

Содержание

1 Введение.....	3
2 Моделирование приёма акустического сигнала на плоскую микрофонную решётку.....	5
3 Формирователи луча.....	9
4 Пространственная фильтрация с применением плоской микрофонной решетки.....	12
5 Методы оценки качества речи.....	17
6 Постановка задачи приёма и фильтрации сигнала с помехой.....	17
7 Объективная оценка качества речи PESQ.....	18
8 Результаты моделирования.....	19
9 Вывод.....	20

1 Введение

Пространственно-частотная фильтрация применяется в различных областях науки для избирательного приёма сигналов от объектов, находящихся в различных точках пространства. Данная фильтрация достигается использованием приёмных систем с характеристиками направленности заданной формы. Одним из таких устройств является микрофонная решетка. Микрофонная решетка – 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, идущих со всех сторон. Совместная обработка полученных реализаций позволяет отфильтровать звуковой сигнал, исходящий из заданного направления. 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 используют для получения объёмной звуковой картины, что нашло своё применение в автомобильной промышленности. Системы микрофонов позволяют определить точное место нежелательных источников шумов. Многие отрасли, в которых шум может стать причиной отказа устройства, например, ветряные турбины, могут извлечь полезную информацию из микрофонной решетки, чтобы точно определить источники проблем. Для сбора речевого сигнала в присутствии шума обычные микрофоны с последующим усилением сигнала не подходят, поскольку вместе с полезным сигналом усиливается и помехи, который приводит к снижению разборчивости и качества речи. Однако микрофонные решетки могут с этим справиться, если источник полезного сигнала и источник шума имеют различные геометрические положения.

Целью данной курсовой работы является изучение основ цифровой обработки акустических сигналов с помощью плоской микрофонной решетки, а именно моделирование приёма сигнала и подавление точечной помехи. 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 разновидности формирователей луча и их взаимосвязь.

2 Моделирование приёма акустического сигнала на плоскую микрофонную решётку

Речевой сигнал является один из основных способов передачи информации между людьми. Приём речевых сигналов необходим для дальнейшего хранения и передачи информации между людьми. Однако речевые сигналы имеют ряд особенностей, которые необходимо учитывать во время моделирования приема акустического сигнала с помощью микрофонной решетки.

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 равномерная микрофонная решетка. В ней элементы МР расположены на равном расстоянии между собой. Схематически это представлено на рис. 1.

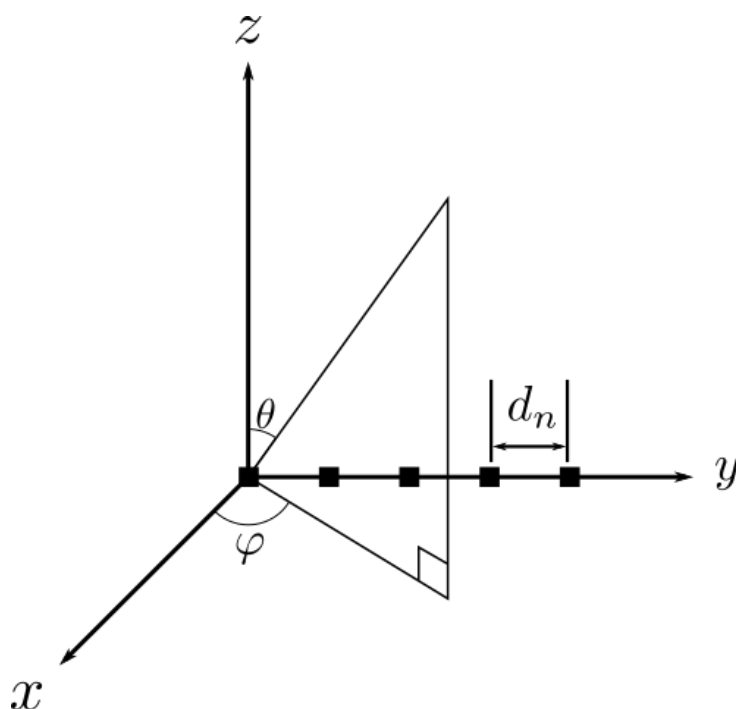


Рис. 1. Прямоугольная равномерная микрофонная решетка

Угол φ – угол азимута, θ – угол подъёма. Расстояния между микрофонами d_x и d_y

должно выполнять следующее условие: $d_x \leq \frac{\lambda_{max}}{2}$ 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: $M \geq \frac{f_{max}}{f_{min}}[1]$. 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. В результате при выборе шага решётки для средней частоты диапазоны на низких частотах решетка оказывается ненаправленной, а на высоких имеет максимумы чувствительности при ложных угловых направлениях поступления сигнала. Для дальнейшего моделирования, использовалась прямоугольная равномерная микрофонная решетка, состоящая из $N=16$ элементов. Расстояние между элементами МР использовали для частоты 4 кГц по следующей формуле: $d_x = \frac{\lambda}{2}$ и $d_y = \frac{\lambda}{2}$. 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, при этом элементы решетки расположены в пространстве упорядоченным образом. На элементы микрофонной решетки речевой сигнал приходит с различной задержкой. Для вычисления задержки сигнала на элементе МР вводят матрицу векторов местоположения элементов МР:

$$\mathbf{p} = [\mathbf{p}_1 \ \mathbf{p}_2 \ \dots \ \mathbf{p}_N], \quad (1)$$

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

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

$$\mathbf{a} = \begin{bmatrix} -\cos \theta \cos \phi \\ -\cos \theta \sin \phi \\ -\sin \theta \end{bmatrix}. \quad (2)$$

Задержка на каждом канале микрофона вычисляется по следующей формуле:

$$\tau = \frac{\mathbf{a}^T \mathbf{p}}{c}, \quad (3)$$

где c – скорость распространения акустической волны в атмосфере.

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 это, используя линии задержки, но в противном случае добиться точной задержки 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:

$$X_{\text{shifted}}(m) = e^{\frac{-j2\pi}{N}nm} X(m), \quad (4)$$

где $X_{\text{shifted}}(m)$ – ДПФ сдвинутой последовательности, $X(m)$ – ДПФ исходной последовательности, n – количество задержанных отсчетов, m – номер отсчёта, N – количество отсчётов[3]. Данная формула работает для гармонических сигналов, но не позволяет производить задержку для речевых сигналов. 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:

$$m = \frac{N}{2n}. \quad (5)$$

После каждому промежутку находим соответствующее центральное значение и умножаем каждый промежуток ДПФ последовательности на экспоненту со своим центральным значением. Результаты моделирования данным способом в

зависимости от числа поддиапазонов представлены ниже на рисунках. Максимальное число диапазонов находится по следующей формуле Это следует из того, что минимальное количество элементов в диапазоне $n_{max} = \frac{N}{2}$ может быть равно единице, подставляя данное значение в формулу (5) получаем максимальное число диапазонов.

3 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 углом 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 местоположений. Формирователь луча выполняет пространственную фильтрацию для разделения сигналов. На рисунке 1 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:

$$y(t) = \sum_{m=0}^{M-1} x_m(t) w_m^*, \quad (6)$$

где $*$ обозначает комплексное сопряжение.

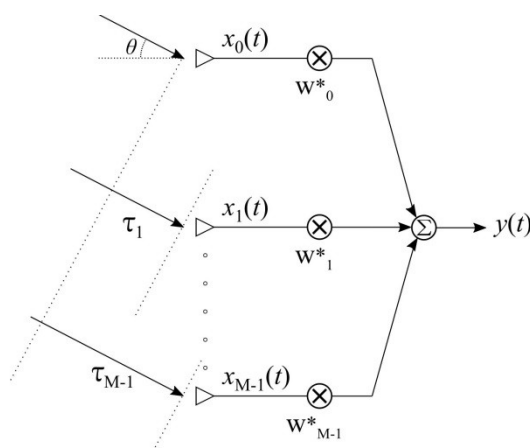


Рис. 1. Узкополосный формирователь луча.

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. Для широкополосного сигнала процедура формирования луча становится более сложной. Поскольку каждый широкополосный сигнал состоит из бесконечного числа различных частотных компонентов, то значения весовых коэффициентов должно быть различным для разных частот. В данном случае вектор весовых коэффициентов можно записать в следующем виде:

$$\mathbf{w} = [w_0(\omega) w_1(\omega) \dots w_{M-1}(\omega)]^T \quad (7)$$

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 2.

