

УДК 681.51; 621.391; 519.21

© А. В. Меркушева, Г. Ф. Малыхина

ОБОБЩЕННЫЕ АДАПТИВНЫЕ НЕЙРОННЫЕ ФИЛЬТРЫ И ФИЛЬТРЫ НА ОСНОВЕ НЕЙРОННОЙ СЕТИ

Рассмотрены структура и свойства группы нелинейных адаптивных фильтров (НАФ), построенных на нейронах и на основе нейронных сетей. Приведены показатели точности: верхняя граница и абсолютное значение средней ошибки фильтров. Такие НАФ являются расширением класса стекковых фильтров, превосходят их по уровню подавления шума и могут использоваться для улучшения качества сигналов и изображений.

Кл. сл.: фильтры, нейронная сеть, точность, адаптация, качество, сигнал, изображение

ВВЕДЕНИЕ

Линейные фильтры (ЛФ) являются наиболее полезным средством для подавления шума, искажающего информативный сигнал. ЛФ оптимальной структуры минимизирует среднюю квадратичную ошибку (СКО) и оказывается лучшим при условии, что шум аддитивный, белый и гауссов [1]. Однако в приложениях такие условия выполняются редко, сигнал бывает нестационарным, а его характеристики часто неизвестны. Поэтому введен класс нелинейных фильтров (НФ), к которому относятся медианный и стекковый НФ [2, 3]. Стекковый НФ легко реализуется в приборном (а позднее и в микропроцессорном) исполнении. Для конфигурации стеккового НФ предложены адаптивные методы [4]. Несколько позже начато изучение методов создания структуры стекковых НФ на основе использования нейронов, а затем изучен метод создания адаптивного нелинейного фильтра с использованием нейронов и адаптивных фильтров на основе использования нейронных сетей (НС) [5, 6]. Вопросы оптимизации НФ на основе использования нейронов и НС отражены в [7] и [8]. Ниже будут проанализированы элементы теоретической основы обобщенных адаптивных нейронных фильтров (ОАНФ) и их связь с методом построения структуры стеккового фильтра, включающего элементы НС.¹⁾

¹⁾ Первоначальная схема для описания свойств НФ, содержащаяся в [5], ограничена примитивом НС: это структура из отдельного нейрона, который состоит из линейной дискриминантной функции и следующего за ней жесткого или мягкого ограничителя [9]. Эта структура интерпретирована соответственно как жесткий или мягкий нейронный фильтр. Большая часть анализа и получение свойств, описанных в нашей статье, выведены для обобщенного адаптивного фильтра, не предпо-

ОАНФ составляют достаточно большую группу ("класс") НФ, которая включает все стекковые фильтры. Оптимальный НФ типа ОАНФ имеет точную верхнюю границу средней абсолютной ошибки (САО), и по критерию САО такая НС функционирует лучше, чем стекковые фильтры.

СТЕКОВЫЕ ФИЛЬТРЫ

Медианный и другие, использующие в качестве аргумента ранговый порядок данных, операторы, обладают двумя важными свойствами: свойством разложения по порогам и стекковое свойство.²⁾ Первое — это свойство суперпозиции содержимого интервалов, ограниченных порогами, которое ведет к новой архитектуре для фильтра. Второе — свойство упорядочивания, которое позволяет эффективно выполнять фильтр в форме микросхемы с высокой интеграцией.

Любой фильтр, который обладает как свойством порогового разложения, так и стекковым свойством, можно условно называть стекковым фильтром (СФ). Можно считать, что СФ представляет

лагающего какой-либо специальной архитектуры нейронной сети. В связи с этим далее использован термин обобщенного адаптивного нейронного фильтра (ОАНФ).

²⁾ Стековым фильтром (СФ) (в узком смысле термина; см. текст ниже) для краткости называют фильтр со стекковой организацией входных данных. При этом выход СФ (на каждом временном шаге) получается как результат обработки некоторого небольшого массива из потока поступающих на СФ данных. Длина этого небольшого массива определяется длиной L стека. Стек обновляется по принципу "первым пришел — первым вышел", т. е. при поступлении нового элемента данных из стека удаляется элемент, который поступил в стек L временных шагов назад.

собой своеобразный "стек" положительных булевых функций, основанных на пороговом разложении и стековом свойстве [10, 11].

Свойство разложения по порогам и стековое свойство

Эти свойства являются определяющими для так называемого стекового фильтра, поэтому целесообразно рассмотреть их подробнее.

1. Бинарные последовательности длины n $X = (x_1, x_2, \dots, x_n)$ и $Y = (y_1, y_2, \dots, y_n)$ считаются равными ($X = Y$), если $x_i = y_i$ для всех i ; если из $x_i = 1$ следует, что $y_i = 1$, то считается, что $X \leq Y$; если же $X \leq Y$ и $X \neq Y$, то $X < Y$.

2. Если полагать, что последовательность $r(n)$: $(r(n_1), r(n_2), \dots, r(n_M))$ размера M , состоящая из отсчетов сигнала, поступает на вход фильтра, то бинарно-пороговое разложение $(T_1(n), T_2(n), \dots, T_{M-1}(n))$ последовательности $r(n)$ этих отсчетов определяется соотношением

$$T_i(n) = \begin{cases} 1, & \text{если } r(n) \geq i, \\ 0 & \text{иначе,} \end{cases} \quad (1)$$

где n стоит для n -го отсчета входной последовательности. Можно заметить, что эти пороговые разложения последовательностей обладают стековым свойством:

$$T_1(n) \geq T_2(n) \geq \dots \geq T_{M-1}(n). \quad (2)$$

3. Для двух бинарных последовательностей X и Y считается, что фильтр, определенный функцией $F(\cdot)$, имеет стековое свойство, только если выполняется условие:

$$F(X) \leq F(Y) \text{ при } X < Y. \quad (3)$$

4. На основе 1–3 выход фильтра $F(x)$ со стековым свойством дополнительно удовлетворяет соотношению:

$$F(T_{M-1}) \leq F(T_{M-2}) \leq \dots \leq F(T_1), \quad (4)$$

где T_i — бинарная последовательность, разложенная на уровне i , которая получена из последовательности $r(n)$, состоящей из M величин.

5. Все упорядочивающие ранг-фильтры выполняются за счет их обеспеченности классом булевых функций, известных как положительные булевы функции. Положительными булевыми функциями (ПБФ) здесь называются булевы функции, которые удовлетворяют стековому свойству, выраженному соотношением (3). При этом фильтр, в котором бинарный оператор в структуре порогового разложения определен положительными булевыми функциями, называется стековым

фильтром.³⁾ Трудность прямого применения метода ПБФ связана с выбором оптимального набора, обладающего нужными свойствами, поэтому желательно получить более эффективную схему конфигурации стекового фильтра.

