ЭКСПЕРТНЫЕ СИСТЕМЫ И АНАЛИЗ ДАННЫХ
(Вычислительные системы)

1991 год Выпуск 144

УДК 621.31:534.4

РЕШЕНИЕ ПРОБЛЕМЫ ВЫДЕЛЕНИЯ СПЕКТРАЛЬНО-КОРРЕЛЯЦИОННЫХ ПРИЗНАКОВ РЕЧЕВОГО СИГНАЛА, ИНВАРИАНТНЫХ К НЕЛИНЕЙНЫМ АМПЛИТУДНЫМ ИСКАЖЕНИЯМ

А.В. Кельманов

Введение

В [1] приведено решение проблемы компенсации нелинейных амплитудных искажений речевого сигнала по оценкам ковариаций при известной (но не всегда обратимой) искажающей функции. Однако компенсация как средство борьбы с искажениями в речевых трактах связи на практике не всегда осуществима из-за отсутствия необходимой информации об искажении. Типичным примером является ситуация, когда об искажении известно лишь то, что оно принадлежит некоторому классу, "выбирается" из этого класса случайно и оценить его не представляется возможным. В этих условиях единственным способом борьбы с искажениями является спектрально-корреляционное оценивание, устойчивое или инвариантное ко всем допустимым искажениям из заданного класса.

Целью данной работы является разработка, обоснование и исследование метода спектрально-корреляционного оценивания, инвариантного к искажениям, характерным для речевых трактов связи. По существу, в работе вводится такое инвариантное к классу амплитудных искажений преобразование, которое приводит к инвариантности оценок корреляций и нормированной спектральной плотности. При этом доказывается, что полученные оценки сходятся в среднеквадратическом к теоретическим корреляциям и нормированной спектральной плотности неискаженного речевого сигнала. Сначала решение задачи проводится для стационарных случайных процессов. Затем результаты интерпретируются на примере гауссовских процессов и, наконец, показывается, каким образом полученные результаты переносятся на обработку речевых сигналов. При этом решение задачи увязывается с проблемой ограниченности выборки. Работа опирается на материалы, приведенные в [1].

1. Постановка задачи

Пусть $x(t)$, $-\infty < t < \infty$, - непрерывный речевой сигнал, а $x_n = x(nT)$, $n = 0, \pm1, \pm2, \ldots$, - дискретные значения (отсчеты) этого сигнала, взятые через равные промежутки времени T. Будем считать, что на коротких участках длительностью T_a, содержащих N отсчетов, сигнал является реализацией стационарного случайного процесса, который полностью описывается своим одномерным распределением P_x (отличным от гауссовского) и ковариационной функцией. Под амплитудным преобразованием или искажением будем понимать нелинейную зависимость $y(t) = f[x(t)]$ или $y_n = f(x_n)$.

Пусть $\mathcal{M}x_n = M_x$, $n = 0, \pm1, \pm2, \ldots$, - математическое ожидание процесса $\{x_n\}$, а $\sigma_x(m)$ и $\rho_x(m)$ - его ковариационная и корреляционная последовательности:

$$\left.\begin{aligned} &\sigma_x(m) = \mathrm{cov}(x_n, x_{n+m}) = \mathcal{M}(x_n - M_x)(x_{n+m} - M_x), \\ &\rho_x(m) = \sigma_x(m)/\sigma_x(0), \quad m = 0, \pm1, \pm2, \ldots \end{aligned}\right\} \qquad (1)$$

Допустим, что спектральная функция процесса $\{x_n\}$ абсолютно непрерывна, а его спектральная плотность есть $s_x(\lambda)$, $-\pi \le \lambda \le \pi$. Тогда [2]

$$\left.\begin{aligned}
\sigma_x(m) &= \int_{-\pi}^{\pi} \cos\lambda m \, s_x(\lambda) d\lambda = \\
&= \int_{-\pi}^{\pi} e^{j\lambda m} s_x(\lambda) d\lambda, \quad m = 0, \pm 1, \ldots, \\
\rho_x(m) &= \int_{-\pi}^{\pi} \cos\lambda m \bar{s}_x(\lambda) d\lambda = \int_{-\pi}^{\pi} e^{j\lambda m} \bar{s}_x(\lambda) d\lambda,
\end{aligned}\right\} \quad (2)$$

где $\bar{s}_x(\lambda) = s_x(\lambda)/\sigma_x(0)$ - нормированная спектральная плотность.

