

УДК 621.391; 519.21; 519.245

© Г. Ф. Малыхина, А. В. Меркушева

СЕТЬ С СИММЕТРИЧНОЙ ФУНКЦИЕЙ ПРЕОБРАЗОВАНИЯ НЕЙРОНОВ НА ОСНОВЕ СТАТИСТИК ВЫСОКОГО ПОРЯДКА ДЛЯ СНИЖЕНИЯ ШУМА НЕСТАЦИОНАРНОГО СИГНАЛА

Проанализирован метод создания сети с симметричной функцией преобразования нейронов, построенный на основе статистик высокого порядка (СВП). Показано, что такая сеть обеспечивает снижение шума нестационарного сигнала. При этом в информационно-измерительной системе хорошо подавляются шумы с распределением Гаусса и шумы с негауссовым, но симметричным распределением. Показана структура сети на основе СВП, сравнение ее с базовой структурой сети с симметричной функцией преобразования нейронов и алгоритмом обучения, использующим критерий МНК.

ВВЕДЕНИЕ

В информационно-измерительных системах (ИИС) для улучшения отношения сигнала к шуму достаточно часто применяются методы адаптивной фильтрации [1–3], однако использование линейных фильтров не обеспечивает желаемой эффективности подавления шума для нелинейных сигналов. Для таких условий лучшие результаты обеспечиваются применением нейронных сетей [4–6]. Опыт использования многослойного персептрона показал, что для аппроксимации нелинейного сигнала он имеет преимущество по числу параметров и сравнительно небольшому количеству обучаемых данных [5, 7]. Однако алгоритм обучения персептрона построен на использовании нелинейной формы оценки параметров; он имеет медленную сходимость, что затрудняет его применение при временном дефиците. Кроме того, чтобы исключить получение локального (а не общего) минимума критерия оптимизации, приходится обучение проводить несколько раз с различными начальными значениями параметров и выбирать наилучшее значение критерия. Более эффективным является использование сети с симметричной функцией преобразования нейронов (СФПН)¹⁾, применявшейся для кластеризации, выравнивания каналов связи и снижения шума [9].

Следующим этапом развития методов применения нейронных сетей для подавления шума, сопутствующего информационно значимому сигналу, является использование статистик высокого порядка (СВП)²⁾. Это направление связано с рабо-

тами Шалдера, Гианакиса и Ли (Shadler, Giannakis, Li [10]), Мо и Шафая (Mo, Shafai [11]), Никиаса и Менделя (Nikias, Mendel [12]). Преимуществом СВП является их непропускание (подавление) гауссова шума и обнаружение шума, имеющего асимметричное или негауссово распределение.

Ниже будет проанализирована структура сети с СФПН на основе СВП и метод ее использования для снижения уровня шума в сигнале, или так называемого "усиления" сигнала в смысле повышения его эффективного отношения сигнал/шум. Важным элементом такой сети является опорный сигнал, образуемый из наращиваемой задержки основного анализируемого сигнала, т. е. из зашумленного информационного сигнала. При этом непропускание на выход блока, формирующего СВП, как гауссова шума, так и симметрично распределенного негауссова шума, снижает влияние этих компонент на кластеризацию, которая входит в алгоритм обучения сети при оценке ее весовых параметров. Анализ сети с СФПН на основе СВП будет проведен в три этапа: оценка характерной структуры и формализма описания простой сети с СФПН; способ образования СВП и структура сети с СФПН, включающая блок формирования СВП; алгоритм обучения сети и формирования весов — свободных параметров сети.

СФПН ДЛЯ УМЕНЬШЕНИЯ ШУМА В СИГНАЛЕ

Если измеряемый сигнал $d(t)$ представить как сумму информационно-значимого сигнала $s(t)$

¹⁾ Общая методология сетей с СФПН, включая особенности их структуры, рассмотрены в [8].

²⁾ Сравнительно недавно используемые СВП связаны с введением распределения группы вероятностных ("случайных") сигналов, построением их совместной харак-

теристической функции и последующим вычислением по общему алгоритму формирования хорошо знакомых первых двух моментов (статистик) — среднего и дисперсии. Формализованное описание СВП дано в следующем разделе.

