

УДК 656.052:681.51.54

Н. М. Быков, к. т. н.; А. В. Денисов; А. Е. Сегеда

МЕТОД БЫСТРОЙ ФИЛЬТРАЦИИ ПОМЕХ В СИСТЕМАХ АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ РЕЧИ

Предложен алгоритм и устройство быстрой фильтрации сигналов помех в речевом сигнале, основывающиеся на модифицированном методе сглаживания скользящими средними.

Ключевые слова: фильтрация, сигналы помех, аппроксимация, коэффициенты фильтра, вычислительная сложность, оптимизированная частотная характеристика.

Постановка проблемы. Проведенные в предыдущих работах авторов исследования [1, 2] показали, что при определенных условиях фильтрация шумов в речевом сигнале способна привести к значительному повышению точности распознавания сигнала. Однако известные методы фильтрации требуют значительных вычислительных затрат, что не позволяет реализовать распознавание в реальном масштабе времени. В связи с этим актуальной задачей является поиск методов фильтрации, требующих минимальных вычислительных затрат компьютера или реализуемых на основе несложных быстродействующих цифровых устройств. Что касается речевого сигнала, то эту задачу можно рассматривать с точки зрения построения сглаживающего фильтра.

Анализ исследований и публикаций по данной проблеме и формулирование целей статьи. Известные классические методы сглаживания находят приближения сигнала с помощью заданных функций [3], в системах распознавания речи нельзя применить по нескольким причинам. Во-первых, заранее неизвестен вид кривой, к которой следует выполнять приближение, во-вторых, такое сглаживание требует наличия информации обо всех отсчетах дискретного сигнала и в реальном времени является невозможным. При таких условиях рациональным является построение сглаживающего фильтра, аппроксимируемого полиномом первой степени, т. е. использующего метод скользящих средних. В данной статье авторы предлагают алгоритм и устройство быстрой фильтрации сигналов помех в речевом сигнале, основывающиеся на модифицированном методе сглаживания скользящими средними.

Математическое обоснование метода фильтрации. Количество отсчетов в "скользящей" группе определяется исходя из требований к точности сглаживания и к сложности фильтра. В качестве критерия оптимальности сглаживания можно использовать, например, минимум суммы квадратов отклонений отсчетов реального сигнала от сглаженного:

$$\sum_{i=1}^m (y_i - y_i^*)^2 = \min, \quad (1)$$

где y_i – текущее значение сигнала на входе фильтра в i -ый момент времени, y_i^* – сглаженный текущий отсчет сигнала, m – количество отсчетов.

Для сглаживания "скользящей" группой из трех точек из условия (1) получаем следующее выражение для отсчетов сглаженного сигнала:

$$y_i^* = \frac{1}{3}y_{i-1} + \frac{1}{3}y_i + \frac{1}{3}y_{i+1}, \quad (2)$$

где y_i^* – сглаженный текущий отсчет сигнала, а y_{i-1} , y_i , y_{i+1} – три последовательные во времени отсчеты сигнала на входе фильтра.

Однако реализация алгоритма сглаживания по данной формуле не является оптимальной

с точки зрения вычислительных затрат, поскольку операции деления требуют наибольших временных и аппаратных затрат. Кроме того, частотная характеристика такого сглаживающего фильтра не оптимизирована, так как он подавляет полезные составляющие, лежащие ниже частоты дискретизации сигнала. Покажем это, для чего найдем z -преобразование выражения (2), а затем передаточную функцию фильтра:

$$Y(z) \cdot z^{-1} = \frac{1}{3} X(z) \cdot (z^{-2} + z^{-1} + z), \quad (3)$$

$$W^*(z) = \frac{Y(z)}{X(z)} = \frac{1}{3} (z^{-1} + z + 1). \quad (4)$$

Частотная характеристика такого фильтра имеет вид:

$$H^*(j\omega) = \frac{1}{3} (1 + e^{-i\omega T} + e^{i\omega T}) = \frac{1}{3} (1 + 2 \cos \omega T). \quad (5)$$

