

УДК: 004.8

# Формирование самоорганизующихся отображений сенсорных сигналов на непрерывные нейросетевые аттракторы

Соловьева К.П.

*Центр оптико-нейронных технологий, Научно-исследовательский институт системных исследований, Российская академия наук, Москва, 119333, Россия  
Московский физико-технический институт (Государственный университет),  
Долгопрудный, Московская область, 141700, Россия  
АНО “Мир 2045”, Москва, 119270, Россия*

**Аннотация.** В работе описывается простая модельная нейронная система, в которой осуществляется самоорганизующееся отображение многомерных рецепторных сигналов  $R$ -мерного пространства (где  $R$  – количество рецепторов) на состояния сети из  $N$  взаимодействующих нейронов. Рекуррентная сеть из  $N$  взаимодействующих нейронов обладает одномерным непрерывным аттрактором («Bump attractor»). В отличие от нейрообразной конструкции Т. Кохонена, отображение входного сигнала осуществлялось на устойчивые состояния рекуррентной сети. В том случае, когда множество входных сигналов представляло собой одномерное циклическое многообразие в  $R$ -мерном пространстве, а нейронная сеть представляла полное кольцо нейронов с локальными возбуждающими связями, в результате процесса обучения формировалось топологически корректное отображение множества входных сигналов на устойчивые состояния нейронной сети. Продемонстрированы свойства системы, устойчивые к детализации модели нейрона.

**Ключевые слова:** самоорганизующееся отображение, непрерывный аттрактор, нейрон МакКаллока–Питтса, обучение, карты Кохонена.

## 1. ВВЕДЕНИЕ

Сигналы внешнего мира, приходя на рецепторы нервной системы, кодируются и обрабатываются на следующих уровнях нервной системы. Расшифровка механизмов этого процесса является одной из основных нерешенных задач в области анализа механизмов работы мозга [1]. Вместе с тем, множественные экспериментальные данные позволяют выявить некоторые основные принципы, одним из которых является сохранение отношений соседства на разных уровнях нервной системы, от рецепторов до коры. Этот принцип проявляется в виде соматотопии в сенсорной, или тонотопии и ретинотопии в слуховой и зрительной системах [2].

Впервые нейрообразная конструкция, производящая подобного рода отображения и удовлетворяющая принципу топологического соответствия, была предложена в работах Тейво Кохонена [3].

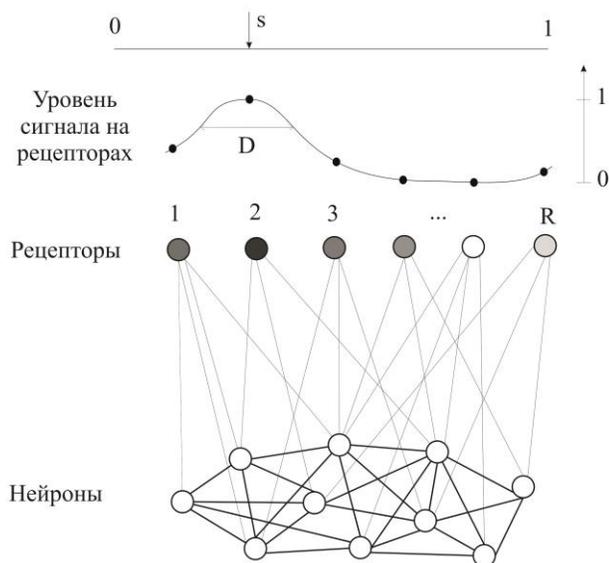
В данной работе мы предприняли попытку реализовать самоорганизующееся отображение активности рецепторов в состояния рекуррентной нейронной сети, содержащей непрерывный одномерный («бугорковый») аттрактор [4,5].

## 2. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ СИСТЕМЫ

Система состоит из двух связанных между собой частей, условно называемых далее рецепторы и нейроны рекуррентной сети (или просто «нейроны») (рис. 1).

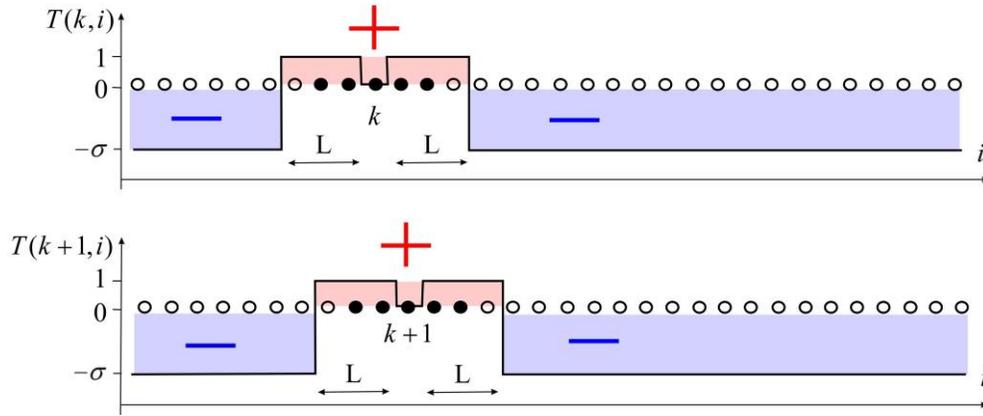
Состояние среды характеризуется  $R$ -мерным вектором  $\mathbf{V}$ , с вещественными координатами в диапазоне от нуля до единицы. Рецепторы между собой никак не связаны. Каждый рецептор связан с каждым нейроном рекуррентной сети. Матрица связей рецепторов с нейронами  $\mathbf{W}$  ( $N \times R$ ).

Рекуррентная сеть состоит из  $N$  связанных симметричными связями бинарных нейронов [6]. Нейронам присвоены номера от 1 до  $N$ . Состояние нейронной сети характеризуется  $N$ -мерным вектором  $\mathbf{X}$  со значениями координат «0» или «1». Матрица связей нейронов между собой  $\mathbf{T}$  ( $N \times N$ ).



**Рис. 1.** Схема выбора сигнала на рецепторах и соединения нейронов рекуррентной сети. Подробности в тексте.

Нейроны расположены по кольцу так, что первый и  $N$ -й нейроны являются соседними. Используемая цикличность не является принципиальной в нейрофизиологическом плане, или в плане получения тех или иных свойств системы, а используется для упрощения, чтобы не учитывать граничные условия. Принцип организации связей между нейронами представляет упрощенный вариант «мексиканской шляпы» [7]. В нашем случае, каждый нейрон связан с  $L$  соседями «слева» и  $L$  соседями «справа» положительными связями, с весом равным единице, и отрицательными связями, с весом  $-\sigma$ , со всеми остальными нейронами (рис. 2).



