

УДК 621.397.3

В. П. Третьяков, С. Г. Михайлов

(Новосибирск)

**АДАПТИВНО-РЕЗОНАНСНАЯ НЕЙРОННАЯ СЕТЬ  
С РАСШИРЕННОЙ ОБЛАСТЬЮ ПРИНЯТИЯ РЕШЕНИЯ**

Предлагается новый нейросетевой алгоритм распознавания бинарных образов. В его основе лежит бинарная версия ART-архитектуры С. Гроссберга. Введенные в нее изменения механизма конкуренции категориальных нейронов основаны на нейрофизиологических данных, указывающих на наличие вставочных нейронов в цепях тормозных обратных связей. Предложенная схема организации конкуренции позволяет добавить к достоинствам ART-архитектур способность разложения комплексного входного паттерна в базисе (часто неортогональном) ранее запомненных образов.

**Введение.** Теория адаптивного резонанса, развиваемая С. Гроссбергом [1], рассматривается ее автором как универсальная теория процессов обучения и адаптации в нервной системе. На основе этой теории разработан ряд нейросетевых алгоритмов, предназначенных для классификации и распознавания образов без учителя (unsupervised learning). Эти нейросетевые архитектуры, разработанные на основе теории адаптивного резонанса (ART), являются пока единственными, которые позволяют решить дилемму «адаптивность — стабильность». Однако все ART-нейросети, описанные в литературе до настоящего времени, лишены способности разложения комплексного входного паттерна в базисе (часто неортогональном) ранее запомненных образов. Обычно предполагается, что это выполняет независимый алгоритм предварительной сегментации. Вместе с тем известно, что в случае высокого уровня помех организация процесса сегментации становится проблематичной.

В настоящей работе предлагается модифицированный вариант бинарной версии ART-архитектуры, названной ее автором ART1 [3] и предназначенной для распознавания бинарных образов. Вводимые в архитектуру изменения механизма конкуренции категориальных нейронов основаны на нейрофизиологических данных, указывающих на наличие вставочных нейронов в цепях обратных связей. Хотя этот факт общеизвестен, в большинстве нейросетевых моделей он не находит применения. Возможно, это вызвано тем, что информационная роль вставочных тормозных нейронов до сих пор сводилась к функции инвертора. В предлагаемой модели мы попытались возложить на эти нейроны значимые вычислительные функции.

**Алгоритм работы ART1.** Класс нейросетевых архитектур ART1 полностью описывается системой простых дифференциальных уравнений двухслойной сети из шунтирующих нейроподобных элементов [1—3]. Для нее доказаны теоремы, гарантирующие самоорганизацию и самостабилизацию сети в ответ на предъявление ей последовательности бинарных образов. Для целей данной работы достаточно рассмотреть функционирование системы в виде последовательного алгоритма (рис 1).

**1. Начальные состояния.**

1.1. Состояние до начала функционирования: устанавливаются параметры системы:  $L$  — параметр нормировки;  $\rho$  — параметр бдительности (vigilance).

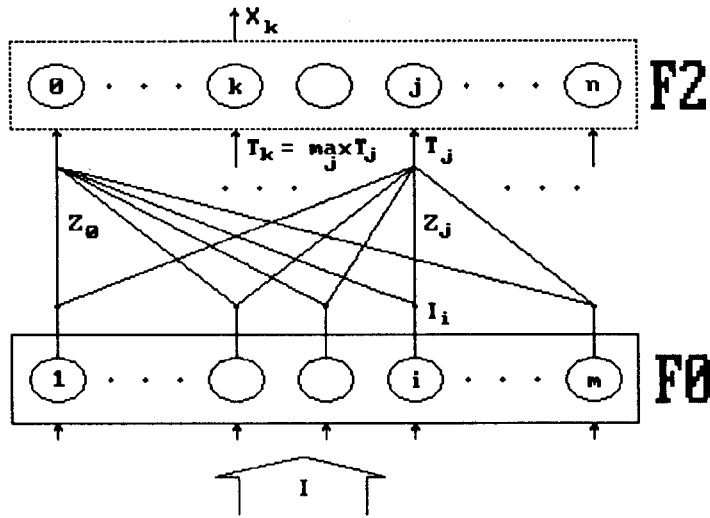


Рис. 1. Схема классической системы ART1

Все категориальные нейроны в поле  $F2$  имеют связи  $Z$  с входным полем  $F0$ , выбранные как