Как известно [2], если

$$\sum_{m=-\infty}^{\infty} |\sigma_x(m)| = \sigma(0) + 2 \cdot \sum_{m=1}^{\infty} |\sigma_x(m)| < \infty, \quad (3)$$

то спектральная плотность непрерывна и

$$\left.\begin{aligned}
s_x(\lambda) &= \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \sigma_x(m) \cos\lambda m = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \sigma_x(m) e^{j\lambda m}, \\
\bar{s}_x(\lambda) &= \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \rho_x(m) \cos\lambda m = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \rho_x(m) e^{j\lambda m}.
\end{aligned}\right\} \quad (4)$$

Для искаженного (преобразованного) процесса $\{y_n\}$ формулы аналогичны.

Символами $c_{xN}(m)$ и $r_{xN}(m)$ обозначим оценки ковариаций и корреляций процесса $\{x_n\}$ (при известных M_x и $\sigma_x(0)$):

$$c_{xN}(m) = \frac{1}{N} \sum_{n=1}^{N-m} (x_n - M_x)(x_{n+m} - M_x),$$

$$r_{xN}(m) = c_{xN}(m)/\sigma_x(0)\,,\ m = 0,1,\ldots,N-1\,, \tag{5}$$

а символом $\hat{s}_{xN}(\lambda)$ - сглаженную оценку нормированной спектральной плотности:

$$\hat{s}_{xN}(\lambda) = \frac{1}{2\pi}\sum_{n=-m_N}^{m_N} k\left[\frac{n}{m_N}\right] r_x(n)\cos\lambda n\,, \tag{6}$$

где $\{m_N\}$ - некоторая последовательность целых чисел, зависящих от N таким образом, что $m_N \to \infty$, $m_N/N \to 0$ и $m_N^2/N \to \infty$ при $N \to \infty$, а $k(z)$ - функция (корреляционное окно), удовлетворяющая определенным требованиям, которые выпишем позднее. Для процесса $\{y_n\}$ формулы аналогичны. Отметим, что при ограничениях на четвертые моменты процесса $\{x_n\}$ оценки (5) и (6) состоятельны [2]. Ниже индекс N, указывающий на зависимость оценок от объема выборки, опущен.

ОПРЕДЕЛЕНИЕ 1. Преобразование $y_n = f^*(x_n)$ будем называть инвариантным относительно класса искажений $\mathcal{F} = \{f_i\}$, $i = 1,2,\ldots$ (инвариантным к классу искажений), если для всех x_n, $n = 0, \pm1, \pm2, \ldots$, и для всех $f_i \in \mathcal{F}$

$$f^*[f_i(x_n)] = f^*(x_n)\,. \tag{7}$$

ОПРЕДЕЛЕНИЕ 2. Если преобразование $y_n = f(x_n)$, $n = 0, \pm1, \ldots$, порождает зависимость между корреляциями

$$\rho_y(m) = \theta[\rho_x(m)]\,,\quad m = 0,\pm1,\pm2,\ldots, \tag{8}$$

для которой существует однозначная обратная функция θ^{-1} такая, что

$$\rho_x(m) = \theta^{-1}[f_y(m)] = H[\rho_y(m)]\,,\ m = 0,\pm1,\pm2,\ldots, \tag{9}$$

то преобразование $\mathfrak{f}$ будем называть корреляционно обратимым.

ОПРЕДЕЛЕНИЕ 3. Оценки типа (5) и (6) будем называть прямыми, т.е. такими, которые позволяют оценивать корреляции и нормированную спектральную плотность непосредственно наблюдаемого временного ряда. При этом индекс x у этих оценок соответствует оцениванию по исходным наблюдениям, а индекс y - по преобразованным или искаженным.

ОПРЕДЕЛЕНИЕ 4. Если $\mathfrak{f}$ корреляционно обратимо, то оценки подстановки вида

$$\tilde{r}_x(m)=\theta^{-1}[r_y(m)]=H[r_y(m)],\ m=1,2,\ldots,N-1, \tag{10}$$

будем называть обратными оценками корреляций, а оценки вида

$$\hat{\bar{s}}_x(\lambda)=\frac{1}{2\pi}\sum_{n=-m_N}^{m_N} k\left(\frac{n}{m_N}\right)\tilde{r}_x(n)\cos\lambda n,\quad -\pi\le\lambda\le\pi, \tag{11}$$

т.е. оценки нормированной спектральной плотности по обратным оценкам корреляций - компенсационными оценками нормированной спектральной плотности.

