

в) лучевая aberrация изображения разрядной ячейки за счет увеличения первичных aberrаций фурье-объективов (17). Требования к согласованию радиусов кривизны фронта пучков зависят от параметров оптической схемы ГЗУ и, как правило, являются высокими.

3. Наиболее жесткие требования предъявляются к сферической aberrации и к кривизне поля объективов. В случае когда входной зрачок объектива восстановления помещается в плоскости матрицы голограмм, а размеры транспаранта данных и матрицы голограмм совпадают, сферическая aberrация и кривизна поля минимальные, а кома и дисторсия компенсируются соответственно частично и полностью. Неточность юстировки объективов приводит к увеличению их первичных aberrаций, причем быстрее растут компенсируемые aberrации (кома, дисторсия) (см. таблицу).

Авторы выражают благодарность Е. Ф. Пену, С. А. Шойдину, Л. Н. Горшковой за полезные обсуждения, а также И. С. Гибину за ценные советы и замечания, сделанные им при чтении рукописи.

ЛИТЕРАТУРА

1. Богданова Е. С., Соскин С. И. Влияние aberrаций оптической системы на емкость голографической памяти.— Автометрия, 1975, № 3.
2. Соскин С. И., Шойдин С. А. Расчет оптической схемы голографического запоминающего устройства.— Опт. и спектр., 1978, т. 44, № 3.
3. Соскин С. И., Шойдин С. А. Оптимизация параметров голографического запоминающего устройства с учетом aberrаций.— Опт. и спектр., 1978, т. 44, № 6.
4. Acos Gy., Kiss G., Varga P. Effect of Lens Aberrations on the Storage Capacity of Holographic Memories.— Opt. Comm., 1977, vol. 20, N 1, p. 63.
5. Гибин И. С. и др. Ассоциативная выборка информации в голограммных запоминающих устройствах.— Автометрия, 1973, № 5.
6. Scamor R. Alignment and Desing Conditional for Optical Memory Systems Using Fouries Holograms.— Optica Acta, 1978, vol. 25, N 2, p. 107.
7. Домбровский В. А., Домбровский С. А., Пен Е. Ф. Влияние параметров гауссового опорного пучка на искажение восстановленного из голограммы изображения.— Опт. и спектр., 1978, т. 45, № 5.
8. Джеррард А., Берч Дж. М. Введение в матричную оптику. М.: Мир, 1978.
9. Когельник Г., Ли Т. Резонаторы и световые пучки лазеров.— ТИИЭР, 1966, т. 54, № 10.
10. Выдрин Л. В. и др. Устройство автоматической записи матриц голограмм.— Автометрия, 1980, № 2.
11. Слюсарев Г. Г. Расчет оптических систем. Л.: Машиностроение, 1975.

Поступила в редакцию 16 декабря 1981 г.

УДК 612.8 : 681.5

И. П. МУРАВЬЕВ, Д. М. РОМАНОВ, А. А. ФРОЛОВ

(Москва)

СТОХАСТИЧЕСКИЙ АССОЦИАТИВНЫЙ ФИЛЬТР

Проектировщики современной вычислительной техники все чаще обращаются к результатам нейрофизиологических исследований, в которых изучаются основные особенности переработки информации в живых системах [1]. Весьма перспективным с точки зрения приложения к технике является стохастический квазиголографический подход к анализу памяти, разрабатываемый начиная с работы Бриндли [2]. В дальнейшем [3—5] этот подход был согласован с данными физиологических экспериментов о функциях и морфологической структуре отдельных областей мозга. В [5] постулировано, что гиппокамп выполняет функцию памяти, обеспечивающей восстановление запомненного ранее полного паттерна нейронной активности после предъявления его части (автоассоциативное воспроизведение [1]). В [1] эта функция обобщена для случая, когда

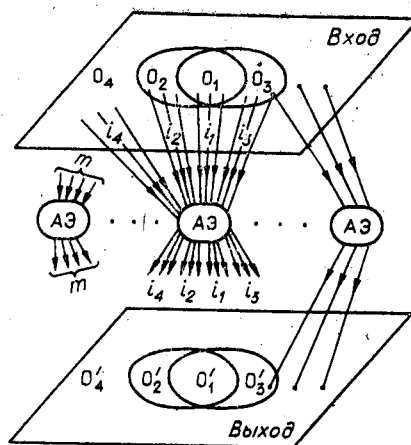


Рис. 1.

возможность использования этого механизма пластичности для построения ассоциативного фильтра.