$$0 < z_{ij} < \frac{1}{L - 1 + m},$$

где  $z_{ij}$  — связь от  $i$ -го входного нейрона к  $j$ -му категориальному нейрону,  $m$  — число нейронов входного поля.

1.2. Состояние в процессе функционирования перед очередным предъявлением входного образа:

- неактивные нейроны имеют прежние связи;
- активные категориальные нейроны  $P_j$  имеют связи, хранящие в себе образы категорий  $C_j$  и веса для каждой категории  $w_j$ ;

$$Z_j = C_j w_j,$$

где  $C_j$  — вектор образа категории,  $w_j$  — весовой фактор для категории.

2. *Восприятие входного образа.*

2.1. На входное поле  $F0$  системы подается бинарный образ  $I$ , представляющий собой вектор с компонентами

$$I_i = \begin{cases} 1, & \text{если } i \in I, \\ 0, & \text{если } i \notin I. \end{cases}$$

2.2. Через связи  $Z_j = C_j w_j$  возбуждение передается на вход слоя категориальных нейронов, при этом активность каждого категориального нейрона равна

$$T_j = (I, z_j) = |I \cap C_j| w_j.$$

2.3. В поле  $F2$  в результате конкурентного взаимодействия между категориальными нейронами остается активным нейрон с максимальным входом:

$$T_k = \max_j (T_j).$$

3. *Контроль достаточности выбора.*

3.1. Вычисляется степень близости выбранной категории  $P_k$  и входного образа  $I$ :

$$d = \frac{|I \cap C_k|}{|I|}$$

А. Выбранная категория удовлетворяет vigilance-параметру:

$$d > \rho.$$

В этом случае происходит модификация существующей категории по правилам п. 5. Если конкуренцию выиграл еще незадействованный нейрон поля  $F2$ , то происходит формирование новой категории по правилам п. 4.

Б. Выбранная категория не удовлетворяет vigilance-параметру:

$$d \leq \rho.$$

В этом случае активность категории  $C_k$  гасится и процесс повторяется с п. 2.3.

4. Формирование новой категории.

4.1. В поле  $F2$  задействуется новый категориальный нейрон  $P_{n+1}$ .

4.2. Связи  $Z_{n+1}$  к  $P_{n+1}$  устанавливаются как

$$C_{n+1} = I,$$

$$w_{n+1} = \frac{L}{L - 1 + |I|}.$$

После создания новой категории система готова воспринимать следующий входной образ.

5. Модификация существующей категории. Модифицируются связи категориального нейрона  $P_m$  с нейронами входного слоя:

5.1. Исчезают некоторые существовавшие связи:

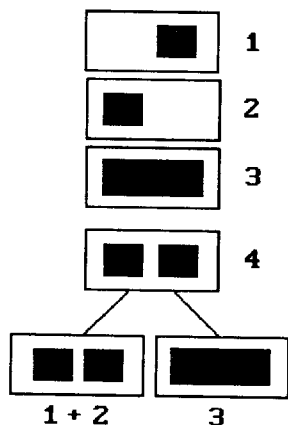
$$C_m = C_m \cap I.$$

5.2. Изменяется весовой фактор категории:

$$w_m = \frac{L}{L - 1 + |C_m|}.$$

После модификации существующей категории система готова воспринимать следующий входной образ.

Рассмотрим работу системы ART1 при распознавании образов, показанных на рис. 2. Предположим, что системе, предварительно обученной распознаванию образов 1—3, предъявлен образ 4. Какой ответ она должна дать?