Зададимся целью найти такое преобразование $\mathfrak{f}^*$, которое, с одной стороны, было бы инвариантно к классу искажений $\mathfrak{T}$, характерных для речевых трактов связи, с другой, - корреляционно обратимо, с третьей, - позволяло бы получать оценки (10) и (11), сходящиеся в среднеквадратическом к теоретическим величинам $\rho_x(m)$ и $\bar{s}_x(\lambda)$, и, наконец, было бы простым в вычислительном плане.

2. Класс допустимых искажений

Ясно, что вид преобразования $\mathfrak{f}^*$ зависит от класса допустимых искажений. Поэтому, прежде чем перейти к решению задачи, приведем описание этого класса.

По имеющимся данным [3-6] для речевых трактов связи характерным является класс амплитудных искажений, сохраняющих неподвижными точки перехода речевого сигнала через нулевой уровень. Поэтому класс искажений можно задать в виде:

$$\mathcal{F}= \{f: |f(x)| \le \mathrm{const},\ x=0 \Rightarrow f(x)=0,$$
$$x>0 \Rightarrow f(x)>0,\ x<0 \Rightarrow f(x)<0\}. \qquad (12)$$

В этот класс входят, например, присущие трактам связи детерминированные полиномиальные искажения [3,4]:

$$\mathcal{F}_Д = \{f: |f(x)| \le \mathrm{const},\ f(x)=\sum_{i=1}^{k} b_i x^i,$$
$$x>0 \Rightarrow f(x)>0,\ x<0 \Rightarrow f(x)<0\}, \qquad (13)$$

стохастические искажения из класса [5]:

$$\mathcal{F}_C = \{f: |f(x)| \le \mathrm{const},\ x=0 \Rightarrow f(x)=0,$$
$$x>0 \Rightarrow f(x) \in P^+,\ x<0 \Rightarrow f(x) \in P^-\}, \qquad (14)$$

где P^+ и P^- - некоторые искажающие одномерные распределения, отличные от P_x, причем P^+ (P^-) обозначает распределение положительной (отрицательной) случайной величины, а также искажения речевых сигналов, вызванные вибропомехами [6], и некоторые мультипликативные искажения [5].

Как показано в [5,6], амплитудные искажения из перечисленных классов могут сохранять разборчивость речи на приличном уровне, но при этом весьма сильно "портят" признаки - оценки корреляций, параметров авторегрессии и спектральной плотности. При этом среднеквадратическое отклонение искаженных оценок от неискаженных может более, чем на порядок превышать среднеквадратическую ошибку оценивания признаков в отсутствие амплитудных искажений, что приводит к заметному снижению надежности распознавания речи по упомянутым характеристикам.

Будем считать допустимыми любые искажения из класса $\mathcal{F}$.

3. Требования к инвариантному преобразованию

Инвариантное преобразование f^* может быть не единственным. Чтобы сузить произвол в его выборе, введем некоторые естественные ограничения или требования к f^*.

Во-первых, f^* не должно быть вырожденным, т.е. $f^*[f(x_n)] \neq \mathrm{const}$ для всех $f \in \mathcal{F}$ и всех x_n, $n = 0, \pm 1, \pm 2, \ldots$ Иначе не удастся измерить корреляционно-спектральные характеристики преобразованного процесса.

Во-вторых, f^* должно быть корреляционно обратимым, чтобы иметь возможность по корреляциям преобразованного процесса $\{y_n\}$ оценить корреляции исходного процесса $\{x_n\}$. В противном случае введение инвариантного преобразования теряет смысл, поскольку инвариантное преобразование может так исказить исходный процесс, что в преобразованном процессе будет утеряна информация, необходимая для надежного распознавания или качественного восстановления речи. Очевидно, что требование корреляционной обратимости исключает вырожденные преобразования f^*.

В-третьих, естественным требованием к f^* является простота вычислений. Это требование диктуется необходимостью иметь наиболее быстрые алгоритмы цифровой обработки речевых сигналов, чтобы обеспечить запас вычислительных ресурсов для решения задачи распознавания.

Наконец, в-четвертых, главным требованием к f^* является среднеквадратическая сходимость обратных оценок корреляций и компенсационных оценок спектральной плотности к истинным значениям корреляций и нормированной спектральной плотности неискаженного сигнала. Нетрудно понять, что корреляционная обратимость является лишь необходимым условием для сходимости в среднеквадратическом. Как будет показано ниже, для обеспечения сходимости преобразованный процесс $\{y_n\}$ должен удовлетво -

рять вполне определенным требованиям. Эти требования являются неявными ограничениями на преобразование f^*.