Оптимальная стековая фильтрация

Метод оптимальной стековой фильтрации основан на минимизации САО между выходом стекового фильтра и желаемым значением выхода при данном распределении шума. Первоначально разработанный вариант этого метода [4] имеет достаточно много ограничений. В нем предполагается совместная стационарность фильтруемого сигнала и желаемого сигнала на выходе фильтра; необходимо знание или оценка коэффициентов функции стоимости (критерия оптимизации фильтра); вычислительная сложность метода экспоненциально возрастает при увеличении размера стека фильтра. Поэтому разработанная в [4] схема оптимизации требует существенного усовершенствования. Чтобы подчеркнуть особенность анализируемого ниже метода, за его основу приняты обобщенные адаптивные нейронные фильтры (ОАНФ).

ОБОБЩЕННЫЕ АДАПТИВНЫЕ НЕЙРОННЫЕ ФИЛЬТРЫ

Для преодоления недостатков и ограничений метода формирования АФ введена новая структура нелинейных фильтров, называемая ОАНФ. Совокупность ОАНФ представляет большую группу нелинейных цифровых фильтров для подавления шумов общего вида.⁴⁾

Полное представление структуры и свойств ОАНФ использует ряд элементов формализации, которые ниже сгруппированы в три раздела.

1) На входной последовательности $r(n)$, где $r(n) \in \{0, 1, \dots, M-1\}$, действует скользящее окно шириной b (временных отсчетов), которое формирует входной вектор $r_b(n)$ для ОАНФ. В свою очередь ОАНФ создает на выходе вектор $\hat{s}(n)$ — оценку желаемого вектора $s(n)$. Вектор $r_b(n)$ определяется соотношением

³⁾ Это определение стекового фильтра (более общее сравнительно с ²⁾) будет в дальнейшем использоваться в статье. Отметим, что число положительных булевых функций (ПБФ) при трех переменных 20, а при пяти около 7.5 тыс. [12], поэтому выбрать оптимальные положительные функции по специальным требованиям достаточно трудно. Это является главным препятствием создания стекового фильтра методом ПБФ.

⁴⁾ Шум общего вида предполагает отсутствие часто делаемых допущений, что шум белый, имеет нулевое среднее и гауссово распределение.

$$\mathbf{r}_b(n) = \left[r\left(n - \frac{b-1}{2}\right), \dots, r\left(n + \frac{b+1}{2}\right) \right],$$

и он может быть представлен пороговым разложением вида:

$$\mathbf{r}_b(n) = \sum_{m=1}^{M-1} T_m[\mathbf{r}_b(n)], \quad (5)$$

где $T_m(\cdot)$ — пороговая функция такая, что

$$T_m[\mathbf{r}_b(n)] = \left\{ T_m\left[r\left(n - \frac{b-1}{2}\right)\right], \dots, T_m[r(n)], \dots, T_m\left[r\left(n + \frac{b+1}{2}\right)\right] \right\}. \quad (6)$$

2) Математическим обозначением для ОАНФ служит выражение $F_{l,b}[\mathbf{X}(n)]$, и оно определяется соотношением

$$\hat{\mathbf{s}}(n) = F_{l,b}[\mathbf{r}_b(n)] \equiv \sum_{i=1}^{M-1} N_i[X_i^{l,b}(n)], \quad (7)$$

где индекс l определяет число соседних уровней выше и ниже текущего уровня, и эти уровни показаны для вычисления выхода фильтра:

$$X_i^{l,b}(n) \equiv \begin{bmatrix} T_{i+l}[\mathbf{r}_b(n)] \\ \dots \\ \dots \\ T_i[\mathbf{r}_b(n)] \\ \dots \\ \dots \\ T_{i-l}[\mathbf{r}_b(n)] \end{bmatrix}. \quad (8)$$

Выход фильтра представляет бинарный массив с $(2l+1) \times b$ входами, полученный из $\mathbf{r}_b(n)$. Тип и структура оператора преобразования (8) неопределены. Это может быть отдельный нейрон, нейронная сеть (НС) с прямым распространением сигнала или НС с радиальными базисными функциями [13]. Но различные аналитические результаты, выведенные для АОФ_НС, не зависят от выбранного оператора, который реализуется нейронной сетью.

3) Распространение стекового свойства на массив, позволяет ввести еще один элемент формализации для последующего представления структуры и свойств ОАНФ. При двух бинарных массивах A и B размера $n \times m$ считается, что A является стеком для B , когда $A \leq B$ и $a(i, j) \leq b(i, j)$ для всех i и j , т. е. из $a(i, j) = 1$ следует что $b(i, j) = 1$.

Оператор порогового разложения, примененный к вектору $\mathbf{r}_b(n)$, обладает стековым свойством:

$$T_m[\mathbf{r}_b(n)] \geq T_l[\mathbf{r}_b(n)] \text{ для } m \leq l. \quad (9)$$

Таким же образом бинарный пороговый массив $X_i^{l,b}(n)$ тоже представляет стековое свойство, как это показывает соотношение (10):

$$X_1^{l,b}(n) \geq X_2^{l,b}(n) \geq \dots, X_{M-1}^{l,b}(n). \quad (10)$$

Рассмотренные выше элементы формализации показывают, что концепция стекового свойства и порогового разложения распространены на операции с массивами. Вместо положительных булевых функций, нейронные операторы адаптируются на каждом бинарном уровне. Далее будут показаны некоторые важные свойства и преимущество такого обобщения.

Структуры обобщенных адаптивных нейронных фильтров

Пример принципиальной структуры ОАНФ с шириной окна 3, использующей три настраиваемых уровня, показан на рис. 1. Для простоты выход каждого нейронного оператора превращается в бинарную форму, проходя ограничитель с "жестким" порогом.

