

МИНОБРНАУКИ РОССИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «ВГУ»)

Физический факультет

Кафедра электроники

Микрофонная решетка с адаптивным диаграммообразованием

Магистерская диссертация

Направление 03.04.03 «Радиофизика»

Профиль «Системы телекоммуникаций и радиоэлектронной борьбы»

Зав. кафедрой



д. ф.-м. н., доц.

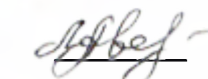
Г. К. Усков _____.20__ г.

Обучающийся



Д. А. Усиков

Руководитель



д. ф.-м. н., доц.

Л. И. Аверина

Воронеж 2023

Реферат

УДК 621.396.946

Усиков Д.А.

Микрофонная решетка с адаптивным диаграммообразованием

Магистерская диссертация по специальности 03.04.03 «Радиофизика», Воронеж, ВГУ, 2023 г. – 62 стр., 46 рис. и 19 источников.

Ключевые слова: микрофонная решетка, формирователь луча, адаптивные алгоритмы, пространственная фильтрация, диаграмма направленности, отношение сигнал/шум, оценка PESQ, переходные процессы, PDM, PCM.

В рамках данной работы проведено моделирование приема речи диктора в присутствии помехового сигнала с помощью микрофонной решетки. Изучены адаптивные временные формирователи луча на основе алгоритмов LC RLS и LC NLMS. Осуществлено адаптивное выделение полезного сигнала из принятой смеси. Проведён сравнительный анализ обозначенных алгоритмов по критериям выигрыша в ОСШ и оценки качества речи PESQ. Проанализирована эффективность работы микрофонной решетки вместе с адаптивным формирователем луча в случае быстро меняющейся помеховой обстановки и в случае отклонения элемента микрофонной решетки от ожидаемой позиции. Предоставлена схема реализации микрофонной решетки на базе микроконтроллера STM32H745ZI и четырёх MEMS микрофонов.

Содержание

Введение.....	4
1. Алгоритмы многоканальной обработки речевых сигналов.....	6
1.1. Моделирование приёма акустического сигнала на плоскую микрофонную решётку.....	6
1.2. Формирователи луча.....	10
1.3. Пространственная фильтрация с применением плоской микрофонной решетки.....	14
1.4. Методы оценки качества речи.....	20
2. Компьютерное моделирование многоканальной системы.....	24
2.1. Постановка задачи приёма и фильтрации сигнала с помехой.....	24
2.2. Сравнительный анализ алгоритмов LC NLMS и LC RLS.....	31
2.3. Эффективность работы микрофонной решетки в быстро меняющейся помеховой обстановке.....	38
2.4. Эффективность работы микрофонной решетки при отклонении её элементов.....	43
3. Аппаратная реализация микрофонной решетки.....	49
Заключение.....	59
Список литературы.....	60

Введение

Пространственно-частотная фильтрация применяется в различных областях науки для избирательного приёма сигналов от объектов, находящихся в различных точках пространства. Такая фильтрация достигается при помощи использования приёмных систем с характеристиками направленности заданной формы. Одним из таких устройств является микрофонная решетка. Микрофонная решетка представляет собой направленный микрофон, реализованный как множество приёмников звука, работающих согласованно. Она состоит из всенаправленных микрофонов, рассредоточенных по периметру пространства. Такие элементы микрофонной решетки принимают сигналы, содержащие информацию о звуках, поступающих со всех сторон. Совместная обработка полученных реализаций позволяет отфильтровать звуковой сигнал, исходящий из заданного направления [1, 2]. Микрофонные решетки нашли своё применение не только в системах речевой связи, но и в сфере безопасности. Так, например, система «Бумеранг» используется для защиты транспортных средств и войск от снайперского огня, позволяя определить положение стрелка. Система «SENTRI», используемая в городах, способна различить звук выстрела, определить его направление и вызвать полицию. Так же микрофонные решетки используются для получения объёмной звуковой картины, что нашло своё применение в автомобильной промышленности. Системы микрофонов позволяют определить точное место нежелательных источников шумов. Для сбора речевого сигнала в присутствии шума обычные микрофоны с последующим усилением сигнала не подходят, поскольку вместе с полезным сигналом усиливаются и помехи, что приводит к снижению разборчивости и качества речи. Однако микрофонные решетки могут с этим справиться, если источник полезного сигнала и источник шума имеют различные пространственные положения.

Целью данной работы является изучение цифровой обработки акустических сигналов с помощью плоской микрофонной решетки и адаптивных формирователей луча.

Исходя из поставленной цели, решались следующие задачи:

- изучить литературу по данной теме
- смоделировать приём полезного сигнала и помехи для микрофонной решетки
- провести выделение речи целевого диктора из принятой акустической смеси
- рассмотреть различные виды адаптивных формирователей луча
- провести сравнительный анализ между адаптивными алгоритмами
- провести моделирование изменения положения источника шума во времени
- изучить методов оценки качества речи

1. Алгоритмы многоканальной обработки речевых сигналов

Речевой сигнал является одним из основных способов передачи информации между людьми. Приём речевых сигналов необходим для дальнейшего хранения и передачи информации. В случае, когда спектры полезного сигнала и шума перекрываются, обычный микрофон не может осуществить фильтрацию без потери качества речи полезного сигнала. Для решения таких задач применяются микрофонные решетки, способные отделить полезный сигнал от шума, в случае разного расположения источников шума и речи. Однако речевой сигнал имеет ряд особенностей, которые необходимо учитывать во время моделирования приема акустического сигнала с помощью микрофонной решетки.

1.1. Моделирование приёма акустического сигнала на плоскую микрофонную решётку

Речевой сигнал является широкополосным, так как ширина его спектра сравнима с центральной частотой. Основная энергия сигнала заключена в диапазоне частот от 20 Гц до 7 кГц. При создании МР необходимо учитывать частотный диапазон звукового сигнала $[f_{min}, f_{max}]$. Для дальнейшего моделирования используется прямоугольная эквидистантная микрофонная решетка. В ней элементы МР расположены на равном расстоянии между собой. Схематически это представлено на рис. 1.1.

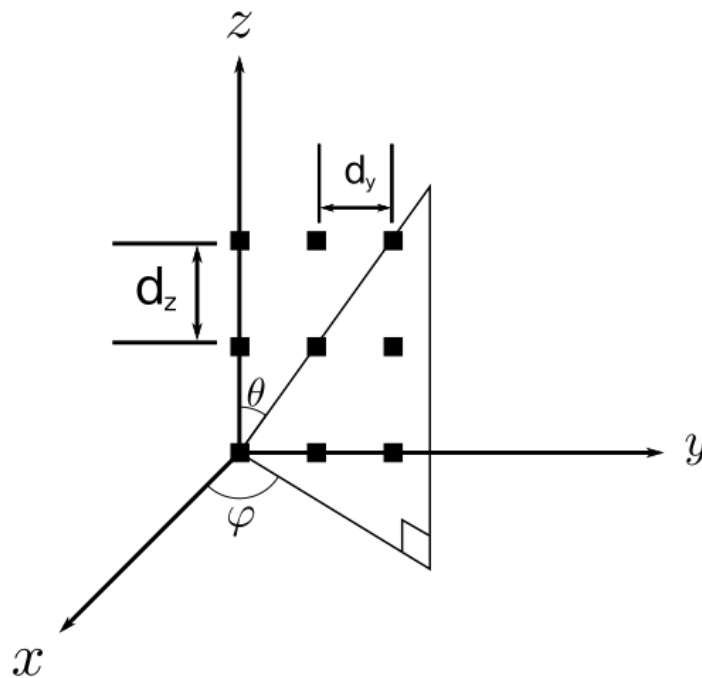


Рис. 1.1. Прямоугольная эквидистантная микрофонная решетка.

Угол ϕ – угол азимута, θ – угол подъёма. Расстояния между микрофонами d_z и d_y должны удовлетворять следующим условиям: $d_z \leq \frac{\lambda_{max}}{2}$ и $d_y \leq \frac{\lambda_{max}}{2}$, где

$\lambda_{max} = \frac{v}{f_{min}}$ – длина волны для f_{min} , v – скорость распространения акустической

волны в атмосфере. Частотный диапазон определяет количество микрофонов:

$M \geq \frac{f_{max}}{f_{min}}$ [3]. Такие соотношения используются для исключения наложения

спектра при дискретизации волны. Однако на разных частотах характерны разные соотношения между длиной волны и фиксированным расстоянием между микрофонами. Использование прямоугольной равномерной микрофонной решетки позволяет осуществить фильтрацию не только по углу азимута, но и углу подъёма.

Микрофонная решетка дискретизирует распространяющуюся в пространстве волну и представляет её в виде последовательности отсчётов. На элементы микрофонной решетки речевой сигнал приходит с различной

задержкой. Для вычисления задержки сигнала на элементе МР вводят матрицу векторов местоположения элементов МР:

$$\mathbf{p} = [\mathbf{p}_1 \ \mathbf{p}_2 \ \dots \ \mathbf{p}_N], \quad (1.1.1)$$

где $\mathbf{p}_i = [x_i \ y_i \ z_i]^T$ – положение i элемента МР, заданное в декартовой системе координат [4].

Так же вводится единичный вектор, характеризующий направление прихода речевого сигнала:

$$\mathbf{a} = \begin{bmatrix} -\cos \theta \cos \phi \\ -\cos \theta \sin \phi \\ -\sin \theta \end{bmatrix}. \quad (1.1.2)$$

Задержка на каждом канале микрофона вычисляется по следующей формуле:

$$\tau = \frac{\mathbf{a}^T \mathbf{p}}{c}, \quad (1.1.3)$$

где c – скорость распространения акустической волны в атмосфере [4].

Частота дискретизации МР – частота взятия отсчётов непрерывного по времени сигнала при его дискретизации. Данная величина измеряется в герцах. Период дискретизации связан с частотой дискретизации следующим соотношением: $T_{\text{диск}} = 1/f_{\text{диск}}$, где $T_{\text{диск}}$ – период дискретизации, $f_{\text{диск}}$ – частота дискретизации. Зная вышеописанные величины, можно получить задержку в количестве отсчетов: $n = \tau/T_{\text{диск}}$. Если данная величина будет целой, то возможно реализовать задержку, используя линии задержки, но в противном случае добиться точного результата не возможно. Для решения этой проблемы используется теорема о сдвиге для дискретного преобразования

Фурье: если последовательность сдвинута на определенное количество отсчетов, то дискретное преобразование Фурье (ДПФ) этой последовательности можно найти из исходной последовательности по следующей формуле:

$$X_{shifted}(m) = e^{\frac{-j2\pi}{N}nm} X(m), \quad (1.1.4)$$

где $X_{shifted}(m)$ – ДПФ сдвинутой последовательности, $X(m)$ – ДПФ исходной последовательности, n – количество задержанных отсчетов, m – номер отсчёта, N – количество отсчётов [5, 6]. Данная формула работает для гармонических сигналов, но не позволяет производить задержку для речевых сигналов. Нулевой элемент отсчётов ДПФ отвечает за постоянную составляющую сигнала. Первая половина отсчётов ДПФ последовательности от 2 до $N/2+1$ называется первой зоной Найквиста, остальные члены последовательности называются второй зоной Найквиста. Для создания задержки широкополосного сигнала мы делили первую зону Найквиста на n равных промежутков. Количество отсчётов в каждом промежутке m можно найти по формуле:

$$m = \frac{N}{2n}. \quad (1.1.5)$$

После каждому промежутку находим соответствующее центральное значение и умножаем каждый промежуток ДПФ последовательности на экспоненту со своим центральным значением. Максимальное число диапазонов находится по формуле $n_{max} = \frac{N}{2}$. Это следует из того, что минимальное количество элементов в диапазоне может быть равно единице, подставляя данное значение в формулу (1.1.5) получаем максимальное число диапазонов.

1.2. Формирователи луча

Сбором речевого сигнала занимается МР, но последующей обработкой для выделения полезного сигнала занимается процессор – формирователь луча (ФЛ). Принцип данного метода состоит в том, чтобы сформировать луч и направить его в нужном направлении. В результате сигналы с углом падения на микрофонную решетку, равным углу направления луча усиливаются, а сигналы от других направлений ослабляются. ФЛ выполняет пространственную фильтрацию для разделения сигналов, которые имеют перекрывающиеся частотные спектры, но исходя из разных пространственных направлений. Формирователь луча выполняет пространственную фильтрацию для разделения сигналов. Главным требованием при реализации любого формирователя луча является синхронность получения данных со всех элементов в один момент времени. Все последующие выкладки основываются на фазовых соотношениях между сигналами с разных каналов. Нарушение условия синхронности приведёт к некорректному результату. На рисунке 1.2 показана простая структура формирования луча на основе прямоугольной равномерной микрофонной решетки, где M датчиков пространственно дискретизируют волновое поле, а выходной сигнал M в момент времени t определяется мгновенной линейной комбинацией пространственных выборок $x_m(t)$, $m=0,1,\dots,M-1$ как:

$$y(t) = \sum_{m=0}^{M-1} x_m(t) w_m^*, \quad (1.2.1)$$

где $*$ обозначает комплексное сопряжение, w_m – весовые коэффициенты.

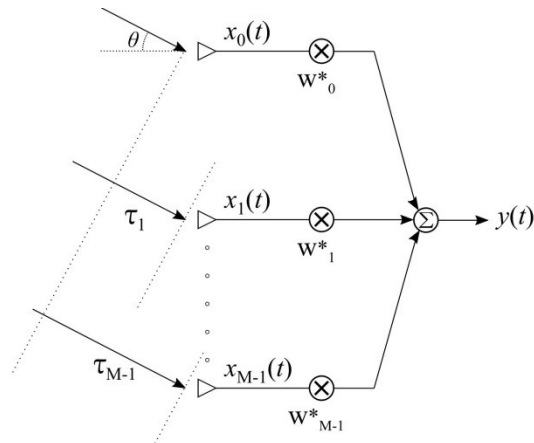


Рис. 1.2. Узкополосный формирователь луча.

Формирователь луча с такой структурой, полезен только для узкополосных сигналов, где «узкополосный» означает, что ширина спектра падающего сигнала достаточно узкая, чтобы гарантировать, что сигналы, принимаемые противоположными концами решетки, всё ещё были коррелированы друг с другом. При фиксированных значениях весовых коэффициентов осуществляется фильтрация с зафиксированного направления. Для широкополосного сигнала процедура формирования луча становится более сложной. Поскольку каждый широкополосный сигнал состоит из бесконечного числа различных частотных компонентов, то значения весовых коэффициентов должны быть различными для разных частот. В данном случае вектор весовых коэффициентов можно записать в следующем виде:

$$\mathbf{w} = [w_0(\omega) w_1(\omega) \dots w_{M-1}(\omega)]^T. \quad (1.2.2)$$

Поэтому узкополосный формирователь луча с постоянным набором коэффициентов для каждого принятого сигнала не будет эффективно работать с широкополосными сигналами. Существует два типа широкополосных формирователя луча: частотный, который разбивает спектр принятых сигналов на диапазоны и умножает каждый диапазон на соответствующие коэффициенты, и временной, который обеспечивает частотно зависимые

коэффициенты с помощью линий задержки КИХ фильтров, расположенных в канале каждого сенсора. Структурная схема временного широкополосного формирователя луча представлена на рисунке 1.3.

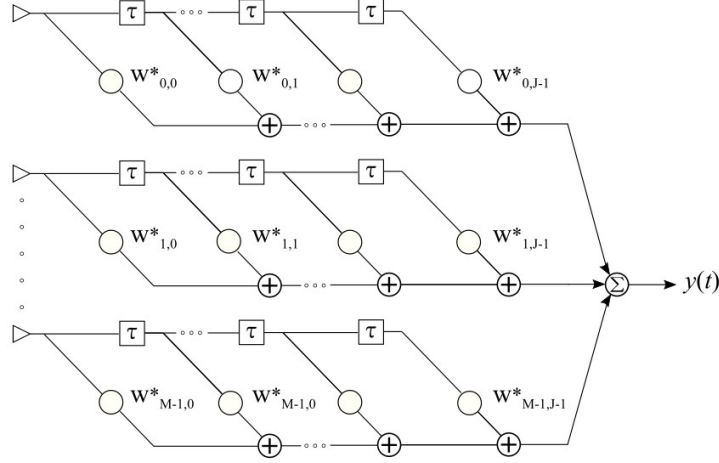


Рис. 1.3. Широкополосный временной формирователь луча.

Здесь J – порядок КИХ фильтров. Далее в работе будет рассматриваться именно широкополосный формирователь луча. Его выходной сигнал описывается следующим выражением:

$$y(t) = \sum_{m=0}^{M-1} \sum_{i=0}^{J-1} x_m(t - iT_s) * w_{m,i}^*, \quad (1.2.3)$$

где T_s – период дискретизации.

