

АНАЛИЗ ДАННЫХ И СИГНАЛОВ

(Вычислительные системы)

1998 год

Выпуск 163

УДК 621.31: 534.4

АЛГОРИТМ ОЧИСТКИ РЕЧЕВОГО СИГНАЛА ОТ КВАЗИПЕРИОДИЧЕСКОЙ ПОМЕХИ¹

О.А. Кутненко

В в е д е н и е

Проблема очистки сигналов, в том числе и речевых, от различного рода помех является одной из важнейших в области автоматической обработки и распознавания сигналов. Разнообразие подходов к решению указанной проблемы, описанных в монографиях, обзорах, отдельных публикациях [1–5], объясняется, с одной стороны, многообразием видов помех и сигналов, а с другой — отсутствием универсального способа подавления (компенсации) помех. Распространенным видом помех являются аддитивные квазипериодические импульсные последовательности. Методы выделения или детектирования полезного (в частности, речевого) сигнала из смеси с указанным видом помех по-прежнему недостаточно изучены. Известные методы [6–8] ориентированы на компенсацию и ослабление аддитивных чисто периодических или аperiodических помех.

В настоящей работе изложен алгоритм подавления аддитивной квазипериодической импульсной помехи, искажающей речевой сигнал. Рассмотрен случай, когда помеха является детерминированной последовательностью заданных одинаковых импульсов, при этом моменты появления импульсов неизвестны, а интервал между началами соседних импульсов лежит в некотором

¹Работа выполнена в рамках проекта № 97-01-00866, поддержанного Российским фондом фундаментальных исследований.

фиксированном диапазоне, ширина которого не меньше длительности самого импульса.

Приведенное решение задачи опирается на результаты работы [9] и относится к апостериорно-последовательным методам решения подобных задач (т.е. к методам обработки сигнала в режиме скользящего кадра фиксированной длины). Для очистки сигнала используется критерий минимума квадрата евклидова расстояния между наблюдаемым вектором и вектором помехи на множестве всевозможных векторов, компоненты которых образуют квазипериодические импульсные последовательности. Минимизация квадрата расстояния сведена к минимизации аддитивного функционала, которая осуществляется методом динамического программирования. В работе приведены формулы пошаговой оптимизации, составляющие сущность вычислительного алгоритма, учитывающие специфику решаемой задачи и обеспечивающие очистку зашумленного сигнала. Дана оценка трудоемкости алгоритма, увязанная с параметрами задачи. Вместе с вычислительным алгоритмом в работе описаны результаты обработки речевых сигналов.

1. Постановка задачи

Пусть $s(t)$ ($s \in R$, $t \in R$, R — числовая прямая) — функция, описывающая изменение речевого сигнала во времени, а $s_n = s(n\tau_s)$, $n = 0, \pm 1, \pm 2, \dots$, — дискретные значения (отсчеты) этой функции (сигнала), взятые через равные промежутки времени $\tau_s = 1/F_s$ (F_s — частота дискретизации сигнала).

Предположим, что речевой сигнал искажается аддитивной помехой, т.е.

$$y_n = s_n + x_n, \quad n = 0, \pm 1, \pm 2, \dots, \quad (1)$$

где x_n , $n = 0, \pm 1, \pm 2, \dots$, — квазипериодическая (почти периодическая) последовательность вида:

$$x_n = \begin{cases} u_n - n_i, & \text{если } n = n_i, n_i + 1, \dots, n_i + q - 1, \\ 0, & \text{если } n = n_i + q, \dots, n_{i+1} - 1, \end{cases} \quad (2)$$

$$i = 0, \pm 1, \pm 2, \dots$$

В формуле (2) $u_n \in R$, $n = 0, \pm 1, \pm 2, \dots$, — числовая детерминированная последовательность или импульс, обладающий свойством:

$$0 < \sum_{n=0}^{q-1} u_n^2 < \infty; \quad u_n = 0, \quad n < 0, \quad n > q - 1; \quad (3)$$

n_i , $i = 0, \pm 1, \dots$, — почти периодическая последовательность моментов времени начала импульсов такая, что

$$q \leq T_{\min} \leq n_i - n_{i-1} \leq T_{\max} < \infty, \quad (4)$$

где T_{\min} , T_{\max} — минимальный и максимальный интервалы между двумя последовательными импульсами.

Вектор $U = (u_0, u_1, \dots, u_{q-1})$ будем называть эталонным импульсом, а q — длиной импульса. Параметры T_{\min} , T_{\max} , q и u_n , $n = 0, 1, \dots, q - 1$, будем считать известными.