Структура и функции ассоциативного фильтра. Структура рассматриваемого ассоциативного фильтра сходна со структурой трехслойного перцептрона [8] и показана на рис. 1. Он состоит из слоев двоичных входных, ассоциативных и выходных элементов, стохастически связанных между собой, причем так же, как в [2, 6, 7], предполагается, что связи между выходными и ассоциативными элементами (АЭ) являются зеркальным отражением связей между входными элементами и АЭ. Число входов и соответственно выходов у каждого АЭ фиксировано и равно m , а сами эти связи распределены равномерно. Ассоциативный фильтр должен обладать способностью восстанавливать на выходном слое любой из L заполненных эталонных паттернов, содержащих фиксированное число l элементов входного слоя, равномерно и статистически независимо распределенных по этим элементам. Каждый предъявляемый к восстановлению зашумленный паттерн должен содержать не менее λl «своих» элементов, равномерно и статистически независимо распределенных по элементам эталонного паттерна, и не более μl «чужих» элементов, равномерно и статистически независимо распределенных по элементам его дополнения; кроме того, элементы предъявляемого паттерна будем считать статистически независимыми от других эталонных паттернов.

Способность к записи и воспроизведению информации основана на модификации передаточной функции АЭ. Предполагается, что у немодифицированных АЭ в процессе воспроизведения передаточная функция всегда равна нулю; у модифицированных она равна нулю, если число активных входов из общего числа m , связанных с данными АЭ, не достигает порогового числа k_1 , в противном случае передаточная функция равна 1. Модификация АЭ происходит, если при записи число активных конвергирующих на него входов не меньше порогового числа k_0 , хотя бы для одного эталонного образа. В режиме воспроизведения возможны ошибки двух типов: активация на выходе элементов, не принадлежащих эталонному паттерну, и потеря его элементов. Целью дальнейшего анализа являются расчет числа АЭ N , необходимого для работы системы с заданными вероятностями ошибок воспроизведения, поиск оптимальных значений параметров m , k_0 и k_1 , обеспечивающих минимум N , и сравнение полученных оценок информационной емкости рассматриваемой стохастической системы с пределами, устанавливаемыми теорией информации.

Анализ свойств ассоциативного фильтра. Предположения о фиксированности числа элементов эталонных паттернов l и числа связей АЭ m

зашумление восстанавливаемого паттерна проявляется не только в потере активности у части его элементов, но и в появлении активности у элементов, не входящих в паттерн. Так как в данном случае функция ассоциативной памяти заключается фактически в выделении и восстановлении ранее запомненных эталонных дискретных сигналов при наличии шума, то такую память можно назвать ассоциативным фильтром.

В работах [6, 7] рассмотрены модели нейронных сетей, способных к замыканию условных рефлексов и осуществляющих функцию автоассоциативного воспроизведения на основе изменения реактивности нервной клетки. В настоящей работе рассмотрена воз-

обеспечивают статистическую независимость модификации и активации каждого в отдельности АЭ несколькими произвольными паттернами и статистическую независимость модификации и активации нескольких произвольных АЭ каждым в отдельности паттерном. Однако статистическая независимость не выполняется для нескольких элементов, активируемых или модифицируемых несколькими паттернами. В приложении проведен расчет корреляций между такими событиями и показано, что для параметров сети, близких к оптимальным, и $\rho = \lambda^2/(\lambda + \mu) > 0,25$ они дают несущественную поправку оценки свойств системы. В дальнейшем анализе они не учитываются. При подаче на входной слой полного паттерна число связей каждого элемента ассоциативного слоя с активированными элементами входного описывается гипергеометрическим распределением, и вероятность модификации некоторого АЭ равна $q =$

$$= \sum_{k \geq k_0}^m C_m^k C_{n-m}^{l-k} / C_n^l.$$

В сделанных предположениях о статистической независимости вероятность модификации некоторого АЭ после записи L паттернов [7]

$$\kappa = 1 - (1 - q)^L \approx 1 - e^{-Q}, \quad Q = qL. \quad (1)$$

Назовем (см. рис. 1) совокупность «зажигшихся» λl элементов вспомняемого кода областью O_1 , «негорящих» $(1 - \lambda)l$ элементов — O_2 , μl «шумовых горящих» — O_3 и все остальные — O_4 , а соответствующие области выходного слоя — O'_k . Вероятность того, что данный АЭ связан с i_1 элементами из области O_1 , i_2 элементами из O_2 и i_3 элементами из O_3 , равна

$$P(i_1, i_2, i_3) = C_{\lambda l}^{i_1} C_{(1-\lambda)l}^{i_2} C_{\mu l}^{i_3} C_{n-i_1-i_2-i_3}^{m-i_1-i_2-i_3} / C_n^m. \quad (2)$$

