
НЕКОТОРЫЕ АЛГОРИТМЫ ОБРАБОТКИ ИЗОБРАЖЕНИЙ И ИХ РЕАЛИЗАЦИЯ НА НЕЙРОСЕТЯХ

И.Н. Айзенберг

Ужгородский государственный университет (Украина)

Аннотация

Статья посвящена реализации некоторых оригинальных алгоритмов обработки изображений на клеточных нейросетях (CNN) из многозначных и универсальных нейроэлементов. Рассматриваются высокоэффективные, т.н. многозначные фильтры, их применение для фильтрации шума и коррекции частот. Рассматриваются также алгоритмы выделения контуров, фильтрации импульсных помех и их реализация на нейросетях. Исследуется также применение нейросетей из многозначных нейроэлементов для решения задач распознавания изображений и экстраполяции ортогонального спектра изображения в область высоких частот с целью повышения разрешающей способности оптической системы.

Ключевые слова: обработка изображений, нейросети, комплекснозначные нейроны.

1. Введение

Клеточные нейросети (CNN), введенные в [1], являются отличным аппаратом для решения различных задач обработки и распознавания изображений. CNN - это двумерный массив нейроэлементов размерности $N \times M$. Каждый нейрон в такой сети соединен только с нейронами из своей ближайшей ($n \times m$) окрестности. Такая конструкция сети очень удобна для реализации алгоритмов фильтрации в пространственной области. Существует целый ряд подобных применений непрерывных CNN (CCNN) [1]. Дискретные CNN (DTCNN), введенные в [2-3], ориентированы на решение задач обработки изображений, описываемых с помощью булевых пороговых функций. Функциональные возможности CNN могут быть расширены путем использования в качестве базовых иных типов нейронов. Алгоритмы, описываемые при помощи не пороговых булевых функций, не могут

быть реализованы на DTCNN, но их можно реализовать на CNN из универсальных логических элементов (универсальных булевских нейронов) - CNN-UBN, которые были впервые рассмотрены в [4-5]. С другой стороны, CCNN и DTCNN не в полной мере позволяют использовать возможности нелинейной фильтрации, хотя некоторые нелинейные фильтры на них реализованы (реализация ряда фильтров исследована в [6]). CNN на многозначных нейронах (CNN-MVN) является отличным средством для реализации многозначных нелинейных фильтров (MVF) [7]. MVF эффективны для фильтрации гауссовского, равномерного, смешанного (с импульсной составляющей) и даже спекл шумов, а также для усиления высоких и средних частот (с целью выделения деталей). В то же самое время другие свойства многозначных нейроэлементов (MVN), рассмотренные в [8-9] - возможность реализации на одном

нейроне любой частично-определенной многозначной функции, быстро сходящийся алгоритм обучения, могут быть использованы при разработке нейросети для распознавания изображений. Подобные сети предлагались в [8-10], но их возможности были ограниченными. Здесь будет предложена новая сеть из MVN, предназначенная для решения задачи распознавания изображений путем анализа низкочастотной части ортогонального спектра. В [8] было рассмотрено применение MVN для экстраполяции временных рядов. Здесь мы рассмотрим проблему повышения разрешающей способности путем экстраполяции ортогонального спектра в область более высоких частот при помощи предсказания соответствующих коэффициентов на MVN.

2. Многозначные и универсальные булевские нейроэлемент

2.1. Математическая модель

Мы рассмотрим здесь два типа нейронов: универсальный булевский (UBN), реализующий произвольную булеву функцию от n переменных и многозначный нейрон (MVN), реализующий полностью определенные пороговые и частично определенные функции k -значной логики (здесь $k > 0$ в общем случае положительное целое. Заметим, что никаких вычислительных проблем не возникает при k порядка 16384). Итак, основная идея состоит в представлении булевых и k -значных функций от n переменных при помощи $n+1$ комплекснозначных весов W_0, W_1, \dots, W_n :

$$f(x_1, \dots, x_n) = P(w_0 + w_1 x_1 + \dots + w_n x_n), \quad (1)$$

где x_1, \dots, x_n - переменные, от которых зависит функция и P - активирующая функция.

