УДК 004.032.26(063)

ИСПОЛЬЗОВАНИЕ АППАРАТА НЕЙРОННЫХ СЕТЕЙ ДЛЯ РЕШЕНИЯ ЗАДАЧИ ВОССТАНОВЛЕНИЯ ГРАФИЧЕСКИХ ИЗОБРАЖЕНИЙ

Е. М. Будкина

Рассматривается применение функции сигмоидного типа в качестве функции активации нейронной сети Хопфилда. Проводится аналогия между функционированием рассматриваемой сети и гомогенным нечетким множеством типа k. Показана эффективность применения сети Хопфилда с функцией гиперболического тангенса для восстановления сильно искаженных изображений.

В последние годы наблюдается повышенный интерес к искусственном нейронным сетям (НС), они находят успешное применение в самых различных областях - бизнесе, медицине, физике, технике.

Как правило, НС используется для решения задач, в которых неизвестен точный вид связей между исходными данными и результатом, но накоплено достаточное число экспериментальных данных (к таким задачам относится в частности задача восстановления графических образов).

Качество классификации изображений в значительной мере зависит от степени их искажения, поэтому задача восстановления графических образов является актуальной. Поэтому исходное изображение должно пройти предобработку восстановлением.

Постановка задачи

Пусть задан набор из m эталонов изображений - n-мерных векторов $\{x^i\}$ $(1 \le i \le m)$. Требуется смоделировать нейронную сеть, которая восстанавливает искаженный эталонный образ вектора x.

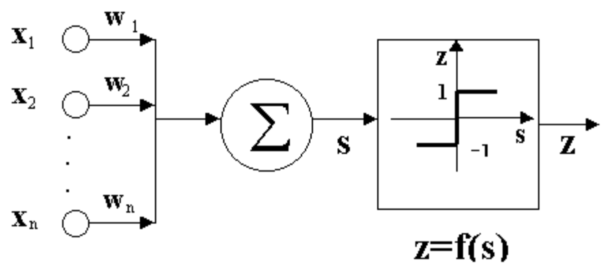
Классическая сеть Хопфилда

Для решения поставленной задачи рассматривается сеть ассоциативной памяти - сеть Хопфилда. Такая сеть позволяет соотносить входное изображение с уже знакомыми ей графическими образами и дополнять его до полного вида, а также отфильтровывать из входного изображения недостоверную информацию [1].

Входные и выходные сигналы сети биполярные: (+1 и -1). Для задачи восстановления графических образов входной сигнал равен 1, если пиксел черного цвета и равен -1 если пиксел белого цвета.

Нейрон получает на вход вектор сигналов x, вычисляет его скалярное произведение на вектор весовых коэффициентов s и преобразует его активационной функцией f (рис. 1) [1, 2, 3]. Сеть Хопфилда имеет жесткие пороговые активационные функции, например sign:

$$f(s) = \begin{cases} 1, & s \geq 0, \\ -1, & s < 0. \end{cases} \tag{1}$$



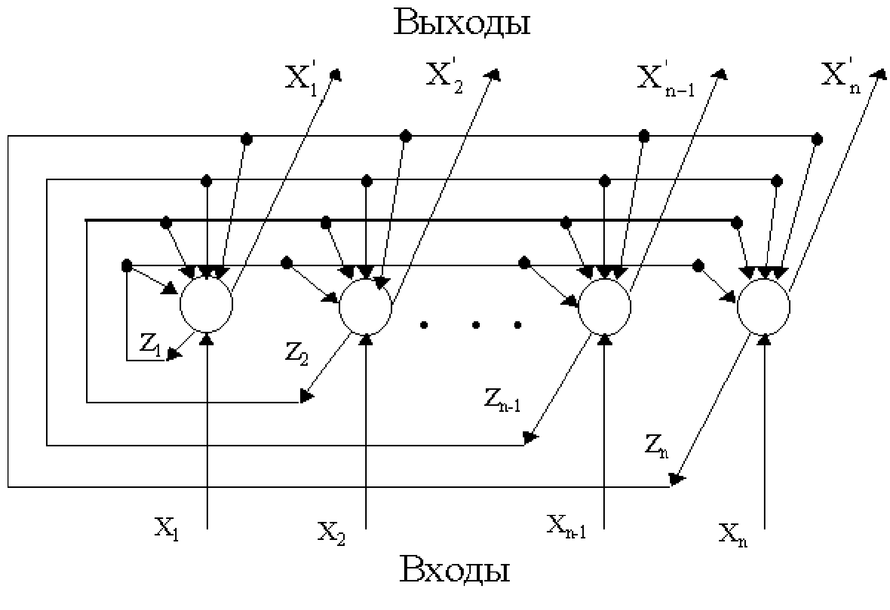
Структурная схема функционирования одного нейрона.

