**МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ**

**БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ**

**ФАКУЛЬТЕТ РАДИОФИЗИКИ И КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ**

**Кафедра интеллектуальных систем**

ФАМИЛИЯ

Имя Отчество

**ВЫДЕЛЕНИЕ ПОЛЕЗНОГО СИГНАЛА НА ФОНЕ ШУМА В ПЕРЕГОВОРАХ АВИАДИСПЕТЧЕРОВ**

Дипломная работа

Научный руководитель:

кандидат физико-математических наук,

доцент ………………..

Допущена к защите

«\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_ 2018 г.

Зав. кафедрой интеллектуальных систем,

кандидат физико-математических наук,

доцент К.В. Козадаев

Минск, 2018

ОГЛАВЛЕНИЕ

|  |  |
| --- | --- |
| ПЕРЕЧЕНЬ УСЛОВНЫХ ОБОЗНАЧЕНИЙ | 2 |
| РЕФЕРАТ | 3 |
| ГЛАВА 1 МЕТОДЫ И ТЕХНОЛОГИИ ШУМОПОДАВЛЕНИЯ В РЕЧЕВЫХ СИГНАЛАХ | 5 |
| 1.1 Классификация методов повышения качества и разборчивости речи | 5 |
| 1.1.1. Адаптивные компенсаторы помех | 6 |
| 1.1.2 Методы, основанные на использовании статистических моделей речевых сигналов во временной области | 7 |
| 1.1.3 Обработка речевого сигнала с использованием аппарата скрытых марковских моделей | 8 |
| 1.1.4 Методы, основанные на использовании, отдельных характерных свойств речевого сигнала. | 10 |
| 1.1.5. Методы, основанные на искусственных нейронных сетях | 12 |
| 1.1.6. Метод оценивания минимальной среднеквадратической ошибки | 14 |
| 1.2 Гейты | 15 |
| 1.3 Экспандеры | 17 |
| 1.4 Цифровые решения | 19 |
| 1.5 Изменяющийся фоновый шум | 22 |
| ГЛАВА 2 ВЕЙВЛЕТЫ И ИХ ПРИМЕНЕНИЕ ДЛЯ ШУМОПОДАВЛЕНИЯ В РЕЧЕВЫХ СИГНАЛАХ | 24 |
| 2.1 Теория вейвлетов | 24 |
| 2.2. Применение вейвлетов для шумоподавления, определения наличия полезного сигнала, повышения отношения сигнал/шум. | 28 |
| 2.3 Типы вейвлетов для решения задач шумоподавления | 31 |
| ГЛАВА 3 ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ ФИЛЬТРОВ И ВЕЙВЛЕТОВ ДЛЯ РЕШЕНИЯ ЗАДАЧИ ШУМОПОДАВЛЕНИЯ В ПЕРЕГОВОРАХ АВИАДИСПЕТЧЕРОВ | 35 |
| 3.1 Особенности задач шумоподавления в речевых сигналах | 36 |
| 3.2 Выбор фильтров для очистки сигнала. Результаты их применения | 38 |
| 3.2.1. Реализация стандартного алгоритма наименьших квадратов | 41 |
| 3.2.2 Фильтр Калмана | 45 |
| 3.2.3 Режекторный фильтр | 48 |
| 3.2.4 Результаты применения фильтров для очистки сигнала | 50 |
| 3.3 Выбор вейвлетов для очистки сигнала и результаты их применения | 52 |
| 3.4 Сравнение результатов | 60 |
| ЗАКЛЮЧЕНИЕ | 64 |
| СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ | 66 |
| ПРИЛОЖЕНИЕ А | 67 |

# ПЕРЕЧЕНЬ УСЛОВНЫХ ОБОЗНАЧЕНИЙ

АБГШ - аддитивный белый Гауссов шум

ВП – вейвлет-преобразование

ДПВ – дискретное вейвлет-преобразование

НВП – непрерывное вейвлет-преобразование

ПФ – преобразование Фурье

ОСШ – отношение сигнал/шум

МНК – метод наименьших квадратов

# РЕФЕРАТ

Дипломная работа: 81 страница, 37 рисунков, 35 источников, 1 приложение.

ШУМОПОДАВЛЕНИЕ В РЕЧЕВЫХ СИГНАЛАХ, ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЕ, РЕЖЕКТОРНЫЙ ФИЛЬТР, ФИЛЬТР КАЛМАНА, ФИЛЬТР МНК.

*Объект исследования* – шумоподавление в переговорах авиадиспетчеров.

*Цель работы* – очистка речевых сигналов от случайных шумов как традиционными методами (с использованием фильтров), так и с использованием вейвлет-преобразования и последующий анализ полученных результатов.

*Методы исследования* – вейвлетные преобразования.

В результате выполнения работы реализованы методы фильтрации речевых сигналов на основе фильтра МНК, режекторного фильтра, фильтра Калмана, на освнове вейвлет-преобразований Добеши, симлета, дискретной вейвлет-функции Мейера. Программная реализация и тестирование выполнена в системе компьютерной математики MATLAB.

**РЭФЕРАТ**

Дыпломная праца, 81 старонкі, 37 малюнкаў, 35 крыніц, 1 прыкладанне.

ШУМАПРЫГЛУШЭННЕ Ў МАЎЛЕНЧЫХ СІГНАЛАХ, ВЭЙВЛЕТ-ПЕРАЎТВАРЭННЕ, РЭЖЭКТАРНЫ ФІЛЬТР, ФІЛЬТР КАЛМАНА, ФІЛЬТР МНК.

*Аб'ект даследавання* – шумапрыглушэнне ў перамовах авіядыспетчараў.

*Мэта працы* – ачыстка маўленчых сігналаў ад выпадковых шумоў як традыцыйнымі (з выкарыстоўваннем фільтраў), так і з выкарыстоўваннем вэйвлет-пераўтварэння і наступны аналіз атрыманых вынікаў.

*Метады даследвання* – вейулетныя пераўтварэнні.

У выніку выканання работы рэалізаваны метады фільтрацыі маўленчых сігналаў на падставе фільтра МНК, рэжэктарнага фільтра, фільтра Калмана, на падставе вэйвлет-пераўтварэння Дабешы, сімлета, дыскрэтнай вэйвлет-функцыі Мэйера. Праграмная рэалізацыя выканана на мове MATLAB.

**ABSTRACT**

The diploma consists of 81 pages. It contains 37 images, 35 sources.

DENOISING IN THE SPEECH SIGNALS, WAVELET TRANSFORM, NOTCH FILTER, LSM FILTER, KALMAN FILTER.

*Object of research* – air traffic controllers negotiations denoising.

*Objective* – denoising of speech signals from random noise as by conventional methods (using filters) and also by using wavelet transform and subsequent analysis of the results.

*Methods of investigation* – wavelet transformations.

The speech signals filtering methods were implemented. These methods based on notch filter, LSM filter, Kalman filter, transform by Dobeshi wavelets, symlets and discrete wavelet function of Meyer. Software implementation is made in the computer mathematics system MATLAB.

# ВВЕДЕНИЕ

Речь – основной переносчик информации при естественном общении между людьми. Велика доля речевой информации и в общем объеме информации, передаваемой радиотехническими системами. Речевые сигналы, с которыми приходится иметь дело на практике, всегда в той или иной степени зашумлены.

При распространении в воздушном пространстве, а также при передаче по радиотехническим системам связи, речевые сигналы подвергаются искажениям. Сильные искажения способны не только ухудшить эстетическое качество воспринимаемых сигналов, но и снизить разборчивость речевого сообщения, т.е. привести к потере части информации. Существует значительное отличие между качеством и разборчивостью речевого сигнала, что приводит к необходимости отдельной оценки этих характеристик речевого сигнала. При наличии внешних шумов значительной интенсивности результаты анализа и распознавания речи существенно ухудшаются, поэтому разработка методов шумоподавления для речевых сигналов является актуальным направлением исследований.

Задача выделения полезного сигнала на фоне шума при обработке звуковых сообщений решается как с применением аппаратных средств, так и программными методами. Одним из наиболее популярных в последние годы программных методов является вейвлет-анализ, который открывает уникальные возможности распознавания тонких особенностей сигналов и функций, а также позволяет с высокой эффективностью проводить шумоподавление в сигналах различной природы.

Одной из особенно важных областей применения методов шумоподавления в речевых сообщениях является сфера авиаперевозок. С момента появления самолетов проблема распознавания речи во время переговоров авиадиспетчеров с пилотами остается актуальной.

Целью дипломной работы является очистка речевых сигналов от случайных шумов как традиционными методами (с использованием фильтров), так и с использованием вейвлет-преобразования, и последующий анализ полученных результатов.

# ГЛАВА 1

# МЕТОДЫ И ТЕХНОЛОГИИ ШУМОПОДАВЛЕНИЯ В РЕЧЕВЫХ СИГНАЛАХ

На сегодняшний день в области цифровой обработки речевых сигналов существует множество различных алгоритмов шумоподавления. Наиболее распространенной моделью шума для рассматриваемых алгоритмов является аддитивный белый гауссовский шум (АБГШ)

Современные системы записи звука работают очень чисто и тихо, и поэтому электронный шум в их цепях – очень редкая проблема. Однако, тот шум, который когда-то был приемлем, поскольку он маскировался шипением ленты, при современных технологиях звукозаписи и звукопередачи будет очень заметен. В процессе своей работы со звуковыми сигналами многие сталкиваются с шумами различной природы. Это может быть шипение и гул гитарных усилителей, какие-то возникающие цифровые щелчки, гудение осветительных ламп, кондиционеров, грохот уличного движения и т.д.

В настоящее время, производители программного обеспечения для обработки аудио предлагают множество вариантов для решения этой задачи [1]. Среди них такие инструменты, как гейты, экспандеры, и многие другие.

При передаче речевых сигналов по каналам связи при их распределенной обработке аддитивная смесь, проходя по трактам передачи, имеющим частотно зависимую передаточную характеристику, претерпевает дополнительные мультипликативные помехи. Известно, что универсального метода обработки, который одинаково эффективно справлялся бы с нестационарными и стационарными, аддитивными и мультипликативными шумами или существенно повышал бы качество и одновременно разборчивость речевых сигналов не существует [2].

Рассмотрим далее основные методы, которые используются при очистке речевых сигналов от шумов и для повышения качества речи.

## 1.1 Классификация методов повышения качества и разборчивости речи

На современном этапе наибольшее распространение получили следующие группы методов повышения качества и разборчивости речи [2]:

- методы адаптивной компенсации помех;

- методы, основанные на использовании математических моделей речевых сигналов во временной области (авторегресионная модель речевого сигнала);

- методы, основанные на использовании математических моделей речевых сигналов в частотной области (оценка минимальной среднеквадратической ошибки, марковские модели сигнала и шума);

- методы, основанные на использовании спектральных характеристик шума (вычитание амплитудных спектров, Винеровская фильтрация);

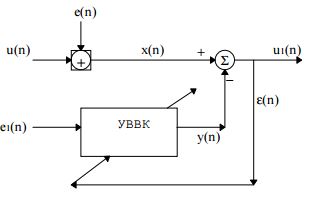
- методы, основанные на использовании моделей искусственных нейронных сетей;

- методы, основанные на моделях восприятия речи человеком;

### 1.1.1 Адаптивные компенсаторы помех

Этот класс методов цифровой обработки зашумленных сигналов основан на использовании одного или нескольких опорных сигналов, коррелированных с шумовым сигналом и некоррелированых (или слабо коррелированых) с полезным сигналом, подлежащим выделению. С помощью опорных сигналов формируется сигнал, который является оценкой помехи. Этот сигнал затем вычитается из зашумленного сигнала, и результат этой операции рассматривается как оценка незашумленного сигнала [3].

На рисунке 1.1 представлена схема адаптивного компенсатора помех, который использует один опорный сигнал.



u(n) – дискретный отсчет полезного сигнала в момент времени n; n=0,1,2…; e(n) – шумовой сигнал; e1(n) – опорный сигнал; ε(n) – сигнал ошибки; u1(n) – выходной сигнал компенсатора;

УУВК – устройство управления весовыми коэффициентами.

Рисунок 1.1 – Схема адаптивного компенсатора помех

Наиболее важной частью адаптивного компенсатора помех является устройство управления весовыми коэффициентами – линейный фильтр, через который пропускается опорный сигнал e1(n). Задача адаптивной компенсации помехи e(n) сводится к подбору коэффициентов фильтра таким образом, чтобы минимизировать энергию сигнала u1(n) на выходе компенсатора. В этом случае будет максимизировано выходное отношение сигнал/шум. Минимизация энергии обычно осуществляется на основе градиентных методов поиска экстремума функций многих переменных. Известно, что адаптивные компенсаторы помех позволяют значительно улучшить качество зашумленных сигналов – на несколько десятков децибел, но требование наличия опорного сигнала существенно сужает область их применения. Во многих приложениях цифровой обработки речевых сигналов (например, при реставрации архивных записей или в криминалистике), опорного сигнала, по крайней мере, в явном виде, не имеется. Поэтому для применения методов адаптивной компенсации помех опорный сигнал в таких случаях приходится получать на основе косвенных соображений, связанных с особенностями речевого сигнала, а сам адаптивный компенсатор в этом случае будет являться одной из составных частей более сложного алгоритма выделения речевого сигнала.

### 1.1.2 Методы, основанные на использовании статистических моделей речевых сигналов во временной области

Класс методов цифровой обработки зашумленных речевых сигналов, который основан на построении математических моделей речевых сигналов и обработке речевых сигналов с использованием этих моделей, быстро развивается, и в настоящее время эти методы приводят к самым успешным результатам. Задача выделения речевого сигнала из смеси с шумом, в случае использования достаточно адекватной модели, сводится к некоторой оценке параметров этой модели с последующим синтезом или фильтрацией речевого сигнала фильтром, построенным на основе или с помощью оцененных параметров [4].

Одними из наиболее перспективных методов в этом классе являются методы статистической фильтрации во временной области. Фильтрация речевого сигнала, моделируемого авторегрессией, осуществляется при этом методами теории оптимального оценивания, например, с помощью построения оптимального линейного фильтра – фильтра Калмана [5].

Вычислительно эффективная, но с менее удачным результатом обработки, реализация алгоритма фильтрации речевого сигнала, моделируемого авторегрессионной моделью с параметрами, связанными в марковскую цепь, предложена в [6]. Совместная оценка сигнала и параметров марковской цепи вычисляются рекуррентным способом с помощью алгоритма максимизации математического ожидания, причем для вычисления условного ожидания сигнала относительно наблюдений используется фильтр Калмана-Бьюси. Представленные в [7] результаты экспериментальных испытаний на речевом сигнале в смеси с некоррелированным аддитивным белым шумом с отношениями сигнал/шум 0, 10 и 20 дБ показывают увеличение отношения сигнал/шум в среднем на 4 дб.

### 1.1.3 Обработка речевого сигнала с использованием аппарата скрытых марковских моделей

Другим классом методов обработки зашумленных речевых сигналов являются методы, в которых речевой сигнал моделируется скрытой марковской цепью. Они основанны на использовании статистических моделей речевого сигнала. Известно, что традиционно используемые методы фильтрации, например, вычитание спектров или фильтр Винера, не используют фонетическую информацию, переносимую речевым сигналом. Исследования, проведенные авторами работы [8] показали, что знание и применение в процессе обработки фонетической структуры сигнала приводит к улучшению качества фильтрации. Поэтому вполне естественным является применение в процессе очистки речевого сигнала от шумов его статистической модели в виде скрытой марковской цепи, которая связана с фонетической структурой сигнала.

Идея реализации такого подхода заключается в том, что первоначально, по записям незашумленного речевого сигнала, строятся статистические модели единиц речевого потока (фонов либо более широких классов звуков). После того, как статистическая модель для множества состояний сигнала построена, по ней можно рассчитать оптимальный фильтр Винера.

При обработке зашумленного речевого сигнала сначала по отфильтрованному на предыдущем шаге сигналу оценивается текущее состояние марковской модели, в соответствии с которым выбирается оптимальный фильтр, который затем используется для фильтрации сигнала и получения очередной оценки.

В исследованиях авторов [9] сначала, используя стандартную базу данных (марковские модели формировались на TIMIT), строились модели состояний для незашумленного речевого сигнала. Для каждого состояния модели β и каждой гауссовской составляющей (реально использовалась только одна составляющая на состояние) оценивался оптимальный фильтр Винера Hβ(θ). В процессе обработки сигнала на каждом шаге (кадре анализа) с помощью процедуры Витерби вычислялись правдоподобия состояний, в соответствии с которыми выбирались весовые коэффициенты Wβ для каждого фильтра Hβ(θ). Очистка сигнала затем производилась в частотной области в соответствии с (1.1):

, 1.1)

где k – номер итерации (кадра анализа), Wβ – вес фильтра Hβ(θ), z(iw) – спектральная компонента зашумленного сигнала и y(iw) – спектральная компонента обработанного сигнала.

Для эффективной обработки нестационарных сегментов отдельно оценивается марковская модель шума. В отличие от простых моделей состояний полезного сигнала шум моделируется набором состояний, каждое из которых содержит несколько гауссовских компонент.

Во время обработки зашумленного сигнала при определении отсутствия полезного сигнала выполняется декодирование сегмента паузы процедурой Витерби для выбора оптимальной модели шума. Модель шума, обеспечивающая максимальное правдоподобие наблюдаемой последовательности используется далее для обработки сигнала.

Следует отметить, что очевидным недостатком подхода является необходимость иметь априорную информацию о возможных типах шумов, например, в виде предварительно обученных марковских моделей состояний. Типов возможных шумов для различных практически важных условий много, и требование наличия заранее вычисленных моделей представляется маловыполнимым. Кроме того, качество обработки сигнала ухудшается в тех случаях, когда помеха имеет существенно нестационарный характер.

В связи с этим дальнейший прогресс в этом направлении может быть достигнут за счет использования более гибких способов моделирования помех. Так, в работе [10] предлагается эффективный алгоритм для переоценки параметров вероятностной модели шума, основанный на адаптивной подстройке элементов кодовой книги, которая состоит из матрицы корреляционных коэффициентов для авторегрессионной скрытой марковской модели, по методу скользящего среднего. Показано, что по сравнению с исходным алгоритмом фильтрации, основанным на марковской модели [11], подобная методика приводит к дополнительному выигрышу примерно в 2,3 дБ на нестационарных шумах и не ухудшает качества обработки сигнала со стационарными помехами.

Использование марковских моделей речевого сигнала оказывается выигрышным при наличии априорной информации синтаксического характера о зашумленном речевом сигнале. Довольно часто, например, при анализе черных ящиков с борта самолета или реставрации аудиозаписей, время обработки сигнала не играет определяющей роли. В тех случаях, когда аудитор может приблизительно указать, дать гипотезы, что именно было произнесено, или на что похоже зашумленное высказывание, можно улучшить качество сигнала с помощью использования аппарата марковских моделей, построенных на предсказанных звуках.

Анализ метода фильтрации речевых сигналов с помощью марковских

моделей при наличии гипотез о произносимом тексте выполнен в работе [8]. Продемонстрировано, что в результате обработки улучшается качество звучания сигнала в условиях различных типов помех и в широком диапазоне отношений сигнал/шум.

### 1.1.4 Методы, основанные на использовании, отдельных характерных свойств речевого сигнала

К методам этого типа относится, прежде всего, класс методов обработки зашумленных речевых сигналов, которые используют квазипериодичность речевого сигнала. Этот класс можно разбить на группы.

Первая группа методов использует периодичность речевых сигналов для построения адаптивного компенсатора помех, с помощью которого обрабатывается зашумленный речевой сигнал. Предполагается, что исходный речевой сигнал s(n) строго периодичен с периодом T, кратным частоте дискретизации, а случайный аддитивный шум v(n) некоррелирован с s(n). В качестве опорного сигнала для адаптивной компенсации помехи используется выражение (1.2):

(1.2)

Результаты применения этого метода приводятся в работе [12]. Показано, что отношение сигнал/шум может быть увеличено на 7 - 10 дБ, однако разборчивость отфильтрованной речи при этом несколько понижается. Под разборчивостью речи в данном случае понимается степень, с которой слушатели могут понять смысл фразы, идентифицировать слова, слоги и фонемы.

Вторая группа методов, использующих периодичность звонких звуков, основана на представлении сигнала в кепстральной области. В этом случае периодический характер речевого сигнала используется для синтеза адаптивной гребенки фильтров [13]*.*

Как известно, периодичность звонких звуков выражается в частотной области в том, что их спектр имеет линейчатый характер, причем соседние пики (спектральные максимумы) отстоят друг от друга на интервал в частотной области равный частоте основного тона. Поэтому, если гребенка фильтров такова, что гармоники основного тона (спектральные пики) попадают в полосы пропускания, то можно рассчитывать на повышение качества речевого сигнала. Во временной области фильтрацию речевого сигнала гребенкой фильтров с равноразнесенными по частоте каналами можно представить соотношением (1.3) [13]:

(1.3)

Исследования адаптивной гребенки фильтров показали, что достигаемое улучшение качества речевого сигнала невелико, а в тех случаях, когда помеха носит структурированный характер, фильтрация такого рода вообще неэффективна. Усовершенствованная модификация рассмотренного выше подхода предложена Кохом и Малахом [12] и заключается в том, что коэффициенты a(k) в выражении (1.3) зависят не только от k, но и от n (modT).

Исследования, проведенные на синтетических гласных звуках, показали, что при надлежащем выборе взвешивающих коэффициентов a(k,n) можно добиться значительного эффекта для улучшения восприятия речи в тех случаях, когда помеха или шум являются структурированными [14]*.*

Обобщение метода, описываемого выражением (1.3) является также известный адаптивный фильтр Фрезиера [15]*.* В этом случае учитывается изменение частоты основного тона на интервале времени, равном длине импульсной характеристики гребенки фильтров. Исследование характеристик такого фильтра показали, что он может дать выигрыш в отношении сигнал/шум до 10 дБ, однако при этом несколько снижается разборчивость речи.

Описанные выше методы имеют ряд недостатков. Так, все методики, кроме фильтра Фрезиера, не учитывают изменений частоты основного тона. Кроме того, эти методы непригодны для фильтрации глухих звуков. Например, существенные для разборчивости речевого сигнала звуки “п”, “т”, “к” не могут быть успешно обработаны такими методами. Наконец, качество обработки сигнала зависит от точности оценки частоты основного тона в зашумленном речевом сигнале, что само по себе не всегда возможно [15].

То обстоятельство, что на разборчивость речи существенно влияет правильное восприятие согласных звуков, в частности, взрывных “п”, “т”, “к” и шумного “с”, было использовано в системе [16]*.* Фильтрация речевого сигнала заключалась в том, что перед взрывными звуками вставлялась короткая пауза – смычка, – а согласный “с” фильтровался специально подобранным фильтром. Испытания показали, что такая обработка повышает разборчивость речевых сигналов. Существенным недостатком предложенной методики является необходимость в априорной информации о местонахождении взрывных звуков и звука “с”.

Работа [16] стала первой работой по фильтрации речевого сигнала, в которой предложено обрабатывать различные типы звуков разными процедурами оценивания параметров и фильтрации. Было выделено пять категорий звуков – фрикативные, взрывные, назальные, гласные и глайды.

### 1.1.5 Методы, основанные на искусственных нейронных сетях

Разработка аппарата искусственных нейронных сетей привела к появлению нового типа алгоритмов для обработки зашумленных речевых сигналов, основанных на использовании моделей нейронных сетей.

Исследование свойств многослойного персептрона как нелинейного фильтра во временной области выполнено в [6]. В качестве помехи рассматривались аддитивный гауссовский шум и нелинейный шум, моделирующий артефакты низкоскоростного CELP кодера. Персептрон обучался на зашумленном сигнале, роль сигнала-учителя выполнял чистый сигнал (доступный на этапе обучения). В качестве речевого материала использовались записи гласного “e” (120 записей от 40 дикторов).

