

РАЗРАБОТКА АЛГОРИТМА УПЛОТНЕНИЯ РЕЧЕВЫХ СИГНАЛОВ ДЛЯ ПЕРСПЕКТИВНЫХ СИСТЕМ ПЕРЕДАЧИ ПАКЕТНЫХ СООБЩЕНИЙ

1. Разработка технических требований для алгоритма уплотнения речевых сигналов

1.2. Назначение и цель выполнения научно-исследовательской работы

Алгоритм речевого уплотнения предназначен для устранения информационной избыточности в речевом сигнале и оптимизации (снижения) скорости работы кодеков перспективных систем передачи сообщений, разрабатываемых СКТБ «Рассвет» при ГУП «Прогресс».

Целью работы является разработка и моделирование алгоритма уплотнения речевых сигналов с дальнейшей выработкой рекомендаций по их применению в перспективных системах передачи пакетных сообщений.

В дальнейшем алгоритмы речевого уплотнения будем называть изделием.

1.2. Основные технические требования к изделию

Проектируемое изделие должно удовлетворять следующим требованиям:

1.2.1. Изделие должно функционировать в составе программного комплекса речевого кодека (включающего управляющий модуль, программы ввода-вывода речевого сигнала и т.д.).

1.2.2. Изделие должно выполнять компрессию (декомпрессию) речевого сигнала по схеме «анализ-синтез» и обеспечивать выполнение следующих основных функций:

- временную сегментацию отсчетов. При этом изделие должно проектироваться с учетом частоты дискретизации речевого сигнала не менее 8кГц и разрядности входного цифрового потока в 16 бит.

- процедуру кратковременного предсказания (при этом порядок p фильтра кратковременного предсказания определяется разработчиком).

- аппроксимацию остаточного сигнала и кодирования.

- декодер должен выполнять функцию восстановления речевого сигнала.

1.2.3. Выбор алгоритмов предсказания, аппроксимации и кодирования сигналов определяется разработчиком изделия с учетом выполнения ниже перечисленных требований:

1.2.3.1. Допустимая временная задержка в процессе сжатия речевого сигнала не должна превышать 100 мсек.

1.2.3.1. Изделие должно обеспечивать максимальную скорость выдачи сообщений в канал связи, не превышающую 12 Кбит/с.

1.2.3.2. Оценочная сложность алгоритма должна определяться информационной мощностью кодека (из учета его быстродействия в 16 млн. кор. оп./сек, выделенного объема ОЗУ - 640 Кб, ПЗУ - не менее 100 Кб).

1.2.3.3. Разборчивость речи при декодировании должна быть не хуже характеристик, предъявляемых к каналу ТЧ.

1.2.4. Требования к программному обеспечению. Моделирование изделия должно осуществляться при помощи программной среды MATLAB с последующей обработкой на эмуляторах сигнального процессора, предоставляемых заказчиком.

1.2.5. Требования к метрологическим характеристикам ввода речевого сигнала не предъявляются.

1.2.6. Требования к надежности, безопасности изделия не предъявляются.

1.2.7. Основные процедуры (кратковременного и долговременного предсказания) изделия должны быть унифицированы со стандартами уплотнения RPE-LTP (GSM), CELP (FS1016), LPC-10e (FS1015).

1.2.6. Требования к защите информации от несанкционированного доступа и помехоустойчивости не предъявляются.

1.3. Порядок приемки, контроля и испытания изделия

1.3.1. Результаты разработки изделия оформляются в виде технического отчета, включающего:

1.3.1.1. Теоретическое описание разработанных процедур сжатия речевого сигнала.

1.3.1.2. Описание алгоритмов разработанного программного обеспечения изделия.

1.3.1.3. Модельные имитационные эксперименты воспроизведения речевых сигналов (для нескольких дикторов).

1.3.1.4. Исполняемые программные модули изделия.

1.3.2. Виды, состав, объем и методы испытаний изделия также согласуются с заказчиком в процессе выполнения научно-технической разработки.

2. Разработка алгоритма речевого кодирования CTCVoCoder1.V1 (Chair of Telecommunications Voice Coder)

2.1. Общая схема алгоритма компрессии и декомпрессии речевого сигнала

Структурная схема алгоритма компрессии речевого сигнала (рис. 1) включает:

- 1) процедуру сегментации и кратковременного кодирования речи (ПККР);
- 2) процедуру долговременного кодирования речи (ПДКР);
- 3) процедуру децимации (прореживание) второго остаточного сигнала (Д);
- 3) аппроксимацию второго остаточного сигнала (АВОС).

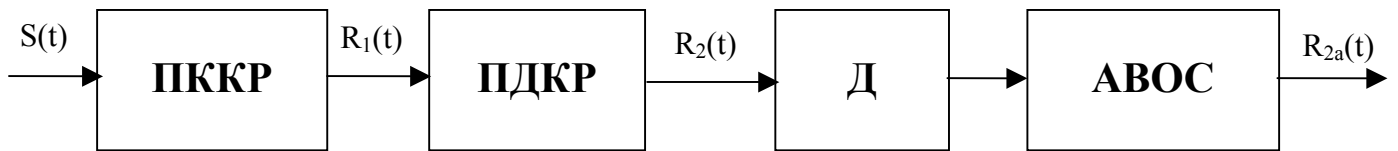


Рис.1

На схеме (рис.1) обозначены $S(t)$ – входной (дискретный) речевой сигнал, $R_1(t)$ – первый остаточный сигнал, $R_2(t)$ – второй остаточный сигнал, $R_{2a}(t)$ – аппроксимированный второй остаточный сигнал. Вместо речевого сигнала $S(t)$ в канал связи передаются параметры фильтров ПККР, ПДКР и значения сигнала $R_{2a}(t)$. Сжатие речевого сигнала происходит за счет устранения его информационной избыточности (которая отражается в параметрах ПККР, ПДКР), децимации сигнала и аппроксимации второго остаточного сигнала (последняя процедура определяет минимальное число бит на кодирование второго остаточного сигнала).

Декомпрессия речевого сигнала аналогична схеме (рис. 1), но в этом случае сигнал подвергается процедурам восстановления, обратным вышеперечисленным.

2.2. Алгоритм кратковременного кодирования речи (LPC – кодер).

При кратковременном кодировании речи удаляется избыточность сигнала, обусловленная квазипериодическими колебаниями на частотах формант. Кратковременное предсказание можно рассматривать в качестве дифференциально-кодовой модуляции. Идея метода состоит в линейной аппроксимации текущего отсчета речевого сигнала его предшествующими значениями с последующим вычислением первого остаточного сигнала. Другими словами, аппроксимированное значение речевого сигнала $\hat{S}(n)$ вычисляется по формуле:

$$\hat{S}(n) = \sum_{k=0}^{p-1} a_k S(n-k), \quad (1)$$

где p – порядок предсказания фильтра, a_k – неизвестные коэффициенты экстраполятора, $S(n-k)$ – задержанные на k – тактов значения речевого сигнала. Первый остаточный сигнал определяется по формуле

$$R_1(n) = S(n) - \hat{S}(n) = S(n) - \sum_{k=0}^{p-1} a_k S(n-k). \quad (2)$$

Из (2) видно, что чем лучше величины $\hat{S}(n)$ и $S(n)$ совпадают, тем меньше по амплитуде второй остаточный сигнал. Следовательно, потребуется меньше бит на его дальнейшее кодирование, что может понизить, в дальнейшем, скорость кодера.

Для восстановления речевого сигнала по первому остаточному сигналу на приемной стороне (в декодере) используется формула

$$S_{восм}(n) = \sum_{k=0}^{p-1} a_k S_{восм}(n-k) + R_1(n), \quad (3)$$

получаемую из (2). В (3) $S_{восст}(n)$ - оценка речевого сигнала на выходе декодера после восстановления.

Как уже было сказано важную роль в сжатии сигнала в LPC кодере играет правильное определение коэффициентов аппроксимации a_k . Известные коэффициенты находят их минимизации квадратичной формы J вида

$$J = \min \left\{ \sum_{n=0}^{L-1} [S(n) - \sum_{k=0}^{p-1} a_k S(n-k)]^2 \right\}, \quad (4)$$

где L – число отсчетов в одном сегменте (L=160).

