



МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ УКРАИНЫ
Севастопольский национальный технический университет

МЕТОДИЧЕСКИЕ УКАЗАНИЯ
к курсовому проектированию по дисциплине
"Системы обработки мультимедиа данных"
для студентов дневной и заочной форм обучения
специальности 7(8).050101 — "Информационные управляющие
системы и технологии"

Севастополь
2012

УДК 004.383.3

Методические указания к курсовому проектированию по дисциплине "Системы обработки мультимедиа данных" для студ. дневной и заочной форм обучения специальности 7(8).050101 — "Информационные управляющие системы и технологии" / Сост. В.Н. Бондарев, А.Л. Овчинников, В.С. Чернега.— Севастополь: Изд-во СевНТУ, 2012. — 95с.

Целью настоящих методических указаний является быстрое введение студентов в проблематику курсового проектирования по дисциплине "Системы обработки мультимедиа данных", а также изложение содержания и порядка работы над курсовым проектом по каждой из предложенных тем.

Методические указания составлены в соответствии с требованиями программы дисциплины "Системы обработки мультимедиа данных" для студентов дневной и заочной форм обучения специальности 7(8).050101 — "Информационные управляющие системы и технологии".

Методические указания утверждены на заседании кафедры информационных систем (протокол N 14 от 26/06/2012 года).

Допущено учебно-методическим центром СевНТУ в качестве методических указаний.

Рецензент: Кожаяев Е.А., канд. техн. наук, доцент кафедры кибернетики и вычислительной техники.

СОДЕРЖАНИЕ

Введение.....	4
1. Тема 1. Распознавание изолированных слов речи.....	6
2. Тема2. Сжатие речи на основе алгоритмов векторного квантования...13	13
3. Тема 3. Верификация диктора по голосу	20
4. Тема 4. Синтез речи по фонологическому тексту.....	27
5. Тема 5. Подавление акустического эхо.....	33
6. Тема 6. Активное управление подавлением шума.....	38
7 Тема 7. Передача дискретных сигналов с фазовой манипуляцией.....	43
8. Тема 8. Передача дискретных сигналов с квадратурной амплитудной манипуляцией.....	49
9. Тема 9. Сжатие речи с помощью канальных и LPC-вокодеров.....	55
10. Тема 10. Сжатие изображений.....	65
11. Тема 11. Реализация звуковых музыкальных эффектов.....	74
12. Тема 12. Сжатие аудиосигналов	89

ВВЕДЕНИЕ

Цель и задачи курсового проектирования

Целью курсового проектирования является систематизация, закрепление и углубление знаний в области технологий обработки мультимедиа сигналов. В процессе выполнения курсового проекта студенты разрабатывают разнообразные технологические цепочки цифрового преобразования сигналов, выполняя программную реализацию элементов систем распознавания и синтеза речи, кодирования речи, сжатия аудио сигналов и изображений, подавления акустического эхо и др.

Задание и этапы проектирования указаны в соответствующих разделах данных методических указаний, в которых описываются предлагаемые темы курсовых проектов.

В ходе выполнения курсового проекта студенты должны приобрести необходимые практические навыки по разработке систем реального времени и закрепить на практике знания, полученные в лекционном курсе.

Содержание пояснительной записки

Пояснительная записка должна выполняться в соответствии с требованиями действующих стандартов и содержать следующие разделы:

- введение;
- обзор существующих систем и методов решения задачи в соответствии с темой курсового проекта;
- обоснование метода решения задачи;
- разработка структурной схемы системы и расчет необходимых системных параметров;
- разработка программной модели системы на языке МАТЛАБ;
- результаты испытаний программной модели системы;
- анализ вычислительной сложности;
- разработка программного обеспечения системы на языке Си и анализ результатов реализации системы в реальном времени на цифровом сигнальном процессоре (выполнение данного раздела уточняется преподавателем) ;
- краткое руководство пользователя;
- заключение;
- библиографический список;
- приложения: тексты программ на языке МАТЛАБ; тексты программ на языке Си.

На защиту студент представляет:

- пояснительную записку;
- электронную презентацию проекта в виде файла формата .ppt или .pdf;
- полностью отлаженное программное обеспечение.

Порядок выбора тем проектирования

Тема курсового проекта может выдаться на группу из 2-х или 3-х студентов (уточняется преподавателем). В пояснительной записке должен быть приведен список авторов с указанием разделов, написанных каждым из них. Возможна выдача одной темы нескольким группам студентов. Однако решения, представляемые каждой из групп, должны быть различными. В случае их совпадения результаты защиты курсового проекта аннулируются для каждой из групп. Защита курсового проекта происходит на 13-14 неделе семестра.

Перед началом проектирования составляется детальный календарный график выполнения курсового проекта, который утверждается руководителем.

1. ТЕМА 1

РАСПОЗНАВАНИЕ ИЗОЛИРОВАННЫХ СЛОВ РЕЧИ

1.1. Цель проектирования

1.1.1. Разработка алгоритма функционирования системы, обеспечивающей распознавание слов речи из ограниченного словаря.

1.1.2. Разработка и исследование программной реализации системы на основе ЦПОС.

1.2. Задание на проектирование

Разработать на ЦПОС TMS320C6711 (5402) систему, выполняющую распознавание изолированных слов речи из словаря, ограниченного следующими словами: “да”, “нет”, “стоп”, “один”, “два”, “три”, “четыре”, “пять”, “шесть”, “семь”, “восемь”, “девять”, “ноль”.

Распознавание слов должно выполняться независимо от диктора в реальном времени. Распознаватель ориентирован на применение в системах, требующих голосового ввода цифровых данных. Распознанные цифры должны отображаться на экране. Слова “да”, “нет”, “стоп” являются командами. Команда “стоп” прерывает текущий сеанс ввода. Команда “нет” отменяет результат распознавания последней цифры. Команда “да” подтверждает результат распознавания и обеспечивает сохранение его в памяти.

1.3. Методические рекомендации по проектированию

Для решения задачи распознавания слов речи из ограниченного словаря могут применяться различные подходы: сопоставление с эталоном, нейронные сети, разнообразные эвристические приемы [2,3,4,5,6].

Рассмотрим распознавание слов речи с помощью сравнения с эталоном. В этом случае структурная схема системы распознавания изолированных слов речи соответствует классической схеме системы распознавания образов [2].

На этапе предварительной обработки речевой сигнал $x(t)$ подвергается фильтрации с целью улучшения качества. Данный этап необходим, если сигнал $x(t)$ содержит помехи.

Процесс формирования образа слов предполагает выполнение двух процедур: сегментации и выделения элементов речи.

В процессе сегментации в потоке речи выделяются участки сигнала $x(t)$, соответствующие отдельным словам. Для этого наиболее часто вычисляют энергию сигнала $x(t)$ на коротких участках длительностью 10-25 мс. Участки с низкой энергией соответствуют паузам между словами.

Действия, выполняемые в ходе процедуры выделения элементов слов, существенно зависят от типов элементов, которые можно положить в основу

акустического представления слова. Если предположить, что акустический образ слова состоит из фонем, то необходимо будет выделять фонемы; если учесть особенности реализации фонем в контексте, то следует выделять аллофоны; если рассматривать акустическую реализацию слова как обычный физический сигнал, то можно его представить в виде последовательности некоторых физических признаков и, следовательно, необходимо будет выделять эти признаки. Словом, выбор здесь довольно большой [3,5,6].

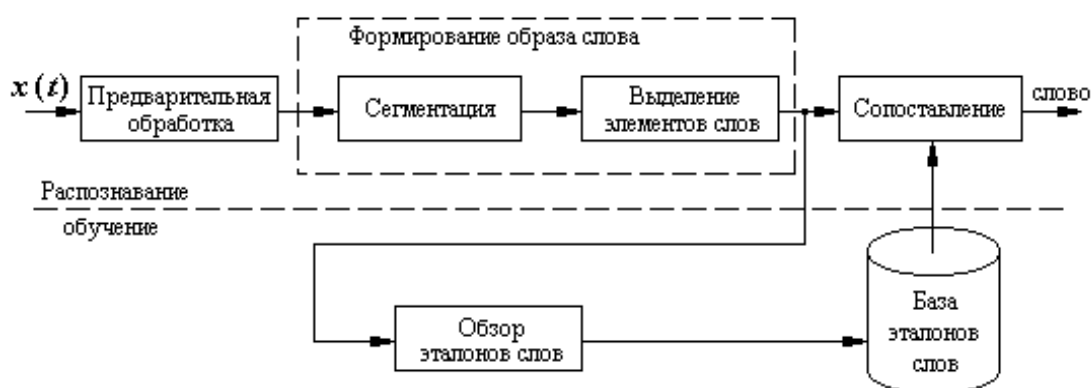


Рисунок 1.1 – Структурная схема системы распознавания

Остановимся на признаковом описании речевого сигнала. Считают, что наиболее перспективны такие описания, которые адекватны процессу речеобразования человека и свойствам его слухового восприятия. Например, формантное описание или описание, основанное на использовании коэффициентов линейного предсказания (ЛПК). Подобные описания подробно изложены в [1].

Пусть для описания акустического образа слова выбрана последовательность кратковременных спектров X_i , $i = 1, 2, \dots, l$ [1]. Тогда слово можно представить в виде спектрограммы

$$x_l = (X_1, X_2, \dots, X_l),$$

где X_i – обозначение спектра $S(f, t_i)$.

Если некоторый спектральный элемент выбросить из последовательности, то получим новую последовательность из $l-1$ элемента. Найдется такое произношение слова (спектрограмма), которое совпадет с преобразованной последовательностью. Различные спектрограммы слова могут быть получены из некоторой базовой спектрограммы путем ее нелинейного сжатия или растяжения вдоль оси времени. Если взять в качестве базовой спектрограмму для самого быстрого и четкого произношения слова, то другие спектрограммы получаются в результате ее растяжения. Обозначим эту спектрограмму последовательностью

$$e_q = (E_1, E_2, \dots, E, \dots, E_q)$$

и назовем ее эталоном.

Если в качестве первого и последнего эталонных элементов выбрать спектральные элементы пауз, то в процессе нелинейного растяжения будут получаться и спектрограммы слов, характеризующиеся различной длиной пауз в начале и конце слова.

Реальные спектрограммы не могут абсолютно точно совпадать с эталонными, объясняется это наличием случайных помех и искажений. Кроме темпа, реальные спектрограммы отличаются также нелинейным изменением интенсивности на длине слова.

Изменения интенсивности учитываются так. Возьмем некоторый эталон e_l . Будем изменять интенсивность эталонных элементов. Для этого первый элемент возьмем с множителем α_1 (вместо E_1 элемент $\alpha_1 E_1$), второй – с множителем α_2 и т.д. Разность $|\alpha_i - \alpha_{i-1}|$ должна быть меньше некоторого порога θ . Теперь имеем новую спектрограмму слова, отличающуюся от e_l интенсивностью спектральных элементов. Это будет допустимая спектрограмма, так как для нее найдется соответствующее произношение слова. Таким образом, из исходной (базовой) спектрограммы можно получить всевозможные спектрограммы слова, которые формируют базу эталонов слов. При этом в базе эталонов хранится исходная (базовая) спектрограмма слова, а эталонные спектрограммы генерируются с помощью указанных правил.

На этапе обучения диктор наговаривает образцы слов, используя быстрое и четкое произношение. Образцы слов преобразуются в последовательности спектральных элементов, которые поступают на вход подсистемы отбора эталонов слов. Данная подсистема выделяет усредненный образец слова, который запоминается в базе данных.

На этапе распознавания происходит следующее. Пусть на вход системы поступило слово, которое представляется в виде спектрограммы x_l длины l . Генерируются по правилам все возможные эталонные спектрограммы e_l и осуществляется поиск наиболее похожей. Сходство спектрограмм оценивается с помощью формулы

$$D(x_l, e_l) = \sum_{i=1}^l g(X_i, \alpha_i E_i),$$

где $g(X_i, \alpha_i E_i)$ - функция, характеризующая сходство наблюдаемого X_i и эталонного $\alpha_i E_i$ спектрального элементов. Часто в качестве функции $g(.)$ используют Евклидово расстояние:

$$g(X_i, \alpha_i E_i) = [X_i - \alpha_i E_i]^2.$$

В приведенной схеме будет генерироваться огромное количество различных эталонных последовательностей e_l .

Для сокращения пространства поиска эталонных слов используют методы динамического программирования. Поясним их суть на примере.

Пусть на вход системы поступило слово $x_8 = (X_1, X_2, \dots, X_i, \dots, X_8)$.

В качестве исходного эталонного образца выбрана последовательность

$$e_5 = (E_1, E_2, E_3, E_4, E_5),$$

на основе которой могут быть сгенерированы эталоны

$$e'_8 = (E_1, E_1, E_2, E_3, E_3, E_4, E_5, E_5)$$

и

$$e''_8 = (E_1, E_2, E_2, E_3, E_3, E_4, E_5, E_5),$$

отличающиеся только во втором элементе. Чтобы ответить на вопрос, какой из эталонов (e'_8 или e''_8) больше похож на реальный образ слова x_8 , нет необходимости в полном сравнении x_8 с эталонами. Достаточно сравнить $g(X_2, E_1)$ и $g(X_2, E_2)$. Если $g(X_2, E_1) < g(X_2, E_2)$, то первый эталон больше похож на образ слова.

Все эталоны, содержащие восемь спектральных элементов и отличающиеся нелинейной деформацией темпа произношения, имеют общие части, поскольку генерируются на основе e_5 . Множество эталонов слова удобно задать в виде графа, изображенного на рисунке 1.2.

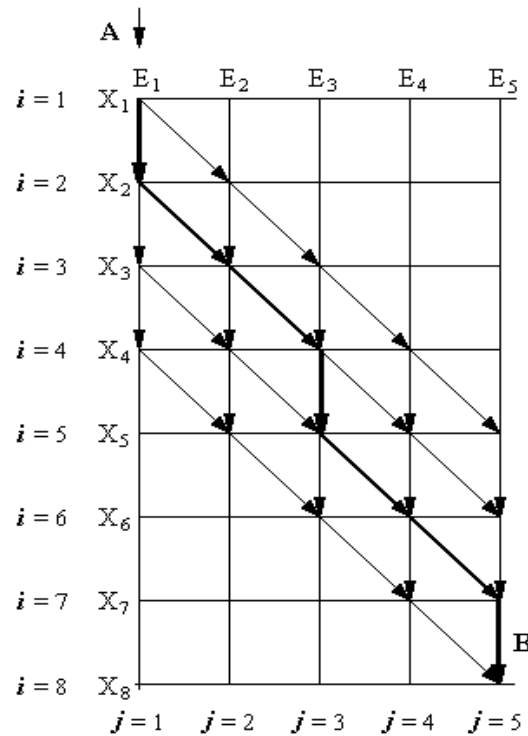


Рисунок 1.2 – Граф деформации эталона

Стрелке, входящей в узел j -го столбца, приписывается j -й эталонный элемент E_j . Тогда каждому пути вдоль стрелок из точки A в точку B будет соответствовать некоторый эталон e_8 слова. Выделенный жирными стрелками путь определяет эталон e'_8 . Все пути из A в B исчерпывающе задают возможный темп произношения.

Пусть $g(X_i, E_j)$ – длина стрелок входящих в узел i, j . Тогда сумма длин стрелок, лежащих на пути из А в В, есть величина, характеризующая сходство x_l с эталоном, т.е. $D(x_l, e_l)$. Таким образом, найти ближайший эталон к распознаваемому образцу означает определить длину кратчайшего пути на графе из точки А в точку В. Для этого вычисляют для каждого узла графа вспомогательные значения F_{ij} .

Примем $F_{11}=g(X_1, E_1)$. Для остальных узлов F_{ij} определяется по формуле

$$F_{ij} = \min\{F_{i-1,j-1}; F_{i-1,j}\} + g(X_i, E_j).$$

Величина F_{lq} и будет расстоянием от распознаваемого сигнала x_l до множества эталонных сигналов слова с различным темпом произношения.

Аналогично можно построить граф, задающий всевозможные эталоны слова, отличающиеся не только темпом произношения, но и изменяющейся интенсивностью произношения.

При распознавании неизвестного слова придется решить столько задач о длине кратчайшего пути сколько слов в словаре, и в качестве решения выдать слово, которое более всего имеет сходство с распознаваемым сигналом.

Рассмотренный алгоритм динамического программирования (ДП) является простейшим. Иные алгоритмы, допускающие не только “растяжение” эталона, но и его сжатие, описаны в [5,6]. В [2] приведен пример программы, реализующий алгоритм ДП.

Для нелинейного согласования темпа речи также могут применяться скрытые марковские модели и алгоритм Витерби [2], алгоритм градиентного спуска [5].

1.4. Порядок проектирования

1.4.1. Изучить способы представления речевых сигналов [1], геометрический метод распознавания образов [2], нормализацию темпа речи на основе метода динамического программирования, в том числе скрытое марковское моделирование и алгоритм Витерби [5,2].

1.4.2. Разработать структурную схему системы, распознающей изолированные слова из ограниченного словаря. Можно воспользоваться схемой, изображенной на рисунке 1.1.

1.4.3. Выбрать и оценить основные системные параметры:

- частоту дискретизации;
- длительность элементарных звуковых единиц, используемых для представления слов;
- максимальную и минимальную продолжительность слов;
- число отсчетов, используемых для представления элементарной звуковой единицы;
- объем памяти, требуемый для хранения эталонов;

- требуемое быстродействие системы (операций/сек).

1.4.4. Разработать в системе MATLAB программные модули, в соответствии со структурной схемой системы. Например, для рисунка 1.1:

- модуль предварительной обработки (центрирование и фильтрация речевого сигнала);
- модуль сегментации (определение пауз и выделение границ слов);
- модуль выделения элементов слов (например, вычисление спектра и его сегментация на частотные полосы);
- модуль отбора эталонов слов (просмотр реализаций слов и их спектральных представлений с целью выделения усредненных образцов слов заданного словаря);
- модуль хранения эталонов (в простейшем случае массив, в котором каждая строка представляет эталон);
- модуль сопоставления (например, функционирующий на основе одного из методов ДП).

1.4.5. Подготовить с помощью разработанных модулей предварительную базу эталонов слов. Для этого необходимо воспользоваться звуковой картой компьютера и микрофоном. При этом необходимо учесть, что звуковая карта компьютера предлагает фиксированные частоты дискретизации в диапазонах 8 КГц, 11 КГц, 22 КГц и 44 КГц. Рекомендуется использовать частоту 8 КГц.

1.4.6. Отладить программные модули системы в пакете MATLAB. В ходе экспериментов подобрать наилучшие параметры модулей, обеспечивающих реализацию функций системы. Отобразить на графиках

- АЧХ и ФЧХ фильтров, используемых в системе;
- примеры временных реализаций слов для разных дикторов;
- примеры спектральных реализаций (спектрограммы) слов для разных дикторов;
- спектрограммы эталонов слов;
- графики изменения энергии речевого сигнала в ходе произношения слова с учетом пауз;
- графики изменения частоты основного тона в процессе произношения эталонов.

1.4.7. Разработать программные модули системы на языке СИ для реализации их на ЦПОС и выполнить их отладку.

1.4.8. Провести тестирование системы, реализованной на ЦПОС для различных дикторов. Оценить вероятности ложного распознавания каждого из слов. Построить графики, указанные в п.6.

1.4.9. Сравнить результаты моделирования системы в пакете MATLAB и реализации ее на ЦПОС. Оценить вычислительную сложность процедур и быстродействие системы.

1.4.10. Сформулировать заключение и предложения по улучшению работы системы.

Библиографический список

1. Бондарев В.Н. Цифровая обработка сигналов: методы и средства/ В.Н.Бондарев, Г.Трестер, В.Н.Чернега.- Харьков: Конус, 2001.-398 с.
2. Бондарев В.Н. Искусственный интеллект/ В.Н.Бондарев, Ф.Г.Аде.- Севастополь: Изд-во СевНТУ, 2002.-616 с.
3. Искусственный интеллект [В 3 кн.] Кн.1. Системы общения и экспертные системы: Справочник / Под ред. Э.В.Попова.- М.: Радио и Связь, 1990.-464с.
4. Компьютеры: справочное руководство [В 3-х т.] / Под ред. Хелмса; Пер. с англ.-М. :Мир, 1986.-403 с.
5. Косарев Ю.А. Естественная форма диалога с ЭВМ/ Ю.А. Косарев.- Л.: Машиностроение,1989.-143с.
6. Методы автоматического распознавания речи [В 2-х кн.] / Под ред. У.Ли.; Пер. с англ.- М.: Мир, 1983.-Кн.1.-328 с. Кн.2.-392 с.

2. ТЕМА 2

СЖАТИЕ РЕЧИ НА ОСНОВЕ АЛГОРИТМОВ ВЕКТОРНОГО КВАНТОВАНИЯ

2.1. Цель проектирования

2.1.1. Разработка алгоритма функционирования системы, обеспечивающей сжатие речи с помощью алгоритмов векторного квантования, и программная реализация алгоритма в системе MATLAB.

2.1.2. Исследование влияния на работоспособность системы аддитивных шумов.

2.1.3. Разработка и исследование программной реализации системы на основе ЦПОС.

2.2. Задание на проектирование

Разработать систему сжатия речи, обеспечивающую сжатие речи до уровня 2400 бит/с и ниже с помощью алгоритмов векторного квантования. Предусмотреть несколько ступеней сжатия. Обеспечить работу системы в двух режимах: дикторо-зависимом и дикторо-независимом. Реализовать систему в пакете MATLAB и подсистему декодирования в реальном времени с помощью ЦПОС TMS320C7711/5402.

2.3. Методические рекомендации по проектированию

На рисунке 2.1,а изображена обобщенная схема сжатия речевого сигнала с помощью алгоритмов векторного квантования.

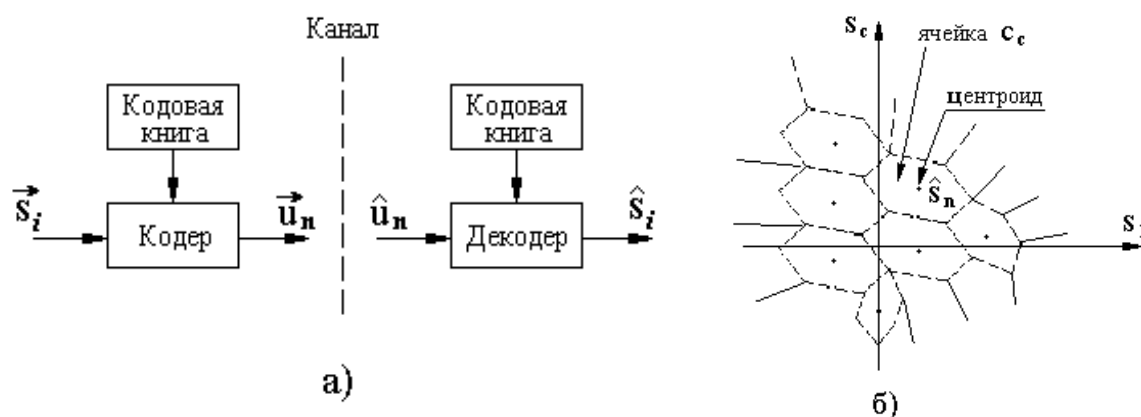


Рисунок 2.1 – Векторное квантование

Входной вектор s_i представляет собой вектор признаков речевого сигнала (например, спектральных),

$$\vec{s}_i = [s_i(0), s_i(1), \dots, s_i(N-1)]^T.$$

Кодер отображает входной вектор $\vec{s}_i$ в выходной символ u_n , $n = 1, 2, \dots, L$ с помощью кодовой книги. Кодовая книга содержит L векторов

$$\hat{\vec{s}}_n = [\hat{s}_n(0), \hat{s}_n(1), \dots, \hat{s}_n(N-1)]^T, n = 1, 2, \dots, L.$$

Предположим, что канал не имеет шумов, т.е. $u_n = \hat{u}_n$.

Векторный квантователь функционирует следующим образом. Входной вектор $\vec{s}_i$ сравнивается с каждым вектором из кодовой книги. В результате из кодовой книги выбирается вектор $\hat{\vec{s}}_n$, ближайший к вектору $\vec{s}_i$, и в канал передается символ u_n , представляющий адрес найденного кодового вектора. На приемной стороне с помощью полученного адреса u_n восстанавливается вектор признаков речевого сигнала $\hat{\vec{s}}_n$, на основе которого синтезируется речевой процесс. В такой интерпретации векторное квантование, по сути, является распознаванием образов, где вектор $\vec{s}_i$ представляет собой входной образ, кодовая книга соответствует базе эталонов (см. рисунок 1.1) [2].

В качестве меры расстояния между входными векторами и векторами из кодовой книги обычно используется сумма квадратов отклонений $s_i(k)$ и $\hat{s}_n(k)$:

$$\varepsilon(\vec{s}_i, \hat{\vec{s}}_n) = \sum_{k=0}^{N-1} (s_i(k) - \hat{s}_n(k))^2. \quad (2.1)$$

Кодовая книга (база эталонов) создается путем деления N - мерного пространства признаков на L непрерывающихся ячеек (областей) (рисунок 2.1,а). Каждая ячейка ассоциируется C_n с вектором-эталом $\hat{\vec{s}}_n$. Если входной вектор $\vec{s}_i$ принадлежит ячейке C_n , то квантователь назначает этому вектору символ u_n , который представляет собой адрес вектора-эталона данной ячейки (центроида).

В простейшем случае, если вектор $\vec{s}_i$ представляет собой блок отсчетов речевого сигнала, рассмотренная схема квантования является обобщением импульсной кодовой модуляции (ИКМ), и называется векторной ИКМ. В векторной ИКМ (ВИКМ) число битов, приходящихся один отсчет речевого сигнала определяется по формуле

$$B = \frac{\log_2 L}{N}. \quad (2.2)$$

ВИКМ имеет преимущество перед различными видами ИКМ [1], если $B \leq 1$.

Процесс проектирования кодовой книги, который связан с обучением, может быть реализован двумя способами. В первом случае кодовая книга разрабатывается на основе алгоритма К-средних [2]. Рекомендуется, чтобы обучающая выборка содержала по 40 примеров векторов признаков для каждого кодового вектора. Вычислительную сложность разработки кодовой книги можно снизить, если определенным образом структурировать кодовую книгу. Действительно, так как в процессе построения кодовой книги выполняется поиск среди L векторов-эталонов, то упорядочение книги может привести к сокращению времени поиска. Для ускорения поиска часто применяют бинарные деревья [2]. Сложность вычислений можно уменьшить, если в кодовой книге отдельно хранить нормализованные векторы $\vec{s}_i$ и масштабный коэффициент G (коэффициент усиления).

Во втором случае кодовая книга создается с помощью алгоритма обучения, в соответствии с которым положение центроидов на каждом шаге уточняется по рекуррентной формуле

$$\vec{s}_n(t+1) = \vec{s}_n(t) + \alpha(t) \cdot [\vec{s}_i(t) - \vec{s}_n(t)], \quad (2.3)$$

где t – номер шага; α - коэффициент обучения, $\alpha \sim 1/\sqrt{t}$. Формула уточняет положение только того центроида, для которого входной вектор $\vec{s}_i(t)$ оказался ближайшим.

Выражение (2.3) соответствует правилу обучения состязательных нейронных сетей, в частности, правилу Кохонена. Подробнее см. в [2].

Существует различные схемы сжатия речи с помощью алгоритмов векторного квантования. Большинство из них основано на схеме “анализ-синтез”. Применяют два варианта таких схем – без обратной связи и с обратной связью [1]. В основе каждой из схем лежит модель синтеза речи на основе коэффициентов линейного предсказания [1]. В соответствии с этой моделью речь может быть получена путем подачи специальным образом подобранного возбуждающего сигнала на вход линейного фильтра, который моделирует резонансные частоты голосового тракта. Передаточная функция фильтра описывается уравнением

$$H(z) = \frac{1}{A(z)} = \frac{G}{1 + \sum_{i=1}^P a_i z^{-i}}, \quad (2.4)$$

где G - коэффициент усиления, a_i - коэффициенты линейного предсказания, P - порядок предсказателя.

Возможная структурная схема системы низкоскоростного кодирования речи с помощью алгоритмов векторного квантования изображена на рисунке 2.2.

