

УДК 007.001.362; 681.327.12.001.362

М.Е. Ваткин

## ОПТИМИЗАЦИЯ СТРУКТУРЫ И АЛГОРИТМОВ НЕЙРОННОЙ СЕТИ «НЕОКОГНИТРОН»

*Рассмотрены особенности построения нейронной сети «неокогнитрон». Приведен анализ эффективности алгоритмов обучения и активации нейронной сети применительно к задаче распознавания изображений. На основе приведенного анализа предложен ряд оптимизаций структуры и алгоритмов нейронной сети, что позволяет значительно повысить точность распознавания.*

### Введение

В 1962 - 65 гг. были опубликованы работы D.H. Hubel и T.N. Wiesel [1, 2], посвященные изучению организации биологических нейронов в зрительной коре мозга млекопитающих. Данные исследования показывают, какие преобразования претерпевает зрительная информация, получаемая сетчаткой глаза в процессе своего продвижения по зрительному тракту мозга. Если обобщить эти исследования, то получим следующие результаты:

1. Поступающее изображение преобразуется в активность нейронов сетчатки глаза, создавая, таким образом, некоторое представление изображения.

2. Далее, по мере продвижения по зрительному тракту, это представление претерпевает ряд послонных преобразований.

3. В каждом слое нейроны группируются в области, которые выделяют одинаковые признаки на изображении, т.е. если изображение равномерно покрыто одним и тем же признаком, то наиболее активной становится группа нейронов, расположенная в некоторой локальной области слоя.

4. При переходе к последующему слою выделяются более крупные признаки изображения, т.е. признаки, занимающие большую площадь по сравнению с признаками, выделяемыми в предыдущем слое.

5. В результате того, что рецепторные поля соседних нейронов из одной группы имеют перекрытие, и того, что эти нейроны находятся в рецепторном поле следующего за ними «обобщающего» нейрона, который активизируется, если в его рецепторном поле присутствует хотя бы один активный нейрон, «обобщающий» нейрон может выделять слегка смещенные признаки. Величина смещения определяется величиной перекрытия рецепторных полей нейронов первого уровня и размерами рецепторного поля «обобщающего» нейрона.

Эти исследования вдохновили Фукушиму на создание искусственной нейронной сети «неокогнитрон», модулирующей процессы, происходящие в зрительном тракте мозга [3-6]. Поскольку автор пытался структурно приблизить свою модель к биологическому аналогу, то в результате получилась модель, которая имеет ряд неоптимальных решений.

Обобщим сущность структурных и алгоритмических решений, применяемых в нейронной сети «неокогнитрон»:

1. Во всех идентифицируемых изображениях выделяют два типа признаков: простые признаки, которые присущи всем изображениям, такие как линии, точки пересечения линий, и комплексные признаки, которые присущи только определенному классу изображений, такие как определенная пространственно-геометрическая комбинация простых признаков.

2. В соответствии с критерием сложности все признаки делятся по уровням, где более сложный уровень признака образуется путем пространственно-геометрической комбинации признаков предыдущего уровня. Каждый уровень образует группу признаков одной сложности.

3. Строится многослойная сеть, где количество слоев совпадает с количеством уровней сложности. Каждый слой делится на подслои в соответствии с количеством признаков в данной группе. Функция отдельного подслоя заключается в обнаружении (выделении) соответствующего признака на всем изображении, т.е. отдельный подслои строит своего рода карту распре-

деления данного признака на изображении. Для обеспечения инвариантности сети к небольшим сдвигам признаков в слои вводятся подслои, которые дублируют карты признаков соответствующих им подслоев, при этом в каждой точке обнаружения признаков производится его «размазывание» в некоторой окрестности.

4. Идентификация неизвестного образа производится путем последовательной активации слоёв: от слоя, выделяющего простые признаки, до слоя, выделяющего признаки всего образа в целом.

Из описанного выше обобщения последовательности задач при создании «неокогнитрона» можно выделить следующие концепции и решения, которые могут рассматриваться как объекты исследования:

- 1) способ выделения признака (нейрон), правило его обучения и активации;
- 2) метод вычисления критерия, оценивающего эффективность набора признаков, входящих в одну группу, который позволяет оптимизировать количество подслоев вручную или автоматически;
- 3) метод, определяющий, как должны сочетаться простые признаки при формировании признака более высокого порядка.

Рассмотрим более детально каждый из объектов исследования.

### 1. Анализ правила активации нейрона

Рассмотрим совокупность правил нейронной сети «неокогнитрон», которые обеспечивают выделение локального признака, инвариантного к смещениям.

1. *Правило выделения признака* (рис. 1). Реализуется с помощью двух нейронных слоев: тормозящего и выделяющего.