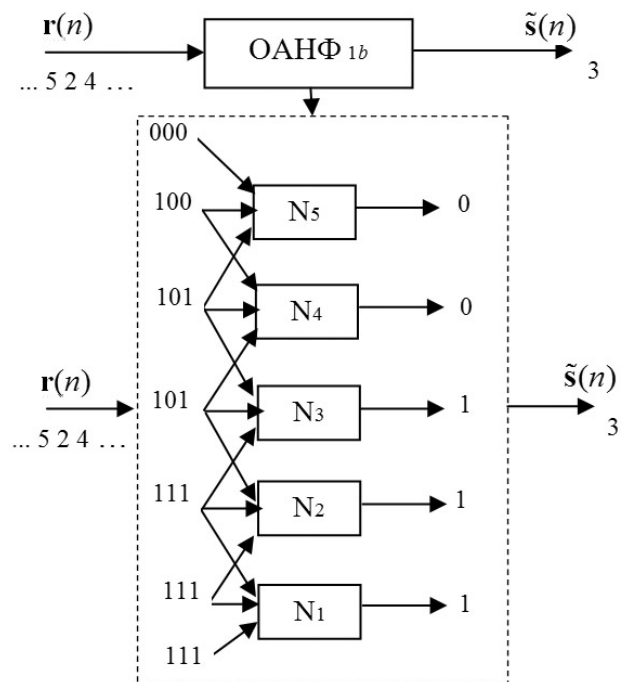


Рис. 1. Обобщенный адаптивный нейронный фильтр с окном ширины 3

Из анализа предыдущего раздела следует, что прохождение M -последовательности сигнала через ОАНФ эквивалентно следующей процедуре.

1. Разложение M значений входного сигнала в набор $M-1$ бинарных сигналов. Бинарный сигнал на уровне i (где i — целое число из набора $\{1, 2, \dots, M-1\}$) получается после порогового ограничения входного сигнала на уровне i , как при стековой фильтрации.

2. Присвоение бинарных последовательностей для уровней из диапазона от 1 до $M-1$. Бинарным последовательностям, которые выше диапазона ($\geq M$), присваивается 0, а тем, которые ниже диапазона (< 1), присваивается 1. Это показано на рис. 1.

3. Передача каждой бинарной последовательности на вход настраиваемых уровней выше и ниже текущего уровня на НС. Выходная величина каждого нейронного оператора является либо непрерывной (т. е. $\hat{s} \in [0, 1]$), либо дискретной (т. е. $\hat{s} \in \{0, 1\}$) в зависимости от нейронного оператора.

4. Сложение выходов нейронного оператора на каждом бинарном уровне. В отличие от стековой фильтрации на выходе ОАНФ стековое свойство может не сохраниться.

Значение термина "обобщенный" применительно к ОАНФ

В [14] показано, что CAO между выходом стекового фильтра $F(\cdot)$ (на входной процесс $(\mathbf{r}_b(n) \in \Omega^b, \text{ где } \Omega \in \{0, 1\})$) и желаемым выходом $s(n)$ для окна ширины b может быть выражен соотношением

$$\begin{aligned} \text{CAO} &= E[|s(j) - F(\mathbf{r}_b(j))|] = \\ &= \sum_{j=1}^{2^b} \sum_{k=1}^{M-1} [\tilde{f}(w_j)\pi_k(1, w_j) + f(w_j)\pi_k(0, w_j)]. \end{aligned} \quad (11)$$

В выражении (11) w_j — бинарный образ размерности b , а f и \tilde{f} — положительная булева функция и ее дополнение соответственно:

$$f(w_j) + \tilde{f}(w_j) = 1. \quad (12)$$

Кроме того, $\pi_i(0, w_j)$ и $\pi_i(1, w_j)$ — это совместные вероятности события, что бинарный образ w_j наблюдается в пороговом разложении входного сигнала (в окне) на уровне i , а величина фактического сигнала не менее чем i соответственно.

Согласно свойству порогового разложения стекового фильтра, максимальная абсолютная ошибка (CAO) стекового фильтра, определенная соот-

ношением (11), на каждом из бинарных уровней может быть представлена суммой CAO положительных булевых функций $f(\cdot)$:

$$\text{CAO} = \sum_{i=1}^{M-1} \text{CAO}_i, \quad (13)$$

где

$$\text{CAO}_i = \sum_{j=1}^{2^b} [\tilde{f}(w_j)\pi_k(1, w_j) + f(w_j)\pi_k(0, w_j)]. \quad (14)$$

Поэтому минимизация CAO стекового фильтра эквивалентна минимизации выхода бинарного оператора на каждом уровне, поскольку на каждом бинарном уровне CAO неотрицательна ($\text{CAO}_i \geq 0$ при $\forall i$):

$$\min \text{CAO} = \min \sum_{i=1}^{M-1} \text{CAO}_i = \sum_{i=1}^{M-1} \min \text{CAO}_i. \quad (15)$$

CAO стекового фильтра была введена и формализована при допущении, что $\pi_i(x, w_k) = \pi_j(x, w_k)$, $x = 0$ или 1 при $\forall i, j$. Следовательно, оптимальный стековый фильтр может быть получен путем минимизации CAO_i на каждом бинарном уровне при условии, что $f(w_j) \geq f(w_k) = 1$ для $w_j \geq w_k$. В общем $\pi_i(0, w_k) \neq \pi_j(0, w_k)$ и $\pi_i(1, w_k) \neq \pi_j(1, w_k)$ при $i \neq j$. Поэтому улучшение фильтра достигается заменой одинаковых положительных булевых функций $f(\cdot)$ на всех уровнях в (14) на различные нейронные операторы $N_i(\cdot)$ (на каждом уровне). В этом случае CAO на каждом уровне выражается соотношением

$$\text{CAO}_i = \sum_{j=1}^{2^b} [\tilde{N}_i(w_j)\pi_k(1, w_j) + N_i(w_j)\pi_k(0, w_j)], \quad (16)$$

где $N_i(\cdot) \in [0, 1]$, а $\tilde{N}_i(\cdot) = 1 - N_i(\cdot)$.

Поскольку $N_i(w_j) \neq f(w_j)$, то

$$\begin{aligned} \min \text{CAO}_i &= \\ &= \min \sum_{j=1}^{2^b} [N_i(w_j)\pi_i(1, w_j) + N_i(w_j)\pi_i(0, w_j)] \neq \\ &\neq \min \sum_{j=1}^{2^b} [\tilde{f}_i(w_j)\pi_i(1, w_j) + f_i(w_j)\pi_i(0, w_j)]. \end{aligned} \quad (17)$$

После того как принимаются нейросетевые операторы, ОАНФ (при оптимизации с $I = 0$) может стать оптимальным стековым фильтром. Но для этого требуется выполнение двух условий (при которых $N_i(\cdot) = f_i(\cdot), \forall i$):

$$1) \pi_l(1, w_i) \geq \pi_l(1, w_j) \text{ только при } w_i \geq w_j; \quad (18)$$

$$2) \pi_l(1, w_i) = \pi_m(1, w_j) \text{ для всех } m \neq l. \quad (19)$$

Практически, выполнение условий (18) и (19) трудно гарантировать. Поэтому лучшее функцио-

нирование ОАНФ следует ожидать, если САО на каждом бинарном уровне может быть минимизировано в соответствии с соотношением (17) (без применения ограничений условиями (18) и (19)). Следует отметить, что стековое свойство (при использовании процедуры минимизации САО на каждом уровне по (17)) может быть утрачено. Тем не менее в следующем разделе будет показано, что у ОАНФ есть САО, которая ограничена величиной САО стекового фильтра, показанного в (13).

