

УДК 519.7:007.52

Е.В. БОДЯНСКИЙ, Е.А. ВИНОКУРОВА

**АДАПТИВНЫЙ ВЭЙВЛЕТ-НЕЙРОННЫЙ ПРЕДИКТОР**

В работе предложен новый оптимальный по быстродействию алгоритм обучения вэйвлет-нейрона, отличающийся тем, что он обеспечивает настройку в реальном времени не только синаптических весов, но и параметров растяжения и сдвига дочерних вэйвлетов. Алгоритм обладает как следящими, так и сглаживающими свойствами, что позволяет эффективно использовать вэйвлет-нейроны как самостоятельно, так и в составе искусственных нейронных сетей для решения задач прогнозирования, фильтрации, компрессии и классификации нестационарных зашумленных сигналов произвольной природы.

**Введение.**

В последние годы для решения широкого класса задач, связанных с обработкой сигналов различной природы широкое распространение получили искусственные нейронные сети, обеспечивающие высокое качество аппроксимации, прогнозирования, фильтрации и т.п. существенно нелинейных процессов в условиях априорной и текущей неопределенности. Параллельно с нейронными сетями развивался аппарат вэйвлетов (всплесков), являющихся весьма эффективным средством локального представления сигналов как во временной, так и в частотной областях [1-3]. На стыке этих двух концепций возникли новые конструкции, получившие название вэйвлет-нейронные сети и сочетающие в себе гибкость и обучаемость нейронных сетей и возможности компактного описания сигналов, присущие вэйвлетам [4-9]. Интересно отметить, что подобно радиально-базисным нейронным сетям, вэйвлет-нейронные сети реализуют идею полиномиальной ядерной аппроксимации [10] и также, как и радиально-базисные нейронные сети [11] являются универсальными аппроксиматорами [12-14].

Необходимость обработки информации в реальном времени привела к появлению адаптивных вэйвлет-нейронных сетей [15-20], эффективность которых, определяется скоростью сходимости используемых алгоритмов обучения, в качестве которых, как правило, используются градиентные процедуры с подбираемыми эмпирически параметрами шага. Естественно, что низкая скорость обычных градиентных алгоритмов и архитектурная громоздкость большинства вэйвлет-нейронных сетей могут вызывать определенные трудности при обработке быстротекущих процессов.

**1. Вэйвлет-нейрон и алгоритм его обучения.**

Введем в рассмотрение структуру вэйвлет-нейрона, приведенную на рис. 1. Как видно, вэйвлет-нейрон достаточно близок по конструкции к стандартному  $n$ -входовому формальному нейрону, однако вместо обычных настраиваемых синаптических весов содержит вэйвлет-синапсы  $WS_i$ ,  $i = 1, 2, \dots, n$ , обучаемыми параметрами которых являются не только веса  $w_{ji}$ , но и факторы растяжения и сдвига дочерних вэйвлетов  $\varphi_{ji}(x_i(k))$ .

При подаче на вход вэйвлет-нейрона векторного сигнала

$$x(k) = (x_1(k), x_2(k), \dots, x_n(k))^T,$$

где  $k = 0, 1, 2, \dots$  - текущее дискретное время.

на его выходе появляется скалярное значение

$$y(k) = \sum_{i=1}^n f_i(x_i(k)) = \sum_{i=1}^n \sum_{j=1}^{h_i} w_{ji}(k) \varphi_{ji}(x_i(k)),$$

определяемое как настраиваемыми весами  $w_{ji}(k)$ , так и используемыми вэйлет-функциями.

Для использования в составе нейросетей в качестве материнских вэйлетов (mother wavelets) весьма привлекательным представляется семейство полиномиальных оконных функций POLYWOG-wavelets [19] (Polynomials WindOwed with Gaussian types of functions wavelets - POLYWOG-wavelets), простейшей из которых является

$$\varphi_{ji}(x_i(k)) = \sqrt{e} x_i(k) \exp\left(-\frac{x_i^2(k)}{2}\right) \quad (1)$$

Среди прочих POLYWOG-вэйлетов следует отметить также весьма популярную в теории искусственных нейронных сетей функцию «Мексиканская шляпа» («Mexican Hat»), имеющую вид

$$\varphi_{ji}(x_i(k)) = (1 - x_i^2(k)) \exp\left(-\frac{x_i^2(k)}{2}\right)$$

Основной особенностью POLYWOG-функций является то, что все их производные также являются вэйлетами, что, как будет показано ниже, позволяет синтезировать достаточно компактные алгоритмы обучения всех параметров вэйлет-нейрона.