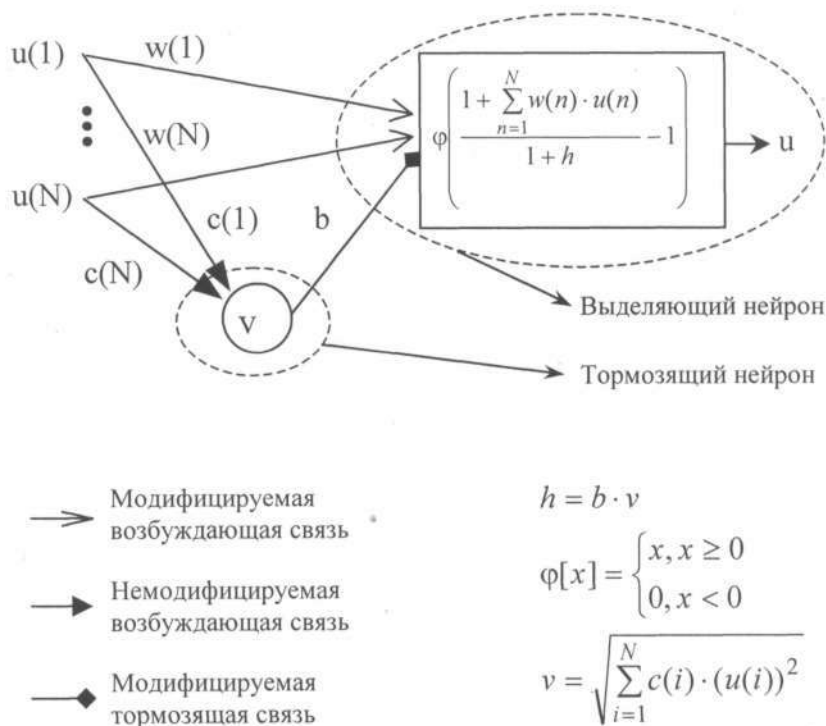


Рис. 1. Структура связей выделяющего слоя

2. *Правило, обеспечивающее инвариантность*. Реализуется с помощью дополнительного обобщающего слоя нейронов, в котором происходит «размазывание» признака (рис. 2).

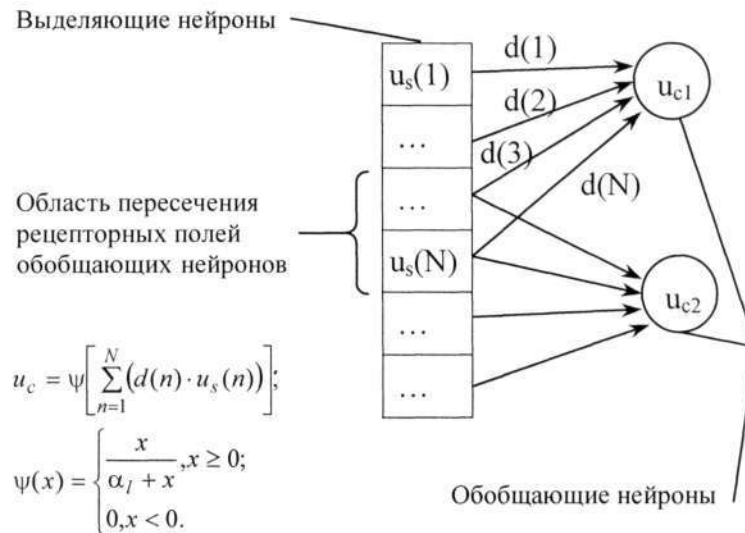


Рис. 2. Структура связей обобщающего слоя

Общий смысл правила активации для обобщающих нейронов таков: если выделенный признак попадает в область пересечения рецепторных полей нейронов, то будут активизироваться оба нейрона. Таким образом, активность одного нейрона из предыдущего слоя переходит в активность нескольких нейронов последующего слоя, что приводит к «размазыванию» признака в данном слое.

Описанная выше модель выделения локального признака, инвариантного к смещениям, близка к ее биологическому аналогу, но с точки зрения машинной реализации (т.е. экономии объема требуемой памяти и машинного времени) возникает ряд спорных вопросов:

- 1) необходимость столь сложного правила для выделения локального признака;
- 2) эффективность применения «размазывания» в качестве правила, обеспечивающего инвариантность;
- 3) необходимость использования дополнительного слоя нейронов для обеспечения инвариантности.