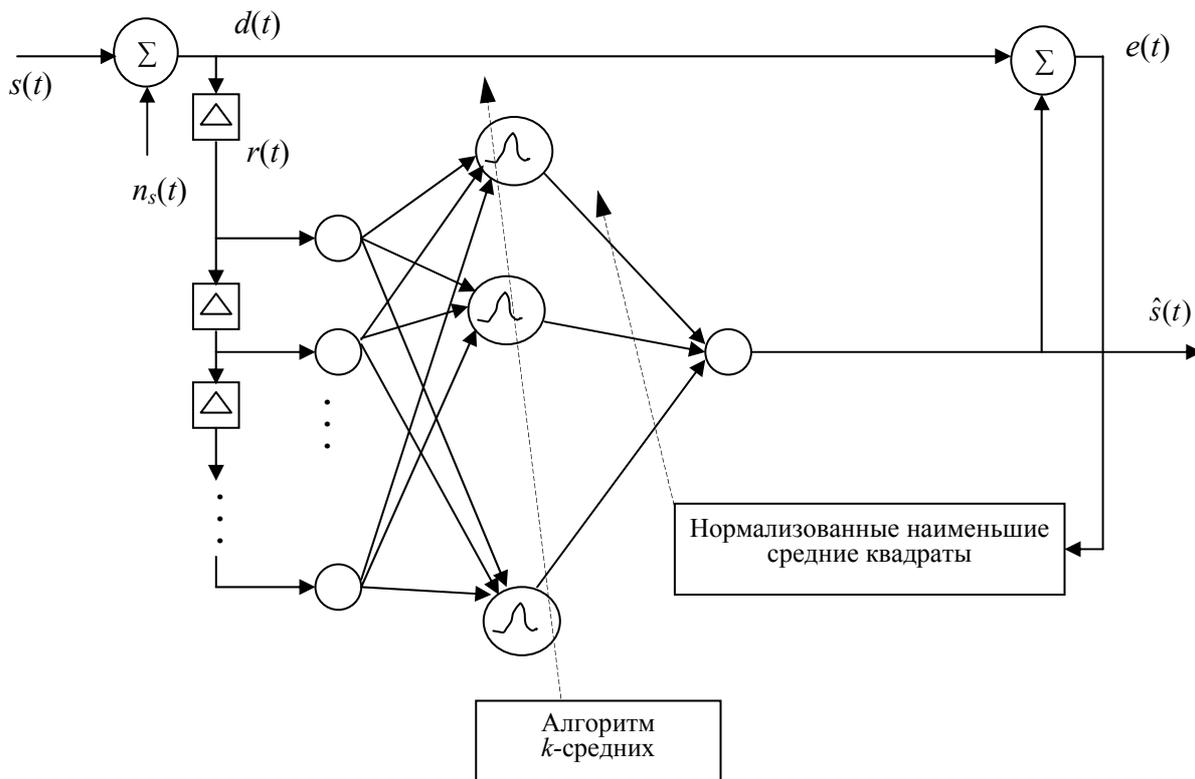


Рис. 1. Сеть с СФПН и элементы алгоритма для снижения шума сигнала. Штриховые стрелки символизируют адаптивную подстраиваемость элементов сети

и некоррелированного с ним гауссова шума $n_s(t)$: $d(t) = s(t) + n_s(t)$, то простейшая форма сети с СФПН может быть представлена структурой, показанной на рис. 1. Опорный сигнал $r(t)$ равен $d(t - \Delta)$ и на каждый следующий узел входного слоя сети поступает с дополнительной задержкой Δ , где в качестве Δ , глубины предсказания, принимается интервал отсчетов (дискретизации) сигнала. Оценка $\hat{s}(t)$ сигнала $s(t)$ получается в этой структуре по логике адаптивной фильтрации [1].

Для сети с N_0 узлами входного слоя, N_1 узлами скрытого слоя и с одним выходным узлом (с линейной ФПН) на выходе сети оценка $\hat{s}(t)$ определяется соотношением

$$\hat{s}(t) = \mathbf{\delta}(t)^T \mathbf{w}(t), \tag{1}$$

где $\mathbf{w}(t)$ — вектор связи, идущей от узлов скрытого слоя к выходному узлу сети; $\mathbf{\delta}(t) = [\delta_1(t), \delta_2(t), \dots, \delta_{N_1}(t)]^T$ — вектор сигналов выхода узлов скрытого слоя. Выход k -го скрытого узла для сети с СФПН "гауссова типа" определяется выражением

$$\delta_k(t) = \exp\left(-\frac{\|\mathbf{r}(t) - \mathbf{v}_k(t)\|^2}{2\rho^2(t)}\right), \tag{2}$$

где $\mathbf{v}_k(t)$ — центр k -го скрытого узла; $\mathbf{r}(t) = [r(t), r(t-1), \dots, r(t-N_0+1)]^T$ — входной вектор сети; $\|\cdot\|$ — евклидова (L_2) норма. Величина $\rho(t)$ — полуширина (аналог σ) СФПН определяется среднеквадратичным рассеянием входного вектора сети:

$$\rho^2(t) = \frac{1}{N_0} \left(\sum_{i=0}^{N_0-1} [r(t-i) - \bar{r}(t)] \right)^2, \tag{3}$$

где $\bar{r}(t)$ — среднее значение компонент вектора $\bar{\mathbf{r}}(t)$.