Дочерним вэйлетом функции (1) является

$$\varphi_{ji}(x_i(k)) = \sqrt{e} \tau_{ji}(k) \exp\left(-\frac{\tau_{ji}^2(k)}{2}\right) = t_{ji}(k) \exp\left(-\frac{\tau_{ji}^2(k)}{2}\right),$$

где  $\tau_{ji}(k) = \frac{x_i(k) - c_{ji}(k)}{\sigma_{ji}(k)}$ ;

$c_{ji}(k), \sigma_{ji}(k)$  - параметры, определяющие положение центра (сдвига) и ширину (растяжение).

Записав производные

$$\frac{\partial \varphi_{ji}(x_i(k))}{\partial c_{ji}} = \sqrt{e} \sigma_{ji}^{-1}(k) (\tau_{ji}^2(k) - 1) \exp\left(-\frac{\tau_{ji}^2(k)}{2}\right) = t_{ji}^c(k) \exp\left(-\frac{\tau_{ji}^2(k)}{2}\right),$$

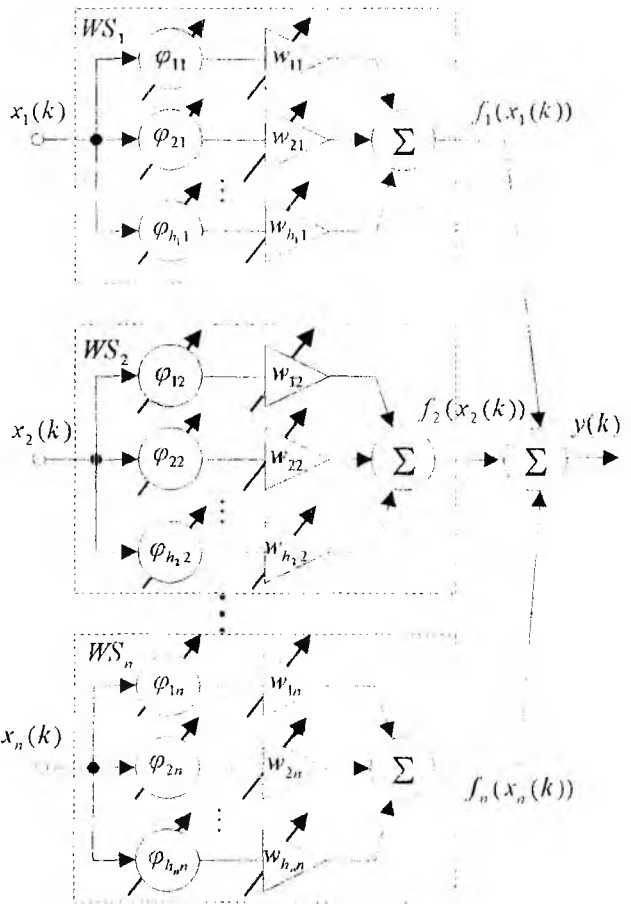


Рис. 1 – Вэйлет-нейрон

$$\frac{\partial \varphi_{ji}(x_i(k))}{\partial \sigma_{ji}^{-1}} = -\sqrt{e}(x_i(k) - c_{ji}(k))(\tau_{ji}^2(k) - 1) \exp\left(-\frac{\tau_{ji}^2(k)}{2}\right) = t_{ji}^{\sigma}(k) \exp\left(-\frac{\tau_{ji}^2(k)}{2}\right),$$

можно заметить, что они также являются вэйвлетами, приведенными на рис. 2.