Таким образом, для машинной реализации одного нейрона из нейронной сети «неокогнитрон» необходимы:

- 1)  $N + 2$  ячеек памяти для хранения весовых коэффициентов и результата активации выделяющего нейрона;
- 2)  $N + 1$  ячеек памяти для хранения весовых коэффициентов и результата активации тормозящего нейрона;
- 3)  $N' + 1$  ячеек памяти для хранения весовых коэффициентов и результата активации обобщающего нейрона;
- 4)  $2N+2$  операций умножения для вычисления активности тормозящего нейрона;
- 5)  $N+3$  операций умножения для вычисления активности выделяющего нейрона;
- 6)  $N' + 2$  операций умножения для вычисления активности обобщающего нейрона.

Операции деления и извлечения квадратного корня считаются эквивалентными операции умножения, а  $N$  операций сложения – эквивалентными одной операции умножения. Итого, необходимы  $2N+N'+3$  ячеек памяти и  $(3N + 5)(N' + 1)$  операций умножения.

## 2. Модификация структуры нейрона

В качестве альтернативы рассмотренному предлагается следующий подход. Пусть есть шаблонное изображение, представленное матрицей весовых коэффициентов нейрона, и анализируемое изображение, представленное матрицей значений яркостей пикселей. Тогда задача

выделения признака ставится как вычисление расстояния между шаблонным изображением и анализируемым.

Поэтому для выделения локального признака предлагается использовать более простую модель нейрона, идентичного нейрону из SOM нейронной сети [7-14], правило обучения которого соответствует функции вычисления динамического среднего:

$$w(t+1) = w(t) + \frac{1}{t+1} \cdot [u(t+1) - w(t)]. \quad (1)$$

Правило активации соответствует некоторой функции расстояния между весовыми коэффициентами нейрона и значениями активностей нейронов предыдущего слоя в области рецепторного поля нейрона. Новая структура нейрона изображена на рис. 3.

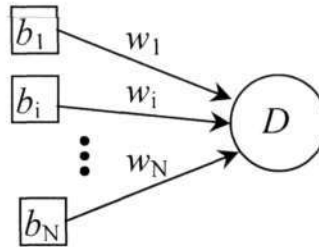


Рис. 3. Новая структура нейрона:  $B = \{b_i\}$  – локальная часть входного изображения;  $W = \{w_i\}$  – соответствующие весовые коэффициенты нейрона;  $N$  – количество пикселей, составляющих локальную часть

### 3. Исследование свойств функции активации нейрона, инвариантной к смещениям признака внутри рецепторного поля

Определим вид искажения изображения, относительно которого функция активации нейрона должна быть инвариантна. Пусть искаженное изображение получается путем смещения пикселей шаблонного изображения. Тогда для каждого пикселя можно определить  $\vec{v}_{x,y}$  – вектор его смещения относительно начального положения. Отметим, что для любых двух соседних пикселей  $x_1, y_1$  и  $x_2, y_2$  справедливо  $|\vec{v}_{x_1,y_1} - \vec{v}_{x_2,y_2}| \leq \theta$ , где  $\theta \leq \max(|\vec{v}_{x_1,y_1}|, |\vec{v}_{x_2,y_2}|)$ . Данное свойство указывает на однородность искажения, т.е. на то, что соседние пиксели будут смещены почти одинаково. Определим радиус искажения  $R$  как среднее значение по всем модулям векторов смещений пикселей изображения.

Определим понятие «отличительный признак изображения» как часть изображения, по которому данное изображение можно отличить от других. Отметим, что такой признак, как правило, является комбинацией более мелких признаков.

Определим понятие «атомарный признак изображения» как признак изображения, который нельзя разделить на более мелкие признаки. Отметим, что граница области изображения, образующая атомарный признак, является почти выпуклой фигурой. Свойство «почти выпуклая фигура» означает, что для такой фигуры любая секущая, проведенная через ее центроид, пересекает границу лишь дважды. Определим параметр «наименьший диаметр атомарного признака» как наименьшую длину секущей, проведенной через центроид фигуры, образованной атомарным признаком. Определим среднее наименьших диаметров атомарных признаков  $evS$ .

Тогда искомая функция расстояния  $D(t, p)$  между шаблонным изображением  $t$  и искаженным изображением  $p$ , инвариантная к описанному выше виду искажений изображения, должна иметь следующие характеристики:

1) область значений функции  $D(t, p)$  ограничена и определена на интервале  $[0 \dots 100] \%$ ;

2) если  $\frac{R}{evS} \rightarrow 0$ , то  $D(t, p) \rightarrow 100$ ;

3) если  $\frac{R}{evS} \rightarrow 1,3$ , то  $D(t, p) \rightarrow 0$ .