Не приводя методику минимизации (4), запишем окончательное решение задачи в матричной форме :

$$r_{p \times p} \bar{a}_{1 \times p} = r_{1 \times p}, \quad (5)$$

где вектор $\bar{a}_{1 \times p} = [a_0, \dots, a_p]^T$ - вектор неизвестных коэффициентов, матрица

$$r_{p \times p} = \begin{pmatrix} r(0) & \dots & r(p-1) \\ \vdots & \ddots & \vdots \\ r(p-1) & \dots & r(0) \end{pmatrix} \text{ и вектор } r_{1 \times p} = [r(1), \dots, r(p)]^T \text{ состоят из значений } r(j)$$

автокорреляционной функции речевого сигнала, вычисляемых по формуле

$$r(j) = \sum_{n=1}^{L-1} S(n)S(n-j). \quad (6)$$

При больших p решение (5) слабо обусловлено и в результате может привести к большим погрешностям вычисления коэффициентов LPC-кодера. Поэтому в силу симметричности матрицы $r_{p \times p}$ используют алгоритм Левинсона-Дурбина, который вычисляет не сами коэффициенты a_k , а их аналоги – коэффициенты отражения P_k . Алгоритм Левинсона-Дурбина выполнен по схеме рекурсивного приближения коэффициентов отражения P_k . Общая схема алгоритма Левинсона-Дурбина содержит p этапов вычислений. В качестве нулевого приближения полагают все коэффициенты $a_k = 0$ ($k=0 \dots p-1$), а среднеквадратичную ошибку $E=1$. Далее выполняют p последовательных вычислений, от $i=0$ до $i=p-1$, определяя на каждом шаге коэффициент отражения по формуле [1]:

$$P_i = [r(i) - \sum_{j=1}^{i-1} a_j^{(i-1)} r(i-j)] / E(i-1),$$

$$a_j^i = K_i,$$

где для $j=1$ до $i-1$ определяют

$$a_j^i = a_j^{i-1} - P_i a_{i-j}^{i-1},$$

$$E(i) = (1 - P_i^2) E(i-1).$$

За конечный результат принимают $a_j = a_j^p, j=1 \dots p$. Блок схема вычисления коэффициентов отражения приведена на рис 2.

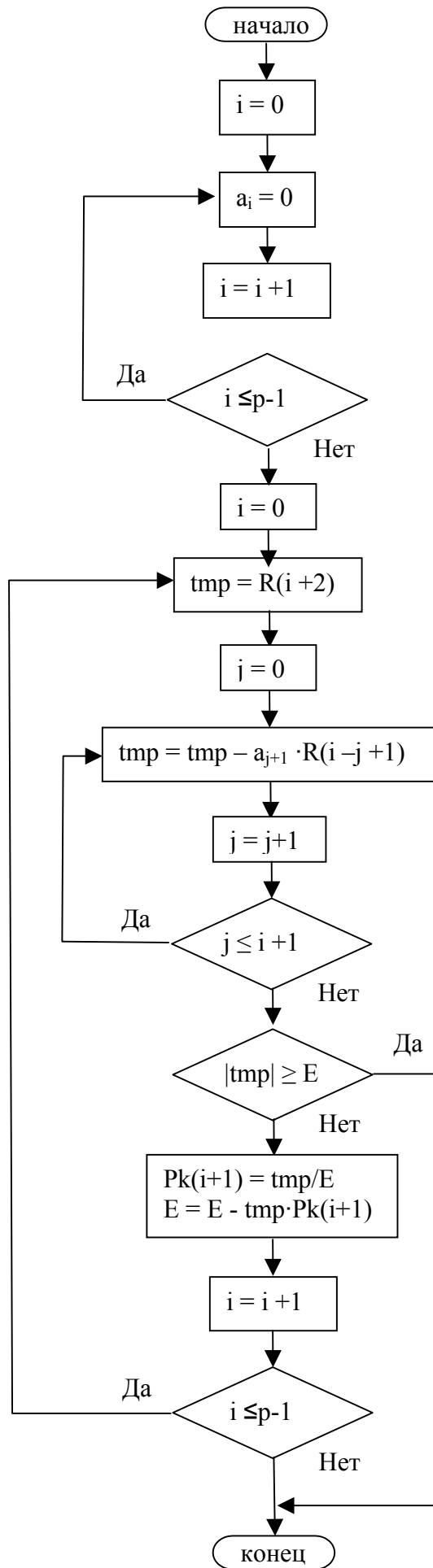


Рис. 2

На рис. 3 приведена блок-схема восстановления коэффициентов аппроксимации a_k по коэффициентам отражения P_k

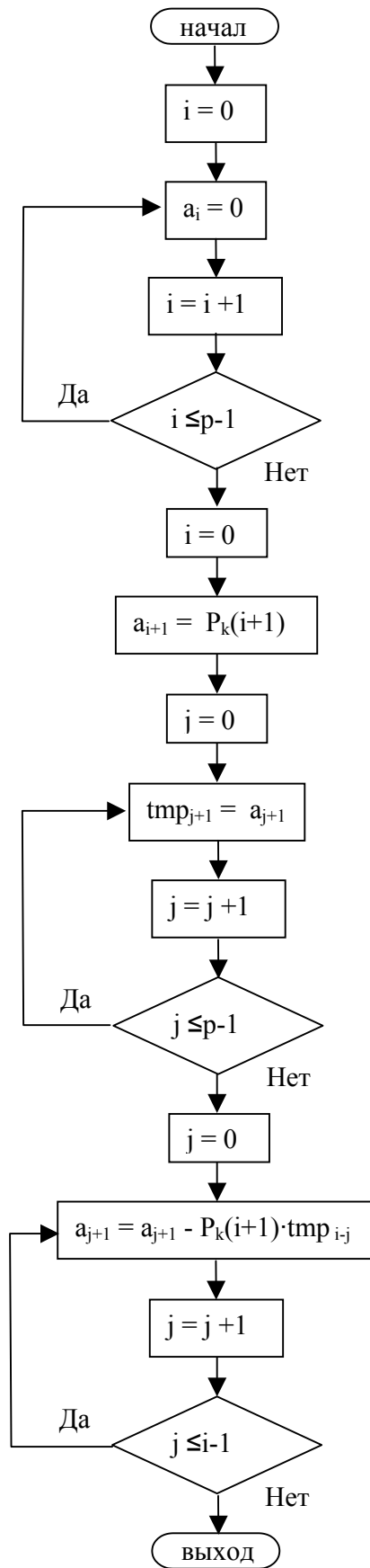


Рис.3

В блок-схемах на рис. 2,3 через tmp обозначена некоторая переменная хранения «промежуточной» информации.

Программная реализация на MATLAB6 процедуры кратковременного предсказания приведена в приложении: модули **myfilter2(3)lar.m**, **mylevinson.m**. Особенностью программной реализации заключается в том, что для каждого сегмента из 160 речевых отсчетов вычисляются 8 коэффициентов отражения и переводятся в LAR (логарифмическое представление) в виде

$$LAR_i = \log \frac{1 - P_i}{1 + P_i}. \quad (7)$$

Это позволит уменьшить количество бит на передачу коэффициентов отражения. В таблице 1 приведена сетка кодирования параметров кратковременного предсказания (LAR)

Таблица 1

Номер LAR	Число бит на LAR
1, 2	6
3, 4	5
5, 6	4
7, 8	3

Для выполнения кодирования предварительно находятся максимумы для каждой из четырех групп LAR (в демонстрационной версии программы поиск максимумов ведется по всем вычисленным LARам, а не только для последнего сегмента).

2.3. Алгоритм долговременного кодирование речи

В первом остаточном сигнале $R_1(t)$ также можно устранить избыточность, выявляя и удаляя колебания с частотой основного тона. Для этого производят аппроксимацию первого остаточного сигнала и определяют второй остаточный сигнал по формуле

$$R_2(n) = R_1(n) - GR_1(n - \alpha), \quad (8)$$

где G, α - неизвестные параметры, характеризующие амплитуду и частоту избыточного квазипериодического сигнала. При этом α - целое число, связанное с периодом основного тона, $\alpha \in [20, \dots, 160]$. Для повышения точности аппроксимации коэффициент G находят из минимизации квадратичной формы J вида

$$J = \min \left\{ \sum_{n=0}^{L^*-1} [R_1(n) - GR_1(n - \alpha)]^2 \right\}, \quad (9)$$

где $L^* = 40$ - число отсчетов в сегменте. Решением (9) является выражение

$$G = \frac{\sum_{n=0}^{L^*-1} R_1(n)R_1(n - \alpha)}{\sum_{n=0}^{L^*-1} [R_1(n - \alpha)]^2}. \quad (10)$$

Далее значение G необходимо подставить в (9), при этом функционал J будет зависеть от неизвестной α . Чтобы преодолеть эту трудность производят

дальнейшую минимизацию (9) путем перебора всех значений α из вышеуказанного диапазона, следя при этом за устойчивостью фильтра долговременного предсказания, т.е. на каждом шаге проверяется условие $|G| \leq 1$. Если же $|G| > 1$, то принудительно устанавливают $|G|=1$ и снова ищут α . В результате на выходе получают минимизированный второй остаточный сигнал.