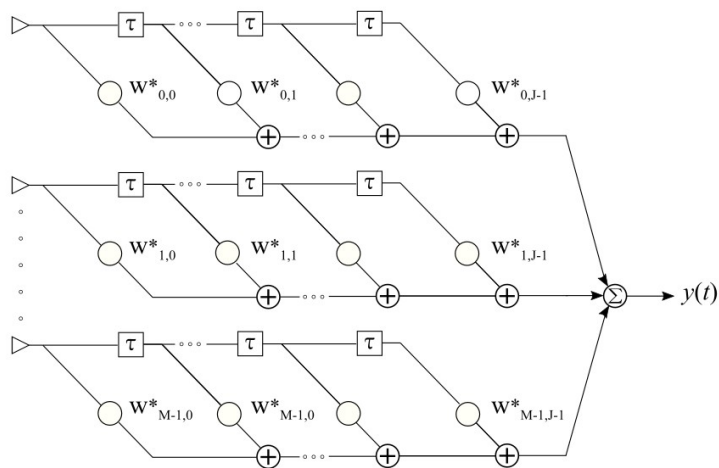


Рисунок 2. Широкополосный временной формирователь луча

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

$$y(t) = \sum_{m=0}^{M-1} \sum_{i=0}^{J-1} x_m(t - iT_s) * w_{m,i}^* \quad (8)$$

Для определения отклика данного формирователя луча вводятся ряд переменных. Сперва вводится волновой вектор, который основывается на формуле (2):

$$\mathbf{k} = \frac{2\pi}{\lambda} \mathbf{a}. \quad (9)$$

Фазирующий вектор имеет следующий вид:

$$\mathbf{v}_k(\mathbf{k}) = \begin{bmatrix} e^{-j\mathbf{k}^T \mathbf{p}_1} \\ e^{-j\mathbf{k}^T \mathbf{p}_2} \\ \dots \\ e^{-j\mathbf{k}^T \mathbf{p}_N} \end{bmatrix}, \quad (10)$$

Отклик данного формирователя луча определяется формулой:

$$P(\omega, \phi, \theta) = \mathbf{W}^T * \mathbf{v}_k(\mathbf{k}), \quad (11)$$

где \mathbf{W} – Фурье образ коэффициентов (7).

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:

$$BP = 20 \log_{10} \left(\frac{|P(\omega, \phi, \theta)|}{\max |P(\omega, \phi, \theta)|} \right). \quad (12)$$

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

$$\mathbf{d}(\phi, \theta, \omega) = [e^{-j\omega \tau} \ e^{-j\omega(\tau+T_s)} \ \dots \ e^{-j\omega(\tau+(J-1)T_s)}]. \quad (13)$$

Используя формулу (13) и (7) отклик формирователя луча можно записать следующим образом:

$$P(\omega, \phi, \theta) = \mathbf{w}^T * \mathbf{d}(\phi, \theta, \omega). \quad (14)$$

4 Пространственная фильтрация с применением плоской микрофонной решетки

Использование формирователя луча с постоянными весовыми коэффициентами неэффективно в случае динамической помеховой обстановке. Настроенные формирователь луча на определённое направление не позволяет качественно отфильтровать шум, пришедший с другого направления. Для таких целей использую адаптивные формирователи луча, которые определяют статистику входных данных и адаптирует весовые коэффициенты к оптимальному с точки зрения определенных критериев решению. Структура адаптивных формирователей лучей представлена на рисунке 3.

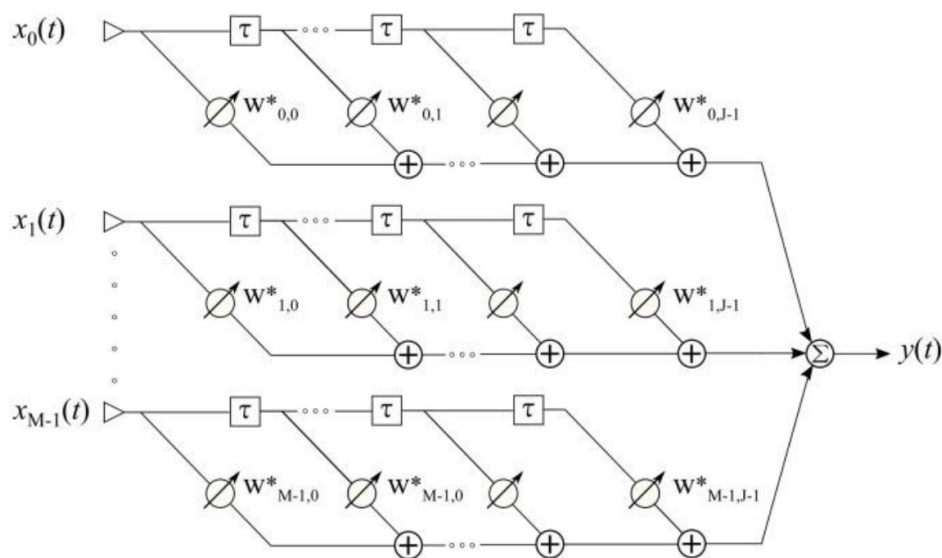


Рисунок 3. Адаптивный формирователь луча

Здесь, M – число элементов МР, J – порядок КИХ фильтра. Существует большое количество адаптивных алгоритмов различающиеся по вычислительной сложности, скорости сходимости, переходными процессами. Алгоритм характеризуется видом используемой целевой функции, методом поиска оптимального решения и природой сигналов ошибок. Для дальнейшего описания алгоритмов, введём общие понятия. Вектор сигналов для каждой итерации имеет следующий вид:

$$\mathbf{X}(k)=[\mathbf{X}_0(k)\mathbf{X}_1(k)\dots\mathbf{X}_{M-1}(k)], \quad (15)$$