Пусть наблюдаемый вектор $Y = (y_0, y_1, \dots, y_{N-1})$ есть сумма двух независимых векторов:

$$Y = S + X, \quad (5)$$

где $S = (s_0, s_1, \dots, s_{N-1})$ — вектор, последовательность компонент которого соответствует N расположенным друг за другом отсчетам речевого сигнала, а компоненты вектора $X = (x_0, x_1, \dots, x_{N-1})$ соответствуют N отсчетам квазипериодической помехи. Вектор Y будем интерпретировать как участок (анализируемый кадр) зашумленного сигнала длительности Nt_s . Учитывая (2), будем считать, что компоненты вектора X связаны формулой

$$x_n = \sum_{i=1}^M u_n - n_i; \quad n = 0, 1, \dots, N - 1; \quad (n_1, \dots, n_M) \in \Omega, \quad (6)$$

где

$$\begin{aligned} \Omega &:= \Omega(N, T_{\min}, T_{\max}, q) = \\ &= \left\{ (n_1, \dots, n_M); N \geq T_{\max} + q - 1; n_1 \geq 1 - q; n_M \leq N - 1; \right. \\ &\quad \left. q \leq T_{\min} \leq n_i - n_{i-1} \leq T_{\max} < \infty, i = 2, \dots, M \right\} - \quad (7) \end{aligned}$$

множество всевозможных наборов моментов времени начала M импульсов. Элементы данного множества полностью определяются указанный в (7) совокупностью ограничений. Причем неизвестное число M импульсов лежит в интервале $[M_{\min}, M_{\max}]$, границы которого являются функциями от N, T_{\min}, T_{\max}, q и задают минимальное и максимальное число импульсов.

Образуем последовательность $\mathbf{Y}(k) = (y_0(k), \dots, y_{N-1}(k))$, $k = 0, 1, 2, \dots$, векторов, где k обозначает момент времени, в который наблюдается вектор \mathbf{Y} . Вектор $\mathbf{Y}(k)$ будем рассматривать как участок или анализируемый кадр зашумленного сигнала, выделенный в момент времени k .

Пусть выделенный в k -й момент времени, $k = 0, 1, \dots$, кадр $\mathbf{Y}(k) = (y_0(k), y_1(k), \dots, y_{N-1}(k))$ зашумленного сигнала содержит неизвестное число $M \geq 1$ импульсов, среди которых имеется хотя бы один полный импульс, т.е. импульс фиксированной длины q , не разбитый на части левой или/и правой границами кадра. Задача состоит в оценивании последовательности векторов $\mathbf{S}(k)$ по наблюдаемой последовательности векторов $\mathbf{Y}(k)$, $k = 0, 1, \dots$

2. Критерий решения задачи

Опустив для простоты зависимость векторов от k , рассмотрим критерий для одного анализируемого кадра. Оценку вектора \mathbf{S} будем искать в виде $\hat{\mathbf{S}} = \mathbf{Y} - \hat{\mathbf{X}}$, где $\hat{\mathbf{X}}$ — оценка вектора помехи.

Обозначим через

$$\Theta = \left\{ \mathbf{X} : x_n = \sum_{i=1}^M u_n - n_i, n = 0, 1, \dots, N-1, (n_1, \dots, n_M) \in \Omega \right\}$$

множество всевозможных векторов, компоненты которых образуют квазипериодические последовательности. Как показано в [10],

мощность множества Ω конечна и каждому элементу (n_1, \dots, n_M) данного множества однозначно соответствует вектор \mathbf{X} . Поэтому множество Θ также конечно.

Воспользовавшись критерием минимума квадрата евклидова расстояния, найдем ближайший к вектору \mathbf{Y} вектор $\hat{\mathbf{X}}$:

$$\hat{\mathbf{X}} = \underset{\mathbf{X} \in \Theta}{\text{Arg min}} \|\mathbf{Y} - \mathbf{X}\|^2. \quad (8)$$

Тогда, в соответствии с (5), $\hat{\mathbf{S}} = \mathbf{Y} - \hat{\mathbf{X}}$.

Обозначим $Q(\mathbf{X}) = \|\mathbf{Y} - \mathbf{X}\|^2$ и заметим, что

$$Q(\mathbf{X}) = \sum_{n=0}^{N-1} \left(y_n - \sum_{i=1}^M u_{n-n_i} \right)^2 = Q(n_1, \dots, n_M/U), \quad (9)$$

$$(n_1, \dots, n_M) \in \Omega.$$

Согласно (8) для нахождения вектора $\hat{\mathbf{X}}$ достаточно минимизировать (9) на множестве Ω . Найденному оптимальному набору

$(\hat{n}_1, \dots, \hat{n}_{\hat{M}}; \hat{M}) = \underset{\Omega}{\text{Arg min}} Q(n_1, \dots, n_M/U)$ однозначно соответствует последовательность компонент вектора $\hat{\mathbf{X}}$:

$$\hat{x}_n = \sum_{i=1}^{\hat{M}} u_{n-\hat{n}_i}, \quad n = 0, 1, \dots, N-1; (\hat{n}_1, \dots, \hat{n}_{\hat{M}}) \in \Omega.$$

3. Алгоритм очистки

1. *Ядро алгоритма.* Обозначив

$$A(n_i) = \sum_{n=0}^{N-1} u_{n-n_i} [u_{n-n_i} - 2y_n], \quad i = 1, \dots, M, \quad (10)$$

сумму (9) можно представить в виде:

$$Q(n_1, \dots, n_M) = \sum_{n=0}^{N-1} y_n^2 + \sum_{i=1}^M A(n_i), \quad (n_1, \dots, n_M) \in \Omega. \quad (11)$$

Из (11) видно, что задача получения оценок $(\hat{n}_1, \dots, \hat{n}_M)$ сводится к минимизации аддитивного функционала

$$D(n_1, \dots, n_M) = \sum_{i=1}^M A(n_i), \quad (n_1, \dots, n_M) \in \Omega, \quad (12)$$

на семействе Ω всевозможных наборов моментов времени начала импульсов с ограничениями на целочисленные переменные n_i , указанные в формуле (7).

