МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО АТОМНОЙ ЭНЕРГИИ

РОССИЙСКАЯ АССОЦИАЦИЯ НЕЙРОИНФОРМАТИКИ

МОСКОВСКИЙ ИНЖЕНЕРНО-ФИЗИЧЕСКИЙ ИНСТИТУТ
(ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ)

НАУЧНАЯ СЕССИЯ МИФИ–2006

НЕЙРОИНФОРМАТИКА–2006

**VIII ВСЕРОССИЙСКАЯ
НАУЧНО-ТЕХНИЧЕСКАЯ
КОНФЕРЕНЦИЯ**

ЛЕКЦИИ ПО НЕЙРОИНФОРМАТИКЕ

По материалам Школы-семинара
«Современные проблемы нейроинформатики»

Москва 2006

УДК 001(06)+004.032.26 (06) Нейронные сети
ББК 72я5+32.818я5
М82

НАУЧНАЯ СЕССИЯ МИФИ–2006. VIII ВСЕРОССИЙСКАЯ НАУЧНО-ТЕХНИЧЕСКАЯ КОНФЕРЕНЦИЯ «НЕЙРОИНФОРМАТИКА–2006»: ЛЕКЦИИ ПО НЕЙРОИНФОРМАТИКЕ. – М.: МИФИ, 2006. – 244 с.

В книге публикуются тексты лекций, прочитанных на Школе-семинаре «Современные проблемы нейроинформатики», проходившей 24–27 января 2006 года в МИФИ в рамках VIII Всероссийской конференции «Нейроинформатика–2006».

Материалы лекций связаны с рядом проблем, актуальных для современного этапа развития нейроинформатики, включая ее взаимодействие с другими научно-техническими областями.

Ответственный редактор
Ю. В. Тюменцев, кандидат технических наук

ISBN 5–7262–0635–5

© *Московский инженерно-физический институт (государственный университет), 2006*

Содержание

Л. Б. ЛИТИНСКИЙ
Институт оптико-нейронных технологий РАН, Москва
E-mail: litin@iont.ru

ПАРАМЕТРИЧЕСКИЕ НЕЙРОННЫЕ СЕТИ И ДРУГИЕ АРХИТЕКТУРЫ НА ИХ ОСНОВЕ (ОБЗОР РАБОТ)

Аннотация

Дается обзор работ по ассоциативным нейронным сетям, выполненных за последние четыре года в ИОНТ РАН. В основу изложения положено описание параметрических нейронных сетей (ПНС), введенных и разработанных авторским коллективом института. ПНС обладают рекордными на сегодняшний день распознающими характеристиками (емкостью памяти, помехоустойчивостью и быстродействием). Акцент сделан на изложении основных идей и принципов.

L. B. LITINSKII
Institute of Optical Neural Technologies, RAS, Moscow
E-mail: litin@iont.ru

PARAMETRICAL NEURAL NETWORKS AND SOME OTHER SIMILAR ARCHITECTURES

Abstract

Some results are considered for associative neural networks (NNs) based on investigations carried out at the Institute of Optical Neural Technologies (IONT) during the last four years. A discussion centers on parametrical neural networks offered and developed by researchers of the IONT. The parametrical NNs have got very high performances (memory capacity, noise stability, operating speed). Basic concepts and features for this kind of neural networks are emphasized.

Введение

Параметрическая нейронная сеть (ПНС) — это вариант ассоциативной памяти, основанный на нелинейно-оптических принципах обработки информации [1, 2]. Элементной базой ПНС являются *параметрические нейроны* — обладающие кубической нелинейностью элементы, способные к преобразованию и генерации частот в процессах параметрического четырехволнового смешения [3]. Нейроны обмениваются между собой квазимонохроматическими импульсами на q различных частотах $\{\omega_L\}_1^q$ с фазами $\phi = \{0/\pi\}$.

Уже первые оценки, сделанные в [1, 2], показали, что при правильно организованный архитектуре такой сети ее емкость памяти в q^2 раз превосходит емкость памяти модели Хопфилда [4, 5]. Резко возрастает и помехоустойчивость сети. Основываясь на этом, был выполнен ряд исследований [6]–[19], в которых удалось, с одной стороны, разобраться в механизмах такого улучшения распознающих характеристик, а с другой — разработать новые эффективные архитектуры.