Третье свойство указывает на то, что если радиус искажения превышает средний диаметр атомарных признаков на 30%, то в результате такого искажения изображение перемешивается настолько, что нельзя узнать атомарные признаки, поэтому невозможно построить отличительный признак, а следовательно, невозможно идентифицировать искаженное изображение как относящееся к классу шаблонного изображения. Для того чтобы реализовать описанные свойства, предлагается использовать структуру функции активации нейрона, изображенную на рис. 4, на котором  $f(B_i, w_i)$  – функция, осуществляющая поиск соответствующего  $w_i$  элемента во множестве  $B_i$ , такая, что  $f(B_i, w_i) = b_k, |b_k - w_i| \rightarrow \min, \forall b_k \subseteq B_i$ ; значение  $K$  определяет диаметр области поиска  $B_i$ ;  $D$  – функция расстояния между весовыми коэффициентами  $W = \{w_i\}$  и соответствующими значениями функции поиска  $f$ . Отметим, что для того чтобы выполнялось третье свойство, необходимо, чтобы  $K = R = 1,3 \cdot evS$ .

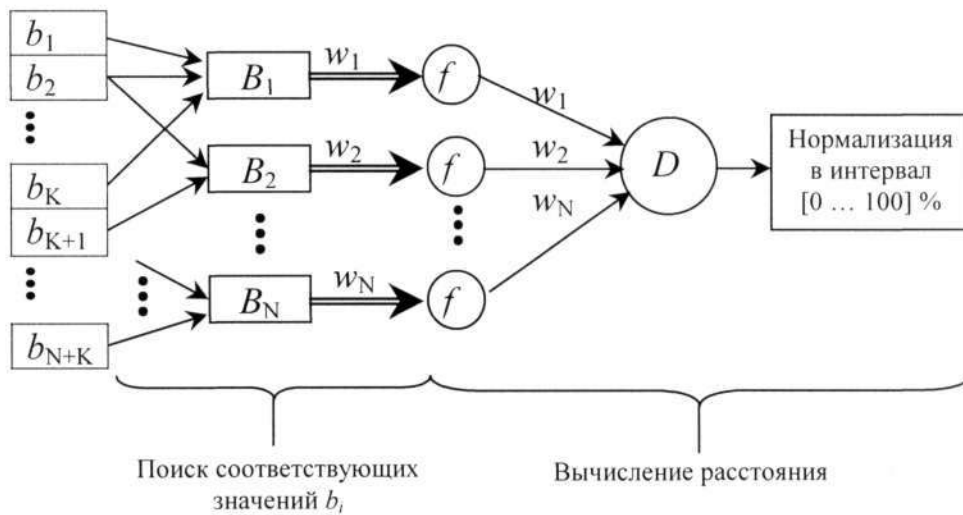


Рис. 4. Структура связей нейрона

Свойства 1, 2, 3 новой функции активации нейрона относятся к случаю, когда искажение изображения получается путем смещения пикселей и при этом сохраняется их изначальная яркость, т.е.  $\forall b_k \subseteq B_i, \min(|b_k - w_i|) = 0$ . Рассмотрим теперь случай, когда дополнительно может меняться еще и яркость смещаемого пикселя, т.е.  $\forall b_k \subseteq B_i, \min(|b_k - w_i|) \geq 0$ , где  $B_i$  – область поиска пикселя, соответствующего коэффициенту  $w_i$ .

Отметим, что любое изображение содержит пиксели, относящиеся к разным информативным признакам изображения. Таким образом, получаем некоторое разбиение изображения в соответствии с отношением пикселей к информативным признакам. Большинство функций расстояния между множествами являются интегральными по этим множествам, т.е. результат таких функций аккумулирует разности соответствующих элементов по всему объему множеств. Если вычислять такую функцию расстояния для двух изображений (множеств пикселей), объединяющих в себе более одного информативного признака, то в результате получим усреднение разностей признаков на площадь изображения, что не оптимально при неравных площадях отдельных признаков. Например, будет получен неправильный результат функции расстояния, стремящийся к минимуму для случая, когда площадь неважного признака гораздо больше важного и разность для неважного признака стремится к минимуму, в то время как для важного стремится к значительному числу. Поэтому предлагается считать функцию расстояния для каждого класса признаков изображения отдельно.