Тогда вероятность P_k ($k = 1, \dots, 4$) того, что этот АЭ активирован и имеет связь с данным элементом из области O_k ,

$$P_k = \sum_{\substack{k_1 < i_1 + i_3 < m, \\ k_0 < i_1 + i_2 < m, \\ i_1 + i_2 + i_3 < m}} P(i_1, i_2, i_3) (i_k / Q_k) + \kappa \sum_{\substack{k_1 < i_1 + i_3 < m, \\ i_1 + i_2 < k_0, \\ i_1 + i_2 + i_3 < m}} P(i_1, i_2, i_3) (i_k / Q_k), \quad (3)$$

где $i_4 = m - i_1 - i_2 - i_3$, а $Q_1 = \lambda l$, $Q_2 = (1 - \lambda)l$, $Q_3 = \mu l$, $Q_4 = n - (\lambda + \mu)l$ — количество элементов в областях O_k .

Предполагая, что N велико, число активированных связей, конвергирующих на элемент выходного слоя, можно положить распределенным по нормальному закону (с параметрами, зависящими от номера области O'_k , которой принадлежит этот элемент). Для элемента из области O'_k среднее число конвергирующих на нем активированных связей будет $M_k = NP_k$. В пренебрежении статистической зависимостью между активациями различных АЭ, обоснованном в приложении, дисперсия числа таких связей $D_k = NP_k(1 - P_k)$.

Хорошее разделение на выходном слое «своих» и «чужих» элементов восстанавливаемого паттерна возможно, если на них слабо пересекаются распределения чисел активных связей от АЭ. Однако оказывается, что это условие не выполняется для элементов областей O'_2 и O'_3 , поэтому для восстановления эталонного паттерна необходимо использовать два порога активации выходных элементов: k' для разделения элементов областей O'_1 и O'_3 (соответствующих «горящим» элементам входного слоя) и k'' для O'_2 и O'_4 (соответствующих «негорящим» элементам). Если

$$k' = M_1 - a_1 \sqrt{D_1} = M_3 + a_3 \sqrt{D_3} \quad (4)$$

и

$$k'' = M_2 - a_2 \sqrt{D_2} = M_4 + a_4 \sqrt{D_4}, \quad (5)$$

то вероятность ошибки воспроизведения для элементов из области O'_k оценивается функцией $\Phi(a_k)$, где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt$. Вероятность появления ошибки хотя бы для одного выходного элемента составляет

$$P_0 \approx 1 - \exp\left(-\sum_{k=1}^4 \Phi(a_k) Q_k\right) \approx \sum_{k=1}^4 \Phi(a_k) Q_k, \quad (6)$$

если $P_0 \ll 1$. В настоящей работе ассоциативный фильтр считается оптимальным, если выбором его параметров m, k_0, k_1, k', k'' минимизирована величина P_0 при фиксированном значении N . Далее рассмотрен случай, когда

$$\lambda m/n \gg 1, \quad (1-\lambda)lm/n \gg 1, \quad \mu lm/n \gg 1, \quad lm/n \gg \ln L,$$

для которого возможна аппроксимация гипергеометрических распределений нормальными. Кроме того, предполагается, что

$$l/n \ll 1, \quad m/n \ll 1, \quad m/n \ll \mu, \quad ml/n^2 \ll 1 - \lambda.$$

В этом случае

$$P_1 - P_3 \approx P_2 - P_4 \approx \Delta = (1 - \kappa) \sqrt{m/2\pi \ln \Phi(c_2)} \exp(-c_0^2/2), \quad (7)$$

где $\Delta \ll P_1 \approx P_2 \approx P_3 \approx P_4 \approx P \approx (\kappa m/n) \Phi(c_1)$.