Адекватным языком описания ПНС оказался векторный формализм — прием, основанный на том, что q различных состояний нейронов изображаются векторами-ортами q-мерного пространства. Правильная организация архитектуры свелась к образованию межнейронных связей по обобщенному хеббовскому правилу с использованием тензорного произведения q-мерных векторов вместо скалярного. Динамика системы векторов-нейронов непосредственно копирует стандартную асинхронную динамику, повсеместно используемую в теории нейронных сетей. Векторный формализм позволил увидеть ресурс модели и создать несколько новых архитектур, обладающих теми или иными полезными свойствами. Обзору полученных результатов посвящено это сообщение.

В следующем разделе излагается векторный формализм и описаны две основные ПНС. В заключительном разделе будет рассказано о том, как ПНС можно использовать для создания других нейросетевых архитектур.

ПНС-архитектуры

Векторный формализм

Для описания q различных состояний нейронов будем использовать векторы-орты $\mathbf{e}_l$ пространства $\mathbf{R}^q$, $q \geqslant 1$:

$$\mathbf{e}_l = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad l = 1, \ldots, q. \tag{1}$$

Состояние i-го нейрона задается вектором $\mathbf{x}_i$,

$$\begin{gathered} \mathbf{x}_i = x_i \mathbf{e}_{l_i}, \quad x_i = \pm 1, \\ \mathbf{e}_{l_i} \in \mathbf{R}^q, \quad 1 \leqslant l_i \leqslant q, \quad i = 1, \ldots, N. \end{gathered} \tag{2}$$

Весь N-мерный образ (с q-нарными координатами) задается набором N q-мерных векторов $\mathbf{x}_i$: $X = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$, а M исходных образов (*паттернов*) — это M наперед заданных подобных наборов:

$$\begin{aligned} &X^{(\mu)} = \left(\mathbf{x}_1^{(\mu)}, \mathbf{x}_2^{(\mu)}, \ldots, \mathbf{x}_N^{(\mu)}\right), \quad \mathbf{x}_i^{(\mu)} = x_i^{(\mu)} \mathbf{e}_{l_i^{(\mu)}}, \\ &x_i^{(\mu)} = \pm 1, \quad 1 \leqslant l_i^{(\mu)} \leqslant q, \quad \mu = 1, \ldots, M. \end{aligned} \tag{3}$$

(Апеллируя к первоначальной нелинейно-оптической модели поясним, что бинарная переменная x_i моделирует наличие у квазимонохроматических импульсов фазы $\phi = \{0/\pi\}$, а орты $\mathbf{e}_l$ отвечают наличию q различных частот ω_l.)

Локальное поле на i-м нейроне имеет, как всегда, вид

$$\mathbf{h}_i = \frac{1}{N} \sum_{j=1}^{N} \mathbf{T}_{ij} \mathbf{x}_j,$$

где $(q \times q)$-матрица $\mathbf{T}_{ij}$ задает межсвязь между i-м и j-м нейронами. Межсвязи выбираются в обобщенном хеббовском виде:

$$\mathbf{T}_{ij} = (1 - \delta_{ij}) \sum_{\mu=1}^{M} \mathbf{x}_i^{(\mu)} \mathbf{x}_j^{(\mu)+}, \quad i, j = 1, \ldots, N, \tag{4}$$

где $\mathbf{x}^+$ означает q-мерную вектор-строку, а через $\mathbf{x}\mathbf{y}^+$ обозначено произведение вектор-столбца $\mathbf{x}$ на вектор-строку $\mathbf{y}^+$, выполняемое по правилам матричного умножения.