Предположим, что инвариантное преобразование f^* найдено. Тогда справедливо следующее легко проверяемое

УТВЕРЖДЕНИЕ 1. *Пусть преобразование* $y_n = f^*(x_n)$ *инвариантно относительно класса искажений* $\mathcal{F}$. *Тогда если существуют величины* $\sigma_y(m)$, $c_y(m)$, *то они также инвариантны к классу искажений* $\mathcal{F}$. *Если, кроме того, преобразование* f^* *корреляционно обратимо и существуют* $\tilde{r}_x(m)$ *и* $\tilde{\tilde{s}}_x(\lambda)$, *то последние инвариантны к классу искажений* $\mathcal{F}$.

4. Сходимость инвариантных оценок

Приведенные ниже теоремы показывают, при каких условиях инвариантные оценки будут сходиться к истинным значениям в среднеквадратическом. Их доказательство аналогично доказательству теорем 4-7 работы [1] и состоит в применении теорем непрерывности и теорем о предельных распределениях статистик [7] к функциям от прямых оценок, свойства которых изложены в [2].

Обозначим через

$$\varkappa_y(m,l,k) = \mathcal{M}(y_n - M_y)(y_{n+m} - M_y)(y_{n+l} - M_y)(y_{n+k} - M_y) -$$

$$- [\sigma_y(m)\sigma_y(l-k) + \sigma_y(l)\sigma_y(m-k) + \sigma_y(k)\sigma_y(m-l)], \quad (15)$$

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

семиинвариант четвертого порядка процесса $\{y_n\}$.

ТЕОРЕМА 1. *Пусть стационарный случайный процесс* $\{x_n\}$, $n = 0, \pm 1, \pm 2, \dots$, *имеющий* $\mathcal{M}x_n^2 < \infty$, *подвергается такому нелинейному корреляционно обратимому преобразованию* $y_n = f^*(x_n)$, *что* $\rho_x(m) = H[\rho_y(m)]$, $m = 0, \pm 1, \pm 2, \dots$, H, *дифференцируема в* *точках* $\rho_y(m)$ *и* $\mathcal{M}y_n^4 < \infty$. *Тогда если*

$$\sum_{m=-\infty}^{\infty} \sigma_y^2(m) < \infty, \qquad \left| \sum_{l=-\infty}^{\infty} \varkappa_y(m,-l,k-l) \right| < \infty ,$$

то при $N \to \infty$ *имеет место сходимость*

$$\sqrt{N}\, (\vec{\tilde{r}}_x - \vec{\rho}_x) \to \vec{\xi}\, H , \tag{16}$$

где $\vec{\tilde{r}}_x = [\tilde{r}_x(1), \dots, \tilde{r}_x(n^*)]$, $\vec{\rho}_x = [\rho_x(1), \dots, \rho_x(n^*)]$, $H = \{h_{mk}\}$ - *матрица производных размера* $n^* \times n^*$, *элементы которой* $h_{mk} = 0$ *при* $m \neq k$ *и* $h_{mk} = H'[\rho_y(m)]$ *при* $k = m$; $\vec{\xi} = [\xi(1), \dots, \xi(n^*)] \in \Phi[0, \Sigma_\xi]$, *где* $\Phi[0, \Sigma_\xi]$ - *нормальное распределение с параметрами* $(0, \Sigma_\xi)$. *При этом предельные ковариации обратных оценок равны*

$$\lim_{N\to\infty} N \operatorname{Cov}[\tilde{r}_x(m), \tilde{r}_x(k)] =$$

$$= H'[\rho_y(m)] H'[\rho_y(k)] \operatorname{Cov}[\xi(k), \xi(m)] =$$

$$= H'[\rho_y(m)] H'[\rho_y(k)] \lim_{N\to\infty} N \operatorname{Cov}[r_y(m), r_y(k)] =$$

$$= H'[\rho_y(m)] H'[\rho_y(k)] \sum_{l=-\infty}^{\infty} [\rho_y(l)\rho_y(l+m-k) +$$

$$+ \rho_y(l-k)\rho_y(l+m) + \frac{1}{\sigma_y^2(0)} \varkappa_y(m,-l,k-l)]. \tag{17}$$

СЛЕДСТВИЕ 1. *Пусть выполнены условия теоремы 1. Тогда если* f^* *инвариантно относительно класса искажений* $\mathcal{F}$, *то оценки* $\tilde{r}_x(m)$, $m = \pm 1, \pm 2, \dots$, *сходятся* *к* $\rho_x(m)$ *в среднеквадратическом и инвариантны к классу искажений* $\mathcal{F}$.