Пусть  $M^t = \{m_i^t\}$  – некоторое разбиение шаблонного изображения  $T$ , а  $\hat{M}^p = \{\hat{m}_i^p\}$  – соответствующее ему разбиение искаженного изображения  $P$ . И наоборот,  $M^p = \{m_i^p\}$  – некоторое разбиение искаженного изображения, а  $\hat{M}^t = \{\hat{m}_i^t\}$  – соответствующее ему разбиение шаблонного изображения. Соответствующее разбиение строится с помощью функции поиска соответствующих пикселей, которая описана выше. Определим функцию расстояния для отдельной пары элементов этих разбиений  $d(m_i, \hat{m}_j)$ , где  $i = j$ ,  $m_i \in M$ ,  $\hat{m}_j \in \hat{M}$ . Поскольку всегда  $i = j$ , то для простоты будем обозначать функцию расстояния для пары элементов как  $d(m, \hat{m})$ . Отметим, что  $d(m, \hat{m})$  может быть несимметричной относительно того, какое изображение разбивается первым, т.е.  $d(m^t, \hat{m}^p) \neq d(\hat{m}^t, m^p)$ . Определим функцию, определенную на множествах  $M$  и  $\hat{M}$ , которая вычисляет расстояние для каждой соответствующей пары элементов  $fd(m^t, m^p) = \max(d(m^t, \hat{m}^p), d(\hat{m}^t, m^p))$ . Затем определим функцию расстояния между изображениями  $T$  и  $P$  как  $D = \max(fd(m_i^t, m_i^p))$ , где  $i = \overline{1 \dots N}$ ,  $N$  – количество элементов разбиений.

Самым простым способом определения классов информативности является разделение гистограммы яркости на две части посередине. Все пиксели, яркость которых ниже установленного порога, относят к первому классу информативности, а пиксели, яркость которых выше, – ко второму. Такой способ разбиения обосновывается тем, что наиболее информативными областями на изображении являются области локальных минимумов и максимумов яркости.

С целью выбора наилучшей функции расстояния для отдельной пары элементов разбиений предлагается исследовать характеристики правил активации, соответствующих различным метрикам. Эффективность применения функции расстояния будет оцениваться с помощью следующих критериев:

- затрат машинного времени для вычисления функции расстояния;
- адекватности вида функции расстояния, которая оценивается с точки зрения свойств инвариантности функции определенному виду искажений.

Затраты машинного времени для вычисления функции расстояния будут оцениваться исходя из количества операций умножения. Операции деления и извлечения корня будут считаться эквивалентными операции умножения, а операции сравнения и вычитания – эквивалентными операции сложения, где  $N$  операций сложения будут эквивалентны одной операции умножения.

Рассмотрим функцию расстояния, соответствующую угловой метрике:

$$d = \arccos \left( \frac{\sum_i (b_i \cdot w_i)}{\sqrt{\sum_i b_i^2 \cdot \sum_i w_i^2}} \right). \quad (2)$$

Для вычисления функции необходимы  $3N+6$  операций умножения. Эта функция является инвариантной к сдвигу в частотном пространстве изображения, т.е. наиболее подходит для случая, когда необходимо сравнивать слабо контрастные и сильно контрастные изображения. Но в то же время эта функция является наиболее требовательной к ресурсу машинного времени.

Рассмотрим функцию расстояния, соответствующую евклидовой метрике. Чтобы снизить вычислительные нагрузки, упростим ее до вида

$$d = \frac{\sum_i |b_i - w_i|}{N}. \quad (3)$$

Для вычисления функции необходимы три операции умножения, что делает ее более привлекательной в смысле затрат. Недостатком является то, что она не имеет никаких свойств инвариантности.

Рассмотрим модификацию последней функции расстояния

$$d = \frac{\sum_i^N \delta(|b_i - w_i|)}{N}, \quad \delta(x) = \begin{cases} 1, & x \leq \varepsilon; \\ 0, & x > \varepsilon. \end{cases} \quad (4)$$

Для вычисления функции необходимы четыре операции умножения. Но она является инвариантной к случайным колебаниям яркости пикселей в диапазоне  $[-\varepsilon; \varepsilon]$ , что наиболее адекватно рассмотренному выше виду искажений, поэтому предлагается использовать именно ее для вычисления функции расстояния.

Итак, для реализации схемы, предложенной на рис. 4, с учетом выбранной функции расстояния необходимо:

1)  $N$  ячеек памяти для хранения весовых коэффициентов, одна ячейка для хранения значения  $\varepsilon$ , одна ячейка для хранения значения допустимого радиуса искажения  $R$ , одна ячейка для хранения результата активации нейрона. Итого  $N+3$  ячеек памяти;

2)  $R$  операций сравнения для вычисления функций сравнения поиска  $f(B_i, w_i)$  или одна операция умножения,  $N/C$  операций умножения и одна деления для вычисления функции  $d(m, \hat{m})$ ,  $(N/C+1)2$  операций умножения для вычисления  $fd(m', m'')$ . Итого, для вычисления активности нейрона  $D = \max(fd(m', m''))$  необходимо  $(N/C+1)2C = 2N+2C$  операций умножения.

Таким образом, данный подход позволяет снизить затраты ресурсов, необходимых для вычисления правила активации нейрона:

- 1)  $2N + N' + 3 > N+3$  – для памяти;
- 2)  $3N + N' + 5 > 2N+2C$  – для времени.