Следуя [9], для решения экстремальной задачи (12) при заданных ограничениях (7), воспользуемся методом динамического программирования. Для каждого $n \in [-q + 1, N + T_{\min} - 1]$ определим множество $\Gamma(n) = \{m : \max[-q + 1, n - T_{\max}] \leq m \leq \min[-q + 1, n - T_{\min}]\}$ и, в соответствии с принципом оптимальности, организуем многошаговый процесс минимизации в виде:

$$\left. \begin{aligned} D_n &= \min_{m \in \Gamma(n)} \{D_m + A(m)\}, \\ I(n) &= \arg \min_{m \in \Gamma(n)} \{D_m + A(m)\}, \\ n &= -q + T_{\min}, \dots, N + T_{\min} - 1, \end{aligned} \right\} \quad (13)$$

$$D_N = D_{\min} = \min_{N \leq n \leq N + T_{\min} - 1} D_n, \quad (14)$$

где через D_n и $I(n)$ обозначены минимальное значение функционала и указатель минимума на n -м шаге. Начальными условиями для вычислений по формулам (13) и (14) являются нулевые значения:

$$D_n = 0, \quad I(n) = 0, \quad n = -q, -q + 1, \dots, T_{\min} - q - 1, \quad (15)$$

а также значения $A(n)$, $n = -q, -q + 1, \dots, N$, подсчитанные по формуле (10) перед началом оптимизации.

Число импульсов и их расположение в анализируемом кадре определяются рекуррентными вычислениями в обратном порядке по указателю $I(n)$ оптимального пути:

$$\left. \begin{aligned} m_0 &= \arg \min_{N \leq n \leq N + T_{\min} - 1} D_n, \\ m_i &= I(m_{i-1}), \quad i = 1, 2, \dots \end{aligned} \right\} \quad (16)$$

Вычисления заканчиваются при таком шаге $i = r$, что $I(m_r) = 0$. В результате получаем последовательность m_r, m_{r-1}, \dots, m_1 такую, что $(\hat{n}_1, \hat{n}_2, \dots, \hat{n}_{\hat{M}-1}, \hat{n}_{\hat{M}}) = (m_r, m_{r-1}, \dots, m_1)$. Величина r дает оценку \hat{M} числа импульсов (включая неполные), попавших в данный анализируемый кадр. При этом, если $\hat{n}_1 \in \{-q + 1, \dots, -1\}$, то в начале кадра имеется неполный импульс. Если $\hat{n}_{\hat{M}} \in \{N - q + 1, \dots, N - 1\}$, то неполный импульс имеется в конце кадра.

Таким образом, для заданного импульса U получение оценок моментов времени начала импульсов и оценок компонент вектора помехи обеспечивают вычисления по формулам (10) и (13)–(16). Оценка компонент вектора очищенного сигнала находится в соответствии с (5): $\hat{s}_n = y_n - \hat{x}_n$, $n = 0, 1, \dots, N - 1$.

2. *Длина кадра.* Обозначим через L длину всего наблюдаемого временного ряда. При обработке временного ряда в режиме скользящего кадра фиксированной длины на N — длину анализируемого кадра — накладываются следующие ограничения:

1) кадр анализа должен содержать хотя бы один полный импульс, что гарантируется, как показано в [10], соблюдением условия $N \geq T_{\max} + q - 1$, заданного формулой (7);

2) длина кадра должна быть не больше длины наблюдаемого временного ряда, т.е. $N \leq L$;

3) длина кадра выбирается с учетом имеющихся в распоряжении ресурсов памяти.

3. *Правило сдвига анализируемого кадра.* Пусть в анализируемом кадре $Y(k)$ длины N , где k — текущий момент времени наблюдения временного ряда, получена последовательность моментов времени начала импульсов $(\hat{n}_1(k), \dots, \hat{n}_{\hat{M}(k)}(k))$, тогда следующий m -й момент времени наблюдения временного ряда определяется по правилу:

$$m = \begin{cases} \hat{n}_{\hat{M}(k)}(k) + T_{\min}, & \hat{n}_{\hat{M}(k)}(k) < N - q + 1 + k, \\ \hat{n}_{\hat{M}(k)-1}(k) + T_{\min}, & \hat{n}_{\hat{M}(k)}(k) \geq N - q + 1 + k. \end{cases} \quad (17)$$

Правило (17) предписывает в следующем кадре $Y(m)$ повторить обработку неполного импульса, если такой импульс будет обнару-

жен в конце текущего кадра. Если обозначить через $\nabla k = m - k$ величину сдвига анализируемого кадра, то из (17) видно, что в зависимости от величины сдвига ∇k кадры могут как накладываться друг на друга, так и не пересекаться.