Для определения отклика данного формирователя луча вводятся ряд переменных. Сперва вводится волновой вектор, который основывается на формуле (1.1.2):

$$k = \frac{2\pi}{\lambda} a. \quad (1.2.4)$$

Фазирующий вектор имеет следующий вид:

$$\mathbf{v}_k(\mathbf{k}) = \begin{bmatrix} e^{-j\mathbf{k}^T \mathbf{p}_1} \\ e^{-j\mathbf{k}^T \mathbf{p}_2} \\ \dots \\ e^{-j\mathbf{k}^T \mathbf{p}_N} \end{bmatrix}, \quad (1.2.5)$$

Отклик данного формирователя луча определяется формулой:

$$P(\omega, \phi, \theta) = \mathbf{W}^T * \mathbf{v}_k(\mathbf{k}), \quad (1.2.6)$$

где \mathbf{W} – Фурье образ коэффициентов (1.2.2).

Зависимость амплитудного отклика $|P(\omega, \phi, \theta)|$ от угла азимута ϕ и угла подъёма θ называется диаграммой направленности (ДН) формирователя луча. Она характеризует чувствительность формирователя луча по отношению к сигналам, поступающим с разных направлений и с разными частотами.

Диаграмма направленности в дБ:

$$BP = 20 \log_{10} \left(\frac{|P(\omega, \phi, \theta)|}{\max |P(\omega, \phi, \theta)|} \right). \quad (1.2.7)$$

Альтернативный способ построения отклика формирователя луча не использует преобразование Фурье над весовыми коэффициентами. Для данного способа изменяется вид фазирующего вектора. Он имеет следующий вид:

$$\mathbf{d}(\phi, \theta, \omega) = [e^{-j\omega \tau} \ e^{-j\omega(\tau+T_s)} \ \dots \ e^{-j\omega(\tau+(J-1)T_s)}]. \quad (1.2.8)$$

Используя формулы (1.2.8) и (1.2.2) отклик формирователя луча можно записать следующим образом:

$$P(\omega, \phi, \theta) = \mathbf{w}^T * \mathbf{d}(\phi, \theta, \omega). \quad (1.2.9)$$

1.3. Пространственная фильтрация с применением плоской микрофонной решетки

Использование формирователя луча с постоянными весовыми коэффициентами неэффективно в случае динамической шумовой обстановки. Настроенный формирователь луча на определённое направление не позволяет качественно отфильтровать шум, пришедший с другого направления. Для таких целей используют адаптивные формирователи луча, которые определяют статистику входных данных и адаптируют весовые коэффициенты к оптимальному, с точки зрения определенных критериев, решению [7 – 9]. Структура адаптивных формирователей лучей представлена на рисунке 1.4.

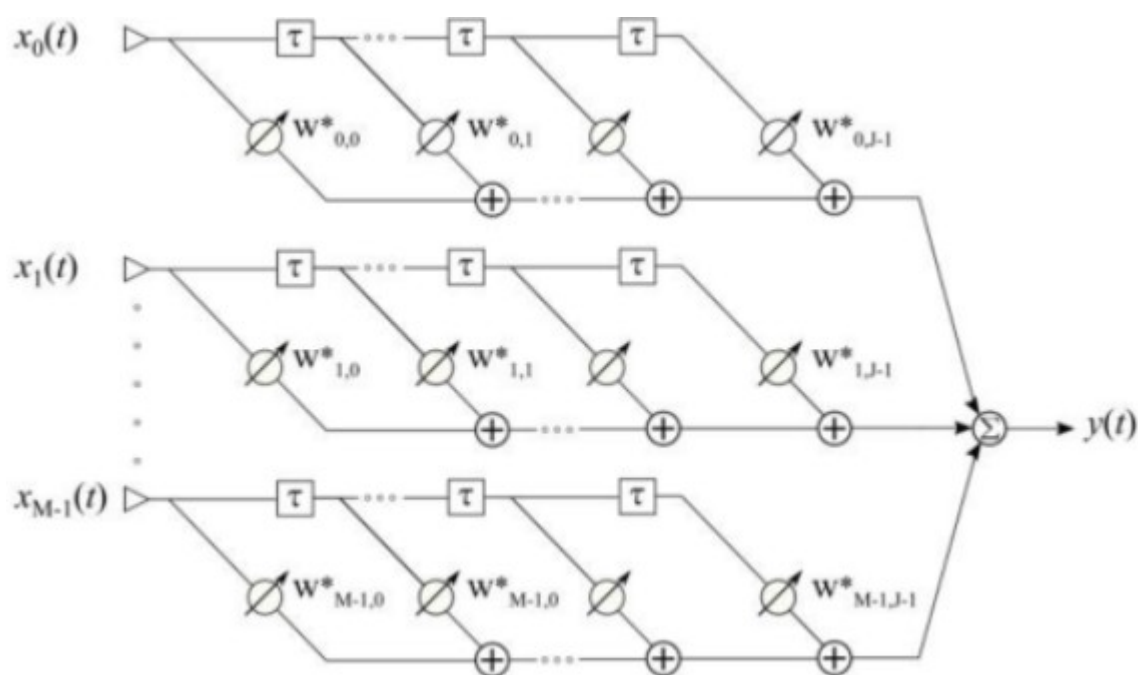


Рис. 1.4. Адаптивный формирователь луча.

Здесь, M – число элементов МР, J – порядок КИХ фильтра. Существует большое количество адаптивных алгоритмов различающиеся по вычислительной сложности, скорости сходимости, переходными процессами [10]. Широкополостность звукового сигнала, а так же многоканальность микрофонной решетки приводит к тому, что для построения адаптивного

фильтра необходимо осуществить многомерную оптимизацию весовых коэффициентов.

Алгоритмы оптимизирующие параметры, представленные на рисунке 1.4, характеризуются методами поиска оптимального решения и природой сигналов ошибок. Для дальнейшего описания алгоритмов, введём общие понятия. Вектор сигналов для каждой итерации имеет следующий вид:

$$\mathbf{X}(k) = [X_0(k) X_1(k) \dots X_{M-1}(k)], \quad (1.3.1)$$

где $X_i(k) = [x_i(k) \ x_i(k-1) \ \dots \ x_i(k-J+1)]$ – вектор сигнала канала.

Матрица весовых коэффициентов КИХ фильтров:

$$\mathbf{W}(k-1) = [\mathbf{W}_0(k-1) \ \mathbf{W}_1(k-1) \ \dots \ \mathbf{W}_{M-1}(k-1)], \quad (1.3.2)$$

где $\mathbf{W}_i(k-1) = [w_{i,0}(k-1) \ w_{i,1}(k-1) \ \dots \ w_{i,J-1}(k-1)]$ – вектор весовых коэффициентов i -ого КИХ фильтра. Предметом оптимизации всех последующих алгоритмов является матрица весовых коэффициентов.

Выходной сигнал с учётом формул (1.3.1) и (1.3.2) можно записать следующим образом:

$$y(k) = \mathbf{W}(k-1)^H \mathbf{X}(k), \quad (1.3.3)$$

где H – знак операции эрмитова сопряжения вектора [10].

В дальнейшем будет рассмотрены два адаптивных алгоритма LC NLMS и LC RLS.

LMS – алгоритм адаптивной фильтрации принадлежащий к классу алгоритмов стохастического градиентного поиска. Данный алгоритм использует упрощенный градиент, что приводит к случайным флуктуациям [10]. Целевой функцией, которую минимизирует данный алгоритм, является

функция среднеквадратической ошибки (MSE) вид которой представлен ниже:

$$F = f\{e(k)\} = E\{|e(k)|^2\}, \quad (1.3.4)$$

где $e(k) = d(k) - y(k)$ – сигнал ошибки, $d(k)$ – требуемый сигнал, $y(k)$ – выходной сигнал адаптивного фильтра. В рамках нашей задачи $d(k) = 0$, так как требуемый сигнал должен обладать минимальной мощностью, то есть стремиться к нулевому. И только линейные ограничения позволяют сохранить сигнал с нужного нам направления [10]. Использование линейных ограничений позволяет принимать сигнал с определённого направления и подавлять все сигналы с направлений, отличающихся от фиксированного. Для описания алгоритма необходимо ввести ряд обозначений. Вектор значений ограничиваемого параметра: f_J . Матрица ограничений C имеет вид:

$$C_{M \cdot J, J} = \begin{bmatrix} c_J & \dots & \mathbf{0} \\ \mathbf{0} & & c_J \end{bmatrix}, \quad (1.3.5)$$

где $c_J = [1 \dots 1]^T$.

Алгоритм LC NLMS в ходе каждой итерации вычисляет выходной сигнал и корректирует весовые коэффициенты с учетом новых значений. Перед началом работы алгоритма необходимо провести инициализацию и задать необходимые значения следующим образом:

$$X(0) = \mathbf{0}, \quad Q_{M \cdot J, J} = C_{M \cdot J, J} [C_{M \cdot J, J}^H C_{M \cdot J, J}]^{-1}, \quad W(0) = Q_{M \cdot J, J} f_J,$$

где $\mathbf{0} = [0 \dots 0]$ – нулевой вектор с числом элементов $M \cdot J$.

В ходе каждой итерации, когда на элементах МР появляются новые значения акустического сигнала, производится фильтрация по формуле (1.3.3). Далее необходимо провести корректировку весовых коэффициентов по следующему алгоритму:

$$\begin{aligned} e(k) &= -\mathbf{W}(k-1)^H \mathbf{X}(k) \\ \tilde{e}(k) &= e(k) [\mathbf{X}(k)^H \mathbf{X}(k) + [\mathbf{X}(k)^H \mathbf{Q}_{M \cdot J, J}][\mathbf{C}_{M \cdot J, J}^H \mathbf{X}(k)] + \delta^2]^{-1} \\ \mathbf{W}'(k) &= \mathbf{W}(k-1) + \hat{\mu} \mathbf{X}(k) \tilde{e}^*(k) \\ \mathbf{W}(k) &= \mathbf{W}'(k) + \mathbf{Q}_{M \cdot J, J} [\mathbf{f}_J - \mathbf{C}_{M \cdot J, J}^H \mathbf{W}'(k)], \end{aligned}$$

где k – номер итерации.

LMS – алгоритм является наиболее широко используемым на практике благодаря тому, что его вычислительная сложность наименьшая среди всех известных алгоритмов адаптивной фильтрации. Вычислительная сложность LC NLMS – алгоритма равна $3M \cdot J + 4(M \cdot J)J$ операций умножения, $4M \cdot J + 5(M \cdot J)J$ операций сложения и одной операции деления. Из-за небольшой вычислительной сложности можно использовать данный алгоритм с большим числом элементов МР и порядком фильтра. Недостатками являются - долгие переходные процессы и малая скорость сходимости. Это не позволяет, в случае частого изменения положения в пространстве источника шума, качественно производить фильтрацию.

Алгоритм RLS так же основан на минимизации квадратичной целевой функции. Данный алгоритм использует входные сигналы для поиска оптимального решения, но этот поиск осуществляется таким образом, что оптимальное решение обеспечивается не только для текущего отсчёта, но и для всех предыдущих отсчётов. Целевой функцией является функция наименьших квадратов (Least Squares, LS), которая представлен ниже:

$$F = f\{e(k)\} = \sum_{i=1}^k |e(i)|^2 \quad (1.3.6)$$

LS-фильтрация использует усреднение по времени, в результате чего находятся весовые коэффициенты оптимального фильтра, которые различаются для каждой реализации стохастического процесса. Перед началом работы алгоритма необходимо провести инициализацию и задать необходимые значения следующим образом:

$$\begin{aligned} X(0) &= \mathbf{0}, \quad R_{M \cdot J}^{-1}(0) = \delta^{-2} I_{M \cdot J}, \\ \Gamma_{M \cdot J, J}(0) &= R_{M \cdot J}^{-1}(0) C_{M \cdot J, J}, \\ \Psi_J^{-1}(0) &= [C_{M \cdot J, J}^H \Gamma_{M \cdot J, J}(0)]^{-1}, \\ W(0) &= \Gamma_{M \cdot J, J}(0) \Psi_J^{-1}(0) f_J \end{aligned}$$

После, в ходе каждой итерации, когда на элементах МР появляются новые значения акустического сигнала, производится фильтрация по формуле (1.3.3). Далее необходимо провести корректировку весовых коэффициентов по следующему алгоритму:

$$\begin{aligned} g_{M \cdot J}(k) &= \frac{R_{M \cdot J}^{-1}(k-1) X(k)}{\lambda + X^H(k) R_{M \cdot J}^{-1}(k-1) X(k)} \\ v_J(k) &= C_{M \cdot J, J}^H g_{M \cdot J}(k) \\ \eta_J^H(k) &= X(k) \Gamma_{M \cdot J, J}(k-1) \\ I_J(k) &= \frac{\Psi_J^{-1}(k-1) v_J(k)}{1 - \eta_J^H(k) \Psi_J^{-1}(k-1) v_J(k)} \\ \Psi_J^{-1}(k) &= \lambda [\Psi_J^{-1}(k-1) + I_J(k) \eta_J^H(k) \Psi_J^{-1}(k-1)] \\ \Gamma_{M \cdot J, J}(k) &= \lambda^{-1} [\Gamma_{M \cdot J, J}(k-1) - g_{M \cdot J}(k) \eta_J^H(k)] \\ \alpha(k) &= -W^H(k-1) X(k) \\ W(k) &= W(k-1) + [g_{M \cdot J}(k) - \lambda \Gamma_{M \cdot J, J}(k) I_J(k)] \alpha^*(k), \end{aligned}$$

где k – номер итерации, $\mathbf{I}_{M \times J}$ – единичная квадратная матрица, δ^{-2} – постоянная величина, λ – параметр экспоненциального взвешивания сигналов. Оценка вычислительной сложности LC RLS алгоритма зависит от способа его реализации [10].

1.4. Методы оценки качества речи

Для оценки эффективности алгоритмов необходимы методы сравнения входного и выходного сигнала между собой. Стандартным методом оценки качества фильтрации является вычисление выигрыша отношения сигнал/шум (ОСШ) на выходе адаптивного фильтра. Если представить входной сигнал, поступающий на фильтр, в виде суперпозиции полезного сигнала и шума $f_{\text{вх}}(k) = s_{\text{вх}}(k) + n_{\text{вх}}(k)$, где $s_{\text{вх}}(k)$ – входной полезный сигнал, $n_{\text{вх}}(k)$ – входной шум. Тогда ОСШ на входе считается по следующей формуле:

$$SNR_{\text{вх}} = 10 \log_{10} \frac{P_{\text{вх.сиг}}}{P_{\text{вх.шум}}}, \quad (1.4.1)$$

где $P_{\text{вх.сиг}} = \frac{1}{N} \sum_{k=1}^N s_{\text{вх}}^2(k)$ – средняя мощность полезного сигнала на входе фильтра, $P_{\text{вх.шум}} = \frac{1}{N} \sum_{k=1}^N n_{\text{вх}}^2(k)$ – средняя мощность шума на входе фильтра.

Аналогичным образом можно посчитать ОСШ на выходе фильтра с учётом того, что выходной сигнал можно записать в виде $f_{\text{вых}}(k) = s_{\text{вых}}(k) + n_{\text{вых}}(k)$, где $s_{\text{вых}}(k)$ – выходной полезный сигнал, $n_{\text{вых}}(k)$ – выходной шум. Тогда формула имеет вид:

$$SNR_{\text{вых}} = 10 \log_{10} \frac{P_{\text{вых.сиг}}}{P_{\text{вых.шум}}}, \quad (1.4.2)$$

где $P_{\text{вых.сиг}} = \frac{1}{N} \sum_{k=1}^N s_{\text{вых}}^2(k)$ – средняя мощность полезного сигнала на выходе фильтра, $P_{\text{вых.шум}} = \frac{1}{N} \sum_{k=1}^N n_{\text{вых}}^2(k)$ – средняя мощность шума на выходе фильтра.

