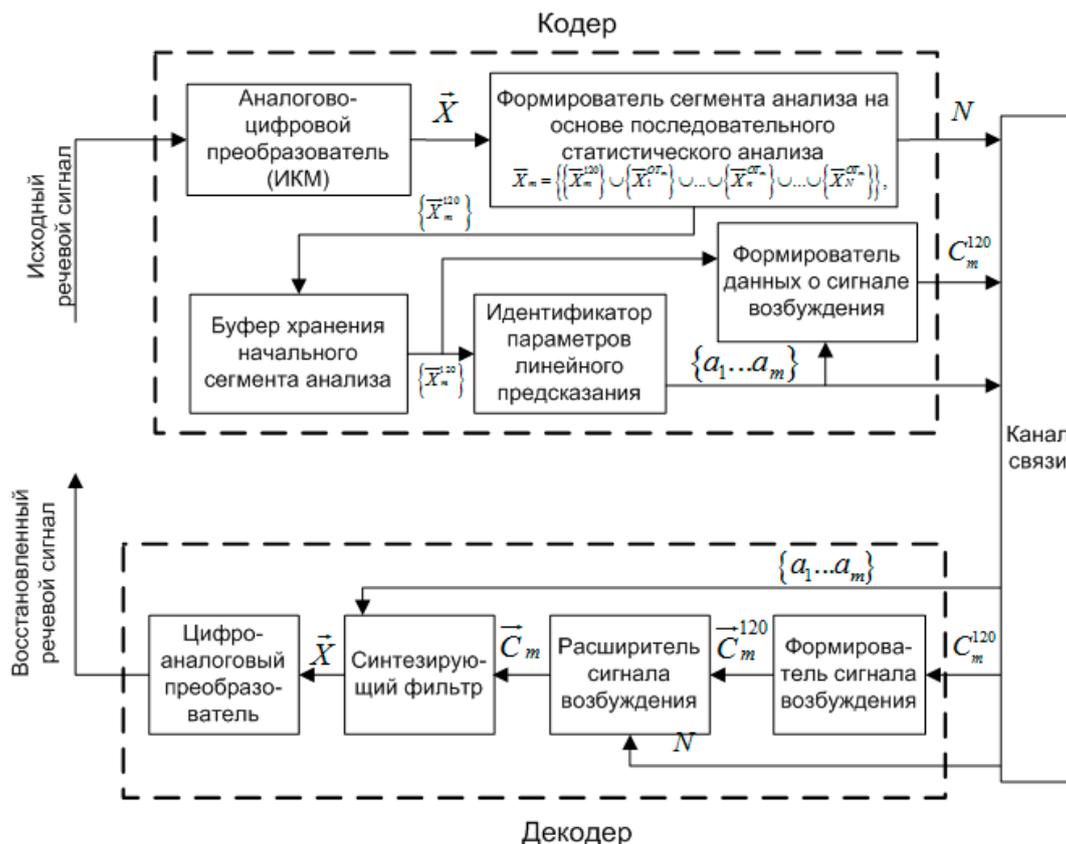


Рис. 6. Система кодирования речи при переменной длительности сегмента анализа

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ И ИХ ОБСУЖДЕНИЕ

С целью подтверждения полученных аналитическим путём заключений о возможности использования параметров предсказания, выделенных на начальных участках, для восстановления расширенного сегмента было проведено имитационное моделирование в среде Matlab R2012a.

Для анализа были представлены 4 набора отсчётов РС длительностью от 18000 до 23000 отсчётов, полученные на выходе аналого-цифрового преобразователя частотой дискретизации 8 кГц. Данные наборы подвергались обработке по двум сценариям:

1. Сегментирование по 20 мс, с выделением параметров предсказания и восстановление с возбуждением белым гауссовым шумом (БГШ).

2. Выделение сегментов с помощью последовательного статистического анализа и восстановление сигнала по параметрам предсказания, рассчитанным на начальных (20 мс) сегментах, с возбуждением от БГШ.