Здесь

$$\begin{aligned} c_0 &= (k_0 - lm/n) / \sqrt{lm/n}, & c_1 &= (k_1 - (\lambda + \mu)lm/n) / \sqrt{(\lambda + \mu)lm/n}, \\ c_2 &= (k_1 - \lambda k_0 - \mu lm/n) / \sqrt{\lambda(1-\lambda)k_0 + \mu lm/n} \approx (c_1 - c_0 \sqrt{\rho}) / \sqrt{1-\rho}, \\ & & \rho &= \lambda^2 / (\lambda + \mu). \end{aligned} \quad (8)$$

При фиксированных P и Δ минимум P_0

$$\begin{aligned} P_0 &= (4 \exp(-B^2/8) / B \sqrt{2\pi}) \{ \sqrt{Q_1 Q_3} \exp(-\ln^2(Q_3/Q_1)/2B^2) / \\ & / (1 - 4 \ln^2(Q_3/Q_1)/B^4) + \sqrt{Q_2 Q_4} \exp(-\ln^2(Q_4/Q_2)/2B^2) / \\ & / (1 - 4 \ln^2(Q_4/Q_2)/B^4) \} \leq P_0 = \sqrt{nl} \exp(-B^2/8) \end{aligned} \quad (9)$$

достигается при

$$\begin{aligned} a_1 &= (1/2)(B - 2 \ln(Q_3/Q_1)/B), & a_3 &= (1/2)(B + 2 \ln(Q_3/Q_1)/B), \\ a_2 &= (1/2)(B - 2 \ln(Q_4/Q_2)/B), & a_4 &= (1/2)(B + 2 \ln(Q_4/Q_2)/B), \end{aligned}$$

где $B^2 = N\Delta^2/P = N(1-\kappa)^2 \Phi^2(c_2) \exp(-c_0^2) / [2\pi \kappa \Phi(c_1)] \gg 1$. Минимум P_0 соответствует максимуму B . В точке максимума по c_1

$$B^2 = (1/f_1 f_2) (N/LL^{2/(1+\rho)}).$$

Здесь $f_1 = Q^{-2/(1+\rho)} e^{\rho}(e^{\rho} - 1)$, $Q = L\Phi(c_0) \approx L \exp(-c_0^2/2) / (\sqrt{2\pi} c_0)$; $f_2 = (\sqrt{\rho}(1-\rho)/(1+\rho)) \sqrt{2(2\pi)^3 \ln(L/Q)}$ при $\rho \leq 1 - 4/\ln(L/Q)$, $c_1 = (2\sqrt{\rho}/(1+\rho))c_0$, $c_2 = (\sqrt{\rho}(1-\rho)/(1+\rho))c_0$; $f_2 \approx 1/\sqrt{2 \ln(L/Q)}$ при $\rho \geq 1 - 4/(4 \ln(L/Q))$, $c_1 = \sqrt{\rho}c_0$, $c_2 = 2 \ln((2/\pi(1-\rho))^{1/2}/c_0)^{1/2}$. Максимум B определяется в основном множителем f_1 . На рис. 2, а приведена зависимость f_1 от Q при различных значениях ρ . Оценка максимума B^2 , как видно из рис. 2, а, будет достаточно точной, если для всех случаев положить $Q \approx 0,5$. В этом случае $c_0^2 \approx 2 \ln(2L)$. Замена гипергеометрических распределений нормальными для получения формул (7) допустима, если $c_0^2 \ll lm/n$, т. е. $\ln L \ll lm/n$, что и предполагалось при их выводе.

На рис. 2, б приведена зависимость функции f_2 от ρ при $Q = 0,5$ и различных значениях L (в прямоугольнике справа на графике $f_2(\rho)$ изменен масштаб).

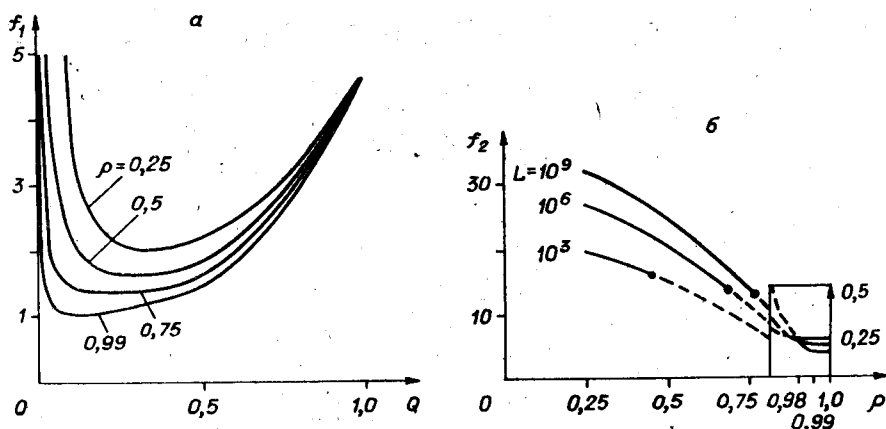


Рис. 2.