Наилучшие результаты были продемонстрированы на трехслойном персептроне (на каждом слое по 60 нейронов): улучшение отношения сигнал/шум составило 3 дБ для шума кодека и 6 дБ для белого шума при начальном уровне сигнал/шум 7,4 дБ. Эти результаты, показывают, что пока выигрыша от использования многослойного персептрона по сравнению со стандартными методиками нет, хотя в более ранней работе [17] утверждалось, что многослойный персептрон с успехом может быть использован для фильтрации гауссовского шума.

### 1.1.6 Метод оценивания минимальной среднеквадратической ошибки

Описываемый далее алгоритм метода оценивания минимальной среднеквадратической ошибки, как и метод вычитания спектров, основан на оценке амплитудного спектра сигнала. Среди других методов фильтрации, предполагающих наличие только одного микрофона, алгоритмы, основанные на минимуме среднеквадратической ошибки, являются одними из наиболее полезных. Их использование приводит к значительному сокращению уровня шума в сигнале без внесения остаточных искажений типа музыкальных тонов. В представленных в работе [18]исследованиях показано, что в значительной мере превосходство метода оценивания минимальной среднеквадратической ошибки над методиками типа Винеровской фильтрации или вычитания амплитудных спектров связано именно с введением априорной оценки сигнал/шум в каждой спектральной полосе. В связи с этим были предложены модификации стандартных подходов, таких как винеровская фильтрация, вычитание амплитудных спектров и оценки максимального правдоподобия, использующие априорные отношения сигнал/шум, что приводит к существенному улучшению результатов фильтрации.

Существуют и другие подходы, например, слепое разделение речевых сигналов (Blind Speech Separation), объединяющее методы получения оценки сигналов источников на основе лишь знания сигналов смесей и некоторых общих предположений о свойствах сигналов источников и процессе смешивания [19]. В ряде случаев эти методы позволяют достичь значительного улучшения качества и разборчивости речи.

Далее рассмотрены различные инструменты и технологии, предназначенные для борьбы с шумами, а также их достоинства и недостатки в применении к тому или иному материалу [2].

## 1.2 Гейты

Гейт (англ. noise gate – шумовые ворота) – шумовой гейт, или просто гейт, это электронное устройство или плагин динамической обработки, используемый для контроля уровня звукового сигнала. Гейт пропускает или глушит сигнал в зависимости от установленного порогового значения. Он очень часто используется для подавления шума в паузах [20].

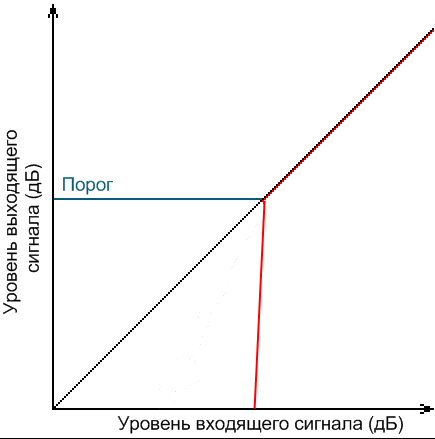


Рисунок 1.3 – Передаточная характеристика гейта

Гейт пропускает сигнал только тогда, когда его уровень выше установленного порога: гейт "открыт". Если уровень сигнала ниже порогового значения, гейт его не пропустит (или значительно ослабит): гейт "закрыт". Гейт рационально использовать тогда, когда уровень полезного сигнала выше уровня шума. Его порог устанавливается выше уровня шума и поэтому, когда нет полезного сигнала (выше порога), гейт закрыт. Гейт не очищает сигнал от шума. Когда гейт открыт, оба сигнала проходят через него, будь то полезный сигнал или шум. Схема гейта представлена на рисунке 1.4.

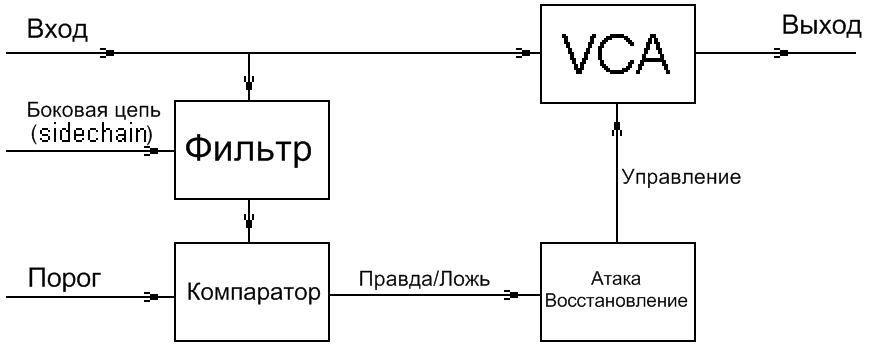


Рисунок 1.4 – Cхема гейта

Классический гейт с минимумом настроек посылает сигнал на компаратор, который сравнивает его уровень с пороговым значением (Threshold) и подает открывающий сигнал на блок управления громкостью, если сигнал выше порога. В результате гейт не пропускает весь лишний шум, идущий в паузах между полезными сигналами.

В настоящее время существуют гейты с иным устройством. Чаще всего перед компаратором стоит фильтр, который позволяет контролировать процесс, отдельной полосой частот спектра. Для более плавного включения/выключения и более точного контроля, в схеме присутствуют огибающие «attack»/«hold»/«release». Также вместо полного закрытия гейта можно указать, на сколько децибел уменьшится уровень шумового сигнала, если компаратор подаёт закрывающий сигнал. На вход компаратора и самого гейта можно подавать разные сигналы, чтобы управлять одним каналом с помощью другого. Есть возможность "переворачивания" гейта, тогда он будет закрываться при превышении порога, а открываться при сигналах с уровнем ниже порогового значения. Эта функция чаще всего применяется для того, чтобы глушить фон в то время, когда говорит диктор [1].

Чтобы возникла маскировка, требуется, чтобы полезный сигнал был громче, чем шум в том же частотном диапазоне. Особенно это актуально для диапазона 500 Гц – 5 кГц, к которому наши уши наиболее чувствительны. Так, если полезный сигнал это низкочастотный звук, а шум – это высокочастотный визг, то шум будет очень заметен. Можно улучшить ситуацию, если воспользоваться ВЧ-фильтром, хотя это обычно повреждает высокочастотные детали полезного сигнала.

Таким образом, видно, что гейт эффективен только лишь в двух случаях: если уровень нежелательных шумов довольно низок или если полезный сигнал маскирует шум в то время, когда он открыт. Чем больше уровень фонового шума или утечек, тем менее естественное звучание будет в итоге.

Стоит заметить, что в большинстве случаев полное удаление шумов нежелательно, поскольку это привлекает внимание и оказывает некоторое раздражающее воздействие. На практике, более тонкая обработка будет эффективнее и менее заметной даже при условии, что шумы будут подавляться не настолько сильно. Например, уменьшая в гейте диапазон подавления до 12 дБ, в результате чего он не закрывается полностью, а просто понижает уровень на это значение, получится более приятное звучание, нежели при полной отсечке шумов.

## 1.3 Экспандеры

Экспандер (англ. expander – расширитель) – устройство, расширяющее динамический диапазон сигнала.

Существует три разновидности экспандеров [20]:

* Повышающий – увеличивает уровень сигнала, превышающего пороговое значение. По принципу работы и устройству идентичен компрессору, соотношение которого меньше единицы. Зачастую эти два прибора объединены с помощью расширенной шкалы соотношения. При превышении порога, как и в компрессоре, сигнал будет сжат в соотношении меньшем единицы, к примеру 0,5:1, то есть если сигнал превысил пороговое значение на 5 дБ, то его уровень будет 10 дБ (5/0,5 = 10).
* Понижающий – уменьшает уровень сигнала, находящийся ниже порога. Данный тип также подобен компрессору, но работающему в обратную сторону. Компрессор сжимает уровень сигнала, когда тот превышает пороговое значение, а экспандер сжимает сигнал, когда он опускается ниже порогового значения. Если такому экспандеру задать крайние значения соотношения (60:1), то он сможет работать и как гейт. На рисунке 1.5 показаны передаточные характеристики экспандеров повышающего и понижающего типов.

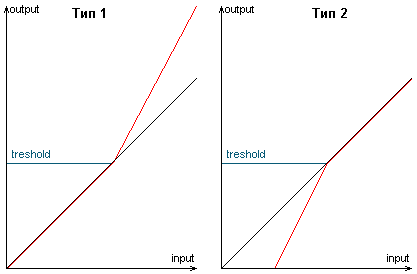


Рисунок 1.5 – Передаточные характеристики двух типов экспандера

* Компандер – компрессор-экспандер, по сути это два прибора в одном, работающие на уменьшение уровня сигнала и выше и ниже порога(-ов) срабатывания.

Сфера применения экспандеров примерно такая же, как и гейтов, с аналогичными ограничениями на использование. Экспандер может быть не настолько резким, как гейт, но как только настроена атака и восстановление, различий в их работе услышано не будет, если только не используется очень низкое отношение.

В более сложных приборах, например, в сайдчейне, имеется специальный детектор, динамически управляющий частотой встроенного НЧ-фильтра. Это повышает эффективность подавления высокочастотных шумов, например, шипения, поскольку детектор понижает частоту фильтра, как только амплитуда полезного сигнала спадает. Как и гейт, эта технология полагается на маскировку шума полезным сигналом.

Динамические фильтры имеются в таких устройствах, как Symmetrix 511, 511A и Drawmer DF320, но нет плагинов, в которых был бы реализован их эквивалент. Можно попробовать воссоздать этот процесс в DAW, используя автоматизацию, управляющую частотой НЧ-фильтра. Аппаратные процессоры включают в себя экспандер, вступающий при очень низких уровнях, чтобы подавить оставшийся во время пауз шум. Этот подход может быть полезен для подавления шипения ленты или электрической гитары, но он не помогает при низкочастотных шумах.

Более сложный подход использует разделение сигнала на множество частотных полос и организацию многополосных экспандеров или гейтов. К примеру, компания Dolby реализовала это в виде четырёхполосного устройства CAT43 (на базе системы шумоподавления Dolby-A). Сейчас этот процессор воссоздан компанией Waves в форме плагина. Четыре слайдера настраивают порог экспандера для каждой полосы, а пятый управляет мастер-секцией. Он плохо работает с широкополосными шумами, но хорошо справляется с «компактными» механическими звуками, такими, как шум камеры или свист осветительных приборов. «Тяжёлая» обработка затронет полезные звуки и создаст различные тональные артефакты, но там, где шумовое загрязнение имеет низкий или умеренный уровень, он может работать хорошо. Этот процессор стал очень популярен для очистки диалогов, записанных в полевых условиях [1].

## 1.4 Цифровые решения

Методики, описанные выше, изначально были реализованы в аналоговой форме, но возможности цифровых технологий позволили выдвинуть системы шумоподавления на новый уровень. Самые ранние цифровые процессоры были подобны уже CAT43, только использовали намного больше полос (до нескольких сотен). Многие известные плагины, включая широкополосную секцию в BIAS Sound Soap Pro, работают аналогичным способом. Серия процессоров CEDAR DNS, очень популярных в сфере вещания и кинопроизводства для обработки диалогов, тоже использует многополосную динамику для удаления шумов [1].

Вручную невозможно настроить пороги для сотен полос экспандера, поэтому многие системы позволяют снимать «отпечаток» шума, взятый из короткой секции аудиозаписи, содержащей только шум. Процессор анализирует его и автоматически настраивает пороги. Для корректного анализа требуется шумовой участок продолжительностью около секунды. Если же в звуковом материале отсутствует такой «солирующий» шум, некоторые системы предлагают набор типичных шумовых профилей, среди которых можно выбрать подходящий и потом подстроить под конкретный материал. У плагинов есть графический дисплей, который показывает шумовой порог и спектр полезного сигнала. Также, могут быть дополнительные инструменты для настройки формы кривой шумового профиля, что позволяет оптимизировать работу шумоподавления и добиться минимальных побочных эффектов.

Вышеупомянутый подход хорошо работает тогда, когда уровень и спектр шума постоянны. В противном случае, шумовой отпечаток быстро теряет свою актуальность, в результате чего шум становится снова заметен, либо теряется часть полезного сигнала. Но эффективность этих процессоров всё же может быть высока. Так, если в определённом частотном диапазоне не будет присутствовать никакого полезного звука, то фильтры в соответствующих полосах будут закрыты. Таким образом, процесс удаления шумов может быть очень эффективен, конечно, при условии, что полезный сигнал способен маскировать шум, находящийся в той же области спектра.

Недостатки, возникающие при обработке сильно зашумлённого сигнала, включают в себя такие вещи, как смазывание и потеря высокочастотных деталей, а также появление новых звуков, таких как журчание (это связано с тем, что большое количество полос в широком диапазоне быстро включается и выключается, затрагивая полезный сигнал). Последний эффект может быть минимизирован при помощи новых алгоритмов, использующих интеллектуальное линкование между частотными полосами. Хотя, всё это завязано на уровень применённого подавления. В большинстве случаев, при подавлении на 10 дБ и менее получается хороший результат [1].

## 1.5 Изменяющийся фоновый шум

Шум, уровень и тональность которого постоянно изменяются, можно убрать, используя методы, которые разработали такие компании, как CEDAR Audio и Sonic Solutions, распознающие шум среди полезного сигнала и создающие динамический профиль, отслеживающий его [1]. Работают данные алгоритмы на основе нейронных сетей. Один из методов, используемых в алгоритмах – это автокорреляция. Суть его в том, что происходит поиск постоянных компонентов в изменяющемся сигнале. И чем дольше время анализа, тем лучше будет итоговое разделение полезного сигнала и шума.

Точно так же, чем быстрее изменяются характеристики шума, тем с худшим качеством программа сможет его отследить. Это особенно характерно для программного обеспечения, работающего в автоматическом режиме: если возникает резкое изменение шумового характера, оно может стать слышимым на несколько секунд, пока система реагирует и корректирует параметры шумового профиля.

Обзор методов и инструментов повышения качества и разборчивости зашумленных речевых сигналов показывает, что существует много различных подходов к обработке зашумленной речи. Такое разнообразие методов обусловлено как важностью проблемы, так и отсутствием достаточно надежных методов ее решения.

Объективное сравнение этих методов, инструментов и выбор наиболее приемлемых сделать весьма затруднительно, так как перед системами коррекции речевых сигналов ставятся различные задачи. Например, можно в качестве главного критерия использовать повышение разборчивости речи, допуская при этом возможность искажений в тембре голоса или появление артефактов в виде структурированного шума. Можно поставить целью понижение утомляемости аудитора или сохранение натуральности голоса диктора, что достигается в основном за счет повышения качества речевого сигнала. Наконец, могут быть известны заранее важные априорные сведения, например тип или параметры шума, характеристики голоса диктора, наконец, гипотезы о произносимом тексте, что также может определяющим образом повлиять на выбор метода фильтрации.

Важно отметить, что универсальных методов обработки, которые одинаково хорошо боролись бы с существенно нестационарными и стационарными, аддитивными и мультипликативными шумами, значительно повышали бы качество и одновременно разборчивость речи, сейчас нет, и в ближайшем будущем, возможно, не будет. Наблюдается и такая тенденция, что если сравнивать системы обработки зашумленной речи по двум показателям – повышению качества звучания речевых сигналов и повышению разборчивости, – то системы, повышающие качество и натуральность звучания, скорее снижают разборчивость, и наоборот, повышение разборчивости приводит к понижению качества и натуральности звучания. Поэтому многие из вышеназванных методов фильтрации следует рассматривать как взаимодополняющие, и в идеальном случае необходимо иметь библиотеку из нескольких методов фильтрации.

Рассматривая последние тенденции в области обработки зашумленных сигналов, следует особенно выделить высокие результаты, полученные за счет использования математических моделей речевых сигналов (авторегрессионные скрытые марковские модели), а также использование вейвлет-преобразований для фильтрации аддитивных стационарных шумов.

# ГЛАВА 2

# ВЕЙВЛЕТЫ И ИХ ПРИМЕНЕНИЕ ДЛЯ ШУМОПОДАВЛЕНИЯ В РЕЧЕВЫХ СИГНАЛАХ

## 2.1 Некоторые сведения из теории вейвлетных преобразований

Вейвлеты представляют собой особые функции в виде коротких волн (всплесков) с нулевым интегральным значением и с локализацией по оси независимой переменной (t или x), способных к сдвигу по этой оси и масштабированию (растяжению/сжатию). Любой из наиболее часто используемых типов вейвлетов порождает полную ортогональную систему функций. В случае вейвлет-анализа (декомпозиции) процесса (сигнала) в связи с изменением масштаба вейвлеты способны выявить различие в характеристиках процесса на различных шкалах, а посредством сдвига можно проанализировать свойства процесса в различных точках на всем исследуемом интервале. Именно благодаря свойству полноты этой системы можно осуществить восстановление (реконструкцию или синтез) процесса посредством обратного вейвлет-преобразования.

Вейвлет-преобразование (ВП) – преобразование, похожее на преобразование Фурье с совершенно иной оценочной функцией. Основное различие лежит в следующем: преобразование Фурье раскладывает сигнал на составляющие в виде синусов и косинусов, т.е. функций, локализованных в Фурье-пространстве, т.е. в частотной области; напротив, вейвлет-преобразование использует функции, локализованные как в реальном (пространственно-временном представлении), так и в в Фурье-пространстве. В общем случае вейвлет-преобразование может быть выражено следующим уравнением:

(2.1)

где \* - символ комплексной сопряженности, функция ψ – некоторая функция уникальная для каждого вида вейвлета [21].

Классическое преобразование Фурье (ПФ) является традиционным математическим аппаратом для анализа стационарных процессов. При этом сигналы разлагаются в базисе косинусов и синусов или комплексных экспонент. Эти базисные функции простираются вдоль всей оси времени.

С практической точки зрения и с позиций точного представления произвольных сигналов ПФ имеет ряд ограничений и недостатков. Обладая хорошей локализацией по частоте, оно не обладает временным разрешением. ПФ даже для одной заданной частоты требует знания сигнала не только в прошлом, но и в будущем, а это – теоретическая абстракция.

Обусловлено это тем, что базисной функцией при разложении в ряд Фурье является гармоническое колебание, которое математически определено на временном интервале от  до . ПФ не учитывает, что частота колебания может изменяться во времени.

Локальные особенности сигнала (разрывы, ступеньки, пики и т.п.) дают едва заметные составляющие спектра, по которым обнаружить эти особенности, и тем более их место и характер, практически невозможно. В этом случае невозможно и точное восстановление сигнала из-за появления эффекта Гиббса [22].

Для получения о сигнале высокочастотной информации с хорошей точностью следует извлекать ее из относительно малых временных интервалов, а не из всего сигнала, а для низкочастотной спектральной информации – наоборот.

Кроме того, на практике не все сигналы стационарны, а для нестационарных сигналов трудности ПФ возрастают многократно.

Часть указанных трудностей преодолевается при использовании оконного ПФ:

 (2.2)

в котором применяется предварительная операция умножения сигнала  на «окно» ; при этом окном является локальная во времени функция (например, прямоугольная, т.е.  при  и  при  и ), перемещаемая вдоль оси времени  для вычисления ПФ в разных позициях . В результате получается текущий спектр, т.е. частотно-временное описание сигнала. Схематично применение оконного преобразования Фурье к сигналу показано на рисунке 2.1.



Рисунок 2.1 – Оконное преобразование Фурье

Если окно, показанное на рисунке 2.1, перемещать скачками (через ) вдоль всего времени существования сигнала , то за некоторое число таких перемещений возможен «просмотр» всего сигнала.

Таким образом, вместо обычной спектрограммы получится набор спектрограмм, схематично представленный в виде прямоугольников на рисунке 2.2, а. Такой спектральный анализ равносилен анализу с помощью набора фильтров с постоянной шириной полосы пропускания, равной .

Поскольку каждое окно выделяет свой небольшой участок во времени, точность представления и разрешающая способность по времени могут быть повышены. Но ввиду принципа неопределенности () невозможно получить одновременно высокое разрешение и по частоте, и по времени. Узкому во времени окну будет соответствовать низкое разрешение по частоте (большая величина ), и наоборот.

Недостаток оконного ПФ состоит в том, что оно использует фиксированное окно и, следовательно, фиксированное разрешение по времени и частоте для всех точек плоскости преобразования, как видно из рисунка 2.2, а, которое не может быть адаптировано к локальным свойствам сигнала.



Рисунок 2.2 – Сравнение оконного ФП и оконного BП

ВП имеет существенное преимущество перед ПФ, прежде всего, за счет свойства локальности у вейвлетов. В вейвлет-преобразовании операция умножения на окно как бы содержится в самой базисной функции, которая сужает и расширяет окно, как показано на рисунке 2.2, б: с ростом параметра *a* увеличивается разрешение по частоте и уменьшается разрешение по времени, а с уменьшением этого параметра уменьшается разрешение по частоте и увеличивается по времени.

Отсюда появляется возможность адаптивного к сигналу выбора параметров окна. Подвижное частотно-временное окно одинаково хорошо выделяет и низкочастотные, и высокочастотные характеристики сигналов. Это свойство ВП дает ему большое преимущество при анализе локальных свойств сигналов.

Можно локально реконструировать сигнал: реконструировать только часть сигнала или выделить вклад определенного масштаба. Если вейвлет-коэффициенты подвержены случайным ошибкам, они будут действовать на реконструируемый сигнал локально вблизи положения возмущения, а ПФ распространяет ошибки по всему восстанавливаемому сигналу.

Именно благодаря выявлению локальных особенностей сигнала, принципиально отсутствующему у ПФ, ВП нашло широкое применение в анализе тонкой структуры сигналов и изображений, их сжатии и очистке от шума, что важно в радиотехнике, электронике, гидроакустике, геофизике, медицине и других областях науки и техники.

Стоит отметить, что ВП не является полной заменой традиционного преобразования Фурье и не умаляет его достоинств и значимости при работе со стационарными процессами, а также когда нет необходимости исследовать локальную структуру сигналов [21].

Как правило, вейвлет-преобразования (ВП) подразделяют на дискретное (ДВП) и непрерывное (НВП). ДВП используется для преобразований и кодирования сигналов, НВП - для анализа сигналов. Вейвлет-преобразования в настоящее время принимаются на вооружение для огромного числа разнообразных применений, включая молекулярную динамику, квантовую механику, астрофизику, геофизику, оптику, компьютерную графику и обработку изображений, анализ ДНК, исследования белков, исследования климата, общую обработку сигналов и распознавание речи.

Вейвлетный анализ представляет собой особый тип линейного преобразования сигналов и физических данных. Базис собственных функций, по которому проводится вейвлетное разложение сигналов, обладает многими специфическими свойствами и возможностями. Вейвлетные функции базиса позволяют сконцентрировать внимание на тех или иных локальных особенностях анализируемых процессов, которые не могут быть выявлены с помощью традиционных преобразований Фурье и Лапласа. К таким процессам в обработке сигналов относят процессы распознавания образов, а, например, в геофизике – поля различных физических параметров природных сред. В первую очередь это касается полей температуры, давления, профилей сейсмических трасс и других физических величин.

По локализации во временном и частотном представлении вейвлеты занимают промежуточное положение между гармоническими функциями, локализованными по частоте, и функцией Дирака, локализованной во времени.