**Рис. 2.** Схема соединения нейронов наподобие метода «мексиканская шляпа». Кружками схематически показаны нейроны. На верхнем рисунке изображён график синаптических весов  $k$ -го нейрона с остальными, на нижнем графике –  $k+1$ -го нейрона со всеми остальными.

В сети задана асинхронная случайная циклическая динамика. В последовательные такты времени случайно выбирается порядок обхода всех  $N$  нейронов и производится  $N$  обновлений состояний нейронов по формуле:

$$x_i(t+1) = \begin{cases} 1, & \text{при } (\mathbf{W}_{i,\cdot}, \mathbf{V}) + (\mathbf{T}_{i,\cdot}, \mathbf{X}(t)) \geq \theta \\ 0, & \text{при } (\mathbf{W}_{i,\cdot}, \mathbf{V}) + (\mathbf{T}_{i,\cdot}, \mathbf{X}(t)) < \theta \end{cases}, \quad \text{для } i = \Xi(t, 1), \dots, \Xi(t, N) \quad (1)$$

где  $\theta$  – порог,  $\mathbf{W}_{i,\cdot}$  – строка матрицы связей нейронов с рецепторами, соответствующая  $i$ -му нейрону, а  $\Xi(t, j)$  – случайная (независимая в разные моменты  $t$ ) перестановка индексов  $\{1, \dots, N\}$ .

При таком законе динамики и отсутствии входа с рецепторов в сети имеется  $N$  устойчивых состояний, в каждом из которых активно некоторое число нейронов; назовём это число шириной точки аттрактора. Ширина точки аттрактора зависит от величины порога и соотношения весов положительных и отрицательных связей. С ростом модуля веса отрицательных связей  $\sigma$  ширина точки аттрактора уменьшается до величины  $L$  (см. выше). При дальнейшем росте  $\sigma$  активность в сети гаснет.

Работа нейронной сети проходит в двух режимах: режим обучения и режим тестирования.

## Обучение

Обучение системы состоит в многократном последовательном повторении следующей процедуры:

1. Подаётся вход на рецепторы  $\mathbf{V}$ .
2. Изначально активность всех нейронов рекуррентной сети равна нулю. Под воздействием сигнала на рецепторах, в рекуррентной нейронной сети запускается асинхронная динамика, приводящая сеть в одно из ее устойчивых состояний.
3. Для всех нейронов, активных в данной точке аттрактора, связи со всеми рецепторами изменяются по правилу Хебба, в модификации, использованной в работе Кохонена:

$$\mathbf{W}_{i,\cdot} = \mathbf{W}_{i,\cdot} + \eta((\mathbf{XV}^T)_{i,\cdot} - \mathbf{W}_{i,\cdot}), \quad (2)$$

где  $i$  – номер нейрона,  $W_{i.}$  – строка матрицы связей нейронов с рецепторами, соответствующая  $i$ -му нейрону,  $V$  – сигнал на рецепторах,  $X$  – вектор установившегося состояния в сети нейронов и  $\eta$  – коэффициент обучения.

Стандартная сеть Кохонена, имеющая на выходе линейную цепочку нейронов, позволяет в том случае, когда в пространстве рецепторов многообразие возможных сигналов имеет одномерную топологию, получать топологически корректное отображение этого многообразия на цепочки элементов конструкции.

Чтобы исследовать свойства самоорганизующихся отображений в нашем случае, мы использовали следующее модельное множество сигналов.

Входные сигналы поступают в дискретные моменты времени. В каждый момент выбирается случайно и равновероятно действительное число  $s$  в полуинтервале  $[0,1)$ . Возбуждения рецепторов при этом определяются значениями функции  $G(D, s, v)$ , где  $G$  вычисляется по формуле:

$$G = \max_p \left[ \exp \left( - \left( \frac{v - (s + p)R}{D} \right)^2 \right) \right], \quad (3)$$

в точках  $v = \{1, 2, \dots, R\}$  (рис. 1),  $p$  принимает значения  $-1$ ,  $0$  или  $1$ . Иначе говоря, значение возбуждения  $v$ -го рецептора определяется как  $V_v = G(D, s, v)$ . Можно сказать, что  $G$  – это дискретная функция Гаусса, заданная на кольце рецепторов, а  $D$  – параметр этой функции, геометрический смысл которого – это ширина «бугорка» этой функции (рис. 1).

## Тестирование

Тестирование системы производится следующим образом.

1. Последовательно, с фиксированным шагом ( $0.01-0.001$ ) задаются значения параметра сигнала  $s$  от  $0$  до  $1$ .

2. То же, что и в процедуре обучения.

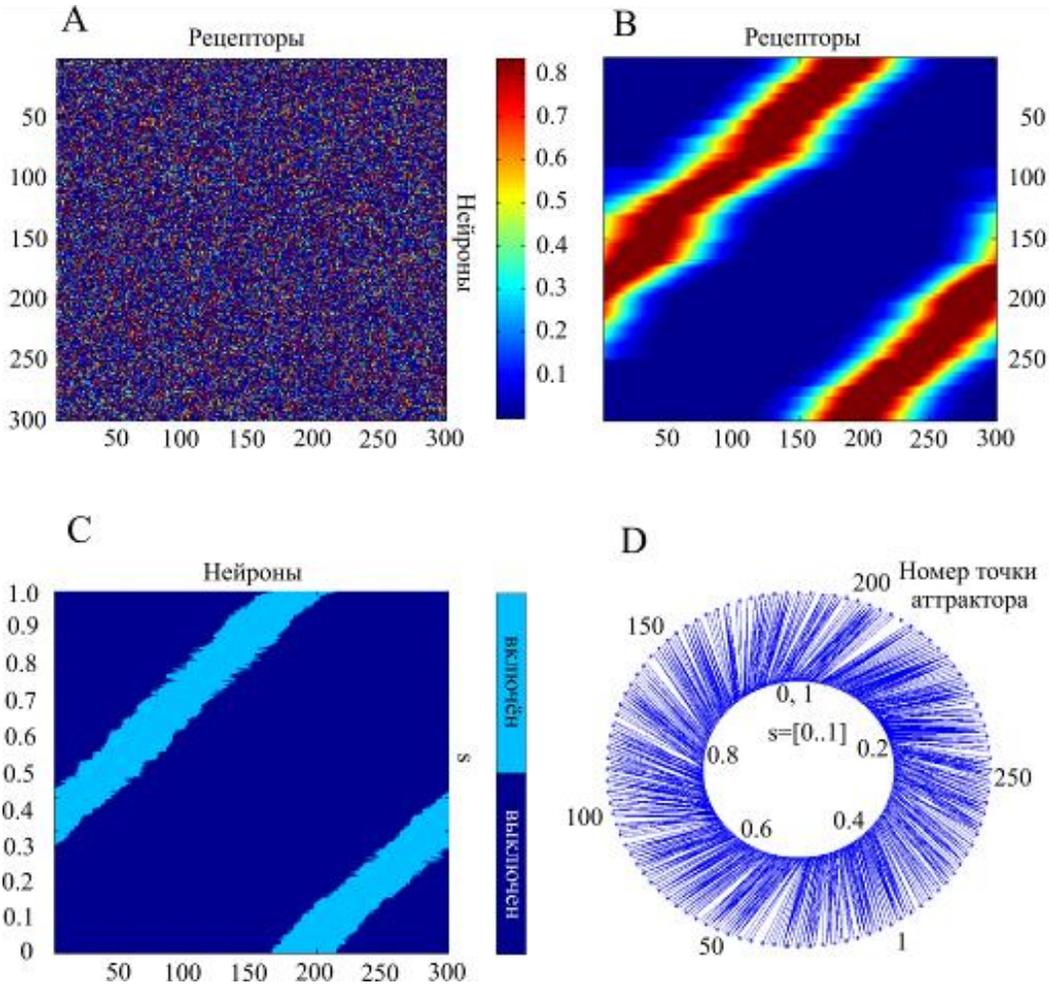
В отличие от процедуры обучения, в процедуре тестирования матрица  $W$  не меняется.