Матрицы $\mathbf{T}_{ij}$ в выражении для локального поля $\mathbf{h}_i$ действуют на векторы $\mathbf{x}_j \in \mathbf{R}^q$. После суммирования по всем j локальное поле на i-м нейроне есть некая линейная комбинация ортов $\mathbf{e}_l$. Динамическое правило, копирующее асинхронную динамику, задается следующим образом: i-й нейрон в момент времени $t+1$ ориентируется в направлении, ближайшем к направлению локального поля $\mathbf{h}_i$ в момент времени t. Или, подробнее: если в момент времени t

$$\mathbf{h}_i(t) = \sum_{l=1}^{q} A_l^{(i)} \mathbf{e}_l,$$

где

$$A_l^{(i)} \sim \sum_{j(\neq i)}^{N} \sum_{\mu=1}^{M} \big(\mathbf{e}_l \mathbf{x}_i^{(\mu)}\big)\big(\mathbf{x}_j^{(\mu)} \mathbf{x}_j(t)\big)$$

и $A_k^{(i)}$ есть наибольшая по модулю амплитуда,

$$\big|A_k^{(i)}\big| = \max_{1\leqslant l\leqslant q} \big|A_l^{(i)}\big|,$$

то

$$\mathbf{x}_i(t+1) = \mathrm{sign}\big(A_k^{(i)}\big)\mathbf{e}_k.$$

Эволюция системы состоит в последовательной ориентации векторов-нейронов по этому правилу. Нетрудно показать, что при этом энергия состояния

$$E(t) \sim -\sum_{i=1}^{N} \mathbf{x}_i^+(t)\mathbf{h}_i(t)$$

монотонно убывает. В конце концов система свалится в локальный минимум по энергии, являющийся неподвижной точкой сети.

Данная модель получила название ПНС-2. При $q=1$ она превращается в обычную модель Хопфилда.

Распознающие характеристики ПНС-2

Оценки распознающих характеристик ПНС-2 даны в [6]–[10]. Их можно получить стандартными теоретико-вероятностными рассуждениями [4], разделяя локальное поле $\mathbf{h}_i$ на две составляющие — *полезный сигнал* и *внутренний шум* — и вычисляя их средние значения и дисперсии. В отличие от (описанной в [4]) стандартной ситуации, парциальные шумовые компоненты будут здесь не независимыми случайными величинами, но только некоррелированными. Это не позволяет использовать центральную предельную теорему для оценки характеристик внутреннего шума. Можно, однако, воспользоваться статистической техникой Чебышева-Чернова [21], которая и в случае некоррелированных случайных величин позволяет получить экспоненциальную часть оценки.

Все вычисления аккуратно проделаны в [9]. Оказывается, что дисперсия внутреннего шума уменьшается по сравнению с моделью Хопфилда в q^2 раз. Связано это только с тем, что громадное большинство произведений $\mathbf{x}_i^{(\mu)}\mathbf{x}_j^{(\mu)+}\mathbf{x}_j$, возникающих при вычислении парциальных шумовых компонент, обращается в ноль; дисперсия внутреннего шума при этом уменьшается. Тогда как при аналогичных вычислениях для модели Хопфилда возникают произведения $x_i^{(\mu)}x_j^{(\mu)}x_j$, каждое из которых может равняться только ± 1 и неизбежно вносит свой вклад в увеличение дисперсии.

Запишем окончательное выражение для вероятности ошибки распознавания поданного на вход сети искаженного паттерна. Пусть a — вероятность искажения координат паттерна по спиновой переменной x_i (по фазе), а b — вероятность искажения его координат по одному из q возможных направлений (на языке параметрических нейронов — вероятность искажения по частоте). Тогда, для рандомизированного набора паттернов (3) получаем оценку вероятности неправильного распознавания паттерна:

$$\Pr_{err} < \sqrt{NM} \cdot \exp\left(-\frac{N(1-2a)^2}{2M} \cdot q^2(1-b)^2\right). \tag{5}$$

При $M, N \to \infty$ эта вероятность стремится к нулю, если число паттернов M не превосходит критического значения

$$M_c = \frac{N(1-2a)^2}{2\ln N} \cdot q^2(1-b)^2. \tag{6}$$

Последнюю величину можно рассматривать как асимптотически достижимую емкость памяти ПНС-2.