Особенностью программной реализации является то, что для получения необходимого качества остаточного сигнала сегмент разбивают на несколько подсегментов и работают с каждым отдельно.

Модуль долговременного предсказания `myfilter2(3)lag.m` приведен в Приложении. При этом для каждого сегмента находятся по 4 коэффициента фильтра долговременного предсказания G (принимают значения от -1 до $+1$) и по 4 порядка фильтра долговременного предсказания `MaxAlpha` (принимает значения от 20 до 147). Каждый коэффициент G кодируется в 4 бита и показатель степени $\alpha = \text{MaxAlpha}$ кодируется с помощью 7 бит.

2.4. Децимация и аппроксимация второго остаточного сигнала

С целью дальнейшего увеличения степени сжатия производят прореживание второго остаточного сигнала, т.к. он «почти» является шумовым порождающим сигналом. Но перед децимацией с помощью фильтра нижних частот подавляется верхняя составляющая спектра второго остаточного сигнала. Фильтрация второго остаточного сигнала производится с помощью КИХ ЦФНЧ с частотой среза $f_{\text{среза}} = 0,24f_d = 1920$ Гц:

$$H_1(z) = -0,0016 - 0,0223z^{-1} + 0,0107z^{-2} + 0,274z^{-3} + 0,4784z^{-4} + 0,274z^{-5} + 0,0107z^{-6} - 0,0223z^{-7} - 0,0016z^{-8}. \quad (11)$$

Для обеспечения нулевой групповой задержки фильтрация производится в прямом и обратном направлении (функция MATLAB `filtfilt`), в результате чего эффективный порядок фильтра увеличивается в 2 раза. Результирующая АЧХ приведена на рис.4.

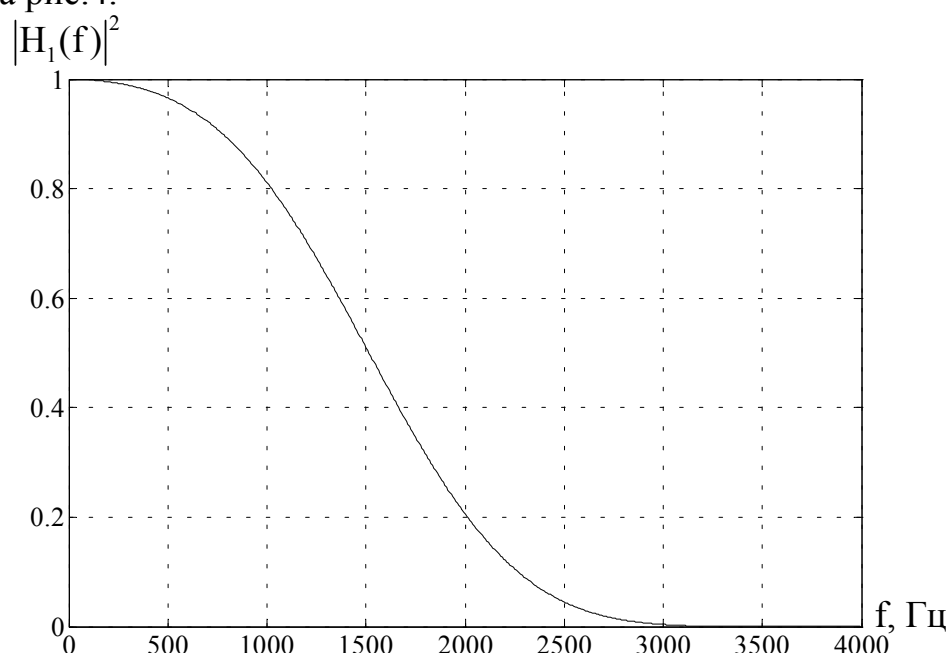


Рис 4. Результирующая АЧХ фильтрации перед прореживанием

Далее производится прореживание, для чего берутся только нечетные отсчеты (1, 3, ...) массива.

Прореженный второй остаточный сигнал кодируется с помощью двухбитной (два бита на один отсчет) IMA ADPCM (imaadpcm_en.m).

В алгоритме IMA ADPCM для хранения каждого отсчёта используется 4 бита. Кодирующий модуль берёт разницу между двумя отсчётами, делит её на текущий *размер шага* и использует получившуюся величину в качестве очередного сжатого 4-битного выходного значения. Декомпрессор, наоборот, берёт это 4-битное значение, умножает его на текущую величину шага и прибавляет получившийся результат к предыдущему отсчёту, получая тем самым очередной отсчёт. Изменяя размер шага, с помощью 4-битного кода можно записать большой диапазон значений.

Для сокращения объёма информации сам размер шага не хранится. Вместо него компрессор сохраняет индекс в таблицу, содержащую 88 возможных значений размера шага. Величины размеров шага следуют приблизительно экспоненциальной кривой.

```
stepSizeTable = [
7,8,9,10,11,12,13,14,16,17,19,21,23,25,28,31,34,37,41,45,50,55,60,66,73,80,88,97,1
07,118,130,143,157,
173,190,209,230,253,279,307,337,371,408,449,494,544,598,658,724,796,876,963,10
60,1166,1282,1411,
1552,1707,1878,2066,2272,2499,2749,3024,3327,3660,4026,4428,4871,5358,5894,6
484,7132,7845,8630,
9493,10442,11487,12635,13899,15289,16818,18500,20350,22385,24623,27086,297
94,32767];
```

Для хранения в таблице индекса размера шага требуется всего 7 бит. 16 бит занимает текущее несжатое значение. Для экономии места современные компрессоры отбрасывают младшие биты текущего несжатого отсчёта, что позволяет размещать информацию о состоянии процесса всего в 16 битах.

Важным элементом IMA ADPCM является механизм управления размером шага. В 4-х битный отсчёт записывается представленное в прямом коде со знаком число в диапазоне от -7 до +7. Поскольку декомпрессору неизвестна величина следующего отсчёта, он должен сделать выбор, основываясь на предыдущих значениях. Если полученное в результате преобразования приращения 4-битное число велико, возможно, что следующее будет ещё больше и для его размещения не хватит четырёх бит. Если 4-битное представление приращения – величина небольшая, значит, вы не в полной мере используете имеющиеся у вас 4 бита информации. В результате, если величина, получающаяся при делении приращения на размер шага, приближается к нулю – размер шага уменьшается, если она возрастает – шаг увеличивается.

В данном проекте входной сигнал IMA ADPCM кодера представляет собой небольшие значения, поэтому младшие биты выходного сигнала равны нулю. Для экономии места, в канал связи они не передаются, а автоматически добавляются на входе декодера.

В результате кодирования создается битовый поток, скорость которого определяется следующим образом:

1. Коэффициенты кратковременного предсказания $(2*6+2*5+2*4+2*3)*50 = 1800$ бит/с. (50 – количество сегментов по 160 отсчетов в одной секунде) – данный расчет не учитывает расходы на передачу четырех максимальных значений коэффициентов. При кодировании звукового файла их достаточно передать 1 раз (выходной файл при этом увеличивается на не более чем $4*32=128$ бит).

2. Коэффициенты долговременного предсказания $(4*4+4*7)*50 = 2200$ бит/с.

3. Второй остаточный сигнал $2*80*50=8000$ бит/с.

Всего скорость работы программы составляет : $1800+2200+8000=12000$ бит/с, что укладывается в основные технические требования.

2.5. Разработка алгоритма декодирования .

2.5.1. Восстановление второго остаточного сигнала.

2.5.1.1. Восстановление прореженных отсчетов (imaadpcm_de.m).