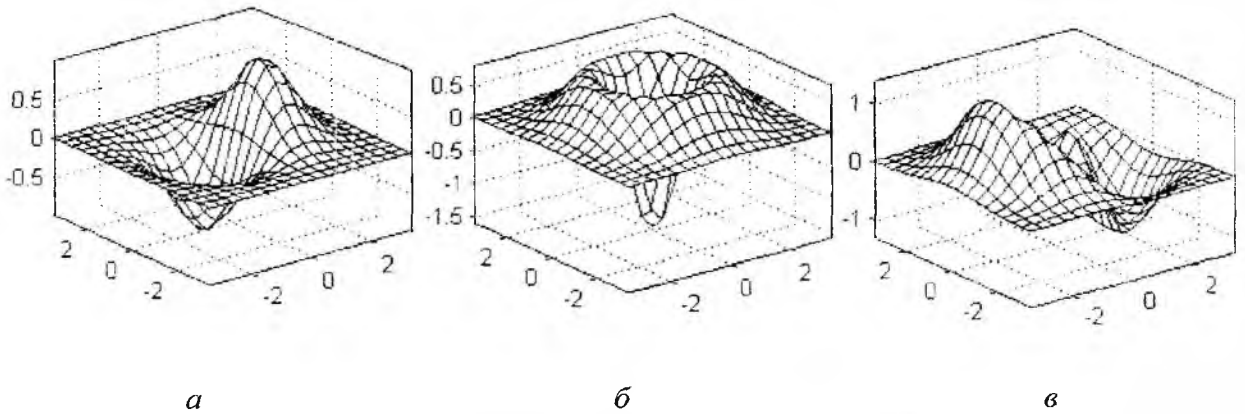


Рис. 2 – POLYWOG-вэйвлет и его производные:  
 а) POLYWOG-вэйвлет,  
 б) производная по сдвигам,  
 в) производная по растяжениям.

В качестве критерия обучения вэйвлет-нейрона будем использовать традиционную квадратичную функцию ошибки

$$E(k) = \frac{1}{2}(d(k) - y(k))^2 = \frac{1}{2}e^2(k) = \frac{1}{2}\left(d(k) - \sum_{i=1}^n \sum_{j=1}^{h_i} w_{ji} \varphi_{ji}(x_i(k))\right)^2, \quad (2)$$

где  $d(k)$  - внешний обучающий сигнал), производные которой по настраиваемым параметрам имеют вид

$$\frac{\partial E(k)}{\partial w_{ji}} = -e(k) \varphi_{ji}(x_i(k)) = -e(k) \sqrt{e} \tau_{ji}(k) \exp\left(-\frac{\tau_{ji}^2(k)}{2}\right) = -e(k) t_{ji}(k) \exp\left(-\frac{\tau_{ji}^2(k)}{2}\right),$$

$$\begin{aligned} \frac{\partial E(k)}{\partial c_{ji}} &= -e(k) w_{ji}(k) \frac{\partial \varphi_{ji}(x_i(k))}{\partial c_{ji}} = -e(k) w_{ji}(k) \sqrt{e} \sigma_{ji}^{-1}(k) (\tau_{ji}^2(k) - 1) \exp\left(-\frac{\tau_{ji}^2(k)}{2}\right) = \\ &= -e(k) w_{ji}(k) t_{ji}^c(k) \exp\left(-\frac{\tau_{ji}^2(k)}{2}\right), \end{aligned}$$

$$\begin{aligned} \frac{\partial E(k)}{\partial \sigma_{ji}^{-1}} &= -e(k) w_{ji}(k) \frac{\partial \varphi_{ji}(x_i(k))}{\partial \sigma_{ji}^{-1}} = e(k) w_{ji}(k) \sqrt{e} (x_i(k) - c_{ji}(k)) (\tau_{ji}^2(k) - 1) \exp\left(-\frac{\tau_{ji}^2(k)}{2}\right) = \\ &= -e(k) w_{ji}(k) t_{ji}^{\sigma}(k) \exp\left(-\frac{\tau_{ji}^2(k)}{2}\right). \end{aligned}$$

Далее, применяя для минимизации (2) градиентную процедуру, можно записать алгоритм обучения взвешенно-нейрона в виде

$$\begin{cases} w_{ji}(k+1) = w_{ji}(k) + \eta^w(k)e(k)t_{ji}(k) \exp\left(-\frac{\tau_{ji}^2(k)}{2}\right), \\ c_{ji}(k+1) = c_{ji}(k) + \eta^c(k)e(k)w_{ji}(k)t_{ji}^c(k) \exp\left(-\frac{\tau_{ji}^2(k)}{2}\right), \\ \sigma_{ji}^{-1}(k+1) = \sigma_{ji}^{-1}(k) + \eta^\sigma(k)e(k)w_{ji}(k)t_{ji}^\sigma(k) \exp\left(-\frac{\tau_{ji}^2(k)}{2}\right), \end{cases} \quad (3)$$

где скалярные коэффициенты  $\eta^w(k), \eta^c(k), \eta^\sigma(k)$  определяют шаг смещения в пространстве настраиваемых параметров.