4. *Стыковка результатов обработки соседних кадров.* При обработке наблюдаемого временного ряда в режиме скользящего кадра фиксированной длины в зависимости от стыковки результатов обработки соседних кадров, некоторые варианты которой описаны ниже, возможно как сохранение, так и нарушение условия квазипериодичности для получаемой последовательности оценок моментов времени начала импульсов. Алгоритмически реализованы три варианта стыковки результатов обработки соседних кадров.

Вариант 1. В следующем анализируемом кадре $Y(m)$, $m > 0$, где m определяется по правилу (17), момент времени начала первого импульса ищется в интервале $[m, m + T_{\max} - T_{\min}]$, т.е. в зависимости от момента начала последнего полного импульса, найденного в текущем кадре $Y(k)$. Данный вариант обеспечивает соблюдение условия квазипериодичности для полученной последовательности оценок моментов времени начала импульсов на всем наблюдаемом временном ряде.

Вариант 2. Момент времени начала первого импульса в следующем анализируемом кадре ищется в интервале $[m, m + T_{\max} - q]$, что не гарантирует выполнения условия квазипериодичности при стыковке результатов обработки соседних кадров.

Для приведенных вариантов стыковки первый импульс анализируемого кадра $Y(k)$, $k > 0$, предполагается полным в соответствии с правилом сдвига (17).

Вариант 3. Каждый кадр анализируется независимо от предыдущего, что, как и во втором варианте, может нарушить условие квазипериодичности для полученной последовательности оценок моментов времени начала импульсов на всем наблюдаемом временном ряде. Однако, данный вариант позволяет избежать возможного "накопления" ошибки оценивания моментов времени начала импульсов вследствие зависимости результатов обработки текущего кадра от результатов обработки предыдущего кадра,

что свойственно для первых двух вариантов при неправильном оценивании моментов времени начала импульсов.

5. *Обработка "хвоста" наблюдаемого временного ряда.* В конце временного ряда может возникнуть ситуация, когда последний анализируемый кадр $Y(k)$ будет иметь длину $N' = L - k$ меньше N , что в случае $N' < T_{\max} + q - 1$ противоречит требованиям, предъявляемым к длине кадра.

Назовем наблюдаемые данные концом (или "хвостом") временного ряда, если $k + \nabla k + N \geq L$, где k — текущий момент времени наблюдения временного ряда, $\nabla k = N - q + T_{\min}$ — максимально возможный сдвиг кадра, определяемый правилом сдвига. И перед обработкой текущего кадра будем проверять данное условие для определения конца наблюдаемого временного ряда.

Можно предложить различные варианты обработки конца наблюдаемого временного ряда. В частности, при возможности увеличения значения N момент времени наблюдения k можно считать началом последнего анализируемого кадра $Y(k)$ длины $N' = L - k$, $N' > N$. При $N \gg T_{\max} + q - 1$ можно обрабатывать "хвост" последовательности кадрами длины $N' \in [T_{\max} + q - 1, N]$.

Алгоритмически реализован вариант, исключаяющий изменение длины N анализируемого кадра. Пусть k — текущий момент времени наблюдения временного ряда, тогда если $k + N > L$, то устанавливается новый момент времени $k' = L - N$ наблюдения временного ряда. При этом последний кадр анализируется независимо от предыдущего. Оценка компонент вектора очищенного сигнала для участка $[k', k - 1]$ повторно не проводится (берется оценка, полученная в предыдущем кадре). Данный подход к обработке "хвоста" наблюдаемого ряда не гарантирует выполнения условия квазипериодичности при стыковке результатов обработки двух последних кадров.

6. *Трудоёмкость алгоритма.* Опираясь на результаты работы [9], трудоёмкость ядра алгоритма можно оценить как величину $\sim [N(q + T_{\max} - T_{\min})]$. Оценивая затраты по времени на обработку последнего кадра так же как и затраты на обработку остальных кадров, общую трудоёмкость обработки наблюдаемого временного ряда можно оценить как величину $\sim [KN(q + T_{\max} - T_{\min})]$, где K — число анализируемых кадров.

Для предлагаемых вариантов обработки "хвоста" наблюдаемого временного ряда можно получить оценки числа K обрабатываемых кадров, в частности, для варианта, запрещающего изменение длины N анализируемого кадра, K удовлетворяет очевидным неравенствам:

$$\left] \frac{L - N}{T_{\min}} \left[+ 1 \leq K \leq \left] \frac{L - N}{N - q + T_{\min}} \left[+ 2,$$

так как, согласно (17), $T_{\min} \leq \nabla k \leq N - q + T_{\min}$. Через $] \times [$ обозначена целая часть x .

4. Экспериментальные результаты

Описанный алгоритм реализован программно и опробован при выделении речевого сигнала из смеси с аддитивной квазипериодической помехой. Обработывались речевые сигналы, введенные в компьютер через динамический микрофон и восьмиразрядный аналого-цифровой преобразователь при частоте дискретизации 10 кГц. Искаженные сигналы формировались программным образом путем наложения сгенерированных квазипериодических импульсных помех на введенный в компьютер речевой сигнал. Моделировались искажения сигнала импульсами различной формы.