Во время имитационного моделирования проводилось сравнение объемов данных необходимых для восстановления каждого набора отсчётов и параметры качества восстановленного сигнала [14, 15], рассчитанных по обоим сценариям:

При этом были применены следующие ограничения и допущения:

1. Объём информации, необходимой для передачи одного сегмента анализа является одинаковым для обоих сценариев.

2. Дополнительными данными, необходимыми для выполнения 2-го сценария, является только информация о количестве отсчётов, составляющих анализируемый сегмент.

Для получения количественных оценок, объем данных, составляющих кадр передачи, был принят в соответствии с алгоритмом низкоскоростного кодирования на основе ЛП с долговременным многоимпульсным возбуждением (LPC-LTP-MPE), что составляет 192 бита на сегмент длиной 20 мс.

С учётом предельного значения длины анализируемого сегмента, накладываемого

требованиями Международного союза электросвязи (МСЭ) объем данных для передачи информации о сигнале возбуждения (L) не превышает значения 640. При проведении эксперимента использовалось максимально-возможное значение L . Для передачи данного значения по цифровым каналам связи необходимо 10 бит.

В результате произведённого моделирования были получены следующие оценки отношения объема информации, необходимой к передаче для синтезирования набора отсчетов (бит), к длительности сигналов, получаемого при обратном цифроаналоговом преобразовании (c). Размерность полученных результатов (бит/с) позволяют считать их приближёнными оценками скоростей кодирования для соответствующих сценариев моделирования. Результаты представлены в табл. 3.

Анализ полученных результатов позволяет сделать вывод, что переход к переменным скоростям кодирования речи, с использованием описанного в данной статье алгоритма приводит к снижению средней скорости кодирования не менее чем на 12 %, по отношению к существующим алгоритмам кодирования с фиксированными сегментами анализа РС.

ЗАКЛЮЧЕНИЕ

В результате проведенных исследований был выработан подход, позволяющий снизить среднюю скорость кодирования РС на основе ЛП, при сохранении качественных показателей синтезированной речи. Достижение поставленной цели осуществляется за счёт наиболее полного использования ста-

тистических взаимосвязей, присутствующих в РС. При этом использование процедуры адаптивного формирования границы анализируемого сегмента приводит к возникновению переменной скорости кодирования речи.

В дальнейшем следует уделить внимание разработке способов представления и формирования вектора сигнала возбуждения, т. к. снижение данной составляющей (параметра L) в кадре передачи позволит достигнуть более эффективного кодирования РС.

Передача речи с переменной скоростью может быть реализована в системах и сетях связи, использующих механизмы коммутации пакетов, протоколы с переменными скоростями, высокоскоростные мультиплексоры потоков с переменной скоростью.

СПИСОК ЛИТЕРАТУРЫ

1. *Быков, С. Ф.* Цифровая телефония: учебное пособие для вузов / С. Ф. Быков, В. И. Журавлев, И. А. Шалимов. – Москва : Радио и связь, 2003. – 144 с.
2. *Иванов, В. Н.* Вычисление линейных спектральных частот / В. Н. Иванов // Электросвязь. – 1997. - № 6. – С. 25–27.
3. *Ланнэ, А. А.* Передача информации о состоянии фильтра-предсказателя с помощью спектральных пар / А. А. Ланнэ, Д. А. Улахович // Радиоэлектроника и связь. 1991. – № 1. – С. 37 – 43.
4. *Ланнэ, А. А.* Новая теория линейных спектральных корней // Цифровая обработка сигналов и ее применение: труды 3-й Международной конференции: труды / А. А. Ланнэ // Москва, 2000. – С. 118 – 125.

Таблица 3

Результаты моделирования

Наборы отсчетов	Скорость кодирования по сценарию 1, бит/с	Средняя скорость кодирования по сценарию 2, бит/с	Снижение качества восстановленного речевого сигнала
Набор 1	9600	8446	нет
Набор 2	9600	8322	нет
Набор 3	9600	8258	нет
Набор 4	9600	8446	нет

5. Рекомендации ITU-T G.729 «Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP)», Женева, 1996.

6. Шалимов, И. А. Теоретико-информационные принципы компрессии речевого сигнала на основе его квазипериодических свойств: дис. д-ра техн. наук – Москва, 2005. – 241 с. РГБ ОД, 71:06-5/528.