Центры $\mathbf{v}_k(t)$ могут быть получены "самоорганизующейся процедурой обучения" с помощью алгоритма кластеризации по k средним [9]:

$$\begin{cases} \mathbf{v}_k(t+1) = \mathbf{v}_k(t) + \mu_v [\mathbf{x}(t) - \mathbf{v}_k(t)] & \text{при} \\ k = \arg \min [\|\mathbf{x}(t) - \mathbf{v}_j(t)\|, 1 \leq j \leq N_1]; & (4) \\ \mathbf{v}_k(t+1) = \mathbf{v}_k(t) & \text{при } 1 \leq i \leq N_1 \text{ и } i \neq k, \end{cases}$$

где μ_v — скорость обучения для весов сети.

Поскольку опорный сигнал $\mathbf{r}(t)$ содержит шум $n_s(t)$, то на центры, полученные по алгоритму k средних, тоже влияет $n_s(t)$, особенно в случае, когда отношение сигнал/шум у входного сигнала имеет небольшую величину.

Весовой вектор $\mathbf{w}(t)$ при обучении сети адаптируется по алгоритму МНК [3], в котором средний квадрат ошибки $d(t) - \hat{s}(t)$ определяется соответствующим критерием. Нормализованный вариант алгоритма МНК выражается соотношением

$$\begin{aligned} \mathbf{w}(t+1) &= \\ &= \mathbf{w}(t) + \frac{\mu_w}{(1 + \delta(t)^T \delta(t))} \delta(t) \cdot (d(t) - \hat{s}(t)), \end{aligned} \quad (5)$$

где μ_w — параметр скорости обучения для вектора весов.

Однако шум $e(t)$ на выходе сети также коррелирован с $n_s(t)$, поскольку

$$e(t) = d(t) - \hat{s}(t) = f(t) + n_s(t) - \hat{s}(t). \quad (6)$$

Поэтому из (6) следует, что в (5) присутствует зависимость от $n_s(t)$, увеличенная за счет коэффициента μ_w . Это можно трактовать как возрастание шума градиента, и использование алгоритма МНК при низких значениях отношения сигнал/шум становится неэффективным.

СЕТЬ С СФПН НА ОСНОВЕ СТАТИСТИК ВЫСОКОГО ПОРЯДКА

Статистики высокого порядка формализованно определяются как коэффициенты разложения в ряд Тейлора совместной характеристической функции произвольного набора вероятностных переменных, в том числе — временных отсчетов стационарного (по крайней мере локально) случайного сигнала [12]. Так, для набора переменных $\{z_1, z_2, \dots, z_n\}$ их кумулянта k -го порядка ($k = k_1 + k_2 + \dots + k_n$) определяется соотношением

$$\begin{aligned} & \text{Кум.} [z_1^{k_1}, z_2^{k_2}, \dots, z_n^{k_n}] \equiv \\ & \equiv (-j)^k \frac{\partial^k \psi(\omega_1, \omega_2, \dots, \omega_n)}{\partial z_1^{k_1} \partial z_2^{k_2} \dots \partial z_n^{k_n}} \Big|_{\omega_1 = \omega_2 = \dots = \omega_n = 0}, \end{aligned} \quad (7)$$

где $\psi(\omega_1, \omega_2, \dots, \omega_n) \equiv E[j \cdot (\omega_1 z_1 + \omega_2 z_2 + \dots + \omega_n z_n)]$ — характеристическая функция и E — символ математического ожидания. Для набора n вероятностных (случайных) стационарных переменных $\{z_i(t)\}_{i=1,2,\dots,n}$ кросс-кумулянта (C) n -го порядка представляется в виде:

$$\begin{aligned} & C_{z_1 z_2 \dots z_n}(\tau_1, \tau_2, \dots, \tau_{n-1}) \equiv \\ & \equiv \text{Кум.} [z_1(t), z_2(t + \tau_1), z_3(t + \tau_2), \dots, z_n(t + \tau_{n-1})]. \end{aligned} \quad (8)$$

Два основных свойства кросс-кумулянты, которые существенны для анализа сети с СФПН на основе СВП, следующие.

1. Если случайные переменные $\{z_i(t)\}_{i=1,2,\dots,n}$ могут быть разделены на две статистически независимые группы переменных, то

$$C_{z_1 z_2 \dots z_n}(\tau_1, \tau_2, \dots, \tau_{n-1}) = 0.$$

2. Если набор случайных переменных $\{z_i(t)\}_{i=1,2,\dots,n}$ соответствует гауссову сигналу, тогда все их кросс-кумулянты порядка более чем 2 равны нулю.

Поскольку $s(t)$ и $n_s(t)$ независимы, то (в предположении, что у опорного сигнала $r(t)$ существует кумулянта n -го порядка) эта кумулянта может быть представлена в виде:

$$\begin{aligned} C_{rr\dots r}(\tau_1, \tau_2, \dots, \tau_{n-1}) &= C_{ss\dots s}(\tau_1, \tau_2, \dots, \tau_{n-1}) + \\ &+ C_{n_s n_s \dots n_s}(\tau_1, \tau_2, \dots, \tau_{n-1}). \end{aligned} \quad (9)$$

Поэтому кумулянты высокого порядка (КВП) опорного сигнала будут фактически близки к кумулянтам сигнала $s(t)$. Это и служит основной идеей построения структуры сети с СФПН на основе СВП. Такая структура, включающая кумулянты, показана на рис. 2.