Согласно (9), для того чтобы вероятность ошибки воспроизведения достигала величины порядка P_0 , нужно положить

$$B^2 = 8 \ln(\sqrt{nl}/P_0), \text{ т. е.}$$

$$N = 8f_1 f_2 l L^{2/(1+\rho)} \ln(\sqrt{nl}/P_0). \quad (10)$$

Это число интересно сопоставить с числом N_0 двоичных элементов памяти, необходимых для записи L паттернов, содержащих по l элементов из общего их количества n , в оптимально построенном детерминированном устройстве. Число N_0 можно оценить по формуле

$$N_0 \approx lL \log_2 n. \quad (11)$$

Видно, что (10) и (11) структурно однотипны, хотя показатель степени для L в формуле (10) больше, чем в (11). Этот результат при $\rho = 0$ совпадает с полученным в [7]. Для достаточно большого диапазона параметров $N/N_0 \sim 10$, т. е. информационная емкость устройства в пересчете на один АЭ составляет величину порядка 0,1 бит/элемент. Такая избыточность стохастического фильтра по числу элементов делает его весьма устойчивым к частичным повреждениям. Если вышла из строя доля ξ от всех ассоциативных элементов, то ошибка воспроизведения составит по формуле (9) всего лишь $P_0(\xi) = (nl)^{\xi/2} P_0^{(1-\xi)}$. Например, при разрушении половины ассоциативных элементов и $n = 10^6$, $l = 10^4$, $P_0 = 10^{-6}$ примерно только у трети воспроизводимых паттернов может появиться ошибка хотя бы в одном элементе. Относительно небольшая избыточность по числу элементов по сравнению с пределами, устанавливаемыми теорией информации, простота конструкции и функционирования, а также высокая устойчивость к частичным повреждениям делают ассоциативный фильтр весьма полезным устройством для решения многих задач обработки сигналов.

П Р И Л О Ж Е Н И Е

Обозначения: a — некоторый входной и симметричный ему выходной элемент; b_α ($\alpha = 1, \dots, N$) — ассоциативные элементы; A_s ($s = 0, \dots, L$) — эталонные паттерны, из них A_0 — восстанавливаемый, A — предъявляемый к восстановлению зашумленный паттерн A_0 ; $A_s \stackrel{i}{\sim} b_\alpha$ означает, что b_α имеет i связей с A_s , $A_s \stackrel{i}{\sim} b_\alpha$ — b_α имеет не менее i связей с A_s ; b_α^+ — b_α модифицирован одним из A_s , $s \neq 0$; b_α^* — b_α активирован; $a \sim b_\alpha$ — a связан с b_α ; $x_\alpha = 1$, если выполняется $a \sim b_\alpha$ (0 в противном случае).

Для числа активированных связей выходного элемента a математическое ожидание и дисперсия имеют вид

$$M\left(\sum_{\alpha} x_{\alpha}\right) = NP, \quad D\left(\sum_{\alpha} x_{\alpha}\right) = D_0 + N(N-1)\Delta P, \quad (\text{П.1})$$

где $P = \mathcal{P}(x_{\alpha} = 1)$, $D_0 = NP(1-P)$, $\Delta P = M(x_{\alpha}x_{\alpha'}) - P^2$.

Используя статистическую независимость эталонных A_s и предъявляемого A паттернов, при вычислении P можно записать:

$$P = \mathcal{P}(a - b_{\alpha}^*) = \mathcal{P}(a - b_{\alpha}, b_{\alpha}^+ \cup (A_0 \overset{h_0}{\sim} b_{\alpha}), A \overset{h_1}{\sim} b_{\alpha}) = \mathcal{P}(a - b_{\alpha}, A_0 \overset{h_0}{\sim} b_{\alpha}, A \overset{h_1}{\sim} b_{\alpha}) + \mathcal{P}(a - b_{\alpha}, \neg(A_0 \overset{h_0}{\sim} b_{\alpha}), A \overset{h_1}{\sim} b_{\alpha}) \mathcal{P}(b_{\alpha}^+). \quad (\text{П.2})$$

$$\text{Здесь } \mathcal{P}(b_{\alpha}^+) = 1 - \mathcal{P}\left(\bigcap_{s \neq 0} \neg(A_s \overset{h_0}{\sim} b_{\alpha})\right) = 1 - \prod_{s \neq 0} \mathcal{P}\left(\neg(A_s \overset{h_0}{\sim} b_{\alpha})\right) =$$