Справедливость следствия 1 вытекает из теоремы 1 и утверждения 1.

Для компенсационных оценок нормированной спектральной плотности справедлива

ТЕОРЕМА 2. *Пусть стационарный случайный процесс* $\{x_n\}$, $n = 0,\pm1,\pm2,\ldots$, *имеющий* $\mathcal{M}x_n^2 < \infty$, *подвергается такому нелинейному корреляционно обратимому преобразованию* $y_n = f(x_n)$, *что* $\rho_x(m) = H[\rho_y(m)]$, $m = 0,\pm1,\pm2,\ldots$, H *дифференцируема в точках* $\rho_y(m)$ *и* $\mathcal{M}y_n^4 < \infty$. *Тогда если*

1) $k(z) = k(-z)$ *- функция, непрерывная на* $[-1, 1]$, $k(0) = 1$, $|k(z)| \le M$ *для некоторого* M *и всех* $|z| \le 1$ *и* $\lim\limits_{z\to0}[1-k(z)]/z^2 = k^* > 0$;

2) $\{m_N\}$ *- некоторая последовательность целых чисел такая, что* $m_N \to \infty$, $m_N/N \to 0$ *и* $m_N^2/N \to \infty$ *при* $N \to \infty$;

3) $\sum\limits_{m=-\infty}^{\infty} |\sigma_x(m)| < \infty$, $\sum\limits_{m=-\infty}^{\infty} |m|^{1/2} |\sigma_y(m)| < \infty$,

$\sum\limits_{m,l,k=-\infty}^{\infty} |\varkappa_y(m,l,k)| < \infty$, *то при* $N \to \infty$ *имеет место сходимость*

$$\sqrt{\frac{N}{m_N}}\,[\,\tilde{\tilde{s}}_x(\lambda) - s_x(\lambda)\,] \to \eta(\lambda) \in \Phi\,[0, \sigma^2_{\eta(\lambda)}], \qquad (18)$$

$$-\pi \le \lambda \le \pi,$$

причем предельные дисперсии и ковариации даются формулами:

$$\lim_{N\to\infty} \frac{N}{m_N} \operatorname{Var}\tilde{\bar{s}}_x(0) = \operatorname{Var}\eta(0) = 2[s^*_y(0)]^2 \int_{-1}^{1} k^2(z)dz, \quad (19)$$

$$\lim_{N\to\infty} \frac{N}{m_N} \operatorname{Var}\tilde{\bar{s}}_x(\pm\pi) = \operatorname{Var}\eta(\pm\pi) =$$

$$= 2[s^*_y(\pi)]^2 \int_{-1}^{1} k^2(z)dz, \quad (20)$$

$$\lim_{N\to\infty} \frac{N}{m_N} \operatorname{Var}\tilde{\bar{s}}_x(\lambda) = \operatorname{Var}\eta(\lambda) = [s^*_y(\lambda)]^2 \int_{-1}^{1} k^2(z)dz, \quad (21)$$

$$\lim_{N\to\infty} \frac{N}{m_N} \operatorname{cov}[\tilde{\bar{s}}_x(\lambda_1), \tilde{\bar{s}}_x(\lambda_2)] =$$

$$= \operatorname{cov}[\eta(\lambda_1), \eta(\lambda_2)] = 0, \ \lambda_1 = \pm\lambda_2 , \quad (22)$$

где

$$s^*_y(\lambda) = H'[0]s_y(\lambda) = H'[0] \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \rho_y(m)\cos\lambda m . \quad (23)$$

Асимптотически нормальным будет и совместное распределение оценок $\tilde{\bar{s}}_x(\lambda_1), \ldots, \tilde{\bar{s}}_x(\lambda_n)$ для любого фиксированного числа значений λ. Условия 1,2 вместе с конечностью второй суммы в условии 3 соответствуют оцениванию с применением наиболее распространенных корреляционных окон $k(z)$ - окон Хэмминга, Хэннига, Даниэля, Блэкмена-Тьюки и Парзена (см.теорему 9.3.3 [2]).

СЛЕДСТВИЕ 2. *Если в условиях теоремы 2* f^* *инвариантно относительно класса искажений* $\mathcal{F}$, *то оценки* $\tilde{\bar{s}}_x(\lambda)$, $-\pi \le \lambda \le \pi$, *сходятся к* $\bar{s}_x(\lambda)$ *в среднеквадратическом и инвариантны к классу искажений* $\mathcal{F}$.

Справедливость следствия 2 следует из теоремы 2 и утверждения 1.