#### 4. Модификация правила обучения нейрона

Поскольку была изменена функция активации нейрона, необходимо также модифицировать правило обучения нейрона SOM нейронной сети. В связи с этим предлагается использовать структуру правила обучения, аналогичную структуре правила активации, изображенной на рис. 4, в которой сохраняется часть, осуществляющая поиск соответствующих значений  $b_i$ , а часть, осуществляющая вычисление расстояния, заменяется правилом вычисления динамического среднего для весовых коэффициентов (1).

#### 5. Трудности подбора значений параметров радиуса геометрических и яркостных искажений для новой функции активации и обучения

Несмотря на видимые улучшения, предложенные выше, для правил обучения и активации нейрона остаются нерешенными несколько вопросов, которые касаются подбора параметров  $\varepsilon$ ,  $R$ . Отметим, что если  $R$  является радиусом искажения изображения в геометрическом смысле, т.е. в смысле перемешивания пикселей, то  $\varepsilon$  можно назвать радиусом искажения изображения в яркостном смысле, т.е. в смысле изменения яркости пикселей. Подбор этих параметров осложняется еще и тем, что они не являются взаимно независимыми. Данный факт поясняет следующий пример: любое искажение можно рассматривать с двух крайних точек зрения: 1) присутствуют только яркостные искажения и не присутствуют геометрические; 2) присутствуют только геометрические искажения и нет яркостных. Таким образом, очевидно, что в любом изображении присутствуют оба вида искажений и они некоторым образом связаны друг с другом. Конечно, данные параметры для определенного вида изображений могут быть получены экспериментально. Но это неоптимально с точки зрения построения автоматизированной системы распознавания образов, где заранее неизвестны количественные параметры искажения, к которому система должна быть инвариантна.

Допустим, что для обучения системы у нас есть одно идеальное изображение и несколько искаженных, тогда из идеального изображения можно найти максимально допустимый радиус геометрических искажений, а сравнивая идеальное и искаженное изображения, можно опреде-

лить радиус искажения яркости. Наиболее трудоемким является процесс определения радиуса геометрических искажений, поэтому рассмотрим сначала его.

Из свойств 2 и 3 функции расстояния между шаблонным и искаженным изображениями видно, что для нахождения радиуса  $R$  необходимо определить среднее наименьших диаметров атомарных признаков  $evS$ . Для того чтобы определить  $evS$ , необходимо построить алгоритм, который разбивает исходное изображение на области пикселей, принадлежащие различным атомарным признакам.

Поскольку построение алгоритма разбиения изображения на области, содержащие отдельные атомарные признаки, является достаточно сложной вычислительной задачей, то предлагается оценивать значение  $evS$  по разбиению изображения на области, содержащие отдельные классы информативности, что является менее трудоемкой задачей.

Далее отдельно для областей каждого класса информативности построим их скелет и определим  $evS = \frac{1}{C} \sum_c \left( \frac{1}{K} \sum_k S_{k,c} \right)$ , где  $S_{k,c}$  – минимальное расстояние от  $k$  точки скелета до границы области класса информативности  $c$ ;  $K$  – количество точек скелета;  $C$  – количество классов информативности. Используя третье свойство, получим

$$R = 1,3 \frac{1}{C} \sum_c \left( \frac{1}{K} \sum_k S_{k,c} \right). \quad (5)$$

Далее, используя  $R$ , определим радиус искажения яркости

$$\varepsilon = \max(|a_i - f(B_i, a_i)|), \quad \forall i, \quad (6)$$

где  $A = \{a_i\}$  – множество точек шаблонного изображения;  $B_i$  – области точек искаженного изображения, которые соответствуют областям поиска соответствующих пикселей для  $a_i$  пикселя шаблонного изображения, причем диаметр  $B_i$  составляет  $2R$  пикселей;  $f$  – функция, осуществляющая поиск соответствующего элемента  $a_i$  во множестве  $B_i$  по формуле  $f(B_i, a_i) = b_k, \forall b_k \subseteq B_i$ , такого, что  $|b_k - a_i| \rightarrow \min$ .

Если у нас есть несколько искаженных изображений, то тогда значение  $\varepsilon$  можно вычислить как

$$\varepsilon = \frac{1}{L} \varepsilon_l = \frac{1}{L} \sum_l \max(|a_i^l - f(B_i^l, a_i^l)|), \quad (7)$$

где  $L$  – количество доступных искаженных образцов изображений.

## 6. Объединение нейронов в группу, выделяющую одинаковый признак на изображении

Для того чтобы была возможность выделения одинакового признака на всем изображении, нейроны организуются в группу, каждый нейрон из которой выделяет один и тот же локальный признак, но расположенный в различных позициях изображения. На рис. 5 приведен пример структуры связей нейронов из одной группы в случае одномерных входных данных.