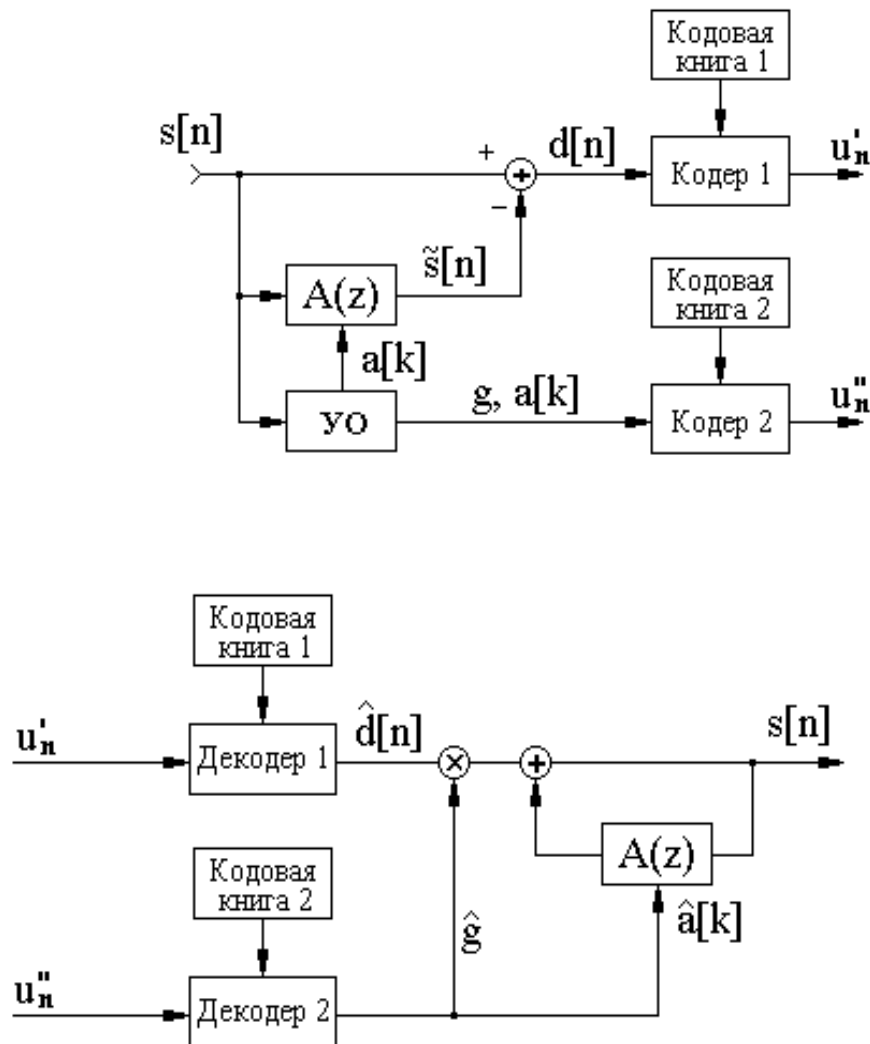


Рисунок 2.2 – Низкоскоростное кодирование речи

Процедура кодирования речи сводится к следующему:

- оцифрованный речевой сигнал $s[n]$ нарезается на сегменты длительностью 20 мс (при $f_g=8$ КГц в каждом сегменте будет по 160 выборок);
- для каждого сегмента вычисляются с помощью устройства оценивания (УО) параметры фильтра линейного предсказания и определяется ошибка предсказания $d[n]$, соответствующая функции возбуждения;
- функция возбуждения и параметры фильтра линейного предсказания кодируются с помощью отдельных векторных квантователей и передаются в канал.

Процедура декодирования заключается в пропуске восстановленного сигнала возбуждения через синтезирующий фильтр (2.4), параметры которого переданы одновременно с функцией возбуждения.

Приведенное описание процессов кодирования и декодирования речи не является исчерпывающим, оно объясняет лишь принцип действия кодера. Практические схемы намного сложнее, и это связано в основном со следующими двумя моментами.

Во-первых, на рисунке 2.2 изображена схема без обратной связи. Лучшего качества синтезируемой речи можно добиться в схемах с обратной связью [1]. Однако такие схемы сложнее.

Во-вторых, описанная выше схема, использует кратковременное предсказание и не обеспечивает в достаточной степени устранения избыточной речи. Поэтому в дополнение к кратковременному предсказанию используется еще и долгосрочное предсказание [1]. Выходной сигнал фильтра кратковременного предсказания используется для оценивания параметров фильтра долгосрочного предсказания – задержки τ и коэффициента предсказания a :

$$H(z) = 1 - az^{-\tau}.$$

При оценке качества кодирования и сопоставлении различных кодеров оцениваются разборчивость речи и качество синтеза речи (качество звучания). Для оценки разборчивости речи используется метод ДРТ (диагностический рифмованный текст). В этом методе подбираются пары близких по звучанию слов, отличающиеся отдельными согласными (“кол-гол-пол”), которые многократно произносятся рядом дикторов, и по результатам испытаний оценивается доля искажений [3,4].

Для оценки качества звучания используется критерий ДМП (диагностическая мера приемлемости) [4]. Испытания заключаются в чтении несколькими дикторами, мужчинами и женщинами, ряда специально подобранных фраз, которые прослушиваются на выходе тракта связи рядом экспертов-слушателей, выставляющих свои оценки по 5-балльной шкале. Результатом является средняя оценка мнений (MOS).

Обратим внимание на следующий факт. Если кодовая книга создается на обучающих данных, принадлежащих только одному диктору, тоне следует ожидать, что она будет обеспечивать хорошее качество звучания для другого диктора. Соответственно, кодовая книга, полученная в лабораторных условиях, не обеспечит того же качества звучания при записи речи в шумовой обстановке, например, в салоне автомобиля. Для построения дикторо-независимой системы необходимо проектировать кодовую книгу на речевых сигналах различных дикторов.

2.4. Порядок проектирования

2.4.1. Изучить способы представления речевых сигналов, обратив особое внимание на линейное предсказание речи [1,3]. Изучить геометрический метод распознавания образов, алгоритм К-средних и алгоритмы обучения состязательных нейронных сетей [2].

2.4.2. Разработать структурную схему системы сжатия речи, обеспечивающей ее передачу со скоростью 2400 бит/с и ниже. Можно воспользоваться схемой, изображенной на рисунке 2.2.

2.4.3. Выбрать и рассчитать необходимые системные параметры:

- частоту дискретизации;
- длительность речевых сегментов;
- порядок предсказывающего фильтра;
- объем кодовой книги;
- разрядность основных параметров;
- быстродействие системы.

2.4.4. Разработать в системе MATLAB программные модули, в соответствии со структурной схемой системы. Например, для рисунка 2.2:

- предсказывающий фильтр и устройство оценивания его параметров;
- подсистему обучения и формирования кодовых книг на основе состязательной нейронной сети (рекомендуется использовать Neural Toolbox);
- подсистему кодирования;
- подсистему декодирования;
- синтезирующий фильтр.

2.4.5. Подготовить с помощью разработанных модулей кодовые книги. Для этого необходимо формировать текстовые файлы, при чтении которых будут многократно, 40-50 раз, встречаться все фонемы русского языка. Предоставить возможность различным дикторам озвучить эти файлы и ввести их в компьютер. Обучить нейронную сеть по методике, описанной в [2], сформировать необходимые кодовые книги. Данная процедура является многоитерационной и может потребовать значительного машинного времени.

2.4.6. Отладить программные модули системы в пакете MATLAB. В ходе экспериментов подобрать наилучшие параметры модулей, обеспечивающих реализацию системы. Отобразить на графиках и рисунках:

- АЧХ синтезирующего фильтра и положение его полюсов;
- примеры реализации функции возбуждения;
- образцы кодовых слов функции возбуждения;
- распределение частот встречаемости кодовых векторов в обучающих данных;
- образцы входных и выходных речевых сигналов, их спектры и спектрограммы.

2.4.7. Разработать программные модули системы на языке СИ для реализации их на ЦПОС и выполнить их отладку.

2.4.8. Провести тестирование системы, реализованной на ЦПОС. Оценить отношение сигнал-шум, оценить разборчивость и качество звучания. Построить графики, указанные в п.6.

2.4.9. Сравнить результаты моделирования системы в пакете MATLAB и реализации ее на ЦПОС. Оценить вычислительную сложность процедур и быстродействие системы.

2.4.10. Сформулировать заключение и предложения по улучшению работы системы.

Библиографический список

1. Бондарев В.Н. Цифровая обработка сигналов: методы и средства/ В.Н.Бондарев, Г.Трестер, В.Н.Чернега.- Харьков: Изд-во Конус, 2001.-398 с.
2. Бондарев В.Н. Искусственный интеллект/ В.Н.Бондарев, Ф.Г.Аде.- Севастополь: Изд-во СевНТУ, 2002.-616 с.
3. Рабинер Л.Р Цифровая обработка речевых сигналов/ Л.Р.Рабинер, Р.В.Шафер.- М.: Радио и Связь. 1981.-495 с.
4. Ратынский М.В. Основы сотовой связи/ М.В. Ратынский; Под ред. Д.Б.Зимины.- М.: Радио и Связь, 1998.- 248 с.
5. Makhoul J. Vector Qnantization // Speech Coding Proceedings of the IEEE, 1985.- Vol. 73. - N 11.- P.1551-1588.

3. ТЕМА 3

ВЕРИФИКАЦИЯ ДИКТОРА ПО ГОЛОСУ

3.1. Цель проектирования

Разработка алгоритма и исследование функционирования системы верификации диктора по голосу, реализованной с помощью цифрового процессора обработки сигналов.

3.2. Задание на проектирование

Разработать систему верификации диктора в реальном времени на цифровом сигнальном процессоре. Система должна предоставлять санкционированный доступ и запрещать несанкционированный доступ. Пользователь (диктор) должен регистрироваться в системе с помощью клавиатуры, вводя номер персональной идентификации. Акустическая модель диктора строится путем сбора образцов особых слов, произносимых пользователем. Система может потребовать нескольких произношений перед тем, как принять решение о доступе или отклонении пользователя. Последовательность контрольных слов не должна быть пользователю известна заранее. Регистрация нового пользователя должна быть удобной. Должно требоваться не больше, чем 50 произношений в течение сеанса регистрации. Сеанс доступа должен требовать не больше, чем 10 произношений, от 2-х до 4-х в среднем.

Пример сеанса регистрации

Система: Ваш персональный PID – 123456.

Запишите его.

Система: Введите Ваш PID

Пользователь: 123456

Система: Ваш номер был назначен.

Начало сеанса регистрации.

Система: Скажите “кот”.

Пользователь: “Кот”.

Система: Скажите “зебра”.

Пользователь: “Зебра”.

И т.д.

Пример сеанса доступа

Система: Введите Ваш PID

Пользователь: PID

Система: Начало сеанса доступа.

Система: Скажите “четыре”.

Пользователь: “Четыре”.

Система: Скажите “кот”.

Пользователь: “Кот”.

Система: Ваша идентификация выполнена.

3.3. Методические рекомендации по проектированию

Распознаванием диктора называется процесс, в ходе которого диктор опознается благодаря характерным свойствам его речи. Существует два вида задач опознавания диктора по голосу – идентификация и верификация диктора. Идентификация диктора по голосу представляет собой процесс, в ходе которого устанавливается ответ на вопрос, кому из зарегистрированных в системе дикторов принадлежит предъявленный речевой сигнал. Верификация диктора – процесс, связанный с подтверждением или отклонением принадлежности речевого сигнала диктору, который идентифицировал себя.

Методы распознавания дикторов по голосу могут быть также разделены на две группы – зависимые от текста и независимые от текста. В первом случае диктор при взаимодействии с системой произносит фразы, которые ранее использовались в ходе регистрации диктора в системе. Во втором случае диктор произносит произвольные фразы.

Методы автоматического опознавания диктора, зависимые от текста, обычно используют операцию сопоставления с образцом (эталоном). Так как образцы речевых сигналов, запомненные в системе, и конкретная акустическая реализация фразы, произносимой диктором при взаимодействии с системой, отличается темпом произнесения, то осуществляется согласование временных масштабов двух фраз, а затем вычисляется некоторая мера их сходства. На основе данной меры принимается решение о принадлежности введенной акустической реализации фразы конкретному диктору.

Методы обработки речевых сигналов, используемые в задаче опознавания диктора по голосу, аналогичны методам, применяемым в системах распознавания речи (см. тему 1). Отличие состоит в том, что в данном случае требуется в ходе обработки выделять такие признаки речи или их совокупность, которые являются уникальными для данного диктора. Обобщенная схема системы автоматической верификации диктора изображена на рисунке 3.1.



Рисунок 3.1 – Обобщенная схема системы верификации диктора

Схема функционирует следующим образом. Диктор, предварительно зарегистрированный в системе, вводит в систему свой идентификационный номер (например, с помощью пластиковой карты, клавиатуры). Затем он произносит в микрофон одну из фраз, которую ему предложила система. Фраза может представляться пользователю либо на экране монитора, либо озвучиваться соответствующими средствами. На этапе преобразования речевого сигнала осуществляются элементарные операции по улучшению введенного речевого сигнала – центрирование, подавление шумов. В ходе выделения из речевого сигнала признаков, используемых для опознавания диктора, каждый сегмент речевого сигнала представляется в виде вектора многомерного пространства признаков. Затем последовательность векторов-признаков $\vec{x}$ сравнивается с моделью диктора на основе операции сопоставления с образцом. В результате формируется значение некоторой меры сходства z . Блок принятия решений сравнивает значение меры с пороговым значением и формирует одно из двух решений – подтвердить соответствие диктора и его идентификатора или отклонить.

Во многих системах автоматической верификации дикторов (АВД) в качестве признаков речевых сигналов используют коэффициенты линейного предсказания (КЛП), коэффициенты отражения, пары частот линейного спектра, кепстр, энергию сигнала и частоту основного тона [2]. Сопоставление может выполняться с помощью методов динамического программирования (ДП, см. задание 1), скрытых Марковских моделей (СММ), искусственных нейронных сетей и векторного квантования [3]. При использовании ДП модель диктора представляет собой некоторый шаблон, в случае СММ модель диктора является статистической, а при векторном квантовании она задается в виде кодовой книги.

На результаты верификации оказывают влияние многие факторы. Перечислим основные из них:

- ошибочное прочтение диктором контрольной фразы;
- эмоциональное состояние диктора (например, стресс);
- положение микрофона;
- окружающие акустические шумы;
- различие в каналах ввода речи при регистрации диктора и при его опознавании;
- болезнь диктора (простуда);
- изменение возраста диктора.

Выделяют два типа ошибок, возникающих в системах АВД, - ложное подтверждение (тип 1 или ЛП) и ложное отклонение (тип 2 или ЛО). Эффективность систем АВД может оцениваться с помощью вероятности правильного и ложного подтверждения. Часто АВД характеризуют ошибкой равного уровня (ОРУ). Такая ошибка соответствует такому значению порога, при котором уровень ложного подтверждения равен уровню ложного отклонения.

При выборе признаков, используемых для опознавания дикторов необходимо руководствоваться следующим:

- они должны быть различными у разных дикторов и примерно одинаковыми у одного и того же диктора;
- должны легко измеряться и вычисляться;
- быть стабильными во времени;
- быть устойчивыми по отношению к изменяющимся свойствам каналов ввода речи;
- быть сложными для имитации другими людьми.

В общем спектральные признаки оказываются более эффективными, чем просодические (интонация, ударение, период основного тона, темп). Наиболее полезными признаками являются формантные частоты: F_2 для назальных звуков, $F_2-F_3-F_4$ для гласных звуков и F_0 для любых звуков [3]. Поэтому в системах АД наиболее часто используют спектральные признаки.

Общим критерием эффективности того или иного признака является F -отношение, которое сравнивает дисперсию признака всех дикторов с дисперсией данного признака у конкретного диктора

$$F = \frac{\text{дисперсия признака разных дикторов}}{\text{дисперсия признака одного диктора}}.$$

Большие значения F говорят о большей разделяющей способности признака. На основе анализа F -отношения установлено, что весьма полезным в ходе решения задачи АД могут оказаться кепстральные коэффициенты. При этом уменьшают значимость коэффициентов кепстра с малыми и большими значениями индексов. Кепстральные коэффициенты с малыми значениями индексов оказываются чувствительными к свойствам канала ввода речи, а кепстральные коэффициенты с большими индексами характеризуются небольшими значениями F -отношения. Для выделения необходимых кепстральных коэффициентов используют схему гомоморфной обработки речи [2], в которой на заключительном этапе используется оконная функция

$$w(k) = 1 + h \cdot \sin(\pi k / L),$$

где $h = L/2$, $k = 1, 2, \dots, L$ и $w(k) = 0$ для всех других k . Обычно $L = 20$. Дополнительно, для исключения влияния свойств канала ввода речи, логарифм спектра речевого сигнала подвергают фильтрации с целью удаления медленноменяющейся аддитивной составляющей. Такая фильтрация может быть выполнена и непосредственно в области кепстральных коэффициентов с помощью фильтра с передаточной функцией:

$$H(z) = \frac{0,2 + 0,1z^{-1} - 0,1z^{-3} - 0,2z^{-4}}{(1 - 0,98z^{-1}) \cdot z^{-4}}.$$

Принятие решений в системах АВД основано на вычислении различных мер (расстояний), определяющих степень сходства между векторами, соответствующими вверенному и образцовому высказываниям. Наиболее часто используют расстояние Махаланобиса [3]:

$$d_i(\vec{x}) = (\vec{x} - \vec{m}_i)^T C^{-1} (\vec{x} - \vec{m}_i),$$

где $\vec{x}$ - входной вектор признаков;

$\vec{m}_i$ - i -ый эталонный вектор (вектор признаков усредненный по всем эталонным произнесениям i -го диктора);

C - ковариационная матрица.

При решении задачи идентификации диктора это правило соотносит введенный речевой сигнал с тем диктором i , для которого $d_i(\vec{x})$ имеет минимальное значение. Обычно используется расстояние Евклида ($C = I$). Однако использование расстояния Евклида не позволяет получать хорошие результаты, если используемые признаки сильно коррелированы. Расстояние Махаланобиса позволяет корректно решать задачу, если вектор признаков характеризуется унимодальной функцией плотности вероятности. В иных случаях рекомендуется использовать расстояние Бхаттачария [7].

В задаче верификации диктора для окончательного решения необходимо сравнить вычисленное расстояние с заранее заданным порогом. Теоретические основы определения величины порога изложены в [1]. Однако рекомендуется оценить значения порога экспериментальным путем. Если важным является минимизировать ошибки первого типа (т.е. пропуск “чужого” диктора), то значение порога должно быть меньше минимального расстояния между эталонами $\vec{m}_i$ и $\vec{m}_j$. Если требуется свести к минимуму ошибки второго типа (т.е. отклонение “своего” диктора), то значение порога выбирается в соответствии с дисперсией эталонного вектора m_i (правило 3σ).

3.4. Порядок проектирования

3.4.1. Изучить способы представления и обработки речевых сигналов [2], изучить основные положения теории распознавания образов [3], изучить метод динамического программирования применительно к проблеме согласования типа речи [5].

3.4.2. Изучить основные подходы, используемые для построения систем автоматического опознавания дикторов по голосу [1,6,7].

3.4.3. Разработать структурную схему системы верификации диктора. Можно воспользоваться схемой, изображенной на рисунке 3.1.

3.4.4. Рассчитать и выбрать основные системные параметры системы АВД:

- частоту дискретизации;
- длительность речевых сегментов;

- параметры частотных характеристик используемых фильтров;
- выбрать систему признакового описания речевого сигнала и определить размерность векторов-признаков;
- определить размер словаря системы и максимальные длительности произнесения слов;
- оценить вычислительную сложность и требуемое быстродействие.

3.4.5. Разработать систему контрольных фраз и принцип регистрации пользователя.

3.4.6. Разработать в системе MATLAB программные модули системы:

- модуль предобработки;
- модуль выделения признаков речевого сигнала;
- модуль сопоставления;
- модуль принятия решений;
- модуль регистрации пользователя и формирования моделей.

3.4.7. Выполнить исследование разработанной системы:

- отобразить на графиках примеры входных речевых сигналов;
- отобразить на графиках примеры эталонных векторов признаков различных дикторов;
- представить в виде таблицы расстояние между эталонными векторами признаков (для 15 дикторов);
- определить для каждого диктора процент ошибок первого и второго типа при трех различных значениях порога (минимальном, среднем и максимальном);
- исследовать влияние шума на качество верификации диктора, добавив к входному сигналу аддитивно белый шум.

3.4.8. Разработать алгоритм и программные модули на языке СИ для реализации системы АД на цифровом сигнальном процессоре TMS320C6711/5402.

3.4.9. Провести исследование системы, работающей в реальном времени, определив процент ошибок первого и второго типа в условиях отсутствия и наличия акустического шума.

3.4.10. Провести сравнение результатов моделирования на MATLAB и реализации системы на цифровом сигнальном процессоре.

3.4.11. Сформулировать выводы.

Библиографический список

1. Атал Б.С. Автоматическое опознавание дикторов по голосам// ТИИЭР.- 1976.- т. 64, № 4.- с. 48-65.
2. Бондарев В.Н. Цифровая обработка сигналов: методы и средства/ В.Н.Бондарев, Г.Трестер, В.Н.Чернега.- Харьков: Конус, 2001.-398 с.
3. Бондарев В.Н. Искусственный интеллект/ В.Н.Бондарев, Ф.Г.Аде.- Севастополь: Изд-во СевНТУ, 2002.-616с.

4. Искусственный интеллект [В 3 кн.] Кн.1. Системы общения и экспертные системы: Справочник/ Под ред. Э.В.Попова.- М.: Радио и Связь, 1990.-464 с.
5. Косарев Ю.А. Естественная форма диалога с ЭВМ/ Ю.А. Косарев.- Л.: Машиностроение.-1989.-143с.
6. Розенберг А.Е. Автоматическая верификация диктора// ТИИЭР, № 4, 1979.- С. 66-79.
7. Cambell J.P. Speaker recognition: a tutorial // Proceedings of the IEEE.- 1997.- Vol. 85.- N 9.- p.1437-1462.

4. ТЕМА 4

СИНТЕЗ РЕЧИ ПО ФОНОЛОГИЧЕСКОМУ ТЕКСТУ

4.1. Цель проектирования

4.1.1. Разработка алгоритма функционирования системы, обеспечивающей синтез речи по тексту, и программная реализация алгоритма в системе MATLAB.

4.1.2. Исследование качества синтезируемой речи.

4.1.3. Разработка и исследование программной реализации системы на основе ЦПОС.

4.2. Задание на проектирование

Разработать систему синтеза речи по тексту, представленному в виде последовательности знаков (фонологического текста), соответствующих фонемам русского языка с указанием продолжительности их звучания и частоты основного тона. Обеспечить удовлетворительную работу системы на подмножестве языка, ограниченного словарем в 1000 слов. Преобразование исходного текста в соответствующее фонемное представление выполняется вручную. Выполнить имитацию системы в пакете MATLAB, а саму систему, функционирующую в реальном времени, реализовать с помощью ЦПОС TMS320C7711/5402.

4.3. Методические рекомендации по проектированию

Речевое сообщение формируется из отдельных элементов речи естественного или искусственного происхождения. Минимальной смысловой единицей речи является фонема. Одна и та же фонема в речи реализуется в виде различных звуков. Последовательности различных звуков, соответствующие одной и той же фонеме, называются *аллофонами*.

В ходе синтеза речевых сообщений используют и иные элементы, формирующую структуру речевого сигнала. Важнейшими из них являются *интонация* и *ударение*, которые определяют *просодию* речи. На уровне акустического сигнала интонация и ударение представляются просодическими характеристиками, к которым относятся [3]: *мелодика* – изменение частоты основного тона; *ритмика* – характеризуется длительностью звуков и пауз; *энергетика* – характеризуется значением интенсивности звука. В процессе синтеза речи орфографический текст сначала преобразуется в последовательность фонем с указанием значений необходимых просодических характеристик. С этой целью текст предварительно разбивается по определенным правилам на интонационно-смысловые группы, важнейшей из которых является *акцентная группа*. Акцентная группа представляет собой одно или несколько слов объединенных общим просодическим контуром (мелодикой, ритмикой, энергетикой), привязанным к единому

групповому ударению. Положение ударения определяется в большей степени длительностью ударного гласного и в меньшей степени его интенсивностью. Для уверенного распознавания положения ударения необходимо, чтобы длительность ударного гласного превышала длительность безударного более чем в 1,5 раза [3].

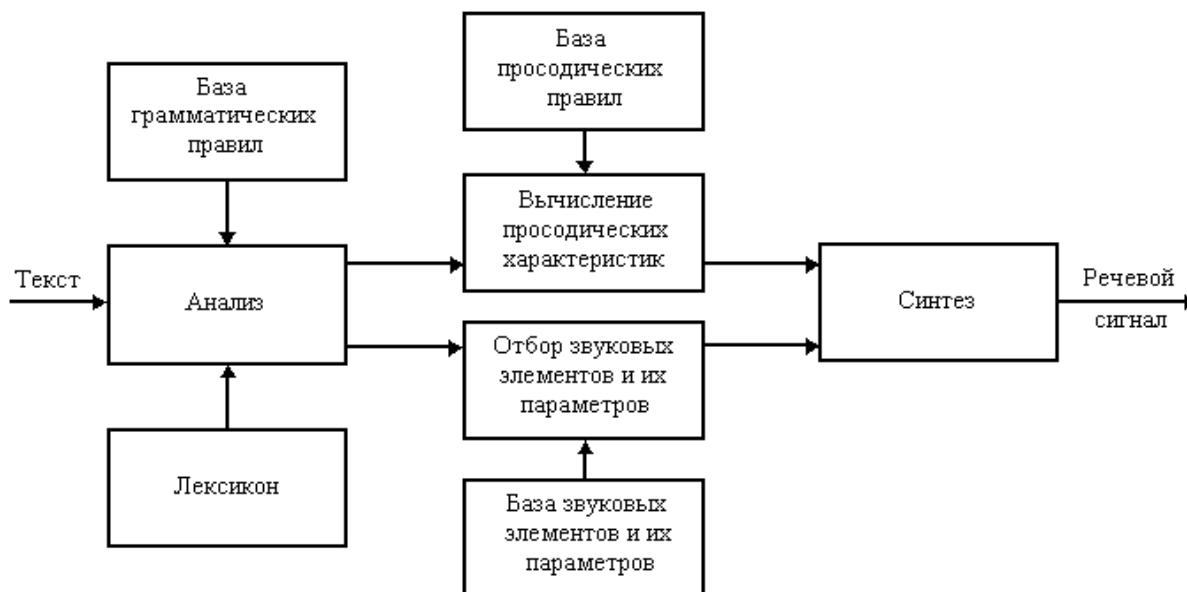


Рисунок 4.1 – Обобщенная схема преобразования текста в речь

На рисунке 4.1 изображена обобщенная схема системы, осуществляющей преобразование текста в речь. В системе предусматривается три этапа обработки. На первом этапе осуществляется морфологический и синтаксический анализ текста с целью получения последовательности фонем и разметки интонационно-смысловых групп предложений. В результате формируется внутреннее фонологическое представление текста. На втором этапе выполняется отбор звуковых элементов (или их параметров), соответствующих фонемному представлению текста, и расчет просодических характеристик: длительностей и интенсивностей звуков, частоты основного тона. Синтез речевого сигнала выполняется на третьем этапе. При этом звуки речи формируются либо из предварительно запомненных звуковых элементов естественного происхождения, либо восстанавливаются на основе параметрической модели речеобразования человека [1].

В системах синтеза речи по тексту применяются три основных метода синтеза речевых сигналов: формантный синтез, синтез на основе коэффициентов линейного предсказания (ЛПК-синтез) и конкатенативный синтез.

Формантный синтез и ЛПК-синтез опираются на модель речеобразования человека. В этом случае образование звуков речи представляется как результат прохождения сигнала источника возбуждения через модель речевого тракта. При формировании гласных звуков источник возбуждения формирует импульсы, частота следования которых (частота основного тона F_0) непрерывно изменяется в соответствии с мелодикой речи. Модель речевого тракта реализуется в виде цифрового фильтра с управляемыми параметрами. Возбуждение цифрово-

го фильтра импульсным источником приводит к формированию гласных звуков. Спектр гласных звуков характеризуется рядом максимумов, соответствующих резонансным свойствам голосового тракта. Частоты максимумов называют *формантными частотами*. Глухие звуки формируются при прохождении через управляемый цифровой фильтр сигнала источника шума. Параметры фильтра могут определяться либо формантными частотами, либо ЛПК-коэффициентами. В первом случае получают формантный синтезатор речи, а во втором – ЛПК-синтезатор речи.

Рассмотрим структурную схему формантного синтезатора [3] (рисунок 4.2). Здесь цифровые фильтры ЦФ1 и ЦФ2 реализуются путем последовательного соединения звеньев второго порядка, моделирующих форманты.