Выигрыш ОСШ фильтра считается следующим образом:

$$SNR_{\text{выиг}} = SNR_{\text{вых}} - SNR_{\text{вх}}. \quad (1.3.3)$$

Однако данная величина не всегда может дать адекватное представление о том, насколько речь стала разборчивее на фоне шумов. Существуют ситуации, когда происходит сильное искажение сигнала, хотя значение ОСШ не изменилось. Для исключения таких ситуаций существуют методы оценки качества речи. Различают субъективную и объективную оценку качества. Первая группа методов заключается в оценке искажений сигнала при помощи субъективно-статистических испытаний, где задействуют группы специально обученных людей. Субъективные оценки качества основываются на статистической обработке субъективных оценок качества достаточно большого числа слушателей-экспертов. Эти оценки существенно зависят от возраста и пола диктора, скорости произнесения фраз и других обстоятельств. Тесты при получении субъективных оценок проводят с имитацией реальных условий, например, посторонний шум, фоновая речь других людей и т. п. Количественные результаты этих тестов отображают усредненное качество, уровень усилий слушателя, разборчивость, естественность звучания. Часто используемой методикой субъективной оценки качества речи стала MOS [11]. Однако такой способ оценки был достаточно затратным по многим аспектам и на смену пришли объективные методы оценки качества речи. Основным принцип таких методов заключается в сравнении исходного сигнала с сигналом, прошедшим через определённую систему. К данному способу относится вышеупомянутое вычисление выигрыша ОСШ, но кроме этого есть объективная оценка качества речи Perceptual Evaluation of Speech Quality (PESQ), определённая в рекомендации МСЭ-Т Р.862 [12]. Для определения качества передачи речи в PESQ предусмотрено сравнение входного, или эталонного, сигнала с его искажённой версией на выходе системы связи. Этот процесс показан на рисунке 1.5:

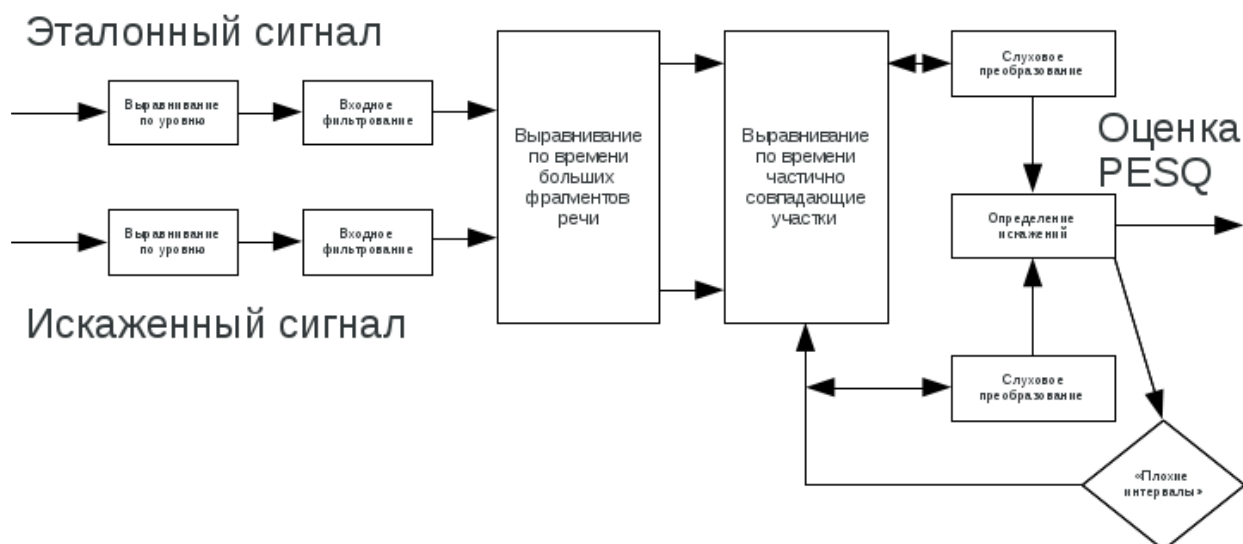


Рис. 1.5. Процесс обработки сигналов в PESQ.

Для вычисления оценки качества речи эталонный и искаженный сигнал выполняется следующий ряд действий. Сперва осуществляется выравнивание мощностей сигналов. После осуществляется входное фильтрование с последующим выравниванием по времени. Далее осуществляется процесс слухового преобразования, которое имитирует определенные особенности человеческого слуха. Это дает информацию о воспринимаемой громкости сигнала в зависимости от времени и частоты, представляемую как поверхность восприятия. Оценка PESQ производится по пятибалльной шкале от 1 до 5. Адаптивные фильтры осуществляют фазовые искажения фильтруемого сигнала, что приводит к тому, что входной сигнал и выходной сигнал нельзя использовать для оценки качества речи при помощи PESQ. Данная проблема решается при помощи отдельной фильтрации исключительно полезного сигнала, а так же отдельно сумму полезного сигнала и шума. Именно эти два сигнала нужно сравнивать между собой для получения корректной оценки. Процесс вычисления оценки PESQ, используемый в работе, представлен на рис. 1.6.

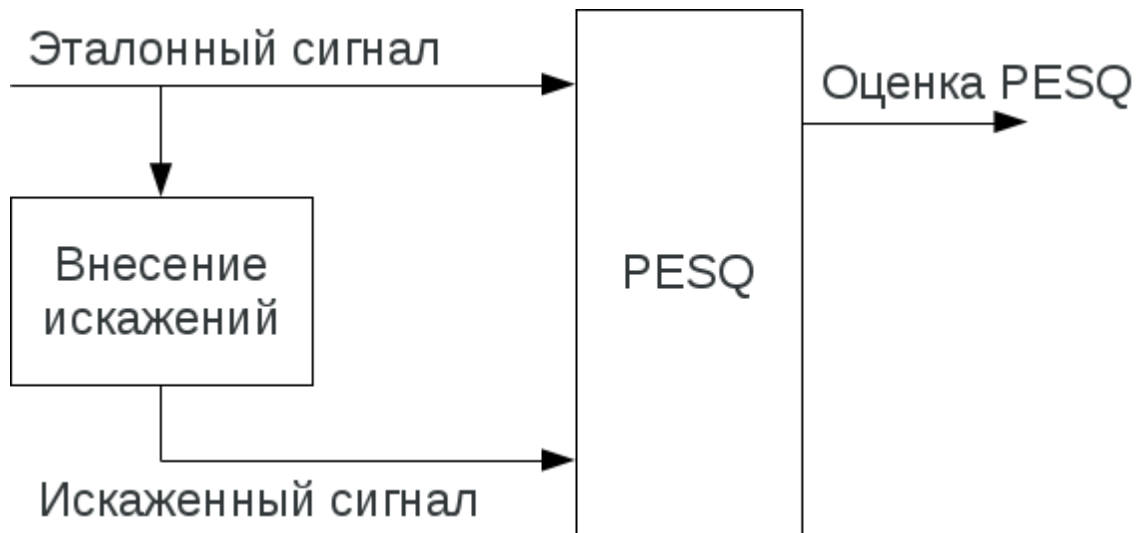


Рис. 1.6. Схема вычисления объективной оценки качества речи PESQ.

Для дальнейшего вычисления эффективности алгоритмов используется как вычисление выигрыша ОСШ, так и объективной оценки качества речи PESQ.

2. Компьютерное моделирование многоканальной системы

Для дальнейшего моделирования микрофонной решетки, необходимо задать её параметры. Моделирование приёма сигнала на МР требует задать параметры прихода сигнала и его характеристики.

2.1. Постановка задачи приёма и фильтрации сигнала с помехой

Прямоугольная эквидистантная микрофонная решетка расположена в плоскости ZoY , как представлено на рисунке 2.1. Число микрофонов $M=9$, что соответствует МР 3×3 .

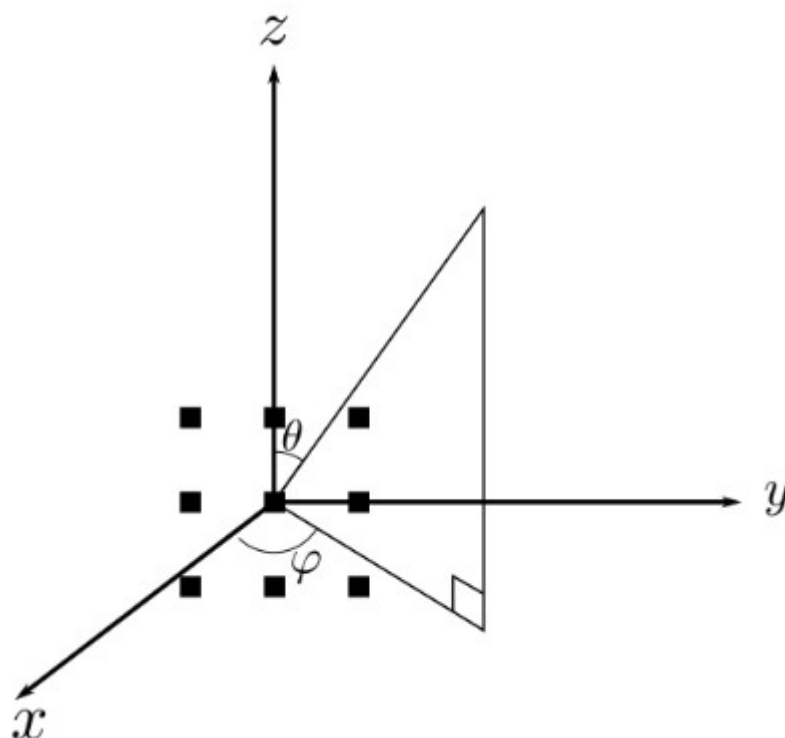


Рис. 2.1. Расположение элементов микрофонной решетки в пространстве. Полезный сигнал приходит с направления $\phi = 0$ град и $\theta = 0$ град. Моделирование приёма шума будет происходить с направления $\phi = 50$ град и $\theta = 0$ град. Порядок фильтров имеет следующее значение: $J=32$. Расстояние между микрофонами $d_z=d_y=4$ см. Основываясь на формулах (1.1.3), (1.1.4) произведён приём сигнала с заданными условиями. При этом полезный сигнал принят на каждый элемент МР без задержки из-за

нулевого направления прибытия этого сигнала. Был реализован алгоритм LC NLMS, при этом использовались линейные ограничения следующего вида: $\mathbf{f}_j = [1 \ 0 \ \dots \ 0]$. Данный набор весовых коэффициентов обеспечивает пропускание сигнала с направления $\phi = 0$ град и $\theta = 0$ град, потому что сигнал приходит без задержки и не требует никакой обработки. На вход адаптивного формирователя луча подаётся принятый сигнал. Результаты фильтрации и входной сигнал представлены на рис. 2.3 и рис. 2.2 соответственно.

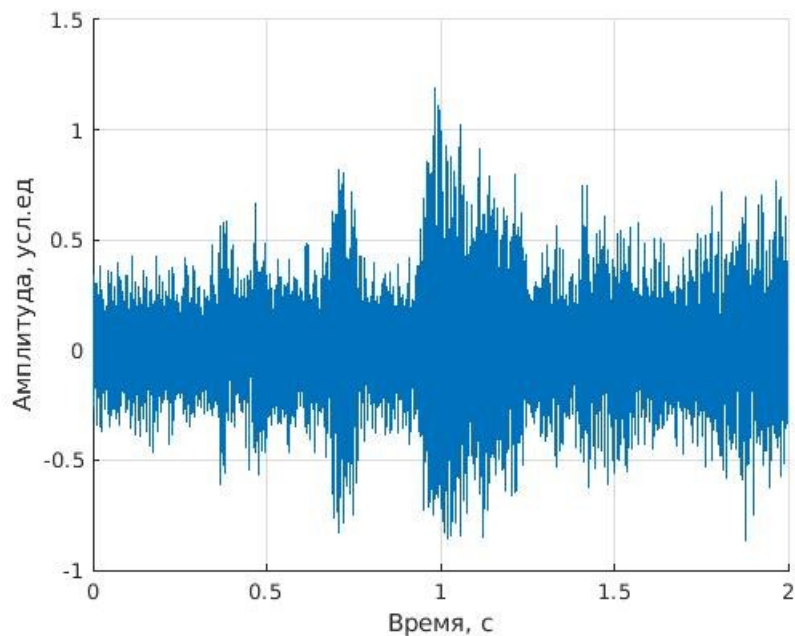


Рис. 2.2. Входной сигнал на одном из каналов.

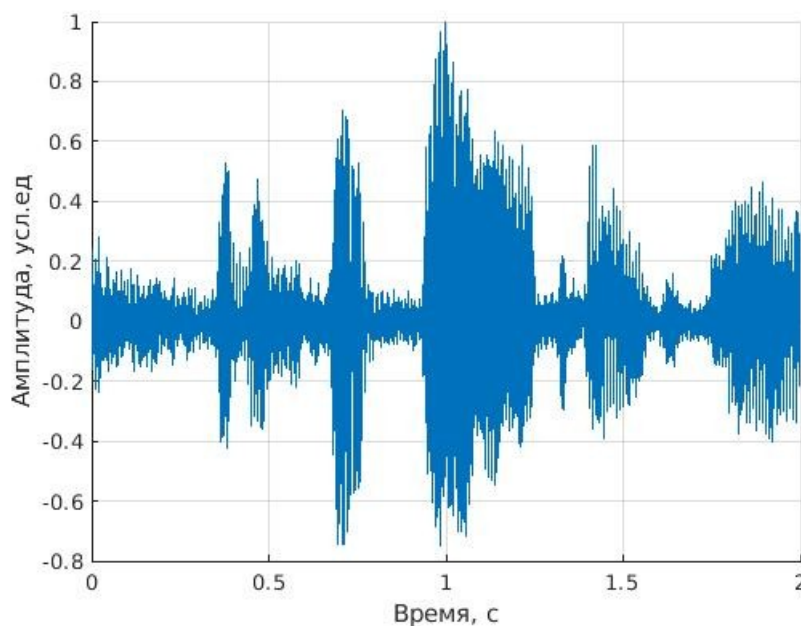


Рис. 2.3. Сигнал на выходе адаптивного формирователя луча LC NLMS.

Из рисунка 2.3 видно, что в начале работы алгоритма выходной сигнал содержит большее количество переходных процессов и формирователь луча производит поиск оптимальных весовых коэффициентов. Переходные процессы выражены в увеличенной дисперсии сигнала в самом начале, относительно среднего уровня шума после. В конце работы алгоритма можно получить весовые коэффициенты и на их основе построить диаграмму направленности от частоты. Из формулы (1.2.6) видно, что отклик формирователя луча зависит от трёх переменных, что не позволяет визуализировать диаграмму направленности. Для этого в дальнейшем происходит фиксирование одного из углов и построение ДН от остальных переменных. Построение диаграммы направленности производится с использованием формул (1.2.6) и (1.2.7), результат которого представлен на рисунках 2.4 и 2.5.

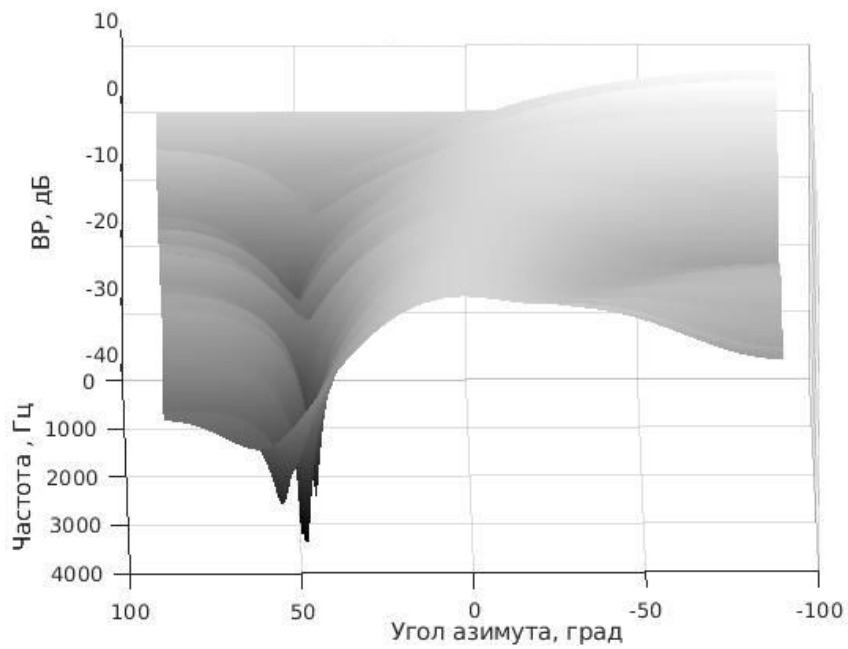


Рис. 2.4. Диаграмма направленности для алгоритма LC NLMS при фиксированном $\theta = 0$ град.

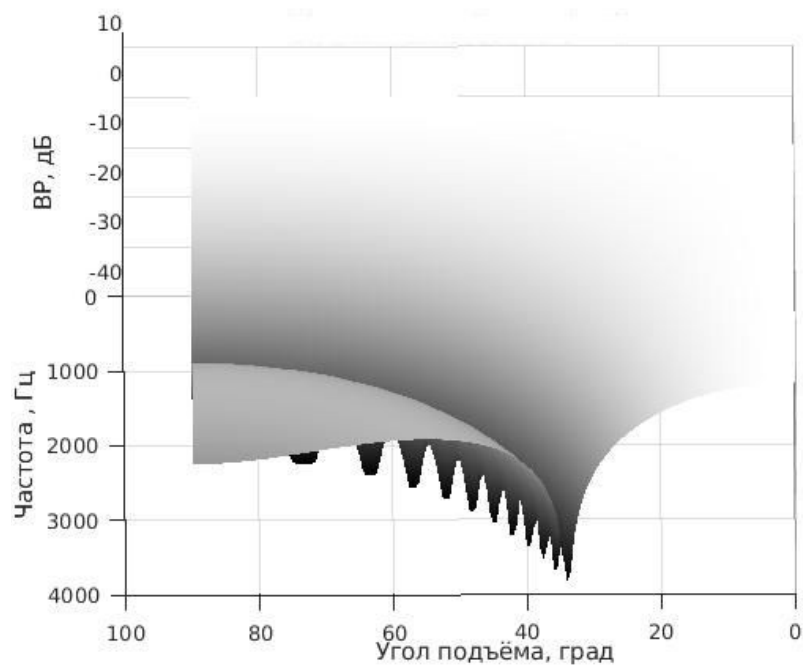


Рис. 2.5. Диаграмма направленности для алгоритма LC NLMS при фиксированном $\phi = 0$ град.