Теория вейвлетов не является фундаментальной физической теорией, но она дает удобный и эффективный инструмент для решения многих практических задач. Основная область применения вейвлетных преобразований – анализ и обработка сигналов и функций, нестационарных во времени или неоднородных в пространстве, когда результаты анализа должны содержать не только частотную характеристику сигнала (распределение энергии сигнала по частотным составляющим), но и сведения о локальных координатах, на которых проявляют себя те или иные группы частотных составляющих или на которых происходят быстрые изменения частотных составляющих сигнала. По сравнению с разложением сигналов на ряды Фурье, вейвлеты способны с гораздо более высокой точностью представлять локальные особенности сигналов, вплоть до разрывов 1-го рода (скачков). В отличие от преобразования Фурье, вейвлет-преобразование одномерных сигналов обеспечивает двумерную развертку, при этом частота и координата рассматриваются как независимые переменные, что дает возможность анализа сигналов сразу в двух пространствах.

Одна из главных и особенно плодотворных идей вейвлетного представления сигналов на различных уровнях разложения заключается в разделении функций приближения к сигналу на две группы: аппроксимирующую – грубую, с достаточно медленной временной динамикой изменений, и детализирующую – с локальной и быстрой динамикой изменений на фоне плавной динамики, с последующим их дроблением и детализацией на других уровнях разложения сигналов. Это возможно как во временной, так и в частотной областях представления сигналов вейвлетами.

**2.1.1 Достоинства и недостатки вейвлетных преобразований**

* Вейвлет-преобразования обладают всеми достоинствами преобразований Фурье.
* Вейвлетные базисы могут быть хорошо локализованными как по частоте, так и по времени. При выделении в сигналах хорошо локализованных разномасштабных процессов можно рассматривать только те масштабные уровни разложения, которые представляют интерес.
* Вейвлетные базисы, в отличие от преобразования Фурье, имеют много разнообразных базовых функций, свойства которых ориентированы на решение задач различного направления. Базисные вейвлеты могут реализоваться функциями различной гладкости.
* Недостатком вейвлет-преобразований является их относительная сложность.

Вейвлеты применяются в общих задачах, связанных с обработкой информации, таких как очистка сигнала от помех, сжатие данных, выявление кратковременных и глобальных закономерностей, спектральный анализ составляющих сигнала, а также статистическая обработка, подавление избыточной информации, криптография и стеганография, обработка мультимедийной информации.

## 2.2 Применение вейвлетов для шумоподавления, определения наличия полезного сигнала, повышения отношения сигнал/шум

При вейвлет-анализе сигнал разлагается на аппроксимирующие коэффициенты, представляющие сглаженный сигнал, и детализирующие коэффициенты, описывающие колебания. Информация о шумовой компоненте в основном содержится в детализирующих коэффициентах, поэтому обычно именно они обрабатываются при очистке от шумов.

В обычной ситуации предполагается, что сигнал шумовой компоненты по модулю меньше основного. При выполнении этого условия простейший способ очистки от шума заключается в присвоении нулевых значений коэффициентам, не превышающим некоторого порогового значения. Эта процедура называется пороговой обработкой (трешолдингом) коэффициентов. Главным преимуществом трешолдинга состоит в широких возможностях вариации параметров обработки путем использования различных типов и способов пороговой обработки и их параметров. При использовании процедуры мягкого вида трешолдинга все коэффициенты детализации бóльшие или равные порогу на i-м уровне разложения уменьшаются на величину порога, все остальные коэффициенты приравниваются к нулю. При использовании процедуры жесткого вида трешолдинга неизменными сохраняются все коэффициенты детализации, бóльшие или равные порогу на этом уровне, а все остальные коэффициенты приравниваются к нулю.

От выбора порогового уровня фона (оценки дисперсии шума) зависит качество шумоподавления, оцениваемое в виде отношения сигнал/шум.

Процедура очистки сигнала от шума состоит из трех шагов [23]:

1. Декомпозиция. Выбирается вейвлет и уровень разложения N. Вычисляется вейвлет-разложение исходного сигнала до уровня N.

2. Детализация. Для каждого уровня от 1 до N выбирается определенный порог и применяется пороговая обработка детализирующих коэффициентов.

3. Реконструкция. Производится вейвлет-реконструкция, основанная на первоначальных аппроксимирующих коэффициентах уровня N и модифицированных детализирующих коэффициентах уровней от 1 до N.

При реализации быстрых алгоритмов разложения сигнала в базисе вейвлет-функций длину выборки целесообразно выбирать равной степени числа 2 (), так как переход от одного уровня разрешения к другому (более детальному) сопровождается уменьшением вдвое длины выборки [24].

Пусть задан временной ряд . При прохождении сигналом  НЧ-фильтра с характеристикой  полученный выходной сигнал представляет собой свертку входного сигнала и характеристики фильтра:

|  |  |
| --- | --- |
|  | (2.3) |

Характеристика взаимосвязанного ВЧ-фильтра  задается следующим образом:

|  |  |
| --- | --- |
|  | (2.4) |

где постоянная  определяет длину области задания вейвлета. Увеличение  позволяет использовать более регулярные функции (которые имеют  нулевых моментов) [25], что обеспечивает возможность более сильного сжатия сигнала, а также сглаживания ошибок фильтрации при его восстановлении. Наряду с этим, увеличение  сопровождается существенным ростом числа коэффициентов фильтра, что также имеет свои недостатки. Наиболее очевидный – существенное увеличение времени вычислений, что может быть нежелательно, например, в задачах кодирования и передачи информации. Осциллирующие остатки – «хвосты» – вейвлет-функций при больших  служат дополнительным недостатком при решении ряда задач [26] . Поэтому выбор вейвлета должен осуществляться в зависимости от приоритетов при проведении цифровой обработки сигналов.

В большинстве практических задач в качестве базисных функций для реализации одномерного дискретного вейвлет-преобразования (1D-ДВП) применяют вейвлеты Добеши. Они были введены Ингред Добеши в конце 1980-х годов и представляют собой семейство функций, часто используемых в различных приложениях. Для проведения расчетов предложены специальные процедуры быстрого вейвлет-преобразования на основе данных вейвлетов.

Для *M ∈* N, вейвлет Добеши *D*2*M* представляет собой функцию вида *ψ* = *M ψ∈ L*2(R), которая определяется выражением:

|  |  |
| --- | --- |
|  | (2.5) |

где *∈* R – постоянные коэффициенты, удовлетворяющие условию:

|  |  |
| --- | --- |
|  | (2.6) |

а для  выполняется требование:

|  |  |
| --- | --- |
|  | (2.7) |

Пример скейлинг-функции и вейвлета Добеши приведен на рисунке 2.1. При малых *М* коэффициенты  могут быть записаны в виде точных выражений, например, вейвлет *D*4 задается коэффициентами *hk*:

|  |  |
| --- | --- |
|  | (2.8) |
|  |

При увеличении *М* определение  проводится путем решения алгебраических уравнений степени *М.* Несмотря на то, что эти уравнения могут быть решены с любой требуемой точностью, значения  с заранее определенным числом разрядов задаются в виде таблиц. Например, набор коэффициентов ВЧ-фильтра, с помощью которого определяется широко применяемый на практике вейвлет *D*8, приведен в выражении (2.9):

|  |  |
| --- | --- |
|  | (2.9) |

В практических вычислениях обычно используются коэффициенты с точностью до 32-х или 64-х десятичных разрядов.

После однократного прохождения сигналом  квадратурных зеркальных фильтров с характеристиками  и , осуществляется прореживание выходных сигналов, при котором выбираются четные или нечетные отсчеты, что соответствует схеме субполосного кодирования [27]. Это прореживание можно провести, поскольку рассматриваемая фильтрация приводит к уменьшению в два раза частотного диапазона сигнала. Последовательности отсчетов, полученных после квадратурных зеркальных фильтров, определяются следующим образом:

|  |  |
| --- | --- |
|  | (2.10) |

Прореженные сигналы вновь поступают на вход фильтров. Схематично процедура многомасштабного анализа на основе 1D-ДВП представлена на рисунке 2.4*а*. На рисунке 2.4*б* показано, как при переходе на следующий уровень разрешения изменяется полоса частот сигнала.

|  |
| --- |
| *а* |
| *б* |
| Рисунок 2.3– Пример скейлинг-функции (*а*) и вейвлета Добеши (*б*) для *M*=10 |
|  |
| fig1  *а* |
| fig2  *б* |
| Рисунок 2.4– Схематическое представление 1D-ДВП в рамках многомасштабного анализа (*а*) и соответствующее уменьшение полосы частот при переходе на следующий уровень разрешения (*б*) |

Несмотря на то, что в результате прореживания каждый из временных рядов будет характеризоваться диапазоном частот вдвое меньшим, чем у сигнала до фильтрации, наличие двух последовательностей на выходе каждого фильтра позволяет однозначно восстановить исходный сигнал при обратном преобразовании.

Коэффициенты разложения по вейвлетам  отражают амплитудные характеристики анализируемых процессов на разных уровнях разрешения. Для фильтрации помех небольшие по абсолютной величине вейвлет-коэффициенты на малых масштабах (наиболее подверженные влиянию флуктуаций) отбрасывают перед проведением обратного преобразования (метод пороговой фильтрации). При этом качество фильтрации существенно зависит от выбора варианта задания пороговой функции [28], на которую умножаются соответствующие коэффициенты перед обратным преобразованием («мягкий» или «жесткий» трешолдинг – рисунок 2.5) и от вейвлет-базиса. Подходящий выбор способствует получению более высокого качества очистки сигнала или изображения от помех.

На рисунке 2.5 изображены три варианта задания пороговой функции  для коэффициентов вейвлет-преобразования. В варианте (*a*) выполняется равенство , которое означает отсутствие корректировок коэффициентов, и в результате обратного преобразования будет получен исходный сигнал. В варианте (*б*) функция задается в виде

|  |  |
| --- | --- |
|  | (2.11) |

При использовании такой пороговой функции остаются неизменными большие по модулю (наиболее значимые) вейвлет-коэффициенты, и обнуляются малые. Наконец, для варианта (*в*) пороговая функция выбирается следующим образом:

|  |  |
| --- | --- |
|  | (2.12) |

В последнем случае уменьшение абсолютных значений всех вейвлет-коэффициентов, включая большие по модулю, может привести к изменению амплитуды восстановленного сигнала. Для тех приложений, где важно сохранить неизменными амплитудные характеристики, такой подход неприменим, однако существуют задачи, где важнее сохранить регулярность сигнала, чем точно воспроизвести его амплитуду. Примером служит фильтрация сигналов от различных помех, где метод «мягкого» задания пороговой функции является широко используемым подходом. Аудио-сигнал после фильтрации может быть усилен, и предварительная очистка его от помех важнее изменения амплитудных характеристик.

|  |
| --- |
| *а*  *б*  *в* |
| Рисунок 2.5 – Задание пороговой функции при вейвлет-фильтрации:  *а* – исходный сигнал, *б* – «жесткий» вариант задания пороговой функции,  *в* – «мягкий» вариант задания пороговой функции |

Применительно к сигналам обычно вводят в рассмотрение величину среднеквадратичной ошибки:

|  |  |
| --- | --- |
|  | (2.13) |

Данная величина позволяет сравнить два сигнала и количественно охарактеризовать степень сходства (или, наоборот, различия) между ними. Если один из сигналов является информационным (не содержащим помех), а второй – сигналом, полученным после фильтрации (и частично содержащим флуктуации в случае неидеального фильтра), то в каждый момент времени оценивается ошибка , которая далее усредняется. Полученная величина позволяет судить о качестве очистки информационного сигнала от помех. Отметим, что  не зависит от временных или пространственных взаимосвязей между выборками исходного сигнала. При анализе сигналов исследователь имеет дело с сильно структурированными объектами, в частности, порядок следования отсчетов отражает важные информационные характеристики. Внесение искажений (например, нарушение корреляций определенной длительности) будет влиять на качество информационного сообщения, однако эти искажения могут не отражаться в величине . По этой причине, например, при анализе аудио-сигналов, содержащих речевые сообщения, применяют дополнительные меры, которые дают более надежную оценку качества по сравнению со среднеквадратичной ошибкой.

Помимо среднеквадратичной ошибки или квадратного корня из величины  в (2.13) при анализе результатов фильтрации рассматривают отношение сигнал/шум (SNR, ОСШ):

|  |  |
| --- | --- |
|  | (2.14) |

В формуле (2.14)  – исходный сигнал, содержащий флуктуации,  – отфильтрованный сигнал, то есть оценка сигнала, «очищенного» от шума, и, соответственно, разность значений  характеризует шумовую составляющую (в случае идеального фильтра). Расчеты количественных критериев целесообразно проводить в дополнение к визуальной оценке качества фильтрации [29].

## 2.3 Типы вейвлетов для решения задач шумоподавления

Вейвлеты, традиционно применяемые для шумоподавления, можно разделить на предвейвлеты, регулярные и дискретные вейвлеты Мейера, ортогональные вейвлеты, биортогональные пары вейвлетов, комплексные вейвлеты.

1. **Предвейвлеты.** Это гауссовы вейвлеты, вейвлеты Морле, мексиканская шляпа. Они имеют минимальные свойства [21]:

* Функция φ не существует;
* Анализ не ортогональный;
* Функция ψ имеет некомпактный носитель;
* Восстановление не гарантируется
* Возможно непрерывное разложение.

Основные достоинства: симметрия, ψ имеет явное выражение.

Основные трудности: быстрые алгоритмы и реконструкция невозможны.

1. **Регулярные вейвлеты Мейера и дискретные вейвлеты Мейера.**

Свойства вейвлетов Мейера:

* Функция φ существует и анализ ортогональный;
* Функции ψ и φ не имеют аналитического выражения;
* Функции ψ и φ не имеют компактного носителя.

Возможный анализ: непрерывное разложение, дискретное преобразование.

Основные достоинства: симметрия, бесконечная дифференцируемость.

Основные трудности: быстрые алгоритмы недоступны.

Дискретные вейвлеты Мейера – это аппроксимация конечной импульсной характеристики к вейвлетам Мейера.

1. **Ортогональные вейвлеты с компактным носителем.** Это вейвлеты Добеши, симлеты, койфлеты.

Общие свойства:

* Функция φ существует и анализ ортогональный;
* Функции ψ и φ имеют компактный носитель;
* Функция ψ имеет некоторое число нулевых моментов.

Возможный анализ: непрерывное разложение, дискретное преобразование с использованием быстрых алгоритмов.

Основные достоинства: компактный нситель, обращающиеся в нули моменты, FIR-фильтры.

Основные трудности: слабая регулярность.

Свойства симметрии:

* Добеши несимметричны;
* Симлеты почти симметричны
* Койфлеты почти симметричны, и функция φ, как и ψ, имеет несколько нулевых моментов.

1. **Биортогональные пары вейвлетов с компактным носителем.** Это В – сплайновые биортогональные вейвлеты.

Свойства:

* Функция φ существует и анализ биортогональный;
* Все функции ψ и φ для декомпозиции и реконструкции имеют компактный носитель;
* Функции ψ, φ для разложения имеют нулевые моменты;
* Функции ψ и φ для восстановления обладают некоторой регулярностью.

Возможный анализ: непрерывное разложение, дискретное преобразование с использованием быстрых алгоритмов.

Основные достоинства: симметрия, FIR-фильтры.

Основные трудности: отсутствие ортогональности.

1. **Комплексные вейвлеты.** Это комплексные гауссовы вейвлеты, комплексные вейвлеты Морле, комплексные вейвлеты Шеннона, комплексные частотные В-сплайновые вейвлеты.

Свойства:

* Функция φ не существует;
* Анализ не ортогональный;
* Функция ψ с некомпактным носителем;
* Свойство реконструкции не обеспечено.

Возможный анализ: комплексное непрерывное разложение.

Основные достоинства: симметрия, ψ имеет явное выражение.

Основные трудности: быстрые алгоритмы и реконструкция невозможны.

Стационарные вейвлеты в ряде случаев представляют собой превосходные средства для решения задач, связанных с очисткой сигналов от шумов. Они хорошо отделяют сигнал от шума и позволяют легко выделять две составляющие сигнала – оригинальный сигнал и шум.

В таблице 2.1 представлено сравнение характеристик основных вейвлетов.

Результаты предварительных исследований показали, что лучшие результаты в очистке сигналов от шумов дают дискретный вейвлет Мейера, вейвлеты Добеши и Симлета. Именно эти семейства вейвлетов было решено использовать для очистки переговоров авиадиспетчеров от шумов. Выбор вейвлетов для очистки осуществлялся из ортогональных и тех, которые обладают возможностью реконструкции сигнала.

Таблица 2.1(а) – Сравнение характеристик вейвлетов

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Хар-ка/Вейвлет** | **Хаара** | **Мейера** | **Добеши** | **Симлеты** |
| **Общие хар-ки** | С компактным носителем. Самый старый и простой вейвлет.  φ =1 на [0;1], иначе 0;  ψ=1 [0;0.5],  ψ=-1 [0,5;1] | - | С компактным носителем.   Вейвлеты с экстремальной фазой и высоким числом нулевых моментов на данной области корр. работы | С компактным носителем. Симметричный с наибольшим числом нулевых моментов на области корр. работы |
| **Ортогональность/**  **биортогональность** | Да/да | Да/да | Да/да | Да/да |
| **Компактная поддержка** | Есть | Есть | Есть | Есть |
| **Дискретное ВП** | Возможно | Возможно без БВП. Базируемое на КИХ, обеспечивает БВП | Возможно | Возможно |
| **Непрерывное ВП** | Возможно | Возможно | Возможно | Возможно |
| **Область корректн. работы** | 1 | Бесконечность | 2N-1 | 2N-1 |
| **Область эффект. работы** | - | [-8;8] |  | - |
| **Регулярность** | Не явл. непрерывным | Бесконечно рег. | ~0,2N при больших N | - |
| **Симметричность** | Да | Да | Нет | Да |
| **Длина фильтра** | 2 | - | 2N | 2N |

*Таблица 2.1(б). Сравнение характеристик вейвлетов*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Хар-ка/Вейвлет** | **Койфлеты** | **Гаусса** | **Мексиканская**  **шляпа** | **Морлеты** |
| **Общие хар-ки** | С компактным носителем. С наибольшим числом нулевых моментов φ и ψ на области корр. Работы | - | - | - |
| **Ортогональность/**  **биортогональность** | Да/да | Нет/нет | Нет/нет | Нет/нет |
| **Компактная поддержка** | Есть | Нет | Нет | Нет |
| **Дискретное ВП** | Возможно | Нет | Нет | Нет |
| **Непрерывное ВП** | Возможно | Возможно | Возможно | Возможно |
| **Область корректн. Работы** | 6N-1 | Бесконечность | Бесконечность | Бесконечность |
| **Область эффект. Работы** | - | [-5;5] | [-5;5] | [-4;4] |
| **Симметричность** | Да | Да.  N четное- симметр,  N нечетное – не симметр. | Да | Да |
| **Длина фильтра** | 6N | - | - | - |

# ГЛАВА 3

# ПРИМЕНЕНИЕ ФИЛЬТРОВ И ВЕЙВЛЕТОВ ДЛЯ РЕШЕНИЯ ЗАДАЧИ ШУМОПОДАВЛЕНИЯ В ПЕРЕГОВОРАХ АВИАДИСПЕТЧЕРОВ

## 3.1 Особенности задачи шумоподавления в речевых сигналах

Голос и речь являются основным средством коммуникации во всех сферах человеческой деятельности, а передаваемая с их помощью информация может иметь исключительную ценность в обеспечении безопасности, например, раскрытии и предотвращении преступлений.

Человеческая речь представляет собой шумоподобный акустический сигнал, несущий амплитудную и частотную модуляции, что схематично показано на рисунке 3.1.



Рисунок 3.1 – Амплитудная и частотная модуляции

Амплитудная модуляция (АМ) – модуляция, при которой незатухающие колебания изменяются по амплитуде в соответствии с модулирующими его колебаниями более низкой частоты, а частотная модуляция (ЧМ) – модуляция, при которой несущая частота сигнала изменяется в соответствии с модулирующим колебанием [30].

Основная энергия акустических колебаний речевого сигнала заключена в диапазоне 70 Гц - 7 кГц, причем более 95% смысловой информации размещается в более узком диапазоне – 200 Гц - 5 кГц.

Акустические колебания выше и ниже этих частот несут информацию об эмоциях и личности говорящего, способствуют узнаваемости и несколько повышают разборчивость речи в условиях повышенных шумов [31].

Основной причиной низкого качества и плохой разборчивости записанной речевых сигналов является присутствие в них искажений и помех.

Искажениями принято называть видоизменения самого полезного речевого сигнала, приводящие к снижению его качества. При искажениях собственные составные части речевого сигнала изменяются по отношению к своему первоначальному виду, приобретают новое, иногда недопустимое звучание [30].

Помехой называется стороннее возмущение, действующее в системе передачи и препятствующее правильному приёму сигналов, а также вызывающее искажение передаваемой информации [30].

Источники помех могут находиться как вне, так и внутри самой системы передачи.

В зависимости от вида информации помехи проявляются в виде:

1. ошибок при передаче телеграмм и передаче данных;
2. шорохов, тресков, в плохой разборчивости речи и слышимости разговоров, ведущихся по соседним каналам, при телефонной связи;
3. искажений команд в системах телемеханики и телесигнализации и т.д.

Действие помехи зависит от множества причин и, как правило, носит случайный характер. Помехи можно разделить на две группы – неаддитивные и аддитивные.

К неаддитивным, относят помехи, вызывающие паразитную модуляцию сигнала. Они возникают из-за нелинейной зависимости характеристик канала связи от параметров сигнала и от времени, и существенно влияют на передачу сигналов в основном в каналах проводной связи большой протяжённости.

К аддитивным относят помехи, которые складываются с сигналом линейно. Аддитивную помеху часто называют шумом.

Шум – звучание, соответствующее восприятию независимого от полезного сигнала источника мешающего звука.

Другими словами, шум – звук нежелательного дополнительного источника, как правило, добавленный к полезному сигналу во время его записи или его передачи по каналам связи. Появление в звуковом сигнале шумов обусловлено раздельным или, чаще, одновременным действием целого ряда факторов процесса записи или передачи сигнала: дополнительного источника звука, находящегося вблизи от полезного источника речи, а также электромагнитных наводок и собственных «технических» шумов различных компонентов канала записи/воспроизведения. То есть шумом может являться как речь другого человека или, например, звук работающего двигателя, так и шум трансформатора или шипение магнитной ленты.

Различают шум стационарный и нестационарный. Стационарный шум характеризуется постоянством средних параметров: интенсивности (мощности), распределения интенсивности по спектру (спектральная плотность). Идеальным стационарным шумом является так называемый "белый шум" - шум с абсолютно равномерным спектром. В реальности такой шум не может существовать, потому что его мощность была бы бесконечной.

Нестационарный шум – это шум, длящийся короткие промежутки времени (меньшие, чем время усреднения в измерителях).

Классификация шума:

1. по источнику образования:
   * механический (работа машин и механизмов) – создается колебаниями твердой и жидкой поверхности;
   * аэро- и гидродинамический – в результате турбулентности газовой или жидкой среды;
   * электродинамический – электрическая дуга, коронные разряды.
2. По частоте:
   * низкочастотный до 300 Гц,
   * среднечастотный от 300 до 800 Гц,
   * высокочастотный свыше 800 Гц.
3. По спектру:
   * широкополосный,
   * тональный.

Спектром шума называется зависимость уровня звукового давления от частоты. Шум считается широкополосным, если его спектр превышает одну октаву, и тональным, если звуковая энергия распределяется неравномерно, с преобладанием большей ее части в области одной октавы.

1. По времени действия:
   * квазистационарный – уровень звукового давления в течение рабочей смены меняется не более чем на 5 дБ,
   * стационарный – меняется в любую сторону более чем на 5 дБ и подразделяется: колеблющийся – уровень звуков непрерывно плавно изменяется во времени; прерывистый – изменяется ступенчато более чем на 5 дБ, оставаясь на ступени не менее 1 сек; импульсный – состоит из одного или нескольких звуковых сигналов продолжительностью менее 1 сек.