Рис. 1.

Сеть функционирует циклически (рис. 2), осуществляя преобразование:

$$x' = f\left(\sum_{i=1}^{n}\left(x, x^i\right)x^i\right), \tag{2}$$

где скалярное произведение (x,x^{i}) - вес i - го эталона.



Структурная схема функционирования сети Хопфилда.

Рис. 2.

Правило формирования весовых коэффициентов следующее:

$$w_{ij}=\begin{cases}\sum_{s=1}^{n} x_i^s x_j^s & ,i\neq j,\\ 0, & i=j\end{cases} \quad .\tag{3}$$

То есть коэффициенты вычисляются, как сумма произведений векторов входа по соответствующим индексам. Матрица весовых коэффициентов W симметрична относительно нулевой главной диагонали.

Функционирование сети происходит рекуррентно:

$$z_j(t+1)=f\left(\sum_{i=1}^{n} w_{ij} z_i(t)\right), \quad 1\leq j\leq n, \tag{4}$$

при начальном условии: $z_j(0)=x_j, \quad 1\leq j\leq n$.

Процесс распознавания завершает свою работу после того, как выходные сигналы нейронов перестают изменяться, то есть значение выхода на некотором шаге T равно значению выхода на предыдущем шаге (T-1):

$$z_j(T)=z_j(T-1), \ \ 1\leq j\leq n. \tag{5}$$

Таким образом, полученный выходной сигнал сети имеет вид:

$$x_j{}' = z_j(T), \ \ 1\leq j\leq n. \tag{6}$$

Сеть Хопфилда имеет существенные недостатки:

1) небольшая емкость, то есть число образов, которое сеть может запомнить не превышает 0,15*n от размерности n входного вектора;
2) запоминаемые образы не должны быть сильно скоррелированы.

Для устранения этих недостатков в настоящей статье для решения поставленной задачи рассматриваются различные преобразования классической сети Хопфилда.

Ортогональная сеть Хопфилда

Степень скоррелированности образов определяется отношением:

$$K=\left|\left(x^i,x^j\right)\right|<\frac{n}{m},\forall i\neq j \tag{7}$$

Сеть Хопфилда с ортогональным преобразованиемпозволяет восстанавливать сильно скоррелированные образы и выполнять следующие преобразования [1]:

$$x'=f\left(\sum_{i=1}^{n}\left(x,v^i\right)x^i\right), \tag{8}$$

где v^i - векторы дуального множества, которые равны:

$$v^i=\sum_{j=1}^{m}\gamma_{ij}^{-1}x^j \ , \tag{9}$$

где γ_{ij}^{-1} - элемент обратной матрицы Грама $\Gamma^{-1}\left(\left\{x^i\right\}\right)$: $\gamma_{ij}=(x^i,x^j)$, ij- ый элемент матрицы Грама равен скалярному произведению i-го эталона на j-ый.

Для работы ортогональной сети Хопфилда необходимо хранить эталоны и обратную матрицу Грама $\Gamma^{-1}\left(\left\{x^i\right\}\right)$.

Ортогональное преобразование снижает зависимость надежности работы сети от степени скоррелированности образов, а также увеличивает емкость сети.

Автокорреляторная сеть Хопфилда

Применение автокорреляторов позволяет обрабатывать визуальные образы, отличающиеся только положением в рамке. Для этого исходное изображение преобразуется в вектор инвариантов, не изменяющийся при сдвиге. Простейший набор инвариантов дают автокорреляторы, которые являются скалярным произведением образа на сдвинутый образ и рассматриваются как функции вектора сдвига.

Автокорреляторная сеть имеет вид:

$$x' = f\left(\sum_{i=1}^{n}\left(Ac(x), Ac(x^i)\right), x^i\right) \tag{10}$$

Вычисление сдвигового автокоррелятора для черно-белых изображений проводится следующим образом. Пусть дан двумерный образ d размером $p\times q=n$. Обозначим точки образа как d_{ij}. Элементами автокоррелятора Ac(d) будут величины: $a_{kl}=\sum_{i=1}^{p}\sum_{j=1}^{q} d_{ij}d_{i+k,j+l}$.