Так как ряд звуков образуется при смешанном характере возбуждения голосового тракта, когда действуют тон и шум, то в рассматриваемой модели цифровой фильтр может возбуждаться сигналами и источника импульсов, и источника шума. Управляющими параметрами модели является:

$A_{рн}$ – амплитуда возбуждения ротового и носового каналов голосового тракта;

A_a – амплитуда аспиративного возбуждения ротового канала (задается при формировании небных согласных).

$A_{ф}$ – амплитуда возбуждения фрикативных формант;

F_o – частота основного тона;

F_i, B_i – соответственно частота и ширина полосы i -ой форманты, $i=1,2,3,4$.

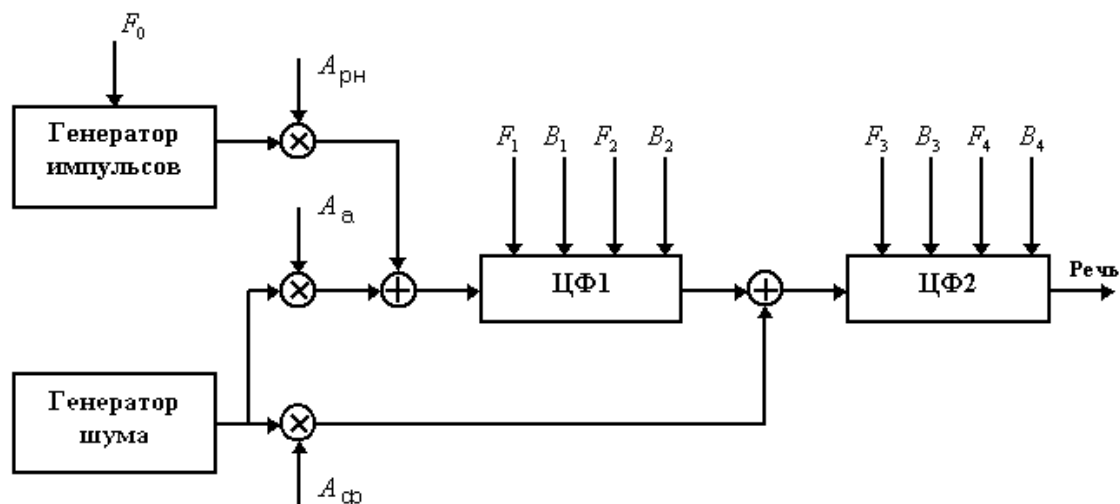


Рисунок 4.2 – Структурная схема формантного синтезатора

При синтезе гласных носовых ('н', 'м' и др.) и аспиративных звуков ('к', 'г' и др.) задаются значения четырех ротовых или двух носовых формант (F_1 - F_4 , B_1 - B_4). При синтезе фрикативных звуков ('с', 'ш' и пр.) фильтр ЦФ2 настраивается на частоты первой и второй фрикативных формант (F_3 - F_4 , B_3 - B_4). Диапазоны значений указанных параметров приведены в [3,4]. Формантные синтезаторы при надлежащем управлении параметрами позволяют получать синтезированную речь высокого качества. Однако их применение требует про-

ведения большой подготовительной работы, связанной с определением параметров формант и их взаимосвязи с фонемами.

В ЛПК-синтезаторах речевой тракт моделируется рекурсивным цифровым фильтром с передаточной функцией

$$H(z) = \frac{1}{1 + \sum_{i=1}^N a_i z^{i-1}}, \quad (4.1)$$

где a_i – коэффициенты, определенные по методу линейного предсказания [1]. Структурная схема ЛПК-синтезатора, подходящего для реализации задания на курсовое проектирование, приведена на рисунке 2.2.

ЛПК-анализ был изначально разработан с целью сжатия и восстановления речевых сигналов. Однако хорошее обоснование ЛПК-модели речеобразования позволили применить ее в задаче синтеза речи по тексту. При этом имеется возможность согласования звуков по энергии, по темпу, по частоте основного тона. Для того чтобы адаптировать схему декодера, изображенную на рисунке 2.2, к решению задачи синтеза речи по фонологическому тексту, необходимо для каждой фонемы (аллофона) определить коэффициенты a_i и обеспечить надлежащее формирование функции возбуждения. Определить типовые коэффициенты a_i для каждой фонемы речи можно с помощью процесса, аналогичного процессу создания кодовых книг при сжатии речи (см. задание 2). В частности, с помощью алгоритма К-средних или состязательных нейронных сетей Кохонена [2]. Функция возбуждения формируется в соответствии с подходом, принятым при построении формантных синтезаторов (рисунок 4.2) [3].

Звуки, формируемые с помощью носового канала голосового тракта, плохо синтезируются фильтром вида (4.1), имеющим только полосы. Кроме этого, расположение пиков в спектре синтезируемой речи оказывается весьма чувствительным к гармонической структуре возбуждающих импульсов (качество синтезируемой речи сильно зависит от спектральной “насыщенности” возбуждающих импульсов). Это приводит к ухудшению качества синтезируемой речи.

Лучшие результаты в отношении разборчивости и натуральности звучания речи удается получить с помощью конкатенативного метода синтеза (от англ. concatenate – сцеплять, сглаживать). Наибольшее распространение получили системы дифонного синтеза. *Дифоном* называют участок речи, расположенный между центрами двух соседних фонем. Дифон содержит переходные элементы речи от одной фонемы к другой, что обеспечивает при синтезе речи ее высокую натуральность. Кроме этого, благодаря тому, что границы дифонов соответствуют серединам фонем, где спектр наиболее стабилен, дифоны легко выделяются. Слова речи при этом синтезируются из предварительно запомненных дифонов. Например, речевой сигнал, соответствующий слову “ива”, образуется из дифонов [/и], [ив], [ва], [а/], которые могли быть получены при сегментации естественной речи при произнесении слов “ил”([/и]), “пиво”([ив]),

“квас”([ва]), “изба”([а/]). Здесь знак “/” обозначает паузу речи. При дифонном синтезе требуется запоминание большого числа дифонов (до нескольких тысяч).

В ходе конкатенативного синтеза возникают две основные проблемы. Во первых, формирование речевого сигнала из сегментов, соответствующих дифонам, требует устранения разрывов, возникающих на границах сегментов. Во вторых, необходимо осуществлять такое независимое управление длительностью дифонов и частотой основного тона, чтобы сохранялись спектральные портреты формант.

Если разрывы на границах сегментов устраняются путем обычного сглаживания, то управление длительностью дифонов и частотой основного тона требует специального рассмотрения [2].

Оценка натуральности звучания и разборчивости синтезированной речи выполняется на основе экспертных оценок. В этом случае эксперты-слушатели, выставляют свои оценки по 5-балльной шкале. Результатом является средняя оценка мнений (MOS).

4.4.Порядок проектирования

4.4.1.Изучить способы представления речевых сигналов, обратив особое внимание на линейное предсказание речи [1,4]. Изучить геометрический метод распознавания образов, алгоритм К-средних и алгоритмы обучения состязательных нейронных сетей [2].

4.4.2.Разработать структурную схему системы синтеза речи, обеспечивающей ее синтез на основе фонологического представления. За основу можно взять схему декодера, изображенную на рисунке 2.2, адаптировав ее в соответствии с рисунками 4.1. и 4.2.

4.4.3.Выбрать и рассчитать необходимые системные параметры:

- частоту дискретизации;
- длительность речевых сегментов, используемых на этапе определения формантных параметров или коэффициентов линейного предсказания;
- порядок (структуру) синтезирующего фильтра;
- количество квазифонем (аллофонов);
- разрядность основных параметров;
- быстродействие системы.

4.4.4.Разработать в системе MATLAB программные модули, в соответствии со структурной схемой системы. Например, для рисунка 2.2:

- предсказывающий фильтр и устройство оценивания его параметров;
- подсистему определения параметров квазифонем на основе состязательной нейронной сети (рекомендуется использовать Neural Toolbox);
- генератор функции возбуждения;
- синтезирующий фильтр.
- подсистему чтения фонологического текста и управления генератором и синтезирующим фильтром.

4.4.5.Подготовить с помощью разработанных модулей базу звуковых элементов (формантных параметров квазифонем или коэффициентов линейного предсказания). Для этого необходимо сформировать текстовые файлы, при чтении которых будут многократно 40÷50 раз встречаться все фонемы русского языка. Предоставить возможность одному диктору озвучить эти файлы и ввести их в компьютер. Обучить состязательную нейронную сеть по методике, описанной в [2], и сформировать необходимую базу звуковых элементов. Определить типовые длительности звуковых элементов, для вокализованных звуковых элементов определить частоту основного тона. Данные процедуры многоитерационные и могут потребовать значительного машинного времени.

4.4.6.Подготовить текст, на основе которого будет выполняться синтез речи. Преобразовать данный текст (объем не менее 1000 слов) в фонологическую форму, выполнить необходимую его разметку. Примеры фонологического представления текста приведены в [2]

4.4.7.Выполнить имитацию системы в пакете MATLAB. В ходе моделирования подобрать наилучшие параметры модулей, обеспечивающих реализацию системы. Отобразить на графиках и рисунках:

- АЧХ синтезирующего фильтра (для всех 42 фонем русского языка);
- примеры реализации функций возбуждения для разных типов звуков;
- образцы реального и синтезированного речевого сигнала (для фиксированного диктора), их спектры и спектрограммы.

4.4.8.Разработать программные модули системы на языке СИ для реализации их на ЦПОС и выполнить их отладку.

4.4.9.Провести тестирование системы, реализованной на ЦПОС. Оценить натуральность речи и ее разборчивость.

4.4.10.Сравнить результаты моделирования системы в пакете MATLAB и реализации ее на ЦПОС. Оценить вычислительную сложность процедур и быстроедействие системы.

4.4.11.Сформулировать заключение и предложения по улучшению работы системы.

Библиографический список

1. Бондарев В.Н. Цифровая обработка сигналов: методы и средства/ В.Н.Бондарев, Г.Трестер, В.Н.Чернега.- Х.: Конус, 2001.-398 с.
2. Бондарев В.Н. Искусственный интеллект/ В.Н.Бондарев, Ф.Г.Аде.- Севастополь: Изд-во СевНТУ, 2002.-616 с.
3. Искусственный интеллект [В 3 кн.] Кн.1. Системы общения и экспертные системы: Справочник/ Под ред. Э.В.Попова.- М.:Радио и Связь, 1990.-464 с.
4. Кейтер Дж. Компьютеры синтезаторы речи/ Дж. Кейтер; Под ред. В.А. Усика.- М.: Мир,1985.—237с.

5. ТЕМА 5

ПОДАВЛЕНИЕ АКУСТИЧЕСКОГО ЭХО

5.1. Цель проектирования

5.1.1. Разработка алгоритма функционирования системы, осуществляющей подавление акустического эхо в процессе телеконференций.

5.1.2. Разработка и исследование программной реализации системы на ЦПОС.

5.2. Задание на проектирование

Разработать систему подавления акустического эхо для организации телеконференций. Считать, что передаваемые сигналы являются речевыми. Расстояние между микрофоном и громкоговорителем 1.0 – 1.5 м.

Исследовать модель системы средствами пакета MATLAB, реализовать систему, работающую в реальном времени, на ЦПОС TMS 320C6711/5402.

5.3. Методические рекомендации по проектированию

Эхоподавители впервые были предложены в 1967 г. Сондхи [5] и предназначались для подавления “электрического” эхо, возникающего в процессе передач телефонных сообщений на большие расстояния. В 1970 г. лабораторией AT&T Bell была разработана первая БИС, осуществляющая эхоподавление, основанная на технологии цифровой обработки сигналов. В настоящее время вопросы эхоподавления регламентируются стандартами G.165 и G.167 МКТТ.

Принципы эхоподавления, разработанные для систем телефонии, стали широко применяться и при решении других задач, в частности, для подавления акустического эхо, возникающего в ходе телеконференций.

Структурная схема системы, осуществляющей подавление акустического эхо, изображена на рисунке 5.1.

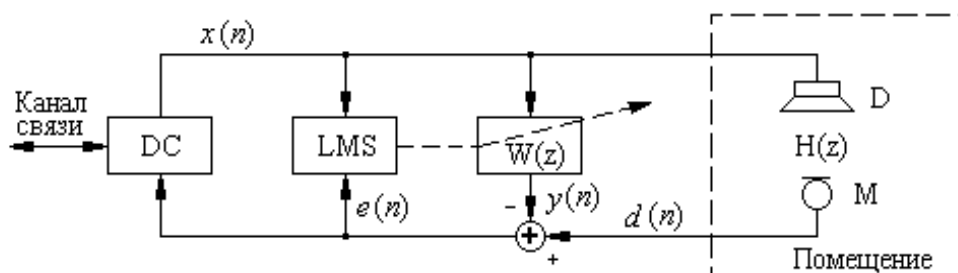


Рисунок 5.1 – Структурная схема системы подавления акустического эхо

Схема содержит следующие блоки: DC - дифференциальная система; D - громкоговоритель (динамик); M - микрофон; LMS - блок, осуществляющий вычисление коэффициентов адаптивного фильтра на основе алгоритма минимизации среднеквадратической ошибки (LMS – least mean square); адаптивный фильтр, характеризуемый передаточной функцией $W(z)$.

Опишем принцип работы схемы. Речевой или аудио сигнал $x(n)$ поступает из канала связи на дифференциальную схему DC, осуществляющую отделение входного сигнала $x(n)$ от выходного сигнала $e(n)$. Входной сигнал $x(n)$ воспроизводится громкоговорителем D и озвучивает помещение. Речь диктора с наложенным на нее акустическим сигналом (пропорциональным $x(n)$), создаваемым громкоговорителем, воспринимается микрофоном M и преобразуется в электрический сигнал $d(n)$. Если сигнал $d(n)$ передать непосредственно в канал связи, то на стороне источника сигнала $x(n)$ будет прослушиваться эхо. Чтобы устранить эхо, необходимо из сигнала $d(n)$ вычесть сигнал, создаваемый на входе микрофона громкоговорителя. Для этого требуется восстановить передаточную функцию акустического канала $H(z)$ между громкоговорителем и микрофоном. Поведение указанного акустического канала моделируется с помощью адаптивного фильтра, уравнение функционирования которого можно записать в виде

$$y(n) = \bar{w}^T(n) \bar{x}(n),$$

где $\bar{w}(n) = [w_0(n), w_1(n), \dots, w_{L-1}(n)]^T$ - вектор весовых коэффициентов адаптивного фильтра; $\bar{x}(n) = [x(n), x(n-1), \dots, x(n-L+1)]^T$ - входной сигнал фильтра; L - порядок адаптивного фильтра. Выходной сигнал адаптивного фильтра $y(n)$ должен соответствовать сигналу, создаваемому громкоговорителем на входе микрофона. В этом случае, вычитая из сигнала $d(n)$ сигнал $y(n)$, можно исключить прохождение “следов” сигнала $x(n)$ в канал, и тем самым устранить эхо. Так как свойства акустического канала, характеризуемого передаточной функцией $H(z)$, меняются в зависимости от уровня акустических шумов в помещении, то коэффициенты фильтра подбираются так, чтобы обеспечить минимум среднеквадратического значения сигнала $e(n)$

$$e(n) = d(n) - y(n).$$

Коэффициенты фильтра удобно вычислять на основе рекуррентного соотношения

$$\bar{w}(n+1) = \bar{w}(n) + \mu e(n) \cdot \bar{x}(n), \quad (5.1)$$

где μ – коэффициент, регулирующий скорость сходимости и устойчивость алгоритма адаптации. Теоретические основы построения адаптивных фильтров и практические примеры их применения изложены в [1,3]. Адаптация фильтра $W(z)$, по сути, представляет собой процесс обучения. Вопросы построения адаптивных фильтров тесно связаны с построением обучающихся систем. Алгоритмы рекуррентного обучения, соответствующие (5.1), рассматриваются в [2,4].

Следует отметить несколько практических моментов, связанных с реализацией схемы, изображенной на рисунке 5.1. Во-первых, если диктор произносит речь близко к микрофону, то алгоритм LMS отключают. С этой целью в систему вводят детекторы такой речи. Во-вторых, если дифференциальная система будет неидеальной, то сигнал $e(n)$ может смешиваться с сигналом $x(n)$. Это приводит к образованию “электрического” эхо. Соответственно, оно должно также подавляться. Кроме этого, в помещении может возникнуть реверберация (многократное отражение акустического сигнала от стен). Это требует включение в состав эхоподавителя адаптивного фильтра высокого порядка. Указанные элементы порождают множество вариантов реализации эхоподавителя.

Рассмотренная система эхоподавления является монофонической. Большой интерес представляют стереофонические системы (рисунок 5.2).

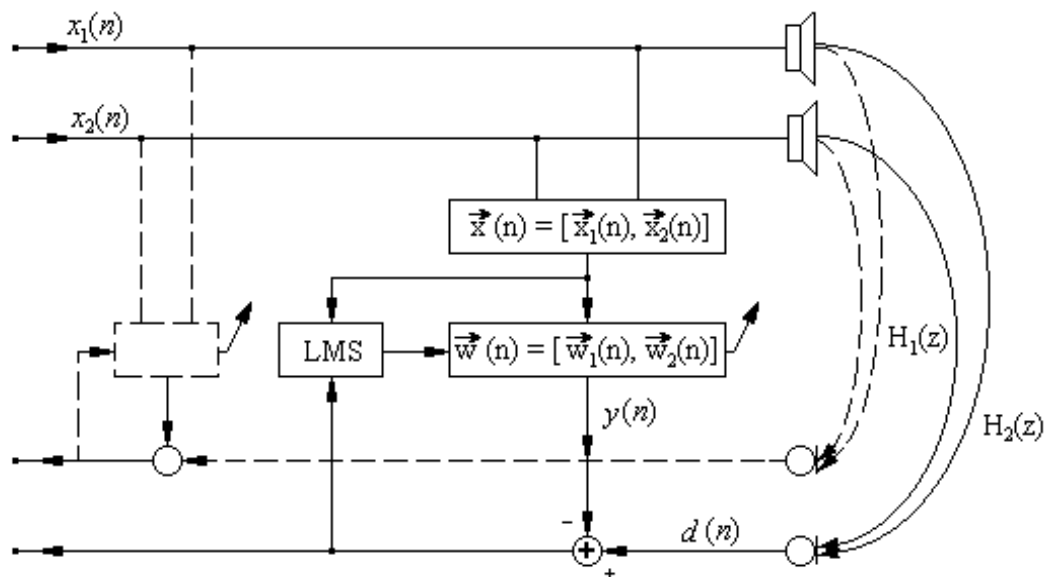


Рисунок 5.2 – Стереофоническая система подавления акустического эхо

Стереофонические системы телеконференций обеспечивают большую реалистичность звучания. Они позволяют слушателям воспринимать пространственную информацию. В стереофонических системах, имеющих два канала, входные сигналы $\bar{x}_1(n)$ и $\bar{x}_2(n)$, а также коэффициенты $\bar{w}_1(n)$ и $\bar{w}_2(n)$ фильтра комбинируются, формируя векторы $\bar{x}(n) = [\bar{x}_1^T(n), \bar{x}_2^T(n)]$ и $\bar{w}(n) = [\bar{w}_1(n), \bar{w}_2(n)]$. Следовательно, стереофоническая система представляет собой линейную комбинацию двух монофонических систем. Однако при создании стереофонических систем возникают специфические проблемы, обусловленные корреляционными связями между каналами, перемещением диктора и др. Это снижает скорость сходимости алгоритма LMS.

Часто для повышения скорости сходимости применяют алгоритмы полосовой адаптивной фильтрации. В этом случае сигналы анализируются в небольших частотных полосах, и в каждой частотной полосе осуществляется независимое подавление эхо. Это позволяет не только повысить скорость сходи-

мости алгоритмов адаптации, но также повышает и их вычислительную эффективность.

В настоящее время системы подавления акустического эхо являются коммерческими продуктами и играют важную роль при организации телеконференций. Реализация систем осуществляется на основе цифровых сигнальных процессоров.

5.4. Порядок проектирования

5.4.1. Изучить способы подавления эхо сигналов, обратив особое внимание на алгоритм настройки адаптивных фильтров [1,2,3,4].

5.4.2. Разработать структурную схему системы подавления акустического эхо. Можно воспользоваться схемой, изображенной на рисунке 5.1 или 5.2.

5.4.3. Выбрать и рассчитать необходимые системные параметры:

- частоту дискретизации;
- порядок адаптивного фильтра;
- разрядность коэффициентов фильтра;
- быстродействие системы.

5.4.4. Разработать в системе MATLAB (Simulink) программные модули в соответствии со структурной схемой системы. Например, для рисунка 5.1:

- адаптивный фильтр;
- блок адаптации коэффициентов фильтра;
- дифференциальную систему (с учетом ее неидеальности);
- модель акустического канала;
- модель сигнала $x(n)$, а также модель речевого сигнала, подаваемого на микрофон.

5.4.5. Отладить программные модули в пакете MATLAB. В ходе экспериментов подобрать наилучшие параметры модулей, обеспечивающих реализацию системы. Отобразить на графиках и рисунках:

- примеры сигнала $x(n)$ и сигнала, подаваемого на микрофон;
- АЧХ, модели акустического канала;
- импульсную характеристику адаптивного фильтра, после завершения процесса адаптации;
- графики схождения параметров адаптивного фильтра;
- АЧХ адаптивного фильтра после завершения процесса адаптации;
- график сигнала ошибки $e(n)$ до начала процесса адаптации и после.

Рекомендуется работу системы проверить как на гармонических сигналах, так и на реальных речевых сигналах. Также перед началом моделирования рекомендуется изучить пример системы адаптивного подавления шума, имеющий в среде MATLAB (DSP block set).

5.4.6. Разработать программные модули на языке СИ для реализации их на ЦПОС и выполнить их отладку.

5.4.7. Провести тестирование системы, реализованной на ЦПОС. Оценить качество работы системы. Построить графики указанные в п. 5.

5.4.8. Сравнить результаты моделирования системы в пакете MATLAB и реализацию ее на ЦПОС. Оценить вычислительную сложность процедур и быстродействие системы.

5.4.9. Сформулировать заключение и предложения по улучшению работы системы.

Библиографический список

1. Адаптивные фильтры: Пер. с англ./ Под ред. К.Ф.Н.Коуэна и П.М.Гранта.- М.: Мир, 1988.- 392 с.
2. Бондарев В.Н. Искусственный интеллект/ В.Н.Бондарев.- Севастополь: СевНТУ, 2002.-616 с.
3. Уидроу Б. Адаптивная обработка сигналов: Пер. с англ./ Б.Уидроу, С. Стирнз.- М.: Радио и Связь, 1989.- 440 с.
4. Цыпкин Я.З. Основы теории обучающихся систем/ Я.З. Цыпкин.- М.: Наука, 1973.- 252 с.
5. Sondhi M. An adaptive echo canceller// Bell. Tech. J. – 1967.- V.46(3).- p.497-511.
6. Widrow B. Adaptive noise cancelling principles and applications// B.Widrow, J.R.Glover, J.M.McCool.// Proc. IEEE, 1975.- V.63.- P.1692-1716 (имеется перевод см. журналы ТИИЭР за 1976 г.)

6. ТЕМА 6

АКТИВНОЕ УПРАВЛЕНИЕ ПОДАВЛЕНИЕМ ШУМА

6.1. Цель проектирования

6.1.1. Разработка алгоритма функционирования системы, осуществляющей подавление акустического шума, и исследование его в системе MATLAB.

6.1.2. Разработка и исследование программной реализации системы на основе ЦПОС.

6.2. Задание на проектирование

Разработать систему активного управления подавлением уровня шума. В качестве источника шума использовать искусственный периодический сигнал с частотой 300 Гц. Обеспечить генерацию такого сигнала с помощью звуковой карты компьютера. Необходимо обеспечить подавление указанного шума в зоне вокруг головы оператора. Расстояние от оператора до источника шума 1.5 – 2.0 М.

6.3. Методические рекомендации по проектированию

Активное управление акустическими шумами было впервые предложено в 1930 г. Несмотря на успешные научные разработки, практическая реализация систем активного управления шумом стала возможным только после появления цифровых сигнальных процессоров.

Активное управление шумом (АУШ) базируется на идее использования вторичного источника акустического сигнала, который формирует акустическое поле, деструктивно интерферирующее с акустическим полем первичного источника шума. С целью требуемой компенсации первичного источника, вторичный источник должен адаптироваться ко всем изменениям, происходящим с первичным источником шума. Системы АУШ имеют некоторую аналогию с системами адаптивного подавления шума [3]. В последних компенсации шума происходит на уровне электрических сигналов, тогда как в системах АУШ – на уровне акустических сигналов. Тем не менее, два указанных класса систем используют, по сути, один и тот же принцип подстройки цифрового адаптивного фильтра. На рисунке 6.1 изображены структурные схемы указанных систем.

В ходе активного подавления шумов подавляются те шумовые компоненты электрического сигнала $d(n)$, формируемого на выходе микрофона, которые линейно коррелируют с сигналом $x(n)$, фиксирующим свойства шума. Для этого подбираются коэффициенты передаточной функции адаптивного фильтра W , так, чтобы достичь минимума среднего квадрата сигнала ошибки $e(n)$.

Целью функционирования системы АУШ является снижение акустического воздействия на микрофон, создаваемого внешним источником шума. Это достигается путем подстройки адаптивного цифрового фильтра W , управляющего громкоговорителем. При этом подстройка выполняется так, чтобы мини-

мизировать значение электрического сигнала $e(n)$ на выходе микрофона, фиксирующего изменения звукового давления в ограниченном пространстве. Весьма важным приложением системы АУШ является подавление периодических звуков, создаваемых подвижными частями различных механизмов. Представляет определенный интерес подавление шумов, создаваемых оргтехникой в учебном помещении.

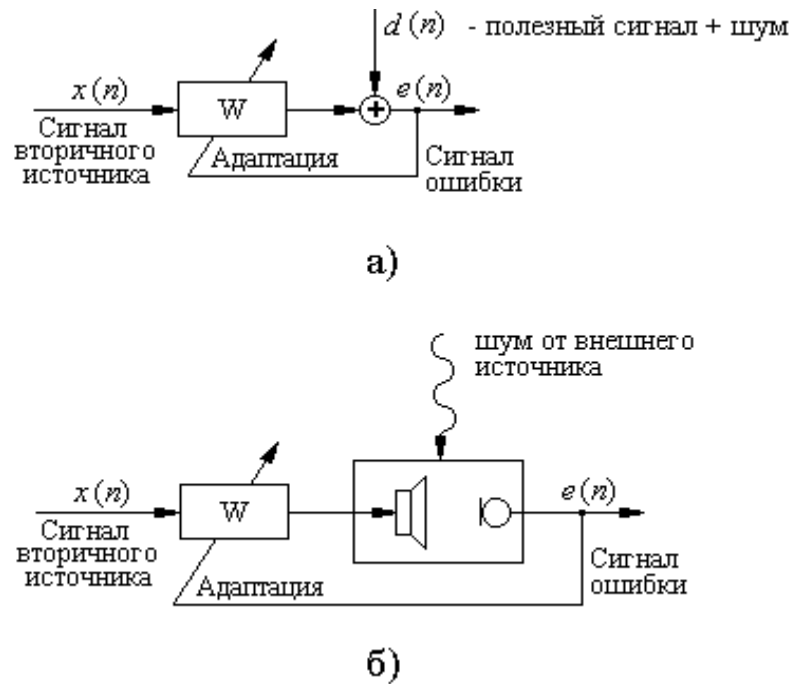


Рисунок 6.1 – Структурные схемы систем адаптивного подавления шума а) и активного управления шумом б)

Отметим, что в системах АУШ эффект подавления шума является локальным. Удастся снизить уровень шума примерно на 10 ДБ в сферическом пространстве вокруг микрофона с диаметром, равным примерно $0,1\lambda$, где λ – длина волны акустического шума. Поэтому, если говорить об улучшении условий труда, специалиста по информационным системам, пользующимся средствами оргтехники, то наиболее подходящим местом размещения акустических элементов АУШ является верхняя часть спинки кресла. Это позволяет создать зону “тишины” вокруг головы оператора.

На рисунке 6.2 изображена практическая схема установки системы АУШ.

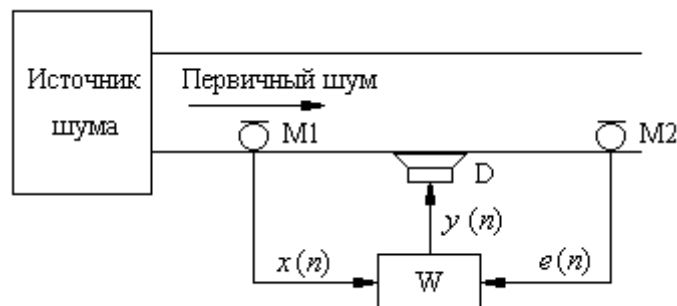


Рисунок 6.2 – Установка системы АУШ

Нежелательный шум от первичного источника фиксируется микрофоном М1 (микрофон М1 устанавливается ближе к источнику шума) и поступает на вход адаптивного фильтра W. Адаптивный фильтр управляется громкоговорителем D, который формирует акустические звуки, подавляющие шум. Остаточный шум фиксируется микрофоном М2. Параметры адаптивного фильтра вычисляются таким образом, чтобы минимизировать остаточный шум.