На рис. 1, где кривая 1 изображает модуль частотной характеристики данного фильтра, видно, что в диапазоне частот $[\frac{2\pi}{3} \dots \pi]$ осуществляется подавление полезных составляющих речевого сигнала. В общем случае частотная характеристика устройства, реализующего сглаживание "скользящей" группой из трех точек, имеет такой вид:

$$H(j\omega) = (a_0 + 2a_1 \cos \omega T). \quad (6)$$

Установим требование, чтобы на нулевой частоте модуль передаточной функции был равен 1, а на самой высокой частоте сигнала был равен 0:

$$\begin{cases} a_0 + 2a_1 \cos(\omega T) = 1, & \text{если } f_c = 0; \\ a_0 + 2a_1 \cos(\omega T) = 0, & \text{если } f_c = \frac{1}{2} f_d = \frac{1}{2T}, \end{cases} \quad (7)$$

где a_0, a_1 – коэффициенты сглаживающего фильтра, ω – угловая частота, T – период дискретизации, f_c – частота сигнала, f_d – частота дискретизации.

Из системы (7) найдем, что $a_0 = \frac{1}{2}$, $a_1 = \frac{1}{4}$. Подставим эти значения в (6), получим выражение передаточной функции фильтра:

$$H^*(j\omega) = \frac{1}{2} + 2 \cdot \frac{1}{4} \cos \omega T. \quad (8)$$

Модуль частотной характеристики полученного фильтра изображает кривая 2 на рис. 1.

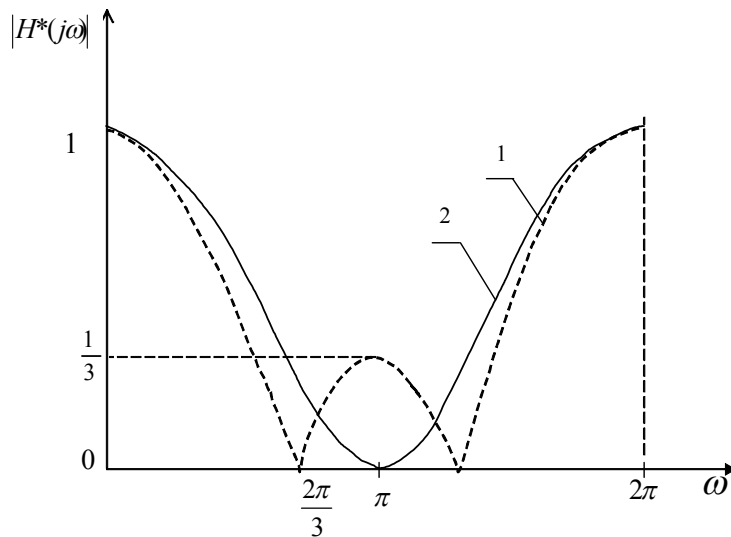


Рис. 1. Частотные характеристики известного (1) и предложенного (2) фильтров

Передаточная функция фильтра, соответствующая данному алгоритму обработки сигнала, имеет такой вид:

$$W(z) = \frac{1}{4}z^{-1} + \frac{1}{2}z^0 + \frac{1}{4}z^{+1}. \quad (9)$$

Во временной области фильтр можно представить в виде свертки последовательных отсчетов сигнала с коэффициентами цифрового фильтра:

$$y_i^* = (y_i + \frac{y_{i-1} + y_{i+1}}{2}) / 2 = \frac{1}{4}y_{i-1} + \frac{1}{2}y_i + \frac{1}{4}y_{i+1}. \quad (10)$$

Количественное уменьшение эффективности сглаживания при использовании предложенного метода по сравнению с классическим можно определить, найдя влияние цифрового фильтра на уровень помех в сигнале. Пусть $y(k) + \varepsilon(k)$ – аддитивная смесь сигнала $y(k)$ и помехи $\varepsilon(k)$, причем математическое ожидание помехи $M[\varepsilon(k)] = 0$, где математическое ожидание M определяется как усреднение по ансамблю сигналов помехи. Предположим также, что для фактических значений смеси $y(k) + \varepsilon(k)$ отсчеты помехи являются некоррелированными, т. е.:

$$M[\varepsilon(k_1), \varepsilon(k_2)] = \begin{cases} \sigma_\varepsilon^2, & k_1 = k_2, \\ 0, & k_1 \neq k_2 \end{cases}, \quad (11)$$

где σ_ε – дисперсия помехи.

Выходной сигнал нерекурсивного фильтра определяют по формуле:

$$y^*(k) = \sum_{i=-N_i}^{N_i} a_i [y(k-i) + \varepsilon(k-i)],$$

где a_i – коэффициенты фильтра.

