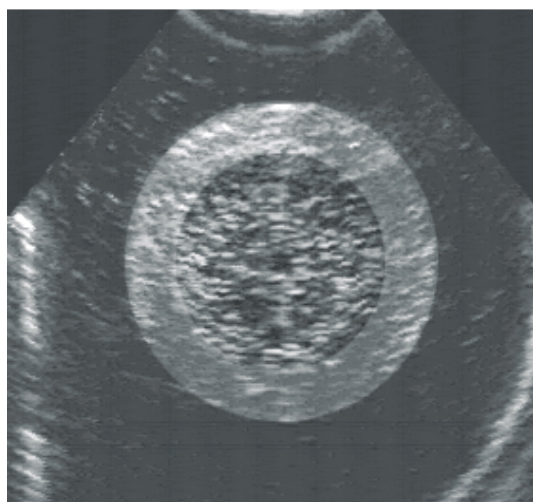


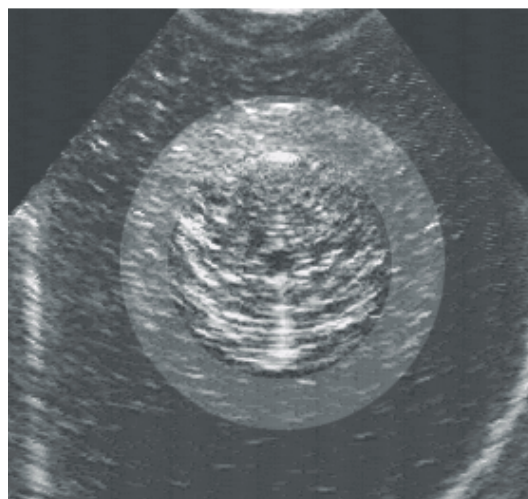
знаваемость низкоконтрастных элементов изображения, а также может использоваться для выделения его отдельных областей.

Цветокodирование исходных изображений выполнялось в привязке к площади контактного пятна, градации яркости которого получали наиболее "горячие" оттенки цвета (соответствующие градациям красного), а по мере удаления от него цвета постепенно "охлаждались".

При переходе от черно-белых изображений к псевдоцвету использовались две цветовые модели. Первая – модель RGB наиболее близка к принципу действия аппаратного обеспечения и основана на трехкомпонентной теории зрения, представляющей цвет в виде трех каналов для красной, зеленой и голубой компонент. Использование этой модели представляется наиболее удобным при введении зависимости оттенка цвета, от близости к контактному пятну. Вторая используемая модель, именуемая HSB (от английского эквивалента слов оттенок - насыщенность - яркость) также представляет цвет в виде трех компонент. Однако за оттенок в этом случае отвечает только первая компонента, а вторая и третья – за глубину цвета, характеризующую насыщенность оттенка при изменении от легкой пастели до сочного цвета, и за его яркость. Следует отметить, что использование данной модели предоставляет более широкие возможности для цветового отображения представляющих интерес элементов изображения, благодаря изменению яркости и насыщенности, при постоянном распределении псевдоцветов по диапазону градаций серого



а)



б)

Рисунок 3. Акустические изображения пар трения, подвергнутые локальной обработке в привязке к их геометрическим и пространственным характеристикам

цвета исходного изображения.

К сожалению, отсутствие возможности публикации цветных иллюстраций не позволяет представить цветокodированные изображения в их естественном виде, а преобразование их для публикации к градациям серого цвета представляется нецелесообразным.

#### СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. А.Я. Акулич, В.Б. Гладышук, М.В. Голуб, Ю.А. Кузавко, Б.Н. Склипус. Ультразвуковой диагностический комплекс томографии машиностроительных изделий//Американо - восточно-европейская конференция «Новые материалы и технологии в трибологии». Минск - Гродно - Варшава. 1997г.
2. Ya. A. Akulich, D. A. Kostiuk, Yu. A. Kuzavko Program improvement of quality of the images in 2D acoustic visualization. / Proceedings of international conference on neural networks and artificial intelligence. Brest, 1999.
3. Прэрт У. Цифровая обработка изображений. М.: Мир, 1982.- Кн. 2 – 480 с.
4. Lim, Jae S. Two-Dimensional Signal and Image Processing. Englewood Cliffs, NJ: Prentice Hall, 1990.

УДК 681.324.01

Головко В.А., Савицкий Ю.В.