При $q = 1$ эти выражения превращаются в известные результаты для модели Хопфилда (в этом случае шум по частоте отсутствует и следует положить $b = 0$). С ростом q экспоненциально спадает вероятность ошибки распознавания — существенно растет помехоустойчивость сети. Одновременно, пропорционально q^2 растет и емкость памяти M_c. Если иметь в виду обработку цветных изображений, то нейроны суть пикселы экрана. Стандарты, принятые при компьютерной обработке изображений, позволяют считать, что число различных состояний нейронов $q \sim 10^2$ (число различных цветов). При таком q емкость памяти ПНС-2 превосходит емкость памяти модели Хопфилда на 4 порядка.

В любом случае, ПНС-2 позволяет хранить число паттернов, во много раз превосходящее число нейронов N. Например, зададимся уровнем надежности распознавания $\mathrm{Pr}_{rec} = 1 - \mathrm{Pr}_{err} = 0.99$. В модели Хопфилда с такой надежностью распознавания можно хранить только $M = N/10$ паттернов, уровень искажений которых не может превышать 30%. В то же время, ПНС-2 при $q = 64$ способна восстанавливать любой из $M = 5N$ паттернов с искажениями до 90% или любой из $M = 50N$ паттернов с искажениями до 65%. Компьютерные эксперименты подтверждают эти оценки.

ПНС-3

С точки зрения реализации ПНС в виде электронно-оптического устройства представляет интерес тот ее вариант, когда у квазимонохроматических импульсов (которыми обмениваются параметрические нейроны) отсутствует фаза. На векторном языке это означает, что для описания различных состояний нейронов используются только орты $\mathbf{e}_l$, а спиновая переменная x_i отсутствует:

$$\mathbf{x}_i = \mathbf{e}_{l_i}, \quad \mathbf{e}_{l_i} \in \mathbf{R}^q, \quad 1 \leqslant l_i \leqslant q, \quad i = 1, \ldots, N.$$

Нейроны могут теперь находиться в одном из q состояний, тогда как в ПНС-2 число различных состояний равно $2q$ (за счет спиновой переменной x_i).

Эта модель получила название ПНС-3 и была рассмотрена в [12, 15]. Оказалось, что для этой модели невозможно провести описанную выше схему рассуждений: парциальные шумовые компоненты становятся здесь коррелированными, что ведет к катастрофическому нарастанию дисперсии внутреннего шума. Прием, который позволяет выйти из затруднения, аналогичен тому, что применяется в *редком кодировании* [22]– [24]: конструируя матрицы межсвязей $\mathbf{T}_{ij}$, из вектор-координат $\mathbf{x}_i^{(\mu)}$ необходимо вычесть среднее значение нейронной активности по паттерну, которое в данном случае равно $\mathbf{e}/q$, где вектор $\mathbf{e}$ есть сумма всех ортов $\mathbf{e}_l$:

$$\mathbf{e} = \sum_{l=1}^{q} \mathbf{e}_l.$$

Иными словами, если вместо выражения (4) для матриц межсвязей $\mathbf{T}_{ij}$ использовать

$$\mathbf{T}_{ij} = (1-\delta_{ij}) \sum_{\mu=1}^{M} \big(\mathbf{x}_i^{(\mu)} - \mathbf{e}/q\big)\big(\mathbf{x}_j^{(\mu)} - \mathbf{e}/q\big)^{+}, \quad i,j = 1,\ldots,N, \qquad (7)$$

парциальные шумовые компоненты становятся некоррелированными и можно осуществить всю описанную в предыдущем пункте программу. Аналогами оценок (5) и (6) здесь будут:

$$\mathrm{Pr}_{err} < \sqrt{NM} \cdot \exp\left(-\frac{N}{2M} \cdot \frac{q(q-1)}{2}(1-\bar{b})^2\right). \qquad (8)$$

и

$$M_c = \frac{N}{2\ln N} \cdot \frac{q(q-1)}{2}(1-\bar{b})^2, \quad \text{где } \bar{b} = \frac{q}{q-1}\, b. \qquad (9)$$

По сравнению с ПНС-2 емкость памяти ПНС-3 уменьшилась вдвое. Это связано с тем, что в ПНС-3 число состояний нейронов в 2 раза меньше, чем в ПНС-2. В целом, обе модели очень близки по своим характеристикам.