где $\mathbf{X}_i(k)=[x_i(k) \ x_i(k-1) \ \dots \ x_i(k-J+1)]$ – вектор сигнала канала.

Вектор весовых коэффициентов КИХ фильтров:

$$\mathbf{W}(k-1)=[\mathbf{W}_0(k-1) \ \mathbf{W}_1(k-1) \ \dots \ \mathbf{W}_{M-1}(k-1)], \quad (16)$$

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.

Выходной сигнал с учётом формул (15) и (16) можно записать следующим образом:

$$y(k)=\mathbf{W}(k-1)^H \mathbf{X}(k), \quad (17)$$

где H – знак операции эрмитова сопряжения вектора.

В дальнейшем будет рассмотрено два основных адаптивных алгоритма LC NLMS и LC RLS.

LMS–алгоритмы адаптивной фильтрации принадлежат к классу алгоритмов стохастического градиентного поиска. Данный алгоритм использует упрощенный градиент, что приводит к случайным флуктуациям [Джиган]. Целевой функцией, которую минимизирует данный алгоритм, является функция среднеквадратической ошибки (MSE) вид которой представлен ниже:

$$F=f\{e(k)\}=E\{|e(k)|^2\}, \quad (18)$$

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 наш требуемый сигнал должен обладать минимальной мощностью, т. е. стремиться к нулевому. И только линейные ограничения позволяют сохранить сигнал с нужного нам направления [статья Димы]. 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: \mathbf{f}_J . Матрица \mathbf{J} ограничений имеет вид:

$$C_{M \cdot J, J} = \begin{bmatrix} c_J & & \mathbf{0} \\ & \dots & \\ \mathbf{0} & & c_J \end{bmatrix}, \quad (19)$$

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 $M \cdot J$.

После в ходе каждой итерации, когда на элементах МР появляются новые значения акустического сигнала производится фильтрация по формуле (17).

Дальше необходимо провести корректировку весовых коэффициентов:

$$e(k) = -W(k-1)^H X(k)$$

$$\tilde{e}(k) = e(k) [X(k)^H X(k) + [X(k)^H Q_{M \cdot J, J}][C_{M \cdot J, J}^H X(k)] + \delta^2]^{-1}$$

$$W'(k) = W(k-1) + \hat{\mu} X(k) \tilde{e}^*(k)$$

$$W(k) = W'(k) + Q_{M \cdot J, J} [f_J - C_{M \cdot J, J}^H W'(k)],$$

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 порядков фильтра. Недостатком является 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 представлена ниже:

$$F = f\{e(k)\} = \sum_{i=1}^k |e(i)|^2 \quad (20)$$

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:

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

После в ходе каждой итерации, когда на элементах МР появляются новые значения акустического сигнала производится фильтрация по формуле (17). Далее необходимо провести корректировку весовых коэффициентов:

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

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

elibrary

5 Методы оценки качества речи

Для оценки эффективности алгоритмов необходимо методы сравнения входного и выходного сигнала между собой. Стандартным методом оценки качества фильтрации является вычисление выигрыша отношения сигнал/шум(ОСШ) на выходе адаптивного фильтра. ОСШ на входе считается по следующей формуле:

$$SNR_{\text{вх}} = 10 \log_{10} \frac{P_{\text{вх.сиг}}}{P_{\text{вх.шум}}}, \quad (21)$$

Где $P_{\text{вх.сиг}} = \frac{1}{N} \sum_{k=1}^N s^2(k)$ – средняя мощность полезного сигнала на входе фильтра,
 $P_{\text{вх.сиг}} = \frac{1}{N} \sum_{k=1}^N n^2(k)$ – средняя мощность шума на входе фильтра, $s(k)$ –
ОСШ на выходе фильтра имеет вид:

$$SNR_{\text{вых}} = \frac{P_{\text{вых.сиг}}}{P_{\text{вых.шум}}} \quad (22)$$

6 Постановка задачи приёма и фильтрации сигнала с помехой

7 Объективная оценка качества речи PESQ

8 Результаты моделирования

9 Вывод