Каждой точке сигнала из пункта 1 ставится в соответствие точка кольцевого аттрактора из пункта 2 (см. режим обучения). Пример представления результатов представлен на рисунке 3.

На рисунке 3 представлено три способа графического представления состояния системы. Это, во-первых, матрица синаптических весов связей между рецепторами и нейронами (рис. 3, А, В), во-вторых, диаграмма активности нейронов при поступлении входного сигнала с последовательными значениями параметра  $s$  от  $0$  до  $1$  с шагом  $0.001$  (рис. 3, С) и, в-третьих, кольцевой график, на котором линией соединяются значение параметра  $s$ , задающего значения входных сигналов, и номер точки аттрактора, к которой сходится активность сети при данном значении  $s$  (рис. 3, D).

Элементы начальной матрицы  $W$  (рис. 3, А) принимают случайные значения в диапазоне от  $0$  до  $1$ . После  $2000$  итераций режима обучения матрица связей приняла приблизительно диагональный вид со сдвигом (рис. 3, В). Далее (рис. 3, С, D) система запускается в режиме тестирования. На вход подаются значения сигнала от  $0$  до  $1$  с шагом  $0.001$  и с порогом  $25$ . Рисунки 3, В и 3, С очень похожи, но имеют разный физический смысл. При предъявлении сигнала на рецепторы с новым значением параметра  $s$ , в течение  $20$  шагов система приходила в состояние равновесия. На рис. 3, С для каждого значения параметра  $s$  показана активность нейронов в течение всех  $20$  шагов, а цвет точки (бледно

или темно-синий) показывает возбуждение или не возбуждение нейрона с номером, обозначенным по горизонтали. В то время, как на рис. 3, В линия соединяет значение параметра  $s$  только с номером точки аттрактора, соответствующей последнему (двадцатому) состоянию сети.



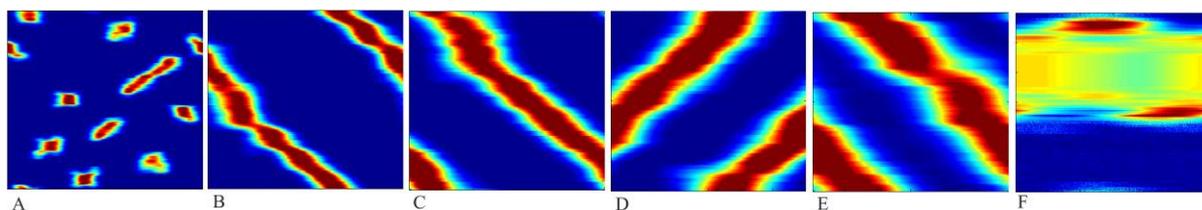
**Рис. 3.** Методы представления состояния системы после обучения. (А) Матрица связей  $W$  между нейронами и рецепторами до обучения. Значения начальной матрицы  $W$  принимают случайные значения в диапазоне от 0 до 1; (В) Матрица связей  $W$  между нейронами и рецепторами после обучения; (С) Диаграмма активности нейронов при поступлении входного сигнала с последовательными значениями параметра  $s$  от 0 до 1 с шагом 0.001 (подробности в тексте). По вертикальной оси – значение параметра  $s$ , по горизонтальной – номера нейронов, каждому значению  $s$  соответствуют 20 строчек матрицы, за которые осуществляется переход в аттрактор при данном  $s$ ; (D) На внутреннем круге отмечается значение параметра входного сигнала рецептора; на внешнем круге – средний номер горящего нейрона в точке аттрактора, полученной после запуска системы с данным входным параметром.  $L = 45$ ,  $D = 45$ ,  $\sigma = 10$ ,  $\eta = 0.1$ . Обучение проводилось при величине порога 20. При такой же величине порога проводилось тестирование. Число нейронов  $N = 300$ , число рецепторов  $R = 300$ .

### 3. РЕЗУЛЬТАТЫ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ

#### Сходимость процесса обучения

Вид матрицы  $\mathbf{W}$ , в котором не наблюдается значительных отклонений от диагонального образа со сдвигом, мы будем называть «правильным видом матрицы  $\mathbf{W}$ » (рис. 3, *B*). Состояние системы с таким видом матрицы связей будет, соответственно, называться «правильным схождением отображения». Была проведена серия вычислительных экспериментов, в которых матрица связей  $\mathbf{W}$  оценивалась периодически через определенное число шагов итераций изменения связей. «Правильность» вида матрицы оценивалась визуально. Шаг просмотра матрицы выбирался так, чтобы можно было определить такое  $i$ , что на  $i$ -м шаге система уже «правильно сошлась», а на  $(i-1)$ -м шаге имеются очевидные отклонения от «правильности». Погрешность такого метода – порядка шага просмотра.

Рис. 4 иллюстрирует результаты данной серии опытов. При фиксированном числе нейронов и рецепторов (по 300), одновременно варьировались значения  $L$  и  $D$ . Величина  $\mu = L/N = D/R$ .