Очевидно, что ART1 распознает это изображение как образ 3. Для того чтобы система ART1 смогла дать ответ 1 и 2, ее надо этому обучить, специально предъявив этот комплексный образ, что приведет к образованию новой «комплексной» категории. Очевидно, что это связано с упрощенной организацией системы конкурентного взаимодействия нейронов поля  $F2$ , которая в предложенном варианте эквивалентна функции  $\max(T_i)$ . Частое использование упрощенных вариантов конкуренции в различных нейронных моделях [4, 5] объясняется отсутствием эф-

Рис. 2. Работа системы ART1 в задаче распознавания комплексного образа

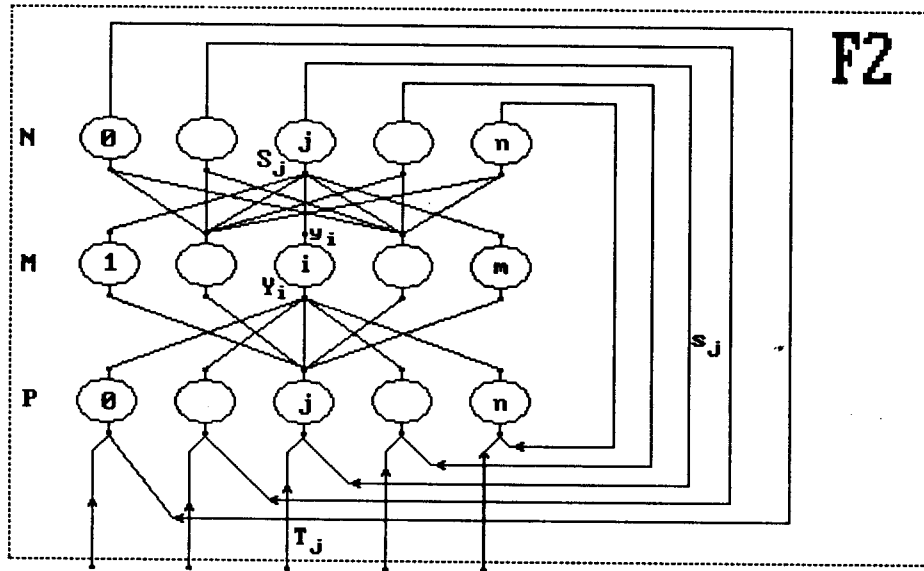


Рис. 3. Схема ART1 с расширенной областью принятия решения

фективных способов обучения обратных связей в итерационных нейросетях.

**Расширение области принятия решения в системе ART1.** Предлагаемая модель внутренних связей в поле  $F2$  изображена на рис. 3. Система конкурентного взаимодействия в поле  $F2$  реализована с помощью двух дополнительных слоев нейронов:  $M$ -слоя, топологически эквивалентного входному слою и называемого далее модельным экраном;  $N$ -слоя тормозных нейронов, топологически эквивалентного слою категориальных нейронов  $P$ .  $M$ -слой обеспечивает вычисление пересечений образов, соответствующих активным категориальным нейронам, а  $N$ -слой служит для нормировки мощности пересечения и преобразования активностей нейронов  $M$ -слоя в тормозной сигнал, поступающий на соответствующий категориальный нейрон. Изменяются также передаточные функции  $P$ -слоя категориальных нейронов: вводится порог чувствительности по принципу «все или ничего», в общем случае различный для каждого категориального нейрона.

**Имитационный алгоритм функционирования системы конкурентного взаимодействия в поле  $F2$ .**

1. *Начальные состояния.* Устанавливаются следующие параметры конкурентной системы, дополнительные к стандартной архитектуре ART1:  $t_j$  — пороги срабатывания категориальных нейронов  $P_j$ ;  $L_1$  — параметр нормировки;  $Q$  — максимальная величина допустимых наложений категорий на входном поле (т. е. отдельно взятая точка входного поля может содержаться в  $Q$ -образах одновременно).