Шум измеряют прибором, который называется шумомер, и состоит из микрофона, усилителя, измерительного прибора и источника питания.

Выделим некоторые типы шумов/помех:

**1. Белый шум** – это шум с постоянной спектральной плотностью в речевом диапазоне частот, который имеет одинаковое распределение мощности для всех частот. Примером белого шума может быть звук ненастроенного телевизора или шум водопада.

Белым шумом является такой шумовой сигнал, у которого на единицу частоты всюду приходится равная энергия.

Поэтому в полосе частот от 100 Гц до 101 Гц сосредоточено энергии столько же, сколько в полосе от 1000 до 1001 Гц. Так как на каждый герц энергии приходится поровну, то в полосе частот 1000 – 5000 Гц её будет в 10 раз больше, чем в полосе 100 – 500 Гц. Таким образом, белый шум звучит для человека менее приятно на высоких частотах.

**2. Розовый шум** имеет одинаковое распределение энергии для каждой октавы (октавные полосы – частотные диапазоны, в которых верхний предел каждой полосы вдвое больше нижнего предела) вместо одинаковой энергии для каждой частоты подобно белому шуму.

Он представляет собой психоакустический эквивалент белого шума. У розового шума на каждую октаву приходится энергии поровну: от 40 до 80 Гц – столько же, сколько от 400 до 800 Гц и от 10 до 20 кГц.

У розового шума энергия, приходящаяся на единицу частоты, уменьшается кратно частоте. Однако на каждую октаву энергии приходится одинаково.

Согласно психоакустической модели слуховой системы, наилучшей маскирующей помехой является комбинация белого и розового шума.

Известно, что белый или розовый шум, применяемый в качестве маскирующего сигнала в устройствах защиты речевой информации, по своей структуре имеет значительные отличия от речевого сигнала [ ].

На знании и использовании этих отличий основаны алгоритмы шумоподавлении речевых сигналов, используемые, например, специалистами технической разведки.

Одним из направлений повышения эффективности защиты речевой информации является использование в качестве помехи, применяемой для зашумления каналов утечки речевой информации, следующих речеподобных сигналов (речевых сигналов низкой разборчивости): окрашенный шум, шумовая речеподобная помеха и комбинированная речеподобная помеха.

**3. Окрашенный шум** – это шум с огибающей амплитудного спектра, подобной речевому сигналу, формирующийся из белого шума в соответствии с огибающей амплитудного спектра скрываемого речевого сигнала.

Для формирования окрашенного шума в пятиоктавных полосах диапазона 100−6000 Гц производится оценка параметров речевого сигнала и осуществляется корректировка уровня шума в тех же полосах с помощью встроенных эквалайзеров.

**4. Речеподобные** помехи формируются путем микширования в различных сочетаниях отрезков речевых сигналов и музыкальных фрагментов, а также шумовых помех, или из фрагментов скрываемого речевого сигнала при многократном наложении с различными уровнями.

а) Шумовая речеподобная помеха – это помеха, формирующаяся как из скрываемого сигнала, так и из некоррелированных со скрываемым сигналом речевых фрагментов.

б) Комбинированная речеподобная помеха, используемая в системе акустической маскировки, формируется путем многократного наложения смещенных на различное время задержек разноуровневых сигналов, получаемых путем умножения и деления частотных составляющих скрываемого речевого сигнала.

Наиболее эффективными являются помехи типа розовый шум, комбинация белого и розового шумов и речеподобная помеха. Помеха типа белого шума по сравнению с помехами типов розовый шум и шумовая речеподобная обладает несколько худшими маскирующими свойствами, проигрывая по энергетике. Значительно более низкими маскирующими свойствами обладает шумовая помеха со спадом спектральной плотности в 6 дБ на октаву в сторону высоких частот (коричневый шум). По сравнению с помехами типов розовый шум и речеподобная она проигрывает по энергетике, а при равной мощности приводит к повышению разборчивости речи.

Для решения практической задачи шумоподавления наиболее важными являются такие специфические свойства слуха человека, как частотная и временная маскировка [32].

Частотная маскировка – явление, возникающее в частотной области, когда сигнал низкого уровня (маскируемый) становится неслышимым (замаскированным) при звучании одновременно с ним более сильного сигнала (маскирующего) при том условии, что маскирующий и маскируемый сигналы достаточно близки друг другу по частоте, как показано на рисунке 3.2.

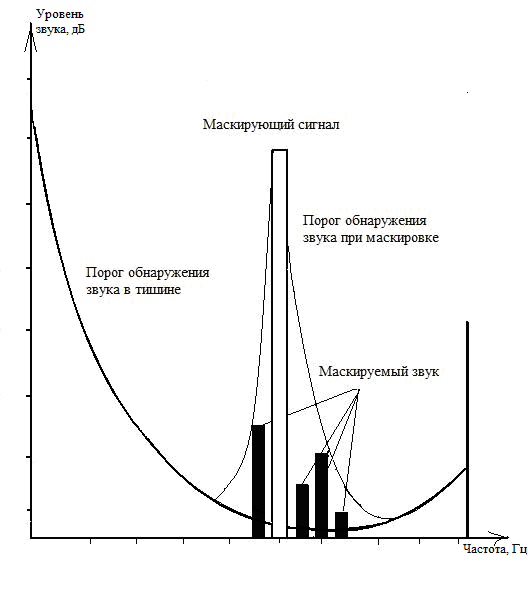


Рисунок 3.2 – Мгновенная маскировка в частотной области для слуховой системы человека

Длинным светло-серым прямоугольником отображен сильный узкополосный сигнал-маскер. Он маскирует (делает неслышимыми) 4 слабых сигнала (черные прямоугольники), расположенных в окрестности сигнала-маскера с уровнями ниже порога маскировки, отмеченного линиями. Другая черная толстая линия показывает кривую порога обнаружения звукового сигнала данной частоты в тишине.

Временная маскировка – это явление, схожее с частотной маскировкой, но происходящее во времени. При прекращении подачи маскирующего звука, маскируемый некоторое время продолжает оставаться неслышимым. В обычных условиях эффект от временной маскировки длится относительно недолго. Время маскировки зависит от частоты и амплитуды сигнала и может достигать 100 мс. Эффект временной маскировки почувствовать очень просто: близко взорвавшаяся хлопушка приводит к временному «закладыванию» одного из ушей (закладывание обоих ушей маловероятно, но тоже возможно).

На рисунке 3.3 представлен график, иллюстрирующий временную маскировку.

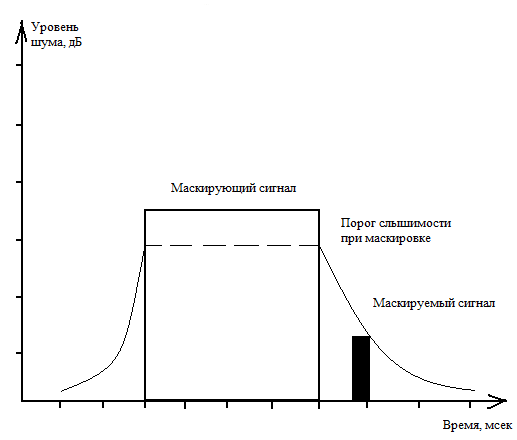


Рисунок 3.3 – Пример временной маскировки

По горизонтальной оси отложено время в миллисекундах, по вертикальной оси отложен уровень звучания сигналов. Белым прямоугольником представлен сигнал-маскер, кривыми линиями - пороги обнаружения звука при маскировке. Маскируются (становятся неслышимыми) слабые сигналы, лежащие ниже порога слышимости при маскировке, отмеченного линиями [32].

В связи с вышесказанным, достаточно часто возникает потребность в проведении специальной обработки или шумоподавления, звукового сигнала, главной целью которой является повышение качества и разборчивости записанной на фонограмме речи, для ее последующего комфортного прослушивания и понимания.

Как правило, грамотно выполненное шумоподавление позволяет уменьшить отрицательное влияние шумов и искажений.

Основной принцип шумоподавления заключается в подчеркивании полезного сигнала или в подавлении компонент зашумленного сигнала в тех областях, где помеха велика, а полезный сигнал мал, и в усилении только компонент полезного сигнала. Наиболее полно эту задачу решают методы цифровой фильтрации, аналоговая обработка обычно позволяет решить эту задачу лишь частично.

Речевые сигналы, передаваемые по каналам связи, всегда в той или иной степени зашумлены. В тех случаях, когда шум имеет значительную интенсивность, его наличие может существенно исказить результаты обработки, анализа или распознавания речи. В целом ряде других случаев, например, при анализе зашумленных записей в криминалистических целях или восстановлении аудиозаписей в архивах, задача очистки сигнала от шума носит самостоятельный характер и является единственной целью работы.

В данной работе для сравнения результатов очистки речевого сигнала от шума использовались записи переговоров авиадиспетчеров.

Данная тематика выбрана для исследований неслучайно. С момента появления самолетов проблема распознавания речи во время переговоров авиадиспетчеров с пилотами стала очень актуальной. Шум двигателя самолета и помехи, возникающие при передаче сообщений на огромные расстояния, мешают, прослушав запись, сразу разобрать речь. Поэтому при возникновении критической ситуации, когда нужно уловить каждый произнесенный звук в переданном сообщении, необходимо прибегать к шумоподавлению.

## 3.2 Выбор фильтров для очистки сигнала. Результаты их применения

Во многих технических областях и приложениях часто возникает такая проблема, как загрязнение полезного сигнала нежелательными, зачастую мощными, сигналами или шумом. Если полезный сигнал и шум занимают строго определенные и разнесенные полосы частот, для извлечения полезного сигнала используются обычные линейные фильтры. Однако существуют ситуации, при которых нужно, чтобы характеристики фильтра были переменными, адаптировались к меняющимся характеристикам сигнала или чередовались, т.е. обладали интеллектуальностью. В подобных случаях, например, когда спектры сигнала и шума перекрываются, как показано на рисунке 3.4, или когда полоса шумовых частот неизвестна или переменна, коэффициенты фильтра должны меняться, и их нельзя задать заранее [22].

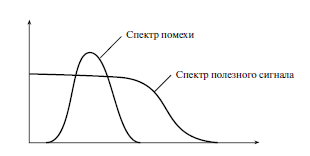


Рисунок 3.4 – Иллюстрация наложения спектров сигнала и мощной помехи

Далее представлены типичные приложения, в которых, согласно [5], неуместны фильтры с фиксированными коэффициентами.

1. Электроэнцефалография (ЭЭГ), где артефакты, или загрязнение сигнала, порожденное движением или морганием глаза, значительно больше подлинной электрической деятельности мозга и занимают ту же полосу частот, что и сигнал, представляющий медицинский интерес. Чтобы удалить артефакты, сохранив при этом нужные сигналы, обычные линейные фильтры не подходят.

2. В цифровой связи присутствует метод расширения спектра, который используется в условиях, когда сильная активная помеха может интерферировать с полезным сигналом. Интерференция часто происходит в узкой, но неизвестной полосе частот в широкополосном спектре, и единственное эффективное решение этой проблемы предлагают адаптивные фильтры.

3. Цифровая передача данных по телефонным каналам на высокой скорости. Искажение сигнала, вызванное плохими амплитудной и фазовой характеристиками канала, приводит к тому, что импульсы, представляющие различные цифровые коды, интерферируют между собой (межсимвольная интерференция), что затрудняет надежное детектирование кодов в приемнике. Для компенсации внесенного каналом искажения, которое может быть переменным или неизвестным, используется адаптивное выравнивание.

Адаптивный фильтр отличается следующим свойством: его частотная характеристика автоматически регулируется или модифицируется для улучшения производительности фильтра в соответствии с некоторым критерием, что позволяет фильтру адаптироваться к изменениям характеристик входного сигнала. Вследствие самонастройки и присущей адаптивным фильтрам гибкости, они используются в таких сферах, как телефонное эхоподавление, обработка сигналов с радаров, системы навигации, выравнивание каналов связи, выделение биомедицинских сигналов.

Итак, адаптивные фильтры стоит использовать, если:

* характеристики фильтра должны быть переменными, адаптироваться к меняющимся условиям;
* существует спектральное перекрытие сигнала и шума;
* полоса шумовых сигналов неизвестна или меняется со временем.

В перечисленных случаях использование обычных фильтров привело бы к неприемлемому искажению полезного сигнала.

Адаптивный фильтр состоит из двух различных частей: цифрового фильтра с регулируемыми коэффициентами и адаптивного алгоритма, который используется для настройки или изменения коэффициентов фильтра (рисунок 3.5). На адаптивный фильтр одновременно подаются два входных сигнала, yk и xk. Сигнал yk загрязнен и содержит полезный сигнал sk и шум nk, причем предполагается, что эти сигналы не коррелируют. Сигнал xk — это мера загрязненного сигнала, который коррелирует с nk. В цифровом фильтре сигнал xk обрабатывается для получения оценки сигнала nk. Затем полезный сигнал оценивается как разность зашумленного сигнала yk и выхода цифрового фильтра :

= yk −= sk + nk − (3.1)

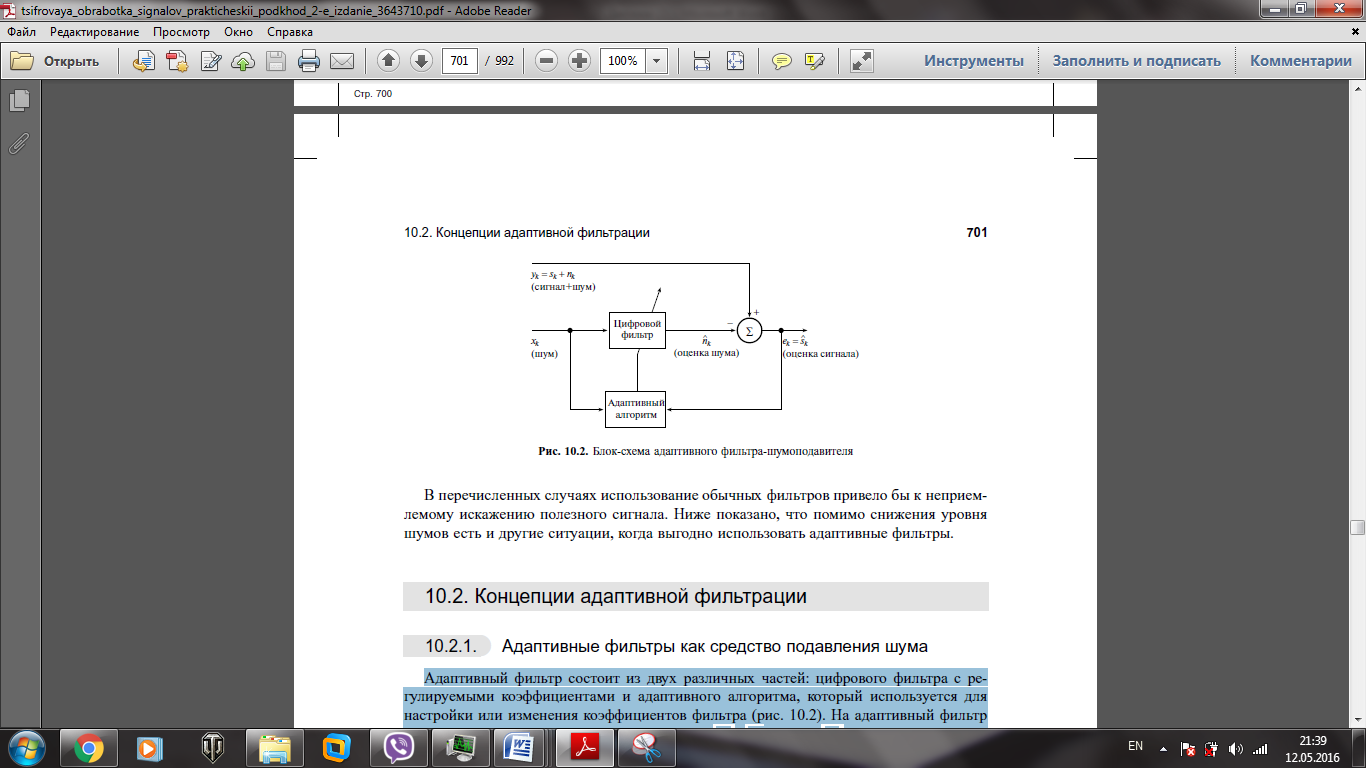


Рисунок 3.5 – Блок-схема адаптивного фильтра-шумоподавителя

Основная цель шумоподавления заключается в получении оптимальной оценки шума в зашумленных сигналах, а следовательно, оптимальной оценки желаемого сигнала. Для этого подается в цепь обратной связи, и на основе этого сигнала и подходящего адаптивного алгоритма регулируются коэффициенты цифрового фильтра для минимизации шума в сигнале . Выходной сигнал используется двояко: 1) как оценка желаемого сигнала и 2) как сигнал ошибки, который применяется для регулировки коэффициентов фильтра.

Стоит иметь в виду, что значение сигналов xk, yk и ek или способа их получения зависит от приложения. Некоторые примеры конфигураций адаптивного фильтра приведены на рисунке 3.6.

В большинстве адаптивных систем цифровой фильтр, показанный на рисунке 3.5, реализован с использованием трансверсальной структуры или схемы с конечной импульсной характеристикой (КИХ) (рисунок 3.7). Иногда используются и другие формы, например, схема с бесконечной импульсной характеристикой (БИХ) или решетчатая структура, но КИХ-структура применяется наиболее широко из-за своей простоты и гарантированной устойчивости. Выход N-точечного фильтра, изображенного на рисунке 3.7, выражается следующим образом:

(3.2)

где (i), i = 0, 1, . . . — регулируемые коэффициенты фильтра (или весовые коэффициенты), а xk(i) и — вход и выход фильтра соответственно. На рисунке 3.6 изображена система с одним входом и одним выходом. В системе с несколькими входами и одним выходом xk может включать одновременные входы N различных источников сигнала.

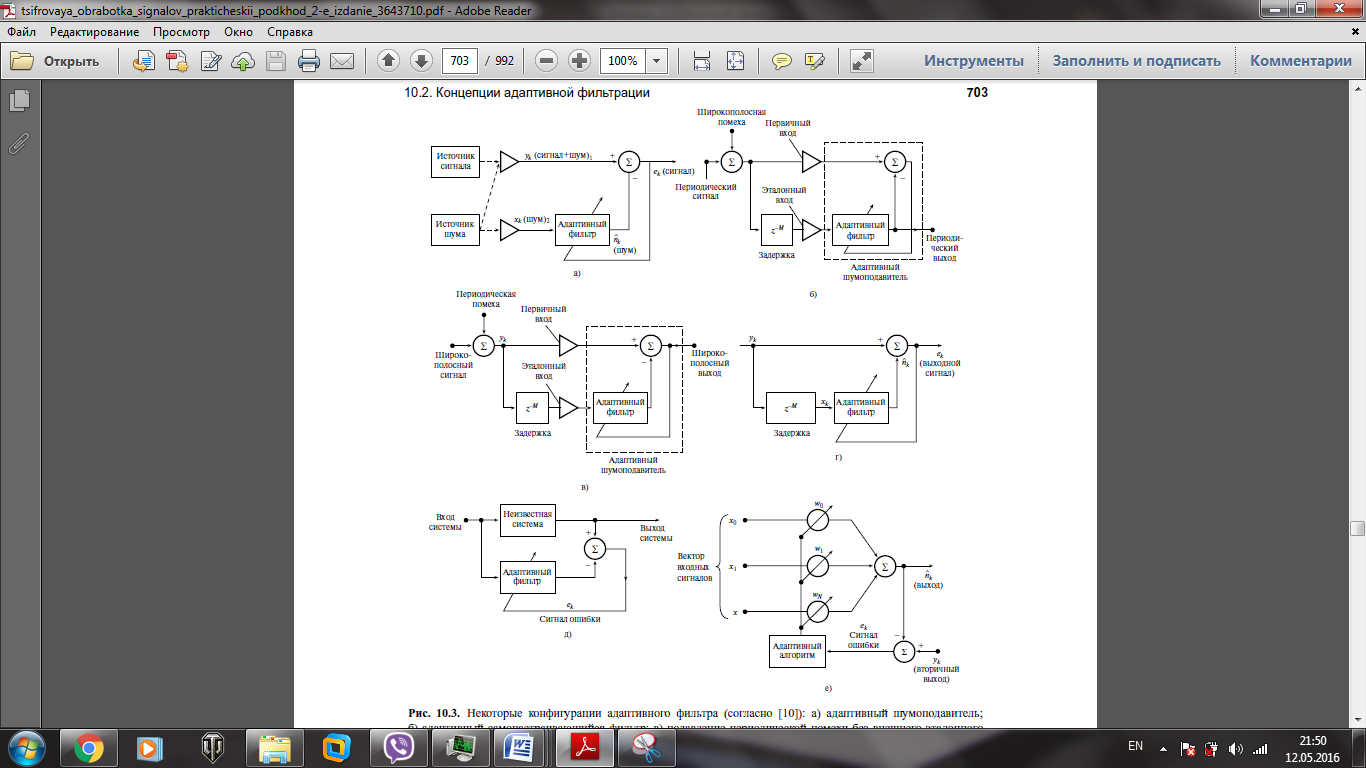


Рисунок 3.6 – Некоторые конфигурации адаптивного фильтра [22]:

а) адаптивный шумоподавитель; б) адаптивный самонастраивающийся фильтр; в) подавление периодической помехи без внешнего эталонного источника; г)адаптивный селективный фильтр; д) моделирование системы; е) подбор линейной комбинации.

Для настройки коэффициентов цифрового фильтра, схема которого представлена на рисунке 3.5, используются адаптивные алгоритмы, в которых сигнал ошибки ek минимизируется согласно некоторому критерию, например, по схеме наименьших квадратов. Наибольшее распространение получили алгоритмы, в которых используются схема наименьших квадратов, рекурсивная схема наименьших квадратов и фильтр Калмана, именно эти фильтры и использовались для очистки сигнала в данной работе. С точки зрения вычислительной сложности и требований к памяти наиболее эффективна схема наименьших квадратов. Более того, для нее несущественна проблема численной неустойчивости, присущая двум другим алгоритмам. По этим причинам во многих ситуациях вначале пробуют схему наименьших квадратов.