Для многозначного нейрона:

$$P(z) = \exp(i2\pi j/k), \quad (2a)$$

$$\text{if } 2\pi(j+1)/k > \arg(z) \geq 2\pi j/k,$$

с целочисленным выходом:

$$P(z) = j, \quad (2b)$$

$$\text{if } 2\pi(j+1)/k > \arg(z) \geq 2\pi j/k,$$

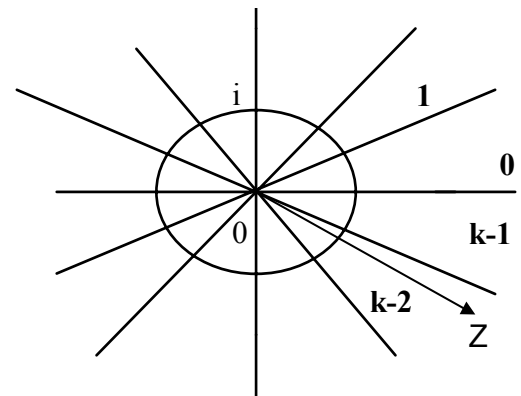
где $j=0,1,\dots,k-1$ - значения k -значной логики, i - мнимая единица, $z = w_0 + w_1 x_1 + \dots + w_n x_n$ - взвешенная сумма, $\arg(z)$ - аргумент комплексного числа z . (значения функции и переменные кодируются комплексными числами, являющимися корнями k -ой степени из единицы: $\varepsilon^j = \exp(i2\pi j/k)$, $j \in [0, k-1]$, иными словами значения k -значной логики представляются корнями k -ой степени из единицы: $j \rightarrow \varepsilon^j$). Формулы (2a-b) иллюстрируется на Рис. 1.

2) Для универсального булевского нейрона:

$$P(z) = (-1)^j \quad (2c)$$

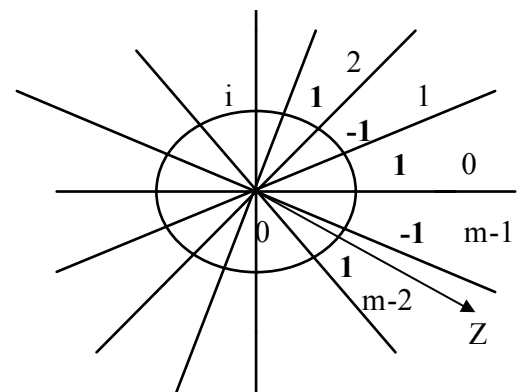
$$\text{if } (2\pi(j+1)/m) > \arg(z) \geq (2\pi j/m)$$

где m - положительное целое $4n \geq m > 0$, j - неотрицательное целое $0 \leq j < m$, $z = w_0 + w_1 x_1 + \dots + w_n x_n$ - взвешенная сумма, $\arg(z)$ - аргумент комплексного числа z . Переменные и функция принимают значения в множестве $\{1, -1\}$. Формула (2b) иллюстрируется на Рис.2. Впервые идея подобного представления многозначных функций была предложена в [11]. Ее дальнейшее развитие позволило предложить аналогичный подход для булевых [4,5,12].



$$P(z) = \varepsilon^{k-2}$$

Рис. 1. Определение активирующей функции многозначного нейрона.



$$P(z) = 1$$

Рис. 2. Определение активирующей функции универсального булевского нейрона.

Главные преимущества MVN и UBN по сравнению с другими типами нейронов (Хопфилда, Чуа, многоуровневый перцептрон, элемент, основанный на свойствах радиально-базисных функций (см. [12]), являются: значительно более высокие функциональные возможности и существование быстро сходящихся алгоритмов обучения. Высокие функциональные возможности означают, что на единственно взятом UBN можно реализовать любые (а не только пороговые) булевы функции любого числа переменных, а на для единственно взятого MVN и заданной многозначной функции всегда можно найти такое k , что данная функция будет реализована на одном нейроне.

2.2. Алгоритмы обучения MVN и UBN

Алгоритмы обучения, которые мы будем использовать здесь, были недавно рассмотрены в [13]. Поэтому мы ограничимся сейчас очень кратким изложением, необходимым для понимания последующих результатов. Очевидно, если взвешенная сумма попадает в “неправильный” сектор на комплексной плоскости (см. Рис. 1, 2), алгоритм обучения состоит в коррекции весов таким образом, чтобы взвешенная сумма попала в один из секторов, соответствующих правильному выходному значению (для UBN), либо в сектор, строго соответствующий ожидаемому выходному значению (для MVN). В [4-5,11] был предложен алгоритм обучения с корректирующим множителем. Он сводится к следующей процедуре коррекции:

$$W_{m+1} = W_m + C \omega \varepsilon^q \bar{X}, \quad (3)$$

где W_m и W_{m+1} - текущий и последующий весовые векторы, \bar{X} - вектор входных сигналов нейрона (комплексно-сопряженный), ω - коэффициент коррекции, ε - первообразный корень m -той или k -той степени из единицы, C - масштабный коэффициент, q - номер “правильного” (ожидаемого) сектора, ω выбирается из соображений, что взвешенная сумма $z = w_0 + w_1 x_1 + w_n x_n$ должна переместиться как можно ближе к “правильному” сектору (или должна сразу попасть в него) после шага обучения, определяемого (3). Для UBN выбор ω основан на следующем соображении. Если взвешенная сумма попала в “неправильный” сектор, то оба соседних к нему сектора будут “правильными”, и двигаться следует именно в тот, к которому текущее значение взвешенной суммы ближе. Для MVN, в случае попадания в неправильный сектор, двигаться нужно в единственно в данном случае “правильный” сектор номер q . ε^q в обоих случаях соответствует правой границе “правильного” сектора (см. Рис. 1-2). Этот алгоритм обучения быстрее сходится для достаточно небольших m и k ($m < 20$, $k < 4096$), но и при произвольных m и k для достаточно гладких функций. Другой алгоритм обучения, рассмотренный в [8, 13-14] быстрее сходится при больших m и k . Он основан на известном правиле “коррекция ошибки”, которое для комплекснозначных нейронов выглядит так:

$$W_{m+1} = W_m + C(\varepsilon^q - \varepsilon^s) \bar{X}, \quad (4)$$

где W_m и W_{m+1} - текущий и последующий весовые векторы, \bar{X} - вектор входных сигналов нейрона (комплексно-сопряженный), ε - первообразный корень m -той, либо k -ой степени из единицы (m и k берутся из (2)), C - масштабный коэффициент, q - номер “правильного” сектора на комплексной плоскости, s - номер “неправильного” сектора, в который попало текущее значение взвешенной суммы. Для MVN q должно равняться ожидаемому выходному значению. Для UBN q выбирается на каждом шаге

из тех же соображений, что в алгоритме (3): если взвешенная сумма попадает в неправильный сектор номер s , ее следует направить в левый ($s+1$)-ый, или правый ($s-1$)-ый соседний сектор в зависимости от того, к какому из них текущее значение взвешенной суммы оказалось ближе.

3. Обработка изображений на клеточных нейросетях из UBN и MVN

3.1. CNN-MVN и CNN-UBN

Мы будем использовать обычную для CNN схему (каждый нейрон соединен только с нейронами из ближайшей к нему прямоугольной (квадратной) окрестности), предложенную в [1], но с UBN, или MVN в качестве базовых нейронов сети. Фрагмент CNN размерности 3×5 с локальными соединениями в окне 3×3 показан на Рис. 3. Динамика каждого нейрона описывается следующим соотношением:

$$Y_{ij}(t+1) = P \left[w_0 + \sum_r w_r^{ij} x_r^{ij}(t) \right], \quad (5)$$

где Y_{ij} - выход нейрона, W_r^{ij} - веса, соответствующие m -тому входу i - j -ого нейрона, X_r^{ij} - m -тый вход i - j -того нейрона, P - активирующая функция нейрона (2), \sum_r означает, что, как обычно в клеточной сети, каждый нейрон соединен только с нейронами из своей ближайшей g -окрестности.

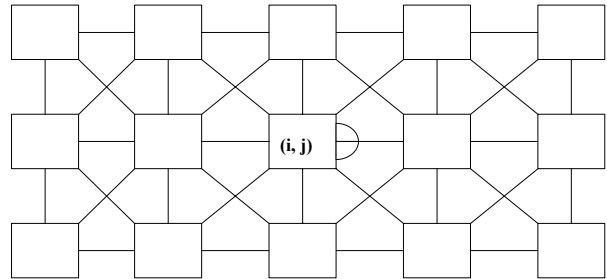


Рис. 3. CNN 3×5 с локальными 3×3 соединениями.

3.2. Многозначная фильтрация и ее реализация на CNN-MVN

Проблема фильтрации изображения в пространственной области сводится к замене текущего значения яркости B_{ij} в i - j -ом пикселе значением \hat{B}_{ij} , которое лучше с некоторой точки зрения. Пусть $B \in \{0,1,\dots,255\}$ - значение яркости. Рассмотрим множество корней k -той степени из единицы ($k=256$). Мы можем поставить в соответствие значению яркости B корень из единицы

$$\varepsilon^B = \exp(i2\pi B/k) = Y \quad (6)$$

Таким образом, мы получим взаимно-однозначное отображение $B \leftrightarrow \varepsilon^B$.