Подводя итоги, можно отметить, что для M значений входного сигнала (в окне "длины" M) ОАНФ может быть создан путем использования $M-1$ нейросетевых операторов для бинарных уровней ($M-1$) значений порогового бинарного разложения таким образом, чтобы была минимизирована средняя абсолютная ошибка на каждом уровне.

Теперь для более полной интерпретации сделанного выше утверждения целесообразно привести три положения.

1. При условии, что оператор на каждом бинарном уровне является линейной функцией дискриминации вида

$$N_i(\mathbf{X}^i) = a_0^i + \sum_{j=1}^b a_j^i x_j^i, \quad (20)$$

выход ОАНФ определяется соотношением

$$\sum_{i=1}^{M-1} N_i(\mathbf{X}^i) = \sum_{i=1}^{M-1} \left[a_0^i + \sum_{j=1}^b a_j^i x_j^i \right], \quad (21)$$

где a_j^i и x_j^i (для $i = 1, 2, \dots, M-1$ и $j = 1, 2, \dots, b$) являются весами и компонентами входного вектора \mathbf{X}^i с окном шириной b на бинарном уровне i . ОАНФ становится суммированием линейных функций всех бинарных уровней.

2. Если для порогового разделения функции дискриминации используется "жесткий" ограничитель, то выполняется соотношение

$$N_i(\mathbf{X}^i) = U \left[a_0 + \sum_{j=1}^b a_j x_j \right], \quad (22)$$

где x_j равно либо 0, либо 1, а $U[\cdot]$ — это функция жесткого ограничения. В этом случае каждый нейронный оператор соответствует определенной булевой функции. Если нейронные операторы образуют стековый набор булевых функций (см. далее в (42)), то результирующий ОАНФ является обобщенным стековым фильтром [15].

3. Если использовано жесткое ограничение и все веса $a_j \geq 0$ для $j = 0, 1, \dots, b$ на всех бинарных уровнях, то каждый нейронный оператор становится положительной булевой функцией. В этом случае ОАНФ обладает свойством упорядочивания в форме (23):

$$\mathbf{r}_b(n) \leq \tilde{\mathbf{r}}_b(n) \Rightarrow F_{l,b}[\mathbf{r}_b(n)] \leq F_{l,b}[\tilde{\mathbf{r}}_b(n)]. \quad (23)$$

Различие между ОАНФ и стековыми фильтрами состоит в отличии ряда их свойств.

– Выход ОАНФ не может обладать стековым свойством.

– Соседние уровни сигнала не могут подаваться на нейронные операторы на каждом уровне.

– ОАНФ создается путем определения нейронных операторов на каждом отдельном уровне, в то время как стековый фильтр создается путем определения булевых функций на каждом уровне.

Помимо этого различия ОАНФ и стековых фильтров оба типа фильтров обладают структурой порогового разложения, и они превращают оператор с M значениями во входной оператор с $M-1$ бинарными значениями.⁵⁾

Конфигурация обобщенных адаптивных нейронных фильтров

Для ОАНФ целесообразно получить оценку средней абсолютной ошибки (САО). При сравнении САО фильтров типа ОАНФ с аналогичной величиной для стековых фильтров оказывается, что САО фильтров типа ОАНФ ограничена сверху значением САО соответствующих стековых фильтров.

Средняя абсолютная ошибка обобщенных адаптивных нейронных фильтров

При обозначении через $V[F(\cdot)]$ критерия оптимизации в виде САО и использовании стековых свойств фильтра справедливо соотношение

$$\begin{aligned} V[F(\cdot)] &\equiv E \|\mathbf{s}(n) - \tilde{\mathbf{s}}(n)\| = \\ &= E \left\{ \sum_{k=1}^{M-1} [s_k(n) - \tilde{s}_k(n)] \right\} = \left\{ \sum_{k=1}^{M-1} E [s_k(n) - \tilde{s}_k(n)] \right\}, \quad (24) \end{aligned}$$

где $\tilde{s}_k(n) \equiv F[\mathbf{X}_k(n)]$ — выход положительной булевой функции при входном векторе \mathbf{X}_k размерности b на k -м бинарном уровне, а $s_k(n) = T_k[s(n)]$ — желаемая бинарная величина на k -м бинарном уровне.

Критерий САО-оптимизации ОАНФ, для которого принято обозначение $C[F_{l,b}(\cdot)]$, определяется соотношением

$$C[F_{l,b}(\cdot)] \equiv E[s(n) - \tilde{\mathbf{s}}(n)] = E \left[\sum_{k=1}^{M-1} [s_k(n) - \tilde{s}_k(n)] \right], \quad (25)$$

где в этом случае $\tilde{\mathbf{s}}(n)$ — выход ОАНФ.

Следует отметить, что ОАНФ не обязательно обладает стековым свойством, т. е., вообще говоря, оно не выполняется:

⁵⁾ Именно это облегчает их реализацию в микроэлектронном исполнении.

$$C[F_{l,b}(\cdot)] \equiv E[s(n) - \tilde{s}(n)] \neq E \left| \sum_{k=1}^{M-1} [s_k(n) - \tilde{s}_k(n)] \right|, \quad (26)$$

где в этом случае $\tilde{s}(n)$ — выход нейронного оператора $N_k[X_k^{l,b}(n)]$.

Верхняя граница средней абсолютной ошибки обобщенного нейронного фильтра

Для получения верхней границы САО ОАНФ целесообразно ввести ряд обозначений, которые будут ниже использоваться на правах определений.

В том случае если критерий САО, обозначенный как $C[F_{l,b}(\cdot)]$, не превышает некоторой константы B , то B считается верхней границей САО, границей $C[F_{l,b}(\cdot)]$. Дополнительно удобно обозначить САО для ОАНФ на уровне i через $G[F_{l,b}(\cdot)]$, т. е. принять, что $G_i[F_{l,b}(\cdot)] = E[|s_i(n) - \tilde{s}_i(n)|]$.