Таким образом, при надлежащем выборе инвариантного преобразования f^* можно получить инвариантные к искажениям из $\mathcal{F}$, асимптотически несмещенные, нормально распределенные оценки $r_x(m)$ и $\tilde{\tilde{s}}_x(\lambda)$, дисперсия которых стремится к 0 при увеличении объема выборки.

Теоремы 1 и 2 и следствия 1, 2 справедливы и для стационарных гауссовских процессов $\{x_n\}$. Приведем полезные для интерпретации результатов аналоги теорем 1,2 для гауссовских процессов, у которых $\rho_x(m) = 0$ и $|m| > m^*$.

ТЕОРЕМА 3. *Пусть стационарный гауссовский процесс* $\{x_n\}$, $n = 0, \pm1, \pm2, \ldots$, *имеющий* $\mathcal{M} x_n^2 < \infty$ *и* $\rho_x(m) = 0$ *при* $|m| > m^*$, *подвергается такому нелинейному корреляционно обратимому преобразованию* $y_n = f^*(x_n)$, *что* $\rho_x(m) = H[\rho_y(m)]$, $m = 0, \pm1, \pm2, \ldots$, H *дифференцируема в точках* $\rho_y(m)$ *и* $\mathcal{M} y_n^2 < \infty$. *Тогда при* $N \to \infty$ *имеют место* (16) *и* (17).

ТЕОРЕМА 4. *Пусть выполнены условия теоремы 3 и условия 1 и 2 теоремы 2. Тогда при* $N \to \infty$ *справедливы* (18)-(20).

Для полного решения проблемы остается указать корреляционно обратимое преобразование, инвариантное к рассмотренному ранее классу амплитудных искажений $\mathcal{F}$ (12), характерных для речевых трактов связи. Это преобразование приводится в следующем параграфе. При этом показывается, что оно корреляционно обратимо для гауссовских процессов. Для речевых сигналов корреляционная обратимость устанавливается эмпирическим путем.

5. Инвариантное преобразование

Рассмотрим нелинейное преобразование (клиппирование):

$$y_n = f^*(x_n) = \begin{cases} 1, & x_n \geq 0, \\ -1, & x_n < 0, \quad n = 0, \pm 1, \pm 2, \ldots \end{cases} \qquad (24)$$

Нетрудно проверить справедливость следующего утверждения.

УТВЕРЖДЕНИЕ 2. *Преобразование (24) инвариантно к классу искажений* $\mathcal{F}$ *(12).*

Дальнейшее решение задачи теперь зависит от того, имеется ли корреляционная обратимость для преобразования (24). При этом от свойств случайного процесса $\{x_n\}$ (а точнее, от вида его двумерной функции распределения) зависит, будет преобразование (24) обратимым или нет.

Проверить наличие корреляционной обратимости можно двумя путями. Если известна двумерная функция распределения случайного процесса, то проверку указанной обратимости можно осуществить теоретически при помощи теоремы Прайса [1]. Ниже это сделано для гауссовских случайных процессов.

В том случае, когда вид двумерной функции распределения неизвестен, можно проверить корреляционную обратимость эмпирическим путем, например так, как это сделано в [9] для речевых сигналов.

УТВЕРЖДЕНИЕ 3. *Для стационарного гауссовского процесса* $\{x_n\}$, $n = 0, \pm 1, \pm 2, \ldots$, *имеющего* $\mathcal{M} x_n = 0$ *и* $\alpha_x(m) = \alpha_x(0)\,\rho_x(m), m = 0, \pm 1, \pm 2, \ldots$, *преобразование (24) корреляционно обратимо.*

Для доказательства заметим, что в соответствии с теоремой Прайса [8] корреляционная функция преобразованного процесса равна

$$\rho_y(m) = \frac{2}{\pi} \arcsin[\rho_x(m)] = \theta[\rho_x(m)], m = 0, \pm 1, \pm 2, \ldots \qquad (25)$$

Поэтому

$$\rho_x(m) = \theta^{-1}[\rho_y(m)] = H[\rho_y(m)] = \sin[\tfrac{\pi}{2}\rho_y(m)], \qquad (26)$$

$$m = 0, \pm 1, \pm 2, \ldots$$

Таким образом, для стационарных гауссовских процессов проблему можно считать решенной. Это следует из теорем 1 и 2 и утверждений 2,3. При этом в теоремах 1 и 2

$$H'[\rho_y(i)] = \cos[\tfrac{\pi}{2}\rho_y(i)] = \sqrt{1-\rho_x^2(i)}, \qquad (27)$$

$$i = \pm 1, \pm 2, \ldots$$

Приведем два полезных для практики следствия.