2.5.1.2. Между каждым отсчетом вставляется по одному нулевому отсчету. В результате спектральная плотность мощности (СПМ) оценки второго остаточного сигнала на участке спектра от 2000 до 4000 Гц будет являться зеркальным отражением СПМ на участке спектра от 0 до 2000 Гц.

2.5.1.3. Т.к. использование неидеального ФНЧ перед прореживанием и применение выбранного варианта ADPCM приводит к некоторому подавлению верхних частот (прореженного второго остаточного сигнала) и с учетом симметричности спектра относительно 2000 Гц, в восстановленном втором остаточном сигнале наблюдается следующая особенность: при движении по спектру от 0 к 2000 Гц и от 4000 к 2000 Гц спектральные компоненты подавлены тем сильнее, чем ближе их частота к 2000 Гц. Для компенсации таких искажений применяется цифровой КИХ-фильтр со следующей передаточной функцией:

$$H_2(z) = -0,0004z^{-1} - 0,3837z^{-2} + 0,0014z^{-3} + 2,2534z^{-4} + \\ + 0,0014z^{-5} - 0,3837z^{-6} - 0,0004z^{-7}$$

Для обеспечения нулевой групповой задержки фильтрация производится в прямом и обратном направлении (функция MATLAB filtfilt), в результате чего эффективный порядок фильтра увеличивается в 2 раза. Результирующая АЧХ приведена на рис 5.

Амплитудно-частотная характеристика фильтра была получена экспериментально при использовании тестового звукового фрагмента fspeech.wav и проверена также при кодировании тестового звукового фрагмента High-ru.wav.

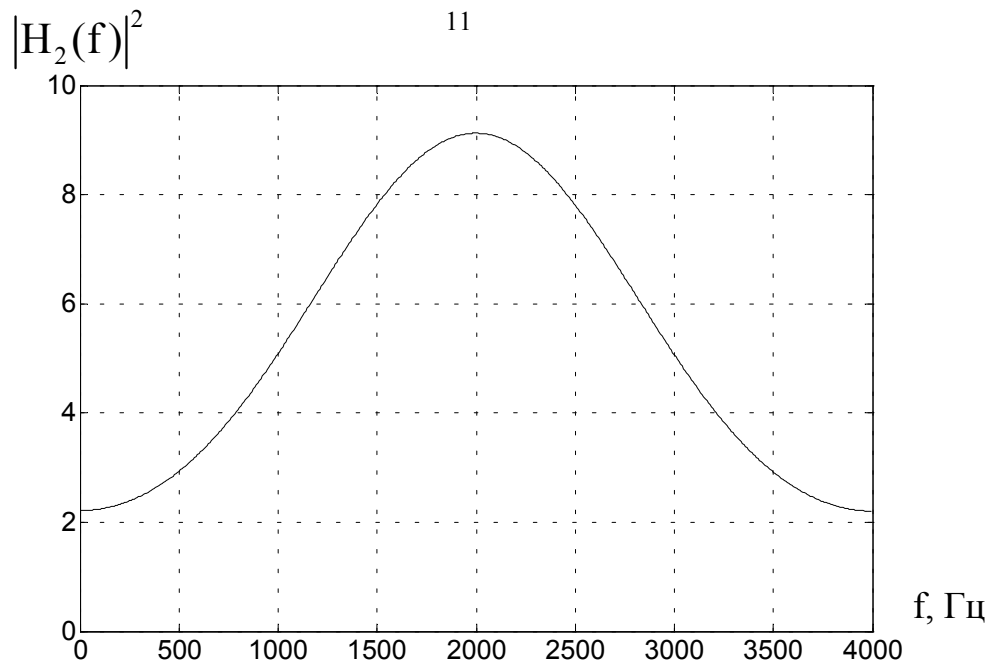


Рис 5. Результирующая АЧХ компенсирующей фильтрации после восстановления второго остаточного сигнала

2.5.2. Декодируются (переводятся в представление с плавающей запятой) коэффициенты долговременного и кратковременного предсказания и с их помощью по формуле (3) восстанавливаются первый остаточный и звуковой сигнал.

2.6.3. Для подавления нежелательных высокочастотных призвуков и шума производится фильтрация звукового сигнала с помощью КИХ ЦФНЧ с частотой среза $0,45f_d=3600$ Гц:

$$H_3(z) = -0,0061 + 0,0185z^{-1} - 0,0508z^{-2} + 0,0856z^{-3} + 0,9055z^{-4} + 0,0856z^{-5} - 0,0508z^{-6} + 0,0185z^{-7} - 0,0061z^{-8}.$$

Для обеспечения нулевой групповой задержки фильтрация производится в прямом и обратном направлении (функция MATLAB `filtfilt`), в результате чего эффективный порядок фильтра увеличивается в 2 раза. Результирующая АЧХ приведена на рис 6.

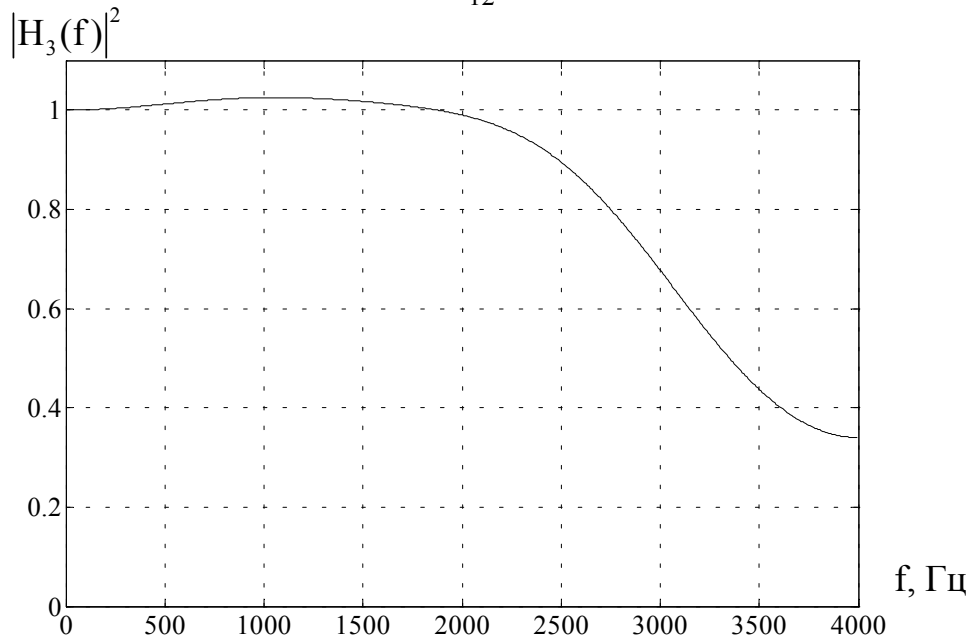


Рис 6. Результирующая АЧХ пост-фильтрации для устранения высокочастотных нежелательных призвуков и шумов

2.6. Экспериментальное исследование процедуры сжатия речевых сигналов

Экспериментальное исследование разработанной программы проводилось для звуковых файлов в формате wav (PCM, формат данных - 16 (рекомендуется) или 8 бит, $f_d=8$ кГц). При этом также была разработана версия программы, позволяющая вначале записать речь диктора с микрофона, и далее ее можно было прослушать. В процессе работы с звуковыми файлами при использовании текущей версии кодека файл должен находиться в одном каталоге с программой. Если длина исходного файла не кратна 160 отсчетам, то он дополняется нулями. Выходные данные заносятся в файл в формате wav (PCM, 16 бит, $f_d=8$ кГц).

Описание основных модулей:

- STCVoCoder1_en – кодирует исходный звуковой файл;
- STCVoCoder1_de – восстанавливает звук из закодированного файла;
- STCVoCoder1_test – тестирование кодирования и восстановления звукового файла.

Краткие пояснения по входным и выходным параметрам перечисленных программ находятся в листингах, приведенных в Приложении.

Экспериментальное исследование процедуры сжатия речевых сигналов проводилось для стандартного звукового файла FSpeech.wav. Фрагменты спектральных характеристик и осциллограмм преобразуемых сигналов приводятся на рис. 7 - 18.