Замечание. ПНС-3 является очищенным от ненужных сложностей вариантом *Поттс-стекольной нейросети*, впервые исследованной 15 лет назад [24], когда активно разрабатывались модели q-нарных ассоциативных нейросетей. Неудачно выбранные обозначения и сжатый стиль изложения создали публикации [24] репутацию заумного и малопонятного текста. На самом деле Поттс-стекольная нейросеть и ПНС-3 почти в точности совпадают. В обеих моделях матрицы межсвязей $\mathbf{T}_{ij}$ определяются с помощью векторов $\mathbf{x}_l = \mathbf{e}_l - \mathbf{e}/q$, которые в Поттс-стекольной модели используются и для описания q различных состояний нейронов; в ПНС-3 для этой цели служат орты $\mathbf{e}_l$. Распознающие характеристики ПНС-3 и Поттс-стекольной модели идентичны.

В заключение раздела еще раз обратим внимание на то, что с ростом числа состояний q распознающие характеристики ПНС улучшаются. Это свойство ПНС будет активно использовано в следующем разделе.

Другие архитектуры на основе ПНС

Замечательные распознающие характеристики ПНС могут быть использованы для создания других архитектур, ориентированных на решение тех или иных специальных задач. *Декоррелирующая ПНС* позволяет хранить полиномиально много бинарных паттернов, и, что особенно важно, даже когда эти паттерны сильно коррелированы. Другой вариант — это *q-нарный идентификатор*, который позволяет на порядки ускорить процесс распознавания и уменьшить физическую память, необходимую для хранения межсвязей.

Декоррелирующая ПНС

Известно, что корреляции между бинарными паттернами катастрофически сказываются на емкости памяти модели Хопфилда. Выходом из положения является *редкое кодирование* [22]–[24]. Декоррелирующая ПНС [11, 13, 14] составляет альтернативу этому подходу.

Основная идея декоррелирующей ПНС такова: бинарные паттерны отображаются в промежуточное, внутреннее представление, использующее векторы-нейроны большой размерности q. Затем на векторно-нейронных паттернах строится ПНС-2. Свойства отображения таковы, что, во-первых, корреляции между векторно-нейронными паттернами становятся пренебрежимо малыми; и, во-вторых, размерность q векторов-нейронов растет экспоненциально по параметру отображения. Поскольку распознающие характеристики ПНС тем лучше, чем больше размерность векторов-нейронов, можно ожидать существенного увеличения емкости памяти сети.

Алгоритм отображения бинарных паттернов в векторно-нейронное представление очень прост. Пусть имеется N-мерный бинарный вектор $Y = (y_1, y_2, \ldots, y_N)$, $y_i = \pm 1$. Разобьем его мысленно на n фрагментов, каждый из которых содержит $k+1$ координат: $N = n(k+1)$. Каждому фрагменту сопоставим снабженное знаком целое число по правилу: 1) первый элемент фрагмента определяет знак целого числа; 2) остальные k элементов фрагмента — абсолютное значение целого числа l:

$$l = 1 + \sum_{i=2}^{k+1} (y_i + 1) \cdot 2^{k-i}, \quad 1 \leqslant l \leqslant 2^k.$$

(Фактически, последние k элементов фрагмента используются как двоичная запись целого числа l.)

После этого сопоставим бинарному фрагменту $(y_1, y_2, \ldots, y_{k+1})$ вектор $\mathbf{x} = \pm\mathbf{e}_l$, где $\mathbf{e}_l$ есть орт q-мерного пространства $\mathbf{R}^q$ размерности $q = 2^k$, а знак вектора определяется первым элементом фрагмента. Иными словами, отображение бинарного фрагмента в q-мерный вектор происходит следующим образом:

$$(y_1, y_2, \ldots, y_{k+1}) \to \pm l \to \mathbf{x} = \pm\mathbf{e}_l \in \mathrm{R}^q, \quad q = 2^k. \tag{10}$$

По этой схеме весь бинарный вектор $Y \in \mathbf{R}^N$ может быть взаимнооднозначно отображен в набор из n q-мерных ортов:

$$Y = (y_1, \ldots, y_{k+1}, \ldots, y_{N-k}, \ldots, y_N) \to X = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n).$$

Результат такого отображения называется *внутренним образом* бинарного вектора Y. Число k называется *параметром отображения*.