Использование обозначений позволяет показать, что сумма САО на каждом уровне ОАНФ (т. е. $G[F_{l,b}(\cdot)]$) является верхней границей для САО у ОАНФ (выше она обозначена, как $C[F_{l,b}(\cdot)]$):

$$C[F_{l,b}(\cdot)] \leq \sum_{i=1}^{M-1} G_i[F_{l,b}(\cdot)] = G[F_{l,b}(\cdot)]. \quad (27)$$

Соотношение (27)⁶⁾ показывает, что величина САО для ОАНФ ограничена сверху значением $G[F_{l,b}(\cdot)]$.

Оптимальным фильтром типа ОАНФ является такой, в котором минимизирована величина $C[F_{l,b}(\cdot)]$. Но минимизировать $C[F_{l,b}(\cdot)]$ непосредственно в структуре порогового разложения N_k очень трудно. Вместо минимизации $C[F_{l,b}(\cdot)]$ оказывается возможным минимизировать САО нейронного оператора отдельно на каждом бинарном уровне.

Принимая, что статистики сигнала, шума и процедура, связанная с окном, известны, и учитывая, что в области $\Omega^{(2^{l+1}) \times b}$ ($\Omega \in \{0,1\}$) имеется $2^{(2^{l+1}) \times b}$ различных состояний, можно получить САО каждого нейронного оператора в виде соотношения

$$G[F_{l,b}(\cdot)] = \sum_{k=1}^{M-1} E[|s_k(n) - \tilde{s}_k(n)|] =$$

⁶⁾ Обоснованием неравенства (27) может служить понятная взаимосвязь последовательности выражений:

$$C[F_{l,b}(\cdot)] = E \left| \sum_{k=1}^{M-1} [s_k(n) - \tilde{s}_k(n)] \right| \leq \\ \leq E \sum_{k=1}^{M-1} |s_k(n) - \tilde{s}_k(n)| = \sum_{k=1}^{M-1} E[|s_k(n) - \tilde{s}_k(n)|] = G[F_{l,b}(\cdot)].$$

$$= \sum_{k=1}^{M-1} \sum_{\mathbf{w}_j \in \Omega^{(2^{l+1}) \times b}} [N_k(\mathbf{w}_j) \pi_k(0 | \mathbf{w}_j) + N_k(\mathbf{w}_j) \pi_k(1 | \mathbf{w}_j)], \quad (28)$$

где $\pi_k(i|\mathbf{w}_j)$ — вероятность того, что истинный сигнал равен i при условии, что \mathbf{w}_j наблюдается с входа на уровне k , а $\pi_k(\mathbf{w}_j)$ — предельная вероятность иметь входное состояние \mathbf{w}_j , наблюдаемое на уровне k .⁷⁾

Средняя абсолютная ошибка (САО) оптимизированного ОАНФ (для любых данных сигналов, шуме и преобразовании окном) не превосходит САО оптимального стекового фильтра (см. Приложение, примечание 1).

Если ОАНФ, $F_{l,b}[\mathbf{r}_b(n)]$, оптимизировано путем минимизации САО на каждом уровне, а также если из $\mathbf{w}_j \geq \mathbf{w}_i$ следует $\pi_k(1, \mathbf{w}_j) \geq \pi_k(1, \mathbf{w}_i)$, то оператор $N_k(\cdot)$ на каждом уровне должен быть положительной булевой функцией для $N_k(\cdot) \in \{0,1\}$ (см. Приложение, примечание 2). Надо отметить, что даже ОАНФ на каждом отдельном уровне можно реализовать на основе положительных булевых функций, но ОАНФ в целом не обязательно обладает стековым свойством.

Упрощение обобщенного адаптивного нейронного фильтра

Средние абсолютные ошибки (САО) идентичны на различных пороговых уровнях. Однако разумно предположить, что статистики на смежных бинарных уровнях сходные, так что дальнейшее упрощение или обобщение структуры ОАНФ возможно состоит в понижении вычислительной сложности и упрощении структуры для последующего выполнения его на микропроцессорном уровне.

Для снижения вычислительной сложности предполагается, что в небольшом интервале со-

⁷⁾ Соотношение (28) следует из определений $\tilde{N}_k(\cdot)$, $N_k(\cdot)$, $\pi_k(\cdot|\cdot)$ и САО. Для данного входного образа \mathbf{w}_j : $N_k(\mathbf{w}_j) + \tilde{N}_k(\mathbf{w}_j) = 1$ при $\forall j$. Согласно (16), САО на уровне i у ОАНФ определяется соотношением:

$$G_i[F_{l,b}(\cdot)] = \sum_{\mathbf{w}_j \in \Omega^{(2^{l+1}) \times b}} [N_i(\mathbf{w}_j) \pi_j(0, \mathbf{w}_j) + \tilde{N}_i(\mathbf{w}_j) \pi_i(1, \mathbf{w}_j)].$$

Так что справедливо соотношение:

$$\min \{G[F_{l,b}(\cdot)]\} = \sum_{k=1}^{M-1} \min \{G_k[F_{l,b}(\cdot)]\}.$$

В общем случае: $\pi_k(1, \mathbf{w}_j) \neq \pi_l(1, \mathbf{w}_j)$ при $k \neq l$. Поэтому нейронный оператор, использованный для минимизации САО на каждом бинарном уровне, для различных уровней не может быть одним и тем же, т. е. $N_k(\mathbf{w}_j) \neq N_l(\mathbf{w}_j)$ при $k \neq l$.

пряженных уровней вероятности $\pi(i, \mathbf{w}_j)$ приблизительно одинаковые, т. е. $\pi_k(i, \mathbf{w}_j) \approx \pi_{k+l}(i, \mathbf{w}_j)$ для $l \leq L$, где $i = 0$ или 1 , а $L \in \{0, 1, \dots, M-1\}$. При этом L представляет число соседних уровней порога, вероятности которых принимаются примерно одинаковыми. Следовательно, выполняется приближенное соотношение

$$N_k(\mathbf{w}_j) \approx N_{k+l}(\mathbf{w}_j) \text{ для } l \leq L,$$

а нейронные операторы идентичны на уровнях от k -го до $(k + l)$ -го. Поэтому приближенно равны между собой значения САО на уровнях от k -го до $(k + l)$ -го, т. е. $G_k(F_{l,b}) \approx G_{k+l}(F_{l,b})$ для $l \leq L$.

Промежуточным итогом этого раздела могут служить три положения (или три факта), которые непосредственно следуют из проведенного выше анализа.

1. Если $L = M - 1$, то $N_1 = N_2 = \dots = N_{M-1}$. В этом случае ОАНФ, $F_{l,b}(\cdot)$, можно считать однородным. Если ОАНФ однородный и нейронный оператор — это положительная булева функция, то этот ОАНФ является стековым фильтром.