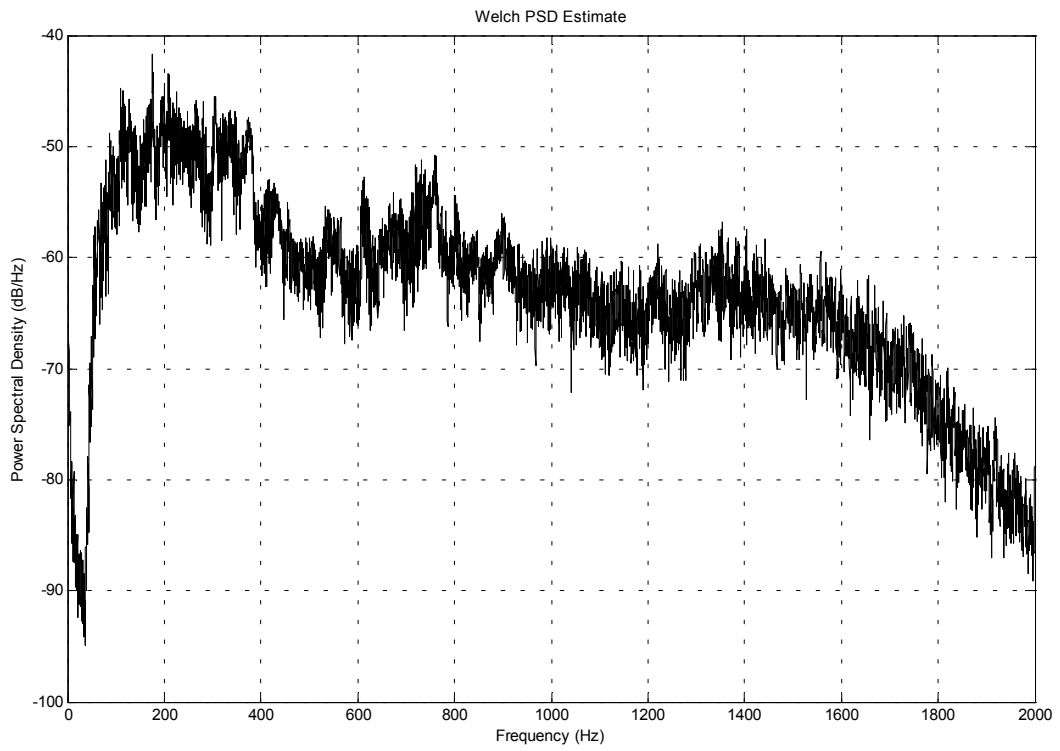


Рис. 7. Спектральная характеристика исходного звукового сигнала

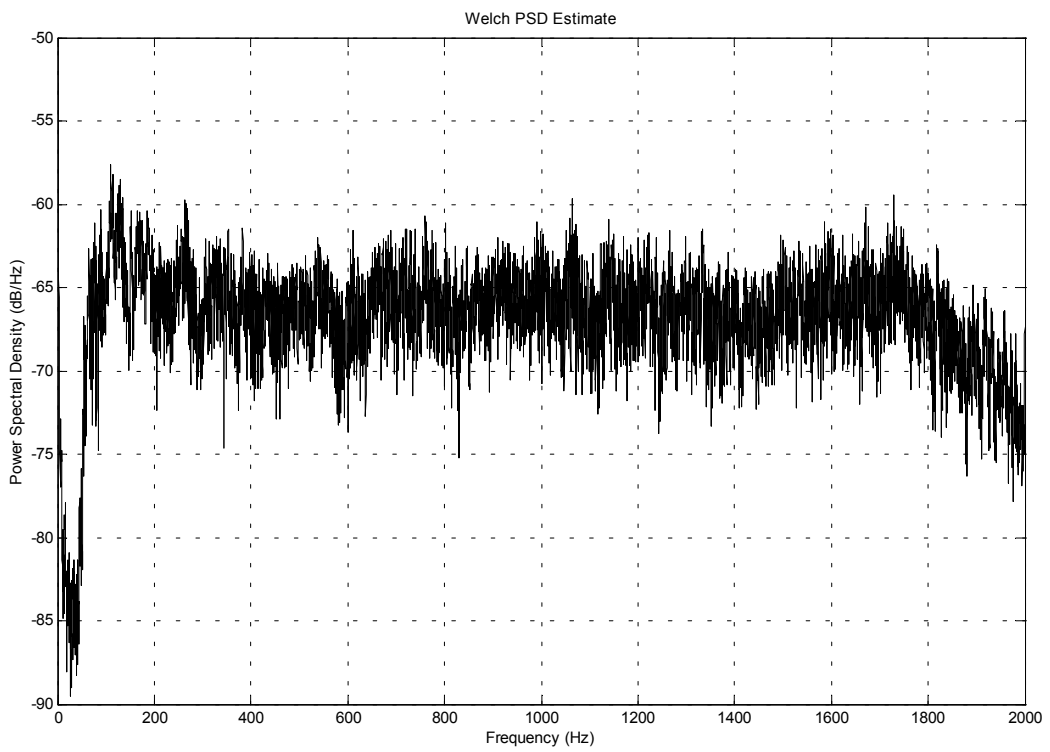


Рис. 8. Спектральная характеристика первого остаточного сигнала

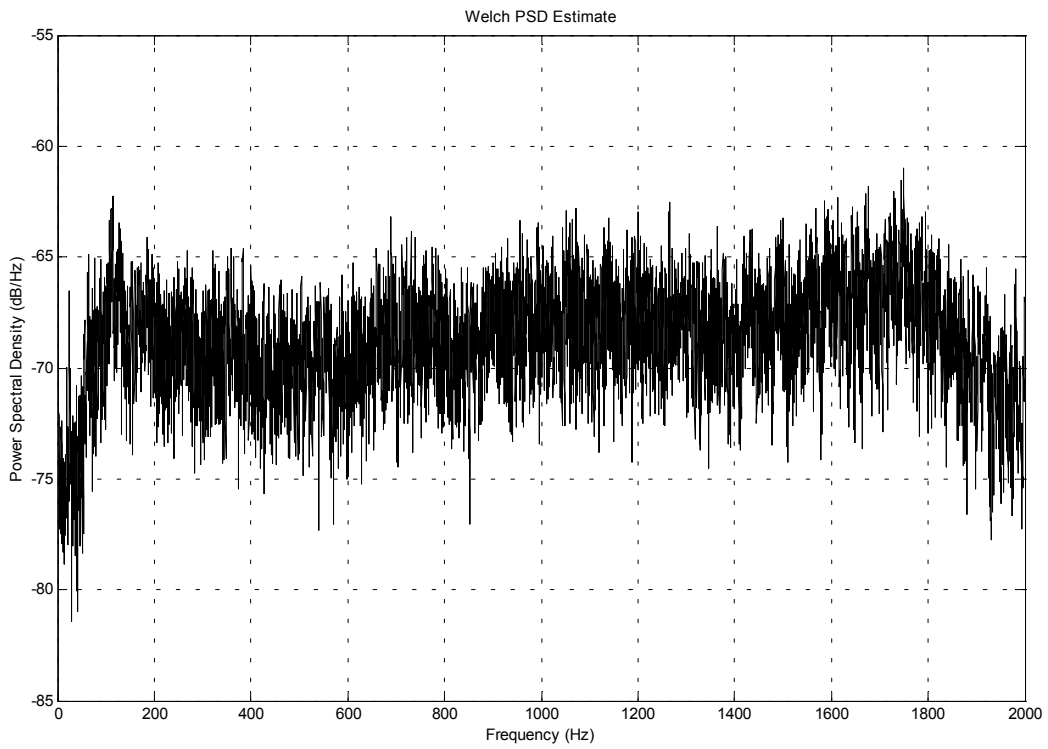


Рис. 9. Спектральная характеристика второго остаточного сигнала