Проделав такое отображение для данного набора бинарных паттернов $\{Y^{(\mu)}\}_1^M$, получим векторно-нейронные образы $\{X^{(\mu)}\}_1^M$, на которых строится ПНС-2. Отметим, что при сколь угодно коррелированных бинарных паттернах $Y^{(\mu)}$ их векторно-нейронные образы можно рассматривать как рандомизированные. Таковы свойства отображения (10) — оно практически устраняет корреляции между паттернами. Действительно, достаточно двум бинарным фрагментам различаться хотя бы в одной координате, и фрагменты отобразятся в два совершенно разных, ортогональных друг другу орта пространства $\mathbf{R}^q$.

Теперь можно воспользоваться приведенными выше оценками для емкости памяти ПНС-2. Если, как и раньше, a есть вероятность искажения бинарных координат (уровень искажений исходных паттернов), то подстановка параметров сконструированной векторно-нейронной сети в (5) дает:

$$M_c \sim \frac{N(1-2a)^2}{2\ln N} \cdot \frac{(2(1-a))^{2k}}{k(1+k/\ln N)}.$$

Первый сомножитель в правой части есть емкость памяти модели Хопфилда (в отсутствие корреляций!), а второй сомножитель дает экспоненциальный по параметру отображения k рост емкости памяти, поскольку всегда выполняется $2(1-a) > 1$.

Необходимо, однако, проявить осторожность: параметр k нельзя брать произвольно большим. Во-первых, число n векторов-нейронов должно быть достаточно велико, чтобы можно было применять теоретико-вероятностные оценки. Это приводит к требованию $k \ll N$. Во-вторых, среди векторов-нейронов искаженного образа должен быть хотя бы один неискаженный — чтобы его вклад в локальное поле позволил восстановить правильное значение координаты паттерна. Из тех же соображений следует, что лучше бы неискаженных векторов-нейронов было не меньше двух. Поскольку число неискаженных вектор-координат образа равно $n(1-a)^{k+1}$, это дает еще одно ограничение: $n(1-a)^{k+1} \geqslant 2$. Аккуратный учет обоих этих требований позволяет выразить максимально возможное значение k_c через N и a. Когда k меняется от 0 до k_c, емкость памяти сети растет экспоненциально. Когда k становится больше k_c, сеть перестает распознавать паттерны. Однако, еще до достижения предельного значения k_c, емкость памяти может быть сделана достаточно большой. Например, при уровне искажений $a = 0.1$ максимально возможная емкость памяти $M_c \sim N^6$, при уровне искажений $a = 0.2$ она составляет $M_c \sim N^3$, а при $a = 0.3$ равняется $M_c \sim N^{1.5}$. Подчеркнем, что эти результаты справедливы при любой степени коррелированности исходных бинарных паттернов.

q-нарный идентификатор

Основная идея данной архитектуры [14, 17, 18] основана на том наблюдении, что когда сеть работает не на пределе своих возможностей (когда число паттернов хотя бы на порядок меньше критического значения M_c), восстановление правильного значения текущей координаты происходит сразу, при первой обработке этой координаты. Иными словами, для восстановления паттерна сеть пробегает по всем координатам всего один раз. Тогда можно попытаться осуществить следующую программу.

Имеем M паттернов, каждый из которых описывается набором N координат, принимающих q различных значений. Пронумеруем все паттерны и запишем их номера в q-ичной системе счисления, для чего потребуется n q-ичных позиций (связь между n, q и N будет проанализирована позже). Отнесемся к этим n q-ичным числам как к дополнительным координатам паттерна: просто автоматически присоединим эти новые n координат к описанию паттерна. Если раньше размерность паттерна была N, то теперь она станет $N + n$. Затем на M паттернах расширенной размерности $N + n$ строится сеть типа ПНС-3, с матрицами межсвязей как в выражении (7). Однако межсвязи эти «прокинуты» только от N «настоящих» нейронов к n «нумерующим», а все остальные $N^2 + n^2$ связей оборваны.