Вводя  $(h_i \times 1)$  - векторы переменных

$$\varphi_i(x_i(k)) = (\varphi_{1i}(x_i(k)), \varphi_{2i}(x_i(k)), \dots, \varphi_{h_i i}(x_i(k)))^T,$$

$$w_i(k) = (w_{1i}(k), w_{2i}(k), \dots, w_{h_i i}(k))^T,$$

$$c_i(k) = (c_{1i}(k), c_{2i}(k), \dots, c_{h_i i}(k))^T,$$

$$\sigma_i^{-1}(k) = (\sigma_{1i}^{-1}(k), \sigma_{2i}^{-1}(k), \dots, \sigma_{h_i i}^{-1}(k))^T \text{ и}$$

$$\tau_i(k) = (\tau_{1i}(k), \tau_{2i}(k), \dots, \tau_{h_i i}(k))^T,$$

можно получить также градиентный алгоритм обучения  $i$ -го взвешенно-синапса:

$$\begin{cases} w_i(k+1) = w_i(k) + \eta^w(k)e(k)\varphi_i(x_i(k)), \\ c_i(k+1) = c_i(k) + \eta^c(k)e(k)w_i(k)\sqrt{e} \odot \sigma_i^{-1}(k) \odot (\tau_i^2(k) - E) \odot \exp\left(-\frac{\tau_i^2(k)}{2}\right), \\ \sigma_i^{-1}(k+1) = \sigma_i^{-1}(k) + \eta^\sigma(k)e(k)w_i(k)\sqrt{e} \odot (x_i(k) - c_i(k)) \odot (\tau_i^2(k) - E) \odot \exp\left(-\frac{\tau_i^2(k)}{2}\right), \end{cases}$$

где  $\tau_i^2(k) = \sigma_i^{-1}(k) \odot \sigma_i^{-1}(k) \odot (x_i(k) - c_i(k)) \odot (x_i(k) - c_i(k))$ ;

$\odot$  - символ прямого (скоттова) произведения;

$E - (h_i \times 1)$  - вектор, состоящий из единиц.

Повысить скорость сходимости процессов обучения можно, переходя от градиентных процедур к алгоритмам второго порядка, среди которых для настройки нейронных сетей наибольшее распространение получил алгоритм Левенберга-Маркварта [21].

Обозначив

$$\begin{cases} \varphi_i^c(x_i(k)) = w_i(k)\sqrt{e} \odot \sigma_i^{-1}(k) \odot (\tau_i^2(k) - E) \odot \exp\left(-\frac{\tau_i^2(k)}{2}\right), \\ \varphi_i^\sigma(x_i(k)) = w_i(k)\sqrt{e} \odot (x_i(k) - c_i(k)) \odot (\tau_i^2(k) - E) \odot \exp\left(-\frac{\tau_i^2(k)}{2}\right), \end{cases}$$

можно записать процедуру обучения вэйвлет-синапса на основе алгоритма Левенберга-Маркварта в виде

$$\begin{cases} w_i(k+1) = w_i(k) + (\varphi_i(x_i(k))\varphi_i^T(x_i(k)) + \alpha^w I_i)^{-1} e(k)\varphi_i(x_i(k)), \\ c_i(k+1) = c_i(k) + (\varphi_i^c(x_i(k))\varphi_i^{cT}(x_i(k)) + \alpha^c I_i)^{-1} e(k)\varphi_i^c(x_i(k)), \\ \sigma_i^{-1}(k+1) = \sigma_i^{-1}(k) + (\varphi_i^\sigma(x_i(k))\varphi_i^{\sigma T}(x_i(k)) + \alpha^\sigma I_i)^{-1} e(k)\varphi_i^\sigma(x_i(k)), \end{cases} \quad (4)$$

где  $\alpha^w, \alpha^c, \alpha^\sigma$  - малые положительные регуляризующие добавки;

$I_i - (h_i \times h_i)$  - единичная матрица.