На рисунке 2.4 видно, что происходит ослабление по всем частотам в направлении $\phi = 50$ град, при этом в направлении $\phi = 0$ град наблюдается

усиление сигнала. Данный вывод подтверждает корректность работы адаптивного алгоритма.

Подобным образом был смоделирован адаптивный алгоритм LC RLS, на вход которого поступила смесь полезного сигнала и шума. Результаты моделирования представлены на рис. 2.6, входной же сигнал изображен на рис. 2.2. Основываясь на полученных весовых коэффициентах, была построена диаграмма направленности при тех же фиксированных углах, что и в случае алгоритма LC NLMS. Результаты построения представлены на рис. 2.7 и рис 2.8.

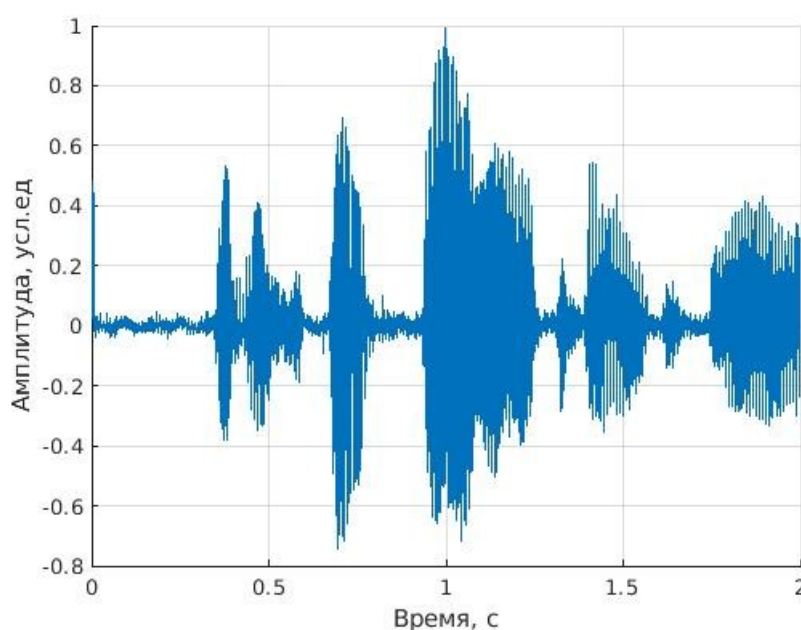


Рис. 2.6. Сигнал на выходе адаптивного формирователя луча LC RLS.

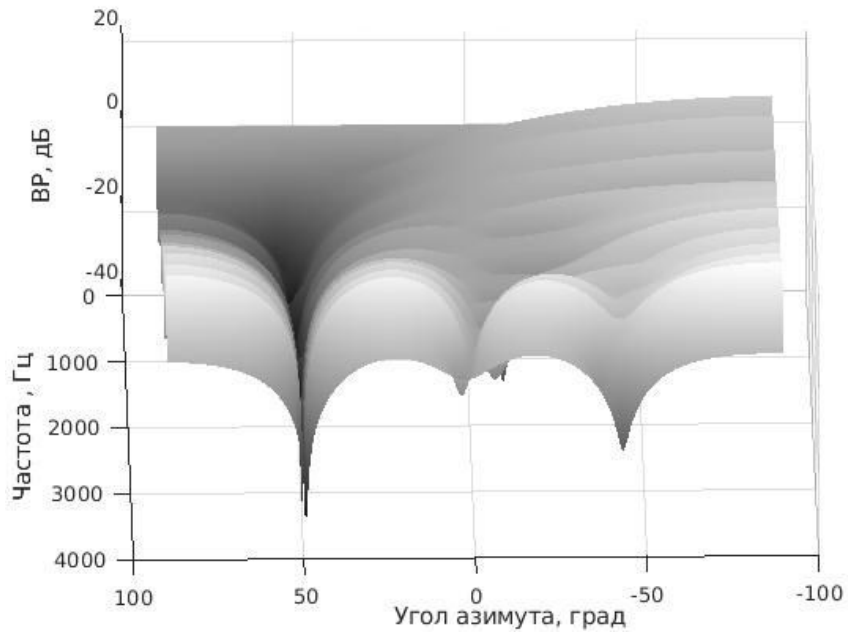


Рис. 2.7. Диаграмма направленности для алгоритма LC RLS при фиксированном $\theta = 0$ град.

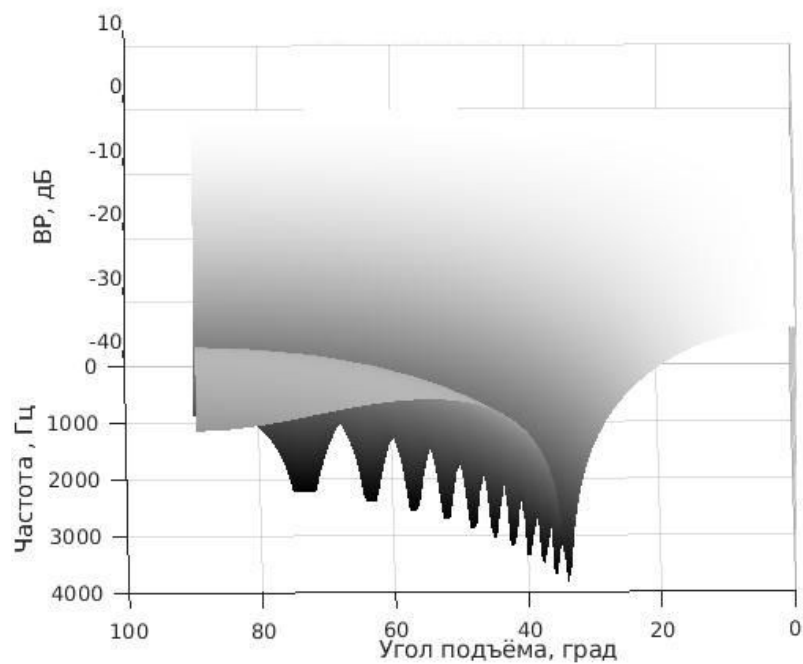


Рис. 2.8. Диаграмма направленности для алгоритма LC RLS при фиксированном $\phi = 0$ град.

Сравнивая рисунки 2.3 и 2.6 видно, что алгоритм LC RLS производит более качественную фильтрацию, при этом переходный процесс занимает

меньше времени. Сравнивая диаграммы направленности видно, что оба алгоритма корректно фильтруют сигнал, оставляя лишь полезный сигнал. Однако в случае алгоритма LC RLS на рис. 2.7 в направлении прихода шума заметны более сильные ослабления в отличие от алгоритма LC NLMS.

2.2. Сравнительный анализ алгоритмов LC NLMS и LC RLS

Проведём сравнительный анализ двух алгоритмов. Рассмотрим эффективность каждого алгоритма и сравним качество фильтрации алгоритмов в зависимости от различных параметров. Критерием оценки будет как значение выигрыша ОСШ, так и субъективная оценка качества речи PESQ. Одним из параметров, который можно менять, является порядок фильтра. Была построена зависимость выигрыша ОСШ от порядка фильтра для двух алгоритмов. Полученная зависимость представлена на рисунке 2.9.

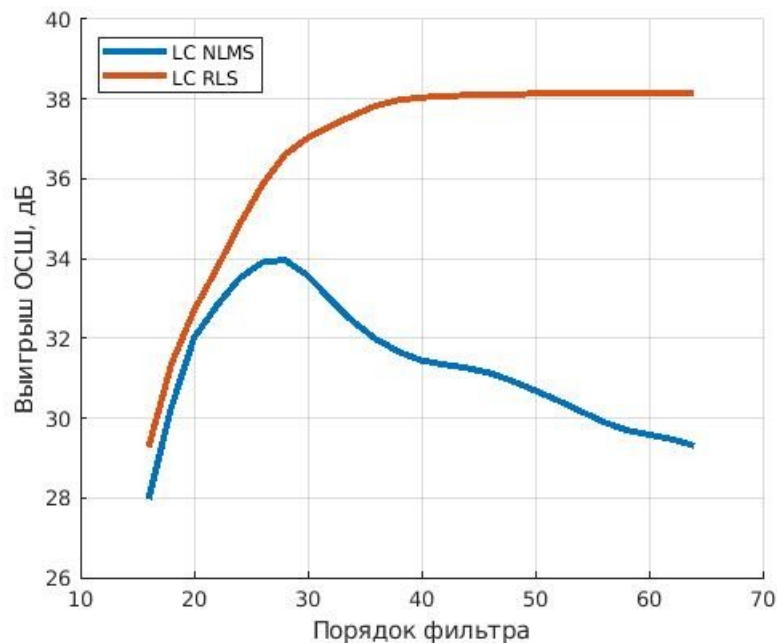


Рис. 2.9. Выигрыш ОСШ от порядка фильтра для двух алгоритмов.

Аналогично была построена зависимость оценки PESQ от порядка фильтров для двух алгоритмов, что представлено на рис. 2.10. Из графиков 2.9 и 2.10 видно, что при превышении порядка фильтров выше определённого значения, не происходит значительного повышения оценки PESQ, что говорит о существовании оптимального порядка фильтра для каждого алгоритма.

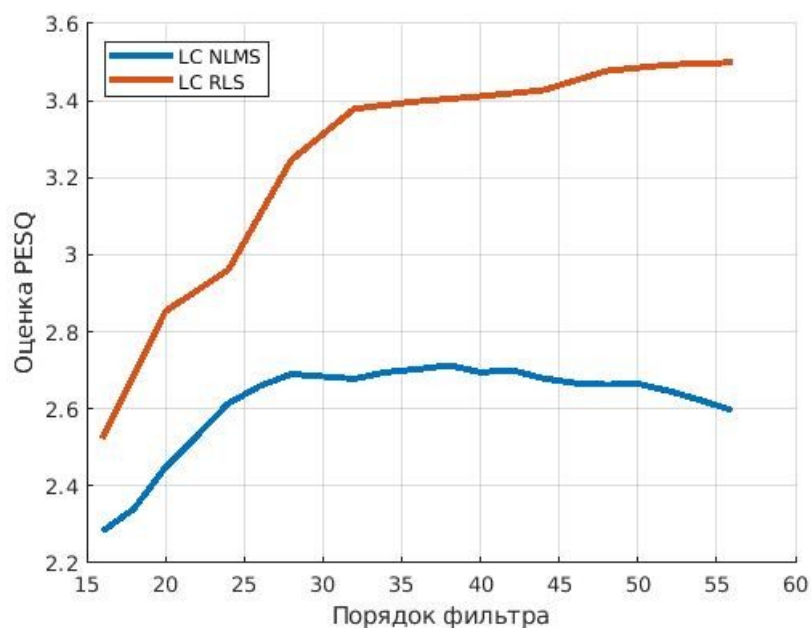
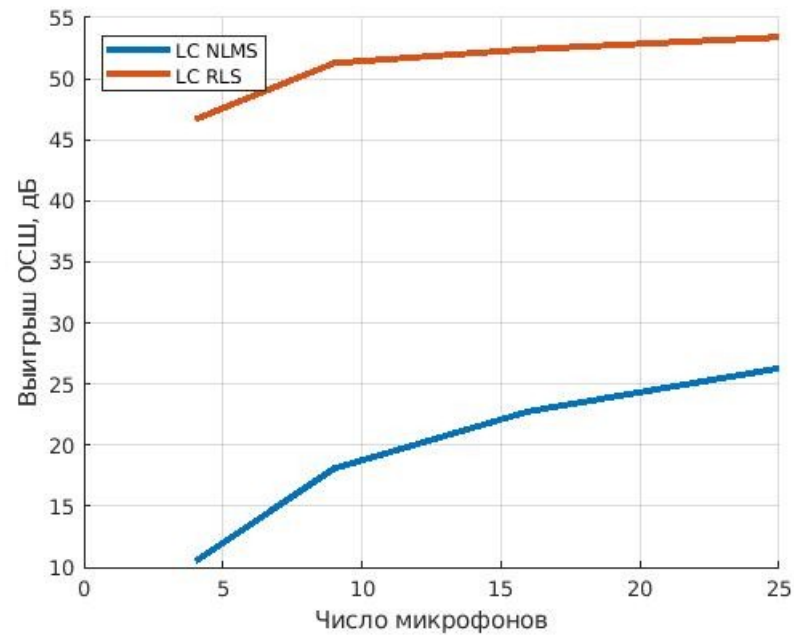


Рис. 2.10. Оценка PESQ от порядка фильтра для двух алгоритмов.

Так же была построена зависимость выигрыша ОСШ и оценки PESQ от числа используемых микрофонов. Поскольку микрофонная решетка прямоугольная, то используются значения числа микрофонов, которые являются квадратом натурального числа, такие как 4, 9, 16, 25. Полученные зависимости представлены на рисунках 2.11 и 2.12. Из рисунка 2.12 видно, что увеличения числа микрофонов в каждом случае приводит к повышению оценки PESQ и соответственно к повышению качества фильтрации.

Построен график зависимости выигрыша ОСШ адаптивных фильтров от входного ОСШ для двух алгоритмов, что представлен на рис. 2.13.



1. Рис. 2.11. Выигрыш в ОСШ в зависимости от числа микрофонов.

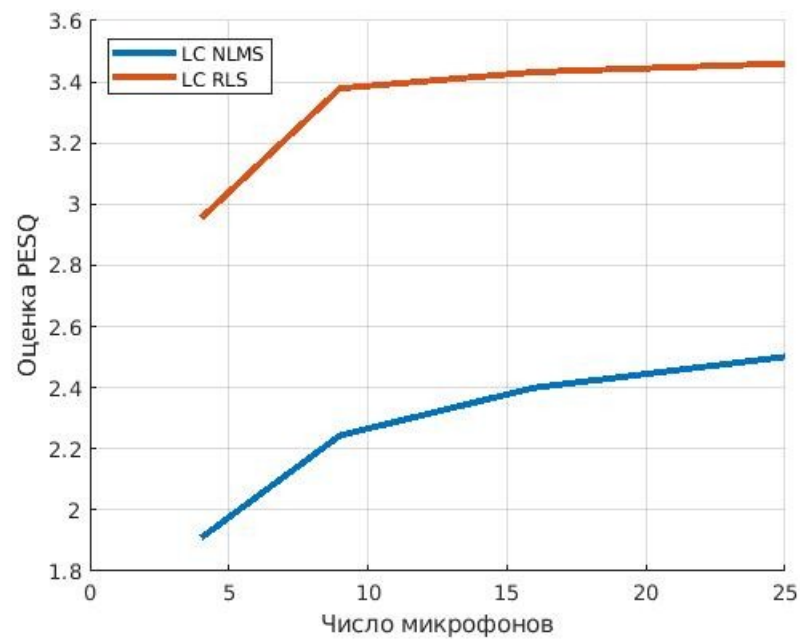


Рис. 2.12. Оценка PESQ в зависимости от числа микрофонов.

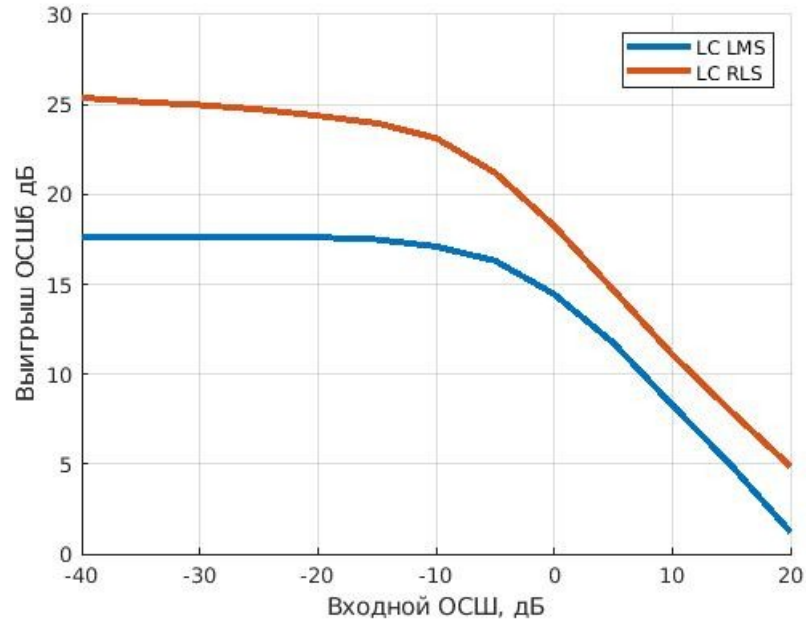


Рис. 2.13. График зависимости выигрыша ОСШ от входного ОСШ.