Кумулянты высокого порядка опорного сигнала $r(t)$ используются как вход сети на основе СВП, для того чтобы снизить влияние шума $n_s(t)$ на кластеризацию, обеспечивающую правильный выбор центров СФПН по (4). Выход $\delta_k(t)$ k -го узла скрытого слоя в такой структуре может быть представлен выражением:

$$\delta_k(t) = \exp \left(- \frac{\|C_{rr\dots r}(t) - \mathbf{v}_k(t)\|^2}{2[\rho(t)]^2} \right), \quad (10)$$

где полуширина $\rho(t)$ СФПН определяется среднеквадратичным рассеянием компонент $C_{rr\dots r}(t)$ аналогично формуле (3).

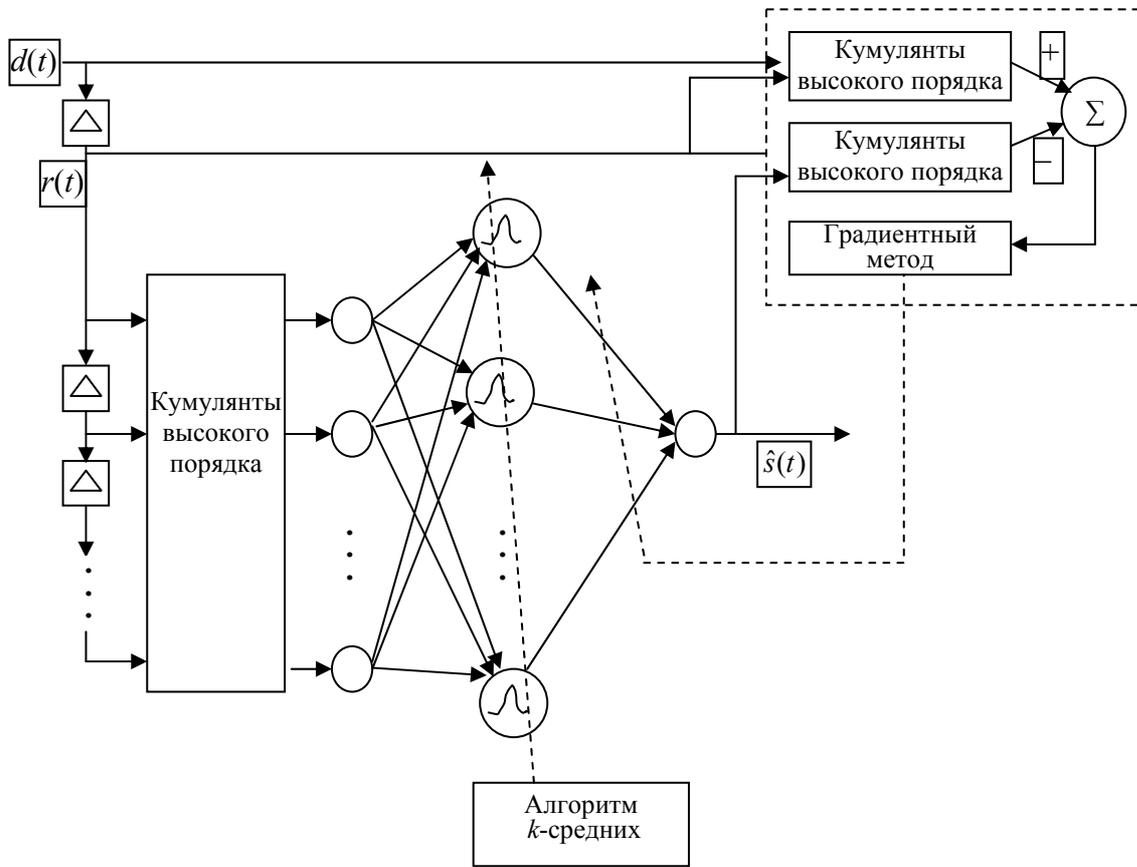


Рис. 2. Схема основной структуры сети с СФПН на основе статистик высокого порядка