Воспользовавшись леммой об обращении суммы двух матриц и проводя последовательность очевидных преобразований [22, 23], алгоритм (4) можно переписать в следующей достаточно простой форме, не требующей обращения матриц, что делает ее весьма удобной для работы в реальном времени:

$$\begin{cases} w_i(k+1) = w_i(k) + \frac{e(k)\varphi_i(x_i(k))}{\alpha^w + \|\varphi_i(x_i(k))\|^2} = w_i(k) + \frac{e(k)\varphi_i(x_i(k))}{\alpha^w(k)}, \\ c_i(k+1) = c_i(k) + \frac{e(k)\varphi_i^c(x_i(k))}{\alpha^c + \|\varphi_i^c(x_i(k))\|^2} = c_i(k) + \frac{e(k)\varphi_i^c(x_i(k))}{\alpha^c(k)}, \\ \sigma_i^{-1}(k+1) = \sigma_i^{-1}(k) + \frac{e(k)\varphi_i^\sigma(x_i(k))}{\alpha^\sigma + \|\varphi_i^\sigma(x_i(k))\|^2} = \sigma_i^{-1}(k) + \frac{e(k)\varphi_i^\sigma(x_i(k))}{\alpha^\sigma(k)}. \end{cases} \quad (5)$$

С целью придания алгоритму (5) сглаживающих свойств, можно ввести его экспоненциально-взвешенную модификацию, при этом

$$\begin{cases} \alpha_i^w(k+1) = \alpha \alpha_i^w(k) + \|\varphi_i(x_i(k))\|^2, \\ \alpha_i^c(k+1) = \alpha \alpha_i^c(k) + \|\varphi_i^c(x_i(k))\|^2, \\ \alpha_i^\sigma(k+1) = \alpha \alpha_i^\sigma(k) + \|\varphi_i^\sigma(x_i(k))\|^2, \\ 0 \leq \alpha \leq 1 \end{cases} \quad (6)$$

здесь  $\alpha$  - параметр забывания.

Несложно видеть, что при  $\alpha = 1$  процедура (5), (6) приобретает свойства стохастической аппроксимации адаптивного алгоритма идентификации Гудвина-Рэмеджа-Кэйнеса [24], а при  $\alpha = 0$  имеет форму популярного в теории искусственных нейронных сетей алгоритма Уидроу-Хоффа.

Как видно, использование модифицированных алгоритмов обучения второго порядка практически не усложняет численную реализацию процедур настройки вэйвлет-синапсов, обеспечивая при этом повышение их скорости сходимости.

### 3. Результаты численного эксперимента.

Эффективность предлагаемого алгоритма исследовалась в процессе решения задачи адаптивного прогнозирования хаотического сигнала Мандельброта [25]

$$x(k+1) = 4(1-x(k))x(k), \quad k = 1, 2, \dots \quad (7)$$

с помощью обучаемого вэйвлет-предиктора, схема которого приведена на рис. 3.