Для получения объективной оценки точности алгоритма в зависимости от параметров q , T_{\min} , T_{\max} , N и формы импульса помехи анализировались следующие характеристики:

$$\alpha_1 = \frac{\|X - \hat{X}\|^2}{\|X\|^2}, \quad \alpha_2 = \frac{\|S - \hat{S}\|^2}{\|S\|^2}.$$

Первая — α_1 — характеризует отношение квадратичной ошибки оценивания помехи к энергии помехи, вторая — α_2 — отношение квадратичной ошибки оценивания сигнала к энергии сигнала. Кроме того проводилась субъективная (экспертная) оценка эффективности алгоритма, т.е. анализ качества восстановленного речевого сигнала на слух.

Каждая точка приведенных ниже графиков кривых нормированных ошибок α_1 и α_2 получена путем усреднения результатов 100 экспериментов. При этом каждый из экспериментов

состоял в очистке речевого сигнала (одной и той же фразы), искаженного случайным образом одной из всевозможных квазипериодических последовательностей.

Были опробованы апостериорный способ обработки наблюдений (т.е. оценивание по всему наблюдаемому временному ряду) и способ обработки сигнала в режиме скользящего кадра фиксированной длины (для трех описанных выше вариантов стыковки результатов обработки соседних кадров). Результаты очистки речевого сигнала различными способами в среднем (по 100 экспериментам) практически совпадают.

На приведенных ниже иллюстрациях представлены результаты экспериментов по очистке сигнала в режиме скользящего кадра фиксированной длины с использованием первого варианта стыковки данных, полученных в соседних кадрах. По оси абсцисс отложено среднее (по 100 экспериментам) значение отношения сигнал/шум: $\eta = 10 \log_{10} \frac{\|S\|^2}{\|X\|^2}$.

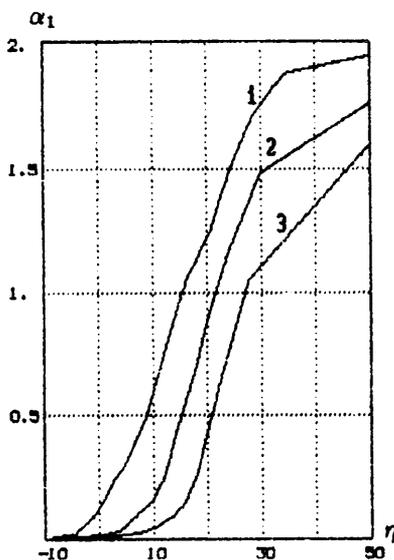


Рис. 1

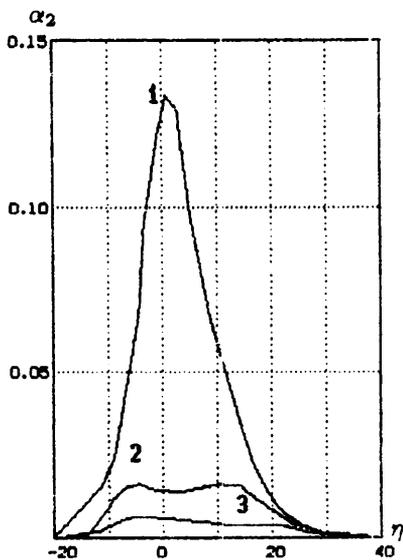


Рис. 2

На рис.1 и 2 представлены экспериментальные кривые нормированных ошибок α_1 и α_2 , полученные для смеси речевого сигнала с аддитивными квазипериодическими помехами при фиксированных $T_{\min} = 60$, $T_{\max} = 100$, $N = 1000$ и трех значениях q : $q = 10$ (кривая 1), $q = 30$ (кривая 2), $q = 50$ (кривая 3). Искажающим импульсом служил один период синусоиды. Рисунки иллюстрируют увеличение значений α_1 и α_2 с уменьшением длины импульса.

По рис.1 видно, что относительная ошибка α_1 оценивания помехи увеличивается при возрастании отношения сигнал/шум. Относительная же ошибка α_2 оценивания сигнала (см. рис.2) имеет максимум в области точки $\eta = 0$, т.е. в том случае, когда энергии сигнала и помехи близки. Подобное поведение α_2 можно объяснить следующим образом. С уменьшением величины η ($\eta < 0$ дБ), т.е. с увеличением мощности помехи, α_2 стремится к нулю за счет более точного оценивания помехи (см. рис.1). С увеличением отношения сигнал/шум ($\eta > 0$ дБ), несмотря на неточное выделение помехи, α_2 стремится к нулю за счет уменьшения мощности помехи (поскольку $\|S - \hat{S}\|^2 = \|X - \hat{X}\|^2$).

На рис.3 и 4 представлены графики, поясняющие зависимость результатов очистки от формы искажающего импульса: прямоугольный импульс (кривая 1), один период синусоиды (кривая 2), затухающая синусоида (кривая 3); значения параметров задачи: $q = 30$, $T_{\min} = 60$, $T_{\max} = 100$, $N = 1000$. При фиксированном значении η искажающие импульсы различной формы имеют одинаковую мощность, но отличающиеся амплитуды: максимальная амплитуда — у затухающей синусоиды, минимальная — у прямоугольника.