Размер выделяемого признака изображения обычно больше размера областей перекрытия соседних нейронов и меньше либо равен размеру области поиска локального признака. Для случая, когда входные данные представляют собой двумерный массив, нейроны тоже организуются в двумерную структуру. Таким образом, активность нейронов из одной группы формирует карту распределения выделяемого признака на входном изображении.



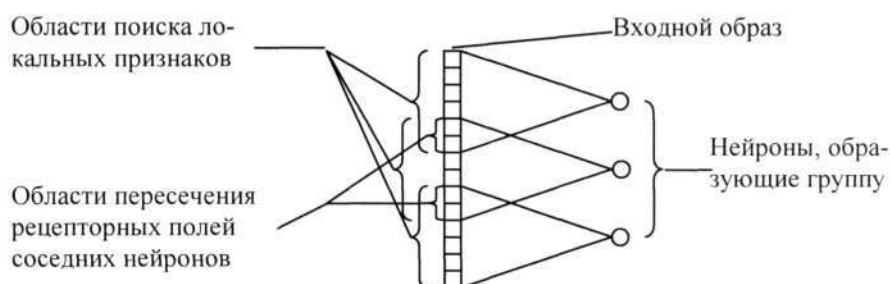


Рис. 5. Схема объединения нейронов в группу

### 7. Объединение нейронов в слое и оптимизация количества групп нейронов в слое

Поскольку процесс послойной обработки информации подразумевает постепенное укрупнение выделяемых локальных признаков, которые как бы перерастают в глобальный признак принадлежности входного изображения к тому или иному классу, то вполне логично выглядит объединение в слое группы нейронов, выделяющие признаки с одинаковыми «габаритными» размерами. Этот метод позволяет построить алгоритм, реализующий генерацию групп нейронов в слое автоматическим или полуавтоматическим способом.

Построение такого алгоритма базируется на следующих предположениях:

1. Так как нейроны, входящие в одну группу, имеют одинаковые матрицы весовых коэффициентов, то всю группу можно представить в виде одного нейрона, а слой представить в виде SOM нейронной сети, в которой каждый нейрон соответствует одной из групп.

2. Для обучения такой сети необходимо использовать образы, получаемые из окна, размер которого равен размеру области поиска выделяемых в этом слое признаков изображения. Причем в автоматическом режиме обучения (без учителя) каждый обучающий образ получается в результате сканирования исходного обучающего изображения с единичным шагом смещения окна. А в полуавтоматическом режиме обучения в обучающую выборку оператор добавляет еще дополнительные изображения признаков с названиями их классов.

Этапы функционирования алгоритма обучения:

1. Определить для начального состояния сети количество нейронов, равное нулю, и применить модифицированный алгоритм обучения SOM нейронной сети [15].

2. После обучения сети использовать такую нейронную сеть как прототип для построения искомого слоя в многослойной сети, в котором количество нейронов в прототипе будет соответствовать количеству групп, а весовые коэффициенты нейронов из прототипа – весовым коэффициентам нейронов из соответствующих групп.

### 8. Общая архитектура сети

Общая архитектура сети изображена на рис. 6, на котором рецепторный слой преобразует яркости пикселей изображения в значения активностей нейронов, внутренние слои производят последовательную обработку поступающей информации, а выходной слой сети генерирует ответ, к какому классу относится входное изображение в виде максимальной активности одного из своих нейронов, причем каждый нейрон в этом слое входит в отдельную группу. Из рисунка видно, что каждый нейрон имеет часть рецепторного поля в каждой из групп предыдущего слоя.