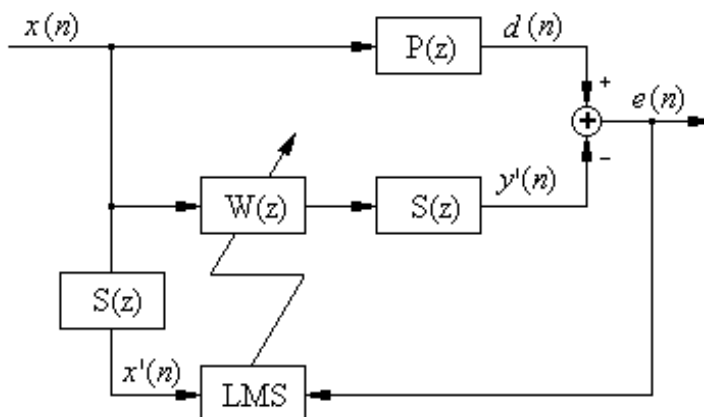


Рисунок 6.3 – Модель системы АУШ

На рисунке 6.3 изображена модель рассмотренной системы АУШ, где $P(z)$ - передаточная функция акустического канала между микрофонами М1 и М2; $s(z)$ - передаточная функция между динамиком D и микрофоном М2; LMS - алгоритм, вычисляющий коэффициенты фильтра путем минимизации среднего квадрата ошибки.

Коэффициенты адаптивного фильтра на каждом шаге работы LMS – алгоритма вычисляются на основе выражения [4]:

$$w(n+1) = w(n) + \mu \cdot e(n) \cdot x'(n),$$

где μ – коэффициент адаптации; $x'(n) = [x'(n)x'(n-1)...x'(n-L+1)]^T$ - сигнал на выходе фильтра с передаточной функцией $s(z)$. Данный сигнал определяется из выражения

$$x'(n) = s(n) * x(n),$$

где $s(n)$ – импульсная характеристика фильтра $s(z)$; * - операция свертки. Импульсная характеристика $s(n)$ обычно определяется экспериментально в результате решения задачи идентификации [2].

6.4.Порядок проектирования

6.4.1.Изучить способы подавления шума и активного управления подавлением шума [1,3].

6.4.2. Разработать структурную схему системы активного управления шумом. Можно воспользоваться схемой, изображенной на рисунках 6.2.

6.4.3. Выбрать и рассчитать основные системные параметры:

- частоту дискретизации;
- ориентировочный диаметр зоны “тишины”;
- разработать геометрическую схему размещения элементов системы;
- задержки прохождения звукового сигнала от источника шума до микрофонов M1 и M2;
- определить порядок L адаптивного фильтра.

6.4.4. Выполнить идентификацию акустического канала между громкоговорителем D и микрофоном M2, определить экспериментально импульсную и амплитудно-частотные характеристики канала. Под задачей идентификации понимают восстановление модели системы (рисунок 6.4) по известным входным и выходным сигналам системы.

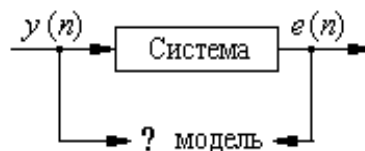


Рисунок 6.4 – Задача идентификации системы

Теоретические вопросы идентификации систем описаны в [2]. При решении задачи идентификации рекомендуется использовать Identification Toolbox пакета MATLAB.

6.4.5. Разработать в системе MATLAB программные модули в соответствии с рисунком 6.3:

- модуль, моделирующий акустический канал с передаточной функцией $P(z)$;
- адаптивный фильтр;
- фильтр с передаточной функцией $s(z)$;
- модуль вычисления коэффициентов адаптации;
- генератор шумового сигнала.

6.4.6. Провести исследование программной модели в системе MATLAB. Отобразить на графиках и рисунках:

- реализацию и спектр шумового сигнала;
- реализацию и спектры сигнала $e(n)$ на разных этапах адаптации;
- график изменения среднего квадрата ошибки во времени;
- АЧХ и импульсные характеристики адаптивного фильтра на разных этапах адаптации;
- АЧХ и импульсные характеристики фильтра $s(z)$.

Рекомендуется внимательно изучить пример системы адаптивного подавления шума, имеющийся в системе MATLAB (см. DSP BLOCK SET).

6.4.7. Разработать программные модули системы на языке СИ для реализации их на ЦПОС. Выполнить отладку данных модулей на ЦПОС TMS320C6711/5402.

6.4.8. Провести тестирование системы, реализованной на ЦПОС. Оценить уровень снижения шума, который обеспечивает разработанная система в качестве источника шума.

6.4.9. Сравнить результаты моделирования системы в пакете MATLAB и с помощью ЦПОС.

6.4.10. Сформулировать заключение.

Библиографический список

1. Адаптивные фильтры: Пер. с англ./ Под ред. К.Ф.Н.Коуэна и П.М.Гранта.- М.: Мир, 1988.- 392 с.
2. Дейч А.М. Методы идентификации динамических объектов/ А.М. Дейч.- М.: Энергия, 1979.-240 с.
3. Уидроу Б. Адаптивная обработка сигналов: Пер. с англ./ Б.Уидроу, С. Стирнз.- М.: Радио и Связь, 1989.- 440 с.

7. ТЕМА 7

ПЕРЕДАЧА ДИСКРЕТНЫХ СИГНАЛОВ С ФАЗОВОЙ МАНИПУЛЯЦИЕЙ

7.1. Цель проектирования

7.1.1. Разработка алгоритма функционирования системы передачи дискретных сигналов (модема) с двухпозиционной фазовой манипуляцией и программная реализация алгоритма в системе MATLAB на цифровом сигнальном процессоре.

7.1.2. Исследование работоспособности системы при воздействии на сигнал аддитивной флуктуационной помехи.

7.1.3. Исследование влияния на работоспособность системы параметров фильтров, входящих в состав передатчика и приемника.

7.1.4.Разработка и исследование программы реализации спроектированного модема на основе цифрового сигнального процессора.

7.2. Задание на проектирование

7.2.1.Скорость манипуляции - 1000 Бод.

7.2.2. Несущая частота - 1800 Гц.

7.2.3. Минимальное отношение эффективных напряжений сигнал/помеха - 2.

7.2.4. Точность подстройки фазы опорного колебания приемника $\pm 5^0$.

7.2.5. Вероятность ошибочной регистрации единичных элементов 10^{-4} .

7.3. Методические рекомендации по проектированию

Типовая упрощенная схема системы передачи дискретных сигналов с фазовой манипуляцией изображена на рисунке 7.1 [1].

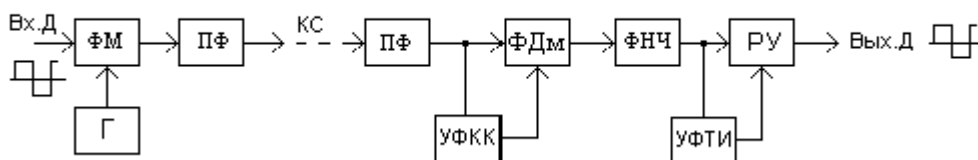


Рисунок 7.1- Упрощённая структурная схема СПДС с ФМ

Передатчик состоит из фазового модулятора (ФМ), генератора несущей частоты (Г) и канального полосового фильтра (ПФ). При фазовой модуляции в случае изменения модулирующего сигнала по закону $f(t)$ и максимальном изменении начальной фазы на величину $\Delta\phi$ фаза сигнала изменяется по закону:

$\theta = \omega_0 t + \varphi_0 + \Delta\varphi f(t)$. Мгновенное значение фазомодулированного напряжения имеет следующий вид:

$$U_{\text{ФМ}} = U_{\text{М}} \cos[\omega_0 t + \varphi_0 + \Delta\varphi f(t)]. \quad (7.1)$$

Здесь $\Delta\varphi$ – индекс фазовой модуляции (девиация фазы). Для двухпозиционных СПДС $\Delta\varphi$ принимают равным 90° . На рисунке 7.2 показаны эюры ФМ-сигнала на выходе передатчика и смесь сигнала с помехой на входе приемника.

При модуляции несущего колебания последовательностью прямоугольных разнополярных импульсов сигнал на выходе модулятора меняется скачкообразно на 180° , т.е. имеет место фазовая манипуляция.

В общем случае спектр ФМ сигнала содержит несущую частоту, симметрично которой располагаются боковые составляющие, отстоящие на частотные интервалы, кратные частоте манипуляции.

В канале связи (КС) передаваемый сигнал искажается вследствие неидеальности АЧХ и ФЧХ канала. На него также воздействуют флуктуационные и импульсные помехи.

Полосовой фильтр приемника выделяет из смеси сигнала с помехой часть, в которой сосредоточена основная энергия сигнала. Из этого процесса в устройстве формирования когерентных колебаний формируется опорное колебание для демодуляции ФМ-сигнала в фазовом демодуляторе (ФДм).

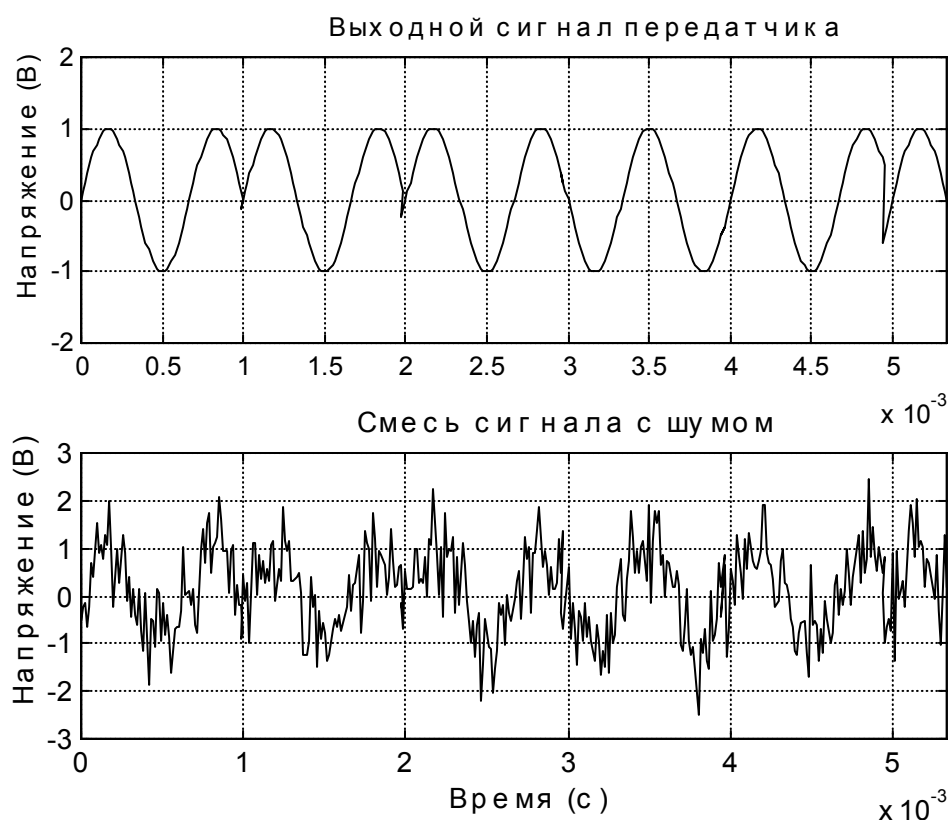


Рисунок 7.2 – Эюры ФМ- сигналов на выходе передатчика и входе приемника

Фазовый демодулятор производит перемножение колебаний, поступающих на его вход с опорным сигналом:

$$u_{оп}(t) = U_{оп} \sin(\omega_{оп}t + \varphi_{оп}).$$

Напряжение на выходе ФДм с коэффициентом передачи 1 и при отсутствии помех и с учетом подавления фильтром НЧ высокочастотных составляющих с частотой $\omega_0 + \omega_{оп}$ определяется выражением

$$U_{ФНЧ}(t) = (U_0 U_{оп} / 2) \cos[(\omega_0 - \omega_{оп}t) + (\varphi_0 - \varphi_{оп})]. \quad (7.2)$$

Здесь U_0 и $U_{оп}$ – амплитуды сигнала и опорного напряжения соответственно.

На основании этого напряжения решающее устройство вырабатывает выходной сигнал данных. Решение в РУ принимается в момент подачи стробирующего сигнала, вырабатываемого устройством формирования тактовых импульсов (УФТИ).

Полосовой фильтр приемника выделяет из смеси сигнала с помехой часть, в которой сосредоточена основная энергия сигнала. Из этого процесса в устройстве формирования когерентных колебаний формируется опорное колебание для демодуляции ФМ- сигнала в фазовом демодуляторе (ФДм). Из пульсирующего напряжения ФДм фильтром нижних частот (ФНЧ) выделяется огибающая на основании которой решающее устройство вырабатывает выходной сигнал данных. Решение в РУ принимается в момент подачи стробирующего сигнала, вырабатываемого устройством формирования тактовых импульсов (УФТИ).

К системным параметрам модема относятся:

- нестабильность частоты генератора несущей частоты;
- полоса пропускания и центральная частота полосовых фильтров передающей и приемной частей системы;
- полоса пропускания фильтров нижних частот СПДС;
- минимальный уровень сигнала на входе приемника.

Нестабильность частоты генерации несущих колебаний в СПДС по каналам тональной частоты регламентируется рекомендациями МККТТ и составляет ± 6 Гц.

При использовании фазовой манипуляции с передачей двух боковых полос ширина полосы пропускания ПФ передатчика определяется по формуле [3]

$$\Delta F_{ПФ} \approx 1,42 B. \quad (7.3)$$

Так как несущая частота равна F_H , то полоса пропускания полосового фильтра равна

$$F_H \pm 0,71 B. \quad (7.4)$$

Полосовой фильтр на входе приемника имеет характеристики, аналогичные ПФ передатчика.

Полосовой фильтр, подключенный на выходе квадратора должен выделять компоненту с удвоенной частотой, т.е. $2 F_H$, и подавлять частотные компоненты, кратные F_H , не менее чем на 20дБ.

Фильтр нижних частот системы фазовой автоподстройки частоты служит для выделения из процесса на выходе фазового детектора постоянной составляющей. Поэтому частота среза такого фильтра выбирается в области 10 Гц.

ФНЧ фазового демодулятора должен выделять огибающую ФМ-сигнала. В связи с этим его полоса пропускания определяется скоростью манипуляции B и должна быть равной

$$\Delta F_{\text{ФНЧ}} \approx 0,71 B. \quad (7.5)$$

Перед моделированием собственно СПДС целесообразно вначале спроектировать цифровые фильтры, входящие в ее состав. Для этого задаются желаемые частотные характеристики фильтров, вычисляются нормированные частоты, определяющие полосу пропускания, а затем производится вычисление коэффициентов соответствующих цифровых фильтров.

В процессе моделирования СПДС следует провести исследование влияние типов и порядка фильтров, формы их характеристик на функционирование системы и помехоустойчивость приема сигналов.

В большинстве случаев целесообразно использовать рекурсивные фильтры относительно низкого порядка (2-5) для минимизации запаздывания, которые вносят фильтры, особенно при старте системы.

При построении модели СПДС необходимо предусмотреть вывод графиков АЧХ и ФЧХ фильтров.

Частоту дискретизации F_d при реализации фильтров следует выбирать таким образом, чтобы при последующем моделировании ФМ-сигналов количество отсчетов на один период несущей было не менее 40, т.е. рекомендуется задать частоту дискретизации $F_d = (40-50)F_H$.

При моделировании ФМ-колебаний целесообразно организовывать цикл формирования единичного элемента сигнала. В таком случае очень просто сгенерировать последовательность произвольной длительности. Следует учесть, что минимальное количество единичных элементов должно быть достаточным для подстройки частоты опорных колебаний в приемнике. Его можно определить экспериментальным путем, наблюдая за ошибкой синхронизации, т.е. процесс подстройки частоты должен продолжаться до тех пор, пока ошибка фазирования не достигнет минимума и стабилизируется (см. рисунок 7.3).

Для генерирования флуктуационной помехи целесообразно воспользоваться функцией `random`, которая формирует вектор помехи в диапазоне значений (0 –1). Изменение диапазона мгновенного значения помехи достигается путем умножения вектора на постоянный коэффициент.

7.4.3. Рассчитать системные параметры СПДС.

7.4.4. Разработать в системе MATLAB программу формирования ФМ-колебаний при поступлении на вход модулятора разнополярных импульсов единичной амплитуды.

7.4.5. Спроектировать цифровой фильтр нижних частот, используемый для формирования спектра передаваемых сигналов и исследовать влияние параметров фильтра (полосы пропускания, порядка и типа фильтра) на ширину спектра канальных сигналов и время запаздывания сигнала на выходе фильтра.

7.4.6. Сгенерировать вектор отсчетов флуктуационной помехи и добавить помеху к передаваемому сигналу.

7.4.7. Разработать в системе MATLAB программу формирования на приемной стороне когерентных опорных колебаний с подстройкой фазы под принимаемый ФМ-сигнал.

7.4.8. Разработать в системе MATLAB программу демодуляции ФМ-сигналов.

7.4.9. В процессе моделирования на MATLAB отобразить на графиках:

- АЧХ- и ФЧХ- характеристики всех фильтров, используемых в модеме;
- выходной сигнал передатчика модема без помехи;
- сигнал с помехой на входе приемника;
- сигнал опорного генератора приемника для каждого шага итерации процедуры подстройки опорного генератора;
- сигнал на выходе демодулятора модема.

7.4.10. Разработать алгоритм и программные модули на языке СИ для реализации модема на цифровом сигнальном процессоре.

7.4.11. Провести исследование программной реализации процедуры передачи и приема ФМ-сигналов цифровым сигнальным процессором.

7.4.12. Отобразить на графиках процессы, отмеченные в 7.4.9.

7.4.13. Провести сравнение результатов моделирования на MATLAB и реализации модема программным способом на ЦПОС.

7.4.14. Сделать выводы по работе.

Библиографический список

1. Боккер П. Передача (Техника связи в системах телеобработки данных). Том 1. Основы: Пер с нем./ Под ред. Д.Д.Кловского. - М.: Связь, 1980. - 264 с.
2. Бондарев В.Н. Цифровая обработка сигналов: методы и средства/ В.Н.Бондарев, Г.Трестер, В.Н.Чернега. - Х.: Конус, 2001. - 398 с.
3. Чернега В.С. Расчет и проектирование технических средств обмена и передачи информации/ В.С. Чернега, В.А. Василенко, В.Н. Бондарев. - М.: Высш. шк., 1990. - 224 с.
4. Шварцман В.О. Теория передачи дискретной информации/ В.О. Шварцман, Г.А.Емельянов. - М.: Связь, 1979. - 424 с.

8. ТЕМА 8

ПЕРЕДАЧА ДИСКРЕТНЫХ СИГНАЛОВ С КВАДРАТУРНОЙ АМПЛИТУДНОЙ МАНИПУЛЯЦИЕЙ

8.1. Цель работы

Разработка алгоритма и исследование функционирования системы передачи дискретных сигналов (модема) с квадратурной амплитудной манипуляцией, реализованной на цифровом сигнальном процессоре.

8.2. Задание на проектирование

8.2.1. Разработать электрическую схему на основе цифрового сигнального процессора и программное обеспечение модема с 16-позиционной квадратурной амплитудной модуляцией (КАМ), соответствующего рекомендации V.29 МККТТ.

8.2.2. Устройство должно содержать:

- преобразователь кода входной последовательности;
- модулятор и демодулятор прямого канала связи;
- каналные фильтры;
- схемы сопряжения с источником и получателем информации;
- схему формирования когерентных колебаний.

8.2.3. Обосновать выбор и рассчитать системные параметры передатчика и приемника (частоту манипуляции и частоту тактовых импульсов, параметры полосовых канальных фильтров и фильтров нижних частот демодулятора и схемы формирования опорных колебаний, минимальный уровень сигнала на входе приемника).

8.2.4. Построить имитационную модель системы передачи дискретных сигналов (СПДС) с использованием пакета “MATLAB” или “LabVIEW” и исследовать функционирование системы со снятием временных и спектральных диаграмм в характерных точках устройства.

8.2.5. Провести исследование влияния типа и порядка используемых в СПДС фильтров на форму и спектры выходного сигнала передатчика и сигналов в характерных точках приемника.

8.2.6. Разработать электрическую принципиальную схему и программное обеспечение реализации СПДС на базе процессора с учетом системных параметров, полученных в предыдущих пунктах.

8.2.7. Произвести отладку разработанной СПДС на отладочном комплексе Code Composer Studio V2 и учебном микропроцессорном наборе с использованием цифровых сигнальных процессоров типа TMS320C5402 либо TMS320C6711 и выполнить исследование работоспособности устройства в реальном времени.

8.2.9. Провести сравнение временных и спектральных характеристик в характерных точках схемы с данными имитационного моделирования.

8.3. Методические рекомендации по проектированию

Типовая упрощенная схема системы передачи дискретных сигналов с квадратурной амплитудной манипуляцией изображена на рисунке 8.1 [1].

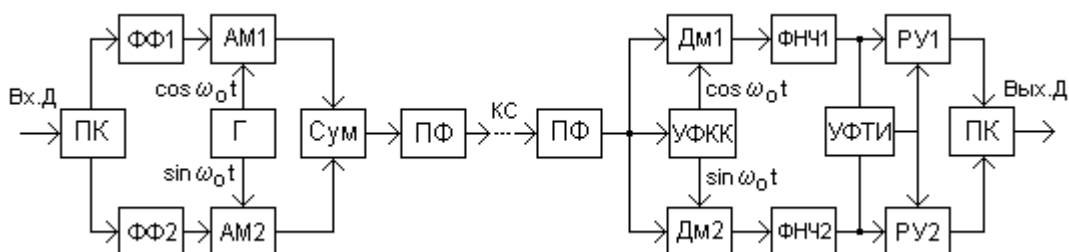


Рисунок 8.1- Структурная схема СПДС с КАМ

Передатчик состоит из двух каналов - синфазного и квадратурного, состоящих соответственно из формирующих фильтров ФФ1 и ФФ2 и амплитудных модуляторов АМ1 и АМ2. При передаче данных со скоростью 9600 бит/с двоичная последовательность входных данных (Вх.Д) разбивается на группы, состоящие из четырех бит (квадрибиты). Каждому квадрибиту в преобразователе кода ПК ставится в соответствие значение амплитуды $A(t)$ для синфазного и $B(t)$ – для квадратурного каналов.

Колебания несущей частоты ω_0 на модуляторы подаются с генератора Г со сдвигом на 90° . Амплитудно-модулированные колебания с выходов обоих модуляторов складываются в сумматоре Сум и, после ограничения спектра в канальном полосовом фильтре, поступают в канал связи КС. Выражение для сигнала на выходе передатчика имеет вид

$$U(t) = A(t)\cos \omega_0 t + B(t)\sin \omega_0 t . \quad (8.1)$$

В результате сложения двух колебаний, находящихся в квадратуре, образуется сигнал с амплитудно-фазовой модуляцией (АФМ). Эпюра АФМ-сигнала показана на рисунке 8.2.

В канале связи к сигналу добавляются флуктуационные и сосредоточенные помехи. На передаваемый сигнал оказывает влияние также неравномерность амплитудно-частотной (АЧХ) и фазо-частотной (ФЧХ) характеристик.

На приемной стороне полосовой фильтр приемника выделяет из смеси сигнала с помехой часть, в которой сосредоточена основная энергия сигнала. Затем сигнал подается на демодулятор синфазной ДМ1 и квадратурной ДМ2 составляющих. В устройстве формирования когерентных колебаний УФКК из принимаемого сигнала с АФМ формируется ортогональные опорные колебания, сдвинутые между собой по фазе на 90° . Для восстановления информации принимаемый сигнал (1) демодулируется путем умножения на ортогональные несущие колебания, когерентные соответствующим колебаниям передатчика.

Напряжения после демодуляторов синфазного и квадратурного каналов описываются соответственно выражениями:

$$U_{\text{СИН}}(t) = A(t) \cos^2 \omega_0 t + B(t) \sin \omega_0 t \cos \omega_0 t = 0,5[A(t) + A(t) \cos 2 \omega_0 t + B(t) \sin 2 \omega_0 t];$$

$$U_{\text{КВ}}(t) = B(t) \sin^2 \omega_0 t + A(t) \cos \omega_0 t \sin \omega_0 t = 0,5[B(t) - B(t) \cos 2 \omega_0 t + A(t) \sin 2 \omega_0 t].$$

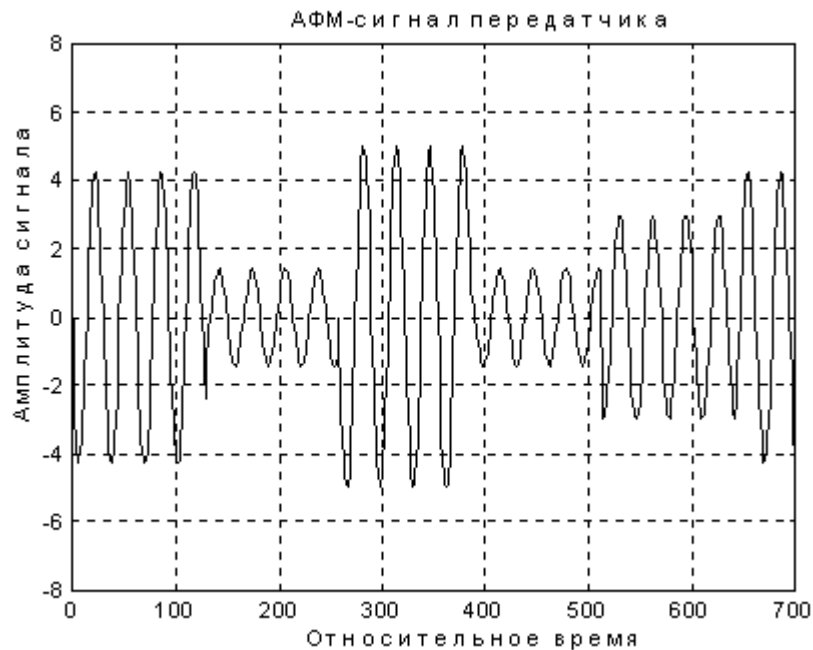


Рисунок 8.2 - Эпюра АФМ-сигнала на выходе передатчика

После подавления высокочастотных компонент в приемных фильтрах ФНЧ1 и ФНЧ2 сигналы обоих каналов полностью разделены:

$$U_{\text{ФНЧ1}}(t) = A(t) \text{ и } U_{\text{ФНЧ2}}(t) = B(t). \quad (8.2)$$

На рисунке 8.3 показаны эпюры напряжений $U_{\text{ФНЧ1}}(t)$ и $U_{\text{ФНЧ2}}(t)$, полученные в результате моделирования работы приемника в системе MATLAB.

Решение в решающих устройствах синфазного и квадратурного каналов РУ1 и РУ2 принимается в момент подачи стробирующего сигнала, вырабатываемого устройством формирования тактовых импульсов (УФТИ). В преобразователе кода приемника ПК производится преобразование двух амплитуд $A(t)$ и $B(t)$ в соответствующий им квадрибит.

К системным параметрам модема относятся:

- нестабильность частоты генератора несущей частоты;
- полоса пропускания и центральная частота полосовых фильтров передающей и приемной частей системы;
- полоса пропускания фильтров нижних частот СПДС;
- минимальный уровень сигнала на входе приемника.

Нестабильность частоты генерации несущих колебаний в СПДС по кана-

лам тональной частоты регламентируется рекомендациями МККТТ и составляет ± 6 Гц.