На рис.4 можно заметить три области различного взаимного расположения экспериментальных кривых относительной ошибки α_2 оценивания сигнала в зависимости от формы импульса помехи: 1) область малых отрицательных значений η ; 2) "окрестность" нуля; 3) область больших положительных значений η . В увеличенном масштабе те же три области различного взаимного расположения экспериментальных кривых относительной ошибки α_1 оценивания помехи выделяются и на рис.3.

В первой области наиболее точно детектируются импульсные помехи с наибольшей амплитудой. Поэтому в этой области речевой сигнал лучше очищается от импульсных помех в виде затухающих синусоид.

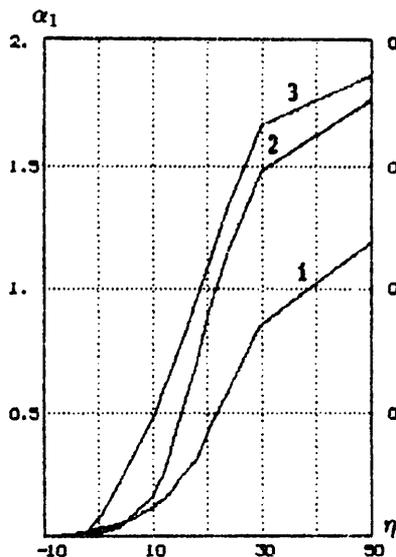


Рис. 3

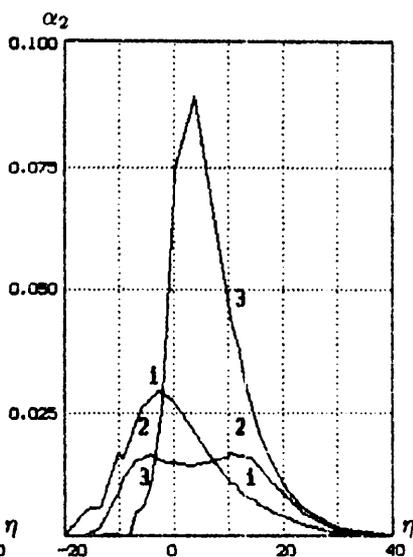


Рис. 4

В третьей области точнее детектируются квазипериодические последовательности наименее похожие на речевой сигнал, представляющий на отдельных участках сумму затухающих синусоид. Поэтому в этой области речевой сигнал лучше очищается от прямоугольных импульсов.

Во второй области наблюдается "переходный" процесс от одного взаимного расположения экспериментальных кривых нормированных ошибок α_1 и α_2 к другому. Характерный для данной области экстремум на графиках α_2 для различных форм импульсов помехи отражает тот факт, что при мощности сигнала близкой к мощности помехи алгоритм очистки работает наименее точно. Для уточнения взаимного расположения кривых относи-

тельных ошибок α_1 и α_2 в данной области требуются дополнительные эксперименты.

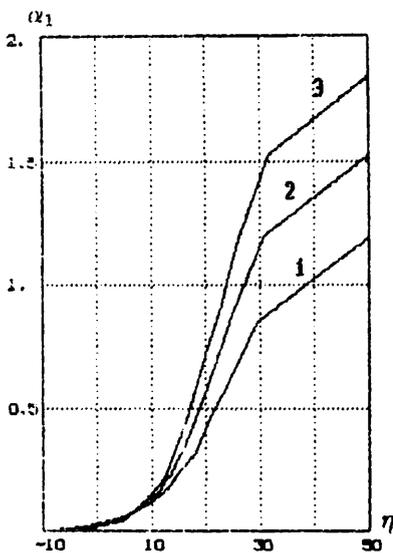


Рис. 5

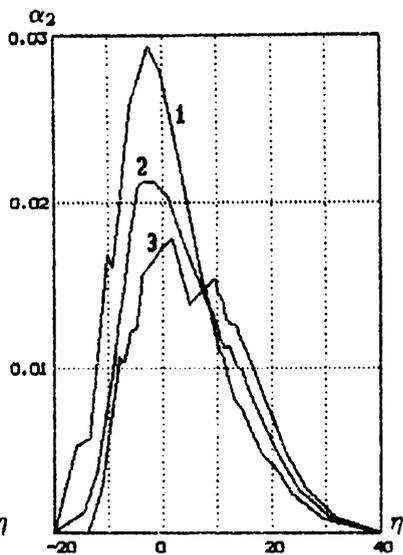


Рис. 6

На рис. 5 и 6 представлены графики, поясняющие зависимость результатов обработки от T_{\max} при фиксированных $q = 30$, $T_{\min} = 60$, $N = 1000$. Искажающим импульсом служил прямоугольник. Значения T_{\max} менялись следующим образом: $T_{\max} = 100$ (кривая 1), $T_{\max} = 150$ (кривая 2) и $T_{\max} = 200$ (кривая 3).

Перед анализом поведения экспериментальных кривых заметим, что при фиксированном значении η для различных значений T_{\max} мощности искажающих импульсов будут отличаться, а именно: большему значению T_{\max} соответствует большая мощность искажающего импульса.