## МЕТОД ОБУЧЕНИЯ РЕЦИРКУЛЯЦИОННЫХ НЕЙРОННЫХ СЕТЕЙ

### ВВЕДЕНИЕ

Одним из наиболее известных методов для сжатия информации без существенных потерь является метод главных компонент (РСА) [1]. Как известно из статистики [4], РСА метод осуществляет линейное сжатие информации. При этом осуществляется такое преобразование входного пространства данных, что выходные значения являются некоррелирован-

ными и основная доля их дисперсии сосредоточена в нескольких первых главных компонентах.

В последнее время большое внимание уделяется нейросетевым алгоритмам сжатия информации. Согласно литературе по РСА существуют следующие нейросетевые алгоритмы: 1) Алгоритм обучения Хебба [2,3]; 2) Антихеббовское обучение; 3) алгоритмы, базирующиеся на методе градиентного спуска

Головко Владимир Адамович. К.т.н., доцент института технической кибернетики НАНБ.

Беларусь, г. Минск, ул. Сурганова, 6.

Савицкий Юрий Викторович. Старший преподаватель каф. ЭВМиС Брестского государственного технического университета.

БГТУ, Беларусь, г. Брест, ул. Московская 267.

[2,6].

В данной статье рассматриваются градиентные алгоритмы обучения. Наиболее известными среди них являются алгоритмы обратного распространения ошибки и кумулятивное дельта-правило. Однако эти алгоритмы характеризуются большой временной сложностью и плохой сходимостью к подходящему решению. Поэтому в работе предлагается новый метод обучения рециркуляционных сетей. Он является нелинейным аналогом метода главных компонент и характеризуется меньшей временной сложностью по сравнению с известными алгоритмами.

### 1. ПОСТАНОВКА ЗАДАЧИ

Рециркуляционные сети характеризуются как прямым, так и обратным преобразованием информации (рисунок 1). Они применяются для сжатия (прямое преобразование) и восстановления исходной (обратное преобразование) информации. Промежуточный слой осуществляет сжатие входного множества данных  $X$ . При этом выходное значение  $j$ -го нейронного элемента промежуточного слоя определяется как

$$y_j = F_p \left( \sum_{i=1}^n \omega_{ij} x_i \right), \quad (1)$$

где  $F_p$  – функция активации нейронных элементов промежуточного слоя,  $\omega_j$  – весовой коэффициент между  $i$ -м входным и  $j$ -м скрытым нейронами.

Выходной слой осуществляет восстановление входного множества данных. Значение  $i$ -го нейрона выходного слоя определяется следующим образом:

$$\bar{x}_i = F_n \left( \sum_{j=1}^p w'_{ji} y_j \right), \quad (2)$$

где  $F_n$  – функция активации нейронных элементов выходного

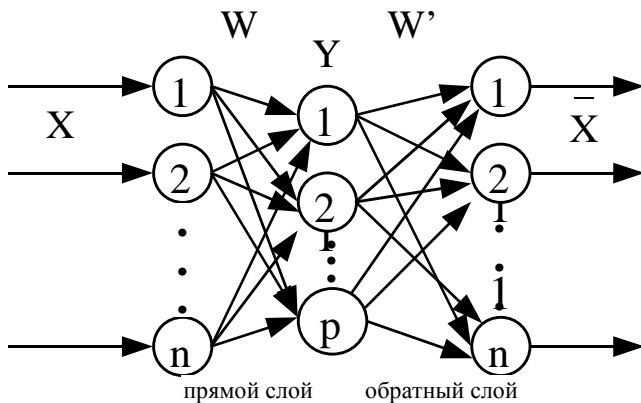


Рисунок 1.

слоя и  $i$ -м нейроном выходного слоя. Таким образом, рециркуляционная нейронная сеть осуществляет следующее преобразование информации:

$$R^n \xrightarrow{F_p} R^p \xrightarrow{F_n} R^n,$$

где  $p < n$ .

Цель обучения состоит в том, чтобы минимизировать суммарную среднеквадратичную ошибку между восстановленными данными  $\bar{X}$  и входными данными  $X$ :

$$E_S = \frac{1}{2} \sum_{k=1}^L \sum_{i=1}^n (x_i^k - \bar{x}_i^k)^2, \quad (3)$$

где  $L$  – общее количество тренировочных образцов. Предлагаемый метод состоит из двух отдельных фаз обучения. На первом этапе осуществляется определение весовой матрицы  $W' = [w'_{ji}]$ ,  $j = \overline{1, p}$ ,  $i = \overline{1, n}$  и матрицы выходных значений скрытого слоя  $\bar{Y} = [\bar{y}_j^k]$ ,  $j = \overline{1, p}$ ,  $k = \overline{1, L}$ . Для этого необходимо минимизировать выражение (3).