Формирующие фильтры ФФ1 и ФФ2 служат для ограничения спектра передаваемых сигналов. Ширина спектра сигнала существенно зависит от вида частотной характеристики ФФ. На практике используют фильтры с характеристикой типа «приподнятый косинус» [3]. Полоса пропускания формирующего фильтра выбирается из соотношения

$$\Delta F_{\text{ФФ}} \approx 0,71 B. \quad (8.3)$$

При использовании амплитудно-фазовой манипуляции с передачей двух боковых полос ширина полосы пропускания ПФ передатчика определяется по формуле [3]

$$\Delta F_{\text{ПФ}} \approx 1,42 B. \quad (8.4)$$

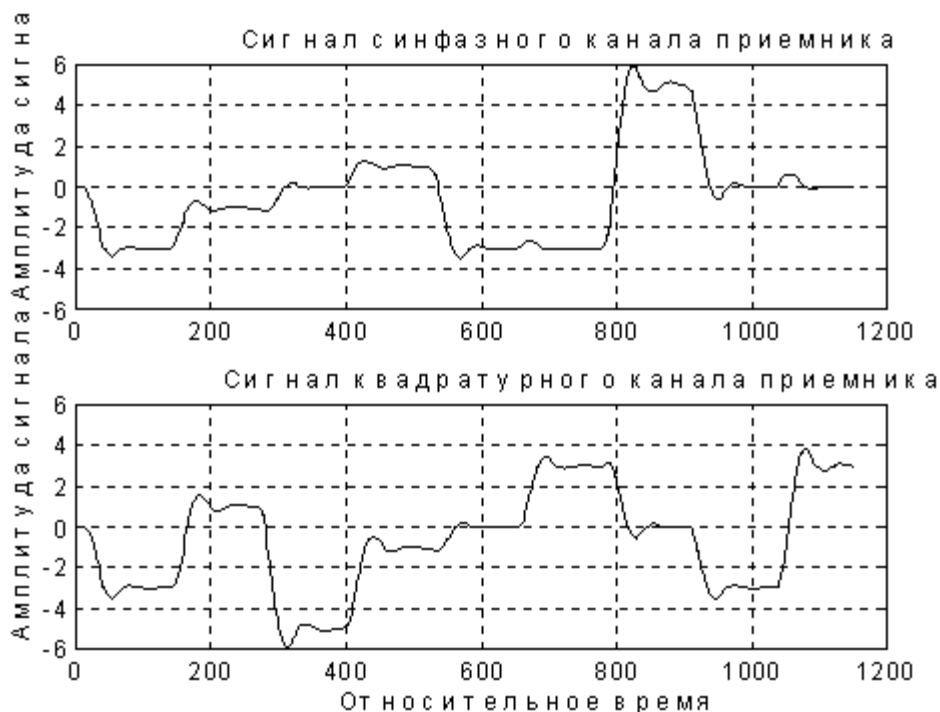


Рисунок 8.3 – Эпюры напряжений на выходах демодуляторов синфазного и квадратурного каналов

Так как несущая частота равна F_H , то полоса пропускания полосового фильтра равна

$$F_H \pm 0,71 B. \quad (8.5)$$

Полосовой фильтр на входе приемника имеет характеристики, аналогичные ПФ передатчика.

ФНЧ амплитудных демодуляторов ФНЧ1 и ФНЧ2 должны выделять огибающую АМ-сигнала. В связи с этим полоса пропускания ФНЧ определяется скоростью манипуляции B и вычисляется по формуле

$$\Delta F_{\text{ФНЧ}} \approx 0,71 B. \quad (8.6)$$

Перед моделированием собственно СПДС целесообразно вначале спроектировать цифровые фильтры, входящие в ее состав. Для этого задаются желаемые частотные характеристики фильтров, вычисляются нормированные частоты, определяющие полосу пропускания, а затем производится вычисление коэффициентов соответствующих цифровых фильтров [2].

В процессе моделирования СПДС следует провести исследование влияния типов и порядка фильтров, формы их характеристик на функционирование системы и помехоустойчивость приема сигналов.

В большинстве случаев целесообразно использовать рекурсивные фильтры относительно низкого порядка (2-5) для минимизации запаздывания, которые вносят фильтры, особенно при старте системы.

При построении модели СПДС необходимо предусмотреть вывод графиков АЧХ и ФЧХ фильтров.

Частоту дискретизации F_d при реализации фильтров следует выбирать таким образом, чтобы при последующем моделировании АМ-сигналов количество отсчетов на один период несущей было не менее 40, т.е. рекомендуется задать частоту дискретизации $F_d = (40-50)F_H$.

При моделировании КАМ-колебаний целесообразно организовывать цикл формирования единичного элемента сигнала. В таком случае очень просто сгенерировать последовательность произвольной длительности. Следует учесть, что минимальное количество единичных элементов должно быть достаточным для подстройки частоты опорных колебаний в приемнике. Его можно определить экспериментальным путем, наблюдая за ошибкой синхронизации, т.е. процесс подстройки частоты должен продолжаться до тех пор, пока ошибка фазирования не достигнет минимума и стабилизируется.

Для генерирования гауссовской флуктуационной помехи целесообразно воспользоваться функцией `randn`, которая формирует вектор помехи в диапазоне значений (0 –1). Для изменения диапазона достаточно умножить вектор на постоянный коэффициент.

В процессе демодуляции АМ-сигналов необходимо формировать на приемной стороне когерентные опорные колебания. Для этого следует реализовать процедуру фазовой автоподстройки частоты (ФАПЧ).

8.4. Порядок проектирования

8.4.1. Повторить материал предшествующей дисциплины «Системы передачи данных», относящийся к принципам построения дискретных каналов, методику проектирования устройств преобразования сигналов [3,4], а также методику проектирования цифровых фильтров [2].

8.4.2. Начертить обобщенную схему системы передачи дискретных сигналов с 16-позиционной КАМ [1,3]. Можно воспользоваться схемой, показанной на рисунке 8.1.

8.4.3. Рассчитать системные параметры СПДС.

8.4.4. Разработать в системе MATLAB программу формирования АФМ-колебаний при поступлении на вход модулятора разнополярных импульсов единичной амплитуды.

8.4.5. Спроектировать цифровой фильтр нижних частот [2], используемый для формирования спектра передаваемых сигналов и исследовать влияние параметров фильтра (полосы пропускания, порядка и типа фильтра) на ширину спектра канальных сигналов и время запаздывания сигнала на выходе фильтра.

8.4.6. Сгенерировать вектор отсчетов флуктуационной помехи и добавить помеху к передаваемому сигналу.

8.4.7. Разработать в системе MATLAB программу формирования на приемной стороне когерентных опорных колебаний с подстройкой фазы под принимаемый АФМ-сигнал.

8.4.8. Разработать в системе MATLAB программу демодуляции сигналов КАМ.

8.4.9. В процессе моделирования на MATLAB отобразить на графиках:

- АЧХ- и ФЧХ-характеристики всех фильтров, используемых в модеме;
- выходной сигнал передатчика модема без помехи;
- сигнал с помехой на входе приемника;
- сигнал опорного генератора приемника для каждого шага итерации процедуры подстройки опорного генератора;
- сигнал на выходе демодулятора модема.

8.4.10. Разработать алгоритм и программные модули на языке СИ для реализации модема на цифровом сигнальном процессоре.

8.4.11. Провести исследование программной реализации процедуры передачи и приема сигналов КАМ на цифровом сигнальном процессоре в системе Code Composer Studio v.2 и на реальном микропроцессорном наборе.

8.4.12. Отобразить на графиках процессы, отмеченные в 4.9.

8.4.13. Провести сравнение результатов моделирования на MATLAB и реализации модема программным способом в ЦСП.

8.4.14. Сделать выводы по работе.

Библиографический список

1. Боккер П. Передача (Техника связи в системах телеобработки данных). Том 1. Основы: Пер с нем./ Под ред. Д.Д.Кловского.-М.: Связь, 1980.- 264 с.
2. Бондарев В.Н. Цифровая обработка сигналов: методы и средства/ В.Н.Бондарев, Г.Трестер, В.Н.Чернега.- Х.: Конус, 2001.-398 с.
3. Чернега В.С. Расчет и проектирование технических средств обмена и передачи информации/ В.С. Чернега, В.А. Василенко, В.Н. Бондарев. - М.: Высш. шк., 1990. - 224 с.
4. Шварцман В.О. Теория передачи дискретной информации/ В.О. Шварцман, Г.А.Емельянов. - М.: Связь, 1979.-424с.

9. ТЕМА 9

СЖАТИЕ РЕЧИ С ПОМОЩЬЮ КАНАЛЬНЫХ И LPC-ВОКОДЕРОВ

9.1. Цель проектирования

9.1.1. Разработка алгоритма функционирования системы сжатия речи с помощью канального и LPC-вокодеров и программная реализация алгоритма в системе MATLAB.

9.1.2. Разработка и исследование программной реализации системы на основе ЦПОС.

9.2. Задание на проектирование

Разработать и реализовать на ЦПОС TMS320C6711(5402) систему сжатия речи (рекомендуется реализовать канальный вокодер или вокодер с линейным предсказанием), осуществляющую сжатие и восстановление речевого сигнала в реальном времени до уровня не более 4800 бит/с. Считать, что исходный речевой сигнал представлен в виде последовательности 16-разрядных отсчетов с частотой дискретизации 8КГц. Необходимо обеспечить удовлетворительное качество восстановленного речевого сигнала (требуется обеспечить разборчивость речи, при этом узнаваемость диктора не обязательна).

9.3. Методические рекомендации по проектированию

9.3.1. Системы сжатия речи

Существует несколько подходов к сжатию речевых сигналов:

- кодирование формы волны речевого сигнала;
- кодирование параметров речевого тракта человека и источника возбуждения;
- кодирование символьной информации (фонем);
- кодирование лингвистической информации (слов, фраз и т.п.).

9.3.1.1. Непосредственное кодирование формы речевого сигнала

Исходный речевой сигнал представляет собой акустическую волну (волна давления в воздухе), которую можно преобразовать в электрический сигнал с помощью микрофона. Будем считать, что спектр речевого сигнала лежит в диапазоне от 100 до 4000 гц. Динамический диапазон изменения амплитуды, достаточный для описания речевых сигналов, составляет 12 двоичных разрядов.

Первым шагом, обеспечивающим сжатие речевого сигнала, является попытка обеспечения равномерной относительной точности измерения значения амплитуды сигнала. Для этого 14-12-ти разрядный динамический диапазон амплитуды разбивают на 8 логарифмических поддиапазонов, в каждом из которых

значение амплитуды кодируют 5 разрядами и, таким образом, достигают сокращения информации до 64000 бит/с (кодирование по μ - и А- законам в соответствии со стандартом ITU -G.711). Следующим шагом является адаптивная дифференциальная импульсно-кодовая модуляция (АДИКМ), (например, в соответствии со стандартами G.721 или G.726 8-40000 бит/с), с помощью которой осуществляют кодирование (аппроксимацию) степени приращения амплитуды сигнала во времени. Таким путем удастся достичь степени сжатия речевого сигнала порядка 32000-16000 бит/сек., причем приемлемое (коммерческое) качество речи (по критерию отношения: полезный_сигнал/шум) обеспечивается до 24000 бит/сек. При более низких скоростях кодирования сохраняется разборчивость речи, но характерны сильные нелинейные и частотные искажения сигнала и ухудшение отношения сигнал/шум. Дальнейшее уменьшение информационной емкости сигнала с помощью данного подхода считается неэффективным.

9.3.1.2. Параметрическое кодирование

Низкоскоростное кодирование складывается из двух основных процессов:

- параметрическое представление речевого сигнала минимальным набором параметров, характеризующих источник возбуждения и акустический фильтр, определяющий передаточную функцию голосового тракта;
- дискретизация речевых параметров для их передачи по каналу связи при использовании минимальной емкости канала.

Для параметрического описания речи обычно используется подход, основанный на вычислении параметров, описывающих передаточную функцию речевого тракта человека и функцию возбуждения. Такими параметрами могут являться: осредненные значения энергии речевого сигнала, разбитого на ряд частотных полос, или коэффициенты линейного предсказания (или, связанные с ними, коэффициенты отражения). Обычно для кодирования речи используются 8-10 параметров (один из вышеперечисленных наборов), вычисляемых на интервалах порядка 5-30 мс (так как на таком интервале речь может считаться стационарным процессом), кроме того, вычисляется параметр, характеризующий изменение амплитуды либо мощности сигнала, период основного тона речи, а также признак типа тон/шум/пауза, характеризующий способ возбуждения речевого сигнала.

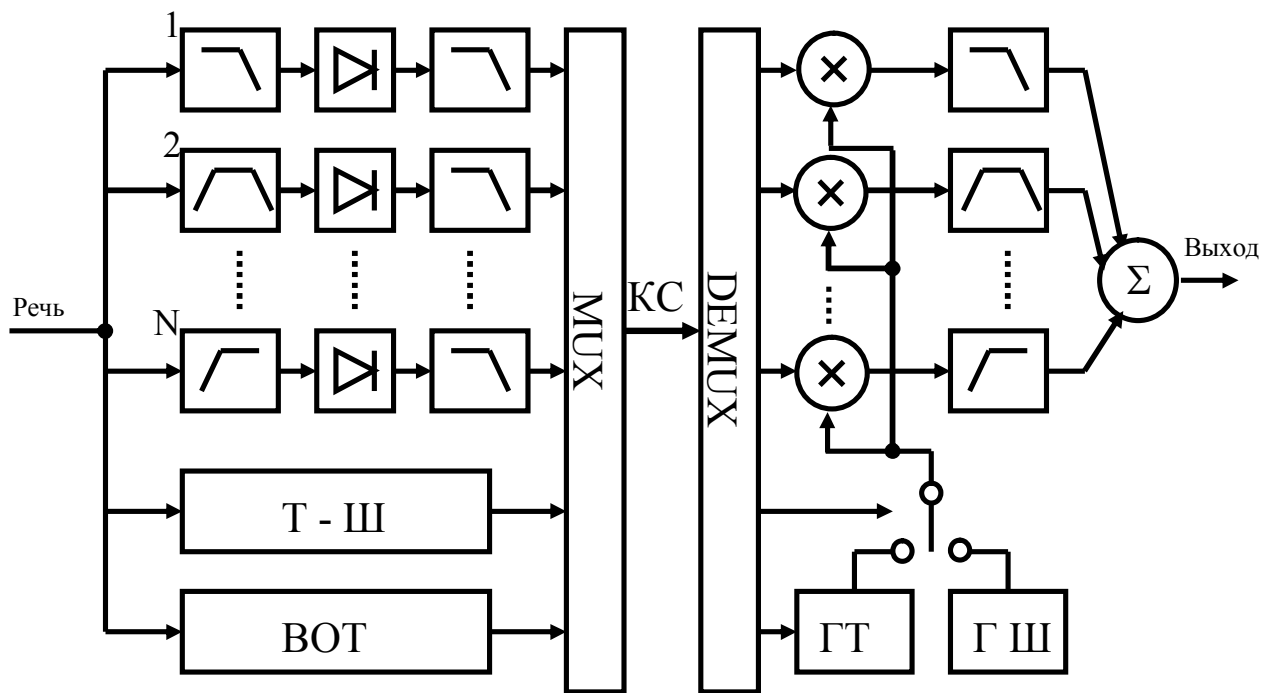
Полученный набор параметров, оптимизированный по критерию точности и минимальной разрядности представления, передается в цифровом виде по каналу связи в реальном времени, а на приемном конце осуществляется синтез речевого сигнала по перечисленным параметрам. Таким путем удастся снизить информационную емкость речевого сигнала до уровня 16000 - 1200 бит/сек, причем с сохранением разборчивости и индивидуальных особенностей речи говорящего.

9.3.1.3. Другие способы кодирования

Следующим шагом в направлении дальнейшего увеличения компрессии является создание фонемного вокодера. Как известно, минимальной слогообразовательной (и словобразовательной) единицей речи является фонема. Поэтому создание устойчивого метода распознавания фонем позволит снизить скорость кодирования речевой информации до 100 бит/сек, что соответствует информационной скорости текста. Следует отметить, что на приемной стороне речь будет восстановлена синтезатором речи по фонемному тексту, при этом информация об индивидуальности диктора будет утрачена.

9.3.2. Канальные вокодеры

Канальный вокодер представляет собой совокупность двух основных частей - анализирующей (передающая сторона) и синтезирующей (приемная), которые содержат идентичные наборы (гребенки) полосовых фильтров, перекрывающих определенный частотный интервал. Структура канального вокодера



представлена на рисунке 9.1:

Рисунок 9.1 - Структурная схема канального вокодера

Фильтры блока анализа обеспечивают тональное разделение спектра сигнала. Для перекрытия всей полосы звуковых частот, наряду с полосовыми, в гребенке используют фильтры НЧ и ВЧ (в самых низкочастотном и высокочастотном каналах). Типовая амплитудно-частотная характеристика гребенки

фильтров, в случае равномерного разделения каналов может иметь следующий вид:

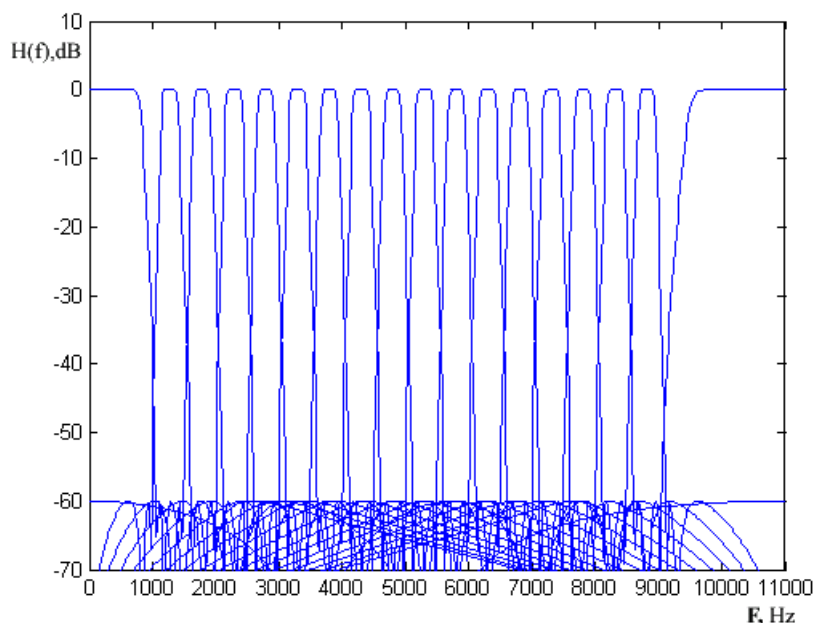


Рисунок 9.2 - АЧХ гребенки фильтров

Детектор и фильтр НЧ в каждом канале выделяют огибающую сигнала данного канала, и каждая из них характеризует энергию речевого спектра в соответствующей полосе частот для фрагмента речи (длина фрагмента 5-30 мс). Для более компактной передачи выходы каждого из каналов могут логарифмироваться и кодироваться с помощью дельта импульсной кодовой модуляции (ДИКМ)[1].

От числа частотных полос зависит разборчивость синтезированной речи. Хорошие результаты получаются при числе каналов 15-16 (полоса 100 Гц... 4 кГц). Для улучшения качества восстановленной речи при выборе центральных частот и ширины полос пропускания канальных фильтров целесообразно использовать критические полосы слуха[1].

При выборе типа фильтров необходимо иметь в виду, что спектральные составляющие сигнала вблизи центральной частоты резонансных фильтров подвергаются существенным фазовым сдвигам, а это приводит к изменению тембра, даже если амплитудные соотношения сохранены. Причем при увеличении порядка фильтров фазовый сдвиг увеличивается, являясь причиной характерной для вокодера неестественности речи. С другой стороны, при недостаточной крутизне спада АЧХ фильтров появляется “смазанность” синтезированной речи. Практика показывает, что оптимальный результат соответствует АЧХ фильтров при крутизне их спада примерно 36 дБ на октаву.

Частоту среза канального фильтра НЧ выбирают в десять раз меньшей центральной частоты полосового фильтра канала, однако возможно использование одинаковых фильтров низких частот с частотой среза около 25 Гц, что несколько снижает качество восстановленного речевого сигнала.

Кроме оценки энергии сигнала в частотных полосах, в блоке анализа также производится оценка характера речевого фрагмента. В детекторе тон/шум(Т-Ш) оценивается: вокализованным или фрикативным является данный фрагмент речевого сигнала. Такой детектор может быть реализован через определение числа переходов через ноль речевого сигнала (для вокализованных фрагментов в отличие от фрикативных число переходов через ноль невелико) или оценкой энергии сигнала в полосах до 800 Гц и более 2КГц (структурная схема такого детектора представлена на рисунке 9.3).

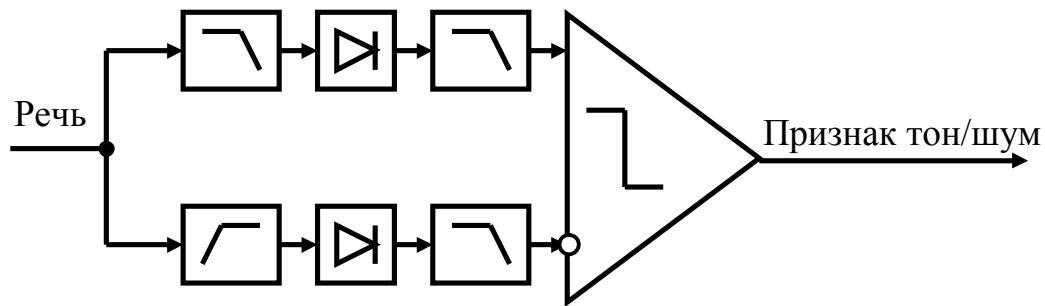


Рисунок 9.3 – Структура детектора тон-шум

Также в блоке анализа происходит определение частоты основного тона для вокализованных фрагментов речи. Для этого в структуру блока анализа вводится выделитель основного тона (ВОТ). Известно, что значение ОТ для разных голосов может изменяться почти в 10 раз - от 2 до 18 мс. Это обстоятельство создает немало трудностей при оценке ОТ, так как слух очень чувствителен к его искажениям. На сегодняшний день известно большое количество алгоритмов оценки ОТ, оперирующих как непосредственно с временным представлением речевого сигнала, так и со спектром речи. Например, метод Голда-Рабинера[2,5], автокорреляционный метод[1,2], SIFT(simplified inverse filter tracking) алгоритм [2].

Таким образом, для каждого фрагмента речевого сигнала на приемную сторону передается признак тон-шум(1 бит), частота основного тона, в случае вокализованного фрагмента(5-6 бит), значения энергии сигнала по каналам (при использовании ДИКМ менее 5 бит на канал).

На приемной стороне, в случае если принят вокализованный фрагмент, в качестве источника возбуждения гребенки фильтров выступает генератор периодических сигналов(ГТ). В качестве сигнала возбуждения может быть выбрана последовательность специального вида импульсов[1] с периодом равным частоте основного тона данного фрагмента (также принимаемой из канала связи). В случае невокализованного фрагмента, в качестве сигнала возбуждения используется белый шум(ГШ), который может быть получен с помощью генератора случайных чисел.

Сигнал возбуждения умножается на значения энергии для каждого из каналов, после чего поступает на фильтр соответствующего канала (гребенка фильтров блока синтеза полностью аналогична гребенке блока анализа).

Для получения выходного синтезированного речевого сигнала выходы всех каналов блока синтеза суммируются.

9.3.3. Синтез цифровых рекурсивных фильтров с использованием пакета MATLAB

Решение задачи синтеза цифровых рекурсивных фильтров (РФ) сводится к нахождению коэффициентов $\mathbf{a}_k$ и $\mathbf{b}_k$ разностного уравнения:

$$y(nT) = \sum_{k=0}^{N-1} b_k x(nT - kT) - \sum_{k=1}^{M-1} a_k y(nT - kT). \quad (9.1)$$

Известны прямые и косвенные методы синтеза рекурсивных цифровых фильтров [4,1]. Прямые методы основаны на непосредственном определении параметров цифровых РФ по заданным временным или частотным характеристикам. Косвенные методы синтеза цифровых РФ основаны на дискретизации аналогового фильтра-прототипа, удовлетворяющего заданным требованиям. Существуют различные способы аппроксимации аналоговых фильтров [6,1]. На практике широко применяется аппроксимация АЧХ аналоговых фильтров с помощью полиномов Баттерворта, Чебышева, Бесселя, Кауэра.

При проектировании фильтров обычно задают следующие требования к их амплитудно-частотным характеристикам (АЧХ):

- границы полос пропускания и задерживания (соответственно ω_1 и ω_2);
- затухание в полосе задерживания (H_2);
- коэффициент передачи в полосе пропускания (H_1 - обычно он равен 1);
- допуск на отклонение реальной АЧХ от желаемой в полосе пропускания (Δ_1).

В пакете MATLAB для синтеза РФ существует ряд функций:

- **butter** – аппроксимация фильтра Баттерворта;
- **cheby1** – аппроксимация фильтра Чебышева первого типа;
- **cheby2** – аппроксимация фильтра Чебышева второго типа (инверсный фильтр Чебышева);
- **ellip** – аппроксимация эллиптического фильтра (фильтр Кауэра).

Например, вызов функции **butter** в пакете MATLAB может иметь следующий синтаксис:

$[b,a] = \text{butter}(n,Wn);$

$[b,a] = \text{butter}(n,Wn,'ftype').$

Для расчета порядка фильтра Баттерворта, обеспечивающего требуемые характеристики (H_1 , H_2 , Δ_1) в пакете MATLAB используется функция **buttord**, имеющая следующий синтаксис:

$[n,Wn] = \text{buttord}(Wp,Ws,Rp,Rs).$

Пример. С помощью пакета MATLAB найти коэффициенты и построить характеристики полосового фильтра Баттерворта с полосой пропускания от 60 до 200 Гц (частота дискретизации 1 кГц), с пульсациями в полосе пропускания не более 3 дБ и затуханием в полосе задерживания не менее 40 дБ.

Ниже представлен текст программы, а на рисунке 9.4 показаны характеристики полученного фильтра.

```
Wp = [60 200]/500; %расчет нормированных граничных частот полосы пропускания
Ws = [50 250]/500; %расчет нормированных граничных частот полосы задерживания
Rp = 3; Rs = 40; % задаем величину пульсаций и затухание
[n,Wn] = buttord(Wp,Ws,Rp,Rs)% определяем порядок фильтра и частоты среза
>>n = 16
>>Wn = [0.1198 0.4005]
[b,a] = butter(n,Wn); %находим коэффициенты фильтра
freqz(b,a,128,1000)% строим характеристики фильтра
title('АЧХ и ФЧХ полосового фильтра Баттерворта')
```

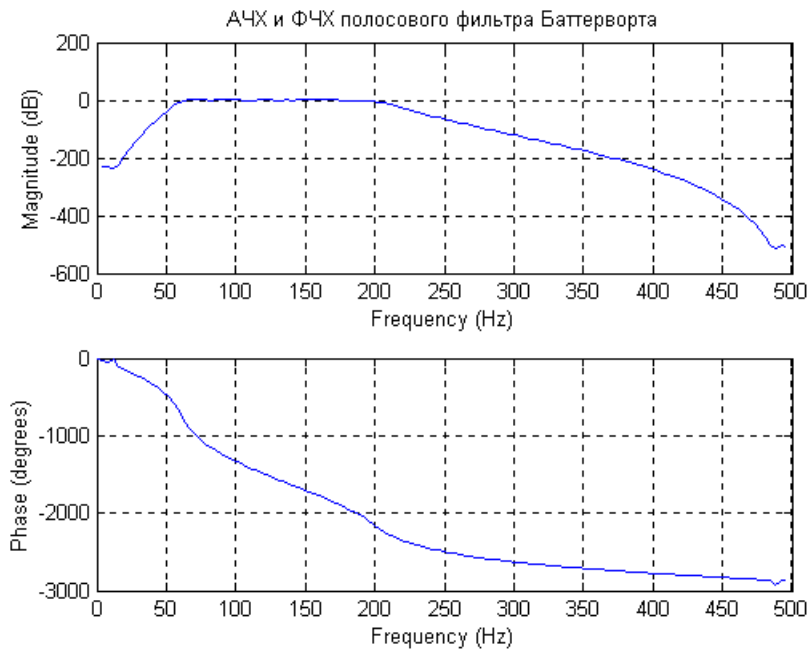


Рисунок 9.4 - АЧХ и ФЧХ полосового фильтра Баттерворта

9.3.4. Вокодеры с линейным предсказанием

Основная идея линейного предсказания (ЛП) [1], как известно, состоит в том, что любой отсчет речевого сигнала $x[n]$ может быть аппроксимирован с помощью линейной комбинации p предыдущих отсчетов речи, где p обычно равно 10-15. Таким образом, предсказанный отсчет может быть получен как:

$$\tilde{x}[n] = -\sum_{k=1}^p \hat{a}[k]x[n-k]. \quad (9.2)$$

Коэффициенты $\hat{a}[k]$ называют коэффициентами линейного предсказания (ЛПК-коэффициентами). Фактически коэффициенты предсказания задают спектральную модель голосового аппарата человека, и чем выше порядок ЛП (P в формуле 9.1), тем точнее модель.