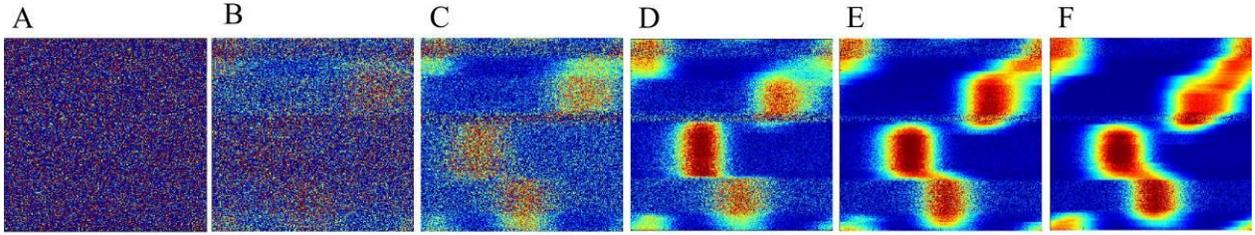


**Рис. 4.** Матрицы связей рецепторов с нейронами  $\mathbf{W}$  в зависимости от параметра и числа итераций процесса обучения. (A)  $\mu = 0.05$ , число итераций 31800; (B)  $\mu = 0.1$ , число итераций 9200; (C)  $\mu = 0.15$ , число итераций 2320; (D)  $\mu = 0.2$ , число итераций 760; (E)  $\mu = 0.25$ , число итераций 200; (F)  $\mu = 0.3$ , число итераций 2050. Число нейронов  $N = 300$ , число рецепторов  $R = 300$ . Обозначения такие же, как на рис. 3, *B*.

Как видно из результатов моделирования, при  $\mu = 0.05$  наблюдалась только фрагментарная сходимость. В матрице связей присутствовало много локальных очагов ответа на внешний стимул, значительно разнесённых друг от друга (рис. 4, *A*). При значениях параметра  $\mu$  от 0.1 до 0.25 отображение сходится (рис. 4, *B–E*), и скорость сходимости возрастает с ростом  $\mu$ . При  $\mu \geq 0.3$  отображение не сходится. Это происходит потому, что зоны реакции от разных сигналов существенно перекрываются, и поэтому на любой стимул реагирует одна и та же группа нейронов (рис. 4, *F*).

Качественно поведение сходимости системы можно описать следующим образом. Разделим процесс обучения на две стадии. На первой стадии определяются зоны ответа для всех рецепторных сигналов. Если изначально матрица связей  $\mathbf{W}$  случайная (рис. 5, *A*), то после прохождения первой стадии для каждого нейрона появляется одна область максимума, где сила связей становится близка к единице, а все остальные связи данного нейрона с рецепторами становятся близки к нулю (рис. 5, *F*). Чисто визуально, первая стадия процесса обучения матрицы заканчивается образованием дискретных компактных пятен больших коэффициентов связей матрицы (рис. 5, *F*).

На второй стадии обучения идет процесс соединения отдельных компактных фрагментов матрицы в конечную «правильную» диагональ. Скорость схождения системы зависит от конфигурации матрицы  $\mathbf{W}$ , полученной на первой стадии.



**Рис. 5.** Матрица связей нейронов с рецепторами  $\mathbf{W}$  в процессе прохождения первой стадии обучения.  $\mu = 0.15$ , состояние матрицы  $\mathbf{W}$  после (A) 0; (B) 20; (C) 40; (D) 60; (E) 80; (F) 100 итераций. Число нейронов  $N = 300$ , число рецепторов  $R = 300$ , обозначения те же, что на рис. 3 С.

Таким образом, процесс сходимости самоорганизующегося отображения к диагональной форме естественным образом делится на два этапа. На первом этапе матрица отображения из некоррелированного случайного состояния приходит к блочно-упорядоченному виду (рис. 5, F). Блоки матрицы выглядят примерно одинаково, их размер определяется параметром  $\mu$ , а их число  $q$  обратно пропорционально параметру  $\mu$ . Скорость сходимости отображения на втором этапе в основном зависит от взаимного расположения упорядоченных блоков на первом этапе. Причём вычислительный эксперимент показывает, что начальное расположение упорядоченных блоков равновероятно для всех возможных комбинаций упорядоченных блоков. Скорость сходимости отображения на втором этапе зависит от того, насколько «удачна» сложившаяся на первом этапе комбинация блоков. Если блоки «выстроились в диагональ», то финальная упорядоченность достигается быстро. Если комбинация блоков на первом этапе далека от диагональной, то достижение диагональной упорядоченности занимает длительный период времени.

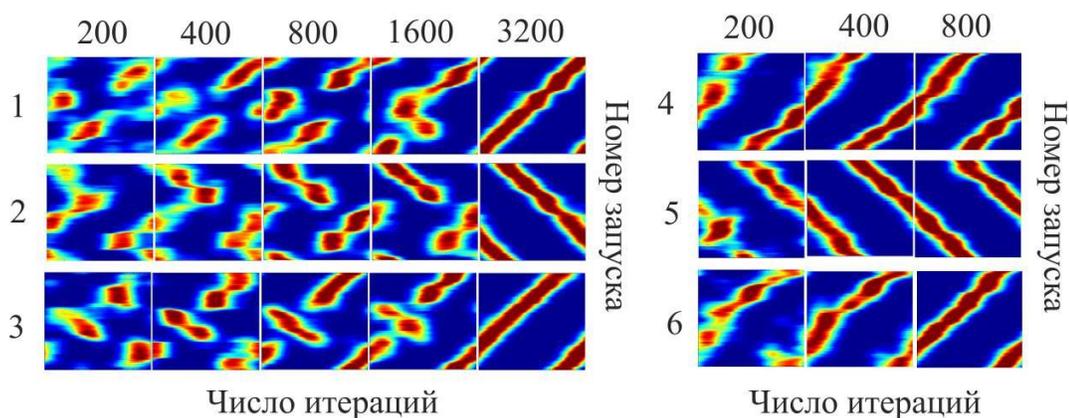
Исходя из упомянутой равновероятности расположения блоков на первом этапе, можно оценить вероятность «выигрышного» для второго этапа расположения упорядоченных блоков матрицы в конце первого этапа, то есть такого варианта расположения, когда блоки выстраиваются в диагональ. Эта вероятность  $P_{\text{выигр}}$  равна