2. Если (когда) $L = 0$, то нейрон на одном уровне не может быть одинаковым с нейроном на любом другом уровне. В этом случае $F_{l,b}(\cdot)$ считается неоднородным ОАНФ.

3. Если L — константа, обозначающая, что для смежных уровней принимаются приблизительно равные априорные вероятности, и $1 < L < M - 1$, тогда $N_1 = N_2 = \dots = N_L$; $N_{L+1} = N_{L+2} = \dots = N_{2L}$; ...; $N_{kL+1} = N_{kL+2} = \dots = N_{M-1}$, где $k = \lfloor (M - 1) / L \rfloor$ — положительное целое число.⁸⁾ В этом случае нейроны, находящиеся в L соседних пороговых уровнях, локально идентичны, и при этом считается, что ОАНФ является полуюднородным.

Структура наиболее простой формы ОАНФ показана на рис. 2.

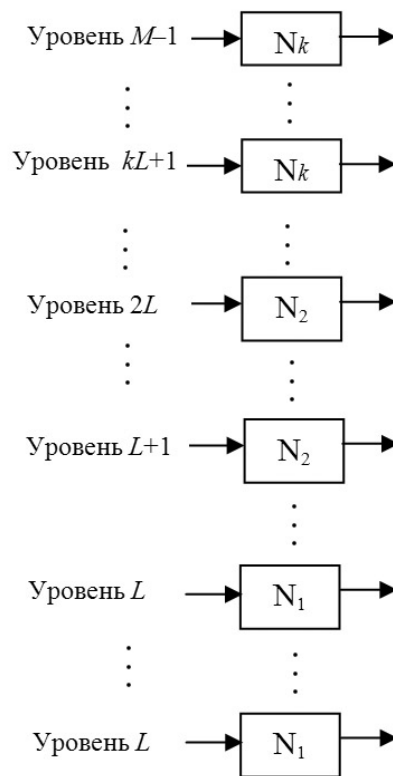


Рис. 2. Упрощенный ОАНФ

РЕАЛИЗАЦИЯ ОАНФ С ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТЬЮ

Рассмотрение целесообразно начать со структуры ОАНФ с одним нейроном (включая ее обучение), поскольку нейронный оператор на каждом уровне ОАНФ в общем случае может быть выполнен в виде гораздо большей нейронной сети. Для простоты на рис. 3 показана структура с нейроном, имеющим квадратичную передаточную функцию, т. к. математически она эквивалентна квадратичной решающей (так называемой дискриминантной) функции для критерия, который при класси-

⁸⁾ Способ определения числа L расположенных соседних уровней у ОАНФ_НС более подробно описан в [8].

фикации образов имеет более высокую способность разделения [16, 17], чем линейная дискриминантная функция.

Удобно обозначить через $\mathbf{X}(n)$ входную матрицу, которая подана в нейрон при n -м отсчете окна с шириной b для входной последовательности $r(n)$. Квадратичная дискриминантная функция критерия (g) определяется соотношением

$$g[\mathbf{X}(n)] = \sum_{j=1}^B \tilde{a}_{jj} x_j^2 + \sum_{j=1}^{B-1} \sum_{k=j+1}^B \tilde{a}_{jk} x_j x_k + \sum_{j=1}^B \tilde{a}_j x_j + \tilde{a}_{B=1}, \quad (29)$$

где B — число элементов в $\mathbf{X}(n)$, а $\tilde{a}_j, \tilde{a}_{jk}$ — веса. Поскольку $x_j \in \{0, 1\}$ для $j = 1, 2, \dots, B$, то вследствие этого $x_j^2 = x_j$.

Соотношение (29) может быть записано в другой форме:

$$g[\mathbf{X}(n)] = a_0 + a_1 f_1 + a_2 f_2 + \dots + a_{K-1} f_{K-1} = \mathbf{A}^T \mathbf{F}(n), \quad (30)$$

где \mathbf{A}^T — вектор-строка, транспонированная форма вектора-столбца с элементами a_0, a_1, \dots, a_{K-1} ; $\mathbf{F}(n) = [1, f_1, f_2, \dots, f_{K-1}]^T$, $f_i \in \{x_j, x_j x_m \text{ для } j \neq m\}$; $K = B(B+1)/2 + 1$; a_i — новые веса. Поэтому для

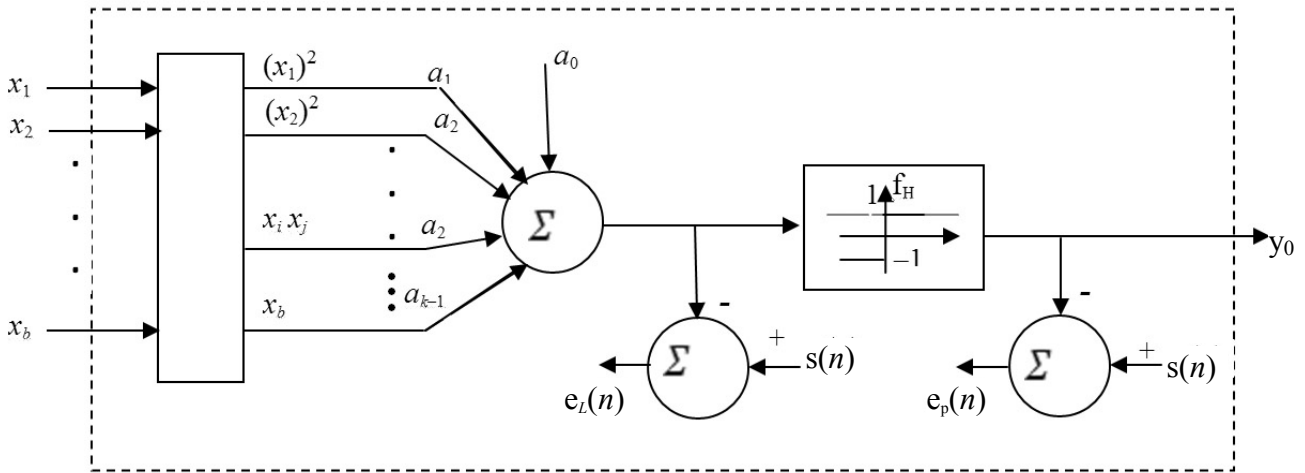


Рис. 3. Использование искусственного нейрона для квадратичной функции критерия

окна с шириной b и входной матрицей размером $(2I + 1) \times b$ количество весов, требуемых для квадратичной дискриминантной функции, составляет K :

$$K = 0.5 \cdot [(2I + 1) \times b] \cdot [(2I + 1) \times b + 1] + 1, \quad (31)$$

где I — число соседних уровней, находящихся выше или ниже текущего уровня.