СЛЕДСТВИЕ 3. *Пусть стационарный гауссовский процесс* $\{x_n\}$, $n = 0, \pm 1, \pm 2, \ldots$, *имеющий* $\mathcal{M}x_n^2 < \infty$ *и* $\rho_x(m) = 0$ *при* $|m| > m^*$, *подвергается преобразованию (24). Тогда оценки* $\tilde{r}_x(m)$ *инвариантны к искажениям из класса* $\mathcal{F}$ *(12) и при* $N \to \infty$ *имеют место (16) и (17).*

СЛЕДСТВИЕ 4. *Пусть выполнены условия следствия 3 и условия 1,2 теоремы 2. Тогда оценки* $\tilde{\tilde{s}}_x(\lambda)$ *инвариантны к искажениям из класса* $\mathcal{F}$ *(12) и при* $N \to \infty$ *имеют место (18)-(22).*

Остается показать, что для речевых сигналов преобразование (24) корреляционно обратимо, и установить достаточность традиционного при цифровой обработке речи объема выборки N (числа отсчетов в окне анализа) для получения удовлетворительной точности.

6. Экспериментальные результаты

Как показано в [1,9], для речевых сигналов преобразование (24) является корреляционно обратимым. Более того, зависимость между корреляциями хорошо аппроксимируется формулой (26), справедливой для стационарных гауссовских процессов. При этом для получения удовлетворительной точности оценок при обработке гауссовских процессов и речевых сигналов достаточно 200-300 отсчетов оцифрованного сигнала (при частоте квантования 10 кГц).

Результаты моделирования и обработки речевых сигналов, подтверждающие эффективность введения преобразования (24), более подробно описаны в [1,9], где на примере этого преобразования демонстрировалась возможность ковариационной компенсации необратимых нелинейных искажений речевых сигналов.

Таким образом, экспериментальные данные показывают, что теоретические результаты, приведенные выше, справедливы и для речевых сигналов.

7. Обсуждение результатов. Выводы

Еще несколько лет назад в широком спектре задач автоматической обработки речи проблема борьбы с нелинейными амплитудными искажениями в трактах связи не имела столь значительного веса как сейчас. Повышенный интерес к решению указанной проблемы возник в результате начавшегося массового внедрения устройств цифровой обработки речевых сигналов в разнообразные автоматизированные системы. Опыт внедрения показал, что нелинейные искажения, присущие каналам связи, могут снизить показатели систем обработки речи до непригодного уровня. Поэтому стало ясно, что без средств борьбы с искажениями не обойтись.

Поскольку речевой сигнал практически полностью описывается своей ковариационной (корреляционной) функцией либо спектральной плотностью и, к тому же, в основе большинства методов компрессии и распознавания речи лежат корреляционно-спектральные характеристики, задача борьбы с искажениями из амплитудно-временной области может быть переведена в корреляционную и спектральную области. В результате такого перевода становится возможным подавлять необратимые амплитудные искажения [1].

В зависимости от объема априорной информации борьба с амплитудными искажениями может осуществляться либо путем ковариационно-спектральной компенсации (как это сделано в [1]), либо путем оценивания корреляционно-спектральных характеристик неискаженного речевого сигнала, инвариантных к классу возможных искажений (как это сделано в данной работе). Очевидно,что инвариантное оценивание является более мощным средством подавления амплитудных искажений.

Решение проблемы выделения или оценивания спектрально-корреляционных характеристик речевого сигнала, инвариантных к нелинейным амплитудным искажениям, характерным для речевых трактов связи, сводится к введению такого амплитудного преобразования, которое приводит к инвариантности оценок корреляций и нормированной спектральной плотности неискаженного речевого сигнала, и к доказательству того, что эти оценки сходятся к оцениваемым величинам в среднеквадратическом.

Введение инвариантного преобразования по своей сути является нелинейным амплитудным искажением сигнала или случайного процесса. Поэтому, опустив проблему инвариантности, в работе сначала доказываются две теоремы (1 и 2), отвечающие на вопрос, при каких условиях по преобразованному процессу можно получить оценки корреляций и нормированной спектральной плотности исходного процесса, сходящиеся к своим истинным значениям в среднеквадратическом и имеющие асимптотически нормальное распределение. Из сформулированных условий вытекают явные и неявные ограничения на вводимое преобразование. Одним из основных явных ограничений является требование корреляционной обратимости преобразования. К неявным ограничениям относятся условия конечности четвертых моментов преобразованного процесса.Если в дополнение к сформулированным условиям амплитудное преобразование инвариантно к некоторому классу искажений,то и полученные оценки также инвариантны к этому классу искажений (следствия 1 и 2). Таким образом, для стационарных случайных процессов проблема инвариантности решена.