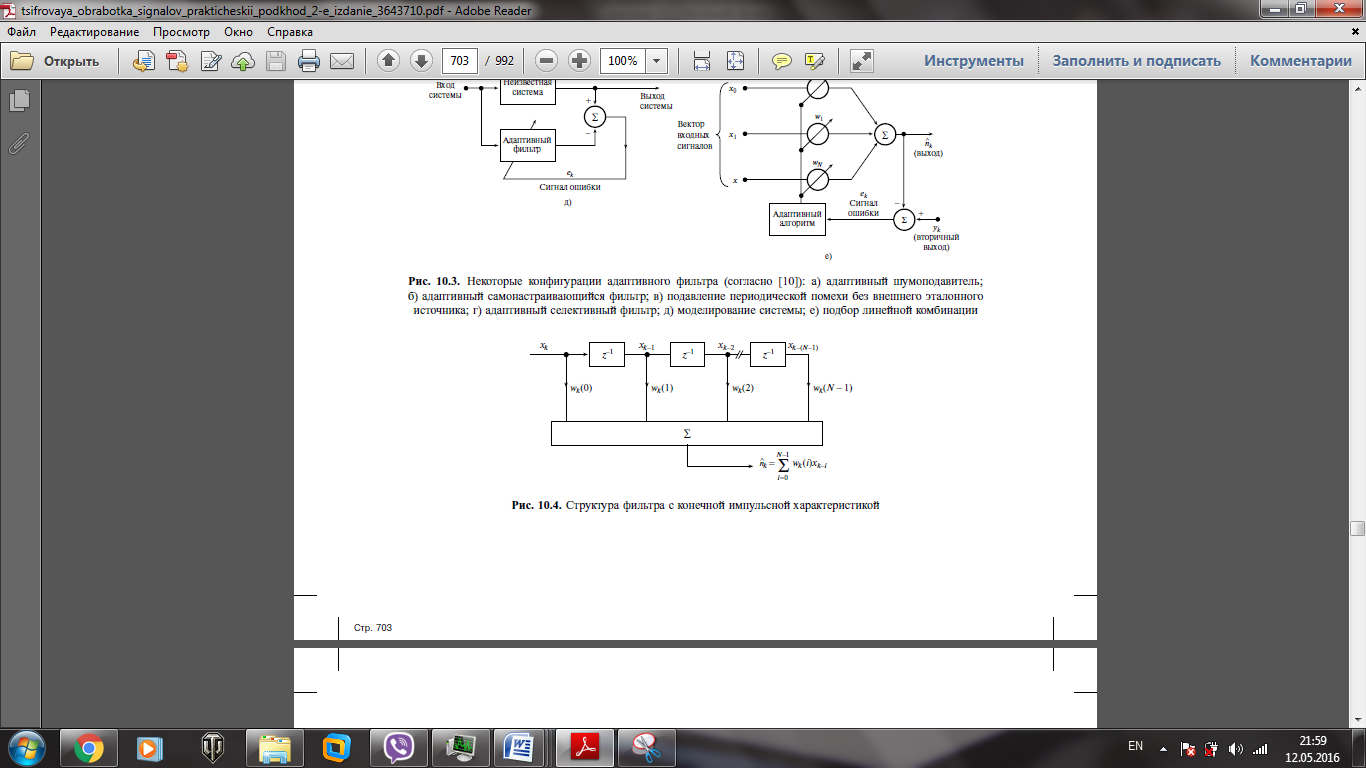


Рисунок 3.7 – Структура фильтра с конечной импульсной характеристикой

Для очистки звукового сигнала от шума применялись следующие фильтры: МНК, Калмана и режекторный фильтр.

### 3.2.1 Реализация стандартного алгоритма наименьших квадратов

Ниже приводится схема вычислений, принятая в алгоритме наименьших квадратов.

1. Изначально каждому весовому коэффициенту (i), i = 0, 1,.., N−1 присваивается произвольное фиксированное значение, например, 0.

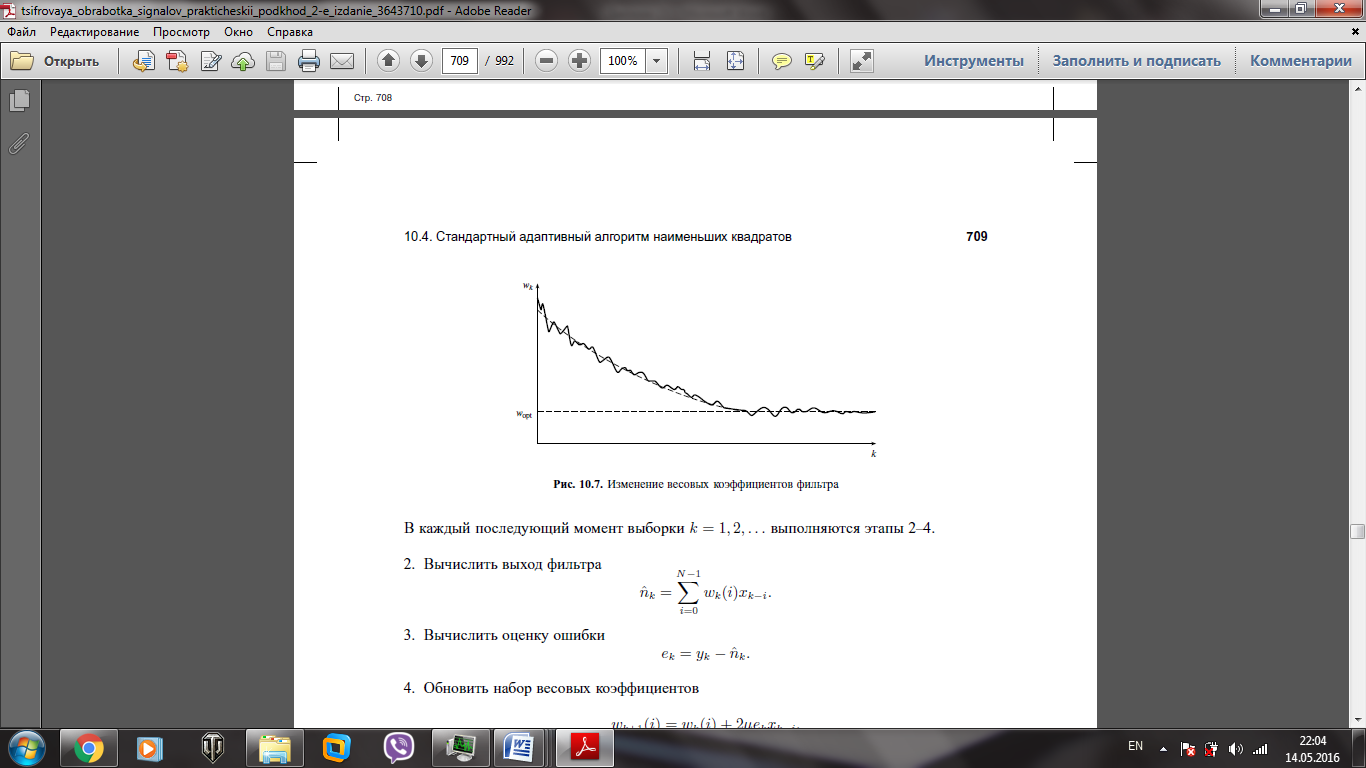


Рисунок 3.8 – Изменение весовых коэффициентов фильтра

В каждый последующий момент выборки k = 1, 2, . . . выполняются этапы 2–4:

1. Вычислить выход фильтра (3.2);
2. Вычислить оценку ошибки:

. (3.3)

1. Обновить набор весовых коэффициентов

(3.4)

Простота алгоритма наименьших квадратов и легкость реализации объясняют, почему данный метод используется в системах реального времени, когда это возможно. Описанный алгоритм требует примерно 2N +1 операций умножения и 2N + 1 операций сложения для каждого нового набора входных и выходных выборок. Основные операции умножения-накопления, фигурирующие в алгоритме, можно выполнять на большинстве процессоров сигналов, т.е. можно использовать прямую реализацию алгоритма.

Применение стандартного алгоритма наименьших квадратов сопряжено с некоторыми проблемами, что приводит к снижению производительности системы.

***Влияние нестационарности***

В стационарной среде поверхность производительности фильтра имеет постоянную форму и ориентацию, а характеристики адаптивного фильтра просто сходятся к оптимальному значению или колеблются около него. Если статистики сигнала меняются после схождения весовых коэффициентов к оптимальному значению, характеристика фильтра должна также перейти в новый набор оптимальных значений путем повторной настройки весовых коэффициентов. При этом считается, что изменение статистик сигнала происходит достаточно медленно, чтобы коэффициенты успевали сходиться до оптимального значения между изменениями. Тем не менее, в нестационарной среде точка минимума непрерывно движется, кроме того, могут меняться ориентация и кривизна траектории ее движения, как показано на рисунке 3.9. Значит, в этом случае алгоритм должен не только найти точку минимума поверхности, но и отследить ее меняющееся положение, что существенно снижает производительность схемы. Следует отметить, что переменная называется нестационарной, если ее статистики (такие как среднее значение, дисперсия, автокорреляция) меняются со временем. Причиной таких изменений могут быть, например, случайные помехи малой продолжительности, как показано на рисунке 3.10, или неверные данные, причем такие изменения часто приводят к расстройке весовых коэффициентов фильтра, что особенно ярко демонстрирует недостатки линейных фильтров.

В настоящее время разработано несколько схем, в которых данная проблема исчезает, но обычно это усложняет стандартный алгоритм наименьших квадратов.

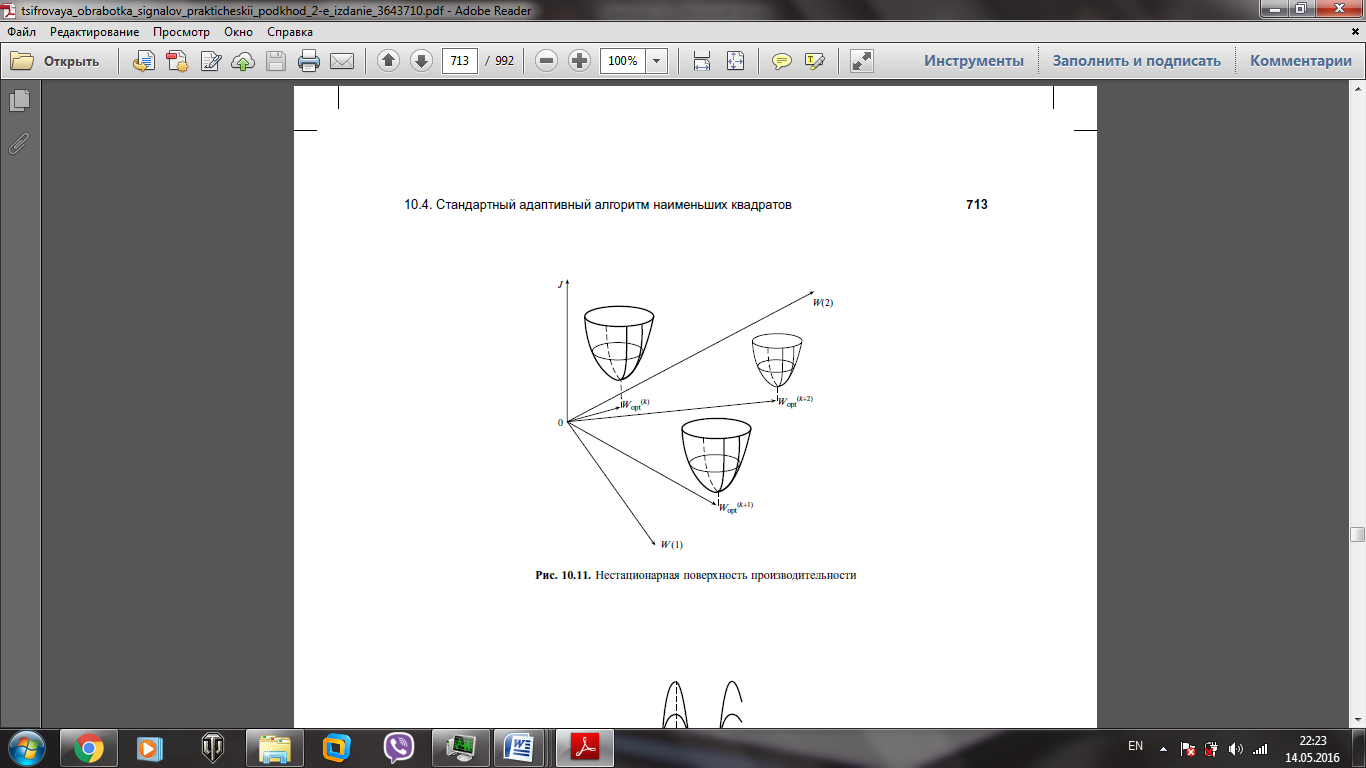


Рисунок 3.9 – Нестационарная поверхность производительности

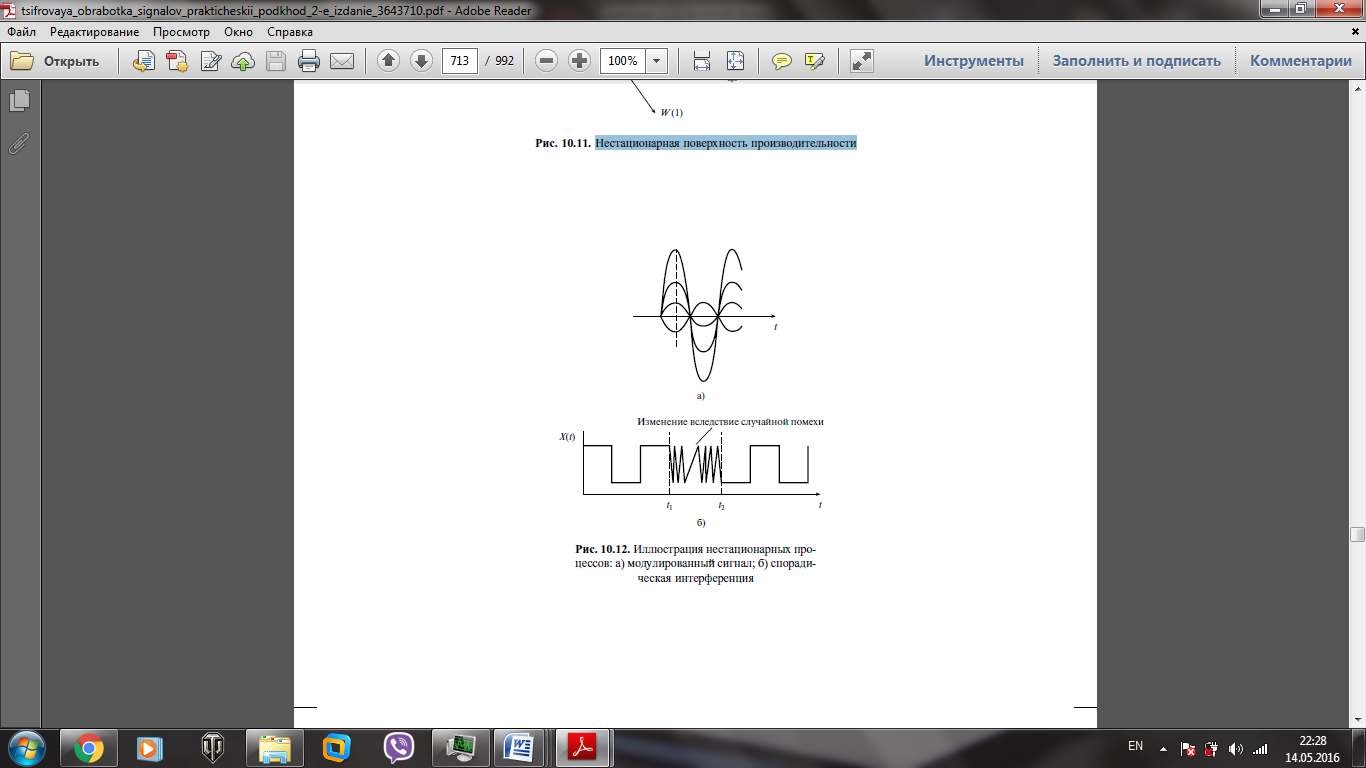


Рисунок 3.10 – Иллюстрация нестационарных процессов:

а) модулированный сигнал; б) спорадическая интерференция.

***Влияние компонентов сигнала на помехи во входном канале***

Производительность алгоритма зависит от измеренного сигнала помехи xk(i), который сильно коррелирует с реальной помехой, но слабо коррелирует (теоретически совсем не коррелирует) с полезным сигналом. В большинстве случаев данное условие не выполняется. В некоторых приложениях зашумленный входной сигнал может включать и нежелательную помеху, и слабые компоненты сигнала. В результате вместе с шумом подавляются некоторые компоненты полезного сигнала. Подобная ситуация иллюстрируется рисунком 3.11. В работе [22] показано, что в этих случаях процесс адаптивного шумоподавления по-прежнему ведет к значительному улучшению отношения полезный сигнал-шум, но только за счет небольшого искажения сигнала. При этом, если xk содержит только сигналы и совсем не содержит шумовых компонентов, полезный сигнал в yk может быть полностью уничтожен.

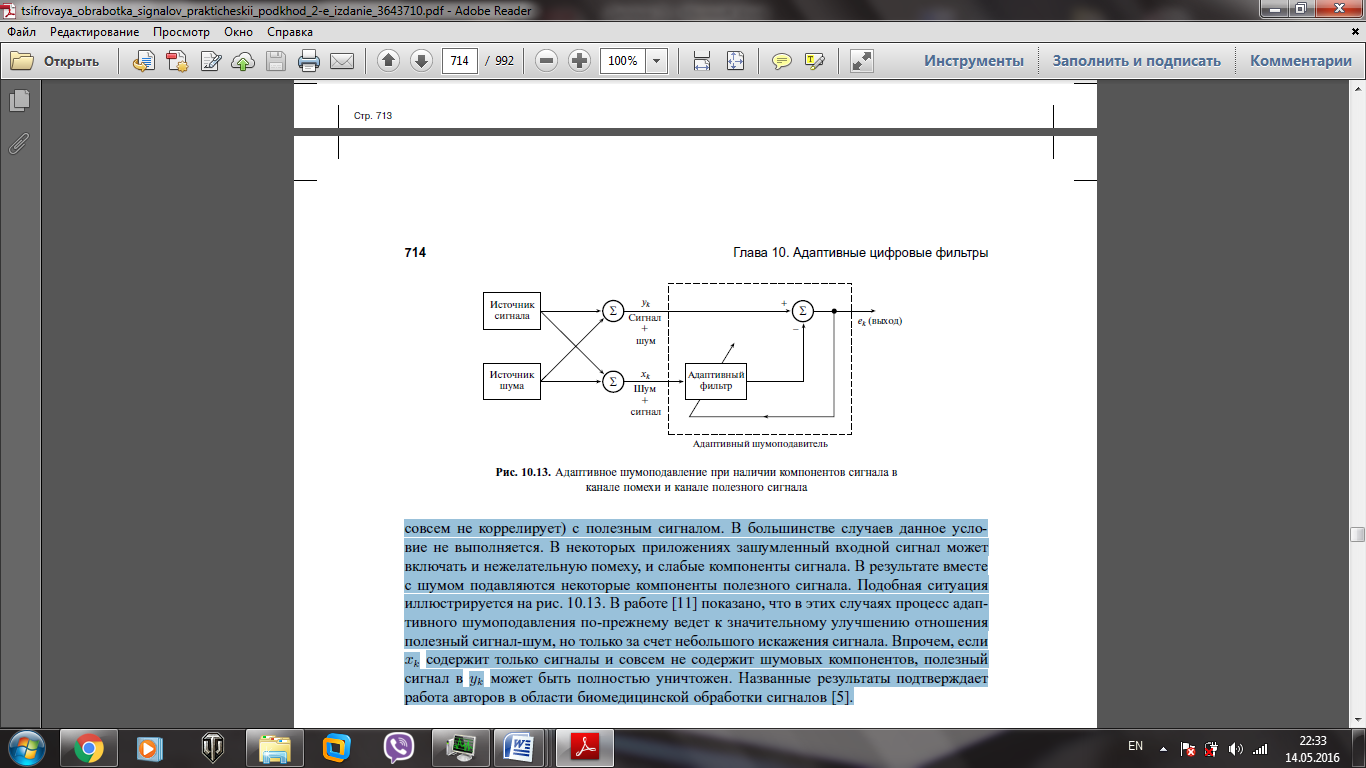
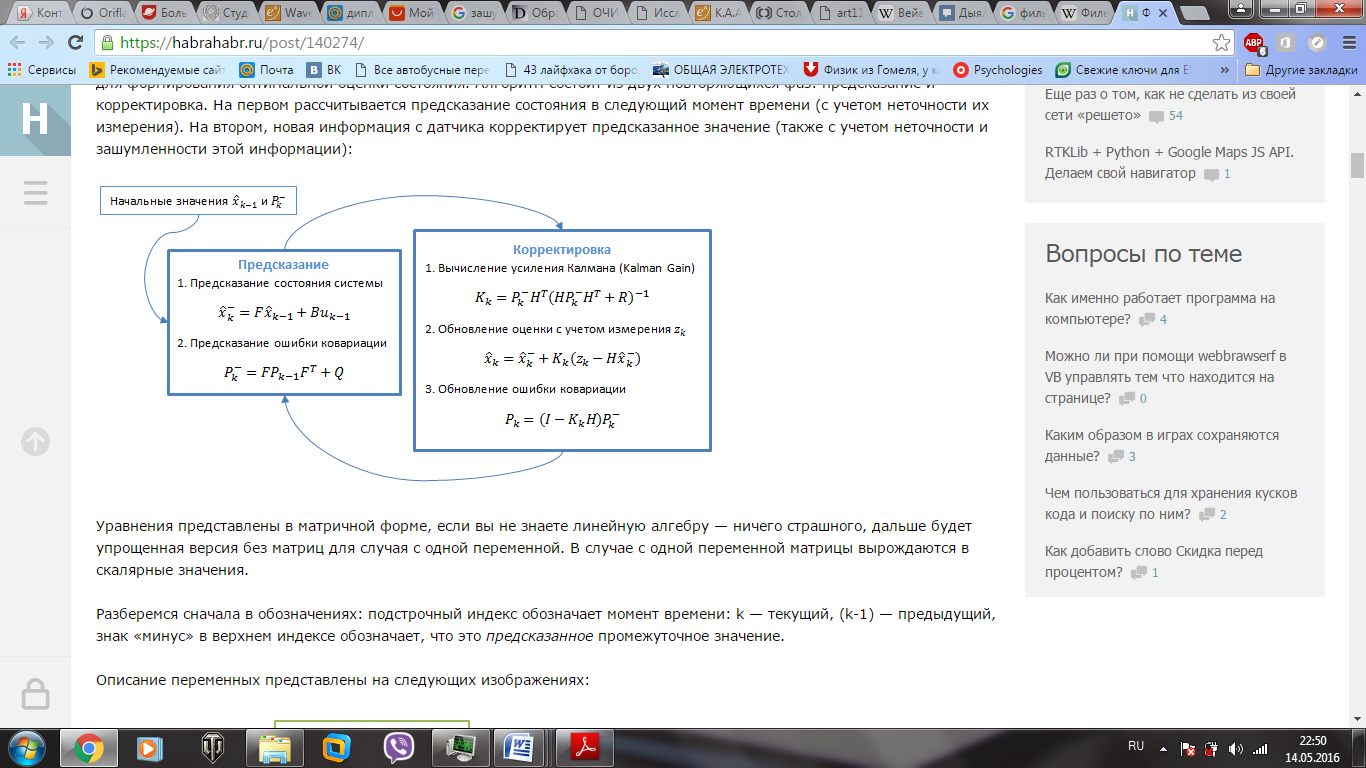


Рисунок 3.11 – Адаптивное шумоподавление при наличии компонентов сигнала в канале помехи и канале полезного сигнала

### 3.2.2 Фильтр Калмана

Фильтр Калмана – это самый популярный сегодня алгоритм фильтрации, используемый во многих областях науки и техники. Благодаря своей простоте и эффективности его можно встретить в GPS-приемниках, обработчиках показаний датчиков, при реализации систем управления и т.д.

В фильтре Калмана есть возможность задать априорную информацию о характере системе, связи переменных и на основании этого строить более точную оценку, но даже в простейшем случае (без ввода априорной информации) он дает отличные результаты. Фильтр Калмана использует динамическую модель системы (например, физический закон движения), известные управляющие воздействия и множество последовательных измерений для формирования оптимальной оценки состояния. Алгоритм состоит из двух повторяющихся фаз: предсказание и корректировка. На первом рассчитывается предсказание состояния датчиков в следующий момент времени с учетом неточности измерений. На втором – новая информация с датчика корректирует предсказанное значение, также с учетом неточности и зашумленности этой информации, как показано на рисунке 3.12.