$= 1 - (1-q)^k = \kappa$. При вычислении ΔP фиксируются связи от входных элементов к элементам b_{α} и $b_{\alpha'}$, тем самым задается число входных элементов v , связанных одновременно с b_{α} и $b_{\alpha'}$. Так как предъявляемый A и все эталонные A_s ($s \neq 0$) паттерны статистически независимы, то при фиксированных связях b_{α} и $b_{\alpha'}$ будут статистически независимы возникновения условий их модификации различными паттернами A_s ($s \neq 0$) и активации паттерном A . Следовательно,

$$\begin{aligned} \mathcal{P}(a - b_{\alpha}^*, a - b_{\alpha'}/v) &= \mathcal{P}(a - b_{\alpha} A \overset{h_1}{\sim} b_{\alpha}, A_0 \overset{h_0}{\sim} b_{\alpha}, a - b_{\alpha'}, A \overset{h_1}{\sim} b_{\alpha'}, A_0 \overset{h_0}{\sim} b_{\alpha'} / \\ & \quad /v) + 2\mathcal{P}(a - b_{\alpha}, A \overset{h_1}{\sim} b_{\alpha}, A_0 \overset{h_0}{\sim} b_{\alpha}, a - b_{\alpha'}, A \overset{h_1}{\sim} b_{\alpha'}, \neg(A_0 \overset{h_0}{\sim} b_{\alpha}), b_{\alpha}^+ /v) + \\ & + \mathcal{P}(a - b_{\alpha}, A \overset{h_1}{\sim} b_{\alpha}, \neg(A_0 \overset{h_0}{\sim} b_{\alpha}), b_{\alpha}^+, a - b_{\alpha'}, A \overset{h_1}{\sim} b_{\alpha'}, \neg(A_0 \overset{h_0}{\sim} b_{\alpha'}), b_{\alpha'}^+ /v) = \\ & = \mathcal{P}(a - b_{\alpha}, A \overset{h_1}{\sim} b_{\alpha}, A_0 \overset{h_0}{\sim} b_{\alpha}, a - b_{\alpha'}, A \overset{h_1}{\sim} b_{\alpha'}, A_0 \overset{h_0}{\sim} b_{\alpha'} /v) + \\ & + 2\kappa \mathcal{P}(a - b_{\alpha}, A \overset{h_1}{\sim} b_{\alpha}, A_0 \overset{h_0}{\sim} b_{\alpha}, a - b_{\alpha'}, A \overset{h_1}{\sim} b_{\alpha'}, \neg(A_0 \overset{h_0}{\sim} b_{\alpha'})) /v) + \\ & + \mathcal{P}(a - b_{\alpha}, A \overset{h_1}{\sim} b_{\alpha}, \neg(A_0 \overset{h_0}{\sim} b_{\alpha}), a - b_{\alpha'}, A \overset{h_1}{\sim} b_{\alpha'}, \neg(A_0 \overset{h_0}{\sim} b_{\alpha'})) /v) (\kappa^2 + \varepsilon(v)), \end{aligned}$$

где $\varepsilon(v) = \mathcal{P}(b_{\alpha}^+, b_{\alpha'}^+ /v) - \kappa^2$. Тогда $\Delta P = \sum \mathcal{P}(v) [\mathcal{P}(a - b_{\alpha}^*, a - b_{\alpha'} /v) - \mathcal{P}^2(a - b_{\alpha}^*)]$, $\mathcal{P}(v) = C_m^v C_{n-m}^{m-v} / C_n^m$ есть вероятность того, что элементы b_{α} и $b_{\alpha'}$ имеют v общих связей с входными элементами. Так как паттерны A и A_0 фиксированы, то условия модификации паттерном A_0 , активации паттерном A и связи с элементом a для элементов b_{α} и $b_{\alpha'}$ статистически независимы (при случайном выборе связей элементов b_{α} и $b_{\alpha'}$ с входными элементами). Поэтому после усреднения по v

$$\begin{aligned} \Delta P &= \mathcal{P}^2(a - b_{\alpha}, A \overset{h_1}{\sim} b_{\alpha}, A_0 \overset{h_0}{\sim} b_{\alpha}) + 2\kappa \mathcal{P}(a - b_{\alpha}, A \overset{h_1}{\sim} b_{\alpha}, A_0 \overset{h_0}{\sim} b_{\alpha}) \times \\ & \times \mathcal{P}(a - b_{\alpha}, A \overset{h_1}{\sim} b_{\alpha}, \neg(A_0 \overset{h_0}{\sim} b_{\alpha})) + \mathcal{P}^2(a - b_{\alpha}, A \overset{h_1}{\sim} b_{\alpha}, \neg(A_0 \overset{h_0}{\sim} b_{\alpha})) \kappa^2 + \\ & + \sum_v \mathcal{P}(v) F(v) \varepsilon(v) - \mathcal{P}^2(a - b_{\alpha}^*) = \sum_v \mathcal{P}(v) F(v) \varepsilon(v). \end{aligned}$$