$$P_{\text{выигр}} = \frac{2}{(q-1)!} \quad (4),$$

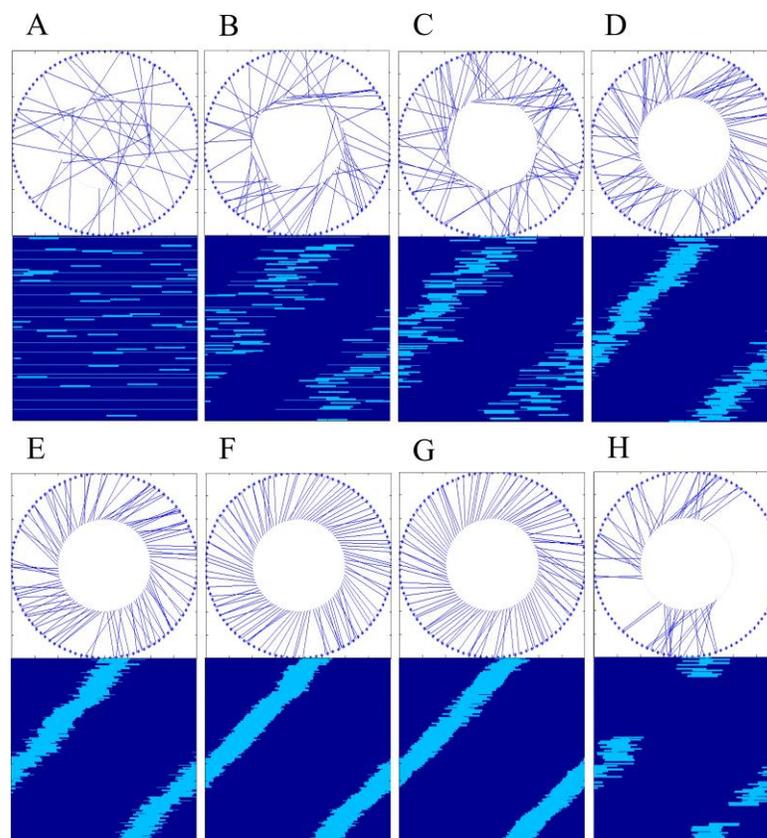
для  $q > 2$ , и  $P_{\text{выигр}} = 1$  для  $q \leq 2$ . При  $q = 3$  эта величина ещё равна 1. Но уже при  $q = 5$   $P_{\text{выигр}} = 1/60$ . При больших значениях  $q$  скорость сходимости процесса обучения медленная. Кроме того, для разных начальных конфигураций блочной матрицы второго этапа, различна скорость сходимости к диагональной упорядоченности. В силу этого, при больших  $q$  велика дисперсия времени достижения финальной упорядоченности.

На рис. 6 представлена динамика изменения матрицы связей рецепторов с нейронами  $\mathbf{W}$  во время «второй стадии» обучения для шести разных реализаций процедуры обучения. Все шесть вариантов процессов отличаются только реализацией случайных величин в ходе эволюции системы. Запуски сгруппированы по скорости сходимости обучения. Первые три запуска (слева: запуски 1–3) проходят стадию, когда половина матрицы  $\mathbf{W}$  упорядочена в одну сторону, а половина в другую. Последние три запуска (справа: запуски 4–6) сразу после первой стадии принимают вид, близкий к диагональному. Таким образом, характерный вид матрицы  $\mathbf{W}$  для такой системы с  $\mu = 0.15$  схож с видом матрицы  $\mathbf{W}$  для

блочной матрицы с  $q = 4$ . Скорость сходимости обучения во второй группе запусков в четыре – восемь раз выше, чем в первой.



**Рис. 6.** Матрицы связей рецепторов с нейронами  $\mathbf{W}$  во время «второй стадии» обучения для шести разных запусков процедуры обучения. Запуски упорядочены. Первые три запуска (слева: запуски 1–3) проходят стадию, когда половина матрицы  $\mathbf{W}$  упорядочена в одну сторону, а половина в другую. Последние три запуска (справа: запуски 4–6) сразу после первой стадии принимают вид, близкий к диагональному. Число нейронов  $N = 300$ , число рецепторов  $R = 300$ ,  $\mu = 0.15$ .



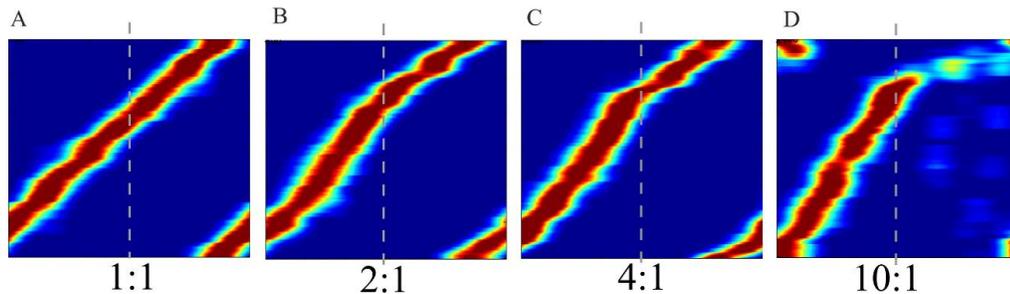
**Рис. 7.** Отображения для разных значений порога при возбуждении нейронов от рецепторов, для матрицы  $\mathbf{W}$  (рис. 3), полученной в процессе обучения (метод отображения как на рис. 3, C, D). Значения порога: (A) 0; (B) 5; (C) 10; (D) 15; (E) 20; (F) 25; (G) 30; (H) 35. Число нейронов  $N = 300$ , число рецепторов  $R = 300$ .

Качество полученного отображения определяется не только видом матрицы связей нейронов с рецепторами  $\mathbf{W}$ , но и параметрами воспроизведения в режиме тестирования. В частности, качество воспроизведения при одной и той же матрице  $\mathbf{W}$  может варьироваться в зависимости от величины порога при передаче сигнала от рецепторов к нейронам (в пункте 2 режима тестирования каждой итерации) (рис. 7). При увеличении порога до некоторого предельного значения, точность отображения возрастает. При превышении предельного значения порога не все точки внешнего сигнала имеют соответствующую точку на непрерывном аттракторе.

### Свойство масштабируемости представления отдельных участков отображаемого сигнала

Одним из основных свойств топических отображений в нервной системе является то, что наиболее часто используемые рецептивные области представлены более широко на более высоких уровнях [8].