Из графика 2.13 видно, что при увеличении входного ОСШ, выигрыш ОСШ становится меньше. Это связано с внешними шумами, которые поступают вместе со смесью полезного сигнала и акустического шума. Такой шум адаптивные алгоритмы воспринимают как суперпозицию волн пришедших с различных направлений, что не позволяет им отфильтровать подобный шум.

Адаптивные пространственные фильтры осуществляют фильтрацию в случае различных геометрических положений источников сигнала и шума. Проанализируем то, насколько близко можно расположить шум по отношению к полезному сигналу, чтобы алгоритм смог отфильтровать лишь шум. Были построены графики зависимости выигрыша ОСШ от угла азимута при фиксированном угле подъёма $\theta = 0$ град. Данная зависимость представлена на рис. 2.14.

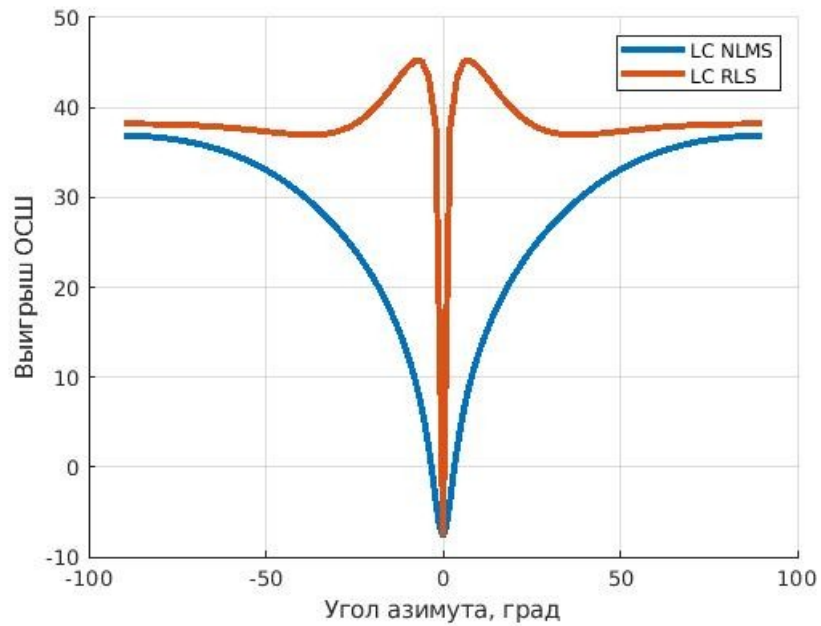


Рис. 2.14. Зависимость выигрыш ОСШ от угла азимута при фиксированном угле подъёма $\theta = 0$ град.

Такая же зависимость построена для оценки PESQ на рис. 2.15.

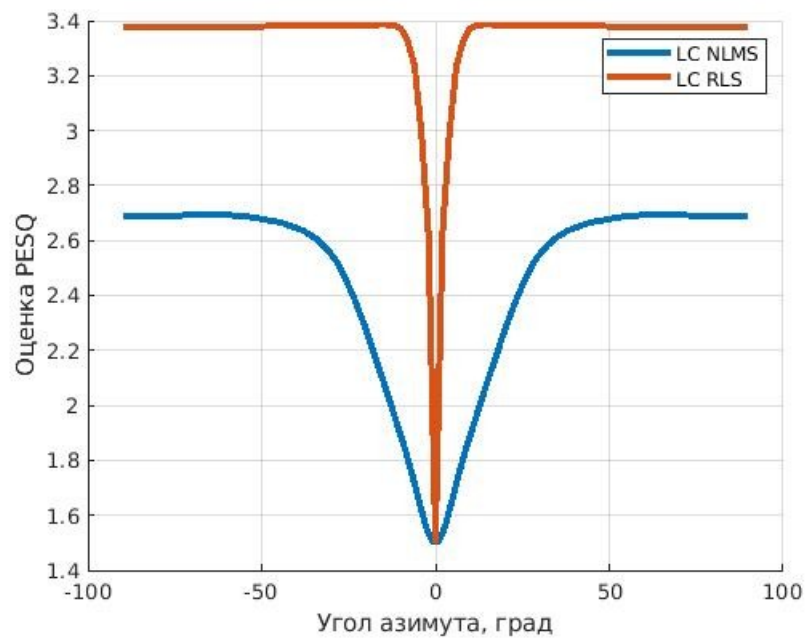


Рис. 2.15. Зависимость оценки PESQ от угла азимута при фиксированном угле подъёма $\theta = 0$ град.

Из рисунков 2.14 и 2.15 видно, что существует некоторый минимальный угол азимута, ниже которого адаптивные алгоритмы не могут отделить сигнал от шума. Особо характерный спад заметен на рис. 2.14. Для алгоритма LC NLMS минимальный угол составляет: $\phi_{\min}=25$ град. Для алгоритма LC RLS минимальный угол составляет: $\phi_{\min}=6$ град. Аналогично были построены графики в случае фиксированного $\phi = 0$ град, что представлено на рис. 2.15 и 2.16.

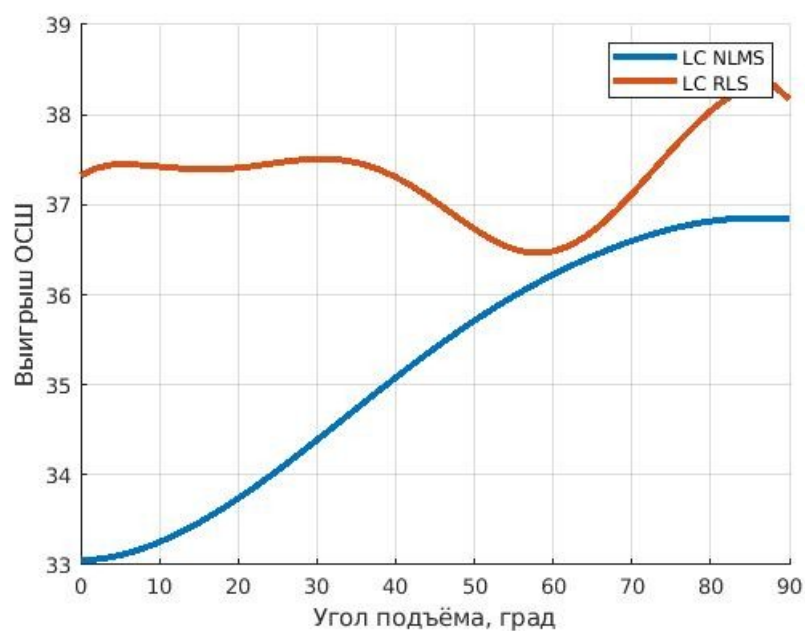


Рис. 2.15. Зависимость выигрыш ОСШ от угла подъема при фиксированном угле азимута $\phi = 0$ град.

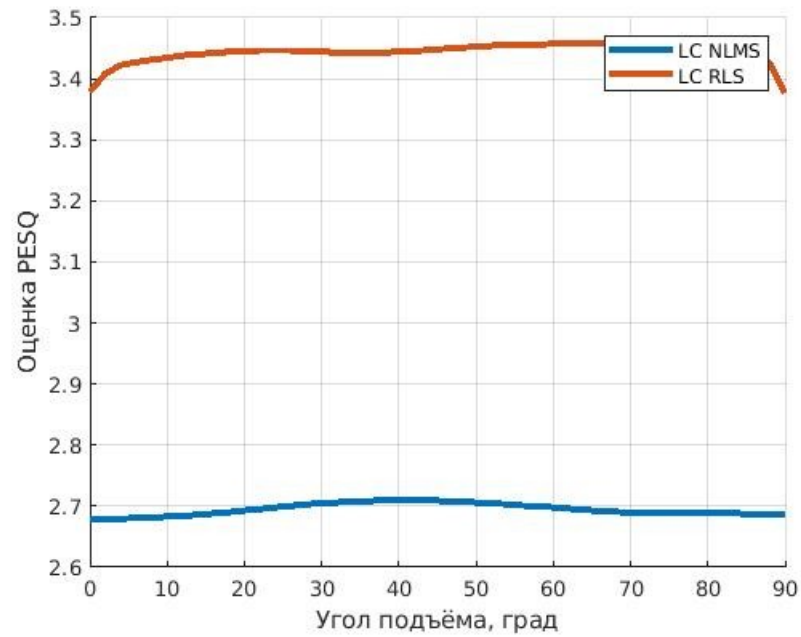


Рис. 2.16. Зависимость оценки PESQ от угла подъёма при фиксированном угле азимута $\phi = 0$ град.

Исходя из рис. 2.16 можно сделать вывод, что в данной случае значительного изменения оценки PESQ от угла подъёма не происходит.

2.3. Эффективность работы микрофонной решетки в быстро меняющейся помеховой обстановке

В случае, когда источник шума находится в одной точке и не меняет своё положение в пространстве, адаптивный алгоритм без труда осуществляет фильтрацию, поскольку его весовые коэффициенты настроены на подавление помех с данного направления. Рассмотрим же ситуацию, когда адаптивный алгоритм производит фильтрацию шума с одного направления и в один момент он перестаёт приходить с этого направления и начинает приходить шум с другого направления. Смоделируем данную ситуацию, когда половину эксперимента шум приходит с направления $\phi = 50$ град и $\theta = 0$ град, а после шум поступает с направления $\phi = -50$ град и $\theta = 0$ град. Входной сигнал и выходной для алгоритмов представлены на рис. 2.17-2.18. Вертикальная линия обозначает момент времени, когда происходит смена направления прихода шума.

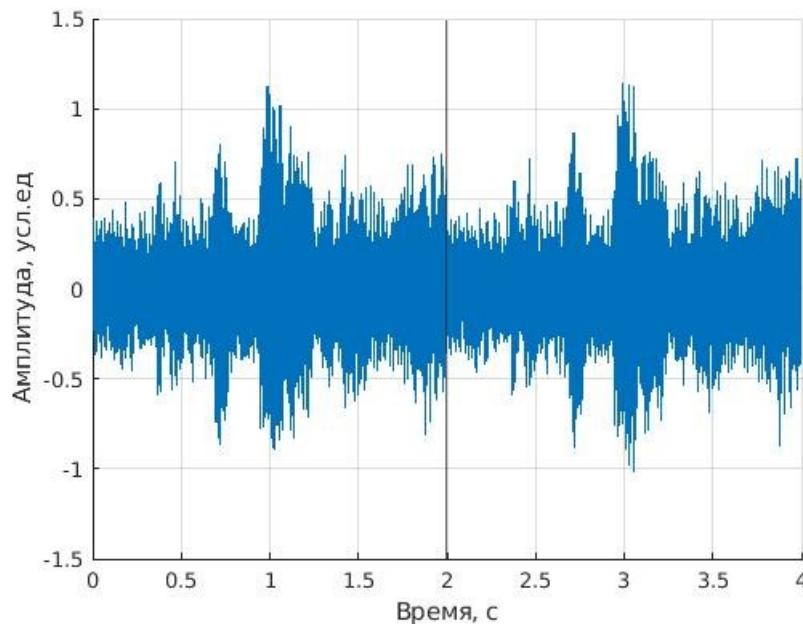


Рис. 2.17. Входной сигнал.

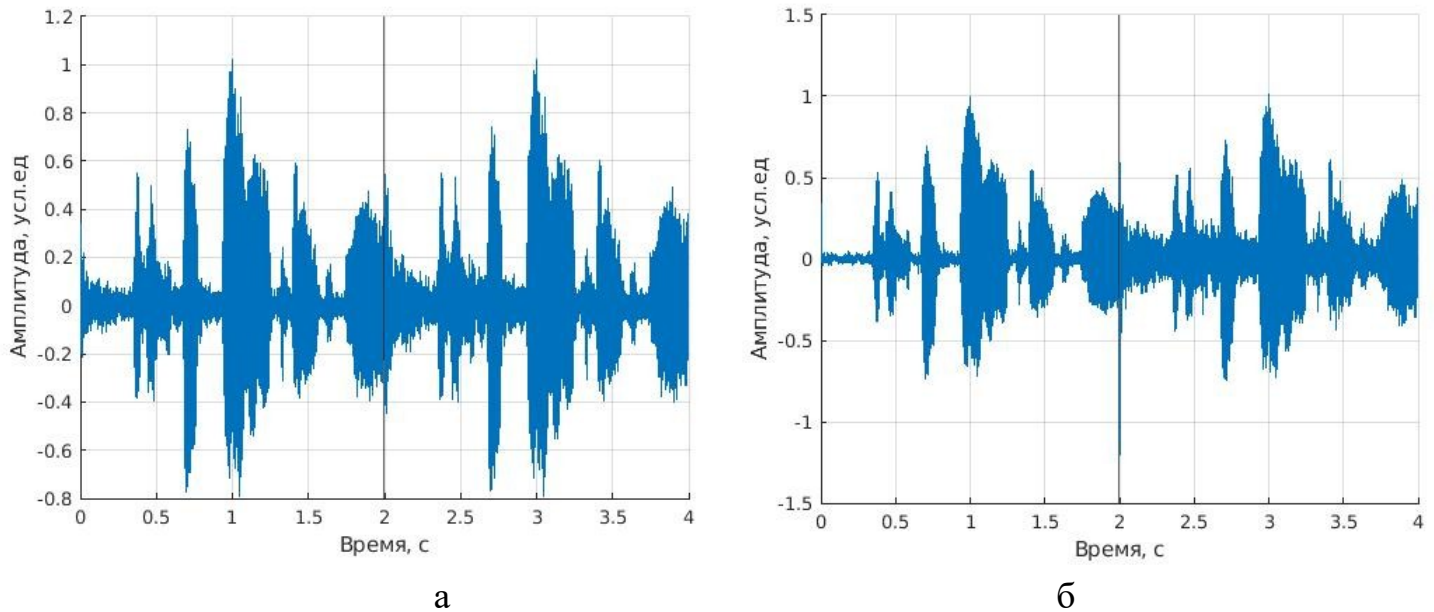


Рис. 2.18. Выходные сигналы для алгоритмов: а) LC NLMS б) LC RLS.

Сравнивая результаты фильтрации рис. 2.18 а и б, видим, что после смены направления прихода сигнала, алгоритмам необходимо определённое время для перестройки весовых коэффициентов. На основе полученных весовых коэффициентов была построена диаграмма направленности для каждого алгоритма при фиксированном $\phi = 0$ град и $\theta = 0$ град. Данные результаты представлены на рис. 2.19-2.22.

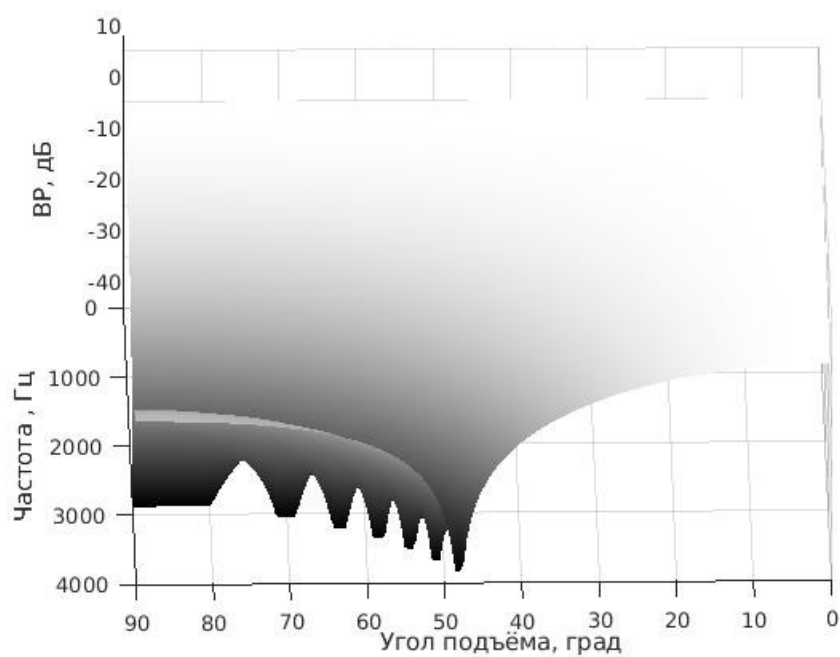


Рис. 2.19. Диаграмма направленности для алгоритма LC NLMS при фиксированном $\phi = 0$ град.