Можно заметить, что $f_i \in \{x_j, x_j x_m\}$ для $i = 1, 2, \dots, K - 1$ и $j, m = i = 1, 2, \dots, b$, как это показано на рис. 3.

Из существующих методов адаптивной подстройки весов элементов нейронной сети целесообразно остановиться на методе наименьших средних квадратов (НСК) и методе обучения персептрона (эти методы основаны на градиентном спуске [18, 19]).

Метод наименьших средних квадратов, обучение с учителем

Для настраиваемых в (30) весов на n -й итерации процесса обучения вводится обозначение в виде вектора-столбца, как

$$[a_0(n), a_1(n), \dots, a_{K-1}(n)]^T.$$

Таким образом, $\mathbf{A}(n) = [a_0(n), a_1(n), \dots, a_{K-1}(n)]^T$.

Во время процесса фильтрации вместе с обычными разветвленными по входу элементами входного сигнала подается дополнительный сигнал $s(n)$, называемый желаемым откликом. Отклик желаемого сигнала обеспечивает опорный фрейм для подстройки весовых коэффициентов фильтра.

Ошибка оценки $e_L(n)$, производимая во время обучения методом НСК (рис. 3), определяется соотношением

$$e_L(n) = s(n) - \mathbf{A}^T(n)\mathbf{F}(n), \quad (32)$$

где $\mathbf{A}^T(n)\mathbf{F}(n)$ — произведение вектора $\mathbf{A}(n)$ (разветвленных по входу весовых коэффициентов) и таких же разветвленных компонент входного вектора $\mathbf{F}(n)$.

Согласно градиентному методу (называемому методом крутого спуска) [18], обновленные величины вектора весов на $(n + 1)$ -й итерации могут быть определены с использованием рекуррентного соотношения

$$\begin{aligned} \mathbf{A}(n + 1) &= \mathbf{A}(n) + \alpha \mathbf{F}(n)[s(n) - \mathbf{F}^T(n)\mathbf{A}(n)] = \\ &= \mathbf{A}(n) + \alpha \mathbf{F}(n)e_L(n), \end{aligned} \quad (33)$$

где α — параметр скорости обучения.

Обучение персептрона, обучение с учителем

Как видно на рис. 3, ошибка $e_p(n)$ образуется после прохождения выхода $y(n)$ через функцию жесткого ограничения f_H . Так что выход y_0 определяется выражением $y_0(n) = f_H(\mathbf{F}^T(n)\mathbf{A}(n))$.

Подобным образом, при методе градиентного спуска правило обучения персептрона выражается достаточно простым рекуррентным соотношением:

$$\mathbf{A}(n + 1) = \mathbf{A}(n) + \alpha \mathbf{F}^T e_p(n), \quad (34)$$

где

$$e_p(n) = s(n) - f_H(\mathbf{F}^T(n)\mathbf{A}(n)). \quad (35)$$

ЗАКЛЮЧЕНИЕ

Рассмотрены основные этапы и их аналитическая основа, определяющие создание нового класса нелинейных адаптивных фильтров, в структуру которых входят нейронные элементы. Проанализированы свойства и особенности разновидностей стековых фильтров, модификация которых предшествовала появлению обобщенных адаптивных нейронных фильтров (ОАНФ).

Совокупность ОАНФ представляет большую группу нелинейных цифровых фильтров для подавления шумов общего вида. Рассмотрена структура ОАНФ и показатель точности его функционирования — средняя абсолютная ошибка (разность между желаемым и фактическим выходом фильтра). Для этой величины приведена оценка ее верхней границы.

Рассмотрены упрощенная форма ОАНФ, реализация его на основе искусственной нейронной сети и методы обучения фильтра.

ОАНФ более эффективны, чем многие нелинейные фильтры (даже в предположении белого гауссова шума). ОАНФ обладают определенной мерой робастности, т. е. могут использоваться для идентификации и управления динамическими системами в нестационарных условиях относительно вида и уровня шума.

Изучение основных этапов построения ОАНФ и сравнение их с другими видами адаптивных фильтров позволяет отметить их основные особенности и характеристики.

– Многообразие ОАНФ включает достаточно обширный класс нелинейных фильтров со скользящим окном, к которому относятся стековые фильтры.

– Средняя абсолютная ошибка (САО) оптимального ОАНФ не превосходит САО оптимального стекового фильтра.

– Алгоритмы получения структуры ОАНФ различной сложности эффективны и в значительной степени независимы от уровня и вида шума. При обучении ОАНФ на основе нейронной сети в условиях реальных уровней и типов шума функционирование фильтра практически робастно.

– При использовании ОАНФ достигается лучшее подавление шума, чем у многих традиционных фильтров.

– Конфигурация ОАНФ может быть выбрана адаптивно для функционирования в различной шумовой среде. Поэтому эти фильтры хорошо приспособлены ко многим прикладным задачам адаптивной фильтрации.

Концепция и метод построения нелинейных адаптивных фильтров с введением в их структуру нейронных элементов принадлежат Йину, Астоле и Нейво (Yin, Astola, Neuvo [5, 6]). Критерии и методы оптимизации фильтров прототипов для

ОАНФ содержат работы Ансари, Цханга (Ansari, Zhang [7]) и Ханека, Ансари (Hanek, Ansari [8]).

ПРИЛОЖЕНИЕ

Примечание 1

Из очевидного неравенства треугольника $|a + b| \leq |a| + |b|$ и использования (25), (27) прямым следствием является цепочка взаимосвязанных соотношений:

$$\begin{aligned} C[F_{I,b}(\cdot)] &= E \left| \sum_{k=1}^{M-1} [s_k(n) - \tilde{s}_k(n)] \right| \leq \sum_{k=1}^{M-1} E |s_k(n) - \tilde{s}_k(n)| = \\ &= \sum_{k=1}^{M-1} \sum_{\mathbf{w}_j \in \Omega(2^{l+1}) \times b} [N_k(\mathbf{w}_j) \pi_k(0 | \mathbf{w}_j) + \\ &\quad + \tilde{N}_k(\mathbf{w}_j) \pi_k(1 | \mathbf{w}_j)] = G[F_{I,b}(\cdot)]. \end{aligned}$$