Здесь обозначено: $F(v) = \mathcal{P}(a - b_{\alpha}, A \overset{h_1}{\sim} b_{\alpha}, \neg(A_0 \overset{h_0}{\sim} b_{\alpha}), a - b_{\alpha'}, A \overset{h_1}{\sim} b_{\alpha'}, \neg(A_0 \overset{h_0}{\sim} b_{\alpha'}) /v) -$ и использовано выражение (П.2) для $\mathcal{P}(a - b_{\alpha}^*)$.

Полагая

$$\varepsilon_1(v) = \mathcal{P}\left(\neg(A_s \overset{h_0}{\sim} b_{\alpha}), \neg(A_s \overset{h_0}{\sim} b_{\alpha'}) /v\right) - (1-q)^2,$$

где $\varepsilon_1(v)$ — монотонно возрастающая функция, имеющая максимум $q(1-q)$ при $v = m$, и учитывая, что вблизи оптимума $Lq = Q < 1$, в разложении $\varepsilon(v)$ по $L\varepsilon_1(v)$ можно ограничиться членами первого порядка. Тогда при $L \gg 1$

$$\varepsilon(v) = 1 - 2\mathcal{P}(\neg b_\alpha^+) + \mathcal{P}(\neg b_\alpha^+, \neg b_\alpha^+/v) - \kappa^2 = 1 - 2(1 - \kappa) + \\ + \left[\mathcal{P}(\neg(A_s^{\overset{k_0}{\sim}} b_\alpha), \neg(A_s^{\overset{k_0}{\sim}} b_\alpha)/v) \right]^L - \kappa^2 \approx L(1 - \kappa)^2 \varepsilon_1(v).$$

Таким образом, $\Delta P = L(1 - \kappa)^2 \sum \mathcal{P}(v) F(v) \varepsilon_1(v)$.

Рассмотрим два крайних случая: $\bar{v} = M(v) = m^2/n \gg 1$ и $\bar{v} \ll 1$. В первом из них, воспользовавшись формулой $M(f_1(x) f_2(x)) \approx M(f_1(x)) \times M(f_2(x)) + M(f_1(x)) M(f_2(x)) D(x)$, справедливой до членов порядка $D(x)$, и тем, что $M(\varepsilon_1(v)) = 0$, ΔP можно представить в виде

$$\Delta P = L(1 - \kappa)^2 D(v) M(F'(v)) M(\varepsilon_1'(v)). \quad (\text{П.3})$$

Для расчета ΔP положим $\varepsilon_1'(v) = \varepsilon_1(v+1) - \varepsilon_1(v)$ и $F'(v) = F(v+1) - F(v)$. Тогда

$$\varepsilon_1'(v) = rR_{1,1} + (1-r)R_{0,0} - r^2R_{1,1} - r(1-r)R_{1,0} - r(1-r)R_{0,1} - \\ - (1-r)^2R_{0,0} = r(1-r)(R_{1,1} - R_{1,0} + R_{0,0} - R_{0,1}),$$

где $r = l/n$ — вероятность попадания некоторой связи от АЭ в паттерн A_s ; $R_{i,j} = \mathcal{P}(\neg(A_s^{\overset{k_0-i}{\sim}} b_\alpha), \neg(A_s^{\overset{k_0-j}{\sim}} b_\alpha)/v)$; A_s^0 — полный паттерн A_s ; A_s^1 — паттерн A_s без одного входного элемента, имеющего фиксированную связь с данным АЭ, а знак точки под чертой означает, что для данного АЭ осталось выбрать $m-1$ случайных связей с оставшимися $n-1$ элементами входного слоя.

Далее, $R_{1,1} - R_{1,0} = -(1 - m/n) \mathcal{P}(\neg(A_s^{\overset{k_0-1}{\sim}} b_\alpha), A_s^{\overset{k_0-1}{\sim}} b_\alpha/v)$, $R_{0,0} - R_{0,1} = (1 - m/n) \mathcal{P}(\neg(A_s^{\overset{k_0}{\sim}} b_\alpha), A_s^{\overset{k_0-1}{\sim}} b_\alpha/v)$. Следовательно,

$$\varepsilon_1'(v) = (l/n)(1 - l/n)(1 - m/n)^2 \mathcal{P}(A_s^{\overset{k_0-1}{\sim}} b_\alpha, A_s^{\overset{k_0-1}{\sim}} b_\alpha/v) \quad (\text{П.4})$$