На втором этапе алгоритма определяется весовая матрица прямого слоя  $W = [\omega_{ij}]$ ,  $i = \overline{1, n}$ ,  $j = \overline{1, p}$ . При этом в качестве эталонных выходов используются значения  $\bar{Y}$ , полученные на первом этапе. Тогда обучение на второй фазе происходит с целью минимизации следующего выражения:

$$E'_S = \frac{1}{2} \sum_{k=1}^L \sum_{j=1}^p (y_j^k - \bar{y}_j^k)^2, \quad (4)$$

где

$$y_j^k = F_p \left( \sum_{i=1}^n \omega_{ij} x_i^k \right). \quad (5)$$

Для минимизации выражений (3) и (4) используется метод градиентного спуска. Таким образом, метод послыного обучения позволяет проводить обучение прямого и обратного слоя раздельно. Рассмотрим основные принципы такого метода.

### 2. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ МЕТОДА

Рассмотрим правила обучения для нелинейного аналога метода главных компонент. Пусть для различных слоев нейронной сети используется одна и та же нелинейная функция.

Тогда

$$F_n = F_p.$$

На первом этапе алгоритма послыного обучения необходимо определить весовые коэффициенты обратного слоя  $W' = [w'_{ji}]$ ,  $j = \overline{1, p}$ ,  $i = \overline{1, n}$ , а также для каждого входного вектора  $X$  определить такой вектор  $Y$ , чтобы обеспечить минимизацию суммарной среднеквадратичной ошибки (3).

Рассмотрим определение вектора  $Y$  для минимизации выражения (3). В соответствии с методом градиентного спуска компоненты вектора  $Y$  должны изменяться с течением времени, как:

$$y_j(t+1) = y_j(t) - \alpha \frac{dE}{d y_j(t)} \quad (6)$$

где  $\alpha$  – скорость обучения.

Ошибка  $E$  определяется при этом следующим образом:

$$E = \frac{1}{2} \sum_{i=1}^n (\bar{x}_i - x_i)^2. \quad (7)$$

Дифференцируя выражение (7) по переменной  $y_j(t)$  получим:

$$\begin{aligned} \gamma_j &= \frac{\partial E}{\partial y_j} = \frac{\partial E}{\partial x_i} \frac{\partial x_i}{\partial S_i} \frac{\partial S_i}{\partial y_j} = \\ &= \sum_i w'_{ji} (\bar{x}_i - x_i) F'(S_i) \end{aligned} \quad (8)$$

где  $S_i$  – взвешенная сумма  $j$ -го нейрона:

$$S_i = \sum_{j=1}^p \omega'_{ji} y_j. \quad (9)$$

Тогда

$$y_j(t+1) = y_j(t) - \alpha_j(t) \gamma_j \quad (10)$$

Определим теперь выражения для настройки весовых коэффициентов  $\omega'_{ji}$  обратного слоя. Тогда в соответствии с методом градиентного спуска

$$\omega'_{ji}(t+1) = \omega'_{ji}(t) - \alpha(t) \frac{\partial E}{\partial \omega'_{ji}(t)}. \quad (11)$$

Производная среднеквадратичной ошибки определяется как:

$$\frac{\partial E}{\partial \omega'_{ji}} = \frac{\partial E}{\partial \bar{x}_i} \frac{\partial \bar{x}_i}{\partial S_i} \frac{\partial S_i}{\partial \omega'_{ji}} = (\bar{x}_i - x_i) F'(S_i) y_j \quad (12)$$

Отсюда следует, что модификация весовых коэффициентов обратного слоя должна производиться следующим образом:

$$\omega'_{ji}(t+1) = \omega'_{ji}(t) - \alpha(t) (\bar{x}_i - x_i) F'(S_i) y_j \quad (13)$$

Таким образом, в результате выполнения первого этапа происходит определение весовых коэффициентов обратного слоя  $\omega'_{ji}$  и эталонных выходов  $\bar{y}_j^k, j = \overline{1, p}, k = \overline{1, L}$  скрытого слоя.

Следует отметить, что значения эталонных выходов скрытого слоя  $\bar{y}_j^k$  могут не принадлежать диапазону от  $-1$  до  $1$ .

Поэтому необходимо проводить нормализацию эталонных выходов в соответствии с используемой функцией активации.