Поэтому в области доминирования помехи (при малых отрицательных значениях η) как помеха, так и сигнал выделяются точнее при большем значении T_{\max} . Когда сигнал доминирует

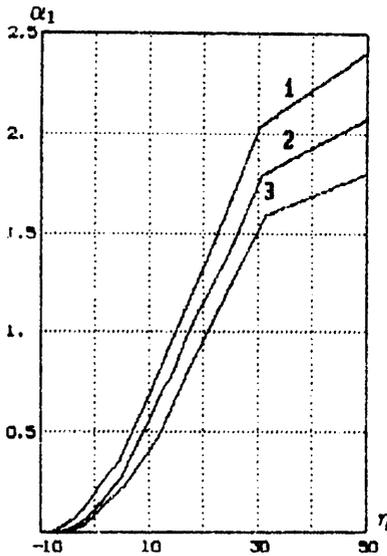


Рис. 7

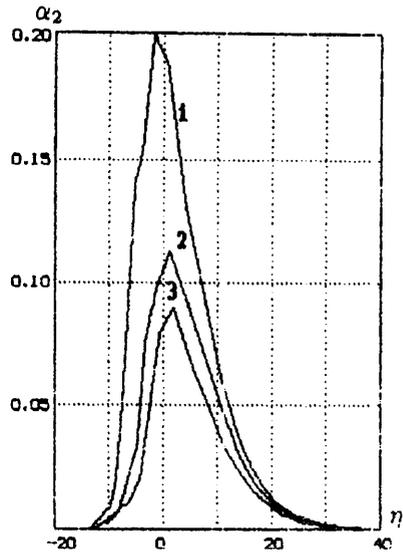


Рис. 8

над помехой (при больших положительных значениях η) детектирование последней становится практически невозможным, и поэтому для фиксированного значения η , лежащего в данной области, значения $\|S - \hat{S}\|^2$ и $\|X - \hat{X}\|^2$ тем больше, чем больше мощность искажающего импульса, т.е. помеха выделяется хуже при большем значении T_{\max} . Этим можно объяснить "перекручивание" экспериментальных кривых 1, 2 и 3, четко выраженное на рис.6. В увеличенном масштабе "перекручивание" кривых заметно и на рис.5.

На рис.7 и 8 представлены экспериментальные кривые нормированных ошибок α_1 и α_2 , полученные для смеси речевого сигнала с аддитивными квазипериодическими последовательностями импульсов в виде затухающей синусоиды при фиксированных $q = 40$, $T_{\max} = 180$, $N = 1000$ и трех значениях T_{\min} : $T_{\min} = 50$ (кривая 1), $T_{\min} = 80$ (кривая 2), $T_{\min} = 120$ (кривая 3). Рисунки иллюстрируют увеличение значений α_1 и α_2 с уменьшением значения T_{\min} .

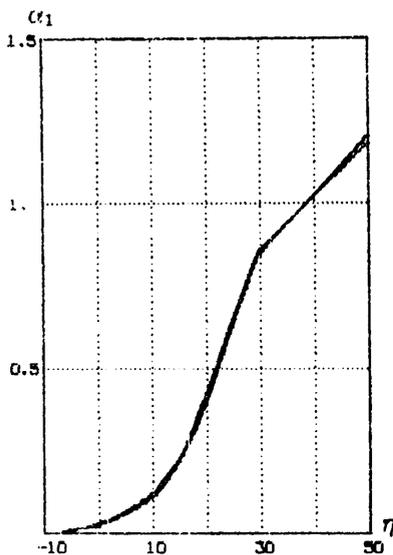


Рис. 9

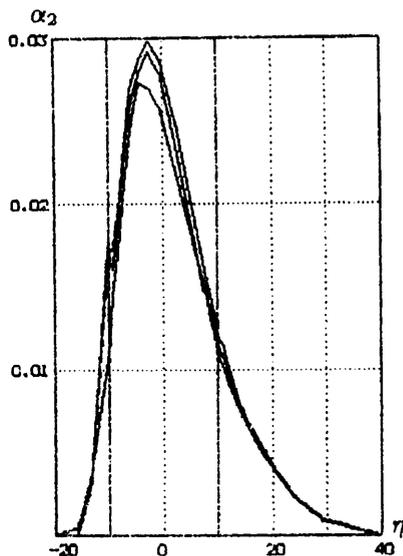


Рис. 10

На рис.9 и 10 совмещены экспериментальные данные, характеризующие зависимость результатов очистки от длины кадра N ($N = 150, 500, 4000$) при фиксированных $q = 30$, $T_{\min} = 60$, $T_{\max} = 100$. Искажающим импульсом служил прямоугольник. Результаты эксперимента подчеркивают тот факт, что α_1 и α_2 практически не зависят от длины кадра и, таким образом, учитывая приведенные выше оценки трудоемкости алгоритма, можно уменьшить время обработки речевых сигналов за счет оптимального выбора длины кадра.