и $M(\varepsilon_1'(v)) = (l/n)(1 - l/n)(1 - m/n)^2 \mathcal{P}^2(A_s^{\overset{k_0-1}{\sim}} b_\alpha)$. Полагая, как и в основном тексте, $m/n \ll 1$, $l/n \ll 1$, $lm/n \gg 1$ и переходя к аппроксимации гипергеометрических распределений нормальными, получим

$$M(\varepsilon_1'(v)) \approx (1/2\pi m) \exp(-c_0^2),$$

где c_0 задается формулой (8). Аналогично в том же приближении

$$M(F'(v)) = (1/n) \Phi^2(c_1) (1 - \Phi(c_3))^2 (1 + (m/n)c_1^2).$$

Здесь c_1 определяется формулой (10), а $c_3 = (c_0 - c_1\sqrt{\rho})/\sqrt{1-\rho}$. Подставляя выражения для $M(\varepsilon_1'(v))$ и $M(F'(v))$ в (П.3) и полагая $D(v) = v$, находим

$$\Delta P = L(1 - \kappa)^2 (m/2\pi n^2) (1 + (m/n)c_1^2) \Phi^2(c_1) (1 - \Phi(c_3))^2 \exp(-c_0^2). \quad (\text{П.5})$$

Можно показать, что при $\bar{v} \ll 1$ ΔP описывается той же формулой. Подставив выражение для ΔP в (П.1) и оценив, как и в основном тексте, величину P , задаваемую (П.2) по формуле $P = \kappa m \Phi(c_1)/n$, получим

$$\eta = (D - D_0)/D_0 = (N - 1) \Delta P/P (1 - P) \approx (NL(1 - \kappa)^2/2\pi n) \times \\ \times (1 + (m/n)c_1^2) \Phi(c_1) (1 - \Phi(c_3))^2 \exp(-c_0^2).$$

В точке оптимума при $\rho < 1 - 4/\ln(2L)$

$$\eta = \gamma_1 (l/n) \ln(\sqrt{ln}/P_0) L^{(1-\rho^2-4\rho)/(1+\rho)^2},$$

где $\gamma_1 = (\sqrt{\rho}(1-\rho)/(1+\rho)) 2^{(7+2\rho+3\rho^2)/2(1+\rho)^2} (2\pi)^{(1+4\rho+\rho^2)/(1+\rho)^2} \times \\ \times (1 - \Phi(c_3)) [(1 + (m/n)c_1^2)/c_1] (2\ln(2L))^{(3+8\rho+3\rho^2)/2(1+\rho)^2}.$

Порядок величины η определяется в основном степенью L , так что $\eta \ll 1$ при $l/n \ll 1$ и $1 - \rho^2 - 4\rho < 0$ (т. е. $\rho \geq 0,25$). В случае $\rho > 1 -$

— $1/(4 \ln(2L))$ условие $\eta \ll 1$ заведомо выполняется. Таким образом, при $\rho \geq 0,25$ добавкой к дисперсии числа активных связей к выходному элементу за счет статистической зависимости условий активации и модификации различных АЭ разными паттернами можно пренебречь.

ЛИТЕРАТУРА

1. Кохонен Т. Ассоциативная память. М.: Мир, 1980.
2. Briendley G. S. Nerve Net Models of a Plausible Size, that Perform Many Simple Learning Tasks.— Proc. Roy. Soc. Lond., 1969, vol. B174, p. 173.
3. Marr D. A Theory of Cerebellar Cortex.— J. Physiol., 1969, vol. 202, p. 434.
4. Marr D. A Theory for Cerebral Neocortex.— Proc. Roy. Soc. Lond., 1970, vol. B176, p. 161.
5. Marr D. Simple Memory: A Theory for Archicortex.— Trans. Roy. Soc. Lond., 1971, vol. B262, p. 23.
6. Литвинов Е. Г., Фролов А. А. Сеть пластических нейронов с немодифицирующимися синапсами, способная к формированию условного рефлекса (мембранная модель обучения).— Биофизика, 1978, т. 23, № 6.
7. Фролов А. А., Харитонов А. П. Модель местного условного рефлекса на сети пластических нейронов.— Автометрия, 1981, № 4, с. 31.
8. Розенблатт Ф. Принципы нейродинамики. М.: Мир, 1965.

*Поступила в редакцию 28 января 1981 г.;
окончательный вариант — 16 марта 1982 г.*