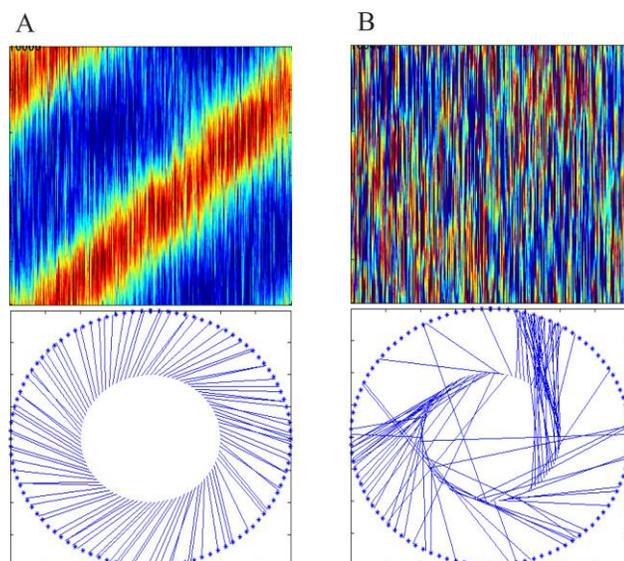
Были проведены модельные эксперименты, в которых сигнал первой половины входного многообразия ( $s < 1/2$ ) подавался в 1, 2, 4 и 10 раз чаще, чем второй ( $s > 1/2$ ) (рис. 8). При этом область, на которую отображается первая половина рецептивного поля, увеличивается согласно увеличению частоты ее использования. При очень редком использовании второй области относительно первой (как в случае 10:1), когда размер потенциальной области ответа нейронной сети на вторую область рецептивного поля становится меньше размера точки непрерывного аттрактора, четкая область ответа в этом поле пропадает совсем (рис. 8,D).



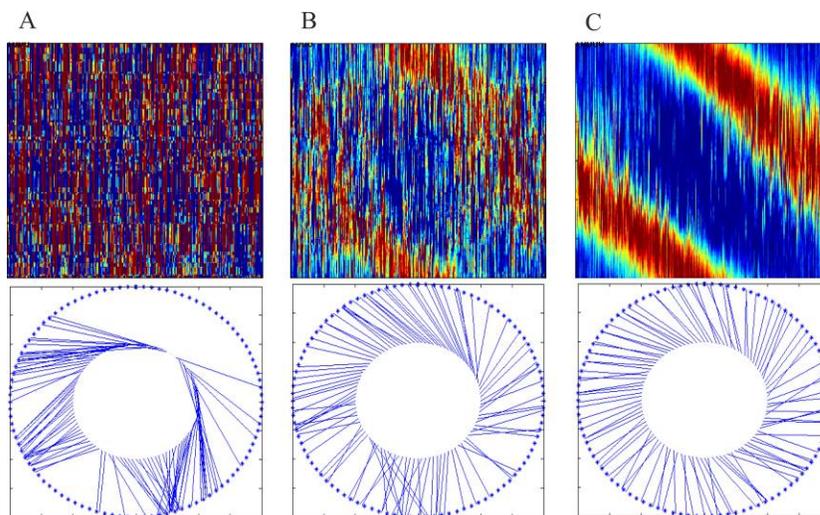
**Рис. 8.** Матрица связей нейронов с рецепторами  $\mathbf{W}$ , для не равновероятного использования первой и второй половины значения входного сигнала. Соотношения вероятностей сигнала в первой половине к сигналу во второй половине: (A) 1:1; (B) 2:1; (C) 4:1; (D) 10:1. Число нейронов  $N = 300$ , число рецепторов  $R = 300$ .

### Свойство устойчивости к шуму отображения

Следующим важным свойством представленного отображения является его высокая устойчивость к шуму входного сигнала. Были проведены модельные эксперименты, в которых в процессе обучения сигнал в систему подавался с белым шумом, в несколько раз превышающим по амплитуде сигнал. Иначе говоря, значение возбуждения  $v$ -го рецептора определяется как  $\mathbf{V}_v = G(D, s, v) + A\xi$ , где  $G$  определяется по формуле (3),  $\xi$  – случайная величина, равномерно распределённая на отрезке  $[-1, 1]$ , независимая в разные моменты времени,  $A$  – постоянная для данного эксперимента величина. При одном и том же протоколе обучения, отображение сигнала с пятикратной амплитудой белого шума (при  $A = 5$ ) принимало правильный вид (рис. 9, A), тогда как отображение сигнала с десятикратной амплитудой белого шума (при  $A = 10$ ) уже не принимало правильного вида (рис. 9, B).



**Рис. 9.** На верхних рисунках показана матрица  $\mathbf{W}$ , на нижних – кольцевой график соответствий параметра  $s$  входного множества и точки непрерывного аттрактора рекуррентной сети (метод изображения, как на рис. 3, *D*). (A) Отображение сигнала с амплитудой белого шума, наложенного на сигнал, в 5 раз превышающей амплитуду сигнала. 1000 итераций обучения с коэффициентом обучения  $\eta = 0.1$ , затем 1000 итераций с  $\eta = 0.01$  и затем ещё 20000 итераций с  $\eta = 0.001$ ; (B) Амплитуда шума, в 10 раз превышающая амплитуду входного сигнала. 1000 итераций обучения с коэффициентом обучения  $\eta = 0.1$ , затем 1000 итераций с  $\eta = 0.01$  и затем ещё 20000 итераций с  $\eta = 0.001$ . Число нейронов  $N = 300$ , число рецепторов  $R = 300$ .



**Рис. 10.** Три стадии обучения системы, отображающей сигнал с уровнем шума, меняющимся синусоидально от нулевого уровня (в центре рецепторного сигнала) до уровня амплитуды шума, в десять раз превышающей уровень сигнала (по краям рецепторного сигнала). На верхних рисунках изображена матрица  $\mathbf{W}$ , на нижних – кольцевой график соответствий параметра  $s$  входного множества и точки непрерывного аттрактора рекуррентной сети (метода изображения, как на рис. 3, *D*). (A) Первая стадия. 1000 итераций с коэффициентом обучения  $\eta = 0.1$ , начиная со случайной матрицы  $\mathbf{W}$ ; (B) Вторая стадия. 1000 итераций с коэффициентом обучения  $\eta = 0.01$  после окончания первой стадии, изображенной на рис. 10, A; (C) третья стадия. 20000 итераций с коэффициентом обучения  $\eta = 0.001$  после окончания второй стадии, изображенной на рис. 10, B. Число нейронов  $N = 300$ , число рецепторов  $R = 300$ .

Если уровень шума меняется вдоль входной кривой ( $A = A(s)$ ), то благодаря быстрой сходимости отображения в участках, соответствующих низкому уровню шума, образуется “центр кристаллизации” и все отображение сходится быстрее. На рисунке (рис. 10) показано три стадии обучения системы, отображающие сигнал с уровнем шума, меняющимся синусоидально от нулевого уровня (в центре) до уровня амплитуды шума, в десять раз превышающей уровень сигнала (по краям). Здесь:  $A = 10 \cdot (\cos(2\pi s) + 1)$ .