На втором этапе метода послойного обучения происходит определение весовых коэффициентов  $w_{ij}$  прямого слоя, где в качестве эталонных выходов  $\bar{y}_j^k$  используются значения, полученные на предыдущем этапе. Для этого необходимо минимизировать выражение (4). В соответствии с методом градиентного спуска

$$w_{ij}(t+1) = w_{ij}(t) - \alpha(t) \frac{\partial E'}{\partial w_{ij}(t)}, \quad (14)$$

где

$$E' = \frac{1}{2} \sum_{j=1}^p (y_j - \bar{y}_j)^2 \quad (15)$$

Дифференцируя (14) по  $w_{ij}$  можно получить, что

$$\frac{\partial E'}{\partial w_{ij}} = \frac{\partial E'}{\partial y_j} \frac{\partial y_j}{\partial S_j} \frac{\partial S_j}{\partial w_{ij}} = (y_j - \bar{y}_j) F'(S_j) x_i \quad (16)$$

где  $S_j$  – взвешенная сумма  $j$ -го нейрона скрытого слоя:

$$S_j = \sum \omega_{ij} x_i$$

Тогда

$$\omega_{ij}(t+1) = \omega_{ij}(t) - \alpha(t) (y_j - \bar{y}_j) F'(S_j) x_i \quad (17)$$

Следует отметить, что в качестве  $\alpha$  может использоваться адаптивный шаг обучения [5]. Приведенные выше выражения могут использоваться для обучения рециркуляционных сетей.

### 3. ИСПОЛЬЗОВАНИЕ РАЗЛИЧНЫХ ФУНКЦИЙ АКТИВАЦИИ

В качестве функций активации для рециркуляционных сетей могут использоваться различные нелинейные функции.

Определим выражение  $F'(S_j)$  для различных функций активации.

*Гиперболический тангенс*

В этом случае выходное значение  $j$ -го нейрона промежуточного слоя определяется следующим образом:

$$y_j = th(S_j),$$

Соответственно для нейронов выходного слоя

$$\bar{x}_i = th(S_i).$$

Тогда

$$F'(S_i) = \frac{\partial \bar{x}_i}{\partial S_i} = (1 - \bar{x}_i^2)$$

$$F'(S_j) = \frac{\partial y_j}{\partial S_j} = (1 - y_j^2)$$

$$F'(0) = 1.$$

*Биполярная сигмоидная функция*

В соответствии с данной функцией активации

$$y_j = \frac{2}{1 + e^{-s_j}} - 1,$$

$$\bar{x}_i = \frac{2}{1 + e^{-s_i}} - 1.$$

Тогда

$$F'(S_i) = \frac{\partial \bar{x}_i}{\partial S_i} = \frac{1}{2} (1 - \bar{x}_i^2)$$

$$F'(S_j) = \frac{\partial y_j}{\partial S_j} = \frac{1}{2} (1 - y_j^2)$$

Используя приведенные выше выражения можно определить алгоритм послойного обучения для различных функций активации.

### 4. АЛГОРИТМ ПОСЛОЙНОГО ОБУЧЕНИЯ

Алгоритм послойного обучения состоит из следующих шагов:

1. Случайная инициализация весовых коэффициентов и задание суммарной среднеквадратичной ошибки сети  $E_m$ .
2. Последовательно подаются  $L$  образов на нейронную сеть. При этом для каждого образа происходит только прямое распространение информации. В результате данного этапа определяется базовый набор сжатых образов  $Y$ , которые будут использоваться на следующем этапе алгоритма.
3. Рассматривается только обратный слой сети. В качестве входной информации используются значения  $Y$ , определенные на предыдущем шаге алгоритма. В качестве эталонных значений выходов обратного слоя принимается вектор исходных данных  $X$ . При этом, производится следующая последовательность действий:
  - 3.1. Для  $L$  входных образов производится настройка весовых коэффициентов обратного слоя  $W'$  по выражению (13).
  - 3.2. Для  $L$  входных образов производится настройка желаемых выходов прямого слоя  $Y$  согласно выражению (10).
  - 3.3. Пункты 3.1 и 3.2 повторяются, пока суммарная среднеквадратичная ошибка обратного слоя не станет меньше заданной  $E_m$ .
4. Производится модификация весовых коэффициентов прямого слоя в соответствии с выражением (14). Для этого входные образы последовательно подаются на сеть и для каждого образа происходит только прямое распространение

Сравнение различных алгоритмов обучения

Алгоритм обучения	$\alpha_1$	$\alpha_2$	$\alpha_3$	Кол-во тестов	Кол-во успешн. тестов	Среднее кол-во итераций	Среднее время, сек	$E_S$
<b>SA</b>	Адапт.	0,1	Адапт.	100	100	15,7	0,0847	1,3E-7
	0,5	0,5	0,5	100	98	16,5	0,0747	1,12E-6
<b>BP</b>	0,05	0,05	—	100	100	129	0,578	8,3E-6
<b>HL</b>	0,03	0,03	—	100	100	19,9	0,0763	3,7E-6

**SA**- алгоритм послойного обучения; **BP**-алгоритм обратного распространения ошибки; **HL**-алгоритм Хебба

информации. В качестве эталонных данных используются значения  $\bar{Y}$ , полученные на предыдущем шаге алгоритма.