**АЛГОРИТМ ОБУЧЕНИЯ СЕТИ С СФПН
НА ОСНОВЕ СТАТИК ВЫСОКОГО ПОРЯДКА**

В алгоритме обучения анализируемой сети на основе СВП в качестве критерия (ξ) используется среднеквадратичная разность ("ошибка") двух кумулянт высокого порядка — для сигналов $d(t)$ и $\hat{s}(t)$ — с целью снизить влияние шума $n_s(t)$ на адаптацию весовых параметров сети. Критерий ξ с использованием приведенных выше определений и обозначений КВП может быть представлен в следующем виде:

$$\begin{aligned}
 \xi &= \sum_{\tau_1} \sum_{\tau_2} \dots \sum_{\tau_{n-1}} \frac{1}{2} [C_{\hat{s}rr\dots r}(\tau_1, \tau_2, \dots, \tau_{n-1})]^2 = \\
 &= \sum \dots \sum_{(\tau_1, \tau_2, \dots, \tau_{n-1}) \in \Gamma} \dots \sum \frac{1}{2} [C_{\hat{s}s\dots s}(\tau_1 - \Delta, \tau_2 - \Delta, \dots, \tau_{n-1} - \Delta) - C_{ss\dots s}(\tau_1 - \Delta, \tau_2 - \Delta, \dots, \tau_{n-1} - \Delta)]^2 = \\
 &= \sum_{(\tau_1, \tau_2, \dots, \tau_{n-1}) \in \Gamma} \sum \dots \sum \frac{1}{2} [\omega_j C_{\delta_j rr\dots r}(\tau_1, \tau_2, \dots, \tau_{n-1}) - C_{dr\dots r}(\tau_1, \tau_2, \dots, \tau_{n-1})]^2.
 \end{aligned}
 \tag{11}$$

Соотношение (11) может быть представлено в матричной форме в виде

$$\xi = \frac{1}{2} [\mathbf{C}_{\delta_{rr\dots r}} \mathbf{w} - \mathbf{C}_{d_{rr\dots r}}]^2, \quad (12)$$

где $\mathbf{C}_{\delta_{rr\dots r}}$ — матрица размера $N_r \times N_1$ и вектор $\mathbf{C}_{d_{rr\dots r}}$ размера $N_r \times 1$; N_r обозначает число точек в наборе Γ .

Для того чтобы минимизировать критерий ξ , используется метод градиентного спуска. Градиент ξ определяется выражением

$$\nabla_{\mathbf{w}} \xi \equiv \frac{\partial \xi}{\partial \mathbf{w}(t)} = -\mathbf{C}_{\delta_{rr\dots r}}^T [\mathbf{C}_{d_{rr\dots r}} - \mathbf{C}_{\delta_{rr\dots r}} \mathbf{w}]. \quad (13)$$

Общее соотношение, которое отражает алгоритм обучения и определяет подстройку — адаптацию весов нейронной сети, имеет вид:

$$\mathbf{w}(t+1) = \mathbf{w}(t) - \frac{\mu_w}{(1 + \text{tr}(\mathbf{C}_{\delta_{rr\dots r}}^T))} \cdot \nabla_{\mathbf{w}} \xi(t). \quad (14)$$

Из выражения (11) следует, что при хорошей аппроксимации $\hat{s}(t)$ для $s(t)$ критерий обучения ξ становится близким к оптимальному значению. Кроме того, (11) показывает, что влияние $n_s(t)$ на $\mathbf{C}_{d_{rr\dots r}}$ может быть ослаблено. Поэтому в алгоритме обучения, основанном на статистиках высокого порядка, градиентное усиление шума эффективно снижается.

Полное представление об алгоритме обучения сети с СФПН на основе СВП дает формализованная сводка его этапов, которая для большей компактности (не уменьшая общности) представлена в форме, ограниченной статистиками третьего порядка. В сводке приведены обозначения переменных, правило инициализации статистик (кросскумулянт) и последовательность вычислительных процедур. Процедуры включают вычисление $\hat{s}(t)$, обновление центров СФПН, оценку кумулянт, образование вектора $\mathbf{C}_{d_{rrr}}$ и матриц $\mathbf{C}_{r_{rrr}}$ и $\mathbf{C}_{\delta_{rrr}}$, а также обновление весов нейронной сети.

Сводка этапов алгоритма обучения сети с СФПН на основе статистик третьего порядка

Определения:

N_0 — число входных узлов; N_1 — число узлов скрытого слоя; λ — фактор "забывания"; μ_w — параметр скорости обучения для весов; μ_v — параметр скорости обучения для центров СФПН; $\mathbf{w}(t)$ — вектор весов; \mathbf{v}_k — центр k -й СФПН в скрытом слое; ρ — полуширина гауссиана СФПН.

1. Начальные условия:

$$\begin{aligned} \tau_1 &= -m_1, (-m_1 + 1), \dots, (m_1 - 1), m_1; \quad \tau_2 = 1, 2, \dots, m_2; \\ \hat{\mathbf{C}}_{r_{rr}}(0; \tau_1, \tau_2) &= 0, \quad i = 1, 2, \dots, N_0; \\ \hat{\mathbf{C}}_{\delta_{rr}}(0; \tau_1, \tau_2) &= 0, \quad i = 1, 2, \dots, N_1; \\ \hat{\mathbf{C}}_{d_{rr}}(0; \tau_1, \tau_2) &= 0; \end{aligned}$$

Набор случайных значений для $\mathbf{v}_k(0)$ при $k = 1, 2, \dots, N_1$.