Субъективная оценка эффективности работы алгоритма была дана пятью экспертами при прослушивании фразы "Вода в луже медленно убывала". Оказалось, что в очищенном сигнале помеха наиболее слышна, когда отношение сигнал/шум лежит в интервале $(-10, 10)$ дБ (максимум слышимости помехи — при значении η близком к нулю), т.е. при мощности помехи близкой к мощности сигнала. Эти данные хорошо согласуются с экспериментальными кривыми нормированных ошибок α_2 (рис.2,4,6,8,10), имею-

ших экстремум в данном интервале. С уменьшением величины η ($\eta < -10$ дБ) качество выделенного сигнала, по мнению экспертов, улучшается. Соответственно ведут себя и экспериментальные кривые нормированных ошибок α_1 (см. рис.1,3,5,7,9): α_1 стремится к нулю, поскольку помеха выделяется более точно с увеличением своей мощности. Так как $\|S - \hat{S}\|^2 = \|X - \hat{X}\|^2$, то и α_2 также стремится к нулю. С увеличением значения отношения сигнал/шум ($\eta > 10$ дБ) при прослушивании экспертами качество выделенного полезного сигнала улучшается. В частности, при $\eta > 30$ дБ помеха не слышна как в зашумленном, так и в выделенном речевом сигнале. С ростом значения η помеха выделяется хуже, однако, с уменьшением мощности помехи она становится менее ощутимой на слух. Соответственно ведут себя и экспериментальные кривые нормированных ошибок: α_1 увеличивается, α_2 стремится к нулю.

Таким образом, введенные характеристики α_1 и α_2 довольно хорошо соотносятся с субъективным восприятием речи человеком.

З а к л ю ч е н и е

В работе предложен новый алгоритм выделения или детектирования полезного (речевого) сигнала из смеси с аддитивной квазипериодической помехой. Разработанный алгоритм опирается на апостериорно-последовательный подход к обработке речевых сигналов, т.е. на обработку в режиме скользящего по сигналу анализируемого кадра фиксированной длины, при этом алгоритмически реализованы три варианта стыковки результатов обработки соседних кадров.

Для очистки сигнала в каждом последовательном кадре используется критерий минимума квадрата евклидова расстояния между наблюдаемым вектором и вектором помехи на множестве всевозможных векторов, компоненты которых образуют квазипериодические импульсные последовательности. Данный критерий является оптимальным, если речевой сигнал рассматривать как некоррелированный гауссовский процесс с нулевым математическим ожиданием и конечной дисперсией. Несмотря на то, что подобная модель речевого сигнала является довольно упрощен-

ной, она вполне приемлема (с практической точки зрения) для рассматриваемой задачи.

В работе показано, что минимизация квадрата расстояния эквивалентна нахождению экстремума аддитивной целевой функции с ограничениями в виде линейных неравенств и, таким образом, подавление квазипериодической помехи может осуществляться методом динамического программирования. Выведены рекуррентные формулы, обеспечивающие нахождение глобального экстремума, учитывающие специфику решаемой задачи и составляющие сущность конструктивного вычислительного алгоритма. Приведены оценки трудоемкости алгоритма, увязанные с параметрами задачи.

Работоспособность алгоритма подтверждена результатами численного моделирования и обработкой зашумленных речевых сигналов. Алгоритм позволяет восстанавливать речевые сигналы полностью утратившие свою разборчивость из-за наличия аддитивной квазипериодической помехи и, кроме того, может использоваться для скрытой передачи речи по каналам связи.

Л и т е р а т у р а

1. САПОЖКОВ М.А. Речевой сигнал в кибернетике и связи. - М.: Связьиздат, 1963. - 452 с.
2. РАБИНЕР Л.Р., ШАФЕР Р.В. Цифровая обработка речевых сигналов. - М.: Радио и связь, 1981. - 495 с.
3. ЛИМ Дж.С., ОППЕНХАЙМ А.В. Коррекция и сжатие спектра зашумленных речевых сигналов // ТИИЭР. - 1979. - Т. 67, № 12. - С.5-26.
4. PARSONS T.W. Separation of speech from interfering speech by means of harmonic selection // J. Acoust.Soc.Amer. - 1976. - Vol.60, Oct. - P.911-918.
5. SAMBUR M.R. Adaptive noise cancelling for speech signals // IEEE Trans.Acoust.,Speech,Signal Proc. - 1978. - Vol.ASSP-26, Oct. - P.419-423.
6. KAUNITZ J. Adaptive filtering of broadband signals as applied to noise cancelling.- Stanford Electronics Lab., Stanford Univ., Stanford Calif., Rep. SU-SEL-72-038, Aug. 1972 (Ph.D.dissertation).

7. Адаптивные компенсаторы помех. Принципы построения и применения / Уидроу, Гловер, Маккул и др.// ТИИЭР.- 1975.- Т.63, № 12. - С.69-98.

8. ТРИФОНОВ А.П., ШИНАКОВ Ю.С. Совместное различение сигналов и оценка их параметров на фоне помех. - М.: Радио и связь, 1986. - 264 с.

9. КЕЛЬМАНОВ А.В., КУТНЕНКО О.А. Алгоритм распознавания квазипериодической последовательности импульсов и обнаружения моментов их начала в гауссовском шуме// Методы обработки информации. - Новосибирск, 1996. - Вып. 157: Вычислительные системы. - С.137-180.

10. КЕЛЬМАНОВ А.В., КУТНЕНКО О.А. Нижняя граница вероятности ошибки распознавания квазипериодической последовательности импульсов в гауссовском белом шуме// Настоящий сборник. - С. 55-77.

Поступила в редакцию
3 июля 1998 года