2. *Реакция на поступившее возбуждение.*

2.1. *Реакция категориальных нейронов.*

А. Первый такт функционирования. Поступившее возбуждение  $T_j$  от входного поля на вход категориальных нейронов генерирует их ответ с учетом установленных порогов:

$$X_j = P_j(T_j) = \begin{cases} 1, & \text{если } T > t_j, \\ 0 & \text{иначе,} \end{cases}$$

где  $X_j$  — ответ  $j$ -го категориального нейрона.

Б. Последующие такты функционирования учитывают тормозящее воздействие, поступающее от  $N$ -слоя:

$$X_j = P_j(T_j - s_j).$$

2.2. Через связи  $C_j$  возбуждение поступает на входы нейронов  $M$ -слоя, давая на входе каждого нейрона величину

$$Y_i = \sum_j C_{ij}X_j.$$

2.3. Нейроны модельного экрана передают возбуждение  $Y_i$  дальше, если оно превысило величину допустимых наложений категорий:

$$y_i = M_i(Y_i) = \begin{cases} 1, & \text{если } Y_i > Q, \\ 0 & \text{иначе.} \end{cases}$$

2.4. Суммарный входной сигнал к нейронам тормозного слоя описывается как

$$S_j = \sum_i C_{ij}y_i.$$

2.5. Нейроны  $N$ -слоя нормируют входные сигналы и затем передают их как тормозящее воздействие на слой  $P$  категориальных нейронов:

$$s_j = N_j(S_j) = \frac{S_j}{L_1 + S_j}.$$

3. *Правило последовательного исключения.* Это правило действует в том случае, когда хотя бы один категориальный нейрон перестает быть активным в результате действия обратной связи. При этом происходит исключение этого (этих) нейрона из дальнейшего функционирования в рекуррентном процессе. Поле  $F2$  при этом начинает рекуррентный процесс заново, без этого нейрона.



Рис. 4. Пример распознавания комплексного образа нейроалгоритмом с расширенной областью принятия решения:  
а — предварительно запомненные эталонные образы; б — входные комплексные образы; в — результат распознавания (на место распознанной буквы подставлен соответствующий эталонный образ)

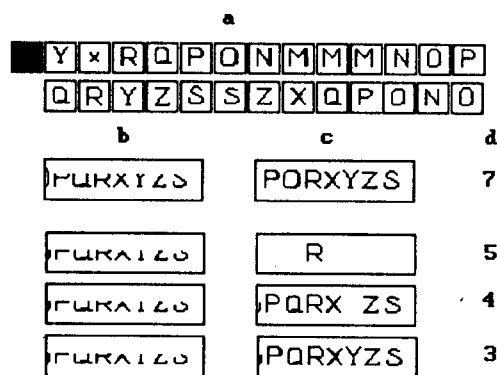


Рис. 5. Пример распознавания комплексного образа нейроалгоритмом с расширенной областью принятия решения.  
Обозначения те же, что и на рис. 4, d — различные значения порога для категориальных нейронов ( $t_j = t_j^*$ )

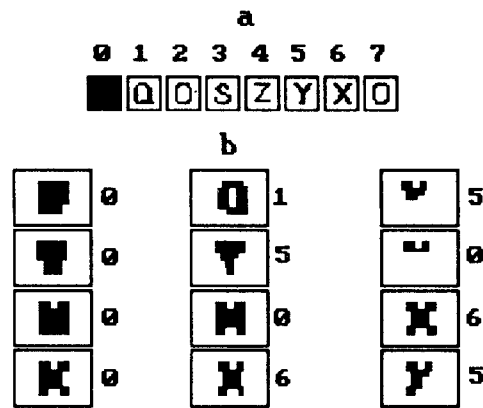


Рис. 6. Пример классификации произвольных изображений: цифры справа от предъявленных образов — номер категории, к которой был отнесен данный образ. (Обозначения те же, что и на рис. 4)

4. *Правило прекращения итерационного процесса.* Процесс конкуренции в поле  $F2$  останавливается, когда все нейроны тормозного слоя получают нулевую активность, т. е.  $S_j = 0$ .