7. Вальд Абрахам. Последовательный анализ: учебное пособие / под редакцией Б.А. Севостьянова. – Москва: Государственное издательство физико-математической литературы, 1960. – 328с.

8. Михайлов В.Г., Златоустова Л.В. Измерение параметров речи / под редакцией М. А. Сапожкова. – Москва.: Радио и связь, 1987. – 168 с.

9. Пат. 2445718 Российская Федерация, МПК G 10 L 19/00 (2006.01). Способ выделения сегментов речи на основе анализа корреляционных зависимостей в речевом сигнале / Афанасьев А. А., Новиков Е. И., Трубицын В. Г., Титов О. Н.; заявитель и патентообладатель Академия Федеральной службы охраны Российской Федерации – № 2010136618/08; заявл. 31.08.2010; опубл. 20.03.2012, Бюл. № 8. – 11 с.

10. Пат. 2684576 Российская Федерация, МПК G 10 L 19/022 (2013.01) / СПК G 10 L 19/022 (2018.08). Способ выделения сегментов обработки речи на основе последователь-

ного статистического анализа / Власов Р. С., Кисляк А. А., Титов О. Н., Любимов Д. В.; заявитель и патентообладатель Академия Федеральной службы охраны Российской Федерации – № 2018103835; заявл. 31.01.2018; опубл. 09.04.2019, Бюл. № 10. – 20 с.

11. Семёнов, В. Ю. Новый метод вычисления линейных спектральных частот речевых сигналов, основанный на универсальном алгоритме решения трансцендентных уравнений / В. Ю. Семёнов // ISSN 1028-7507 Акустический вестник – 2004. – Том 7, № 3. – С. 55–64.

12. Семёнов, В. Ю. Новый подход к вычислению линейных спектральных частот речевых сигналов, основанный на свойстве межфреймовой упорядоченности / В. Ю. Семёнов // ISSN 1028-7507 Акустический вестник – 2004. – Том 7, № 3. – С. 55–64.

13. Афанасьев, А. А. Синтез сигналов возбуждения в низкоскоростных вокодерах с линейным предсказанием / А. А. Афанасьев // Телекоммуникации. – 2010. – № 11. – С. 8–13

14. ГОСТ Р 50840-95. Передача речи по трактам связи. Методы оценки качества, разборчивости и узнаваемости. – Москва: ИПК издательство стандартов, 1996. – 230 с.

15. ГОСТ Р 51061-97. Системы низкоскоростной передачи речи по цифровым каналам. Параметры качества речи и методы измерений. – Москва : Госстандарт России, 1997. – 230 с.

Афанасьев Андрей Алексеевич – канд. техн. наук, доцент, Академия Федеральной службы охраны Российской Федерации, e-mail: fromnet@yandex.ru

Власов Роман Сергеевич – сотрудник, «Академия Федеральной службы охраны Российской Федерации», e-mail: vlasrsv@mail.ru

А. А. Афанасьев, Р. С. Власов

SYSTEM LOW SPEED ENCODE THE SPEECH SIGNAL AT A VARIABLE DURATION SEGMENT ANALYSIS

A. A. Afanasjev, R. S. Vlasov

Academy FGS of Russia (Orel)

Annotation. In article, materials on research of a speech signal such properties, as uniformity (similitude) of the speech signal instant values distribution are stated. This property is used in a way of the speech processing segments allocation on the basis of the consecutive statistical analysis. The parameters of a linear prediction counted on each step of expansion undergo insignificant (within a quantization step) changes in relation to an initial segment. This circumstance allows to use results of the analysis of an initial site for restoration of the site expanded in the above-stated way that allows to lower essentially both computing complexity, and speed of a speech signal coding.

Keywords: речевой сигнал, сегмент анализа, линейные спектральные частоты, шаг квантования, расширение сегмента.

Afanasyev Andrey Alekseevich – Candidate of Technical Sciences, Assistant professor, Academy FGS of Russia, city Orel, e-mail: fromnet@yandex.ru

Vlasov Roman Sergeevich – Employee, Academy FGS of Russia, city Orel, e-mail: vlasrsv@mail.ru