Поскольку операция усреднения применяется только к $\varepsilon(k)$, то наиболее вероятное значение сигнала на выходе фильтра равно:

$$M[y^*(k)] = \sum_{i=-N_i}^{N_i} a_i \{y(k-i) + M[\varepsilon(k-i)]\} = \sum_{i=-N_i}^{N_i} a_i y(k-i).$$

Вычислим дисперсию входного сигнала:

$$\begin{aligned}
 M\left\{\left[\sum_{i=-N_i}^{N_i} a_i \{y(k-i) + \varepsilon(k-i)\} - M(y^*(k))\right]^2\right\} &= M\left[\sum_{i=-N_i}^{N_i} a_i \varepsilon(k-i)\right]^2 = \\
 &= M\left\{[a_i \varepsilon(k-i)]\left[\sum_{i=-N_i}^{N_i} a_i \varepsilon(k-i)\right]\right\}.
 \end{aligned}$$

Умножив выражение в скобках и учитывая (11), получим:

$$\sum_{i=-N_i}^{N_i} a_i^2 M(\varepsilon^2) = \sum_{i=-N_i}^{N_i} a_i^2 \sigma_\varepsilon^2 = \sigma_\varepsilon^2 \sum_{i=-N_i}^{N_i} a_i^2. \quad (12)$$

Из выражения (12) видно, что подавление шума фильтром определяется суммой квадратов его коэффициентов, поэтому снижение эффективности фильтрации фильтром (10) по сравнению с фильтром (2) определим выражением:

$$E_f = \frac{\sigma_\varepsilon^2 \sum_{i=-N_i}^{N_i} a_{ip}^2}{\sigma_\varepsilon^2 \sum_{i=-N_i}^{N_i} a_{io}^2}, \quad (13)$$

где a_{ip} – коэффициенты данного фильтра, a_{io} – коэффициенты оптимального фильтра.

Подставляя в выражение (13) значения коэффициентов a_{ip} – предложенного цифрового фильтра и a_{io} – классического фильтра, получим количественное значение снижения эффективности сглаживания:

$$E_f = \frac{(1/2)^2 + 2 \times (1/4)^2}{3 \times (1/3)^2} = 1,68.$$

Фильтр (10) реализует вычисления коэффициентов путем сдвигов последовательных отсчетов сигнала и тем самым исключает операции умножения и деления, необходимые для фильтра (2). Несложные подсчеты показывают, что быстродействие фильтрации при этом повышается на порядок и выше. Таким образом, предложенный нами фильтр имеет быстродействие на порядок выше оптимального фильтра при незначительном снижении эффективности фильтрации. Аппаратно такой нерекурсивный фильтр можно реализовать в виде цепочки буферных регистров, выходы которых подключены ко входу сумматора (рис. 2).

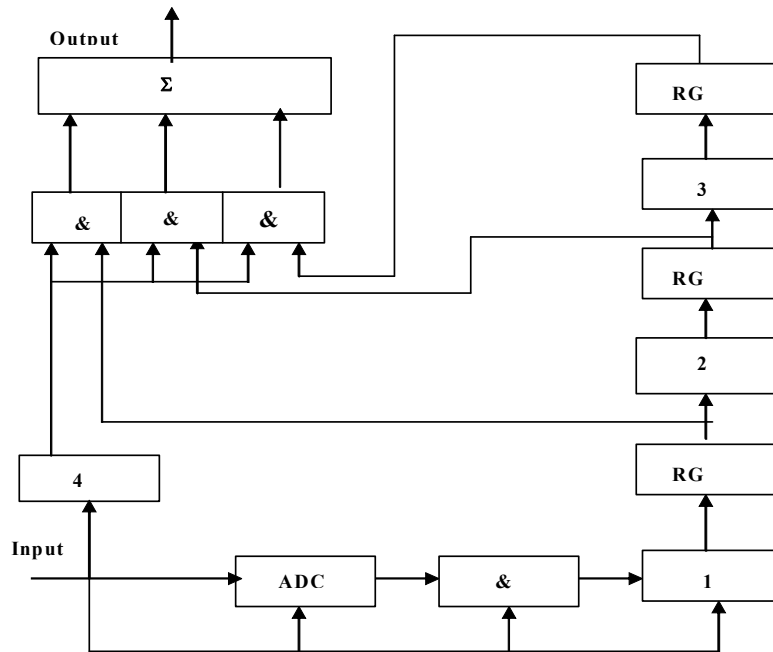


Рис. 2. Структурная схема устройства фильтрации предложенным методом

На рис. 2 приняты следующие обозначения: ADC – аналого-цифровой преобразователь, RG – регистры, 1 – цепь смещения двоичного слова на 2 разряда вправо, 2 – цепь смещения на один разряд влево, 3 – цепь смещения на один разряд вправо, 4 – блок разрешения вывода информации, Σ – двоичный сумматор, & – конъюнкторы.