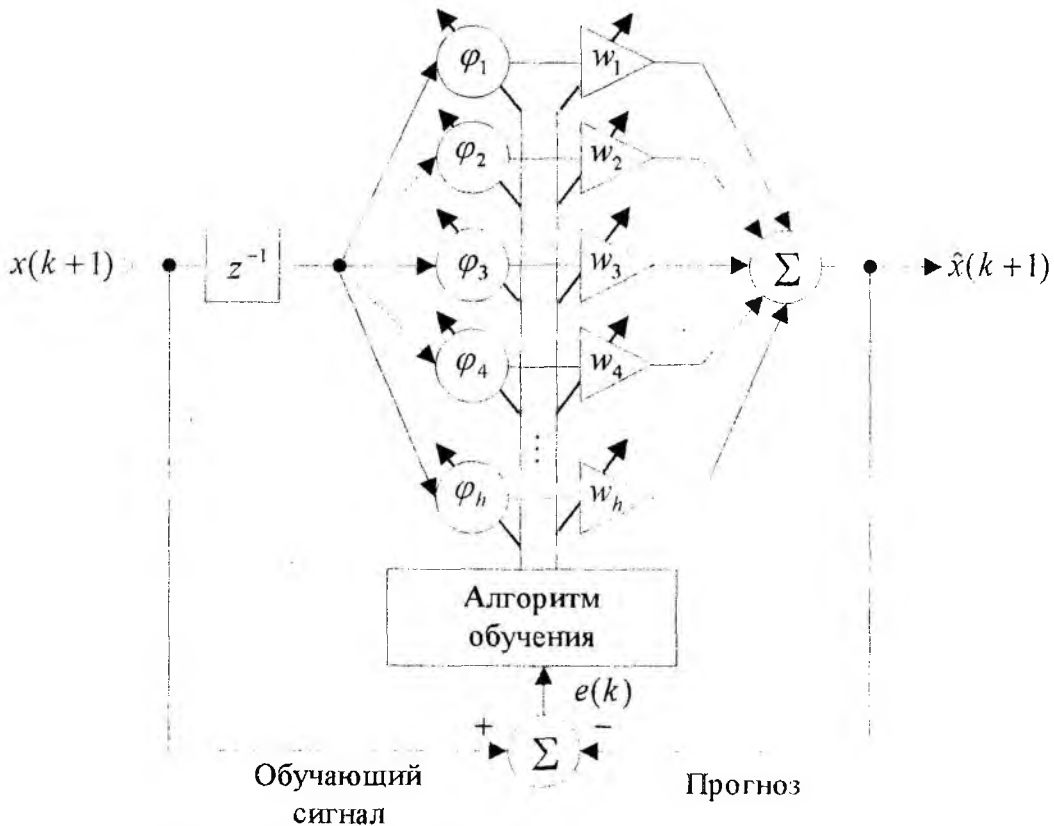


Рис. 3. Адаптивный вэйвлет-предиктор реального времени

В качестве начального условия было принято  $x(1) = 0.2$ , объем обучающей выборки  $x(k)$  составил 1500 точек. Базисные функции были созданы с помощью процедуры кластеризации, которая обнаружила 10 кластеров, параметры которых были использованы для инициализации 10 нейронов ( $h = 10$ ). Начальные веса выходного слоя были приняты нулевыми.

Настройка одноходового вэйвлет-нейрона производилась с помощью градиентного алгоритма (3) с постоянными параметрами шага  $\eta^w = 0.5$ ,  $\eta^c = 0.5$ ,  $\eta^\sigma = 10$  (эти коэффициенты могут выбираться эвристическим методом проб и ошибок и зависят от многих факторов, таких как размер сети, дисперсия входов и т.д.) и предложенной процедуры (5), (6) с коэффициентами забывания  $a^w = \alpha^c = \alpha^\sigma = 0.99$ .

На рис. 4 а приведены графики реального процесса (7) (пунктирная линия), прогноза на основе алгоритма (3) (сплошная линия) и график ошибки прогноза, а на рис. 4 б приве-

дено изменение среднеквадратичной ошибки (RMSE) при прогнозировании алгоритмом (3), которая составила  $RMSE = 0.012512$ .

Аналогично на рис. 5 а приведены графики реального процесса (7) (пунктирная линия), прогноза на основе алгоритма (5), (6) (сплошная линия) и ошибки прогноза, а на рис. 5 б приведено изменение среднеквадратичной ошибки при прогнозировании с помощью алгоритма (5), (6), которая составила  $RMSE = 0.009592$ .

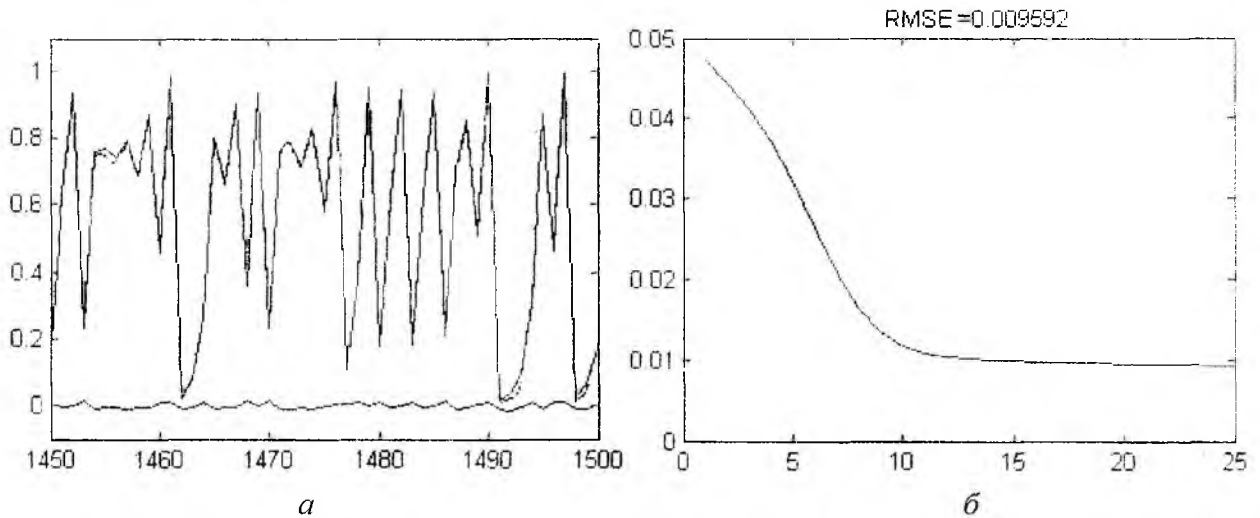


Рис. 4. Прогнозирование хаотического процесса на основе алгоритма (3)