Ошибка предсказания будет равна:

$$e[n] = x[n] - \tilde{x}[n] = x[n] - \sum_{k=1}^P \hat{a}[k] x[n-k]. \quad (9.3)$$

Следует отметить, что ошибка предсказания будет похожа на функцию возбуждения голосового тракта $u[n]$. Поэтому по $e[n]$ можно определить период основного тона речи.

Существуют различные методы определения коэффициентов линейного предсказания. Один из них основан на минимизации среднеквадратического значения ошибки $e[n]$.

В [1] (рисунок 4.13,а) приведена схема сжатия речи без обратной связи, основанная на анализе по методу линейного предсказания. Здесь речевой сигнал $s[n]$ разбивается на сегменты длительностью 20-30 мс. На каждом из сегментов с помощью устройства оценивания (УО) определяются коэффициенты линейного инверсного фильтра анализа $\Phi 1$ 10-го порядка (ЛПК коэффициенты). Кроме этого, на этапе сжатия с помощью выделителя основного тона (ВОТ) и анализатора тон-шум (Т-Ш) по полученной после фильтрации ошибке предсказания $d[n]$ определяются соответствующие параметры функции возбуждения. В кодере выполняется кодирование коэффициентов фильтра и параметров функции возбуждения, которые затем передаются по каналу связи.

В восстанавливающем устройстве [1] (рисунок 4.13,б) сначала происходит декодирование коэффициентов фильтра и параметров функции возбуждения, а затем выполняется синтез речевого сигнала. Для этого в зависимости от значения признака тон-шум (ТШ) на вход фильтра-синтеза $\Phi 2$ подается сигнал либо с выхода генератора тона (ГТ), либо с выхода генератора шума (ГШ). Для кодирования периода основного тона используют 6 бит, для коэффициента усиления - 5 бит, для признака тон/шум - 1 бит, для коэффициентов линейного предсказания - 8-10 бит. С учетом того, что для каждого сегмента речи оценивается десять коэффициентов предсказания, получим 97-117 бит на один сегмент. Скорость передачи при длительности сегмента 30 мс составит примерно 3000 бит/с.

9.4. Порядок проектирования

9.4.1. Проектирование канального вокодера

9.4.1.1. С использованием пакета MATLAB произвести расчет гребенки полосовых фильтров блоков анализа и синтеза, состоящей из фильтра нижних

частот, 12-18 полосовых фильтров и фильтра верхних частот. Рекомендуется использование рекурсивных фильтров, с порядком, обеспечивающим крутизну в переходной полосе не менее 36 дБ на октаву. Построить АЧХ рассчитанных фильтров, исследовать разработанную гребенку фильтров и построить ее АЧХ.

9.4.1.2. С использованием пакета MATLAB произвести расчет гребенки фильтров нижних частот блока анализа и построить АЧХ рассчитанных фильтров.

9.4.1.3. Разработать с использованием пакета MATLAB детектор тон-шум, реализующий определение характера фрагмента речи (длина фрагмента 5-30 мс). Сравнить различные алгоритмы реализации такого детектора и выбрать наиболее быстродействующий и эффективный.

9.4.1.4. Разработать с использованием пакета MATLAB выделитель частоты основного тона (BOT) фрагмента речи. Сравнить различные алгоритмы реализации BOT и выбрать наиболее быстродействующий и эффективный.

9.4.1.5. С помощью пакета MATLAB реализовать блок анализа канального вокодера, осуществляющий сжатие wav-файлов (частота дискретизации 8000 Гц, 16 бит/отсчет), содержащих речевые сообщения, и сохраняющих сжатую речь во внешнем файле.

9.4.1.6. С помощью пакета MATLAB реализовать блок синтеза канального вокодера, восстанавливающий сжатую речь.

9.4.1.7. Оптимизировать разработанную в пунктах 9.4.1.5 и 9.4.1.6 систему по критериям быстродействие, качество (разборчивость речи). Система должна осуществлять сжатие речевого сигнала до уровня не менее 4800 бит/с.

9.4.1.8. С помощью Code Composer Studio v2 реализовать канальный вокодер на языке C для сжатия/восстановления речевого сигнала в реальном времени на ЦПОС TMS320C5402 или TMS320C7611.

9.4.1.9. Отладить разработанную в пункте 9.4.1.8 систему и оптимизировать ее по критерию качества восстанавливаемой речи.

9.4.2. Проектирование вокодера с линейным предсказанием

9.4.2.1. Реализовать с помощью пакета MATLAB устройство оценивания (УО), реализующего определение коэффициентов линейного инверсного фильтра (ЛПК коэффициенты). Убедиться в правильности работы УО, сравнив получаемые коэффициенты с коэффициентами, рассчитываемыми с помощью функции **lpc** пакета MATLAB.

9.4.2.2. Разработать с использованием пакета MATLAB детектор тон-шум, реализующий определение характера фрагмента речи (длина фрагмента 5-30 мс) по выходному сигналу инверсного фильтра.

9.4.2.3. Разработать с использованием пакета MATLAB выделитель частоты основного тона (BOT), реализующий определение частоты основного тона фрагмента речи по выходному сигналу инверсного фильтра.

9.4.2.4.Реализовать с помощью пакета MATLAB кодер с линейным предсказанием (рис. 3.4,а), осуществляющий сжатие wav-файлов (частота дискретизации 8000Гц, 16 бит/отсчет), содержащих речевые сообщения, и сохраняющих сжатую речь во внешнем файле.

9.4.2.5.В соответствии со структурной схемой приведенной на рис. 3.4,б реализовать декодер сжатого речевого сигнала.

9.4.2.6.Оптимизировать разработанную в пунктах 9.4.2.4 и 9.4.2.5 систему по критериям быстродействие, качество (разборчивость речи). Система должна осуществлять сжатие речевого сигнала до уровня не менее 2400 бит/с.

9.4.2.7.С помощью Code Composer Studio v2 реализовать вокодер с линейным предсказанием на языке С для сжатия/восстановления речевого сигнала в реальном времени на ЦПОС TMS320C5402 или TMS320C7611.

9.4.2.8.Отладить разработанную в пункте 9.4.1.7 систему и оптимизировать ее по критерию качества восстанавливаемой речи.

Библиографический список

1. Бондарев В.Н. Цифровая обработка сигналов: методы и средства / В.Н. Бондарев, Г. Трестер, В.С. Чернега.–Севастополь: Изд-во СевГТУ, 1999. – 388с.
2. Рабинер Л.Р. Цифровая обработка речевых сигналов / Л.Р. Рабинер, Р.В. Шафер. – М.: Радио и связь, 1981.-495с.
3. Применение цифровой обработки сигналов / Под ред. Э. Оппенгейма. - М.:Мир, 1984. -550с.
4. Гольденберг Л.М. Цифровая обработка сигналов: Справочник / Л.М. Гольденберг, Б.Д. Матюшкин, М.Н. Поляк.- М.: Радио и связь, 1985. - 312с.
5. Мошиц Г. Проектирование активных фильтров: Пер. с англ./ Г.Мошиц, П.Хорн.- М.:Мир,1984.-320с.
6. Вокодерная телефония. Методы и проблемы. /Под ред. А.А. Пирогова. – М.: Связь, 1974
7. Рэйдер Ч. Цифровая обработка сигналов/ Ч. Рэйдер. - М.: Советское радио, 1973.
8. Оппенгейм А.В. Цифровая обработка сигналов/ А.В. Оппенгейм, Р.В. Шафер.- М.: Связь 1979.
9. Сапожков М.А. Вокодерная связь/М.А. Сапожков, В.Г. Михайлов. – М.: Радио и связь, 1983.

10. ТЕМА 10 СЖАТИЕ ИЗОБРАЖЕНИЙ

10.1. Цель проектирования

Закрепление навыков практического решения задач цифровой обработки сигналов в реальном времени с использованием ЦПОС TMS320C6711(5402).

10.2. Задание на проектирование

Разработать и реализовать на ЦПОС TMS320C6711(5402) систему сжатия полноцветных фотореалистичных изображений на основе дискретного косинусного преобразования(ДКП). Считать, что исходное изображение представлено в виде матриц RGB(три матрицы интенсивностей: красного, зеленого и синего цветов), содержащих однобайтные значения. Необходимо обеспечить изменяющийся коэффициент сжатия изображения в диапазоне 2-20 раз. Рекомендуется использование стандарта сжатия изображений - JPEG.

10.3. Методические рекомендации по проектированию

10.3.1. Методы сжатия изображений

Существует две основных группы методов сжатия изображений – сжатие без потерь и сжатие с потерями.

К методам сжатия изображений без потерь относятся:

- групповое кодирование;
- кодирование методом LZW;
- кодирование методом Хаффмана;
- кодирование с помощью алгоритма JBIG;
- кодирование с помощью алгоритма Lossless JPEG.

Используя один из алгоритмов сжатия без потерь, можно обеспечить коэффициент сжатия изображения примерно в два раза. В то же время алгоритмы сжатия с потерями позволяют достигать коэффициента сжатия: 10-200 раз.

К методам сжатия с потерями относятся:

- алгоритмы использующие ДКП(в том числе JPEG);
- методы фрактального сжатия;
- сжатие с использованием вейвлет(wavelet) преобразований.

10.3.1.1. Алгоритмы сжатия без потерь

10.3.1.1.1. Групповое кодирование

Групповое кодирование - Run Length Encoding (RLE) - один из самых старых и самых простых алгоритмов архивации графики. Изображение в нем (как и в нескольких алгоритмах, описанных далее) вытягивается в цепочку байт по строкам раstra. Само сжатие в RLE происходит за счет того, что в исходном изображении встречаются цепочки одинаковых байт. Замена их на пары "счетчик, значение" уменьшает избыточность данных. Лучший, средний и худший коэффициенты сжатия - $1/32$, $1/2$, $2/1$.

К положительным сторонам алгоритма можно отнести то, что он не требует дополнительной памяти при работе, и не требует больших вычислительных затрат. Ориентирован алгоритм на изображения с небольшим количеством цветов: деловую и научную графику. Применяется в форматах PCX, TIFF, BMP.

10.3.1.1.2. Сжатие методом LZW

Название алгоритм получил по первым буквам фамилий его разработчиков - Lempel, Ziv и Welch. Сжатие в нем, в отличие от RLE, осуществляется уже за счет одинаковых цепочек байт. Существует довольно большое семейство LZ-подобных алгоритмов, различающихся, например, методом поиска повторяющихся цепочек. Коэффициенты сжатия: $1/1000$, $1/4$, $7/5$. Коэффициент $1/1000$ достигается только на одноцветных изображениях размером больше 4 Мб. Этот метод сжатия изображений реализован в форматах GIF, TIFF и TGA.

10.3.1.1.3. Алгоритм Хаффмана

В этом алгоритме сжатия используется частота появления одинаковых байт в изображении. Кодер сопоставляет символам входного потока, которые встречаются большее число раз, цепочку бит меньшей длины. И напротив - встречающимся редко - цепочку большей длины. Для сбора статистики требуется два прохода по изображению. Коэффициенты сжатия: $1/8$, $2/3$, 1. Требуется записи в файл таблицы соответствия кодируемых символов и кодирующих цепочек. На практике используются его разновидности. Так, в некоторых случаях возможно либо использовать постоянную кодовую таблицу (в этом случае изображение кодируется в два "прохода" – при первом проходе определяется статистика изображения и строится кодовая таблица, а при втором – изображение кодируется используя построенную), либо строить ее "адаптивно", т.е. в процессе сжатия/восстановления (этот прием избавляет от двух проходов по изображению и необходимости хранения таблицы вместе со сжатым изображением).

Близкая модификация алгоритма используется при сжатии черно-белых изображений. Последовательности подряд идущих черных и белых точек заменяются числом, равным их количеству с признаком цвета. А этот ряд уже, в свою очередь, сжимается по Хаффману с фиксированной таблицей. Алгоритм реализован в формате TIFF

10.3.1.1.3. JBIG

Алгоритм разработан группой экспертов ISO (Joint Bi-level Experts Group) специально для сжатия однобитных черно-белых изображений. Например, факсов или отсканированных документов. В принципе может применяться и к 2-х, и к 4-х битовым картинкам. При этом алгоритм разбивает их на отдельные битовые плоскости. JBIG позволяет управлять такими параметрами, как порядок разбиения изображения на битовые плоскости, ширина полос в изображении, уровни масштабирования. Последняя возможность позволяет легко ориентироваться в базе больших по размерам изображений, просматривая сначала их уменьшенные копии. Характерной особенностью JBIG является резкое снижение степени сжатия при повышении уровня шумов входной картинки.

10.3.1.1.4. Lossless JPEG

Этот алгоритм, разработан группой экспертов в области фотографии (Joint Photographic Expert Group). В отличие от JBIG, Lossless JPEG ориентирован на полноцветные 24-битные безпалитровые изображения. Он представляет собой специальную реализацию алгоритма JPEG, который будет рассмотрен ниже. Коэффициенты сжатия: 1/20, 1/2, 1. Lossless JPEG рекомендуется применять в тех приложениях, где необходимо побитовое соответствие исходного и восстановленного изображений.

10.3.1.2. Алгоритмы сжатия с потерями

10.3.1.2.1. Алгоритм JPEG

Алгоритм JPEG – используется для сжатия как полутоновых многоградиентных, так и для полноцветных фотореалистических изображений. Оперирует алгоритм фрагментами изображения 8x8, на которых яркость и цвет меняются сравнительно плавно. Вследствие этого, при выполнении над матрицей такой области двумерного ДКП значимыми оказываются только первые коэффициенты (левый верхний угол). Таким образом, сжатие в JPEG осуществляется за счет малой величины значений амплитуд высоких частот в реальных изображениях.

Коэффициент архивации в JPEG может изменяться в пределах от 2 до 200 раз. Как и у любого другого алгоритма сжатия с потерями, у JPEG свои особенности. Наиболее известны "эффект Гиббса" и дробление изображения на квадраты 8x8. Первый проявляется около резких границ предметов, образуя своеобраз-

разный "ореол". Он хорошо заметен, если, допустим, поверх фотографии сделать надпись цветом, сильно отличающимся от фона. Разбиение на квадраты происходит, когда задается слишком большой коэффициент сжатия для конкретной картинке.

JPEG реализован в форматах JPG и TIFF.

10.3.1.2.2. Фрактальное сжатие

Коэффициент сжатия у фрактальных алгоритмов варьируется в пределах 2-2000. Причем большие коэффициенты достигаются на реальных изображениях, что, вообще говоря, нетипично для предшествующих алгоритмов. Кроме того, при разархивации изображение можно масштабировать. Уникальная особенность этого алгоритма заключается в том, что увеличенное изображение не дробится на квадраты. Во фрактальном сжатии используется принципиально новая идея - не близость цветов в локальной области, а подобие разных по размеру областей изображения.

Алгоритм ориентирован на полноцветные изображения и изображения в градациях серого цвета. Его особенностью является потребность в колоссальных вычислительных мощностях при архивации. При этом распаковка требует меньше вычислений, чем у JPEG, т.е. этот алгоритм является несимметричным. Причем, если у предыдущих алгоритмов коэффициент симметричности (отношение времени архивации ко времени разархивации) не превышал 3, то у фрактального алгоритма он колеблется в пределах 1000-10000. Фрактальное сжатие реализовано в формате FIF.

10.3.1.2.3. Сжатие с использованием wavelet

Ориентирован алгоритм на цветные и черно-белые изображения с плавными переходами. Идеален для картинок типа рентгеновских снимков. Коэффициент сжатия задается и варьируется в пределах 5-20 раз. При попытке задать больший коэффициент, на резких границах, особенно проходящих по диагонали, проявляется "лестничный эффект" - ступеньки разной яркости, размером в несколько пикселей.

10.3.2. Описание алгоритма JPEG

Сжатие изображения в соответствии с алгоритмом JPEG выполняется в несколько этапов. Структурная схема кодера JPEG приведена на рисунке 10.1:

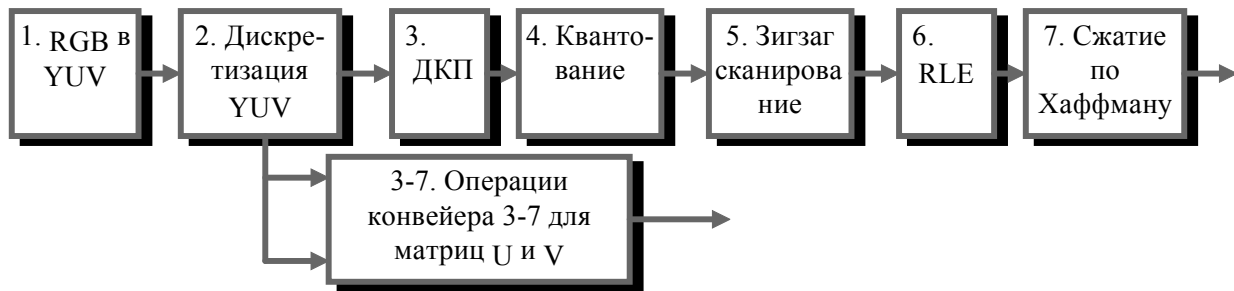


Рисунок 10.1 – Структурная схема кодера JPEG

Рассмотрим подробнее этапы сжатия.

10.3.2.1. Переводим изображение из цветового пространства RGB, с компонентами, отвечающими за красную (Red), зеленую (Green) и синюю (Blue) составляющие цвета точки, в цветовое пространство YCrCb (иногда называют YUV).

В нем Y — яркостная составляющая, а Cr, Cb — компоненты, отвечающие за цвет (хроматический красный и хроматический синий). За счет того, что человеческий глаз менее чувствителен к цвету, чем к яркости, появляется возможность архивировать массивы для Cr и Cb компонент с большими потерями и, соответственно, большими коэффициентами сжатия. Подобное преобразование уже давно используется в телевидении. На сигналы, отвечающие за цвет, там выделяется более узкая полоса частот.

Упрощенно перевод из цветового пространства RGB в цветовое пространство YCrCb можно представить так:

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.5 & -0.4187 & -0.0813 \\ 0.1687 & -0.3313 & 0.5 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix} \quad (10.1)$$

Обратное преобразование осуществляется умножением вектора YUV на обратную матрицу:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.402 \\ 1 & -0.34414 & -0.71414 \\ 1 & 1.772 & 0 \end{bmatrix} * \left(\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} - \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix} \right) \quad (10.2)$$

10.3.2.2. Разбиваем исходное изображение на матрицы 8x8. Формируем из каждой три рабочие матрицы ДКП — по 8 бит отдельно для каждой компоненты. При больших коэффициентах сжатия, этот шаг может выполняться чуть сложнее. Изображение делится по компоненте Y — как и в первом случае, а для компонент Cr и Cb матрицы набираются через строку и через столбец. Т.е. из исходной матрицы размером 16x16 получается только одна рабочая матрица ДКП. При этом, как нетрудно заметить, мы теряем 3/4 полезной информации о цветовых составляющих изображения и получаем сразу сжатие в два раза. Мы

можем поступать так благодаря работе в пространстве YCrCb. На результирующем RGB изображении, как показала практика, это сказывается не сильно.

10.3.2.3.Применяем ДКП к каждой рабочей матрице. Дискретное косинусное преобразование преобразует матрицу пикселей размером 8x8 в матрицу частотных коэффициентов(трансформант) соответствующего размера. Формулы прямого и обратного дискретного косинусного преобразования представлены ниже:

$$F(i, j) = \frac{1}{\sqrt{2N}} C(i)C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} I(x, y) \cos \frac{(2x+1)i\pi}{2N} \cos \frac{(2y+1)j\pi}{2N}, \quad (10.3)$$

$$I(x, y) = \frac{1}{\sqrt{2N}} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} C(i)C(j)F(i, j) \cos \frac{(2x+1)i\pi}{2N} \cos \frac{(2y+1)j\pi}{2N}, \quad (10.4)$$

где $C(i)$ и $C(j) = \frac{1}{\sqrt{2}}$ для $i, j=0$ и $C(i), C(j)=1$ для $i, j>0$;

$I(x, y)$ - значение пикселя;

$F(i, j)$ - значения трансформант.

В получившейся матрице коэффициентов $F(i, j)$ низкочастотные компоненты сгруппированы в левом верхнем углу, а высокочастотные - в правом нижнем. Следует отметить, что высокочастотные компоненты не сильно влияют на качество изображения, поэтому в дальнейшем некоторой их частью можно пренебречь, не внося серьезных искажений в изображение.

Как видно из формулы (10.3), по определению дискретного косинусного преобразования, для его реализации требуется два вложенных цикла, и тело циклов будет выполняться $N \times N$ раз для каждого элемента матрицы дискретного косинусного преобразования. Значительно более эффективный вариант вычисления коэффициентов дискретного косинусного преобразования может быть реализован через перемножение матриц. При таком подходе формула дискретного косинусного преобразования может быть записана в следующем виде:

$$F = K\Pi * I * K\Pi^T,$$

где $K\Pi$ - матрица косинусного преобразования размером $N \times N$, элементы которой определяются по формуле:

$$C(i, j) = \begin{cases} \frac{1}{\sqrt{N}}, & i = 0 \\ \sqrt{\frac{2}{N}} \cos \left[\frac{(2j+1)i\pi}{2N} \right], & i > 0, \end{cases}$$

$K\Pi^T$ - транспонированная матрица $K\Pi$.

При перемножении матриц для вычисления одного элемента результирующей матрицы необходимо N умножений и N сложений, при вычислении матрицы дискретного косинусного преобразования - $2 \times N$ соответственно. По сравнению с $N \times N$ это заметное повышение производительности.

Следует также отметить, что так как дискретное косинусное преобразование является разновидностью преобразования Фурье, то все методы ускорения преобразования Фурье могут быть применены и в данном случае[1].

10.3.2.4. Производим квантование. В принципе это просто деление рабочей матрицы на матрицу квантования поэлементно. Для каждой компоненты (Y , U и V), в общем случае, задается своя матрица квантования (МК) $Q[u,v]$:

$$Yq[u, v] = \text{IntegerRound} \left(\frac{Y[u, v]}{q[u, v]} \right).$$

На этом шаге осуществляется управление степенью сжатия, и происходят потери. Понятно, что, задавая МК с большими коэффициентами, мы получим больше нулей и, следовательно, большую степень сжатия.

В стандарт JPEG включены рекомендованные МК, построенные опытным путем[4].

Для увеличения или уменьшения коэффициента сжатия исходная матрица Q умножается на некоторое число γ .

С квантованием связаны и специфические эффекты алгоритма. При больших значениях коэффициента γ потери в низких частотах могут быть настолько велики, что изображение распадется на квадраты 8×8 . Потери в высоких частотах так же могут проявиться в так называемом “эффекте Гиббса”, когда вокруг контуров с резким переходом цвета образуется своеобразный “нимб”.

10.3.2.5. Переводим матрицу трансформант 8×8 в 64-элементный вектор при помощи “зигзаг”-сканирования, т.е. берем элементы с индексами $(0,0)$, $(0,1)$, $(1,0)$, $(2,0)$... , как показано на рисунке 10.2:

$a_{0,0}$	$a_{0,1}$	$a_{0,2}$	$a_{0,3}$	$a_{0,4}$	$a_{0,5}$	$a_{0,6}$	$a_{0,7}$
$a_{1,0}$	$a_{1,1}$	$a_{1,2}$	$a_{1,3}$	$a_{1,4}$	$a_{1,5}$	$a_{1,6}$	$a_{1,7}$
$a_{2,0}$	$a_{2,1}$	$a_{2,2}$	$a_{2,3}$	$a_{3,0}$			
$a_{3,0}$	$a_{3,1}$	$a_{3,2}$	$a_{3,3}$				
$a_{4,0}$	$a_{4,1}$	$a_{4,2}$					
$a_{5,0}$	$a_{5,1}$						
$a_{6,0}$	$a_{6,1}$						
$a_{7,0}$	$a_{7,1}$						

Рисунок 10.2 – “Зигзаг”- сканирование

Таким образом, в начале вектора мы получаем коэффициенты матрицы, соответствующие низким частотам, а в конце — высоким.

10.3.2.6. Сворачиваем вектор с помощью алгоритма группового кодирования[2]. При этом получаем пары типа (пропустить, число), где “пропустить” является счетчиком пропускаемых нулей, а “число” — значение, которое необходимо поставить в следующую ячейку. Так, например, вектор 42 3 0 0 0 -2 0 0 0 0 1 ... будет свернут в пары (0,42) (0,3) (3,-2) (4,1)

10.3.2.7. Сворачиваем получившиеся пары кодированием по Хаффману с фиксированной таблицей[2].

Процесс восстановления изображения в алгоритме JPEG полностью симметричен процессу сжатия.

10.4. Порядок проектирования

10.4.1. С использованием пакета MATLAB разработать процедуры, осуществляющие преобразование изображения из цветового пространства RGB в пространство YCrCb и обратно. Отладить процедуру и сравнить получаемые результаты с результатами функций *rgb2ycbcr()* и *ycbcr2rgb()* пакета MATLAB.

10.4.2. С использованием пакета MATLAB разработать процедуру, выполняющую двумерное ДКП над входной матрицей размера 8x8. Отладить процедуру и сравнить получаемые результаты с результатами функции *dct2()* пакета MATLAB.

10.4.3. Разработать с использованием пакета MATLAB процедуру, реализующую квантование матрицы трансформант, получаемой в п. 4.2

10.4.4. Разработать с использованием пакета MATLAB процедуру, реализующую «зигзаг»-сканирование квантованной матрицы и кодирование полученного вектора с помощью алгоритма группового кодирования.

10.4.5. С помощью пакета MATLAB реализовать двухпроходный кодер Хаффмана, реализующий сжатие полученных в п. 4.4 пар значений.

10.4.6. С помощью пакета MATLAB реализовать, программу выполняющую все этапы алгоритма сжатия изображений JPEG. Отладить разработанную программу.

10.4.7. С помощью пакета MATLAB реализовать, программу выполняющую все этапы алгоритма восстановления изображений.

10.4.8. Оптимизировать разработанную в пунктах 10.4.6 и 10.4.7 систему по критерию быстродействия.

10.4.9. С помощью Code Composer Studio v2 разработать программу сжатия и восстановления изображений в соответствии с алгоритмом JPEG на языке C на ЦПОС TMS320C5402 или TMS320C7611.

10.4.10. Отладить разработанную в п. 4.9 систему и оптимизировать ее по критерию быстродействия.

10.4.11.Используя возможности профайлера(profiler) Code Composer Studio v2 сравнить быстродействие разработанной функции, реализующей двумерное ДКП с функцией **fdct_8x8()** из библиотеки **imglib** для TMS320C62x (подключение библиотеки и вызов библиотечных функций описаны в л. р. №8 часть 2).

Библиографический список

1. Бондарев В.Н. Цифровая обработка сигналов: методы и средства/ В.Н. Бондарев, Г. Трестер, В.С. Чернега.- Севастополь: Изд-во СевГТУ, 1999. – 388с.
2. Чернега В.С. Сжатие информации в компьютерных сетях/ В.С. Чернега. – Севастополь: Изд-во СевГТУ, 1997. – 214с.
3. Ватолин Д.С. Алгоритмы сжатия изображений. Методическое пособие/ Д.С.Ватолин.-М.: Издательский отдел факультета Вычислительной Математики и Кибернетики МГУ им. М.В.Ломоносова, 1999.-76с.
4. Wallace G. The JPEG Still Picture Compression Standart// Communication of the ACM. –1987. –Vol. 30 №6, p. 520-540

11. ТЕМА 11

РЕАЛИЗАЦИЯ ЗВУКОВЫХ МУЗЫКАЛЬНЫХ ЭФФЕКТОВ

11.1. Цель проектирования

Закрепление навыков практического решения задач цифровой обработки сигналов в реальном времени с использованием ЦПОС TMS320C6711(5402).

11.2. Задание на проектирование

Разработать на ЦПОС TMS320C6711(5402) систему, реализующую звуковые музыкальные эффекты в реальном времени. Необходимо реализовать эффекты: однократной и многократной цифровой задержки, эффект детонации, эффект хора, эффект вибрато, эффект реверберации, компрессор и экспандер.

11.3. Методические рекомендации по проектированию

11.3.1. Звуковые музыкальные эффекты

В современной музыкальной индустрии повсеместно используются различные звуковые эффекты, обогащающие палитру музыкальных произведений и добавляющие к ним дополнительные краски. Эффекты могут использоваться не только при обработке музыкальных произведений, но и при воспроизведении речевого материала для снижения монотонности при прослушивании длительных фрагментов.