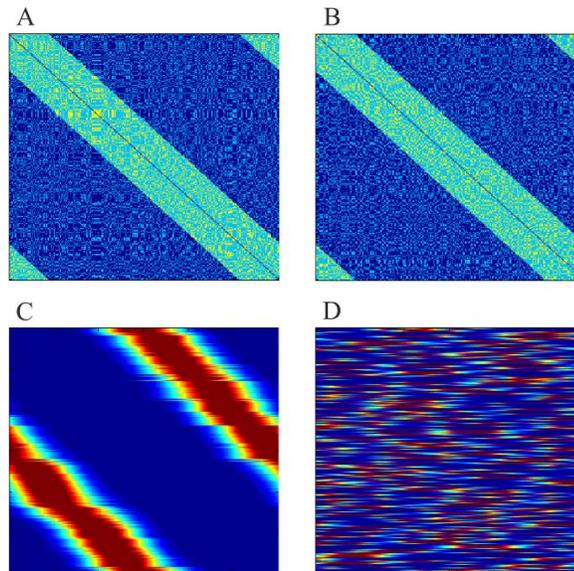
### Свойство связности отображения в сетях, содержащих непрерывный аттрактор из нескольких компонент связности

Поскольку нейронные сети, содержащие в себе аттракторные структуры, предполагают возможность существования одной или нескольких компонент связности аттракторов, интересно изучить вопрос о том, как отображается связное множество на систему, содержащую непрерывный аттрактор с более чем одной компонентой связности.

Далее рассматривается отображение пространства рецепторов, содержащее одно однопараметрически заданное кольцевое множество, на нейронную сеть, содержащую два кольцевых непрерывных аттрактора [4].

Сеть, обладающая заявленным свойством, формируется следующим образом:

На сети нейронов вводятся две случайные нумерации – первая и вторая, где «нумерация» – это перестановка номеров нейронов. Если у двух нейронов номера отличаются меньше чем на  $2L$  в обеих нумерациях, то между ними устанавливается сила синаптической связи 2. Если в обеих нумерациях номера нейронов отличаются больше, чем на  $2L$ , то вес синаптической связи между ними будет равен  $-\sigma$ . В случае, когда номера нейронов отличаются меньше чем на  $2L$  только в одной из двух нумераций, вес синаптической связи между ними устанавливается равным 1. Иначе говоря, связи нейронов задаются как суперпозиция синаптических весов в двух нумерациях. В каждой из этих нумераций связи задаются наподобие «мексиканской шляпы»: нейрон связан положительными связями с  $L$  соседями с обеих сторон и отрицательными со всеми остальными.

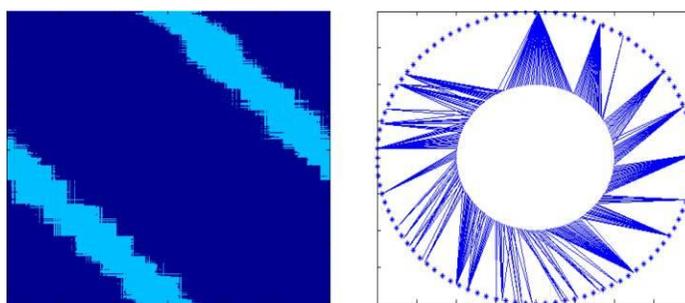


**Рис. 11.** (A) Матрица связей нейронов  $T$  в первой нумерации; (B) Матрица связей нейронов  $T$  во второй нумерации. На (A) и (B) цветом отмечены силы связей: тёмно-синее  $-10$ , голубое 1, жёлтое 2; (C) Матрица  $W$  после обучения в первой нумерации; (D) Вид той же матрицы  $W$  после обучения во второй нумерации. Число нейронов  $N = 300$ , число рецепторов  $R = 300$ .

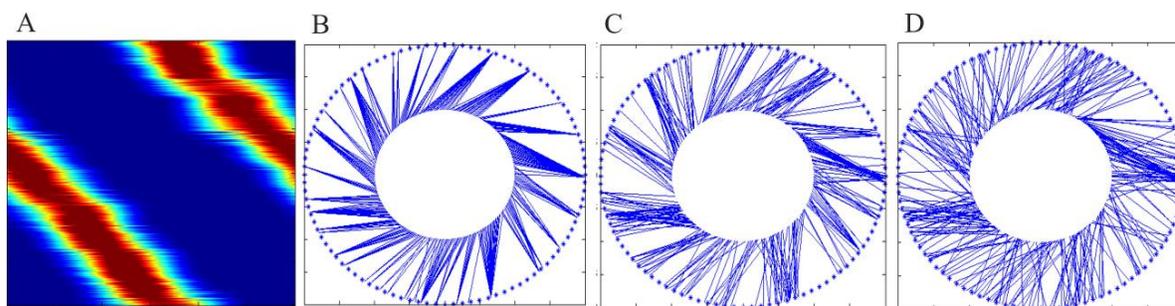
В первой нумерации связи второй нумерации выглядят как постоянный шум, и наоборот (рис. 11, *A, B*).

После процесса обучения матрица связей нейронов с рецепторами в одной нумерации приобретает упорядоченный, а в другой – случайный вид (рис. 11, *C, D*). Таким образом, вход с рецепторов полностью отображается на один из аттракторов сети. На рисунке 12 показана активность нейронов в первой нумерации при постепенном изменении  $s$  и соответствующий график соответствия параметра входа с номером аттрактора в первой нумерации. Как можно заметить, при отображении происходит «склейка» – когда используется небольшое число возможных аттракторов сети. Это происходит из-за влияния второго кольца.

Для того, чтобы сгладить неровности энергетического рельефа, мы добавили в систему во время обучения «синаптический шум» – шум синаптических весов между нейронами рекуррентной сети (рис. 13). При этом воспроизведение происходило без шума (рис. 13, *B*), с шумом, меньшим, чем при записи (рис. 13, *C*), и при таком же уровне шума (рис. 13, *D*). С увеличением уровня шума воспроизведения количество используемых точек аттрактора растет, но точность воспроизведения при этом ухудшается.



**Рис. 12.** Отображение, полученное по матрице  $W$  (рис. 11) в первой нумерации. Слева: активность нейронов, метод отображения как на рис. 3, *C*, Справа: кольцевая диаграмма, метод отображения как на рис. 3, *D*. Число нейронов  $N = 300$ , число рецепторов  $R = 300$ .



