

УДК 681.332 : 778.38

В. И. Козик, Е. С. Нсжевенко, О. И. Потатуркин

(Новосибирск)

**ОПТОЭЛЕКТРОННЫЙ НЕЙРОННЫЙ ПРОЦЕССОР
ДЛЯ РАСПОЗНАВАНИЯ МАЛОРАЗМЕРНЫХ ИЗОБРАЖЕНИЙ**

Рассматривается концепция построения нейросетевых систем распознавания малоразмерных изображений в относительно большом поле наблюдения. Она основана на использовании пространственно-инвариантных операторов, что позволяет резко снизить число синаптических связей, упростить и ускорить процедуру обучения нейронной сети. Предложенная нейросетевая система, построенная на основе применения оптико-электронного конвольвера, обеспечивает эквивалентную производительность $\sim 10^{10}$ операций сложения/умножения и может быть выполнена в виде нейрочипа размером $\sim 2 \times 2$ см с потребляемой мощностью $\sim 0,05$ Вт.

Эффективность нейрокомпьютерных методов распознавания изображений отмечается многими исследователями [1]. Основным преимуществом этих методов является возможность создания на их основе универсальной системы обучения, обеспечивающей эффективное разбиение изображений на классы. Использование нелинейных операторов в тракте многослойной нейрокомпьютерной обработки позволяет строить в пространстве признаков весьма сложные гиперплоскости и соответственно реализовать в нейрокомпьютере достаточно сложные алгоритмы распознавания. Все это свидетельствует о перспективности таких методов. В то же время в литературе встречается не слишком много конкретных вариантов реализации подобного подхода, так как при попытке создания системы распознавания реальных изображений возникают технические трудности в основном из-за большого объема информации, обрабатываемой в процессе распознавания. В настоящей работе рассмотрены: одна из весьма распространенных задач распознавания малоразмерных изображений в большом поле наблюдения, нейрокомпьютерная схема ее решения и вариант технической реализации, основанный на оптико-электронных технологиях.

Пусть требуется распознать малоразмерный объект в поле, существенно превышающем его размеры, причем координаты объекта не известны. Стандартный способ использования нейронной сети для решения этой задачи – спроецировать сцену на сетчатку, в которой число рецепторов соответствует количеству пикселей в сцене, использовать нейронную сеть с количеством входов, равным числу рецепторов, и обучить ее, предъявляя объект при всевозможных преобразованиях, включая сдвиг. Типичная нейронная сеть для случая, когда необходимо распознать в сцене $N \times N = R$ пикселей объект

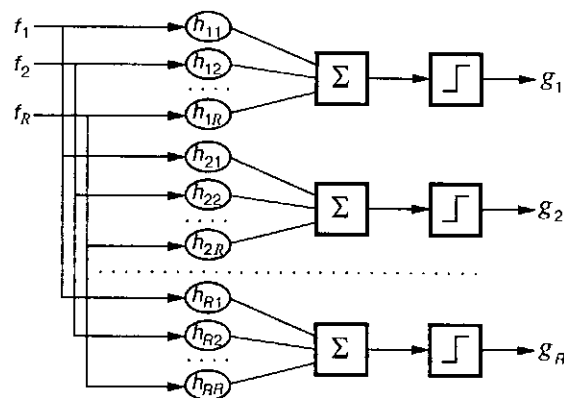


Рис. 1. Принцип построения нейронной сети

меньшей размерности, представлена на рис. 1. Здесь f_1, \dots, f_R — входы рецепторов, h_{11}, \dots, h_{RR} — синаптические коэффициенты, g_1, \dots, g_R — отклик нейронной сети после нелинейного преобразования. При наличии полного набора $N^2 \times N^2$ связей требуется время обучения, необходимое для их настройки, а реализация алгоритма распознавания выполняется за N^4 операций умножения и N^4 операций сложения. В то же время для распознавания объекта размерностью $M \times M$ пикселей достаточно M^4 связей, однако при этом необходимо нейронную сетчатку смещать в поле наблюдения с шагом в один элемент по обеим координатам.