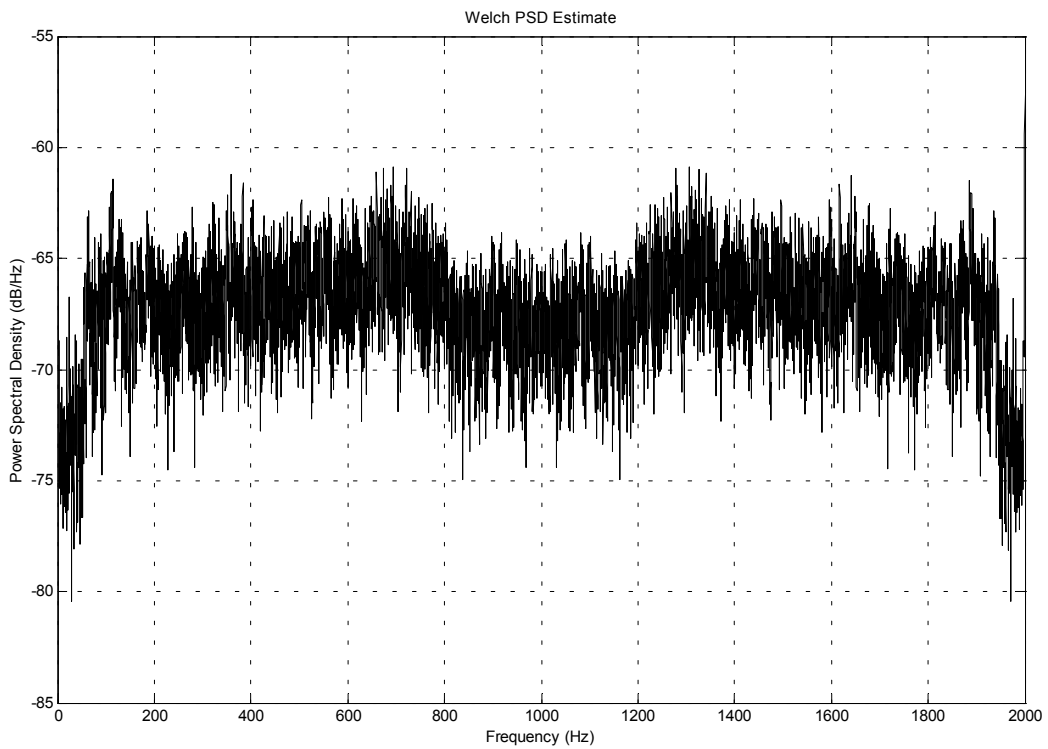


Рис. 10. Спектральная характеристика восстановленного (оценки) второго остаточного сигнала

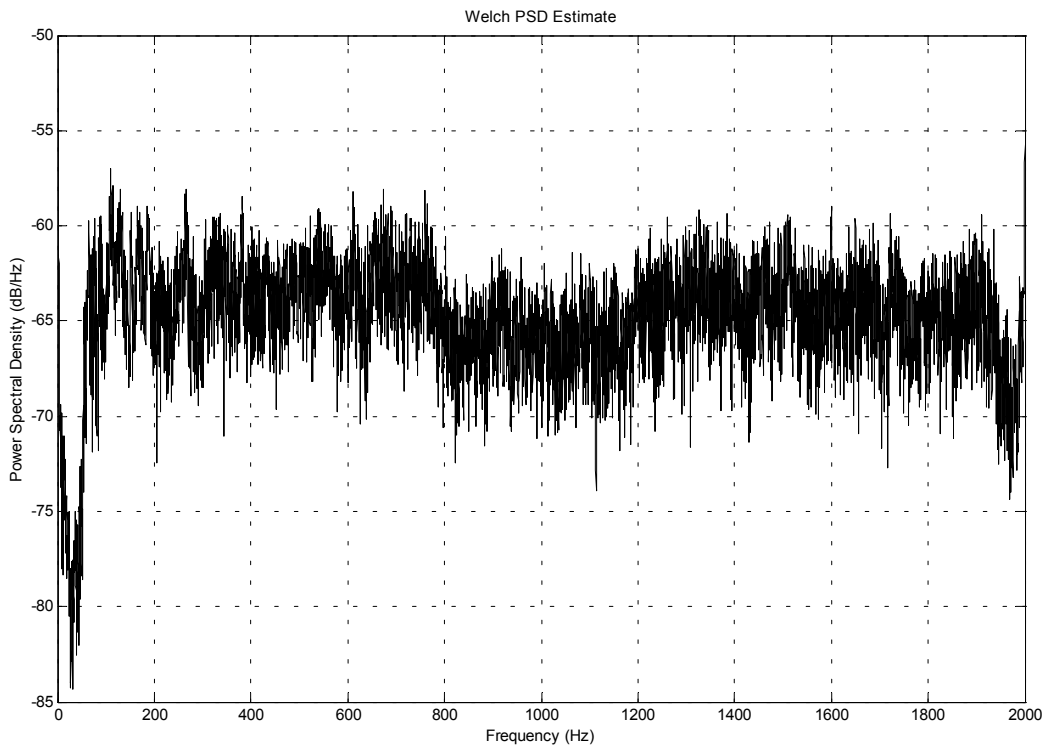


Рис. 11. Спектральная характеристика восстановленного (оценки) первого остаточного сигнала