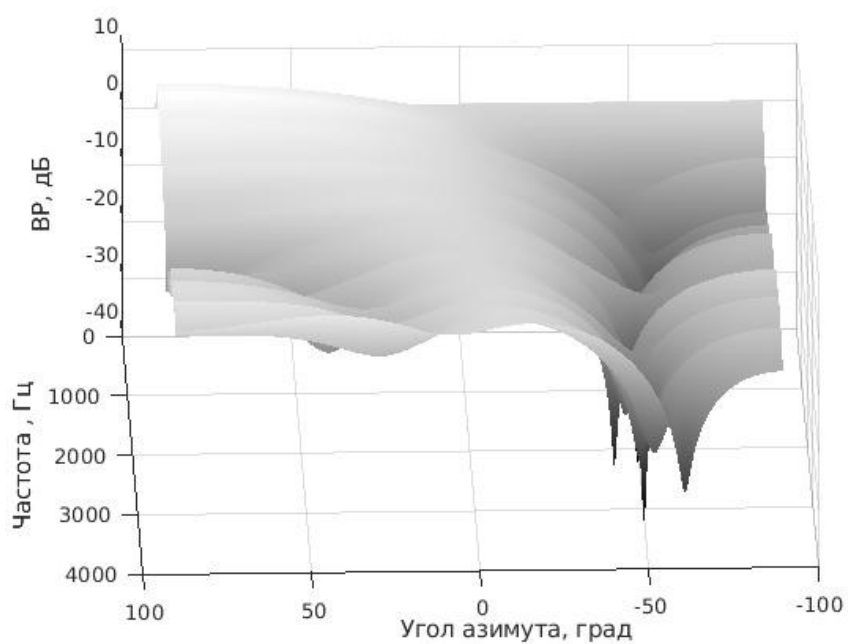


Рис. 2.20. Диаграмма направленности для алгоритма LC NLMS при фиксированном $\theta = 0$ град.

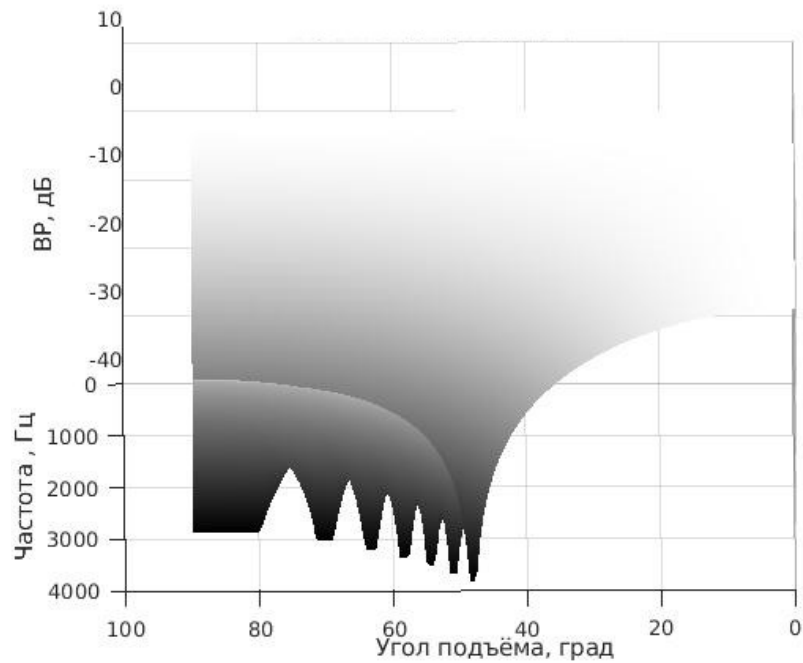


Рис. 2.21. Диаграмма направленности для алгоритма LC RLS при фиксированном $\phi = 0$ град.

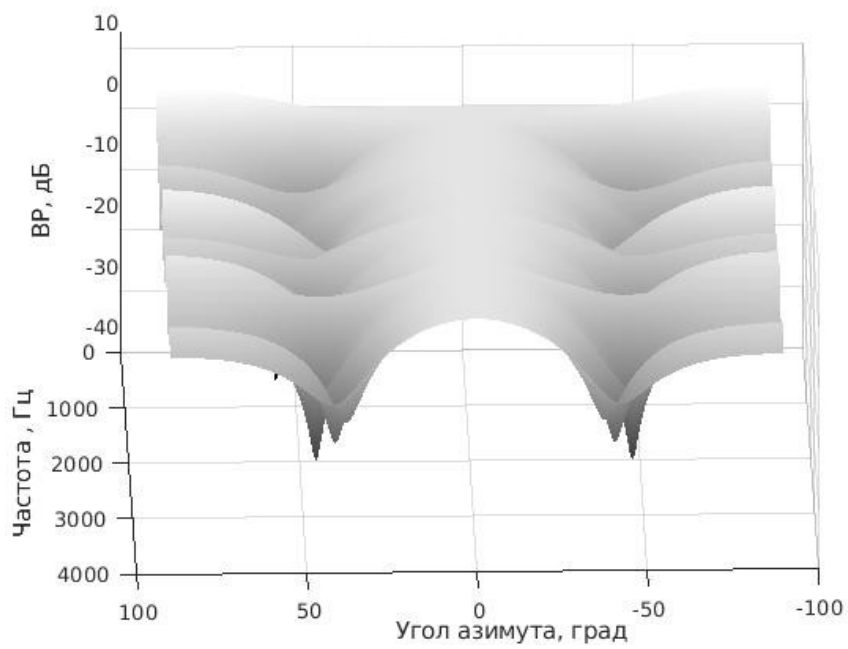


Рис. 2.22. Диаграмма направленности для алгоритма LC RLS при фиксированном $\theta = 0$ град.

Из рис. 2.19-2.20 видно, что алгоритм LC NLMS подстроил свои весовые коэффициенты на подавление помехи с направления $\phi = -50$ град.

Однако же исходя из рис. 2.21-2.22 видно, что алгоритм LC RLS адаптировал весовые коэффициенты на подавление помехи пришедшей с направления $\phi = 50$ град и с направления $\phi = -50$ град. Из-за этого качество фильтрации становится хуже, что заметно на рис. 2.19 после смены направления. Это объясняется тем, что данный алгоритм подстраивает весовые коэффициенты так, что бы они удовлетворяли всем предыдущим значениям. Решением данной проблемы является изменения коэффициента λ при моделировании алгоритма, поскольку он отвечает за „память“ алгоритма.

2.4. Эффективность работы микрофонной решетки при отклонении её элементов

Рассмотрим ситуацию, когда один из элементов МР сместился от своего исходного положения. В таком случае микрофонная решетка перестаёт быть прямоугольной и равномерной. Расположение микрофонов в проекции ZoY представлено на рис. 2.23.

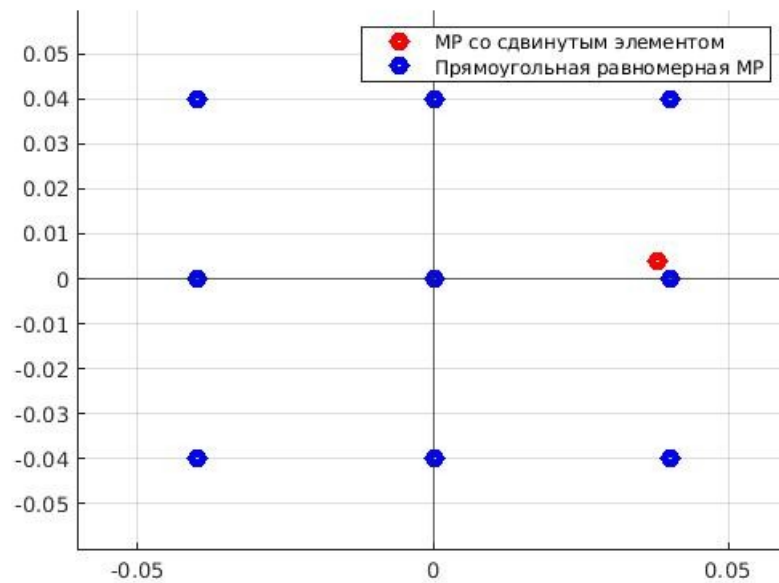


Рис. 2.23. Положение микрофонов в проекции ZoY .

Смоделируем сигнал принятый на такую решетку и подадим его на вход каждого алгоритма. Результаты фильтрации и входной сигнал представлены на рис. 2.24-2.26.

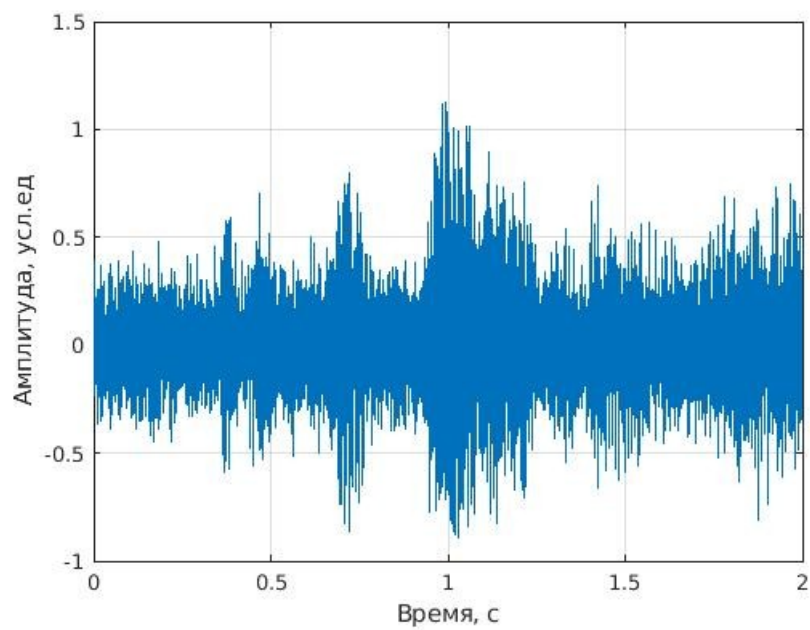


Рис. 2.24. Входной сигнал на одном из элементов.

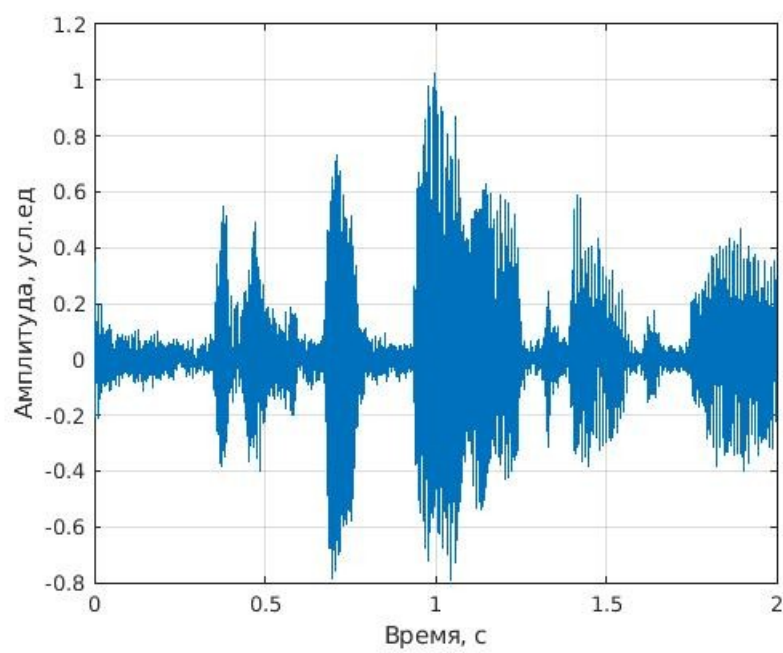


Рис. 2.25. Выходной сигнал адаптивного LC NLMS.

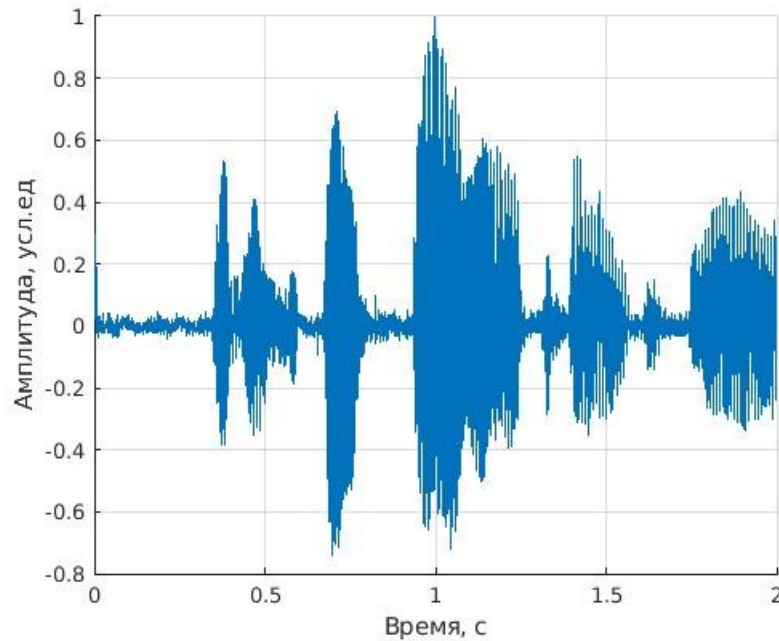


Рис. 2.26. Выходной сигнал адаптивного LC RLS.

Из рис. 2.24-2.26 видно, что на качество фильтрации отклонение элемента МР от исходного положения не повлияло. Это произошло из-за того, что линейные ограничения настроены на приём сигнала с направления $\phi = 0$ град и $\theta = 0$ град. Построим зависимость оценки PESQ и выигрыша ОСШ от порядка фильтра для случая отклонения элемента и для случая прямоугольной эквидистантной микрофонной решетки. Данные графики представлены на рис. 2.27-2.30.

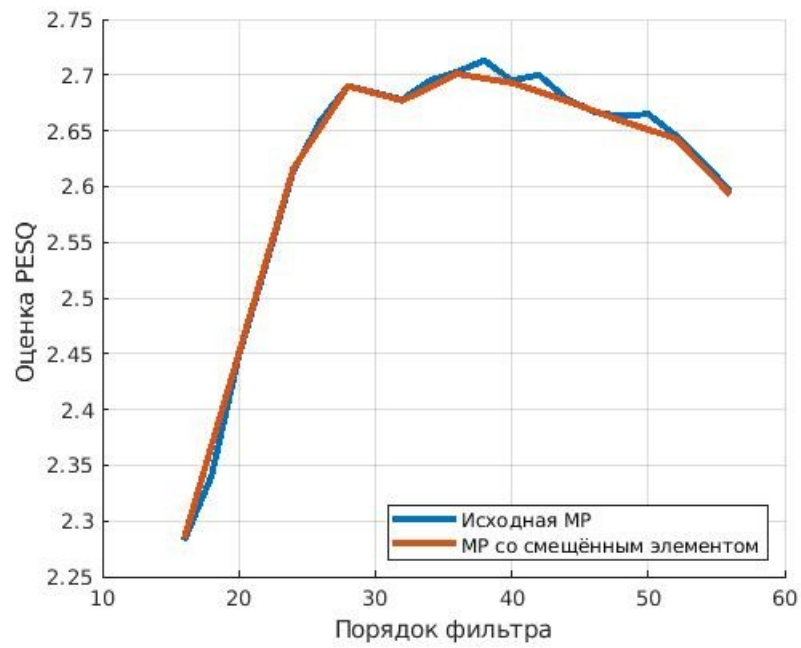


Рис. 2.27. Оценка PESQ для LC NLMS.

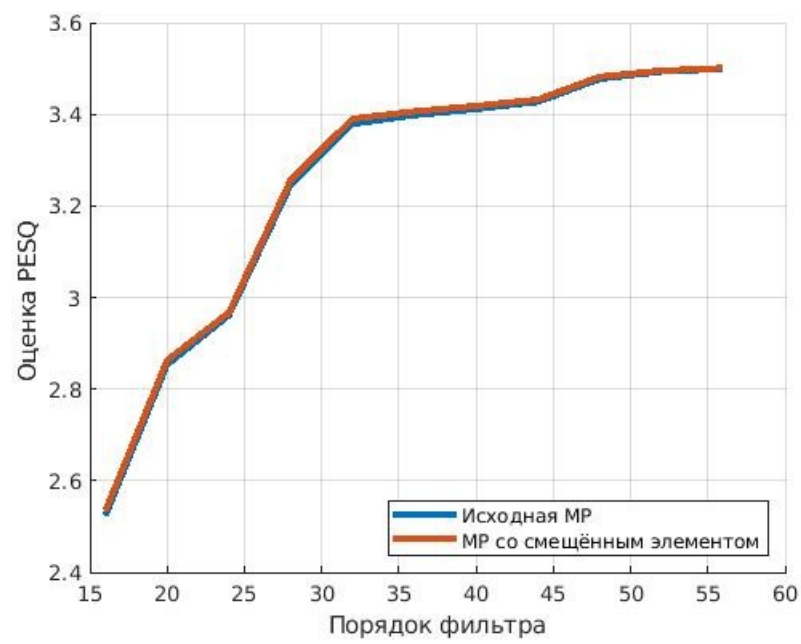


Рис. 2.28. Оценка PESQ для LC RLS.