2. Вычисления:

$$t = 1, 2, \dots; \quad \hat{s}(t) = [\delta_1(t), \delta_2(t), \dots, \delta_{N_1}(t)] \cdot \mathbf{w}(t), \quad \text{где}$$

$$\delta_k(t) = \exp\left(-\frac{\|\mathbf{C}_{r_{rr}}(t) - \mathbf{v}_k\|^2}{2\rho^2}\right).$$

2.1. Обновление центров:

$$\begin{cases} \mathbf{v}_k(t+1) = \mathbf{v}_k(t) + \mu_v [\mathbf{C}_{r_{rr}}(t) - \mathbf{v}_k], \\ k = \arg \min [\|\mathbf{C}_{r_{rr}}(t) - \mathbf{v}_j\|^2, 1 \leq j \leq N_1]; \\ \mathbf{v}_i(t+1) = \mathbf{v}_i(t), \quad 1 \leq i \leq N_1 \text{ и } i \neq k. \end{cases}$$

2.2. Оценка кумулянт третьего порядка:

Для $\tau_1 = -m_1, (-m_1 + 1), \dots, (m_1 - 1), m_1;$
 $\tau_2 = 1, 2, \dots, m_2$

$$\hat{\mathbf{C}}_{d_{rr}}(t; \tau_1, \tau_2) = \lambda \hat{\mathbf{C}}_{d_{rr}}(t-1; \tau_1, \tau_2) + (1-\lambda)d(t)r(t+\tau_1)r(t+\tau_1+\tau_2).$$

Для $\tau_1 = -m_1, (-m_1 + 1), \dots, (m_1 - 1), m_1;$
 $\tau_2 = 1, 2, \dots, m_2; \quad i = 1, 2, \dots, N_0$

$$\hat{\mathbf{C}}_{r_{rr}}(t; \tau_1, \tau_2) = \lambda \hat{\mathbf{C}}_{r_{rr}}(t-1; \tau_1, \tau_2) + (1-\lambda)r(t-i+1)r(t+\tau_1)r(t+\tau_1+\tau_2).$$

Для $\tau_1 = -m_1, (-m_1 + 1), \dots, (m_1 - 1), m_1;$
 $\tau_k = 1, 2, \dots, m_2; \quad i = 1, 2, \dots, N_1$

$$\hat{\mathbf{C}}_{\delta_{rr}}(t; \tau_1, \tau_2) = \lambda \hat{\mathbf{C}}_{\delta_{rr}}(t-1; \tau_1, \tau_2) + (1-\lambda)\delta_i(t)r(t+\tau_1)r(t+\tau_1+\tau_2).$$

2.3. Построение $C_{drr}(t)$, $C_{\delta rr}(t)$ и $C_{rrr}(t)$:

$$C_{drr}(t) = \left[\hat{C}_{drr}(t; -m_1, 0), \hat{C}_{drr}(t; -m_1, 1), \dots, \hat{C}_{drr}(t; m_1, m_2 - 1), \hat{C}_{drr}(t; m_1, m_2) \right]^T;$$

$$C_{rrr}(t; \tau_1, \tau_2) = \begin{bmatrix} \hat{C}_{r_{1rr}}(t; -m_1, 0) & \hat{C}_{r_{1rr}}(t; -m_1, 1) & \dots & \hat{C}_{r_{1rr}}(t; m_1, m_2) \\ \hat{C}_{r_{2rr}}(t; -m_1, 0) & \hat{C}_{r_{2rr}}(t; -m_1, 1) & \dots & \hat{C}_{r_{2rr}}(t; m_1, m_2) \\ \dots & \dots & \dots & \dots \\ \hat{C}_{r_{N_0rr}}(t; -m_1, 0) & \hat{C}_{r_{N_0rr}}(t; -m_1, 1) & \dots & \hat{C}_{r_{N_0rr}}(t; m_1, m_2) \end{bmatrix};$$

$$C_{\delta rr}(t; \tau_1, \tau_2) = \begin{bmatrix} \hat{C}_{\delta_{1rr}}(t; -m_1, 0) & \hat{C}_{\delta_{1rr}}(t; -m_1, 1) & \dots & \hat{C}_{\delta_{1rr}}(t; m_1, m_2) \\ \hat{C}_{\delta_{2rr}}(t; -m_1, 0) & \hat{C}_{\delta_{2rr}}(t; -m_1, 1) & \dots & \hat{C}_{\delta_{2rr}}(t; m_1, m_2) \\ \dots & \dots & \dots & \dots \\ \hat{C}_{\delta_{N_1rr}}(t; -m_1, 0) & \hat{C}_{\delta_{N_1rr}}(t; -m_1, 1) & \dots & \hat{C}_{\delta_{N_1rr}}(t; m_1, m_2) \end{bmatrix}.$$