Автокорреляторное преобразование сети позволяет обрабатывать различные образы, отличающиеся только положением в рамке [1].

Нечеткая сеть Хопфилда с сигмоидной функцией активации

Использование пороговой функции в качестве активационной функции сети Хопфилда значительно ограничивает возможности сети. Качество работы такой сети зависит от степени искажения исходного изображения. Для повышения надежности восстановления в настоящей работе предлагается метод с использованием в качестве функции активации функции гиперболического тангенса:

$$f(s)=th(\alpha s)=\frac{1-e^{\alpha s}}{1+e^{\alpha s}} \tag{11}$$

Функция гиперболического тангенса (рис. 3) с большим значением характеристики α приближается по внешнему виду к пороговой функции (рис. 1).

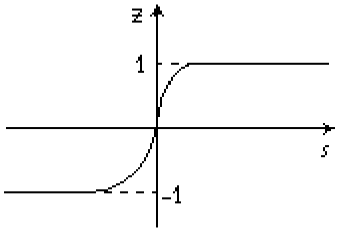


График функции гиперболического тангенса.

Рис. 3.

Применение такой активационной функции позволяет увеличить способность сети запоминать и точно воспроизводить сильно искаженные эталоны.

Таким образом, преобразованная сеть Хопфилда будет иметь вид:

$$x' = th\left(\sum_{i=1}^{n}\sum_{j=1}^{m}\gamma_{ij}^{-1}\left(Ac(x), Ac(x^{i})\right)x^{i}\right). \qquad (12)$$

Утверждение. НС Хопфилда с функцией гиперболического тангенса реализует гомогенное n–мерное нечеткое множество (НМ) типа k, где n – число элементов сети Хопфилда, k – число тактов функционирования сети.

Доказательство.

1. Введем понятие НМ А для сети Хопфилда. Каждый нейрон НС соответствует одному пикселу изображения, который может иметь только одно из двух значений цветов: черный или белый. Сеть с функцией активации (1) однозначно определяет цвет каждого пиксела. Сеть с функцией активации (11) каждому пикселу ставит в соответствие число из интервала [-1, 1], которое можно рассматривать как степень близости к черному или белому цвету. Таким образом для каждого элемента $x \in X$ один нейрон сети сопоставляет нечеткое множество $A = \{(x, \mu_A(x))\}$, где μ_A – число, характеризующее степень принадлежности элемента x множеству А, осуществляя отображение $\mu_A: X \to [-1, 1]$. По определению [4] введенное множество является НМ типа 1.

2. Так как НС состоит из n нейронов, то каждый элемент $x \in X$ характеризуется вектором значений принадлежности $(\mu_1(x), \mu_2(x), \ldots, \mu_n(x))$, где $\mu_i(x)$ –значение функции принадлежности i-го нейрона, (i=1,2,…,n). Таким образом, строится функция, отображение которой представляется в виде:

$$\mu: X \to S_1 \times S_1 \times \ldots \times S_n, \qquad (13)$$

где S_i (i=1,2,…,n) – ограниченные линейно упорядоченные множества на интервале [-1, 1].

По признаку однородности области значений функций принадлежности, построенное НМ является гомогенным n-мерным и имеет вид:

$$\mu: X \to [-1,1]^n. \tag{14}$$

3. Процесс функционирования сети (4) в обозначениях НМ имеет вид:

$$\mu^{t+1}(X) = \mu(\mu^t(X)), \tag{15}$$

при начальном условии $\mu^0(X)=X$.

Таким образом на каждом шаге t ($1 \le t \le k$) функционирования НС реализуется НМ типа t [4]:

$$\begin{aligned} &\mu^1(X) = \mu(\mu^0(X)) = \{\mu \mid \mu: X \to [-1,1]^n\}, \\ &\mu^2(X) = \mu(\mu^1(X)) = \{\mu \mid \mu: X \times [-1,1]^n \to [-1,1]^n\}, \\ &\ldots, \\ &\mu^k(X) = \mu(\mu^{k-1}(X)) = \{\mu \mid \mu: X \times [-1,1]^n \times \ldots \times [-1,1]^n \to [-1,1]^n\}. \end{aligned} \tag{16}$$

Утверждение доказано.