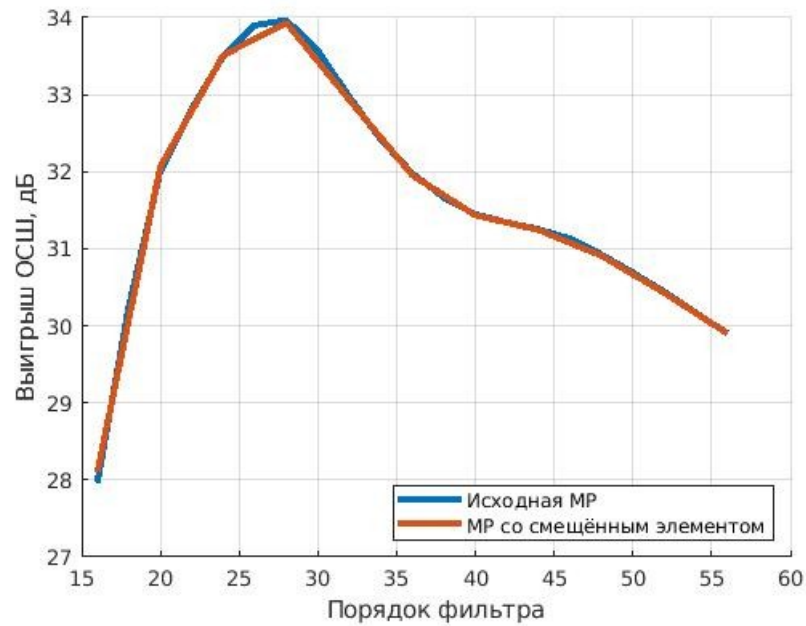


Рис. 2.29. Выигрыш ОСШ для алгоритма LC NLMS.

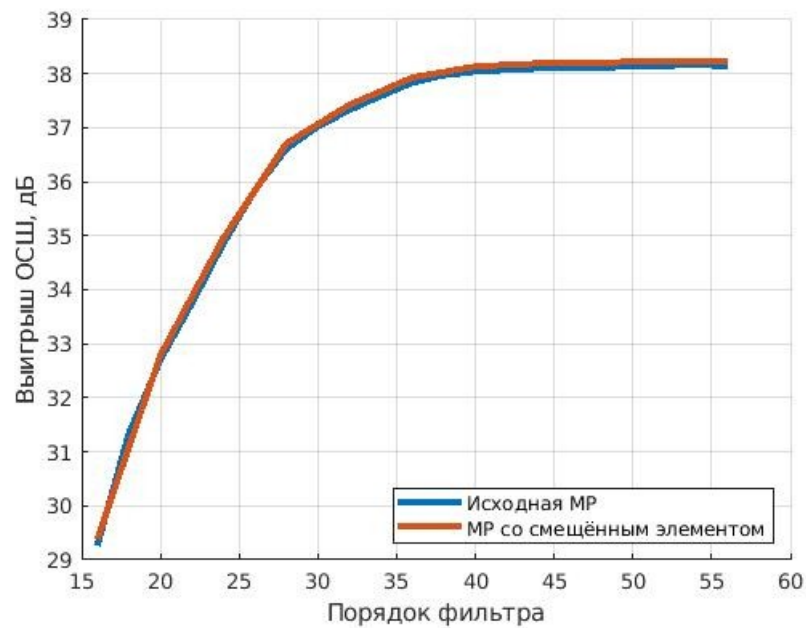


Рис. 2.30. Выигрыш ОСШ для алгоритма LC RLS.

Из рис. 2.27-2.30 видно, отклонение элемента МР не приводит к значительному ухудшению качества фильтрации. На каждом рисунке видно, что различие между данными микрофонными решетками незначительно. Можно сделать вывод о том, что отклонение элемента МР не приводит к ухудшению работы по выделению полезного сигнала, пришедшего с

направления $\phi = 0$ град и $\theta = 0$ град. При этом если ограничения на расстояния между микрофонными решетками не будут выполняться, то фильтрация будет произведена с ошибками.

3. Аппаратная реализация микрофонной решетки

В ходе работы была реализована собственная микрофонная решетка с адаптивным диаграммообразованием. Звуковые сигналы с каждого микрофонного элемента поступают на микроконтроллер для последующей обработки. Результирующий сигнал передаётся на компьютер для дальнейшего анализа. Структурная схема системы, реализующую МР представлена на рис. 3.1.

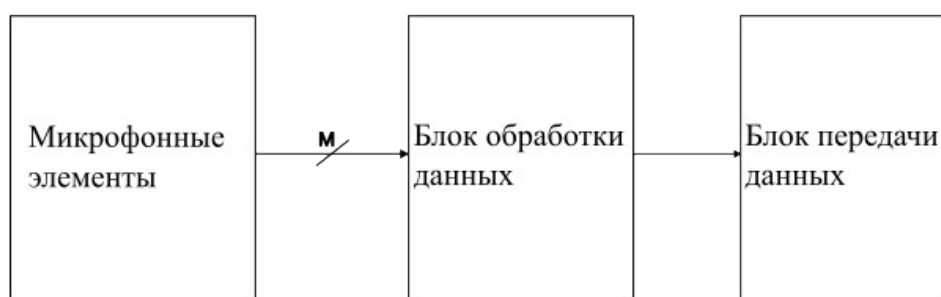


Рис. 3.1. Структурная схема микрофонной решетки.

Для реализации адаптивных алгоритмов, которые были представленных выше, требуется выполнить условие синхронности получения данных с каждого элемента. Классические аналоговые микрофоны не обеспечивают выполнение данного требования. Для решения данной проблемы в работе используются цифровые MEMS микрофоны, звуковые данные которых представлены при помощи импульсно-плотностной модуляцией (PDM). PDM используется для представления аналогового сигнала в цифровой домен, представляя собой высокочастотный однобитный поток цифровых импульсов. В сигнале PDM относительная плотность импульсов соответствует амплитуде аналогового сигнала. Большой кластер логических единиц соответствует высокой (положительной) величине амплитуды, в то время как большой кластер логических нулей соответствует низкой (отрицательной) величине амплитуды, и смена лог. 1 и лог. 0 одинаковой длительности соответствует нулевому значению амплитуды.

Одной из отличительных особенностей таких микрофонов является возможность объединить два сигнальных выхода с двух элементов в один провод, тем самым обеспечивая меньшее количество проводов необходимых для получения данных со всех микрофонов. На рис. 3.2 представленная диаграмма сигналов с каждого микрофона. Видно, что при положительном фронте тактируемого сигнала данные поступают с первого микрофона, а при отрицательном фронте данные поступают со второго. Временное разделение сигналов в течении одного периода позволяет объединить эти данные на один провод.

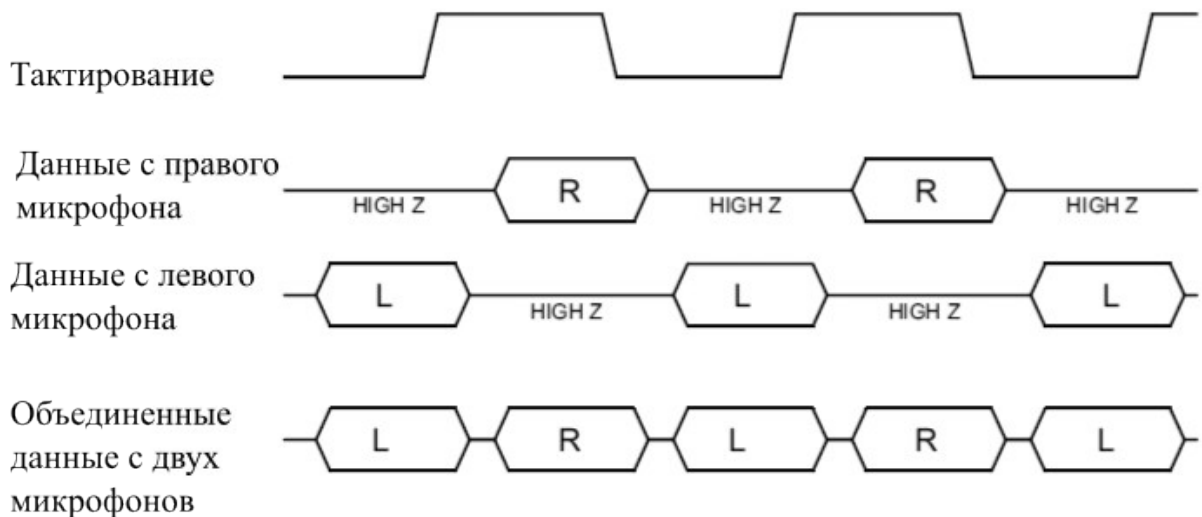


Рис. 3.2. Диаграммы сигналов микрофонов в конфигурации стерео

Для дальнейшей обработки полученных сигналов необходимо конвертировать их в формат PCM (импульсно-кодовая модуляция). Чтобы преобразовать поток PDM в выборки данных PCM, поток PDM нужно отфильтровать и подвергнуть децимации. На стадии децимации частота дискретизации сигнала PDM снижается до необходимой частоты дискретизации звука (например, 16 кГц). Децимация осуществляется взятием одной выборки из M выборок, при этом частота дискретизации понижается в M раз. Таким образом, частота данных PDM (которая равна частоте тактов микрофона) понижается в M раз, чтобы частота выборок звука соответствовала нужной частоте выборок в приложении, и коэффициент M

называется коэффициентом децимации. Стандартным значением коэффициента децимации является диапазон от 32 до 64. Синхронизация данных со всех элементов обеспечивается за счёт общего источника тактирования. Как представлено выше, на рис. 3.2 пара микрофонов выдают данные с разницей в пол периода, что нарушает требование синхронности данных со всех элементов, однако из-за повышенной частоты PDM формата разница между двумя отчётами в одном такте несоизмерима с периодом дискретизации микрофонов и это не приводит к искажению фаз сигналов. Реализация процесса фильтрации с последующей децимацией на микроконтроллере описана в статье [13].

Для принятия звуковых данных с последующей обработкой в работе использовался двухъядерный микроконтроллер STM32H745ZI на базе отладочной платы NUCLEO-H745ZI-Q, которая представлена на рис 3.3.

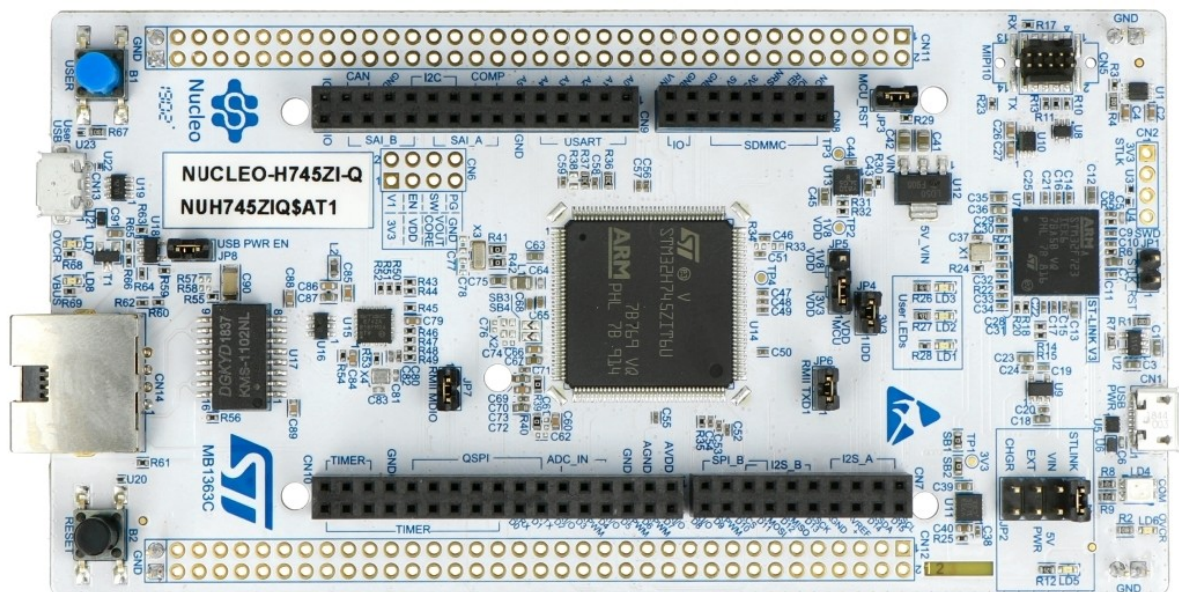


Рис. 3.3. Микроконтроллер STM32H745ZI на отладочной плате NUCLEO-H745ZI-Q.

Для захвата PDM данных со всех микрофонов можно использовать такие периферии как SPI, I2S и SAI. Однако при помощи них возможно осуществить синхронный приём данных лишь с двух микрофонов

одновременно. Для использования большего количества элементов потребуется использовать несколько таких периферий, однако таким способом невозможно осуществить синхронный приём из-за отсутствия механизма синхронизации двух периферий между собой. Для решения данной проблемы в работе используется специализированная для таких задач периферия цифровых фильтров для сигма-дельта модуляторов (DFSDM). Данная периферия реализует вышеописанный процесс конвертации PDM данных в PCM формат. Цифровые MEMS-микрофоны предоставляют выходные данные в формате PDM, которые могут быть напрямую направлены в DFSDM. DFSDM предоставляет отфильтрованные и прошедшие децимацию выборки. Каждый фильтр имеет свой собственный канал DMA, вследствие чего выборки каждого канала разделяются. Это позволяет приложению избежать тяжелой нагрузки операций фильтрации и удаления чередования бит. Структурная схема показывающая работу данной периферии с четырьмя микрофонами представлена на рис. 3.4.

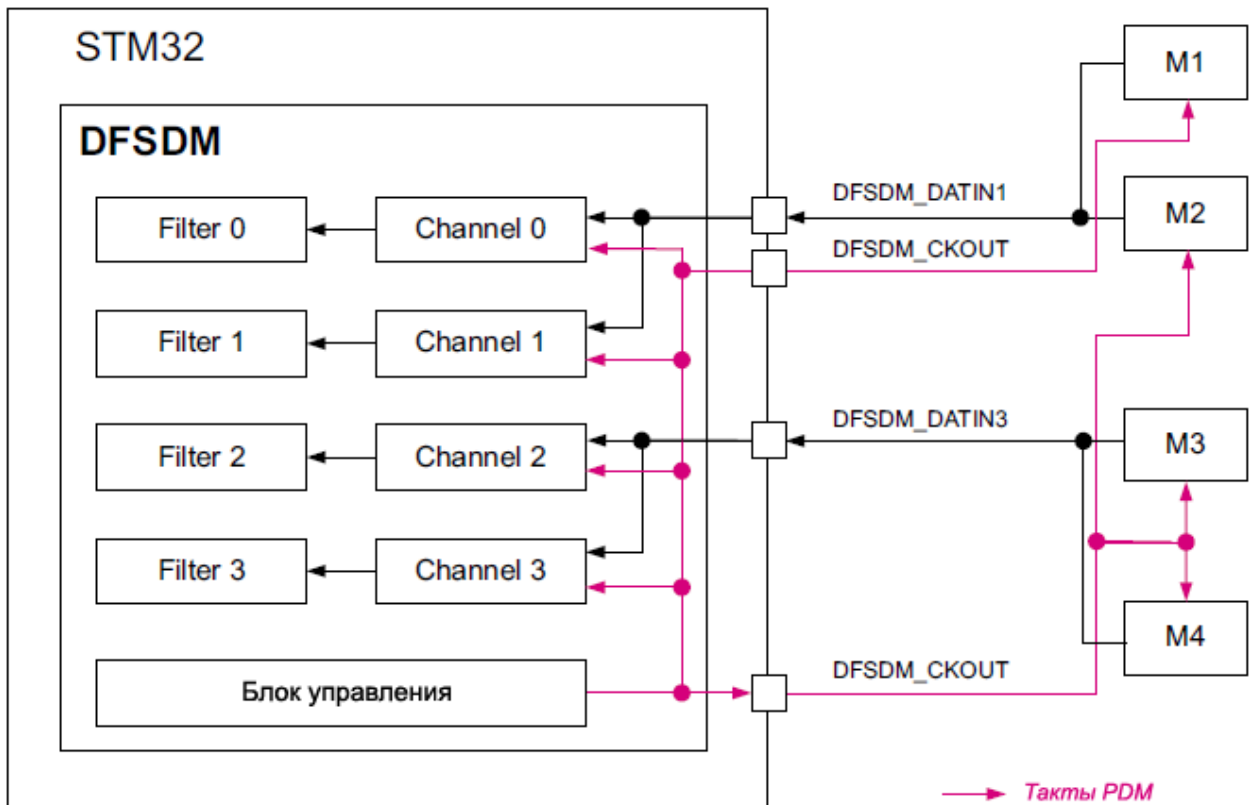


Рис. 3.4. Структурная схема подключения четырёх микрофонов к периферии DFSDM.