2.4. Обновление весов:

$$\nabla_w(t) \equiv \frac{\partial \xi}{\partial \mathbf{w}(t)} = -C_{\delta rr}^T(t) \cdot [C_{drr}(t) - C_{\delta rr}(t) \cdot \mathbf{w}(t)];$$

$$\mathbf{w}(t+1) = \mathbf{w}(t) - \frac{\mu_w}{\text{tr}[C_{\delta rr}^T(t)C_{\delta rr}(t)]} \cdot \nabla_w(t).$$

СРАВНЕНИЕ АЛГОРИТМОВ НА ОСНОВЕ СВП И НА ОСНОВЕ МНК ДЛЯ СЕТИ С СФПН

Сравнение показателей функционирования сети с СФПН при использовании для обучения двух критериев — критерия на основе статистик высокого порядка и обычно применяемого нормализованного критерия наименьших квадратов (МНК) полезно проследить на сравнительно несложном нестационарном сигнале. В качестве него принят сигнал электрокардиограммы (ЭКГ), имеющий хорошо знакомую форму с квазипериодическим чередованием Q-R-S-T-комплексов, с небольшой вариацией амплитуд Q-R-S-T-пику, со спектром в пределах 150 Гц, а также с удобным и наглядным способом внесения различных уровней зашумленности.

Сравнение показателей функционирования проводилось на сети (при одинаковых значениях параметра μ_w), обученной по критерию МНК и критерию СВП, включающему статистику 3-го порядка (с $m_1=10$ и $m_2=3$). При внесении гауссова шума и обычном определении отношения сиг-

нал/шум (ОСШ)³⁾ алгоритмом на основе СВП удается снизить шум в 1.5 раза более, чем с помощью алгоритма на основе МНК, т. е. на 30 и на 20 дБ соответственно. Кроме того, алгоритм на основе СВП значительно меньше зависит от величины μ_w — параметра скорости обучения. Особенно значительно преимущество алгоритма на основе СВП при очень низких значениях ОСШ. Так в диапазоне ОСШ от -10 до -20 дБ СВП обеспечивают в 3–7 раз более низкую среднеквадратичную ошибку получаемого сигнала на выходе сети с СФПН.

Примерно такое же преимущество алгоритм на основе СВП имеет перед алгоритмом с критерием МНК для случая, когда сигналу сопутствует нестационарный гауссов шум, у которого может со временем изменяться как величина, так и текущее значение средней частоты. Для низких значений ОСШ до -10 дБ алгоритм на СВП (сравнительно с МНК) обеспечивает вдвое меньшую среднеквадратичную ошибку сигнала.

Следует отметить большую вычислительную нагрузку, которая требуется для выполнения алгоритма на СВП. Поэтому при не слишком малых уровнях ОСШ, где достигаются почти эквивалентные показатели функционирования обоих алгоритмов, предпочтение может быть отдано более

³⁾ ОСШ по амплитуде, в децибелах (дБ):

ОСШ = $20 \log_{10} \left(\frac{A_{\text{сигн.}}}{A_{\text{шум}}} \right)$ дБ, где $A_{\text{сигн.}}$ и $A_{\text{шум}}$ — соответственно амплитуды сигнала и шума.

простою из них, т. е. нормализованному МНК [3].

ЗАКЛЮЧЕНИЕ

Рассмотрен метод применения статистик высокого порядка (СВП), так называемых кумулянт, при обучении нейронной сети с симметричной функцией преобразования нейронов (сети с СФПН) и применение такой сети для увеличения информационно-значимого сигнала на основе подавления шума, т. е. повышения величины отношения сигнал/шум (ОСШ) в тракте измерительной или управляющей системы.

СВП являются сравнительно новой статистической концепцией⁴⁾, нашедшей применение в области обработки сигналов [12, 17]. Развитие свойств СВП и методов применения выполнено Белоучрани, Мераим и Джеан-Франкоис (Belouchrani, Meraim, Jean-Francois) [18], Инойе и Хирано (Inouye, Hirano) [19], Йеунг и Яу (Yeung, Yau) [20], причем в двух последних работах получен ряд основополагающих теоретических положений, расширяющих возможности применения СВП для анализа смеси сигналов и снижения уровня шума у нестационарных сигналов.

- В статье проанализирован метод создания сети с СФПН, показана ее структура, обеспечивающая снижение шума нестационарного сигнала, и критерий обучения, построенный на основе СВП.

- Показано формализованное определение СВП (совместных кумулянт группы стохастических переменных), их основные свойства и способ формирования из временных отсчетов компонент вектор-сигнала в ИИС.

- Приведен алгоритм обучения сети с СФПН на основе СВП, где в качестве целевого критерия обучения использована среднеквадратичная разность двух СВП, и дана подробная сводка всех вычислительных этапов алгоритма.

- Дано сравнение показателей функционирования сети с СФПН при использовании для обучения двух критериев — критерия на основе статистик высокого порядка и обычно применяемого нормализованного критерия наименьших квадратов (МНК).

- Показано, что сети с СФПН на основе СВП обеспечивают большее снижение компоненты шума в информационно-значимом сигнале, и это преимущество сети особенно усиливается при низкой величине отношения сигнал/шум.