Когда на вход такой сети поступит искаженный паттерн, сеть будет восстанавливать только его нумерующие координаты. Если мы находимся в условиях, когда правильное значение координаты восстанавливается сетью сразу, без многократных пробегов по всему паттерну, номер паттерна будет восстановлен сразу. Ну а если мы знаем номер паттерна, больше нам ничего и не требуется.

Число нумерующих координат паттерна n должно быть таким, чтобы n^q равнялось критической емкости памяти ПНС-3 (9), уменьшенной, скажем, в 10 раз. Из этого требования легко получить асимптотическую оценку для n:

$$n \sim 2 + \frac{\ln N}{\ln q}, \quad N, q \gg 1.$$

Например, при числе пикселов экрана $N \sim 10^4$ и числе цветов $q \sim 10^2$ получаем, что $n \approx 4$. Иначе говоря, число искусственно введенных нумерующих координат пренебрежимо мало.

Достоинства q-нарного идентификатора очевидны. Во-первых, существенно ускоряется процесс распознавания — сеть должна восстановить правильные значения всего $n \sim$ 4–5 координат. И, во-вторых, резко уменьшаются затраты на физическую память, необходимую для хранения межсвязей $\mathbf{T}_{ij}$: таких матриц теперь требуется только Nn, а не $(N+n)^2$.

Сеть в данном случае перестает быть авто-ассоциативной и работает как q-нарный персептрон, межсвязи которого устроены по обобщенному хеббовскому правилу (7). Отсутствует и процедура распознавания паттерна (восстановления правильных значений его координат). Вместо этого происходит идентификация номера паттерна. Поэтому данная архитектура получила название q-нарный идентификатор. Компьютерное моделирование показало ее высокую работоспособность.

В заключение необходимо сказать, что освоение ассоциативных нейронных сетей для практического применения фактически еще не начиналось. По-видимому, это вызвано невысокими распознающими характеристиками модели Хопфилда. В этом плане ПНС оказались очень плодотворной и многообещающей архитектурой. Можно надеяться, что разработка ПНС и нейросетей на их основе послужит активизации интереса как к теории ассоциативных нейронных сетей, так и к их практическому применению.

Работа выполнялась в рамках проекта «Интеллектуальные компьютерные системы» (программа 2.45) при финансовой поддержке грантом РФФИ 04–07–90038.