11.3.1.1. Эффекты, основанные на задержке сигнала

Одним из самых распространенных классов являются звуковые эффекты, основанные на задержке сигнала[1]. В реальной ситуации задержка звукового сигнала может возникнуть при отражении звуковой волны от твердой преграды, например, от стены. Величину возникающей при этом задержки легко определить, зная пройденный звуковой волной путь и учитывая, что скорость звука на уровне моря можно считать равной 330 м/с. Таким образом, за одну миллисекунду звуковая волна проходит 33 сантиметра. Для расчетов удобнее помнить, что один метр она проходит за 3 мс.

Рассмотрим сначала реализацию однократного отражения звуковой волны. Этот простой эффект является базовым для реализации более сложных эффектов, таких, как имитация реверберации и хора.

Как уже говорилось выше, отражение звуковой волны от твердой преграды приводит к появлению в точке регистрации задержанной версии излучаемого сигнала. До появления цифровых процессоров обработки сигнала и других цифровых устройств для задержки основного сигнала использовались специ-

альные магнитофоны с несколькими подвижными головками. При цифровой реализации линии задержки входной аналоговый сигнал преобразуется в цифровую форму, записывается в циклический или любой другой буфер задержки, и сигнал с выхода этого буфера задержки преобразуется обратно в аналоговую форму и воспроизводится.

Блок-схема реализации однократной задержки звукового сигнала приведена на рисунке 11.1

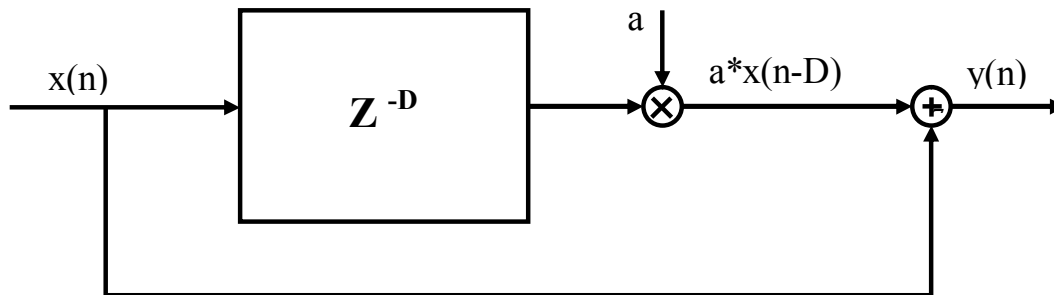


Рисунок 11.1 - Блок-схема реализации однократной задержки звукового сигнала

Входной сигнал $x(n)$ одновременно поступает на линию задержки и на сумматор, формирующий выходной сигнал. Задержанный на величину D сигнал $x(n)$ умножается на масштабный коэффициент a и также поступает на сумматор, формирующий выходной сигнал. Преобразования, осуществляемые данной схемой, могут быть выражены следующей формулой:

$$y(n) = x(n) + ax(n - D). \quad (11.1)$$

Данная схема является линейным фильтром с передаточной функцией:

$$H(z) = 1 + az^{-D}. \quad (11.2)$$

Поскольку при любом отражении звуковой сигнал ослабляется, то $-1 < a < 1$. Отрицательные значения соответствуют отражению в противофазе. Эта модель является достаточно грубой, поскольку как распространение звуковой волны, так и ее отражение являются частотно-зависимыми процессами. Однако в большинстве случаев эта модель является адекватной. Величина D определяется разностью времени между приходом прямой и отраженной звуковой волны.

Передаточная характеристика (11.2) является характеристикой гребенчатого фильтра. Отличительной особенностью данного класса фильтров является выполнение для них равенства

$$H(j\omega) = H(j\omega + \frac{2\pi}{D}). \quad (11.3)$$

Однократная цифровая задержка часто используется для оживления "сухой" музыки. В этом случае как основной, так и задержанный аудиосигналы складываются с одинаковой амплитудой. Для того чтобы при выполнении этой

операции не возникало переполнения разрядной сетки цифрового устройства, оба сигнала складываются с коэффициентом $1/2$. В этом случае используется задержка порядка 15—40 мс. Введение этой задержки приводит к появлению у слушателя впечатления, что два исполнителя играют в унисон. При уменьшении величины задержки ниже 10 мс в сигнале могут возникнуть нежелательные биения, поскольку сумма двух гармонических сигналов с одинаковой амплитудой может быть представлена как амплитудная модуляция гармонического сигнала с частотой, равной сумме частот исходных сигналов, гармоническим сигналом с частотой, равной разности частот исходных сигналов.

Область использования однократных задержек сигнала очень ограничена, поскольку в реальной жизни редко встречается ситуация, когда в акустическом поле присутствует только одна отражающая поверхность. Кроме того, как отмечалось выше, даже в этом случае для повышения адекватности модели необходимо использовать формирующий фильтр, учитывающий частотную характеристику отражающей поверхности и, возможно, частотную характеристику среды распространения сигнала. Для моделирования акустических полей с несколькими отражающими поверхностями и учета частотных характеристик процессов отражения и распространения звуковой волны используются многократные задержки.

Многократная задержка сигнала может быть достигнута двумя путями: доступом к различным элементам, хранящимся в одной линии задержки, или созданием нескольких отдельных линий, имеющих различное время задержки. Первый способ используется в цифровой обработке сигналов и в магнитофонах с несколькими считывающими головками, а второй — при использовании механических и электрических линий задержки.

Многократная задержка сигнала может быть реализована как с использованием трансверсальных (с конечной импульсной характеристикой), так и с использованием рекурсивных фильтров (с бесконечной импульсной характеристикой).

Блок-схема реализации многократной задержки сигнала с использованием трансверсального фильтра приведена на рисунке 11.2.

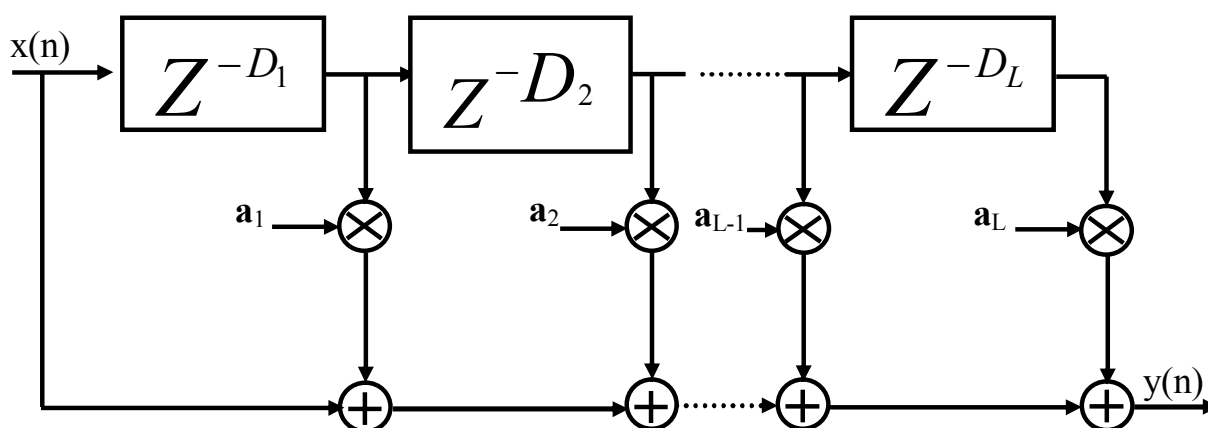


Рисунок 11.2 - Блок-схема реализации многократной задержки с использованием трансверсального фильтра

Разностное уравнение трансверсального фильтра может быть записано в виде:

$$y(n) = x(n) + \sum_{i=1}^L a_i x(n - \sum_{k=1}^i D_k) \quad (11.4)$$

В данном фильтре используются различные величины задержек входного сигнала, что не позволяет считать этот фильтр гребенчатым. Использование конечного числа линий задержки дает возможность учесть влияние ограниченного числа отражающих поверхностей.

Поскольку в трансверсальном фильтре может быть учтено только ограниченное число отражений сигнала, то его нельзя использовать для моделирования переотражений звуковой волны. Для этого следует использовать рекурсивные фильтры.

Блок-схема реализации многократной задержки с использованием простейшего рекурсивного фильтра приведена на рисунке 11.3.

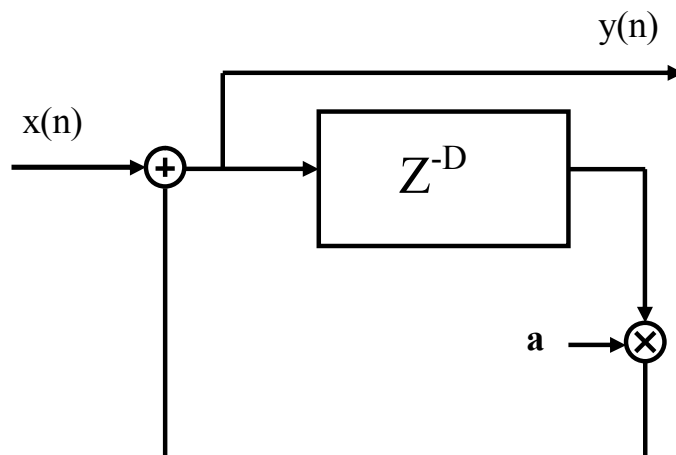


Рисунок 11.3. Блок-схема реализации многократной задержки с использованием рекурсивного фильтра

Этот фильтр может быть использован для описания переотражений звуковой волны между двумя бесконечными параллельными стенами, на одной из которых расположен излучатель плоской волны. Приемник может располагаться на любой из стен. При отражении от стены энергия звуковой волны падает в a раз. Импульсная характеристика данного фильтра представляет собой последовательность одиночных импульсов, расположенных на расстоянии D отсчетов друг от друга, причем импульс с номером n имеет амплитуду a^n .

Данный фильтр описывается разностным уравнением:

$$y(n) = ay(n - D) + x(n) \quad (11.5)$$

и имеет передаточную функцию:

$$H(z) = \frac{1}{1 - az^{-D}}. \quad (11.6)$$

Для имитации более сложных преотражений может быть использован фильтр, блок-схема которого приведена на рисунке 11.4.

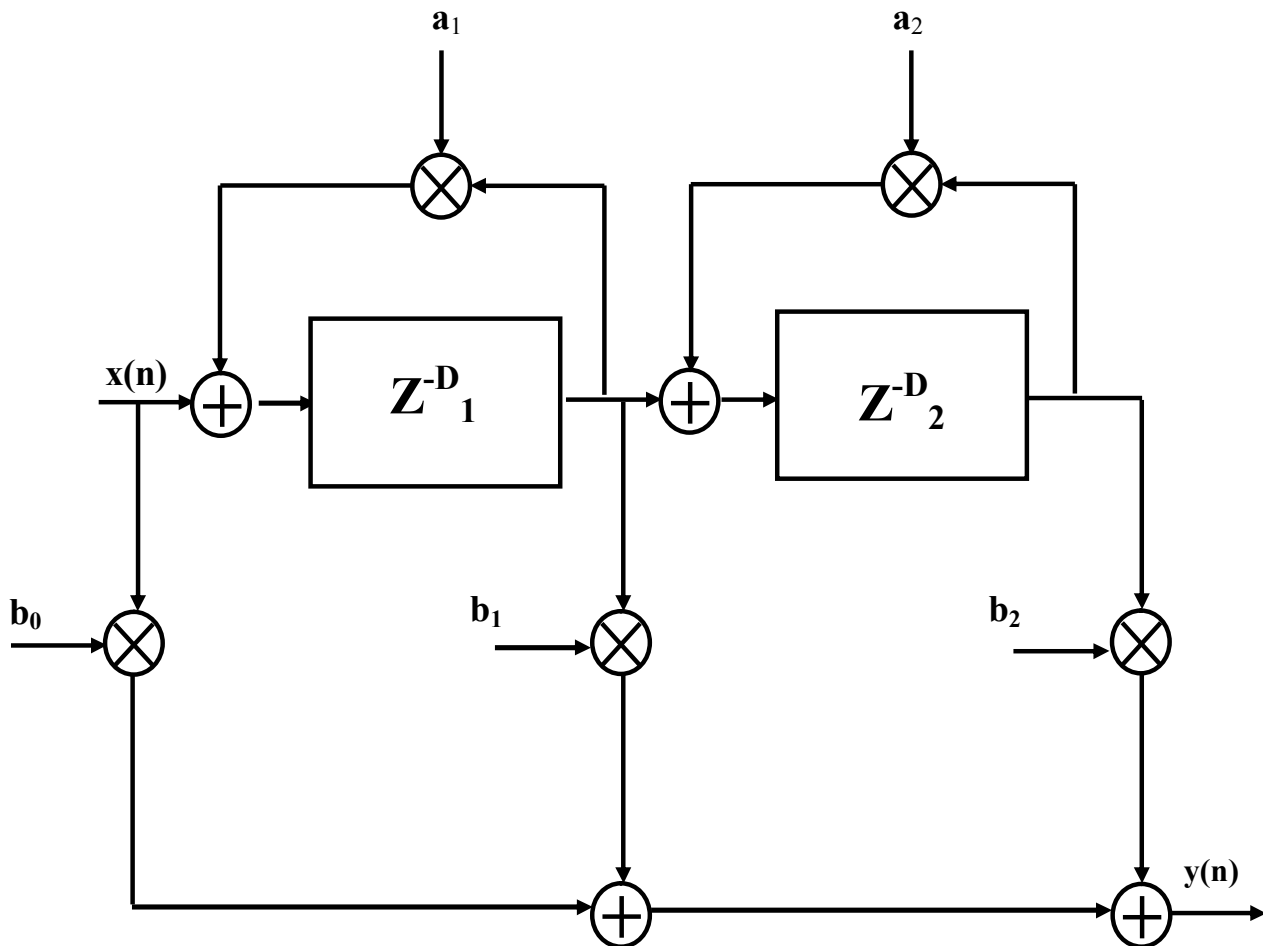


Рисунок 11.4 - Блок-схема реализации многократной задержки с использованием двух линий задержки

Передаточная функция данного фильтра может быть записана в следующем виде:

$$H(z) = b_0 + b_1 \left[\frac{z^{-D_1}}{1 - a_1 z^{-D_1}} \right] + b_2 \left[\frac{z^{-D_1}}{1 - a_1 z^{-D_1}} \right] \left[\frac{z^{-D_2}}{1 - a_2 z^{-D_2}} \right]. \quad (11.7)$$

11.3.1.2. Более сложные звуковые эффекты

Кроме описанных выше простейших звуковых эффектов существуют и более сложные, основанные на перемещении спектральных составляющих исходного сигнала. Такой эффект может быть, например, достигнут при изменении величины задержки сигнала по определенному закону в базовой блок-схеме, приведенной на рис. 11.5

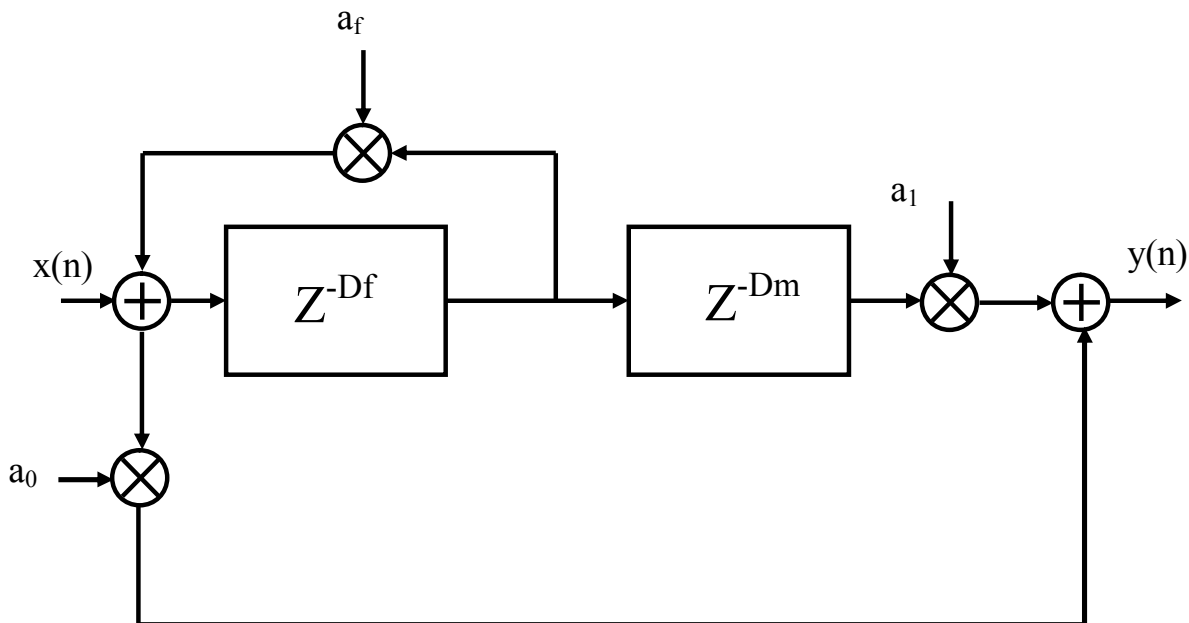


Рисунок 11.5 - Базовая блок-схема

В этой блок-схеме величина D_f является постоянной, а величина $D_m = f_m(t)$ представляет собой некоторую функцию времени. Поэтому данная блок-схема реализует гребенчатый фильтр с изменяющейся во времени частотной характеристикой. Эта блок-схема имеет слишком сложную структуру, чтобы описать ее одним разностным уравнением. Выбор закона изменения величины D_m , а также значения постоянных величин a_0 , a_1 , a_f и D_f позволяют реализовать различные эффекты.

11.3.1.2.1. Эффект детонации

Эффект детонации был открыт при попытке использования выходных сигналов с двух катушечных магнитофонов для создания эффекта хора, который будет описан ниже. В результате возник эффект, напоминающий звук взлетающего реактивного самолета.

Для реализации эффекта детонации в цифровых устройствах может быть использовано устройство, блок-схема которого приведена на рисунке 11.6.

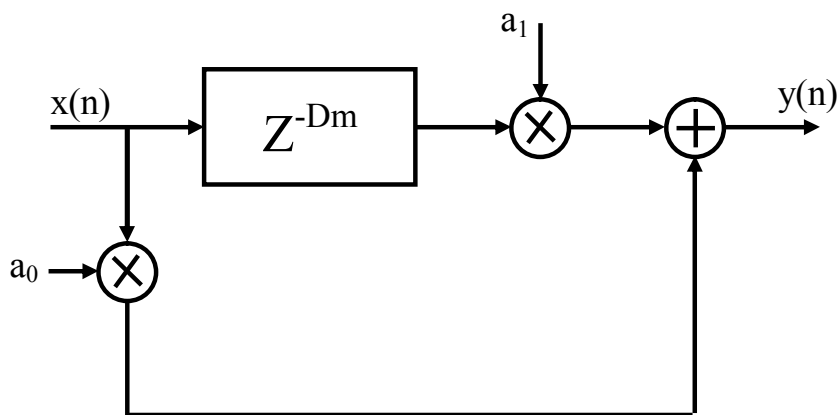


Рисунок 11.6- Блок-схема устройства, реализующего эффект детонации

Как видно из рисунка 11.6, эта блок-схема является частным случаем базовой блок-схемы, изображенной на рисунке 11.5. Данной блок-схеме соответствует следующее разностное уравнение:

$$y(n) = a_0 x(n) + a_1 x(n + f(n)) , \quad (11.8)$$

где

$$f(n) = \frac{D}{2} \left[1 + \cos\left(\frac{2\pi}{N} n\right) \right] .$$

Для предотвращения возникновения переполнений разрядной сетки при вычислении выходного значения коэффициент a_0 выбирается равным 0,5, а коэффициент a_1 выбирается от -0,5 до 0,5. Значение параметра D , определяющего диапазон изменения задержки, обычно выбирается от 0.25 до 25 мс. Значение параметра N , определяющего частоту детонации, выбирается исходя из требований, предъявляемых к выходному сигналу.

Уменьшение длины линии задержки приводит к пропорциональному подъему частот спектральных составляющих исходного сигнала, а ее уменьшение — к понижению этих частот.

11.3.1.2.2. Эффект хора

Другим эффектом, обогащающим звуковую палитру произведения, является эффект хора. Он позволяет создать у слушателя иллюзию, что данное произведение одновременно исполняется на нескольких одинаковых музыкальных инструментах. Хотя музыканты стараются играть синхронно, они не могут избежать небольших отклонений от общего темпа. Кроме того, неизбежны различия в настройке отдельных инструментов и в громкости исполнения.

Для создания эффекта одновременного исполнения произведения на M инструментах может быть использовано устройство, блок-схема которого приведена на рисунке 11.7.

Как правило, все коэффициенты при реализации эффекта задержки выбираются равными ($1/M$), а сами величины задержки не превышают 15—35 мс. Естественно, что каждая линия задержки должна изменять свои параметры независимо от всех остальных.

При таком выборе коэффициентов моделируются только различия в темпе исполнения произведения и в настройке инструментов. То что эти различия связаны друг с другом, на слух не определяется, поскольку в реальной ситуации такая зависимость отсутствует, и слух не приспособлен для ее выделения. Для моделирования различного уровня воспроизведения достаточно сделать переменными коэффициенты фильтра, добавив к их постоянному значению некоторый случайный низкочастотный сигнал малого уровня, поскольку исполнители стараются сохранять средний уровень громкости при воспроизведении. Для простоты в качестве случайного низкочастотного сигнала может быть использован тот же самый сигнал, что и в линиях задержки.

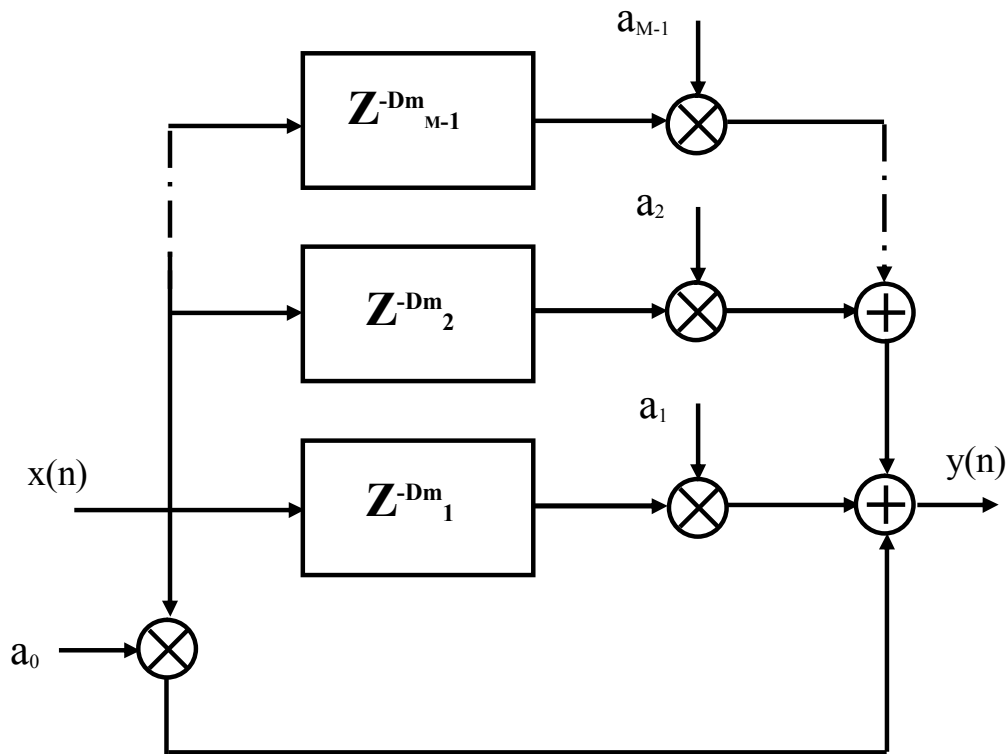


Рисунок 11.7 - Блок-схема устройства, реализующего эффект хора

11.3.1.2.4. Эффект вибрато

Эффект вибрато имитирует вибрато человеческих голосов, определяющий их тембр, теплоту и выразительность, и соответствующий прием исполнения на струнных музыкальных инструментах, периодически изменяющий в небольших пределах высоту звуков. Эффект вибрато может быть получен в устройствах, реализующих эффекты детонации и хора при полном удалении входного сигнала из выходного. Величина задержки при этом выбирается таким образом, чтобы частотный сдвиг основных спектральных составляющих сиг-

нала составлял несколько герц. Этот эффект часто путают с эффектом тремоло, где сигнал подвергается не частотной, а амплитудной модуляции. Оба эффекта используют одинаковый модулирующий сигнал. Эффекты тремоло и вибрато могут совместно использоваться для создания эффекта вращающегося диктора.

Блок-схема устройства, реализующего эффект вибрато, приведена на рисунке 11.8:

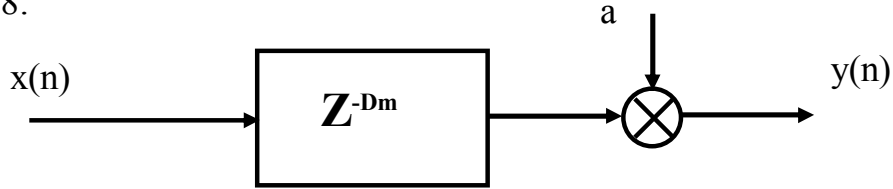


Рисунок 11.8 - Блок-схема устройства, реализующего эффект вибрато

Устройство, блок-схема которого приведена на рисунке 11.8, может использоваться также и для изменения основного тона диктора (например, чтобы сделать голос говорящего неузнаваемым). В этом случае величина задержки изменяется от нуля до некоторой предельной величины по линейному закону. По достижении задержки своей предельной величины она сбрасывается в ноль, и процесс возобновляется. Изменение задержки по такому закону позволяет обеспечить работу устройства в реальном времени, но в момент обнуления величины задержки слышен щелчок. Дело в том, что как раз в данный момент происходит разрыв воспроизводимого сигнала при переходе к повторному воспроизведению некоторого его участка.

Если необходимо не повысить, а понизить частоту основного тона, то величина задержки постоянно уменьшается и при достижении нуля устанавливается в свое максимальное значение. При таком переходе теряется часть входного сигнала. Для устранения щелчков, возникающих при изменении величины задержки, рекомендуется использовать фильтры нижних частот.

Более простым способом устранения щелчков является использование двух подобных устройств, фазы пилообразных модулирующих сигналов которых сдвинуты на половину периода. При этом необходимо задать весовое окно (например, окно Ханна) на периоде модулирующего сигнала. Выходные сигналы каждого устройства умножаются на отсчет весовой функции, соответствующий текущей длине линии задержки, и складываются друг с другом, что обеспечивает устранение щелчков при переходах. Если обозначить период модулирующего сигнала как N , отсчет модулирующего сигнала как $f_m(n)$, а отсчет весовой функции как $W(n)$, то, с использованием функции получения остатка от целочисленного деления числа M на число N , обозначаемой как $\text{mod}_N(M)$, процедура преобразования основного тона может быть записана в форме следующего разностного уравнения:

$$y(n) = W \{ \text{mod}_N [f_m(n)] \} x(n - f_m(n)) + \\ + W \left\{ \text{mod}_N \left[f_m(n) + \frac{N}{2} \right] \right\} x(n - f_m(n) - \frac{N}{2}). \quad (11.10)$$

11.3.1.2.4. Эффект реверберации

Эффект реверберации возникает в больших помещениях вследствие множественных переотражений звуковой волны от различных преград. Реверберация не может быть сведена к нескольким повторениям исходного сигнала с использованием конечной линии задержки. Время реверберации, т. е. время затухания сигнала на 60 дБ, колеблется для различных помещений от нескольких сотен миллисекунд до нескольких секунд.

В реверберационном сигнале можно различить три основных компонента:

- основной сигнал — сигнал, пришедший к слушателю, не претерпев по пути отражений;
- первичные отражения — различимые отражения основного сигнала от отдельных преград, поступающие с задержкой от 10 до 100 мс;
- шум переотражений — возникает вследствие множества переотражений звуковой волны. В данном шуме не могут быть выделены результаты отражения звуковой волны от определенной преграды.

Основная энергия сигнала, как правило, приходится на основной сигнал и первичные отражения. В случаях сильно реверберирующих помещений суммарный уровень отраженных составляющих может превышать уровень основного сигнала на несколько десятков децибел. Это связано с тем, что препятствия, встречающиеся на пути звуковой волны, поглощают лишь малую ее часть. Например, коэффициент отражения звуковой волны от поверхности пресной воды составляет 0,9994, т. е. при отражении поглощается только 0,06% энергии звукового сигнала. Поэтому, чтобы поглощаемая энергия равнялась излучаемой, общая энергия звукового сигнала должна быть во много раз больше излучаемой энергии. Кроме того, следует учесть, что суммарный уровень отраженных сигналов примерно одинаков во всем помещении, а уровень основного сигнала спадает пропорционально квадрату расстояния от источника.