- С помощью описанной структуры сети с СФПН и алгоритма на основе СВП осуществляет-

ся подавление аддитивного гауссова шума как стационарного, так и нестационарного.

СПИСОК ЛИТЕРАТУРЫ

1. Уидроу Б., Стернз С. Адаптивная обработка сигналов. М.: Радио и связь, 1989. 439 с.
2. Zeidler J.R. Performance Analysis of LMS Adaptive Prediction Filters // Proceedings of IEEE. 1990. V. 78, N 12. P. 1781–1806.
3. Haykin S. Adaptive Filter Theory. Englewood Cliffs, N.J.: Prentice Hall, 1991. 560 p.
4. Michalopoulou Z., Nolte L., Alexandrou D. Performance Evaluation of Multilayer Perceptron in Signal Detection and Classification // Neural Networks. 1995. V. 6, N 2 (March). P. 381–386.
5. Gandhi P.P., Ramamurti V. Neural Networks for Signal Detection in Non-Gaussian Noise // IEEE Transactions on Signal Processing. 1997. V. 45, N 11. P. 2846–2851.
6. Cha I., Kassam S.A. Interference Cancellation Using Radial Basis Function Networks // Signal Processing. 1995. V. 47. P. 247–268.
7. Watterson J.W. Optimum Multilayer Perceptron Neural Receiver for Signal Detection // IEEE Transactions on Neural Networks. 1990. V. 1, N 4 (December). P. 298–300.
8. Меркушева А.В., Малыхина Г.Ф. Методология сетей с симметричными функциями преобразования нейронов // Научное приборостроение, 2006. Т. 16, № 2. С. 34–45.
9. Chen S., Mulgrew B., Grant P.M. A Clustering Technique for Digital Equalization Using Radial Basis Function Networks // IEEE Transactions on Neural Networks. 1993. V. 4, N 4. P. 570–578.
10. Shadler B.M., Giannakis G.B., Li K.-S. Estimation and Detection in Non-Gaussian Noise Using Higher Order Statistics // IEEE Transactions on Signal Processing. 1994. V. 42, N 10. P. 2729–2741.
11. Mop S., Shafai B. Blind Equalization Using Higher Order Statistics and Neural Network // IEEE Transactions on Signal Processing. 1994. V. 42, N 11. P. 3209–3217.
12. Nikias C.L., Mendel J.M. Signal Processing with Higher Order Spectra // IEEE Signal Processing Magazine. 1993. V. 10, N 3 (July). P. 10–37.
13. Гнеденко Б.В. Курс теории вероятности. М.: Изд. Физ.-мат. литературы, 1961. 406 с.
14. Кендалл М.Дж., Стюарт А. Многомерный статистический анализ и временные ряды. М.: Изд. Наука, 1976. 736 с.
15. Справочник по теории вероятностей и математической статистике / Под ред. акад. В.С. Королука. Киев: Наукова думка, 1978. 582 с.
16. Хастингс Н., Пикок Дж. Справочник по статистическим распределениям. М.: Статистика,

⁴⁾ Методология СВП и даже упоминание о них отсутствуют в авторитетных монографиях по статистике и теории вероятностей [13–16].

1980. 94 с.
17. *Nikias C.L., Petropulu A.P.* Higher Order Spectral Analysis: A Nonlinear Signal Processing Framework. Englewood Cliffs, N.Y.: Prentice Hall, 1993. 140 p.
18. *Belouchrani A., Meraim K.A.* A Blind Source Separation Technique Using Second-Order Statistics // IEEE Transactions on Signal Processing. 1997. V. 45, N 2. P. 434–443.
19. *Inouye Y., Hirano K.* Cumulant-Based Blind Identification of Linear Multi-Input-Multy-Output Systems Driven by Colored Inputs // IEEE Transactions on Signal Processing. 1997. V. 45, N 6. P. 1543–1552.
20. *Yeung K.L., Yau S.F.* A Cumulant-Based Super-Exponential Algorithm for Blind Deconvolution of Multi-Input-Multi-Output Systems // Signal Processing. 1998. P. 141–162.

Санкт-Петербург

Материал поступил в редакцию 5.09.2008.

NETWORK WITH SYMMETRICAL NEURONS TRANSFORMATION FUNCTION BASED ON HIGHER-ORDER STATISTICS FOR LOWERING THE NOISE OF UNSTATIONARY SIGNAL

G. F. Malychina, A. V. Merkusheva

Saint-Petersburg

The method for network designing with symmetrical neurons transformation functions (SNTF) based on higher-order statistics (HOS) was analyzed. This network was shown to provide noise lowering for unstationary signal. In the information-measurement systems noises with Gauss distribution and noises with non-Gauss but symmetric distribution are well suppressed. The structure of the network based on HOS is shown, and is compared with the network structure (having SNTF) and common learning LMS algorithm.