Эту операцию можно реализовать, используя операторы, инвариантные к сдвигу, в частности оператор свертки (корреляции):

$$g(k, l) = \sum_{i=1}^M \sum_{j=1}^M f(i, j) h(k - i, l - j), \quad (1)$$

где $k = 1, \dots, N - M, l = 1, \dots, N - M$ — параметры сдвига, при этом одновременно выполняются и другие преобразования, необходимые для реализации нейросетевого алгоритма.

Одним из существенных преимуществ данного оператора является возможность применения для его реализации принципов, схемотехники и элементной базы конвольверов (корреляторов), в частности оптико-электронных. Эти устройства интенсивно разрабатывались в 1970–80-е гг. для решения задач распознавания изображений, однако широкого применения так и не нашли, поскольку в качестве значений решающей функции распознавания в них использовались коэффициенты корреляции, эффективные в весьма ограниченном числе случаев, прежде всего при распознавании детерминированных объектов, искаженных аддитивным шумом. В нашем случае конвольверы используются только как системы, обеспечивающие пространственную инвариантность обработки, а все преимущества нейрокомпьютерного распознавания при этом сохраняются. Если сопоставить нейросетевую и корреляционную системы распознавания, то легко видеть, что одному пикселу эталона в конвольвере соответствует один синаптический коэффициент в нейросетевой системе.

Выясним, при каких соотношениях между параметрами объекта и поля метод использования оператора свертки имеет преимущества. Выше были определены параметры вычислительной системы, требуемой для реализации традиционной нейросетевой системы распознавания. Определим эти же параметры для предлагаемого метода. В этом случае нейросетевая система требует M^4 синаптических связей, а распознавание объекта во всех $(N - M)^2$ положениях производится за счет использования механизма сдвига. Тогда размерность нейронной сети, подлежащей обучению, равна M^4 , а число сложений и умножений, необходимых для распознавания, составит $M^4(N - M)^2$. Очевидно, что по фактору скорости обучения предлагаемый метод выгоден всегда, поскольку $N > M$, а по числу операций при распознавании — когда $N > M^2$.

Рассмотрим организацию вычислений в нейронной сети с использованием конвольверов. Пусть, как и выше, требуется в сцене размерностью $N \times N$ пикселей распознать объект размерностью не более $M \times M$ пикселей. Схема обработки изображения в нейронной сети с применением пространственно-инвариантных операторов представлена на рис. 2. Обучение нейронной сети проводится в этом случае путем настройки M^4 связей, причем обучающая выборка должна включать распознаваемый объект (объекты) со всевозможными искажениями и преобразованиями, которые могут встретиться при распознавании. Алгоритмы обучения (настройки связей) общеизвестны [2] и в данной работе рассматриваться не будут. В процессе обучения формируется набор из $P \leq M^2$ эталонов (синаптическая матрица) размерностью M^2 каждый. В процессе распознавания в соответствии с (1) конвольвер вычисляет $(N - M)^2$ векторов $g(k, l)$, каждый из которых является при фиксированных k, l P -мерным вектором (с элементами $g^p(k, l)$) первичной реакции нейронной сети на фрагмент поля наблюдения размером $M \times M$ пикселей с координатами k, l . Далее в соответствии со стандартной процеду-

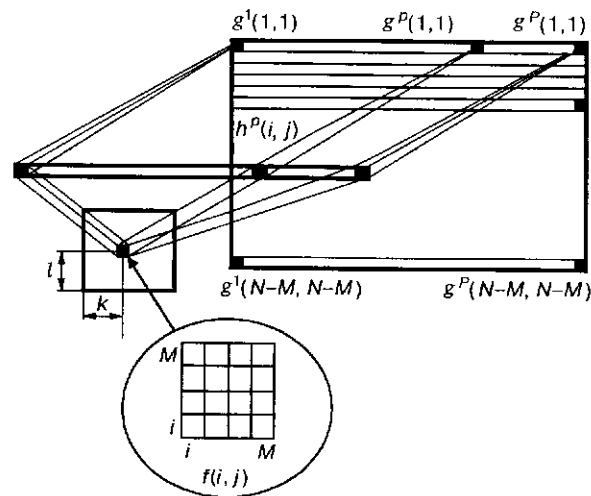


Рис. 2. Схема обработки изображения в нейронной сети с применением пространственно-инвариантных операторов