Численный эксперимент

Апробация метода с использованием функции гиперболического тангенса, осуществлялась с помощью численного эксперимента, который заключался в восстановлении отпечатков пальцев. В качестве исходных данных взяты растровые черно-белые изображения отпечатков пальцев размером 40x50 пикселей. Таким образом входной вектор $\bar{x} = \{x_1, x_2, \ldots, x_m\}$ и выходной вектор $\bar{x}' = \{x'_1, x'_2, \ldots, x'_m\}$ имеют размерность m=2000.

Исходными данными для формирования весовых коэффициентов сети являются векторы-образцы, содержащиеся в базе эталонов. При восстановлении изображения, сеть осуществляет преобразование входного вектора x так, чтобы выходной вектор x’ максимально соответствовал эталону, который рассматривается как правильный ответ.

Сравнительная характеристика работы классической и преобразованной сети Хопфилда приведена табл. 1.

Сравнительный анализ полученных результатов позволяет сделать следующие выводы: качество восстановления изображения НС с пороговой функцией в значительной мере зависит от степени искажения изображения, такая сеть практически не восстанавливает значительно смещенные от мнимого центра образы, имеющие частичную потерю фрагментов. НС с функцией гиперболического тангенса не восстанавливает только те изображения, потеря фрагментов которых составляет более 85% при несмещенном изображении и более 75% при смещенном от мнимого центра изображения.

Таблица 1

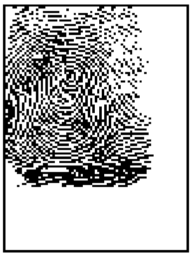
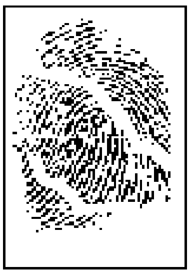
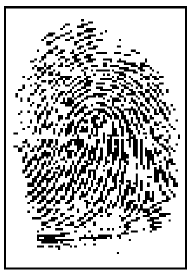
Сравнение преобразованных сетей Хопфилда

Модель НС	Качество распознавания (в %)		
	потеря информации 20%	потеря информации от 20% до 50%	потеря информации более 50%
Преобразованная сеть Хопфилда с пороговой функцией активации sign	100	73	51
Преобразованная сеть Хопфилда с сигмоидной функцией активации th	100	100	97

В табл. 2 приведены результаты восстановления искаженных отпечатков пальцев преобразованной сетью Хопфилда с функцией гиперболического тангенса. Приведены примеры сдвинутых отпечатков, отпечатков с частичной потерей фрагментов.

Таблица 2.

Примеры восстановления отпечатков пальцев.

Искаженные отпечатки	Восстановленные отпечатки
  	
  	

Дальнейшие исследования предполагают рассмотреть восстановление полутоновых и цветных изображений, обработку аэрокосмических изображений (сжатие, восстановление, сегментация, констатирование и обработка текстур), поиск и восстановление на изображении объектов заданной формы, обработка серии или потоков изображений.

Выводы

В настоящей статье для задачи восстановления графических изображений предлагается использовать нейронную сеть Хопфилда с функцией активации – гиперболический тангенс. Анализируется связь предложенной НС с теорией нечетких множеств. Результаты, полученные при реализации численного эксперимента, показали эффективность применения НС Хопфилда с функцией гиперболического тангенса для восстановления сильно искаженных изображений.

СПИСОК ЛИТЕРАТУРЫ

1. Горбань А.Н. Обучение нейронных сетей. – М.: СП ПараГраф, 1990.- 160 с.
2. Горбань А.Н., Россиев Д.А. Нейронные сети на персональном компьютере. – Новосибирск: Наука, 1996.- 276 с.

3. Уоссермен Ф. Нейрокомпьютерная техника. Теория и практика./ Пер. с англ. под ред. Галушкина А.И. - М.: Мир, 1992. – 238 с.
4. Нечеткие множества в моделях управления и искусственного интеллекта./ Под ред. Поспелова Д.А. – М.: Наука, 1986. – 395 с.

СВЕДЕНИЯ ОБ АВТОРЕ

Будкина Елена Михайловна, старший преподаватель кафедры математического и программного обеспечения информационных систем комплексов и сетей филиала «Восход» Московского государственного авиационного института (технического университета), аспирантка кафедры дифференциальных уравнений Московского государственного авиационного института (технического университета). E-mail: lenok@baikonur.ru