Подстрочный индекс обозначает момент времени: k — текущий, (k-1) — предыдущий, знак «минус» в верхнем индексе обозначает, что это предсказанное промежуточное значение.

Рисунок 3.12 – Алгоритм реализации фильтра Калмана

Описание переменных представлено на рисунках 3.13 и 3.14.

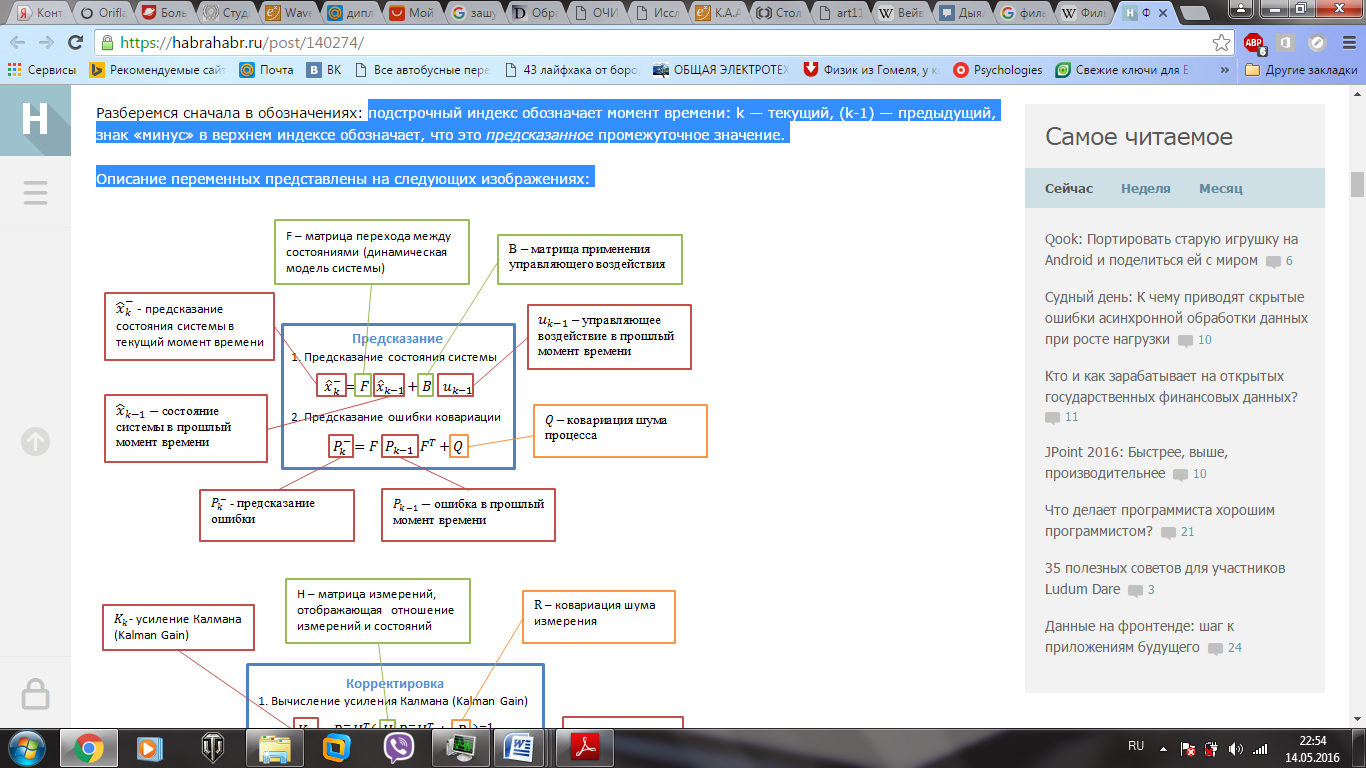


Рисунок 3.13 – Описание переменных предсказания для случая, представленного на рисунке 3.12

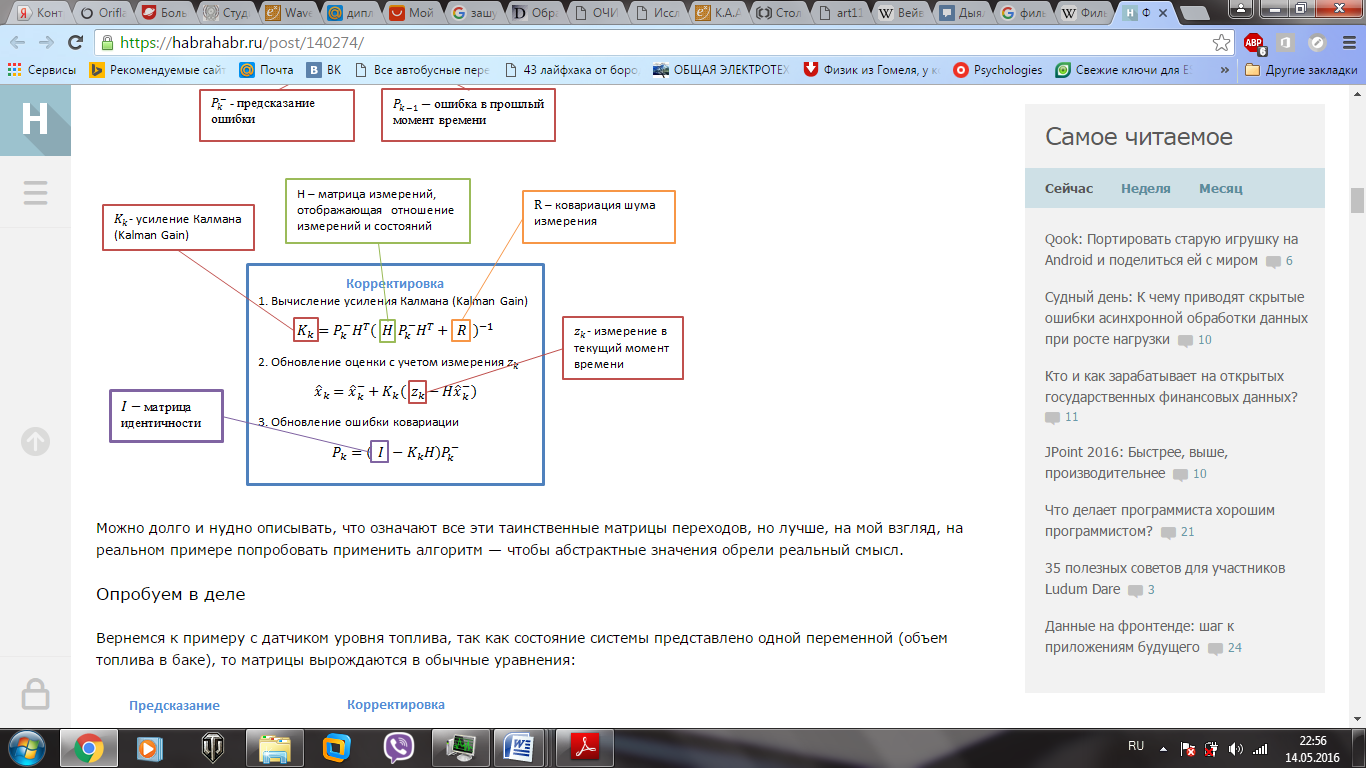


Рисунок 3.14 – Описание переменных корректировки для случая, представленного на рисунке 3.12

### 3.2.3 Режекторный фильтр

Режекторный [фильтр](https://ru.wikipedia.org/wiki/%D0%A4%D0%B8%D0%BB%D1%8C%D1%82%D1%80_(%D1%8D%D0%BB%D0%B5%D0%BA%D1%82%D1%80%D0%BE%D0%BD%D0%B8%D0%BA%D0%B0)) – [электронный](https://ru.wikipedia.org/wiki/%D0%AD%D0%BB%D0%B5%D0%BA%D1%82%D1%80%D0%BE%D0%BD%D0%BD%D1%8B%D0%B9_%D1%84%D0%B8%D0%BB%D1%8C%D1%82%D1%80) или любой другой фильтр, выделяющий [колебания](https://ru.wikipedia.org/wiki/%D0%9A%D0%BE%D0%BB%D0%B5%D0%B1%D0%B0%D0%BD%D0%B8%D0%B5) некоторой определённой полосы [частот](https://ru.wikipedia.org/wiki/%D0%A7%D0%B0%D1%81%D1%82%D0%BE%D1%82%D0%B0). Эта полоса подавления характеризуется шириной полосы заграждения и расположена приблизительно вокруг центральной частоты ω0 (рад/с) подавления, или fо=ω0/(2π) (Гц). Для реальной амплитудно-частотной характеристики частоты ωL и ωU представляют собой нижнюю и верхнюю частоты полосы подавления, как показано на рисунке 3.15.

Для очистки сигнала был выбран фильтр второго порядка, так как фильтры более высоких порядков подразумевают наличие в сигнале нескольких гармоник. Таким образом, использовался адаптивный режекторный фильтр второго порядка с передаточной функцией вида:

(3.5)

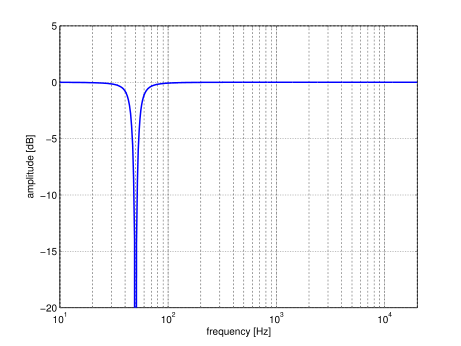


Рисунок 3.15 – АЧХ заграждающего фильтра промышленной частоты 50 Гц, применяемого в аудиоустройствах

В данном случае передаточная функция имеет один адаптивный параметр θ, представляющий собой искомую частоту. Если параметр подстройки r стремится к 1, то полоса пропускания фильтра может быть представлена в виде 2(1 – r) радиан [33], что позволяет параметру фильтра θ стремиться к искомой частоте. Таким образом, r должен удовлетворять условию 0 < r < 1, для обеспечения необходимой полосы пропускания фильтра. Тогда результат фильтрации y[n] для входной последовательности x[n] можно описать как:

Для уменьшения влияния ошибки применяется метод наименьших квадратов:

где μ – есть параметр сходимости фильтра, а функция градиента β(n) представима в виде:

### 3.2.4 Результаты применения фильтров для очистки сигнала

В качестве объектов исследований были выбраны речевые сообщения, составленные из отдельных частей переговоров авиадиспетчеров. Такие сообщения, содержащие различные аддитивные шумы, позволяют выбрать универсальный метод шумоподавления для данного типа речевых сигналов. Исходные сигналы зашумлены настолько, что с первого прослушивания невозможно четко распознать смысл речевого сообщения. Критерием для определения наилучшего метода шумоподавления являлись разборчивость речевого сообщения с первого прослушивания и определение отношения сигнал/шум, рассчитанного по формуле (2.14).

Исходный зашумленный сигнал представлен на рисунке 3.16. На всех представленных ниже графиках ось ординат – амплитуда сигнала, ось абсцисс – время, мс. Ниже представлены результаты очистки речевого сигнала с помощью МНК-фильтра, режекторного и фильтра Калмана.

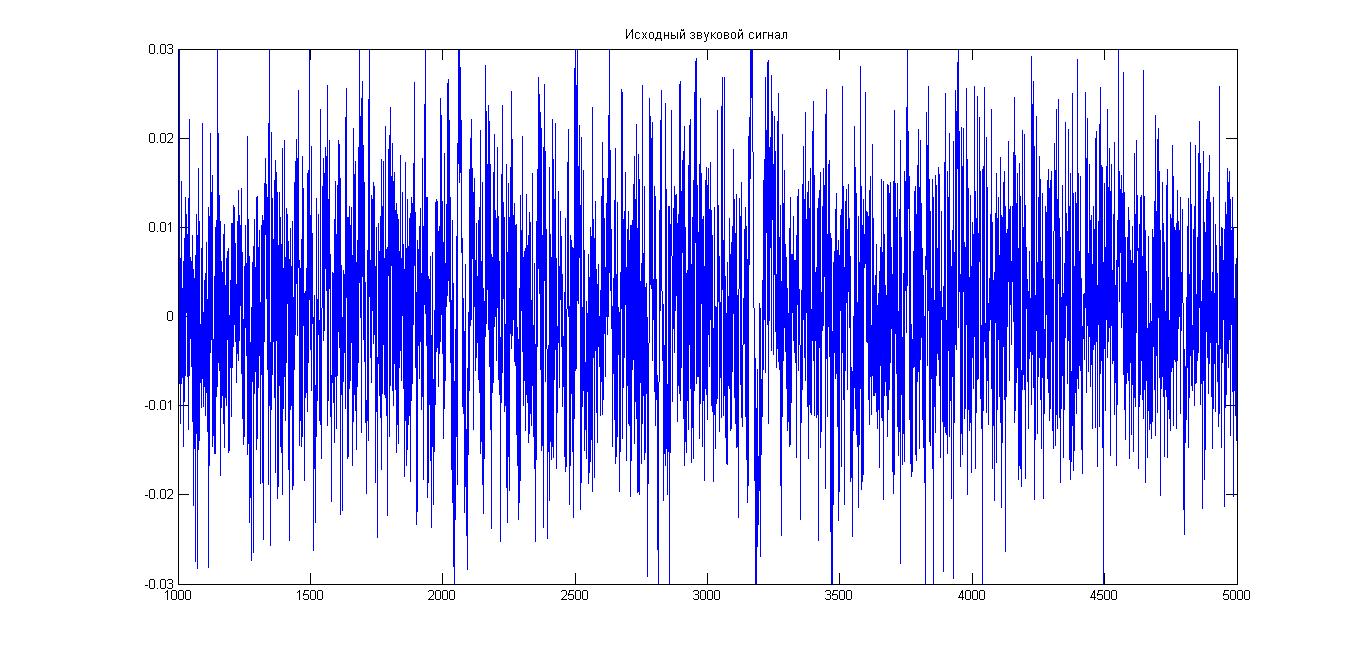


Рисунок 3.16 – Исходный звуковой сигнал

1. Результат применения МНК-фильтра представлен на рисунках 3.17, 3.18. Как упомянуто выше, одним из недостатков МНК-фильтра является подавление полезного сигнала вместе с шумом. Так же на рисунке 3.17 видно: сигнал не является гладким, что говорит о наличии шумов и недостаточной разборчивости сигнала. Это подтверждается полученными результатами (ОСШ 11.5135 дБ).

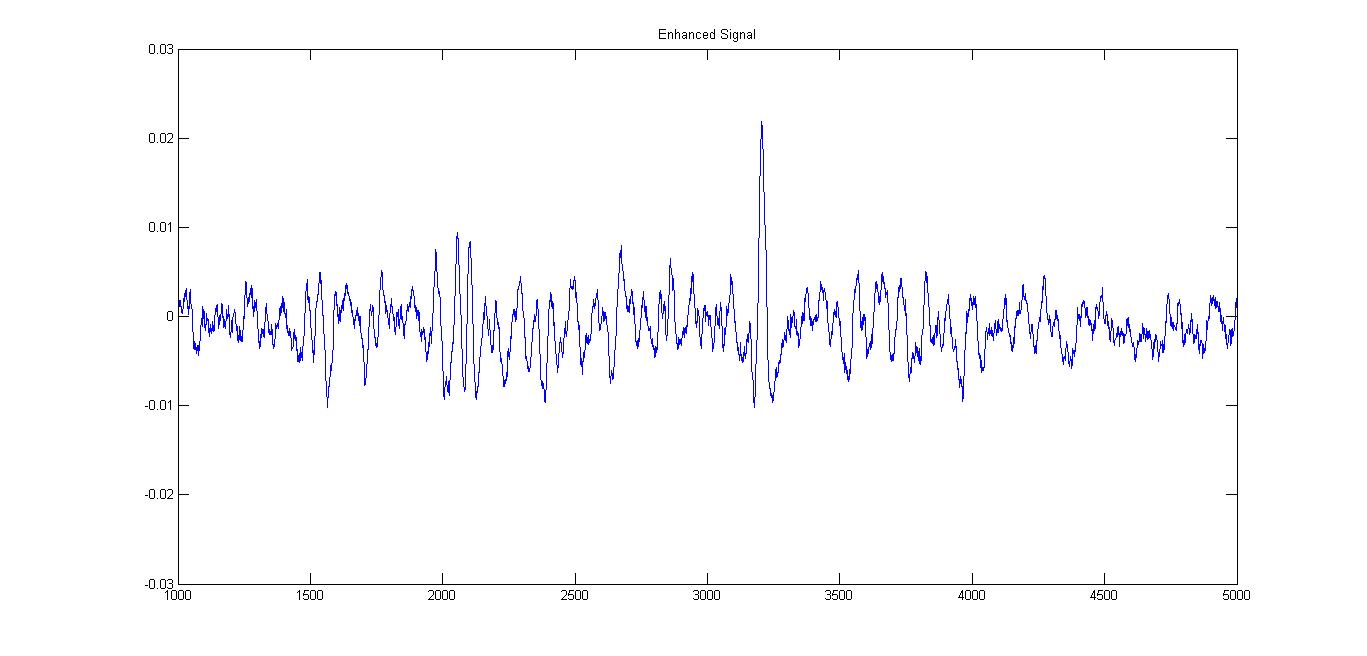


Рисунок 3.17 – Результат применения фильтра МНК

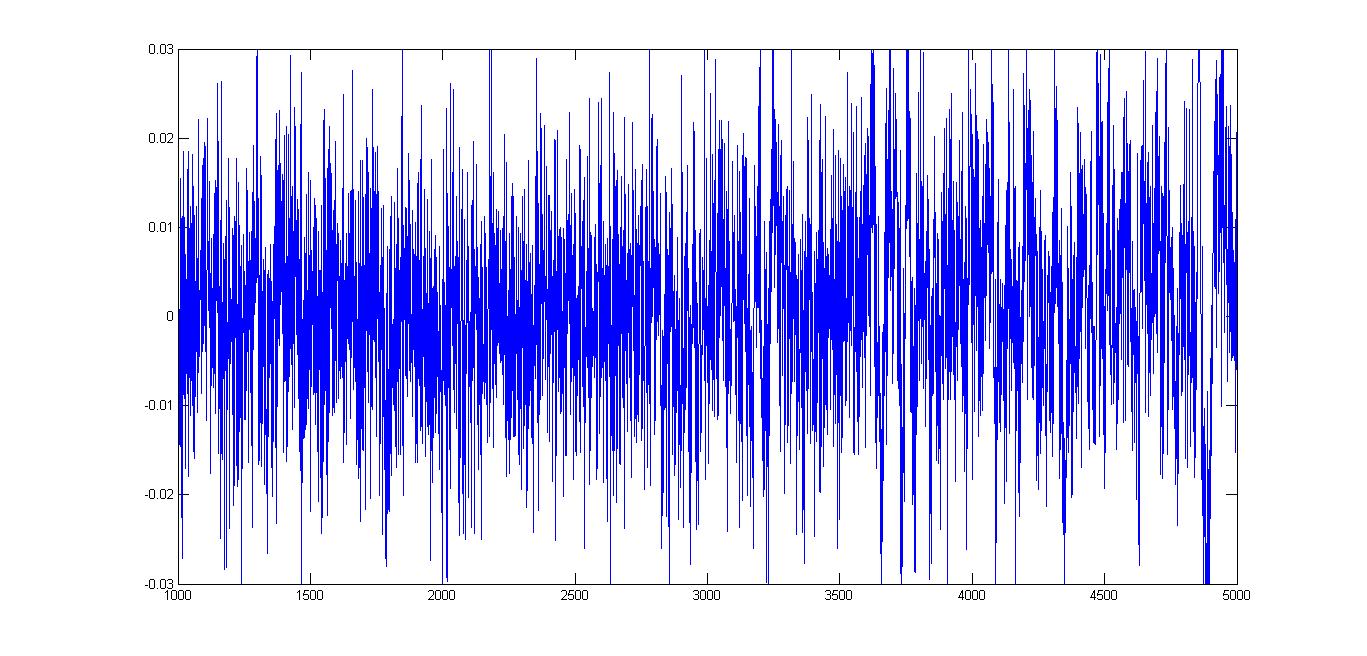


Рисунок 3.18 – Шум, удаленный МНК-фильтром

1. Результат применения режекторного фильтра представлен на рисунках 3.19 и 3.20. Результат работы фильтра выглядит зашумлено, однако если прослушать полученный результат и оценить ОСШ (13.2057 дБ), видно, что фильтр вырезал шумовую компоненту сигнала. Т.к. частота речи меняется при разговоре, а в сигнале присутствует преимущественно белый шум, который не подвергается законам изменения частоты.

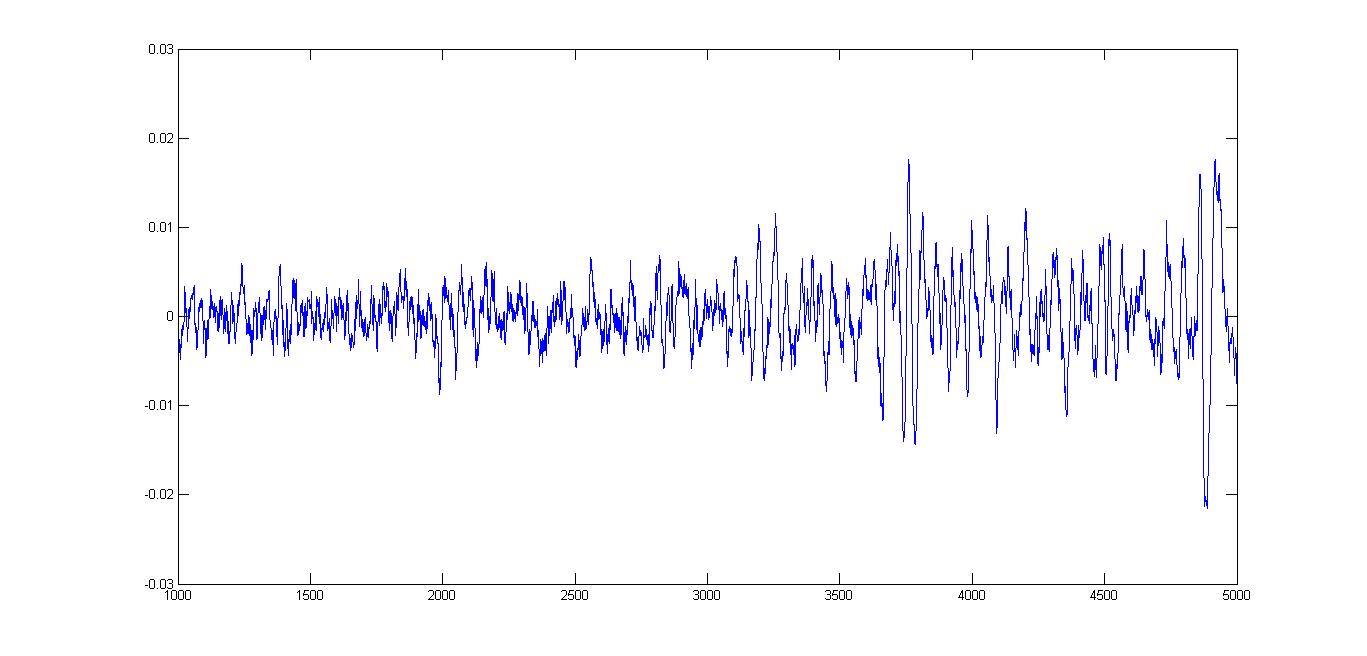


Рисунок 3.19 – Результат применения режекторного фильтра

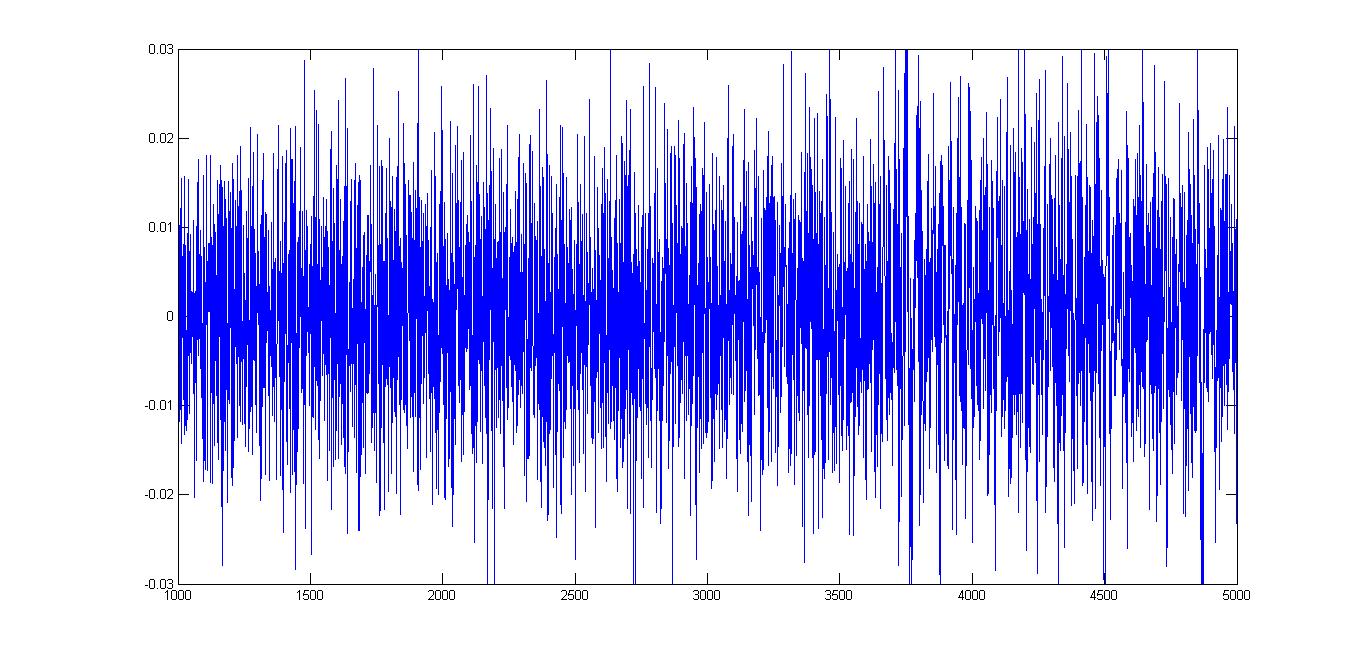


Рисунок 3.20 – Шум, вырезанный режекторным фильтром

1. Результат применения фильтра Калмана представлен на рисунках 3.21 и 3.22. Результат, полученный с применением фильтра Калмана, является лучшим по сравнению с результатами фильтров, представленных выше. Сигнал после применения фильтра Калмана является более гладким, а ОСШ составляет 16.7684 дБ. Это подтверждает универсальность данного фильтра и обосновывает его популярность.