рой нейросетевой обработки проводится пороговая обработка массива $g^p(k, l)$, в результате чего он превращается в бинарный:

$$r(p, k, l) = 0,5\{\text{sign}[g^p(k, l) - h_0] + 1\}. \quad (2)$$

В случае однослойной нейронной сети при ее обучении с обучающей выборкой $\{h^p\}$, определяющей классы распознаваемых изображений, сопоставляется набор эталонных векторов h^p . В процессе распознавания $r(p, k, l)$ анализируется, и, если вектор r с координатами (k_0, l_0) принадлежит множеству $\{h^p\}$, фиксируется наличие объекта p -го класса. Если же сеть многослойная, для каждого вектора первого слоя вычисляется вектор реакции второго слоя

$$g^q(k, l) = \sum_{p=1}^P r(p, k, l)h(p, q)$$

и далее по аналогии с (2) проводится пороговая обработка

$$r(q, k, l) = 0,5\{\text{sign}[g^q(k, l) - h_1] + 1\}.$$

При большом количестве слоев аналогично вычисляется реакция третьего слоя и т. д.

Однако даже в случае использования принципа свертки для реализации нейросетевых алгоритмов распознавания изображений при технической реализации возникают большие трудности. Действительно, при поле наблюдения 512×512 элементов и распознаваемом изображении 32×32 элемента для обработки в нейросетевом процессоре при наличии всех связей требуется выполнить $5,12 \cdot 10^{11}$ операций (сложений и умножений), что влечет значительные временные затраты для современных вычислительных систем обработки изображений. Покажем, что предлагаемый метод может быть эффективно реализован в оптико-электронном варианте. При этом главный недостаток оптико-электронных конвольверов, связанный с аналоговым способом вычислений и соответственно сравнительно низкой точностью, особой роли не играет, что характерно для нейросетевой обработки.

Рассмотрим реализацию предложенного метода с использованием принципа построения оптико-электронного сигнального процессора [3]. Главной особенностью этого процессора является то, что, имея одну и ту же базовую структуру, он может быть настроен на решение самых различных задач обработки сигналов. Структура такого процессора, настроенного на решение описанной выше задачи, представлена на рис. 3, где 1 – панель источников излучения; 2 – первый блок переадресации, формирующий из точечных источников набор световых полос; 3 – пространственно-временной модулятор света (ПВМС); 4 – второй блок переадресации, проецирующий изображение с ПВМС на многоэлементное фотоприемное устройство (ФПУ) 5. В качестве ФПУ используется специализированная ПЗС-матрица размером $N \times N$, состоящая из $P = N/M$ полос по $M \times N$ элементов каждая (где $N \times N$ – размер поля входных изображений; $M \times M$ – размер эталонов) и имеющая соответственно P параллельных независимых выходов.

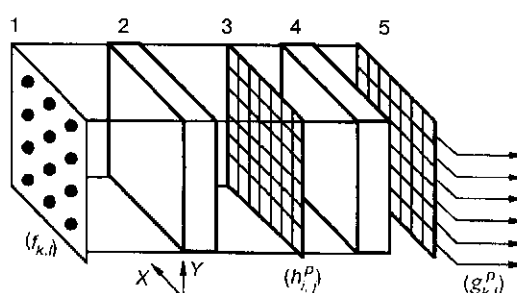


Рис. 3. Структура процессора

При использовании данной системы может проводиться распознавание изображений, поступающих как в виде последовательности строк, так и сформированных в двумерном виде. В первом случае строки f_l распознаваемых изображений (РИ) $f_{k,l}$ (k, l – дискретные пространственные координаты элементов РИ) последовательно вводятся посредством линейки источников излучения, а эталоны $h_{i,j}^p$ (i, j – дискретные пространственные координаты внутри эталонов; p – условный номер эталона) в виде набора двумерных массивов данных хранятся на ПВМС. По координате Y модулятор условно делится на P полос по $M \times N$ элементов. Каждая полоса содержит (N/M) одинаковых эталонов $M \times M$ элементов, равномерно распределенных по координате X , а по координате Y располагаются различные эталоны. Элемент линейки источников излучения освещает один столбец ПВМС. На ПВМС производится перемножение элемента входной строки на столбец элементов каждого эталона, результатом является дискретная функция

$$g_{k,l,i,j}^p = f_{k,l} h_{i,j}^p \quad (3)$$

для одного l всех i, j, k, p , при этом k и i связаны соотношением $k = i + nM$, где n – порядковый номер эталона в полосе.