- а) Реальный процесс и прогноз;
- б) Изменение среднеквадратичной ошибки

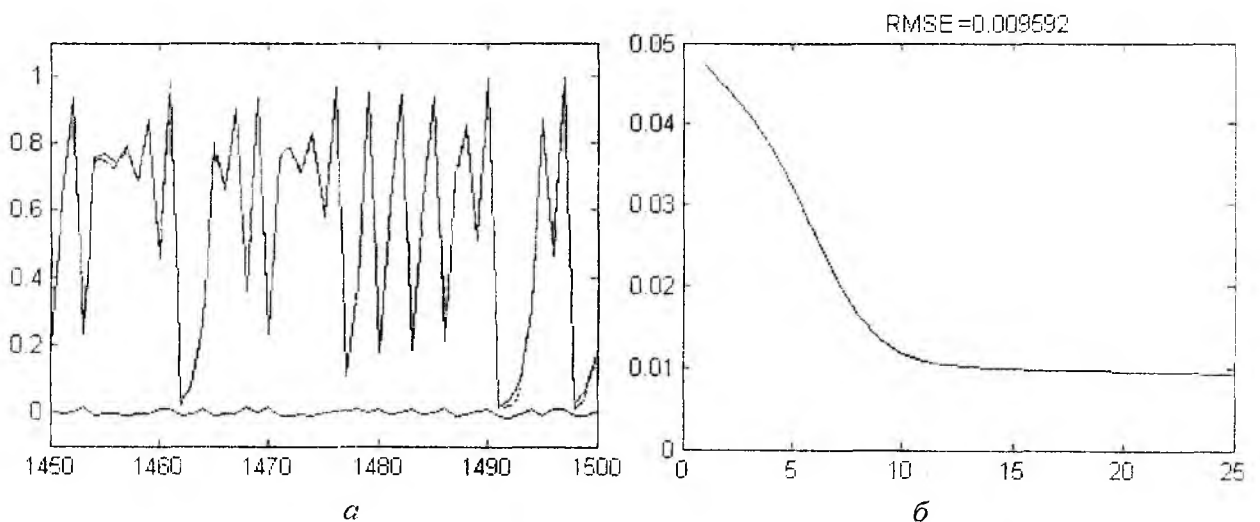


Рис. 5 – Прогнозирование хаотического процесса на основе алгоритма (5), (6)

- а) Реальный процесс и прогноз;
- б) Изменение среднеквадратичной ошибки

Видно, что оба алгоритма обеспечивают примерно одинаковую точность прогнозирования, однако предлагаемая процедура обеспечивает более высокую скорость сходимости, что иллюстрируется рис. 4, б и рис. 5, б.

**Выводы.**

Предложен новый оптимальный по быстродействию алгоритм обучения вэйвлет-нейрона, позволяющий настраивать все его параметры и обладающий как следящими, так и фильтрующими свойствами. Алгоритм прост в реализации и обеспечивает высокое качество обработки сигналов, что подтверждено результатами экспериментов. Возможность работы в реальном времени расширяет функциональные свойства вэйвлет-нейронных сетей.