Каждый фильтр имеет возможность независимой настройки своих параметров, а так же выбор источника входного сигнала. Синхронность получения данных в данной периферии обеспечивается за счёт возможности запуска всех фильтров вместе с первым. Таким образом, настроив 3 фильтра на приём данных в режимом синхронного запуска относительно первого фильтра, запуск первого приведёт к синхронной обработке 4 микрофонов. Для уменьшения процессорных ресурсов на обработку данных используется канал DMA в циклическом режиме вместе с включённым прерыванием, обеспечивающим непрерывную обработку данных. Для подтверждения синхронности работы периферии DFSDM для каждого канала при помощи цифровых ножек микроконтроллера и логического анализатора была построена временная зависимость, показывающая моменты времени, в которых происходят прерывания по DMA в каждом канале. Вышеописанная

зависимость представлена на рис 3.5, где каждый фронт обозначает момент времени, когда на определённый канал пришло прерывание.



Рис. 3.5. Временная диаграмма каждого канала.

Из рис. 3.5 видно, что данные с микрофонов поступают синхронно. Небольшие разницы во времени присутствуют из-за времени, необходимого на обработку каждого прерывания.

Полученные данные поступали на блок адаптивной фильтрации, выходные данные которого поступали на блок передачи данных на компьютер. Последний блок реализован посредством внешней микросхемы W5500 от компании WIZnet. Микросхема W5500 представляет собой встроенный Ethernet-контроллер с жестким подключением TCP/IP, который упрощает подключение к Интернету для встроенных систем с использованием SPI [14, 15]. Данная микросхема берёт на себя обработку низкоуровневых уровней модели OSI, тем самым позволяет не тратить процессорное время на обработку этих данных. Интерфейс взаимодействия с W5500 является SPI с частотой тактирования вплоть до 30 МГц. На основе

этой микросхемы был поднят UDP сокет при помощи готовой библиотеки от самих производителей [16, 17]. Микросхема имеет встроенный буфер объемом 32 Кбайт, который используется для хранения данных на отправку. Всего в микросхеме можно поднять одновременно 8 отдельных сокетов с реализацией различных протоколов. Передача данных с микроконтроллера на W5500 осуществляется в блокирующем режиме, однако приём данных с хоста основывается на прерывании микросхемы по событию получения данных с хоста. На основе статьи об анализе пропускной способности данной микросхемы было подтверждено, что она обеспечит непрерывный поток данных [18].

После получения всех данных с каждого канала при помощи DFSDM, осуществляется обработка аудио сигналов при помощи адаптивного фильтра. В работе реализованы алгоритмы LC NLMS и LC RLS для 4 микрофонов. Полученные на выходе адаптивного фильтра данные отправляются на компьютер посредством UDP протокола. Для получения аудиоданных использовалось готовая плата с четырьмя микрофонами, расположенными вдоль одной линии. Плата представлена на рис. 3.6.



Рис. 3.6. Плата с микрофонами

В ходе припаивания микрофонов из позиция может измениться относительно ожидаемой. Это может внести изменения в ход работы алгоритма, однако в главе 8 рассмотрен такой случай и было установлено, что при тех алгоритмах,

что рассматривались в работе, смещение положения элемента не приводит к значительному ухудшению качества фильтрации. Расстояние между микрофонами на данной плате составляет 2 см. Напряжение питания составляет диапазон от 1.8 В до 3.6 В. В реализации подаётся 3.3 В, совпадающее с логическим уровнем микроконтроллера. На плате, представленной на рис. 3.6 установлены MEMS микрофоны ADMP521. Схема подключения пары микрофонов представлена на рис. 3.7. Таким образом к плате на рис. 3.6 необходимо подключить пять проводов следующего типа: питание, земля, тактирование, два выходных провода с данными. Остальные провода используются для управления светодиодами на плате, их функционал не задействован в работе.

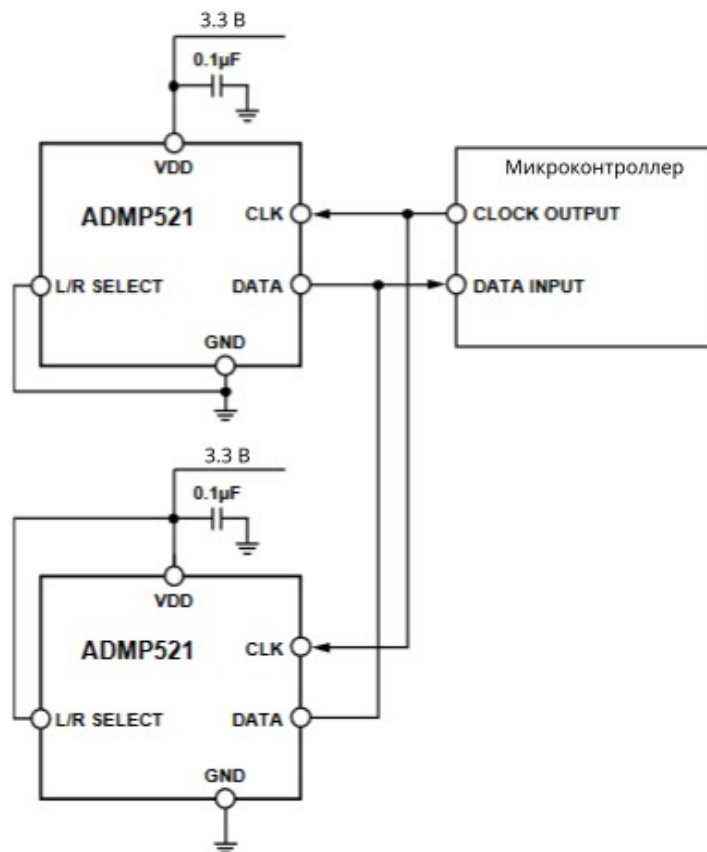


Рис. 3.7. Схема подключения пары микрофонов к микроконтроллеру.

Для объединения двух микрофонов на один провод необходимо контакт L/R SELECT одного микрофона подсоединить к напряжению питания, а другой к земле.

Конечная схема реализации микрофонной решетки представлена на рис. 3.8.

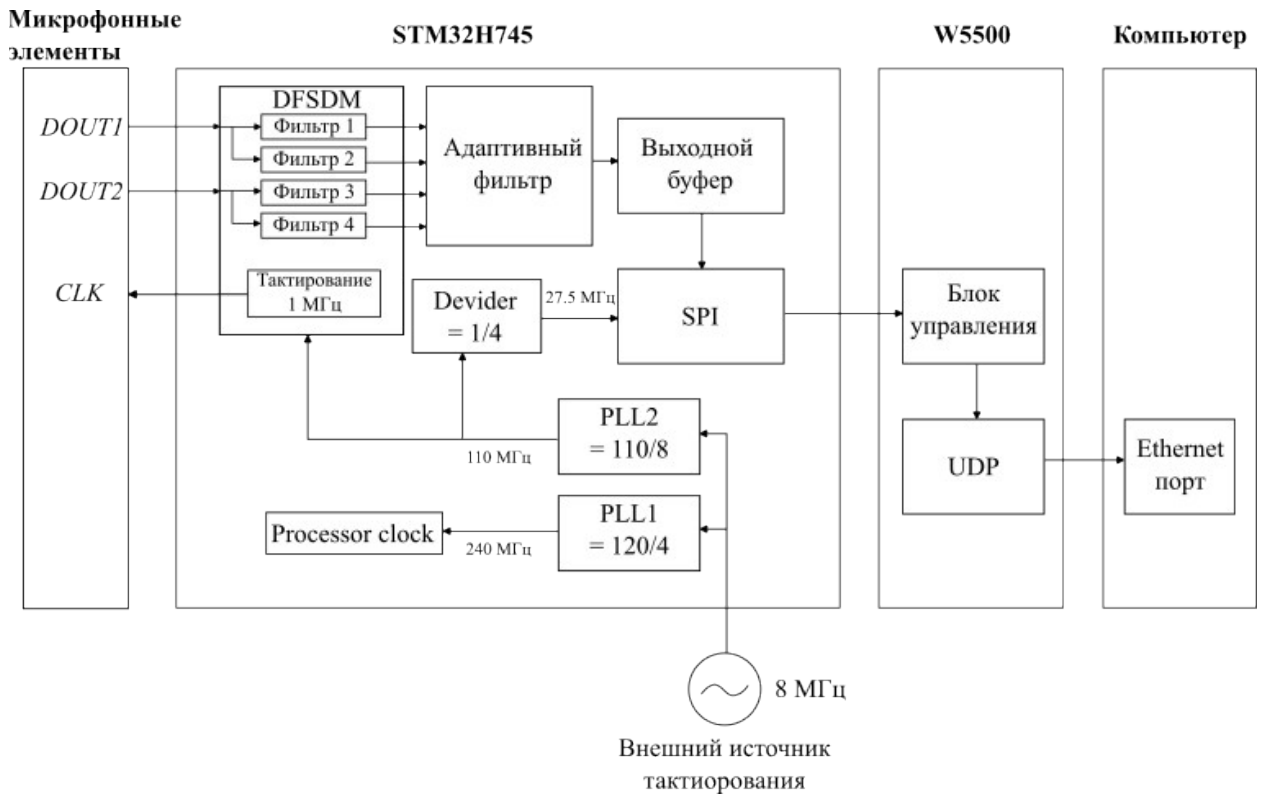


Рис. 3.8. Схема реализации МР на базе STM32H745ZI

Тактирование микроконтроллера осуществляется от внешнего кварцевого резонатора с частотой 8 МГц. Два отдельных модуля PLL используются для генерации частоты тактирования процессора, а так же периферий SPI и DFSDM. Внутри периферии SPI используется делитель частоты со значение 4, таким образом SPI тактируется на частоте 27.5 МГц. Эта частота задаёт скорость передачи данных между микроконтроллером и микросхемой W5500. Процессор тактируется с частотой 240 МГц. Код программы для микроконтроллера был написана с использование операционной системы реальной времени FreeRTOS [19]. Модуль DFSDM

тактируется с той же частотой, что и модуль SPI, однако частота тактирования внешнего микрофона равна 1 МГц. Все внутренние фильтры предназначенные для преобразования PDM данных в PCM реализуют sinc-фильтр третьего порядка. Передаточная функция такого фильтра имеет следующий вид:

$$H(z) = \left(\frac{1 - z^{-Fosr}}{1 - z^{-1}} \right)^3, \quad (9.1)$$

где $Fosr$ - коэффициент децимации фильтра.

В данной реализации коэффициент децимации фильтра равен 125. Преобразованные данные с каждого фильтра передаются в пред выделенный буфер при помощи периферии DMA. При накоплении в буфере первых 128 отсчётов, DMA вызывает прерывание в котором выставляется флаг готовности этих данных. Отдельный поток проверяет готовность всех каналов для обработки и осуществляет адаптивную фильтрацию. Полученные на выходе адаптивного фильтра данные накапливаются для передачи по SPI в микросхему W5500. Программа на компьютере осуществляет приём обработанных данных и накопление их для последующего сохранения их в wav файл.

Заключение

Плоские равномерные микрофонные решетки вместе с адаптивным формирователем луча позволяют создавать гибкие системы приёма полезного речевого сигнала. В ходе работы был изучен способ моделирования приёма сигнала на микрофонную решетку. Так же были реализованы два адаптивных алгоритма LC NLMS и LC RLS. Их анализ показал, что они подстраивают свои весовые коэффициенты, подавляя шумы с направлений, отличных от направления полезного сигнала. Был проведён анализ двух адаптивных формирователей луча. Можно сделать вывод, что алгоритм LC RLS даёт большие значения выигрыша ОСШ и оценки PESQ при равных условиях, в сравнении с LC NLMS. При этом вычислительная сложность первого алгоритма больше, чем у второго. Так же было установлено, что существует минимальный угол между направлением полезного сигнала и шума, при котором адаптивные формирователи луча могут корректно отфильтровать полезный сигнал. Помимо этого, микрофонная решетка вместе с адаптивным формирователем луча может осуществлять фильтрацию звука в случае резко изменяющейся шумовой обстановки. Отклонение одного из элементов МР при фильтрации сигнала не приводит к значительному ухудшения качества сигнала на выходе. Данный факт был использован в дальнейшем при проектировании микрофонной решетки, ведь в случае механического смещения положения элемента, функционал такого устройства не изменится. Все выше перечисленные факты говорят о том, что микрофонная решетка с адаптивным формирователем луча является эффективным устройством для записи речи в шумовой обстановке.

Список литературы

1. McCowan, I. A. Robust Speech Recognition using Microphone Arrays / I. A. McCowan. PhD Thesis, Queensland University of Technology, Australia, 2001.
2. Аверина, Л. И. Выделение речи целевого диктора в помеховой обстановке за счёт амплитудно-фазового синтеза линейной микрофонной решётки / Л. И. Аверина, Д. С. Кузнецов // Теория и техника радиосвязи. 2019. № 3. С. 30–36.
3. Столбов, М. Б. Применение микрофонных решеток для дистанционного сбора речевой информации / М. Б. Столбов // Научно-технический вестник информационных технологий, механики и оптики. – 2015. – Т. 15, № 4. – С. 661 – 675.
4. Liu, Wei Wideband beamforming : concepts and techniques / Wei Liu, Stephan Weiss // Proceedings of the IEEE. – 2010. – P. 286.
5. Ричард Лайонс Цифровая обработка сигналов. - 2-е изд. - 2006: 000 «Бином-Пресс» , 536 с.
6. Баскаков С.И. Радиотехнические цепи и сигналы / С.И. Баскаков. – М. : Высш. шк., 1983. – 536 с.
7. Fante R.L., Vaccaro J.J. Wideband cancellation of interference in a GPS receive array. IEEE Transactions on Aerospace and Electronic Systems, 2000, vol. 36, no. 2, pp. 549-564
8. Rasekh M., Seydnejad S.R. Design of an adaptive wideband beamforming algorithm for conformal arrays. IEEE Communications Letters, 2014, vol. 18, no. 11, pp. 1955–1958.
9. El-Keyi A., Champagne B. Adaptive linearly constrained minimum variance beamforming for multiuser cooperative relaying using the kalman filter. IEEE Transactions on Wireless Communications, 2010, vol. 9, no. 2, pp. 641–651.

10. Джиган, В. Адаптивная фильтрация сигналов / В. Джиган. — М. : Техносфера, 2013. — 528 с.
11. Gutierrez, Elijah, Oplustil, Pilar, Lai, Catherine PY. Location, Location: Enhancing the Evaluation of Text-to-Speech Synthesis Using the Rapid Prosody Transcription Paradigm - 2021/07/06 — P. 6
12. A. W. Rix, J. G. Beerends, M. P. Hollier and A. P. Hekstra, "Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs," *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No. 01CH37221)*, Salt Lake City, UT, USA, 2001, pp. 749-752 vol.2, doi: 10.1109/ICASSP.2001.941023.
13. Кузнецов Д.С., Аверина Л.И. Микрофонные решетки с адаптивным формирователем луча // Вестник ВГУ. Серия: Физика. Математика. — 2021. — № 1. — С. 5 – 15.
14. Leens. F, "An Introduction to I2C and SPI Protocols" IEEE Instrumentation & Measurement Magazine, pp. 8-13, February 2009.
15. Motorola Inc., "SPI Block Guide V03.06," February 2003
16. Bova, T. & Krivoruchka, T. "Reliable UDP Protocol," IETF Internet-Draft, 25 February 1999.
17. He, E.; Leigh, J.; Yu, O.; DeFanti, T.A., "Reliable blast UDP: Predictable high performance bulk data transfer," In Proceedings of IEEE Cluster Computing, pages 317-324, September 2002
18. Жмудь Вадим Аркадьевич, Незванов Александр Игоревич, Трубин Виталий Геннадьевич СРАВНЕНИЕ СКОРОСТИ ПЕРЕДАЧИ ПО ПРОТОКОЛАМ TCP/UDP С STM32F103 И WIZNET W5500 // Автоматика и программная инженерия. 2019. №4 (30). URL: <https://cyberleninka.ru/article/n/sravnenie-skorosti-peredachi-po-rotokolam-tcp-udp-s-stm32f103-i-wiznet-w5500-1>

19. Курниц Андрей FreeRTOS - операционная система для микроконтроллеров // Компоненты и Технологии. 2011. №11 (124).
URL: <https://cyberleninka.ru/article/n/freertos-operatsionnaya-sistema-dlya-mikrokontrollerov-6>