Соответствующее световое распределение считывается ПЗС-матрицей. Далее на ПЗС-матрице производится сдвиг зарядов, на линейку излучателей подается следующая строка и вновь считывается результат перемножения (3). После $2M$ тактов обработки выходной сигнал на каждом канале ПЗС-матрицы соответствует

$$g_{k,l,i}^p = \sum_{j=1}^{2M} f_{k,l+j} h_{i,j}^p, \quad (4)$$

что представляет собой одномерную свертку $2M$ строк РИ с каждой из полос эталонов для $k \in [1, N], l \in [l_0, l_0 + 2M]$, где l_0 – номер начальной строки РИ. В данном случае размер изображения по координате Y взят в 2 раза больше размера объекта, чтобы избежать «сшивки» для обнаружения объектов, попавших на границу фрагмента $[l_0, l_0 + M]$.

Полученные данные запоминаются, затем изображения смещаются на один элемент вдоль линейки источников излучения, и операция повторяется.

Двумерные (для каждого p) функции (4) суммируются, в результате чего после выполнения $2M$ последовательных сдвигов строк РИ получаем

$$g_{k,l}^p = \sum_{i=1}^{2M} \sum_{j=1}^{2M} f_{k+i,l+j} h_{i,j}^p, \quad (5)$$

что представляет собой полный набор P -размерных векторов g^p реакции нейронной сети на фрагменты поля наблюдения размером $M \times M$ пикселей с координатами k, l для всех $k \in [1, N], l \in [l_0 + M/2, l_0 + 3M/2]$.

В случае когда РИ заданы в виде двумерного массива данных, они могут вводиться при помощи ПВМС, а эталоны (имеющие такую же структуру, как и в предыдущем варианте) – через излучатели. Процедура свертки проводится аналогично представленной выше. При этом, однако, появляются дополнительные сложности, связанные с необходимостью обеспечения «сшивки» выходного поля, так как объекты, находящиеся на границе отдельных каналов ПЗС-матрицы, будут давать отклик соответственно в окрестности начала предыдущей полосы и конца следующей. В этом случае свертка по координате Y для каждого эталона осуществляется за M тактов сдвига заряда, а после проведения полной операции свертки по обеим координатам для одного эталона результирующие сигналы по всем каналам ПЗС-матрицы будут представлять фрагмент функции (5) для одного p и всех $k, l \in [1, N]$.

Общее время, затраченное на выполнение свертки изображения размером $N \times N$ с набором эталонов, складывается из $2N \times 2M$ тактов накопления/сдвига на ПЗС-матрице. При тактовых частотах сдвига строк $\sim 4 \cdot 10^4$ Гц [4], $N = 512$, $M = 32$ общее время обработки по 16 эталонам составит менее 1 с, что для аналогичной нейронной сети соответствует эффективной производительности $\sim 10^{10}$ операций сложения/умножения.

Представленная процедура может быть использована для решения достаточно широкого класса задач обработки изображений и сигналов, однако, как показано выше, она требует больших вычислительных мощностей для реализации множества однородных операций. Поэтому имеет смысл создавать нейросетевую процессор на основе интегральных технологий, обеспечивающих ускорение процедуры обработки при взаимодействии со стандартными вычислительными средствами. Его вариант представлен на рис. 4. Для ввода строк РИ в нем используется линейка лазерных или светодиодов с протяженными по одной координате элементами. В качестве устройств ввода также могут быть использованы: матрица лазерных или светодиодов [5], элементы которой включены параллельно по столбцам; миниатюрный дисплей [6] или однородный осветитель [7] с линейным модулятором света. В этом случае оба блока переадресации представляют собой фокусирующие элементы и для их реализации можно применить дифракционные решетки. Модулятор – матрица из жидкокристаллических элементов, при этом минимальный размер его светоклапанного элемента может составлять $\sim 10 \times 10$ мкм [8]. Тогда при скважности размещения элементов 2 общий размер микросхемы составит $\sim 2 \times 2$ см.