**Список литературы:** 1. *Mallat S. G.* A theory for multiresolution signal decomposition: The wavelet representation // *IEEE Trans. on Pattern Anal. Machine Intell.* 1989. 11. P. 674-693. 2. *Daubechies I.* The wavelet transform, time-frequency localization and signal analysis // *IEEE Trans. on Information Theory.* 1990. 36. P. 961-1005. 3. *Rioul O., Vetterli M.* Wavelets and signal processing // *IEEE Signal Processing Mag.* 1991. 8. № 4. P. 14-38. 4. *Zhang Q., Benveniste A.* Wavelet networks // *IEEE Trans. on Neural Networks.* 1992. 3. P. 889-898. 5. *Casarent D. P., Smokelin J. S.* Neural net design of macro Gabor wavelet filters for distortion-invariant object detection in clutter // *Optical Engineering.* 1994. 33. - № 7. P. 2264-2270. 6. *Chang P.-R., Yeh B.-F.* Nonlinear communication channel equalization using wavelet neural networks // *IEEE Tran. on Neural Networks.* 1995. 6. P. 923-934. 7. *Aussem A., Murtagh F.* Combining neural networks forecasts on wavelet transformed time series // *Connection Sci.* 1997. 9. P. 113-121. 8. *Aussem A., Campbell J., Murtagh F.* Wavelet-based feature extraction and decomposition strategies for financial forecasting // *J. Comput. Intell. Finance.* 1998. March. P. 5-12. 9. *Zhang B. L., Coggins R., Jabri M. A., Dersch D., Flower B.* Multiresolution forecasting for futures trading using wavelet decomposition // *IEEE Trans. on Neural Networks.* 2001. 12. P. 765-774. 10. *Дзядык В. К.* Введение в теорию равномерного приближения функций полиномами. М.: Наука, 1977. 512 с. 11. *Park J., Sandberg I. W.* Universal approximation using radial-basis-function networks // *Neural Computation.* 1991. 3. P. 246-257. 12. *Kobayashi K., Torioka T.* A wavelet neural network for function approximation and network optimization / Eds. C. H. Dagli, B. R. Fernandez, J. Ghosh, R. T. S. Kumara "Intelligent Engineering Systems Through Artificial Neural Networks". 1994. 4. P. 505-510. 13. *Marar J. F., Filho E. C. B. C., Vasconcelos G. C.* Function approximation by polynomial wavelet generated from powers of sigmoids/ Ed. H. H. Szu "Wavelet Applications III" Proc SPIE Int. Soc. Optical Engineering. 1996. 2762. P. 365-374. 14. *Jiao L., Pan J., Fang Ya.* Multiwavelet neural network and its approximation properties // *IEEE Trans. on Neural Networks.* 2001. 12. P. 1060-1066. 15. *Resnikoff H. L.* Wavelets and adaptive signal processing / Ed. S. Haykin "Adaptive Signal Processing." Proc. SPIE-Int. Soc. Optical Engineering. 1991. 155. P. 370-382. 16. *Szu H. H., Telfer B. A., Kadambe S.* Neural network adaptive wavelets for signal representatin and classification // *Optical Engineering.* 1992. 31. № 9. P. 1907-1916. 17. *Kadambe S., Srinivasan P.* Applications of adaptive wavelets for speech // *Optical Engineering* 1994. 33. P. 2204-2211. 18. *Jouny I., Kanapathipillai M.* Neural network adaptive wavelet classification of radar targets // Proc. 1994 Int. Geoscience and Remote Sensing Symp. 1994. 4. P. 1889-1891. 19. *Lekutai G., van Landingham H. F.* Self-tuning control of nonlinear systems using neural network adaptive frame wavelets // Proc. IEEE Int. Conf. on Systems, Man and Cybernetics. - Piscataway, N.J. 1997. 2. P. 1017-1022. 20. *Бодянский Е. В., Винокурова Е. А., Плисс И. П.* Алгоритм обучения искусственной всплеск-нейронной сети // Сб. науч. тр. 1-го Международного радиоэлектронного форума «Прикладная радиоэлектроника. Состояние и перспективы развития» МРФ-2002. Часть 2. Харьков АН ПРЭ, ХНУРЭ, 2002. С. 87-89. 21. *Shepherd A. J.* Second-Order Methods for Neural Networks. London: Springer-Verlag, 1997. 145 p. 22. *Бодянский Е. В., Михальов О. І., Плисс І. П.* Адаптивне виявлення розладнань в об'єктах керування за допомогою штучних нейронних мереж. Дніпропетровськ: Системні технології, 2000. 140 с. 23. *Bodyanskiy Ye., Kolodyazhniy V., Stephan A.* An adaptive learning algorithm for a neuro-fuzzy network / Ed. by B. Reusch "Computational Intelligence. Theory and Applications." Berlin Heidelberg New York: Springer, 2001. P. 68-75. 24. *Goodwin G. C., Ramadge P. J., Caines P. E.* A globally convergent adaptive predictor // *Automatica.* 1981. 17. № 1 P. 135-140. 25. *Mandelbrot B. B.* Die fractale Geometrie der Natur. Basel: Birkhaeuser Verlag, 1991. 491 s.

Поступила в редколлегию 11.10.2003