Очевидно, что проблема решена и для стационарных гауссовских процессов. Однако для преобразованных гауссовских процессов обременительное условие конечности четвертых моментов может быть опущено (теоремы 3 и 4), если предположить, что корреляционная функция гауссовского процесса равна нулю, начиная с некоторой задержки. Поскольку на практике указанное предположение практически всегда выполняется, приведенные теоремы будут полезны для интерпретации результатов.

Так как обработка речевых сигналов ведется на коротких участках анализа, где сигнал можно рассматривать как реализацию стационарного случайного процесса, проблема инвариантности оказывается решенной и для речевых сигналов. Иными словами,если имеется класс искажений и для него найдено подходящее инвариантное корреляционно обратимое амплитудное преобразование,то и для речевых сигналов будут справедливы теоремы 1 и 2 и их следствия, т.е. будет обеспечена среднеквадратическая сходимость оценок.

Перечисленные результаты относятся к разряду обобщающих, так как при их формулировке не описывается класс искажений и не указывается вид инвариантного преобразования. Поскольку вид инвариантного преобразования зависит от класса допустимых искажений, для решения конкретной задачи в работе дано формальное описание класса искажений, характерных для речевых трактов связи. Требования к инвариантному преобразованию вместе с описанием класса искажений позволили указать инвариантное преобразование (утверждение 2) - клиппирование, которое является корреляционно обратимым для гауссовских процессов (утверждение 3) и для речевых сигналов (по результатам экспериментов [1,9]). Следовательно,для рассмотренного класса искажений проблема теоретически решена. В целях иллюстрации решения проблемы приведены следствия 3 и 4 для $\mathbf{m}$-связных гауссовских процессов.

Теоретические результаты, приведенные в работе, подтверждены экспериментальными данными. Моделирование показало, что зависимость между корреляциями речевого сигнала до и после введенного инвариантного преобразования (клиппирования) практически совпадает с аналогичной зависимостью при ведении такого же инвариантного преобразования для гауссовских процессов. Среднеквадратическая ошибка инвариантного оценивания корреляций и нормированной спектральной плотности есть величина $O(1/N)$. При этом речевой сигнал, восстановленный по инвариантным оценкам, на слух почти не отличается от речевого сигнала, восстановленного по оценкам корреляций, полученным по непреобразованному сигналу.

Литература

1. КЕЛЬМАНОВ А.В. Решение проблемы компенсации необратимых нелинейных амплитудных искажений речевого сигнала по оценкам ковариаций //Настоящий сборник. - С. 22-42.

2. АНДЕРСОН Т. Статистический анализ временных рядов: Пер. с англ. /Под ред. Ю.К.Беляева. - М.: Мир, 1976. - 755 с.

3. САПОЖКОВ М.А. Речевой сигнал в кибернетике и связи. - М.: Связьиздат, 1963. - 450 с.

4. РЕПИНА О.И. Искажения в телефонном тракте. - М.: Связь, 1978. - 174 с.

5. КЕЛЬМАНОВ А.В., ХАЙРЕТДИНОВА А.Г. Исследование свойств искаженных речевых сигналов //Анализ данных и знаний в экспертных системах. - Новосибирск, 1990. - Вып. 134: Вычислительные системы. - С. 140-160.

6. КЕЛЬМАНОВ А.В., ХАЙРЕТДИНОВА А.Г., ХАМИДУЛЛИН С.А.Цифровая обработка речевых сигналов, искаженных вибропомехами // Анализ сигналов и символьных последовательностей. - Новосибирск, 1991. - Вып. 141: Вычислительные системы. -С.102-116.

7. БОРОВКОВ А.А. Математическая статистика. - М.: Наука, 1984. - 472 с.

8. ЛЕВИН Б.Р. Теоретические основы статистической радио - техники. Т. 1. - М.: Сов. радио, 1974. - 552 с.

9. КЕЛЬМАНОВ А.В., ХАМИДУЛЛИН С.А. Статистическое оценивание зависимости между первыми и вторыми моментами речевого сигнала до и после нелинейных искажений //Анализ сигналов и символьных последовательностей. - Новосибирск, 1991.-Вып. 141: Вычислительные системы, -С. 117-131.

Поступила в ред.-изд.отд.

17 сентября 1991 года