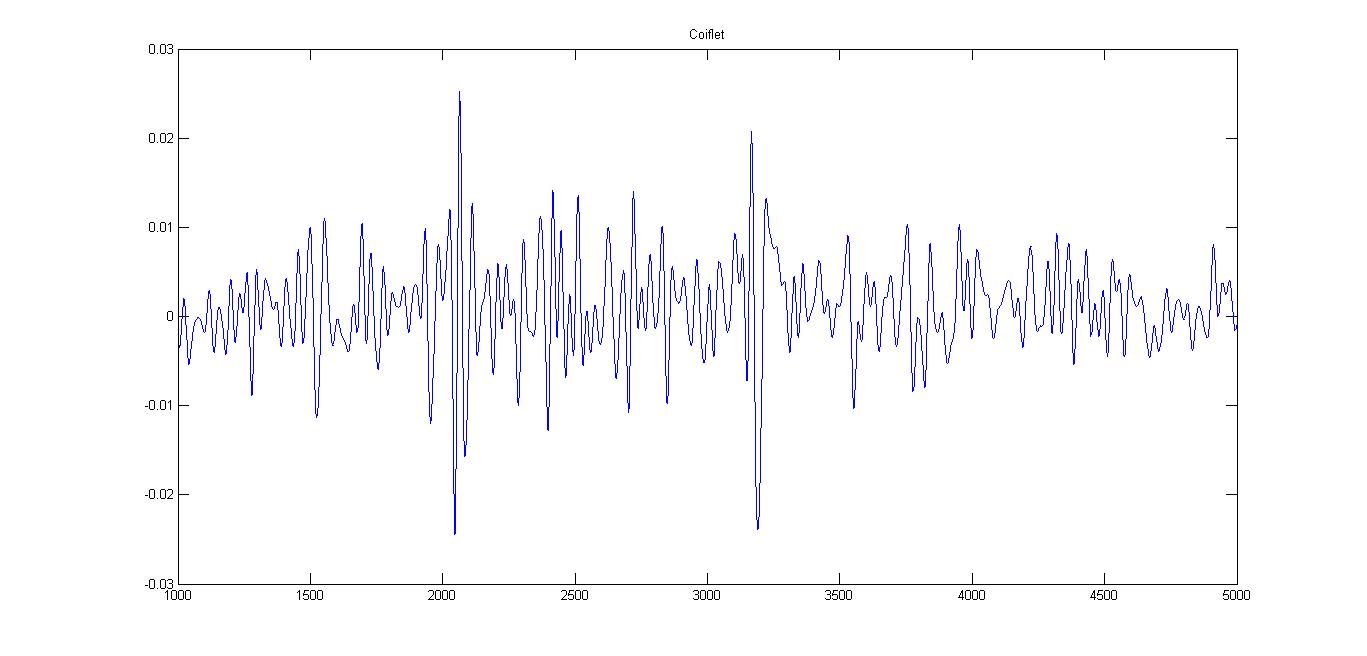


Рисунок 3.21 – Результат применения фильтра Калмана

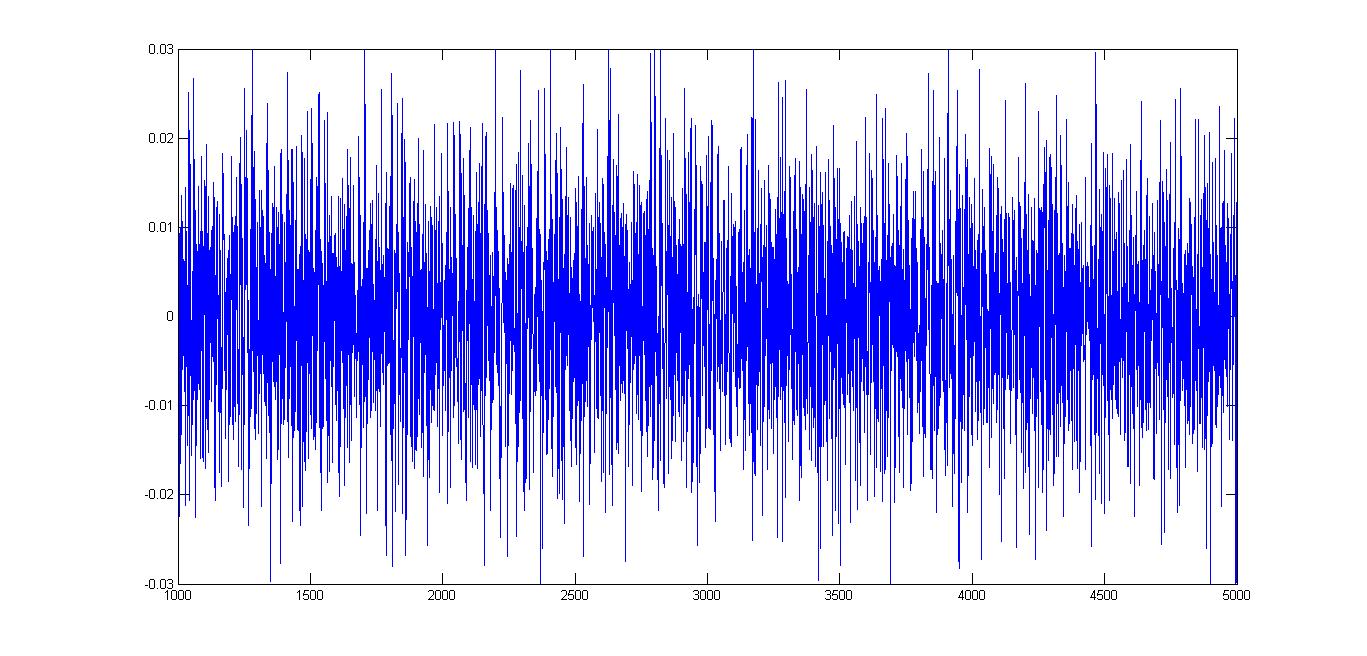


Рисунок 3.22 – Шум, вырезанный фильтром Калмана

Результаты вычисления отношения сигнал/шум (на входе и на выходе), полученные с помощью адаптивных фильтров представлены в таблице 3.1.

Таблица 3.1 – Сравнение результатов работы вейвлетов

|  |  |  |
| --- | --- | --- |
| **Фильтр** | **Отношение сигнал/шум на входе, дБ** | **Отношение сигнал/шум на выходе, дБ** |
| МНК | 8.4134 | 11.5135 |
| Режекторный | 8.4134 | 13.2057 |
| **Калмана** | **8.4134** | **16.7684** |

## 3.3 Выбор вейвлетов для очистки сигнала и результаты их применения

Как было упомянуто в главе 2.3, в качестве вейвлетов для решения задачи очистки переговоров авиадиспетчеров от шума были выбраны дискретный вейвлет Мейера, вейвлеты Добеши (db4) и симлет (sym8).

Исходный звуковой сигнал представлен на рисунке 3.16. Ниже представлены результаты применения вышеуказанных вейвлетов.

1. Для начального выбора порога фильтрации был использован минимаксный критерий с вейвлет-функцией Добеши, которая определяется выражением (2.5). Применение минимаксного критерия обычно позволяет хорошо отфильтровать сигнал, когда малые составляющие сигнала находятся в полосе шума [34]. Тем не менее, как показано на рисунках 3.17 и 3.18, применение этого критерия с вейвлетом db4 к сигналам в данном исследовании показало, что часть полезного сигнала также была отфильтрована, что снизило разборчивость исходного сообщения, хотя ОСШ повысилось и его значение составило 10.7868 дБ.

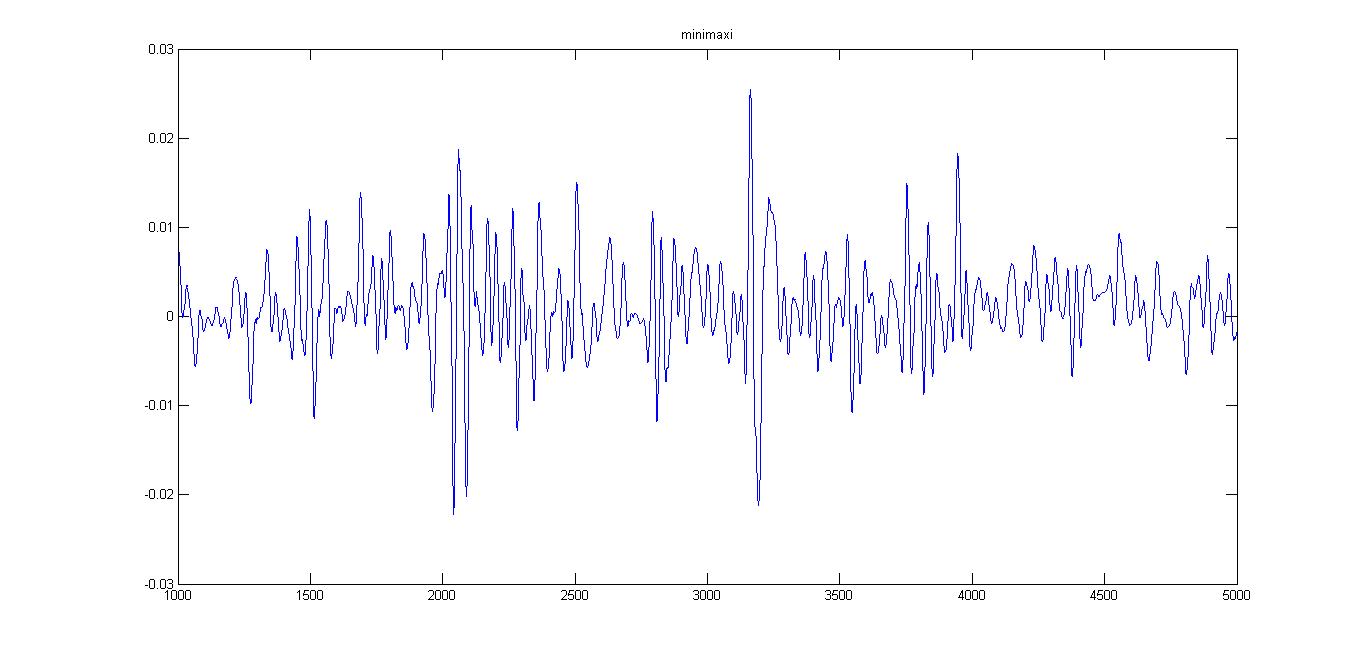


Рисунок 3.17 – Результат применения вейвлета Добеши db4

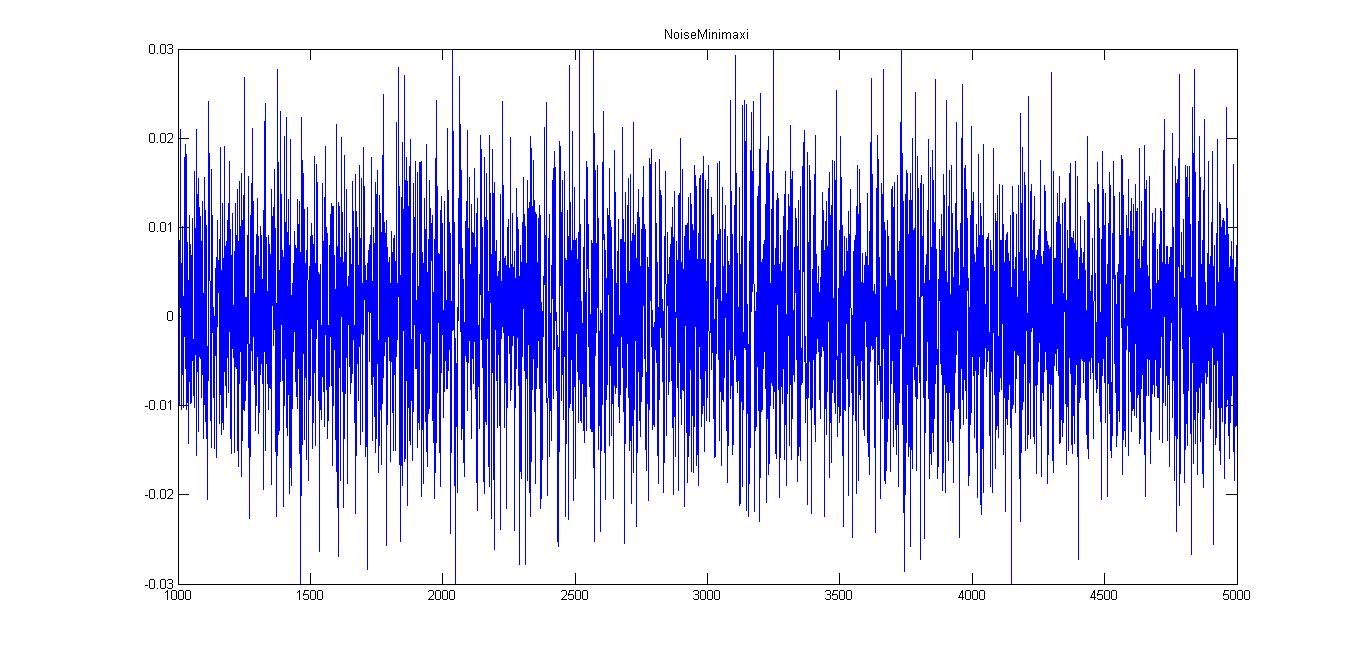


Рисунок 3.18 –Шум, вырезанный вейвлет-преобразованием Добеши db4

1. Применение симлета sym8 (коэффициенты симлета рассчитываются аналогично коэффициентам фильтра Добеши) с универсальным критерием Донохо-Джонстона ‘sqtwolog’ (порог равный , где х – длина сигнала), с жестким видом трешолдинга позволило отфильтровать высокочастотные компоненты шума, однако гладкость полученного сигнала не улучшилась, что видно из рисунков 3.19 и 3.20. Это связано с выбранным порядком симлета, но увеличение порядка ведет к значительно большим вычислительным затратам, поэтому именно восьмой порядок считается наиболее оптимальным [35]. Однако и применение к исходному сигналу более высокого порядка симлета не привело к улучшению результата. Таким образом, применение симлета (как порядка 8, так и 10) не повысило разборчивость речевого сообщения по сравнению с применением вейвлета Добеши (db4).

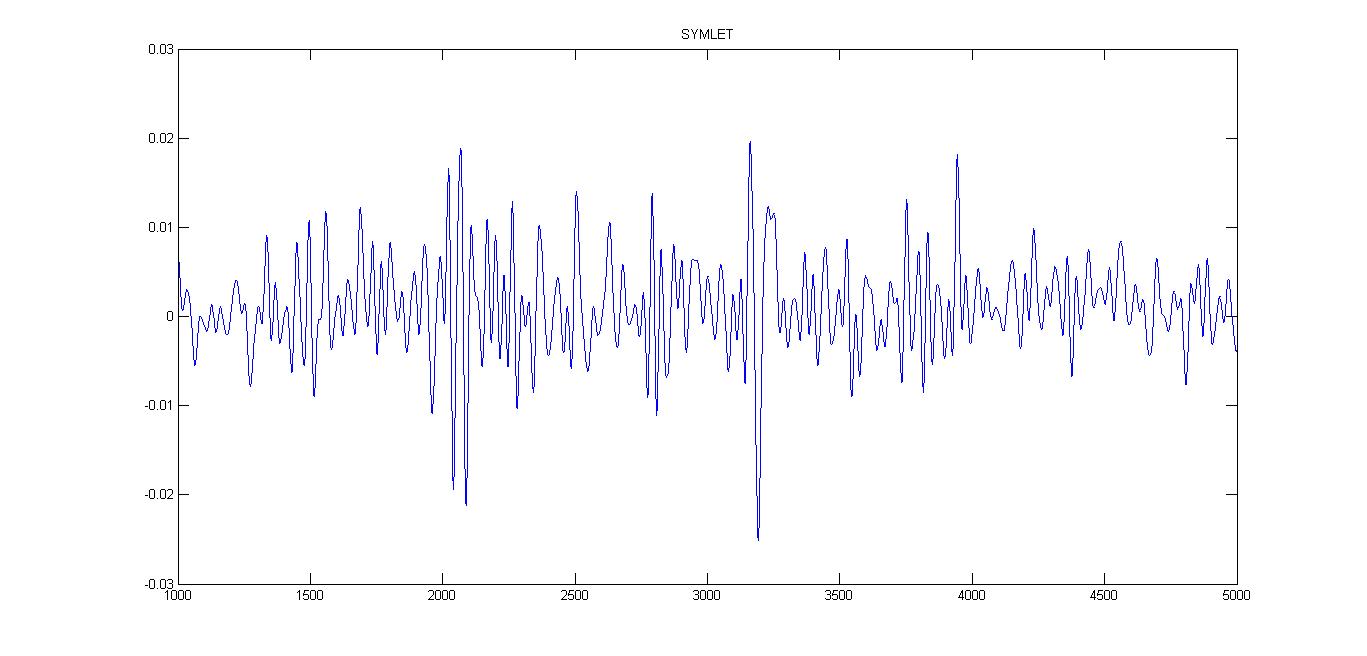


Рисунок 3.19 – Результат применения симлета sym8

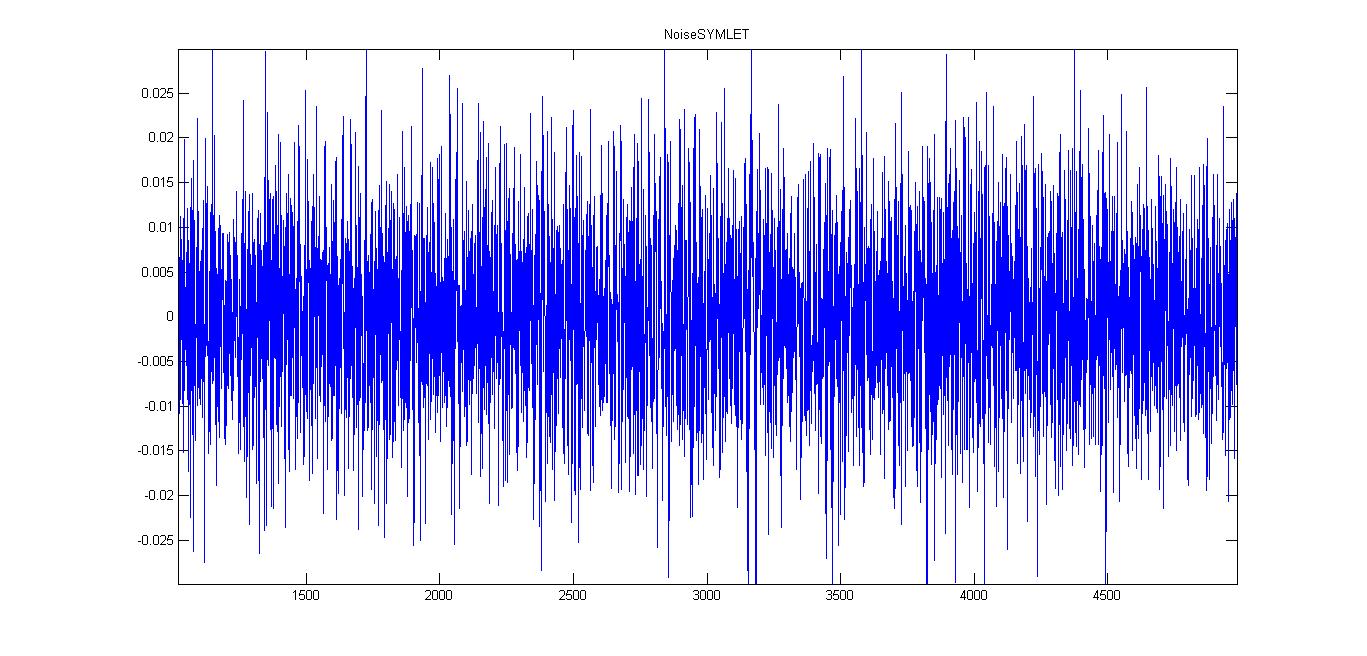


Рисунок 3.20 – Шум, вырезанный симлетом sym8

1. С помощью дискретного вейвлета Мейера с модифицированным критерием Штайна ‘heursure’ с мягким видом трешолдинга удается достигнуть большей по сравнению с симлетом sym8 гладкости наряду с хорошим отделением высокочастотной составляющей шума, а также сохранением информативной составляющей исходного сигнала. Результаты применения вейвлета Мейера представлены на рисунках 3.21 и 3.22.

Вейвлет-функции Мейера определены в частотной области следующим образом:

 (3.9)

 (3.10)

 (3.11)

Результат применения вейвлета Мейера – повышение разборчивости речевого сообщения до уровня понимания с первого прослушивания. Было также установлено, что применение в этом случае жесткого вида трешолдинга не приводит к улучшению разборчивости по сравнению с мягким трешолдингом.