Матрица лазерных диодов может обеспечивать мощность до 10 Вт и содержать порядка 10^6 элементов на кристалле [5]. При пороговой чувствительности ПЗС-матрицы $\sim 10^{-13}$ Дж/эл., тактовой частоте лазерных диодов $\sim 5 \times 10^4$ Гц и потерях света в системе 90 % при всех включенных элементах

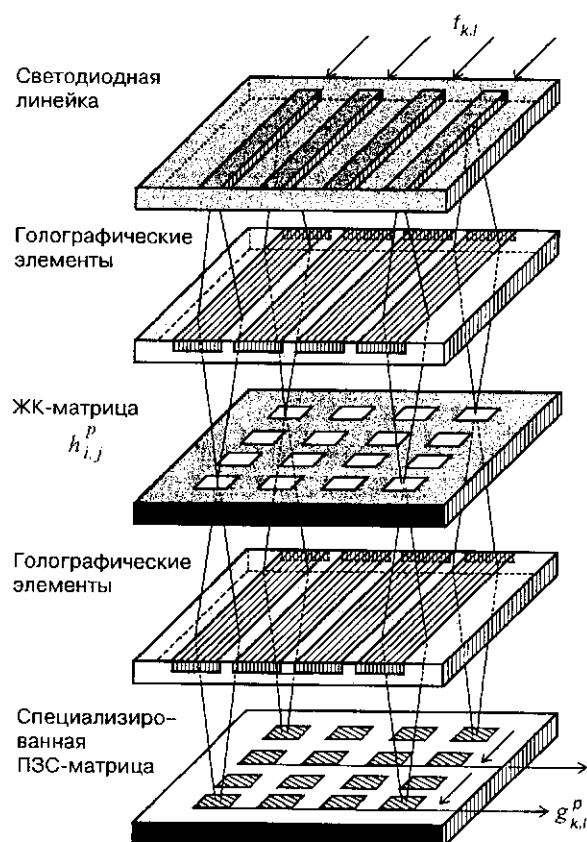


Рис. 4. Интегральный вариант исполнения опико-электронной системы

на входе для рабочего поля 10^6 элементов требуется мощность $\sim 0,05$ Вт. Реально при распознавании предварительно обработанных изображений, характеризующихся контурами, это значение может снизиться более чем на порядок.

Таким образом, показано, что применение в нейронных сетях операторов, инвариантных к сдвигу (в принципе могут рассматриваться и другие инвариантные операторы), позволяет значительно упростить процедуру обучения нейронной сети, существенно уменьшить число синаптических связей и использовать для ее реализации опико-электронные средства, весьма эффективные с точки зрения вычислительной производительности. Предложенная нейросетевая опико-электронная система с эквивалентной производительностью $\sim 10^{10}$ операций сложения/умножения может быть выполнена в виде нейрочипа размером $\sim 2 \times 2$ см с потребляемой мощностью $\sim 0,05$ Вт.

СПИСОК ЛИТЕРАТУРЫ

1. Neifeld M. A. Optical dual-scale architecture for neural image recognition // Appl. Opt. 1995. 34, N 26.

2. Горбань А. Н., Россиев Д. А. Нейронные сети на персональном компьютере. Новосибирск: Наука, 1996.
3. Гофман М. А., Неженко Е. С., Фельдбуш В. И. Оптико-электронные сигнальные процессоры // Автометрия. 1989. № 3. С. 53.
4. High-Resolution Camera // Laser Focus World. 1997. 33, N 1. P. 156.
5. Lerner E. J. Laser diodes and LEDs light optoelectronic devices // Ibid. N 2. P. 109.
6. Miniature display // Ibid. P. 212.
7. Shaped-output Lamps // Optoelectronics World (Suppl. to Laser Focus World). 1998. P. 18.
8. Козик В. И., Твердохлеб П. Е. 3-D оптические интегральные схемы ассоциативной памяти // Автометрия. 1993. № 3. С. 44.

Поступила в редакцию 19 мая 1999 г.

Реклама продукции в нашем журнале – залог Вашего успеха!