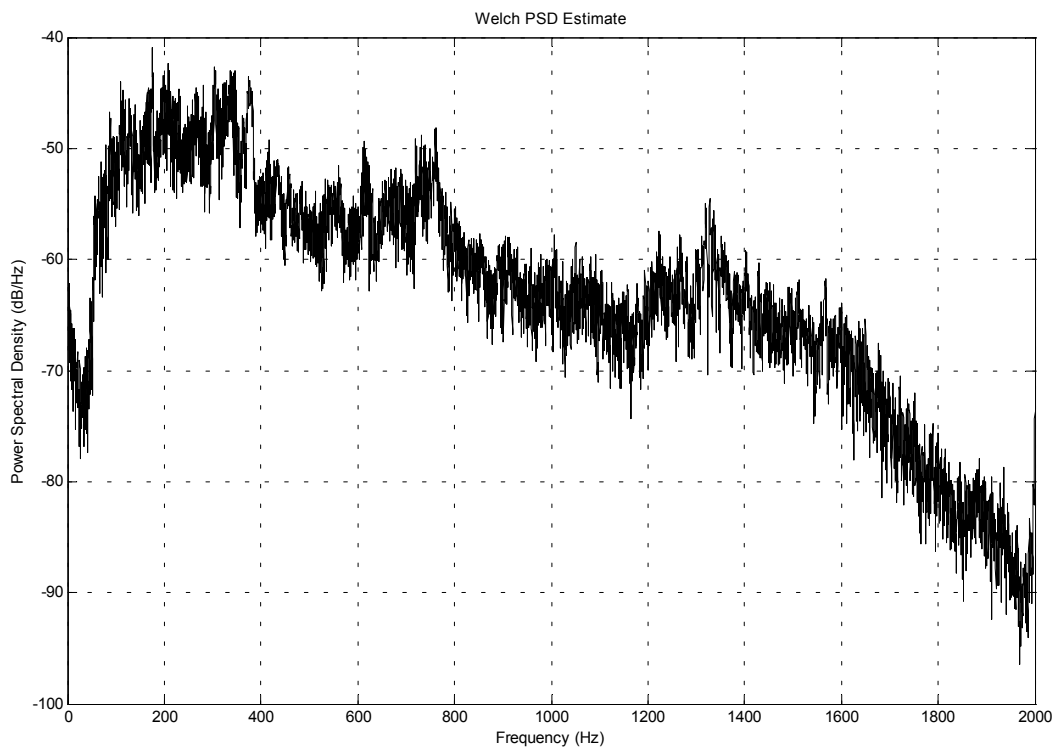


Рис. 12. Спектральная характеристика восстановленного (оценки) звукового сигнала

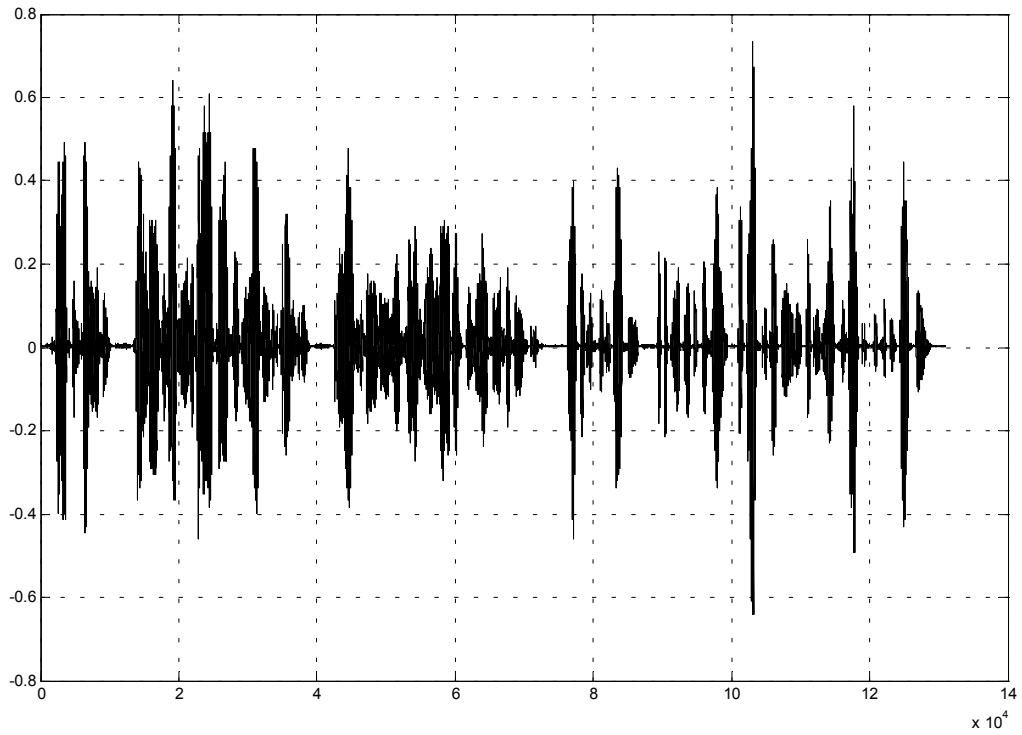


Рис. 13. Осциллограмма (трек) исходного звукового сигнала

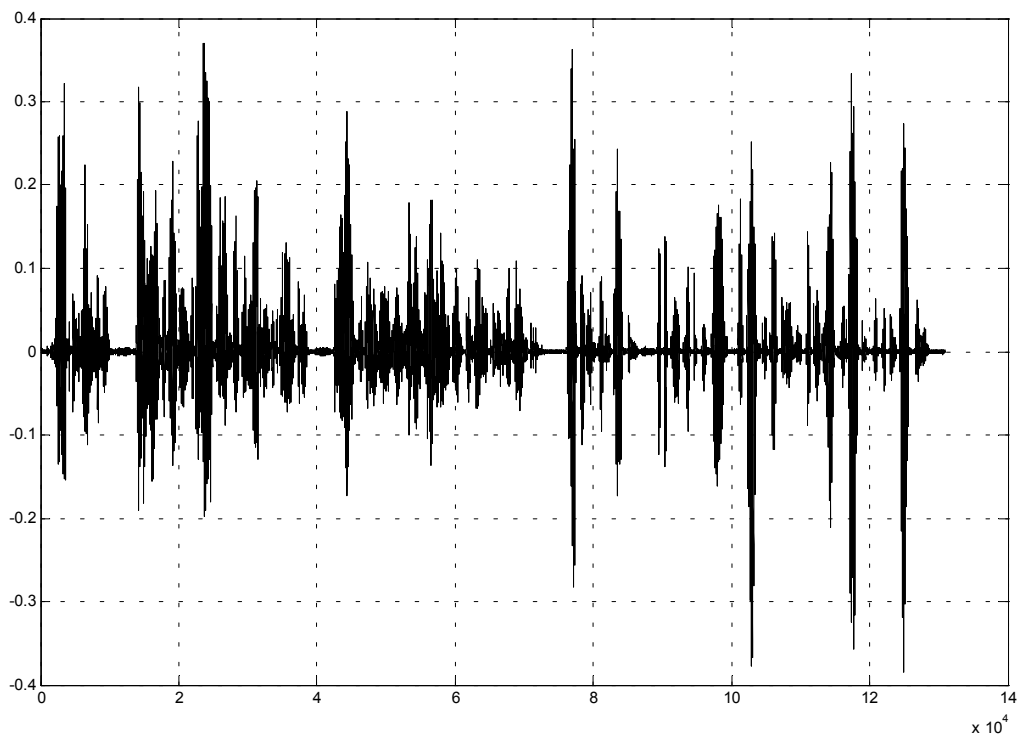


Рис. 14. Осциллограмма (трек) первого остаточного сигнала

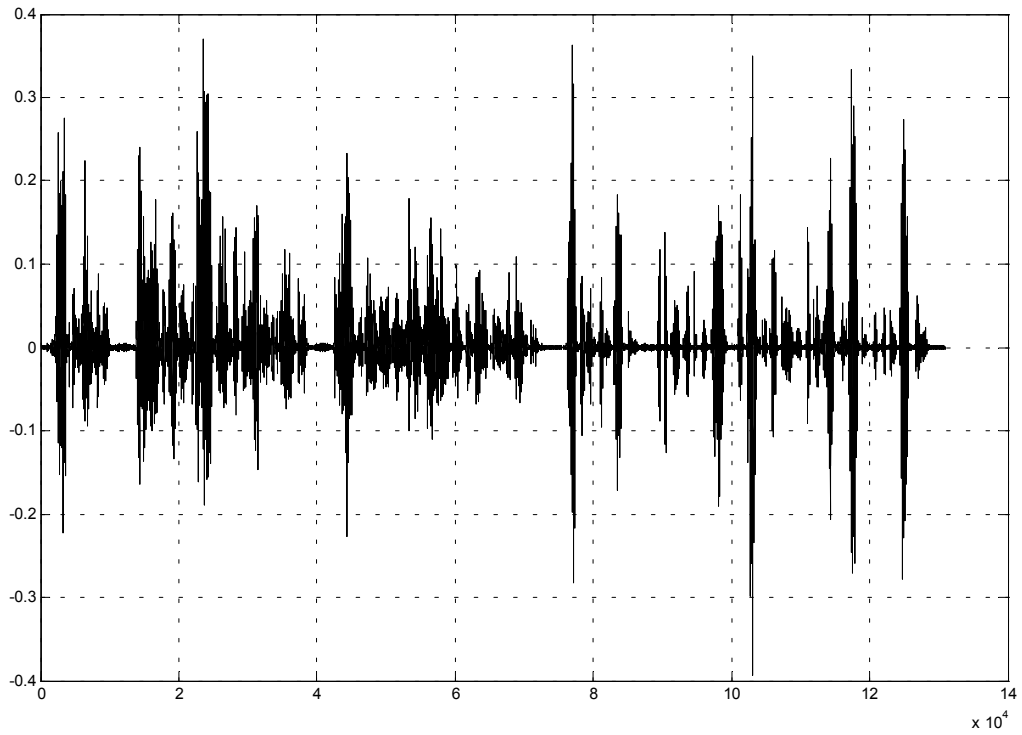


Рис. 15. Оциллограмма (трек) второго остаточного сигнала

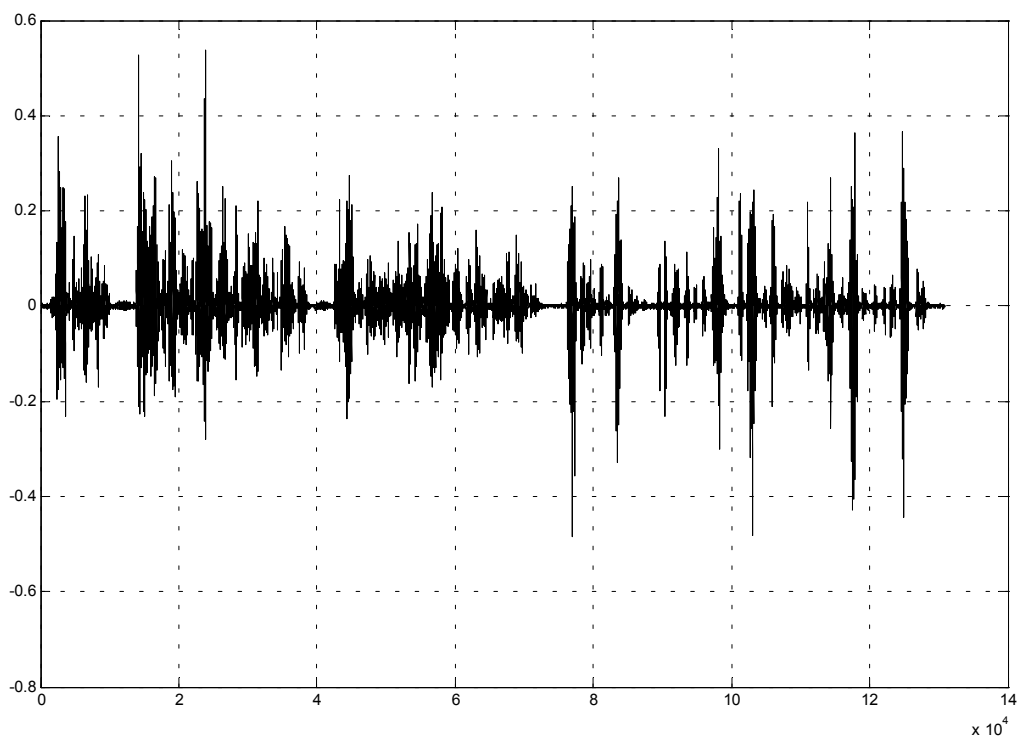


Рис. 16. Оциллограмма (трек) восстановленного второго остаточного сигнала

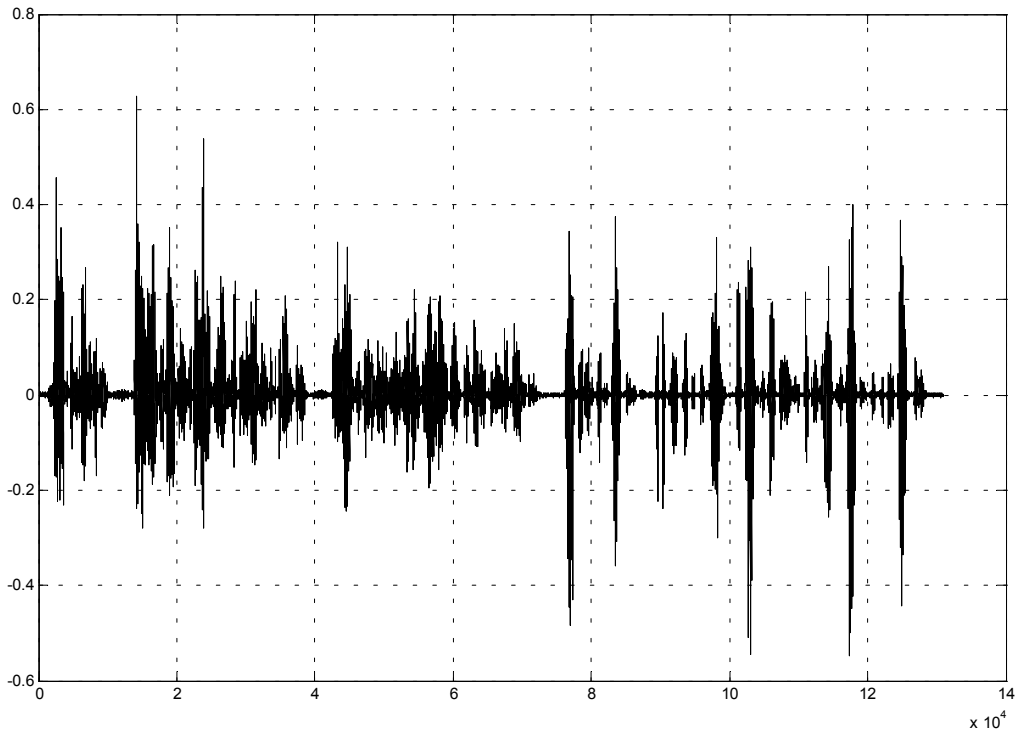


Рис. 17. Оциллограмма (трек) восстановленного первого остаточного сигнала

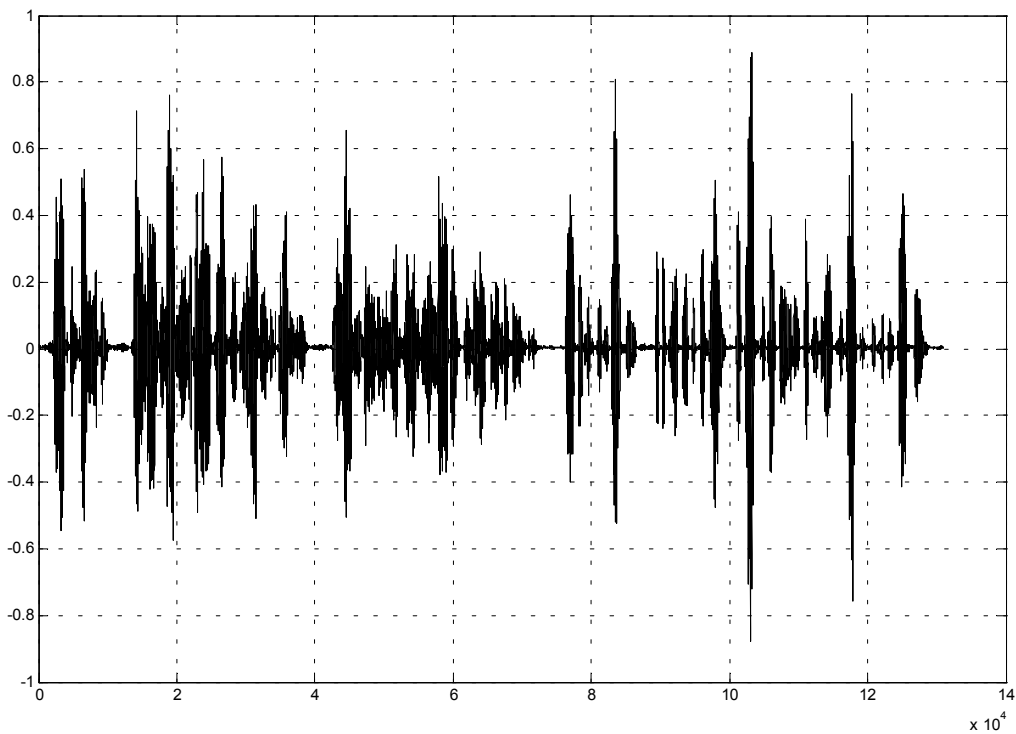


Рис. 18. Оциллограмма (трек) восстановленного звукового сигнала

Из полученных экспериментальных результатов исследования алгоритма сжатия речевых сообщений можно сделать вывод, что качество

восстановленного звукового сигнала в целом удовлетворяет поставленным требованиям технического задания.

2.7. Заключение

1. Разработанный экспериментальный алгоритм сжатия речевых сигналов в целом удовлетворяет поставленным требованиям и может явиться базой для разработки прикладных программ кодирования пакетных сообщений для перспективных телекоммуникационных систем.

2. Необходимо провести дальнейшую поисковую научно-исследовательскую работу по улучшению качества кодирования речи. Предлагаются и ведутся работы над следующими методами алгоритма сжатия речевых сигналов:

- Адаптивное уменьшение длины передаваемого речевого сегмента на основе использования математического аппарата предсказания временных рядов.
- Аппроксимация второго остаточного сигнала с использованием кодовых книг.
- Динамическое изменения порядков фильтров кратковременного и долговременного предсказания на основе использования нестационарного спектрального анализа или вейвлет-анализа.
- Использование принципа формантного сжатия, т.е. путем частичного перевода речевого сообщения в текстовое с последующим использованием методов сжатия без потери информации, например, методами Шеннона-Фано, Лемпела-Зива.

ИСПОЛНИТЕЛИ

Наименование организации	Должность исполнителя	Фамилия имя, отчество	Подпись	Дата
УГАТУ	доцент каф. ТС, к.т.н.	Кузнецов И.В.		
	Ст. преподаватель	Городецкий И.И.		
	Ассистент	Тулбаев Д.Д.		

Используемая литература

1. Бабков В.Ю., Вознюк М.А.,Петраков В.А.,Рыжков А.Е.,СиверсМ.А. Передача информации в системах подвижной связи/ СПбБГУТ,СПб, 1999,с.1-152.
2. Кинтцель Т. Руководство программиста по работе со звуком = A Programmer's Guide to Sound: Пер. с англ.- М.: ДМК Пресс, 2000.- 432с.
3. Рабинер Л.Р., Шафер Р.В. Цифровая обработка речевых сигналов: Перевод с англ. под ред. М.В. Назарова и Ю.Н. Прохорова. М.: Радио и связь. 1981.
4. Kondo A.M. Digital speech. Coding for low rate communication systems. Chichester, 1994. 442p.