В результате конкуренции останутся активными несколько категориальных нейронов, описывающих разложение входного образа на составляющие компоненты. При задании  $Q = 1$  эти компоненты не будут взаимно пересекаться, т. е. входной образ разложится на составляющие его взаимно ортогональные категории.

Система сдвиг-инвариантного распознавания двумерных бинарных образов. На основе предложенного варианта конкуренции можно построить систему сдвиг-инвариантного распознавания двумерных бинарных образов. Для этого достаточно вместо каждого категориального нейрона сформировать семейство нейронов, каждый из которых будет отвечать за свою область входного поля. Система конкуренции при этом остается без изменений.

Результаты моделирования процесса распознавания комплексных образов. Моделирование алгоритма производилось в программной среде, позволяющей предъявлять на вход сети как синтезированные, так и реальные изображения, накладывать на них аддитивный шум, стирать произвольные фрагменты входного образа.

Модель нейронной сети реализована в соответствии с вышеизложенным описанием алгоритма. Для отслеживания внутреннего состояния сети в процессе принятия решения активность нейронов  $M$ -слоя нормировалась на диапазон градаций яркости монитора и отображалась на экране в отдельном окне. Поскольку этот слой топологически эквивалентен входному, приход сети в устойчивое состояние сопровождается стабилизацией этого изображения, а само изображение является результатом распознавания. Обучение сети производилось однократным предъявлением эталонных образов, в качестве которых можно было интерактивно выбрать любой фрагмент изображения на экране. После запоминания сетью набора эталонов на вход сети предъявлялся комплексный образ, составленный из произвольно расположенных модифицированных эталонных образов. Размеры эталонных образов лежали в пределах  $20 \times 20$  пиксел, а тестовый комплексный образ включал в себя до 20 элементарных.

На рис. 4 и 5 приведены результаты распознавания букв русского и латинского алфавитов и продемонстрировано, что система способна распознавать сильно поврежденные буквы, степень разрушения которых достигает двух третей. Другой род тестирования состоял в предъявлении синтезированных образов, только отдаленно напоминавших эталонные. Результат такого эксперимента приведен на рис. 6.

**Закключение.** Предложенная схема организации конкуренции обладает следующими преимуществами перед альтернативными способами:

— обеспечивает разложение комплексных образов на составляющие их эталонные образы без специального обучения всем возможным сочетаниям эталонов;

— соответствует нейрофизиологическим данным, указывающим на наличие вставочных нейронов в цепях тормозных обратных связей;

— обеспечивает существенную экономию оперативной памяти, необходимой для хранения весов обратных связей, что немаловажно для имитации сети на компьютере.

Авторы выражают искреннюю признательность Ю. Н. Золотухину за поддержку проводимых исследований.

#### СПИСОК ЛИТЕРАТУРЫ

1. Grossberg S. Adaptive resonance in development, perception and cognition // SIAM-AMS Proc.—1981.—13.—P. 107.
2. Carpenter G. A., Grossberg S. ART2: Self-organization of stable category recognition codes for analog input patterns // Appl. Opt.—1987.—26, N 23.—P. 4919.
3. Carpenter G. A., Grossberg S. A massive parallel architecture for a self-organization neural pattern recognition machine // Comput. Vision, Graphics, and Image Process.—1987.—37, N 1.—P. 54.
4. Hopfield J. Neural networks and physical systems with emergent collective properties // Proc. Nat. Acad. Sci.—1986.—79.—P. 2554.
5. Kohonen T. The "Neural" phonetic typewriter // Computer.—1988.—21, N 3.—P. 11.

*Поступила в редакцию 16 февраля 1993 г.*