Поскольку для функции $N_k(\cdot)$ нет ограничения, что она положительная булева функция или стековый набор булевых функций, то

$$\begin{aligned} \min MAO_k &= \min \sum_{\mathbf{w}_j \in \Omega(2^{l+1}) \times b} [N_i(\mathbf{w}_j) \pi_i(0, \mathbf{w}_j) + \\ &\quad + \tilde{N}_i(\mathbf{w}_j) \pi_i(1, \mathbf{w}_j)] \leq \\ &\leq \min \sum_{\mathbf{w}_j \in \Omega(2^{l+1}) \times b} [f_k(\mathbf{w}_j) \pi_i(0, \mathbf{w}_j) + \tilde{f}_k(\mathbf{w}_j) \pi_i(1, \mathbf{w}_j)] \leq \\ &\leq \min \sum_{\mathbf{w}_j \in \Omega(2^{l+1}) \times b} [f(\mathbf{w}_j) \pi_i(0, \mathbf{w}_j) + \tilde{f}(\mathbf{w}_j) \pi_i(1, \mathbf{w}_j)]. \end{aligned}$$

Здесь $f_k(\cdot)$ соответствует k -й функции стекового набора булевых функций обобщенного стекового фильтра, который удовлетворяет условию: $f_k(\mathbf{w}_k) \leq f_{k-1}(\mathbf{w}_{k-1})$, где \mathbf{w}_k и \mathbf{w}_{k-1} в этом случае — наблюдаемые образы на k -м и $(k-1)$ -м уровне соответственно. Кроме того, поскольку $f(\cdot)$ является положительной булевой функцией стекового фильтра, то $f(\mathbf{w}_j)$ нежно действует только на b -й бит y $[(2l-1)b]$ -битового образа \mathbf{w}_j . Таким образом,

$$\min C[F_{I,k}(\cdot)] \leq \min G[F(\cdot)] \leq \min H[F(\cdot)] \leq \min B[F(\cdot)],$$

где $[F(\cdot)]$ и $B[F(\cdot)]$ — САО обобщенного стека и стековый фильтр соответственно.

Примечание 2

На основании соотношения (27) можно заключить, что если $G_k[F_{I,b}(\cdot)]$ минимизировано, то из неравенства $\pi_k(1, \mathbf{w}_j) \geq \pi_k(1, \mathbf{w}_i)$ следует $N_k(\mathbf{w}_j) \geq N_k(\mathbf{w}_i)$. При условии, что $\mathbf{w}_j \geq \mathbf{w}_i$, $N_k(\cdot)$ является положительной булевой функцией.

СПИСОК ЛИТЕРАТУРЫ

1. Haykin S. Adaptive filter Theory. Englewood Cliffs, N.Y., Prentice Hall, 1991. 854 p.
2. Бендат Дж., Пирсол А. Прикладной анализ случайных данных. М.: Мир, 1989. 540 с.
3. Wendt P.D., Coile E.J., Gallagher N.C. // IEEE Transactions on Acoustics: Speech and Signal Processing. 1986. V. 34, N 8. P. 898–911.

4. *Lin J., Sellke T.M., Coile E.J.* Adaptive stack filtering under mean absolute error criterion // *IEEE Transactions on Acoustics: Speech and Signal Processing*. 1990. V. 38, N 6. P. 938–954.
5. *Yin L., Astola J., Neuvo Y.* A new class of filters – neural filters // *IEEE Transactions on Signal Processing*. 1993. V. 41, N 3. P. 1201–1222.
6. *Yin L., Astola J., Neuvo Y.* Adaptive multistage weighted order statistic filters based on the backpropagation algorithm // *IEEE Transactions on Signal Processing*. 1994. V. 42, N 2. P. 419–422.
7. *Ansari N., Zhang Z.* Generalized adaptive neural filters // *IEEE Electron Letters*. 1993. V. 20, N 4 (February). P. 342–343.
8. *Hanek H., Ansari N.* Speeding up the generalized adaptive neural filters // *IEEE Transactions on Image Processing*. 1996. V. 5, N 5. P. 705–712.
9. *Меркушева А.В.* Фильтрация нестационарного сигнала (речи) в вейвлет-области с адаптацией к виду и динамике шума // *Научное приборостроение*. 2003. Т. 13, № 2. С. 73–87.
10. *Gilbert E.N.* Lattice – theoretic properties of frontal switching functions // *Journal of Mathematical Physics*. 1954. V. 33. P. 57–67.
11. *Hampson S.E., Volper D.J.* Disjunctive models of Boolean category learning // *Biological Cybernetics*. 1987. V. 56. P. 121–137.
12. *Coile E.G., Lin G.H.* Stack filters and the mean absolute error criterion // *IEEE Transactions on Acoustics: Speech and Signal Processing*. 1988. V. 36, N 6. P. 1244–1254.
13. *Малыхина Г.Ф., Меркушева А.В.* Сеть с симметричной функцией преобразования нейронов на основе статистик высокого порядка для снижения шума нестационарного сигнала // *Научное приборостроение*. 2008. Т. 18, № 4. С. 129–136.
14. *Zeng H., Zhou H., Neuvo Y.* FIR stack filters // *Optical Engineering*. 1991. V. 30, July. P. 965–975.
15. *Lin J., Coile E.J.* Minimum mean absolute error estimation over the class of generalized stack filters // *IEEE Transactions on Acoustics: Speech and Signal Processing*. 1990. V. 38, N 4. P. 663–678.
16. *Ту Дж., Гонсалес Р.* Принципы распознавания образов. М.: Мир, 1978. 411 с.
17. *Фукунага К.* Введение в статистическую теорию распознавания. М.: Наука, 1979. 367 с.
18. *Haykin S.* *Neural Networks. A comprehensive foundation.* Prentice Hall, 1999. 842 p.
19. *Меркушева А.В.* Нейросетевые методы обработки сигналов в информационных системах: Элементы структуры, принципы обучения и мера многообразия отображения информации // *Информационные технологии*. 2005. № 3. С. 9–20.

СПбГТУ, Санкт-Петербург

Контакты: *Малыхина Галина Фёдоровна,*
g_f_malychina@mail.ru

Материал поступил в редакцию 22.06.2011.

GENERALIZED ADAPTIVE NEURAL FILTERS AND FILTERS BASED ON NEURAL NETWORK

A. V. Merkusheva, G. F. Malychina

SPbSPU, Saint-Petersburg

The structure and properties of nonlinear adaptive filter group (NAF) are considered. The filters are based on single neurons and on neural networks. The precision parameters — the upper level and mean absolute error of filters are analyzed. Such NAF extend the class of stack filters and have better performance in noise suppression.

Keywords: filters, neural network, precision, signal, noise, adaptation, performance, quality