Работа устройства осуществляется следующим образом: в блоке ADC измеренные в дискретные моменты времени значения сигнала превращаются в цифровой двоичный код (двоичное слово). Также в эти же моменты времени ADC производит тактовые импульсы для синхронизации всего устройства. При передаче двоичных слов, определяющих отсчеты сигнала, из регистра в регистр они сдвигаются цепями сдвига таким образом, что содержание первого и третьего регистров равно $y_{i-1}/4$ и $y_{i+1}/4$ соответственно, а содержимое второго регистра равно $y_i/2$.

Сглаженный отсчет сигнала получаем путем суммирования на сумматоре содержания трех регистров на каждом тактовом импульсе. Блок разрешения вывода информации 4 в начальный момент времени запрещает чтение содержимого регистров до момента заполнения всех трех регистров. В данном устройстве операция деления выполняется одновременно смещением кодов всех составляющих уже при передаче отсчетов сигнала из регистра в регистр. Поскольку отпадает необходимость в выполнении операции деления, то арифметическое устройство значительно упрощается по сравнению с арифметическим устройством для выполнения вычислений по формуле сглаживания (2), поскольку в его состав не нужно вводить устройства для выполнения деления.

Выигрыш в быстродействии предлагаемого сглаживающего фильтра по сравнению с фильтром (2) можно рассчитать по формуле:

$$e_w = \frac{k_d(m_d - n_d)T}{T}, \quad (14)$$

где k_d – количество микроопераций в микропрограмме, реализующей алгоритм деления; n_d – разрядность делителя; T – продолжительность такта на выполнение микрооперации.

Примем $m_d = 8$, $k_d = 6$, $n_d = 2$ (разрядность числа 3), получим $e_w = 6(8-2)=36$.

Основываясь на теории аппроксимации функций [4] и теории цифровых фильтров [5], можно показать, что путем соединения таких элементарных фильтров можно построить быстродействующий фильтр с заданными характеристиками. Доказательство этого положения выходит за рамки данной работы.

Выводы. В данной статье предложен алгоритм и устройство быстрой фильтрации сигналов помех в речевом сигнале, основанные на модифицированном методе сглаживания скользящими средними. Предложенный фильтр реализует вычисления коэффициентов путем сдвигов последовательных отсчетов сигнала и тем самым исключает операции умножения и деления, необходимые для классического фильтра. Математически доказано, что полученный с учетом оптимизации частотной характеристики сглаживающий фильтр имеет быстродействие на порядок выше по сравнению с известным, что позволяет применить его для фильтрации речевого сигнала в реальном масштабе времени. При этом наблюдается незначительное уменьшение эффективности фильтрации.

СПИСОК ЛИТЕРАТУРЫ

1. Биков М. М. Математична модель впливу завад на точність розпізнавання мови / М. М. Биков, І. В. Кузьмін, Л. В. Проценко. – Львів: ДНДІ, 2002. – 290 с.
2. Быков Н. М. Методы и средства измерения и преобразования информации в системах машинного распознавания речи / М. М. Быков. – Винница: ВПИ, 1985. – 243 с.
3. Widrow B. Adaptive Noise Cancelling: Principles and Applications / B. Widrow // IEEE Transactions on Audio and Electroacoustics. – October 1975. – Vol. 63. – № 12. – P. 1672 – 1716.
4. Rabiner L. R. An Approach to the Approximation Problem for Nonrecursive Digital Filters / L. R. Rabiner, B. Gold, C. A. McGonegal. // IEEE Transactions on Audio and Electroacoustics. – October 1971. – Vol. 11. – № 7. – P. 56 – 65.

Быков Николай Максимович – к. т. н., профессор кафедры компьютерных систем управления, nkbykov@mail.ru, тел.: (0432)-598-430.

Денисов Артем Валерьевич – студент кафедры компьютерных систем управления, assd18@yandex.ru.

Сегеда Алена Евгеньевна – студентка кафедры компьютерных систем управления, fate2107@yandex.ru.

Винницкий национальный технический университет.