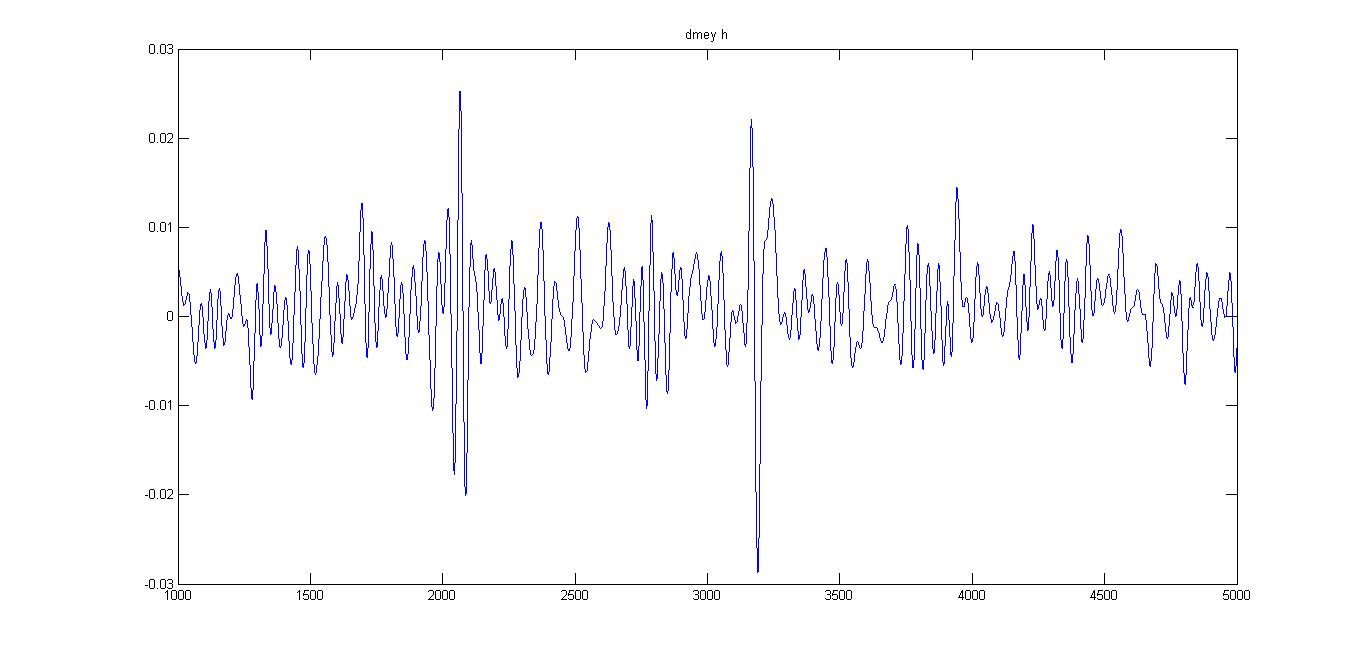


Рисунок 3.21 – Результат применения дискретной вейвлет-функции Мейера

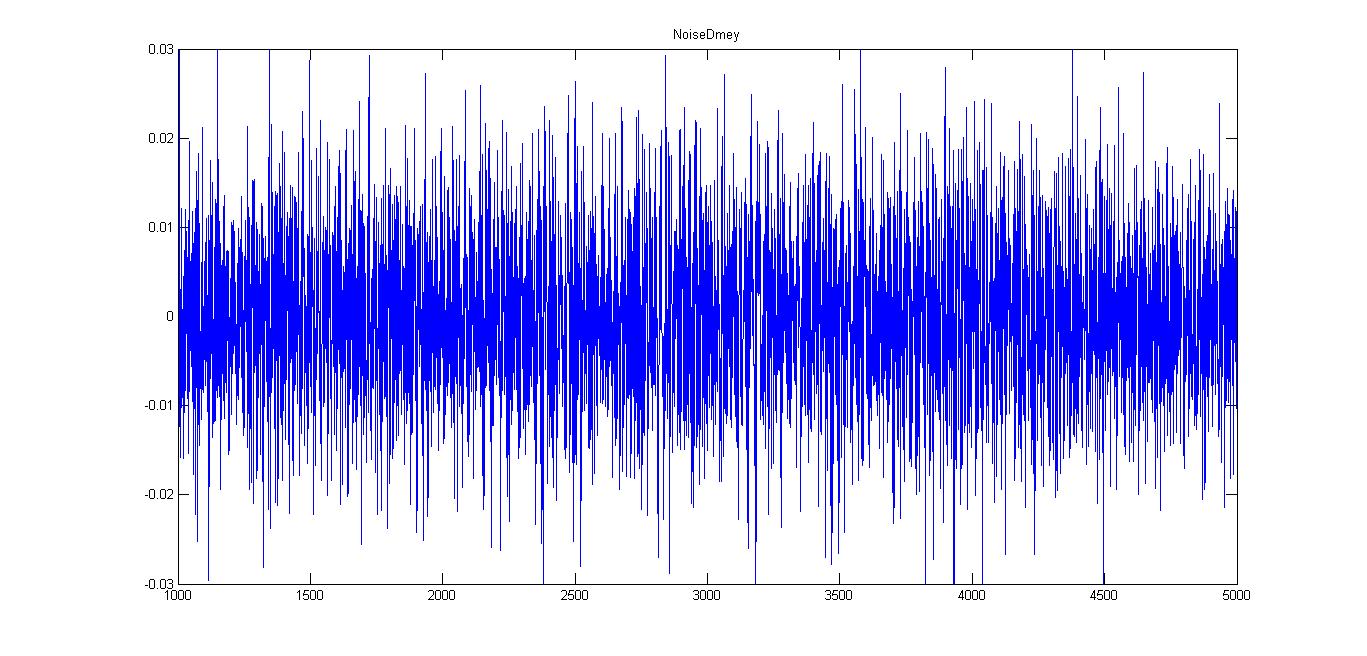


Рисунок 3.22 – Шум, вырезанный дискретной вейвлет-функцией Мейера

Полученные результаты также подтверждены вычислением отношения сигнал/шум на входе и на выходе.

## 3.4 Сравнение результатов

Проведенные исследования показывают, что лучшими средствами из выбранных методов очистки речевого сигнала со случайными помехами является использование дискретной вейвлет-функции Мейера и фильтра Калмана. Они увеличили отношение сигнал/шум на выходе на 10.7632 дБ и 8.355дБ соответственно, тем самым увеличив разборчивость речи в несколько раз.

Это говорит об универсальности выбранных методов для шумоподавления случайных шумов (звук двигателя и иные помехи) в записях разговоров авиадиспетчеров, т.к. в записи был не один переговор, а нарезка из нескольких случайно выбранных переговоров.

В таблице 3.2 представлены значения отношения сигнал/шум на входе и выходе системы шумоподавления при применении разных вейвлетов.

Таблица 3.2 – Сравнение результатов работы вейвлетов

|  |  |  |
| --- | --- | --- |
| **Вейвлет** | **Отношение сигнал/шум на входе, дБ** | **Отношение сигнал/шум на выходе, дБ** |
| Добеши | 8.4134 | 10.7868 |
| Симлет | 8.4134 | 15.3636 |
| **Дискретная вейвлет-функция Мейера** | 8.4134 | **19.1766** |

# ЗАКЛЮЧЕНИЕ

В результате выполнения дипломной работы были получены следующие результаты:

1. Исследованы существующие современные методы и технологии шумоподавления.
2. Изучена теория вейвлет-преобразования и рассмотрена возможность его применения в решении задачи шумоподавления в речевых сигналах.
3. Рассмотрены и реализованы методы фильтрации речевых сигналов оптимальными линейными фильтрами: МНК фильтром, режекторным фильтром и фильтром Калмана. Наилучший результат из данных фильтров показало применение фильтра Калмана (отношение сигнал/шум увеличилось на 8.355 дБ). МНК фильтр прост в использовании и не требует дополнительной памяти, однако недостатком МНК-фильтра было выявлено подавление полезного сигнала вместе с шумом. Недостатком режекторного фильтра является наличие шумовой компоненты в сигнале на выходе, однако разборчивость речи по сравнению с результатом, полученным МНК фильтром, повысилась, ОСШ увеличилось на 1,69дБ.
4. Проведен сравнительный анализ реализованных методов при решении задачи шумоподавления в речевых сигналах. Полученные результаты свидетельствуют, что использование дискретной вейвлет-функции Мейера и фильтра Калмана дает наилучшие результаты в очистке подобных речевых сигналов.

Результаты проделанной работы показывают, что для очистки речевых сигналов от шумов могут использоваться как традиционные методы очистки (фильтр Калмана), так и вейвлеты (дискретная вейвлет-функция Мейера). Однако недостатком применения вейвлетов является использование большой вычислительной мощности.

# СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Paul White “Sound On Sound”, Noise‑reduction Tools & Techniques, January 2012.
2. Чучупал В.Я., Чичагов А.С., Маковкин К.А. Цифровая фильтрация зашумленных речевых сигналов: Сообщения по программному обеспечению ЭВМ. М.: 28 Вычислительный центр РАН, 1998. 52 с.
3. Widrow B., at al. Adaptive Noise Candelling: Principles and Applications. Proc IEEE,Vol63, No. 12, 1975
4. Санников В.Г., Журавский Ю.И. Прохоров Ю.Н. Формирование банка априорных данных о речи диктора. Материалы всесоюзного семинара АРСО-12,Киев, 1982.
5. Сейдж Э., Мелс Дж. Теория оценивания и ее применение в теории связи и управления. М., Связь, 1976.
6. Lee K.Y., Lee B.-G., Song I. etc. Recursive Speech Enhancement Using the EM Algorithm with Initial Conditions Trained by HHM’s. Int Conf on Acoustics, Speech and Signal Proc, ICASSP-96, 1996.
7. Yoo C. Selective All Pole Modeling of Degraded Speech Using M-Band Decomposition. Int.Conf on Acoustics, Speech and Signal Proc. ICASSP-96, 1996.
8. Hansen G.H.L, Pellom B.L. Text-directed speech enhancement employing phone class parsing and feature map constrained vector quantizaion. Speech Communication, Vol 21, 1997.
9. Sheikhzadeh H., Sameti H., Deng L. Comparative Performance of Spectral Subrtraction and HMM Based Speech Enhancement Strategies with Application to Hearing Aid Design. Proc. ICASSP-94.
10. McKinley B.L., Whipple G.H. Noise Model Adaptation in Model Based Speech Enhancement. Int.Conf on Acoustics, Speech and Signal Processing, ICASSP-96, 1996.
11. Ephraim Y. Statistical model-based speech enhancement systems” Proc. IEEE, Vol 8, No 10, 1992.
12. Sambur M.R. Adaptive Noise Cancelling for Speech Signals. IEEE Trans. ASSP,vol.ASSP-26, 1978.
13. Lyon R.F. A computational Model of Filtering, Detection and Compression in the Cocklea. Proc, Int.Conf. on Acoust.Speech and Signal Proc. ICASSP-1982.
14. Malah D., Cox R.V. A Generalized Comb Filtering Technique for Speech enhancement. Proc 1982 IEEE Int. Conf ASSP.
15. Frasier R.E. , etc. Enhancement of Speech by Adaptive Filtering. Proc.1976 IEEE Int. Conf. ASSP, ICASSP-76.
16. Drucker H. Speech Processing in a High Ambient Noise Environment. IEEE Trans. On ASSP, ICASSP-76.
17. Fechner T. Nonolinear noise filtering with neural networks: comparison with Wiener optimal filtering” Proceedings of IEE conf. On Artificial Neural Networks, 1993.
18. Scalart P., Benamar A. A system for speech enhancement in the context of hands-free radiotelephony with combined noise reduction and acoustic echo cancellation. Speech Communication, Vol.20, 1996.
19. Новоселов С.А.. Подавление шума в речевых сигналах на основе метода нелокального усреднения. Цифровая Обработка Сигналов №1/2012.
20. <http://wikisound.org>.
21. Яковлев А.Н. Введение в вейвлет-преобразования: Учеб. пособие. –Новосибирск: Изд-во НГТУ, 2003. – 104 с.
22. Айфичер, Эммануил С., Джевирс, Барри У. Цифровая обработка сигналов: практический подход, 2-е издание.: Перевод с англ. – М. Издательский дом «Вильямс», 2008.
23. Московский С.Б., Сергеев А.Н., Лалина Н.А. Очистка сигнала от шумов с использованием вейвлет-преобразования // Universum: Технические науки : электрон. научн. журн. 2015. № 2 (15).
24. Meyer, Y. Wavelets and operators / Y. Meyer. – Cambridge: Cambridge University Press, 1993.
25. Дремин, И. M. Вейвлеты и их применение / И. М. Дремин, О. В. Ива-нов, В. А. Нечитайло // Успехи физических наук. – 2001. – Т. 171.].
26. Addison, P. S. The illustrated wavelet transform handbook: applications in science, engineering, medicine and finance / P. S. Addison. – Bristol; Philadelphia: IOP Publishing, 2002.
27. Vetterli, M. Wavelets and subband coding / M. Vetterli, J. Kovacevic. – NJ: Prentice Hall, 1999.
28. Jansen, M. Noise reduction by wavelet thresholding / M. Jansen. – New York: Springer-Verlag, 2001.
29. ЯСИН Алаулдин Салах Ясин. Фильтрация зашумленных сигналов и изображений с помощью вейвлет-преобразования: диссертация на соискание ученой степени кандидата физико-математических наук. Саратовский государственный университет им. Н.Г. Чернышевского, Саратов, 2016.
30. Назаров М.В., Прохоров Ю.Н. Методы цифровой обработки и передачи речевых сигналов.- М.: Радио и связь, 2005.
31. В.Е. Гантмахер, Н.Е. Быстров, Д.В. Чеботарев Шумоподобные сигналы. Анализ, синтез, обработка. – СПб.: Наука и Техника, 2005.
32. Рабинер Л.Р., Шафер Р.В. Цифровая обработка речевых сигналов. -М.: Радио и связь, 2011.
33. L. Tan, J. Jiang, L. Wang, Adaptive Harmonic IIR Notch Filters for Frequency Estimation and Tracking. InTech, 2011.
34. Дьяконов В.П.MATLAB и SIMULINK для радиоинженеров. – М.:ДМК Пресс, 2011.
35. Смоленцев Н.К. Основы теории вейвлетов. Вейвлеты в MATLAB. -М.: ДМК Пресс, 2005.- 300с.

# ПРИЛОЖЕНИЕ А

1. **Реализация МНК фильтра**

clear all

close all

filename = 'NEW\_CUT\_Signal.wav';

x = wavread(filename);

Signal = x(:,1);

SignalPower = mean(Signal.^2);

%program to illustrate adaptive filtering using the LMS algorithms

% X delayed input data vector

% Y measured signal

% W coefficient vector

% E enhanced signal

N=30; % filter length

M=0; % delay

w0=1; % initial value for adaptive filter coefficients

SF=2048; % factor for reducing the data samples - 11 bit ADC assumed

mu=0.05;

X = zeros(N,1);

delay = zeros(1,M+1);

W = w0\*ones(N,1);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%in = fopen('ADF.dat','r'); %read input data from specified data file

Y =Signal./SF;

if w0==0

sf = SF; % scaling factor for display

else

sf = SF/N/w0;

end

for i=1:length(Y)

if M>0

delay(2:M+1) = delay(1:M); % shift data for delay

end

delay(1) = Y(i);

X(2:N) = X(1:N-1); % update buffer

X(1) = delay(M+1);

E(i) = Y(i)-W'\*X; % the enhanced signal

W = W + 2\*mu\*E(i)\*X; % update the weights

end

%subplot(2,1,1),

figure;

plot(Y\*SF); title('Input Signal'); %1:length(Y),

axis([1000,5000,-0.03,0.03]);

%subplot(2,1,2),

figure;

plot(E\*sf); title('Enhanced Signal'); %1:length(E)

axis([1000,5000,-0.03,0.03]);

figure;

Noise = Y-E';

plot(Noise); title('Noise');

axis([1000,5000,-0.03,0.03]);

NoisePower = mean(Noise.^2);

1. **Реализация режекторного фильтра**

clear all

close all

% X delayed input data vector

% Y measured signal

% W coefficient vector

% E enhanced signal

filename = 'NEW\_CUT\_Signal.wav';

x = wavread(filename);

Signal = x(:,1);

SignalPower = mean(Signal.^2);

N=1; % filter length

M=0; % delay

w0=1; % initial value for adaptive filter coefficients

SF=16; % factor for reducing the data samples

mu=0.005;

r = 0.95;

X = zeros(N,1);

delay = zeros(1,M+1);

W = w0\*ones(N,1);

W1 = w0\*ones(N,1);

W2 = w0\*ones(N,1);

W3 = w0\*ones(N,1);

x1 = zeros(N,1);

x2 = zeros(N,1);

c = 1;

t0 = 60;

count = t0;

J = ones(1,3);

w = [1 ; 1 ; 2\*pi/60];

lam = 1e2;

angle = zeros(1,1000);

Y=Signal;

Y = Y./SF;

theta(3) =0;

E(1) = 1;

E(2) = 1;

beta(2) =0;

beta(1) = 0;

for i = 3:length(Y)

E(i) = Y(i) - 2\*cos(theta(i))\*Y(i-1) + Y(i-2) + 2\*r\*cos(theta(i))\*E(i-1) - r\*r\*E(i-2);

beta(i) = 2\*sin(theta(i))\*Y(i-1) - 2\*r\*sin(theta(i))\*E(i-1) + 2\*r\*cos(theta(i))\*beta(i-1) - r\*r\*beta(i-2);

dw = 2\*mu\*beta(i)\*E(i);

theta(i+1) = theta(i) + dw;

end

% plot(theta);

figure

% subplot(3,1,1),

hold on

plot(1:length(Y),Y\*SF);

axis([1000,5000,-0.03,0.03]);

hold on

% plot(1:length(E),E\*SF,'r');

plot(E(200:end)\*SF, 'y');

axis([1000,5000,-0.03,0.03]);

% title('Enhanced');

% subplot(3,1,3),

% hold on

figure

delta = Y(1:end)\*SF - (E(1:end)).'\*SF;

plot(delta(50:end),'black');

axis([1000,5000,-0.03,0.03]);

hold on

plot(Y(50:end)\*SF, 'gr');

axis([1000,5000,-0.03,0.03]);

1. **Реализация фильтра Калмана**

*Файл TEST\_CN.m:*

clc;close all;clear;

speech= wavread('NEW\_CUT\_Signal.wav') ;

%noise=data;

signal\_clear=speech;

%%%%%% Defines

N=length(speech);

sigmaq=0.2;

psignal=signal;

%%Kalman Filtering

output=KalmanDenoisingCN(signal,signal\_clear,44100);

rsignal=output';

Lo=length(output);

figure;

subplot(311);plot(psignal);title('noisy signal');

subplot(312);plot(signal\_clear);title('clear speech');

subplot(313);plot(rsignal);title('denoised signal');

%Calculate SNR

SNR\_noisysignal=10\*log10(sum(signal\_clear(1:Lo).^2)/sum((psignal(1:Lo)-signal\_clear(1:Lo)).^2))

SNR\_denoisedsignal=10\*log10(sum(signal\_clear(1:Lo).^2)/sum((rsignal-signal\_clear(1:Lo)).^2))

wavwrite(psignal,16000,'noisy signal CN');

wavwrite(rsignal,16000,'denoised signal CN');

*Файл KalmanDenoisingCN.m:*

function output=KalmanDenoisingCN(Noisy,Clean,fs)

fs=44100;

single='in.wav'

W=fix(0.040\*fs);

SP=1;

p=8;

q=6;

Window=ones(W,1);

a = Noisy;

s = Clean;

y = Noisy;

n = Noisy-Clean;

V=var(Noisy-Clean);

R=var(n);

FsUpper=[zeros(p-1,1) eye(p-1)];

FnUpper=[zeros(q-1,1) eye(q-1)];

I=eye(p+q);

Hs=[zeros(1,p-1) 1];

Gs=Hs';

Hn=[zeros(1,q-1) 1];

Gn=Hn';

G=[Gs zeros(p,1);zeros(q,1) Gn];

H=[Hs Hn];

P=diag(repmat(V,1,p+q));

o=zeros(1,W\*size(s,2));%

hwb = waitbar(0,'Please wait...','Name','Processing');

start=p+1;

X=[y(1:p,1);ones(q,1).\*0.005];%іхЦµ

t=p+1;

[B E]=lpc(Noisy-Clean,q);

[A D]=lpc(y,p);

for k=1:3

for n=1:size(s,2)

waitbar(n/size(s,2),hwb,['Please wait... ' num2str(fix(100\*n/size(s,2))) ' %'])

Fs=[FsUpper; fliplr(-A(n,2:end))];

Fn=[FnUpper; fliplr(-B(2:end))];

F=[Fs zeros(p,q);zeros(q,p) Fn];

for i=start:W

%Time update(predict)

X\_=F\*X;

% Q=[D(n) 0;0 V];

Q=[D(n) 0;0 R(n)];

P\_=F\*P\*F'+G\*Q\*G';%

K=P\_\*H'\*inv(H\*P\_\*H');

X=X\_+K\*(y(i,n)-H\*X\_);

P=(I-K\*H)\*P\_;

o(t-p+1:t)=X(1:p,1)'; %Notice that the previous p-1 output samples are updated again

t=t+1;

end

start=1;

end

z=frame(o,W,SP,Window);

[A D]=lpc(z,p);

t=p+1;

start=p+1;

X=[y(1:p,1);ones(q,1).\*0.005];

figure;

plot(o);

end

close(hwb)

output=o;

1. **Реализация вейвлет-преобразований**

clear all

close all

filename = 'NEW\_CUT\_Signal ';

x = wavread(filename);

Signal = x(:,1);

%sound(Signal,44100);

figure;

plot(Signal); title('Исходный звуковой сигнал');

axis([1000,5000,-0.03,0.03]);

SignalPower = mean(Signal.^2);

lev = 4;

figure;

SignalSqtwolog = wden(Signal, 'sqtwolog', 'h', 'one', lev, 'sym8');

plot(SignalSqtwolog), title('SYMLET');

axis([1000,5000,-0.03,0.03]);

SignalSqtwologPower = mean(SignalSqtwolog.^2);

%sound( xds,44100);

%

figure;

NoiseSqtwolog = Signal-SignalSqtwolog;

plot(NoiseSqtwolog); title ('NoiseSYMLET');

axis([1000,5000,-0.3,0.3]);

NoiseSqtwologPower = mean(NoiseSqtwolog.^2);

figure

lev =4;

SignalHeursure = wden(Signal, 'heursure', 'h', 'one', lev, 'dmey');

plot(SignalHeursure), title('dmey h');

axis([1000,5000,-0.03,0.03]);

SignalHeursurePower = mean(SignalHeursure.^2);

figure;

NoiseHeursure = Signal-SignalHeursure;

plot(NoiseHeursure); title ('NoiseDmey');

axis([1000,5000,-0.03,0.03]);

NoiseHeursurePower = mean(NoiseHeursure.^2);

figure

lev=4;

SignalRigrsure = wden(Signal, 'minimaxi', 's', 'sln', lev, 'db4');

plot(SignalRigrsure), title('minimaxi');

axis([1000,5000,-0.03,0.03]);

SignalRigrsurePower = mean(SignalRigrsure.^2);

%sound(SignalRigrsure,44100);

figure;

NoiseRigrsure = Signal-SignalRigrsure;

plot(NoiseRigrsure); title ('NoiseMinimaxi');

axis([1000,5000,-0.03,0.03]);

NoiseRigrsurePower = mean(NoiseRigrsure.^2);