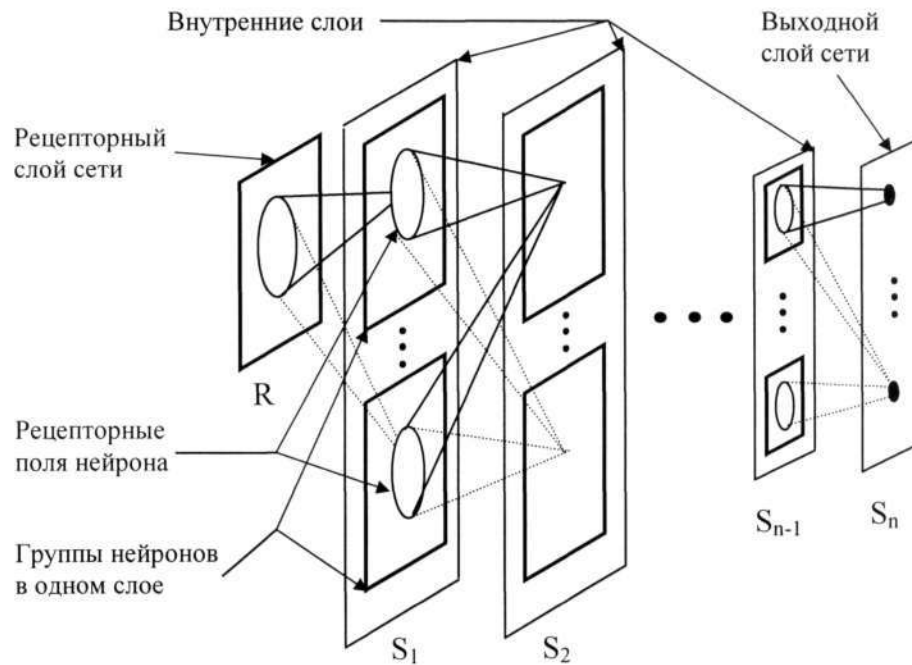


Рис. 6. Структура нейронной сети

### Заключение

В статье проводится анализ структуры, а также алгоритмов нейронной сети «неокогнитрон» с целью улучшения характеристик скорости обработки информации и точности распознавания. Архитектура разработанной нейронной сети представляет собой новый класс многослойных нейронных сетей, который использует самоорганизующиеся отображения и сочетает в себе структурные решения нейронной сети «неокогнитрон» и нейронной сети Кохонена. В основе правила активации нейронов лежит новый метод сравнения матриц яркостей изображений, который позволяет сравнивать матрицы с нечетким позиционированием соответствующих элементов матриц. Данный метод, в свою очередь, основывается на расчете геометрических и яркостных искажений по набору изображений объекта. Особенностью данного правила активации является инвариантность результатов сравнения к локальным искажениям формы объекта.

### Список литературы

1. Hubel D.H. and Wiesel T.N. Receptive fields, binocular interaction and functional architecture in cat's visual cortex // *J.Physiol (London)*. - V. 160. - Jan. 1962. - P. 106-154.
2. Hubel D.H. and Wiesel T.N. Receptive fields and functional architecture in two nonstriate visual area (18 and 19) of the cat // *J. Neurophysiol.* - V. 28. - 1965. - P. 229-289.
3. Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position // *Biol. Cybern.* - V. 36. - 1980. - P. 193-202.
4. Fukushima K., Miyake S. Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position // *Pattern Recognition*. - V. 15. - 1982. - P. 455-469.
5. Fukushima K., Miyake S., Ito T. Neocognitron: A neural network model for a mechanism of visual pattern recognition // *IEEE Trans. Syst., Man, Cybern.*, V. SMC-13. - 1983. - P. 826-834.
6. Fukushima K., Wake N. Handwritten alphanumeric character recognition by the Neocognitron // *IEEE Trans. on Neural Networks*. - V. 2. - № 3. - 1991. - P. 355-365.
7. Moody J. and Darken C.J. Fast learning in networks of locally-tuned processing unit // *Neural computation*. - 1(2). - 1989. - P. 281-294.
8. Powell M.J.D. Radial basis function for multivariable interpolation: a review // *Algorithm for approximation*. - Oxford: Clarendon Press, 1987. - P. 143-167.

9. Foor W. Adaptive optical radial basis function neural network for handwritten digit recognition // Applied optics. - V.34. - №32. - 1995. - P. 7545-7555.
10. Teuvo Kohonen. Self-Organizing Maps. - Springer-Verlag, 1995.
11. Neifeld M.A., Psaltis D. Optical implementation of radial basis classifiers // Applied Optics. - V. 32. - №. 8. - 1993. - P. 1370-1379.
12. Bishop C.M. Neural networks for pattern recognition. - Oxford: Clarendon press, 1997. - P. 165-191.
13. Simpson P.K. Neural Networks Applications. - IEEE; 1st edition, 1996. - 963 p.
14. Hassoun, Mohamad H. A Bradford book. - Fundamentals of artificial neural networks. - Cambridge; London: Massachusetts inst. of technology, 1995. - 511 p.
15. Ваткин М.Е. Распознавание рукописных символов на основе модифицированной SOM нейронной сети // Мат. I междунар. конф. – Мн.: БГУ, 2002. - Ч. 2. - С. 201-206.

Поступила 12.01.04

*Объединенный институт проблем  
информатики НАН Беларуси,  
Минск, Сурганова, 6  
e-mail: vatkin@lsi.bas-net.by*

**M.E. Vatkin**

#### **OPTIMIZATION OF «NEOCOGNITRON» NEURAL NETWORK STRUCTURE AND ALGORITHMS**

Characteristics of creation of «neocognitron» neural network are considered. The analysis of the efficacy of the training and activation rule is considered applying to the image recognition task. Using the analysis the optimizations are proposed that allow reducing the performance costs and increasing the recognition accuracy.