Такое низкое отношение сигнал/шум, казалось бы, противоречит жизненному опыту, поскольку даже в сильно реверберирующих помещениях разборчивость речи достаточно высока. Этот факт объясняется тем, что отраженные сигналы, задержка которых относительно основного сигнала не превышает 50—70 мс, не ухудшают, а даже несколько повышают разборчивость. Кроме того, за счет бинаурального прослушивания человеку удастся локализовать в пространстве источник информации и отсеять сигналы, приходящие с других направлений. Роль бинаурального прослушивания хорошо демонстрируется при сравнении разборчивости речи при непосредственном прослушивании в реверберирующем помещении и при прослушивании монофонической магнитофонной записи, проведенном в том же помещении.

Каждый из компонентов реверберационного сигнала должен моделироваться отдельно. Проще всего обстоит дело с основным сигналом: его необходимо только задержать и изменить уровень. Эта процедура существенно упрощается, если исключить задержку основного сигнала. Возможность исключения задержки основного сигнала связана с тем, что слушатель воспринимает сигнал не в абсолютной, а в относительной шкале времени, привязанной не ко

времени излучения акустического сигнала, а ко времени прихода его первых компонентов, т. е. времени прихода основного сигнала. Первичные отражения сигнала могут быть промоделированы с использованием трансверсального фильтра. А для генерации шума переотражений должны использоваться рекурсивные фильтры.

Структура устройства, предназначенного для моделирования реверберации и содержащего в себе все вышеперечисленные типы фильтров, была предложена Мурером (James A. Moorer) в 1979 году. Блок-схема этого устройства приведена на рисунке 11.9.

Данное устройство состоит из трансверсального фильтра, моделирующего первичные отражения, шести рекурсивных гребенчатых фильтров, моделирующих переотражения звуковой волны, и сумматора, объединяющего три основных компонента сигнала реверберации.

Трансверсальный фильтр представляет собой линию задержки на 80 мс, имеющую 18 отводов. Для получения выходного сигнала фильтра сигнал с каждого из отводов умножается на свой масштабный коэффициент и суммируется с масштабированными сигналами с других отводов линии задержки. Значения задержек и масштабных коэффициентов приведены в таблице 11.1.

Таблица 11.1- Параметры трансверсального фильтра

Номер элемента	D, мс	A
1	4,3	0,841
2	17,2	0,504
3	1,00	0,490
4	4,3	0,379
5	0,2	0,380
6	2,8	0,346
7	16,0	0,289
8	2,7	0,272
9	8,7	0,192
10	1,5	0,193
11	0,8	0,217
12	1,7	0,181
13	9,5	0,180
14	0,1	0,181
15	1,8	0,176
16	1,5	0,142
17	1,2	0,167
18	4,4	0,134

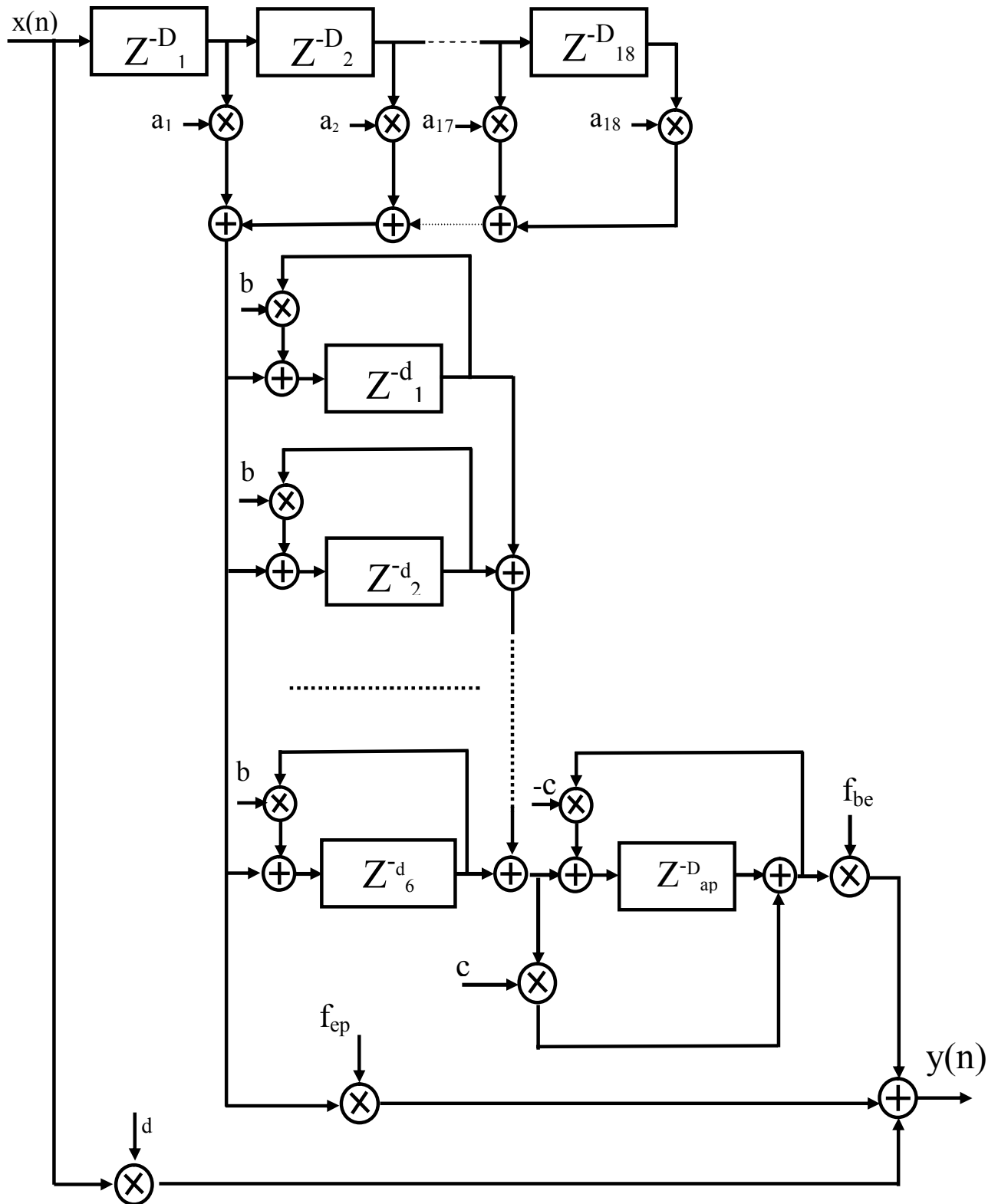


Рисунок 11.9 - Блок-схема устройства, реализующего эффект реверберации

Все рекурсивные фильтры имеют одинаковый коэффициент обратной связи, равный 0.3, и различаются только величиной задержки. Величины задержек рекурсивных фильтров приведены в таблице 11. 2.

Таблица 11.2- Величины задержек рекурсивных фильтров

Номер фильтра	Величина задержки, мс
1	40
2	44
3	48
4	52
5	56
6	60

Для повышения качества выходного сигнала устройства полученный шум переотражений пропускается через фазовращатель, т. е. фильтр, частотная характеристика которого имеет одинаковое значение модуля во всем диапазоне частот. В данном фильтре использована задержка $D_{ap} = 7$ мс. Переходная функция данного фильтра имеет вид:

$$H(z) = \frac{1 + cz^{-D_{ap}}}{1 + cz^{-D_{ap}}}, \quad (11.11)$$

где $c = 0,7$.

Коэффициенты f_d , f_{ep} и f_{be} , определяющие вклад в выходной сигнал основного сигнала, первичных отражений и шума переотражений, выбираются исходя из конкретных условий.

11.3.1.2.5. Компрессоры, экспандеры и ограничители

Компрессоры и экспандеры используются для изменения динамического диапазона сигнала, т. е. для уменьшения различия уровней самого тихого и самого громкого сигналов. Для этого при превышении текущим значением сигнала некоторого порога устанавливается коэффициент передачи, меньший, чем используемый до превышения порога. Отношение старого коэффициента передачи к новому называется ослаблением. Все коэффициенты передачи задаются в логарифмической шкале. То есть при ослаблении 2:1 приращению входного сигнала на 20 дБ будет соответствовать приращение выходного сигнала на 10 дБ. Подобный компрессор, изменяющий свою передаточную характеристику в точке a , может быть реализован по следующей формуле:

$$y = \begin{cases} x, & |x| \leq \alpha \\ \text{sgn}(x) \sqrt{\frac{|x|}{\alpha}}, & |x| < \alpha \end{cases}. \quad (11.12)$$

Применение компрессоров позволяет передавать по линиям связи сигналы, имеющие больший динамический диапазон, чем передающий канал. На приемном конце этой линии устанавливается экспандер, имеющий передаточную характеристику, обратную характеристике компрессора.

Другой областью применения компрессоров является звукозаписывающая индустрия и радиовещание. Дело в том, что при записи музыкальных произведений для более точной передачи звучания всех групп инструментов оркестра устанавливается несколько микрофонов. Звуковое давление создает на выходе микрофона электрический сигнал, имеющий динамический диапазон 60—70 дБ. Однако большинство каналов неспособно передать весь диапазон изменения сигнала вследствие ограничений, накладываемых помехами в канале и его нелинейностями. Поэтому возникает задача сжать исходный динамический диапазон сигнала до динамического диапазона канала, обычно составляющего порядка 40 дБ.

При уровне ослабления выше чем 10:1, компрессор превращается в ограничитель. Его основное отличие заключается в том, что применение компрессора предполагает последующее использование экспандера, а после ограничителя какие-либо разговоры о восстановлении сигнала, как правило, бессмысленны. Поэтому наиболее часто используются ограничители с уровнем ослабления $\infty : 1$, что соответствует ограничению сигнала на уровне порога.

Поскольку передаточные характеристики компрессоров и экспандеров отличаются большой нелинейностью, их использование приводит к существенному искажению сигнала. Так, например, при использовании аналоговых компрессоров и экспандеров неизбежно возникает вопрос об обеспечении равенства амплитуд передаваемого и принимаемого сигналов, поскольку в противном случае точки изменения коэффициентов передачи компрессора и экспандера не совпадут, что приведет к возникновению нелинейных искажений. Как отмечалось выше, во многих случаях необходимо производить компрессию сигнала с опережением, т. е. до фактического превышения сигналом порогового уровня. В этом случае при медленном росте или спаде сигнала несоответствие точек изменения коэффициентов передачи приведет к растягиванию во времени описанных выше нелинейных искажений.

11.4. Порядок проектирования

11.4.1.С использованием пакета MATLAB реализовать эффект однократной задержки сигнала (рисунок 11.1). Входные данные вводить из монофонического звукового файла формата wav (частота дискретизации – 8кГц, разрядность - 16 бит).

11.4.2.С использованием пакета MATLAB реализовать эффект многократной задержки сигнала с использованием трансверсального фильтра (рису-

нок 11.2) и рекурсивных фильтров (рисунок 11.3 и рисунок 11.4). Сравнить полученные результаты и выбрать наиболее подходящий (по критериям быстродействие, качество выходного сигнала).

11.4.3. С использованием пакета MATLAB реализовать эффект детонации в соответствии со схемой, приведенной на рисунке 11.6.

11.4.4. С использованием пакета MATLAB реализовать эффект хора в соответствии со схемой, приведенной на рисунке 11.7. Произвести анализ результатов при различных значениях M .

11.4.5. С помощью пакета MATLAB реализовать компрессор и экспандер в соответствии с формулой (11.11). Исследовать работу полученных устройств при различных значениях α .

11.4.6. С помощью Code Composer Studio v2 разработать программу на языке C, реализующую эффекты разработанные в пунктах 11.4.1–11.4.5 для обработки звукового сигнала в реальном времени на ЦПОС TMS320C5402 или TMS320C7611. Для управления выбором типа эффекта, а также параметрами выбранного эффекта при обработке сигнала в реальном времени разработать Windows-приложение, использующее для этой цели возможности RTDX-каналов ЦПОС TMS320C5402(6711) (см. документацию Code Composer Studio v2).

11.4.10. Отладить разработанную в пункте 11.4.9 систему и оптимизировать ее по критериям быстродействия и качества выходного сигнала.

Библиографический список

1. Секунов Н.Ю. Обработка звука на PC/ Н.Ю. Секунов. – СПб.: БХВ-Петербург, 2001. – 1248 с.
2. Бондарев В.Н. Цифровая обработка сигналов: методы и средства/ В.Н. Бондарев, Г. Трестер, В.С. Чернега.- Севастополь: Изд-во СевГТУ, 1999. – 388с.

12. ТЕМА 12

СЖАТИЕ АУДИОСИГНАЛОВ

12.1.Цель проектирования

Закрепление навыков практического решения задач сжатия звуковых сигналов.

12.2.Задание на проектирование

Разработать и реализовать на языке C/C++ систему сжатия аудиосигналов, осуществляющую сжатие и восстановление звукового аудиосигнала, представленного в виде последовательности 16-разрядных отсчетов с частотой дискретизации 44100Гц. Обеспечить коэффициент сжатия на уровне 5:1 без заметной потери качества сигнала.

12.3.Методические рекомендации по проектированию

12.3.1.Особенности восприятия аудиоинформации человеком

При оценке восприятия аудиоинформации приходится постоянно помнить о двояком подходе к одному и тому же звуковому колебанию — объективном и субъективном.

С одной стороны, звук — это объективное явление, описываемое совокупностью физических характеристик — интенсивностью, частотой, спектром и другими характеристиками. Но при оценке его восприятия нас интересует не сам акустический сигнал, а воздействие, оказанное им на слуховой анализатор, т. е. раздражение, величина и характер которого зависят от всех перечисленных выше физических характеристик звука, которые в этом смысле могут быть названы параметрами раздражения.

При анализе параметров раздражения необходимо, прежде всего, определить его пороговые величины, т. е. такие величины данного параметра, при увеличении или уменьшении которых восприятие аудиоинформации прекращается.

Одной из базовых пороговых величин является порог слышимости. Дело в том, что не всякие изменения акустического давления воспринимаются слухом как звук: существуют определенные границы слухового ощущения как по величине давления, так и по его частоте (которые в данном случае являются параметрами раздражения).

Минимальное звуковое давление, при котором еще существует слуховое ощущение, называется порогом слышимости. Он характеризует чувствительность слуха: чем выше порог, тем меньше чувствительность. Опыты по определению порога слышимости показали, что его величина зависит от условий опыта, частоты и особенностей звукового сигнала. Под условиями опыта по-

нимается характер звукового поля: создается ли оно одним громкоговорителем, помещенным перед слушателем, или многими источниками, равномерно распределенными вокруг головы; имеются ли отражения от границ помещения (так называемое диффузное поле) или приняты меры по их устранению; производились ли измерения минимального давления непосредственно около ушной раковины или в этой же точке при отсутствии слушателя (т. е. в свободном поле).

Порог слышимости существенно зависит от частоты. Наибольшей чувствительностью ухо обладает в области частот 3000—3500 Гц, где порог слышимости имеет наименьшую величину.

При воздействии на слуховой аппарат человека длительного звука большой интенсивности воспринимаемая слушателем громкость звука постепенно уменьшается. Это означает, что под воздействием звука большой громкости чувствительность слуха со времени уменьшается. После прекращения действия звука или уменьшения его интенсивности слух не сразу адаптируется к изменившейся громкости, что приводит к маскировке звуков, следующих за звуком большой громкости. Этот вид маскировки называют *временной маскировкой*. Поскольку восстановление чувствительности слухового аппарата может занять несколько сотен миллисекунд, то вследствие временной маскировки может быть “потерян” фрагмент сигнала.

Слуховой аппарат человека является нелинейной системой. Нелинейность слуха проявляется в том, что при воздействии на барабанную перепонку достаточно громкого синусоидального звука с частотой f в слуховом аппарате зарождаются гармоники этого звука с частотами $2f$, $3f$ и т. д. Поскольку в первичном воздействующем тоне этих гармоник нет, они получили название *субъективных гармоник*. Наличие субъективных гармоник обнаруживается методом звукового зонда. Суть метода заключается в следующем. Если к звуку, под действием которого возникают субъективные гармоники, добавить второй звук небольшой интенсивности, частоту и уровень которого можно плавно изменять, то при неточном совпадении частоты этого звука с частотой субъективной гармоники слушатель будет воспринимать биения, возникшие в результате взаимодействия зондирующего звука и субъективной гармоники.

С нелинейностью слуха непосредственно связано явление, называемое маскировкой. Оно возникает в том случае, когда звуки полезные, т. е. несущие определенную информацию, приходится слушать одновременно с мешающими, или маскирующими звуками. В результате возникает потеря части информации.

Маскирующее действие различных звуков выявляют путем определения порога слышимости испытательных синусоидальных сигналов или узкополосных шумов в присутствии мешающего звука, имеющего определенный уровень интенсивности. Изменяя частоту испытательного тона и определяя на каждой частоте уровень звукового давления, при котором он начинает прослушиваться наряду с мешающим звуком, можно получить кривую порога слышимости при наличии маскировки.

Анализ этих кривых показывает, что маскировка испытательных сигналов, имеющих частоту, лежащую ниже частоты мешающего тона, проявляется значительно слабее, чем для вышележащих частот.

12.3.2. Критические полосы слуха

Проведя более подробные исследования маскировки гармонического сигнала узкополосным шумом, Цвикер пришел к выводу о том, что слуховой анализатор разделяет весь спектр частот на отдельные частотные группы. При этом было сделано предположение о том, что в частотной области до 500 Гц органы слуха воспринимают интенсивность звука, разделяя ее на участки постоянной абсолютной ширины, а свыше 500 Гц — на участки постоянной относительной ширины.

Задолго до этого к тому же выводу пришел Флетчер, он назвал найденные им частотные интервалы критическими полосками слуха. Критические полосы слуха и частотные группы отличаются друг от друга, поскольку Флетчер исходил из маскирующего действия шума, а Цвикер — из соотношений громкости.

Критические полосы слухового восприятия представлены в [1] (таблица 4.1).

12.3.3. Структура системы сжатия звука

Обобщенная структурная схема системы сжатия звука приведена на рисунке 12.1. Кодер состоит из трех компонентов. В блоке временного и частотного анализа входной речевой сигнал преобразуется в спектральные коэффициенты, группируемые в блоки переменной длины (БПД). В блоке назначения разрядности значениям БПД выделяется определенное количество бит, с учетом которого в блоке спектрального квантования происходит квантование спектральных отсчетов.

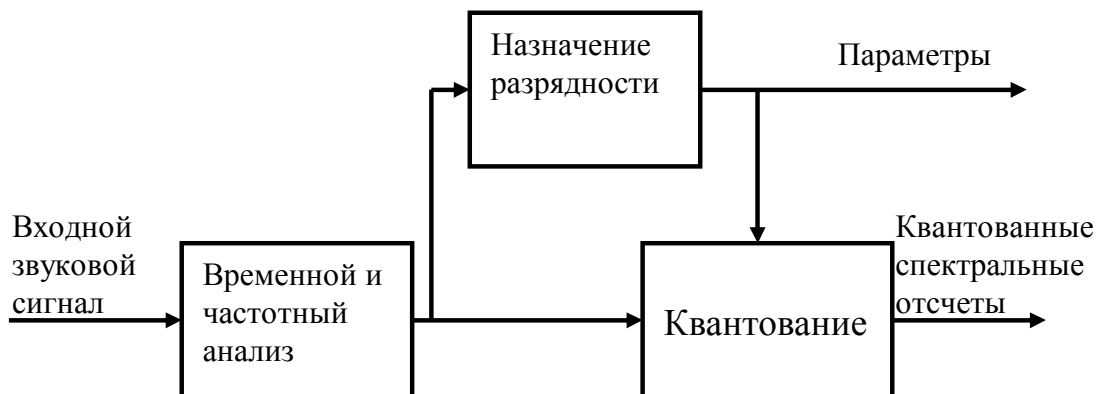


Рисунок 12.1 – Обобщенная структурная схема системы

В блоке временного и частотного анализа (рисунок 12.2) сначала сигнал разделяется на три частотные полосы: 0 – 5.5 кГц, 5.5 – 11 кГц и 11 – 22 кГц. Затем сигнал в каждой из этих полос переводится в частотную область, для этого используется модифицированное дискретное косинусное преобразование (МДКП), которое основано на 50% перекрытии блоков входных отсчетов и позволяет уменьшить краевые эффекты, проявляющиеся в точках сопряжения блоков.

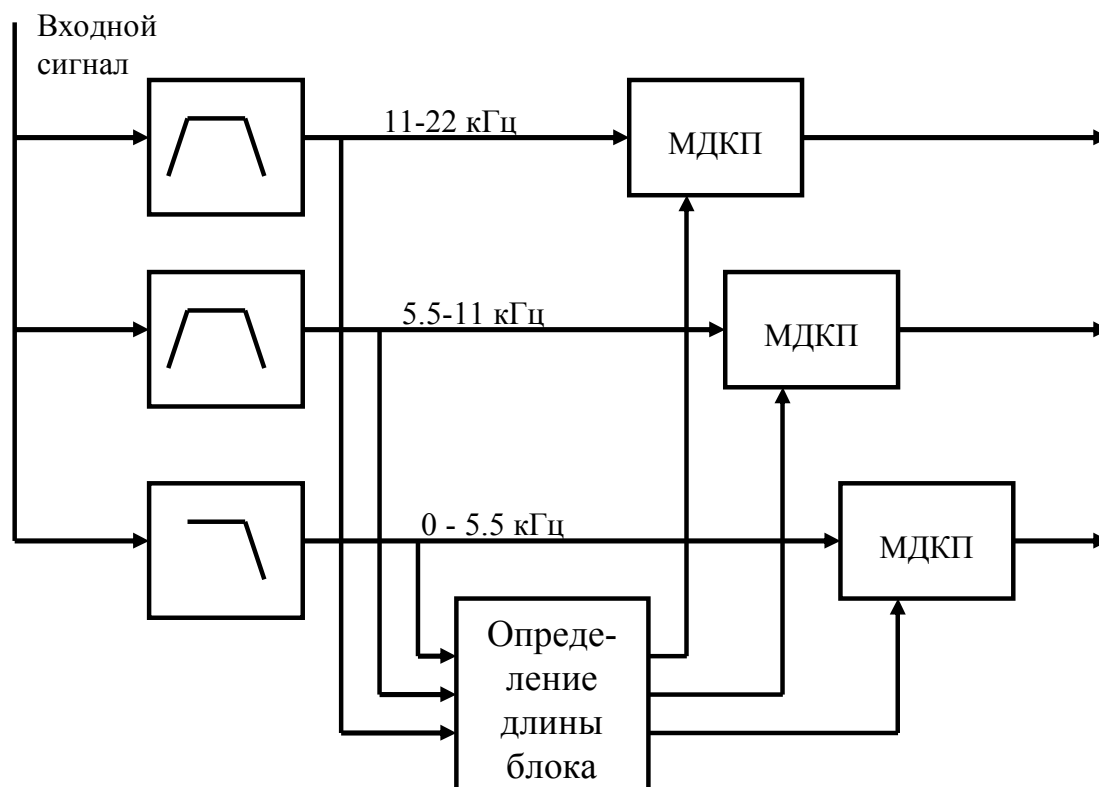


Рисунок 12.2 – Структурная схема блока временного и частотного анализа

Длина блока МДКП для каждого из частотных диапазонов определяется в соответствующем блоке на основании характеристик сигналов каждой из полос. Используется два типа блоков: длинный (11.6 мс) и короткий (1.45 мс для высокочастотной полосы и 2.9 мс для двух остальных). Переключение между данными типами осуществляется на основе анализа стационарности участка аудиосигнала. Короткие блоки используются на нестационарных участках. Это позволяет уменьшить эффект предварительного эхо, который наблюдается, когда в пределах одного блока имеется участок молчания и участок со значительным уровнем звука[1]. Следует отметить, что в том случае, когда за участком с высоким уровнем звука следует участок молчания, в использовании короткого блока нет необходимости, из-за явления временного маскирования.

Трансформанты на выходе каждого из трех каналов анализируются в блоке определения разрядности. Последовательность трансформант разбивает-

ся на группы в соответствии с критическими полосами слухового восприятия, и для каждой из групп, на основании максимального значения трансформанты в данной группе, определяется количество разрядов, необходимое для хранения каждой из трансформант этой группы.

На основании рассчитанной разрядности производится квантование трансформант. Таким образом, в выходной поток для каждого из каналов записываются: длина блока, значения разрядности квантованных трансформант для каждой из групп и, наконец, сами квантованные трансформанты.

12.3.4. Структура декодера

Обобщенная структурная схема декодера звука приведена на рис. 12. 3. Сначала в декодере происходит восстановление квантизированных значений трансформант, используя значения разрядности для каждой из групп. Затем с использованием обратного модифицированного дискретного косинусного преобразования трансформанты преобразуются во временную область, с учетом длины блока. И, наконец, сигналы из трех каналов смешиваются, и получается восстановленный аудиосигнал.

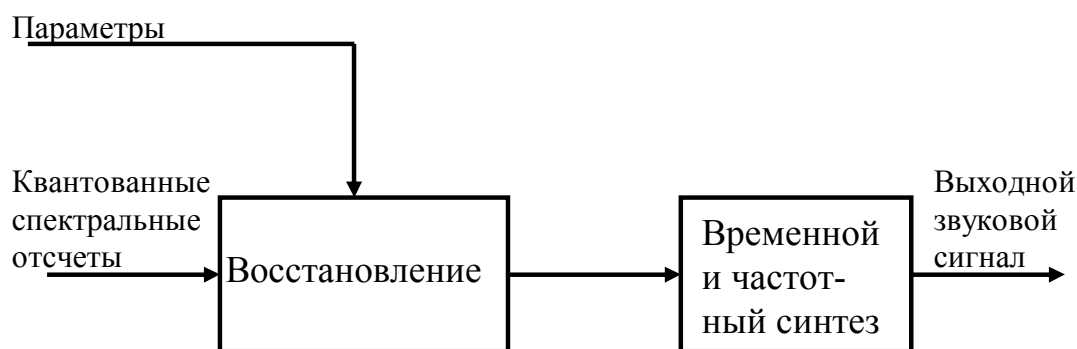


Рисунок 12.3 – Обобщенная структурная схема декодера

12.4. Порядок проектирования

12.4.1.С использованием пакета MATLAB произвести расчет фильтров, разделяющих входной звуковой сигнал на три частотные полосы: 0-5.5, 5.5-11, 11-22 кГц и построить характеристики рассчитанных фильтров. Входной сигнал вводить из файла формата wav, частота дискретизации - 44100 Гц, разрядность – 16 бит.

12.4.2.Разработать с использованием пакета MATLAB процедуру, реализующую модифицированное дискретное косинусное преобразование. Сравнить получаемые результаты с результатами, возвращаемыми функцией **mdct()** пакета MATLAB.

12.4.3.Разработать с использованием пакета MATLAB процедуру временного анализа, определяющую длину блока речевого сигнала в каждой из частотных полос.

12.4.4.С использованием пакета MATLAB реализовать блок назначения разрядности, выполняющий определение необходимого количества разрядов для хранения трансформант в каждой из критических полос слуха.

12.4.5.С помощью пакета MATLAB реализовать блок квантования значений трансформант в соответствии с определенной ранее разрядностью.

12.4.6.С использованием пакета MATLAB реализовать систему сжатия аудио сигнала. Система должна осуществлять сжатие сигнала до уровня не менее 5:1.

12.4.7.С использованием пакета MATLAB реализовать систему восстановления аудио сигнала.

12.4.8.Оптимизировать разработанную в пунктах 12.4.6 и 12.4.7 систему по критериям быстродействие, качество.

12.4.9.Реализовать на языке С систему для сжатия/восстановления речевого сигнала в реальном времени. Ввод аудио информации осуществлять со входа звуковой карты, используя возможности библиотеки DirectSound (технологии DirectX).

12.4.10. Отладить разработанную в пункте 12.4.9 систему и оптимизировать ее по критерию качества восстанавливаемого сигнала.

Библиографический список

1. Бондарев В.Н. Цифровая обработка сигналов: методы и средства/ В.Н. Бондарев, Г. Трестер, В.С. Чернега. - Севастополь: Изд-во СевГТУ, 1999. – 388с.
2. Секунов Н.Ю. Обработка звука на РС/ Н.Ю. Секунов. – СПб.: БХВ-Петербург, 2001. – 1248 с.
3. Цифровая обработка сигналов: Справочник / Л.М. Гольденберг, Б.Д. Матюшкин, М.Н. Поляк.- М.: Радио и связь, 1985. - 312с.