Таблица 2

5. Пункт 4 продолжается до тех пор, пока суммарная среднеквадратичная ошибка прямого слоя не станет меньше заданной.

Сжатый массив

0,765	0,816	-0,285
0,128	0,0386	-0,556
0,482	0,395	-0,798
0,0498	-0,115	-0,894
0,439	0,464	-0,189
0,374	0,404	-0,11
0,347	0,327	-0,355
0,376	0,238	-0,99
0,156	0,0427	-0,702
0,177	0,112	-0,466

### 5. ЭКСПЕРИМЕНТЫ

Анализ возможностей метода послойного обучения проводился на различных числовых массивах и изображениях. Как показали эксперименты, эффективность его зависит от входного массива данных. Один из примеров входного массива чисел показан в таблице 1. В этом случае нейронная сеть состоит из 6 входных, 3 скрытых и 6 выходных нейронных элементов. В качестве функции активации использовалась функция гиперболического тангенса. В таблице 2 представлен сжатый массив данных, который получается после обучения нейронной сети.

Таблица 1

Входной массив

0,9	0,1	0,9	0,9	0,9	0,9
0,6	0,7	0,6	0,6	0,6	0,6
0,9	0,8	0,9	0,9	0,9	0,9
0,7	0,9	0,7	0,7	0,7	0,7
0,7	0,1	0,7	0,7	0,7	0,7
0,6	0,0	0,6	0,6	0,6	0,6
0,7	0,4	0,7	0,7	0,7	0,7
0,9	0,9	0,9	0,9	0,9	0,9
0,7	0,8	0,7	0,7	0,7	0,7
0,6	0,6	0,6	0,6	0,6	0,6

Пусть  $\alpha_1$  – шаг обучения для модификации весовых коэффициентов прямого слоя  $w_{ij}$ ,  $\alpha_2$  – шаг обучения для модификации весовых коэффициентов обратного слоя  $w_{ji}'$  и  $\alpha_3$  – скорость модификации выходов скрытого слоя  $y_j$ . В таблице 3 представлена информация о временной сложности различных алгоритмов обучения. Как следует из таблицы, при использовании алгоритма послойного обучения получаются лучшие результаты по сравнению с другими алгоритмами.

### ЗАКЛЮЧЕНИЕ

В работе предложен новый метод обучения рециркуляционных нейронных сетей. Данный алгоритм называется методом послойного обучения. Результаты экспериментов с предложенным методом показали, что он позволяет достичь более высокой скорости обучения по сравнению с другими алгоритмами.

### СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Jolliffe I.T. Principal Component Analysis. Springer-Verlag, 1986
2. P. Baldi and K. Hornik. Learning in linear networks; A Survey // IEEE Trans. Neural Networks, Vol. 6, pp. 837-857, 1995
3. C. Fyfe and R. Baddley. Non-linear data structure extraction using simple hebbian networks. Biological Cybernetics, 72(6); 533-541, 1995.
4. E. Oja. Principal components, minor components and linear networks // Neural Networks, vol.5, pp.927-935, 1992
5. V.Golovko, V.Gladyschuk. Unsupervised training algorithm for recirculation neural networks // Proceedings of ICNNAI'99, Brest, 1999, p.19-25
6. H.Ando, S.Suzuki, T.Fujita. Unsupervised visual learning of three-dimensional objects using a modular network architecture // Neural Networks, vol. 12, pp. 1037-1051, 1999

УДК 681.324.01

**Головко В.А., Савицкий Ю.В., Фоменкова Н.Е.**

## ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ПРОГНОЗИРОВАНИЯ И МОДЕЛИРОВАНИЯ НЕЛИНЕЙНЫХ СИСТЕМ

### ВВЕДЕНИЕ

Способность нейронных сетей к обобщению результатов создает предпосылки для построения на базе их различного рода прогнозирующих систем. Пусть дан временной ряд  $x(t)$  на промежутке  $t=1, m$ . Тогда задача прогнозирования состоит

в нахождении продолжения временного ряда на неизвестном промежутке времени, т.е. необходимо определить значения  $x(m+1)$ ,  $x(m+2)$  и так далее.

Для прогнозирования временных рядов используется ме-

**Фоменкова Наталья Евгеньевна.** Аспирант каф. ЭВМиС БГТУ. БГТУ, Беларусь, г. Брест, ул. Московская 267.