**Рис. 13.** (A) Матрица связей рецепторов с нейронами  $W$  после обучения системы, содержащей два непрерывных кольцевых аттрактора, при наличии «синаптического шума»; (B) Тестирование отображения по матрице рис. 13, A без шума; (C) Тестирование отображения по матрице рис. 13, A с меньшим уровнем шума, чем при обучении; (D) Тестирование отображения по матрице рис. 13, A с таким же уровнем шума как и при обучении. «Синаптический шум» представляет собой прибавление случайной величины к 20% случайно выбранным связям нейронов рекуррентной сети на каждом шагу. Значения этой случайной величины: (A) При обучении от  $-1$  до  $1$ ; (C) При тестировании с меньшим уровнем шума – от  $-0.5$  до  $0.5$ ; (D) При тестировании с уровнем шума, как при обучении – от  $-1$  до  $1$ . Число нейронов  $N = 300$ , число рецепторов  $R = 300$ .

#### 4. ЗАКЛЮЧЕНИЕ

В данной работе была представлена конструкция, состоящая из формальных нейронов, которая осуществляет самоорганизующееся отображение  $R$ -мерного рецепторного пространства в пространство состояний нейронной сети. В том случае, когда в пространстве рецепторов многообразие возможных сигналов имеет одномерную топологию, данная система позволяет получать топологически корректное отображение этого многообразия на непрерывный аттрактор рекуррентной сети.

Был исследован характер поведения скорости сходимости отображения от значения параметров системы.

Показано, что представленная конструкция обладает следующими свойствами:

- 1) Масштабируемость представления отдельных участков отображаемого сигнала.
- 2) Устойчивость формирования всего и отдельных частей отображения к шуму входного сигнала.
- 3) Односвязность отображения в сетях, содержащих непрерывный аттрактор из нескольких компонент связности.

Представленная конструкция схожа с нейрообразной конструкцией Тейво Кохонена [3], но обладает рядом существенных отличий.

Во-первых, наша конструкция обладает физиологическим правдоподобием. В частности, в то время, как в системе Кохонена отображение точки входного пространства фактически происходит на один нейрон (принцип «winner takes all»), в нашей системе «победителями» оказывается группа нейронов. В данной работе рассматривались непрерывные аттракторы в сети длины не больше  $N$ . Описаны нейронные сети с непрерывными аттракторами длины  $2N$  и более [4]. При использовании подходящего закона обучения и аттрактора длины больше  $N$ , возможно сделать разрешающую способность системы, то есть число дискретов, на которые отображается непрерывный внешний сигнал, больше  $N$ , где  $N$  – число нейронов рекуррентной сети.

Во-вторых, реверберационные процессы в аттракторе позволяют системе самостоятельно восстанавливать нейрональную окрестность, которая в данном случае и является точкой аттрактора. В отличие от системы Кохонена, где окрестность определяется «вручную», по соотношению топологической близости нейрообразных ячеек, в данной системе нейрональная окрестность появляется в качестве устойчивого состояния, содержащего данный нейрон.

Таким образом, имитационное моделирование, проведённое в нашей работе, показало, что сети, содержащие непрерывный аттрактор, могут быть использованы для представления непрерывных переменных в нейронных системах.

Существуют и другие работы, где непрерывные аттракторы используются для представления таких непрерывных величин, как, например, направление головы [9] или координаты в пространстве [10]. Но, как правило, в таких работах обходится стороной вопрос о том, как в таких системах точки аттрактора принимают значения, соответствующие тому или иному внешнему стимулу. Тем не менее, в работе С. Штрингера и других авторов [11] рассмотрено самоорганизующееся формирование непрерывного аттрактора под действием внешнего входа. В этой работе связи в сети, на которую осуществляется отображение, изначально случайные. В отличие от этой работы, в нашей работе рассматривалась такая сеть, в которой непрерывный аттрактор генетически «зашит», или предформирован, а не формируется ассоциативно. В такой системе в процессе обучения изменяются связи между нейронами двух различных слоёв, а не между нейронами отдельного слоя.

Автор выражает благодарность В.Л. Дунину-Барковскому за постановку задачи, Я.М. Карандашеву, В.Б. Котову и Р.Р. Алиеву за обсуждение рабочих моментов и результатов моделирования.

Статья подготовлена при поддержке РФФИ, грант № 13-07-01004.

### СПИСОК ЛИТЕРАТУРЫ

1. Dunin-Barkowski W. Data formats in multi-neuronal systems and brain reverse engineering. Invited talk, Conference *Biologically Inspired Cognitive Architecture 2011*. URL: <http://2045.ru/pdf/ARLINGTON20111106wldb.pdf> (дата обращения 10.04.2013)
2. Kass J.H. Topographic Maps are Fundamental to Sensory Processing. *Brain Research Bulletin*. 1997. V. 44. № 2. P. 107–112.
3. Kohonen T. Self-Organized Formation of Topologically Correct Feature Maps. *Biological Cybernetics*. 1982. V. 43. № 1. P. 59–69.
4. Dunin-Barkowski W., Osovets N. Hebb-Hopfield neural networks based on one-dimensional sets of neuron states. *Neural Processing Letters*. 1995. V. 2. № 5. P. 28–31.
5. Knierim J.J., Zhang K. Attractor dynamics of spatially correlated neural activity in the limbic system. *Annu. Rev. Neurosci.* 2012. V. 35. P. 267–285.
6. Мак-Каллок У.С., Питтс В. Логическое исчисление идей, относящихся к нервной активности. В: *Автоматы*. Под ред. Шеннона К.Э. и Маккарти Дж. М.: Изд-во иностр. лит., 1956. С. 363–384. (Перевод английской статьи 1943 г.)
7. Amari S.I. Dynamics of pattern formation in lateral-inhibition type neural fields. *Bio. Cyber.* 1977. V. 27. P. 77–87.
8. Candia V., Rosset-Llobet J., Elbert T., Pascual-Leone A. Changing the brain through therapy for musicians' hand dystonia. *Ann N Y Acad Sci.* 2005. V. 1060. P. 335–342.
9. Zhang K. Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: A theory. *Journal of Neuroscience*. 1996. V. 16. P. 2112–2126.
10. Tsodyks M. Attractor neural network models of spatial maps in hippocampus. *Hippocampus*. 1999. V. 9. № 4. P. 481–489.
11. Stringer S.M., Rolls E.T., Trappenberg T.P., de Araujo I.E.T. Self-organizing continuous attractor networks and path integration: two-dimensional models of place cells. *Network: Comput. Neural Syst.* 2002. V. 13. P. 429–446.

Материал поступил в редакцию 01.04.2013, опубликован 27.05.2013.