Литература

1. *Крыжановский Б.В., Микаэлян А.Л.* О распознающей способности нейронной сети на нейронах с параметрическим преобразованием частот // *ДАН (мат.-физ.)*, т. 383(3), стр. 318-321, 2002.
2. *Fonarev A., Kryzhanovsky B.V. et al*. Parametric dynamic neural network recognition power // *Optical Memory & Neural Networks*, vol. 10(4), pp. 31- 48, 2001.
3. *Бломберген Н*. Нелинейная оптика. М.: Мир, 1966.
4. *Hertz J., Krogh A., Palmer R.* Introduction to the Theory of Neural Computation. NY: Addison-Wesley, 1991.
5. *Ежов А.А., Шумский С.А.* Нейрокомпьютинг и его приложения. М.: МИФИ, 1998.
6. *КрыжановскиЙ Б.В., Литинский Л.Б.* О векторной модели параметрической нейросети // *Искусственный интеллект*. – 2002, т. 4, стр. 710-718.
7. *Kryzhanovsky B.V., Litinskii L.B., Fonarev A.* Parametrical neural network based on the four-wave mixing process // *Nuclear Instruments and Methods in Physics Research. Section A*. – 2003, v. 502. – pp. 517–519.
8. *Крыжановский Б.В., Литинский Л.Б.* Векторные модели ассоциативной памяти // V Всероссийская н/т конференция. *Лекции по нейроинформатике*, т. 1. – М: МИФИ, 2003. – с. 72–85.
9. *Крыжановский Б.В., Литинский Л.Б.* Векторные модели ассоциативной памяти // *Автоматика и телемеханика*. – 2003, № 11. – с. 152–165.
10. *Kryzhanovsky B.V., Litinskii L.B., Mikaelian A.L.* Parametrical neural network // *Optical Memory & Neural Networks*. – 2003, v. 12(3). – pp. 138–156.
11. *Крыжановский Б.В., Микаэлян А.Л.* Ассоциативная память, способная распознавать сильно коррелированные образы // *ДАН (информатика)*. – 2003, т. 390(1). – с. 27–31.
12. *Алиева Д.И., Крыжановский Б.В.* Бесфазовая модель параметрической нейронной сети // *Международная конференция по искусственному интеллекту* (IEEE AIS'2003). М.: Физматлит, 2003. – с. 511–517,
13. *Kryzhanovsky B., Litinskii L., Fonarev A.* An effective associative memory for pattern recognition // In: *Advances in Intelligent Data Analysis V.* 5th International Symposium on Intelligent Data Analysis (IDA 2003), Berlin: Springer, 2003. – pp. 179–186.
14. *Kryzhanovsky B.V., Kryzhanovsky V.M., Fonarev A.B.* Decorrelating parametrical neural network // In: *Proceedings of International Joint Conference on Neural Network–2005* (IJCNN05), Montreal, 2005, pp. 153–157.

15. *Крыжановский Б.В., Литинский Л.Б., Микаэлян А.Л.* Векторно-нейронные модели ассоциативной памяти // *Информационные технологии и вычислительные системы*. – 2004, № 1. – с. 68–81.

16. *Крыжановский Б.В., Крыжановский В.М.* Быстрая система распознавания и принятия решений на основе векторной нейросети // *Искусственный интеллект*. – 2004, № 4. – с. 534–541.

17. *Kryzhanovsky B.V., Mikaelian A.L. Fonarev A.B.* Vector neural net identifying many strongly distorted and correlated patterns // *Int. Conf. on Information Optics and Photonics Technology*, Photonics Asia-2004, Beijing-2004. Proc. of SPIE, vol. 5642, pp. 124–133.

18. *Kryzhanovsky B.V., Kryzhanovsky V.M., Magomedov B.M., Mikaelian A.L.* Vector perceptron as fast search algorithm // *Optical Memory & Neural Network*. – 2004. – vol. 13, No. 2. – pp. 103–108.

19. *Alieva D.I., Kryzhanovsky B.V., Kryzhanovsky V.M., Fonarev A.B.* *q*-valued neural network as a system of fast indentification and pattern recognition // Pattern Recognition and Image Analysis. – 2005. – Vol. 15, No. 1. – pp. 30–33.

20. *Kryzhanovsky B.V., Mikaelian A.L., Koshelev V.N. et al.* On recognition error bound for associative Hopfield memory // *Optical Memory & Neural Networks*. – 2000. – vol. 9(4). – pp. 267–276.

21. *Perez-Vicente C.J., Amit D.J.* Optimized network for sparsely coded patterns // *Journal of Physics A*. – 1989. – vol. 22. – pp. 559–569.

22. *Palm G., Sommer F.T.* Information capacity in recurrent McCulloch-Pitts networks with sparsely coded memory states // *Network*. – 1992. – vol. 3. – pp. 1–10.

23. *Frolov A.A., Murav'ev I.P.* Informational characteristics of neural networks capable of associative learning based on Hebbiean plasticity // *Network*. – 1995. – vol. 4. – pp. 495–536.

24. *Kanter I.* Potts-glass models of neural networks // *Phys. Rev. A*. – 1988, vol. 37. – pp. 2739–2742.

Леонид Борисович ЛИТИНСКИЙ, старший научный сотрудник, кандидат физико-математических наук, заведующий сектором динамических нейронных сетей в Институте оптико-нейронных технологий РАН. Область научных интересов: ассоциативные нейронные сети, распознавание образов, многоэкстремальные задачи. Автор более 30 научных публикаций.