Двумерный многозначный фильтр (MVF) определяется следующим нелинейным преобразованием [7]:

$$\hat{B}_{ij} = P(w_0 + \sum_{\substack{i-n \leq k \leq i+n \\ j-m \leq l \leq j+m}} w_{kl} Y_{kl}), \quad (7)$$

где P - это функция (2а-б), Y связан с B соотношением (6), i, j - координаты фильтруемого пиксела, n и m - апертура фильтра, w_{kl} - коэффициенты фильтра (комплексные в общем случае). Очевидно, что, (7) задает класс фильтров, каждый из которых определяется набором весовых коэффициентов w .

Следует отметить, что без нелинейной функции P и комплексного сложения (если Y будет также заменен соответствующим B), а также при действительных и неотрицательных весовых коэффициентах (7) сведется к хорошо известной низкочастотной линейной фильтрации в пространственной области. Таким образом, нелинейности, вносимые комплексным сложением и функцией P принципиальны, очень важны и дают, как будет показано ниже, замечательный эффект. Заметим, что из соотношений (2а-б), (5) и (7) следует, что многозначные фильтры могут быть легко реализованы на клеточной нейросети из многозначных нейронов.

Специальным случаем MVF является фильтр, также определяемый (7), но одновременно соединяемый с ранговым фильтром. Тем самым мы получим рангово-многозначный фильтр (ROMVF), для которого в (7) будут браться лишь слагаемые, удовлетворяющие ранговому критерию. Иначе говоря, ROMVF - это ранговый фильтр [14], в котором значения сигнала преобразованы в комплексную форму в соответствии с (6), действительное сложение заменено комплексным и функция P , определяемая формулой (2а) применяется к вычисленной взвешенной сумме:

$$\hat{B}_{ij} = P(w_0 + \sum_{\substack{i-n \leq k \leq i+n \\ j-m \leq l \leq j+m}} Y_{kl}). \quad (8)$$

(Взвешивающий ROMVF получится путем соединения фильтра(7) с ранговым критерием [14]).

Проблема реализации MVF на CNN - MVN сводится к нахождению соответствующих наборов весов. Мы рассмотрим здесь наборы весов, определяющие трех представителей семейства фильтров (7). Первый из этих наборов очень эффективен для собственно фильтрации шума. Два других набора решают задачу выделения деталей путем коррекции высоких и средних пространственных частот. Отметим, что мы рассмотрим здесь только действительные веса (но оперирующие с сигналами, преобразованными в комплексную форму в соответствии с (6)).

Очень простой, но весьма эффективный набор весов, обеспечивающий фильтрацию гауссовского, равномерного, смешанного (с импульсной составляющей) и даже спекл-шума, полученный исходя из

эвристических соображений обеспечения мягкого сглаживания высоких частот:

$$W_0 = C; W = \begin{pmatrix} 1 & 1 & 1 \\ 1 & w_{22} & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad (9)$$

где C - константа (может быть равна нулю), W_{22} - параметр. Чтобы удалить шум, содержащий импульсную компоненту, следует рекомендовать $W_{22}=1$. Чтобы фильтровать гауссовский, равномерный, либо спекл-шум (возможно в ходе итеративного процесса) наилучший диапазон значений для W_{22} следующий: $1 \leq W_{22} \leq 10$ (большие значения обеспечивают более надежное сохранение не сглаженными границ деталей на изображении). CNN - MVN с локальным окном 3x3 реализует фильтр, определяемый набором весов (9). Рассмотрим пример применения MVF для фильтрации сложного неоднородного шума.

Рис. 4а представляет исходное изображение, на Рис. 4б представлено то же изображение, но искаженное гауссовским шумом с нулевым средним и импульсным шумом с динамическим диапазоном [0, 255] и искажением 10%. Результат фильтрации набором весов (9) ($W_{22}=1$) представлен на Рис. 4с. Результат фильтрации при помощи ROMVF (8) (окрестность по ближайшим по рангу отсчетам, ранг равен 8) представлен на Рис 4д. Результат чисто ранговой фильтрации (апертура 3x3, окрестность по ближайшим по рангу отсчетам, ранг равен 8) представлен на Рис. 4е для сравнения. Очевидно, что границы на изображении гораздо лучше сохраняются при помощи MVF и ROMVF. В Таблице 1 для сравнения эффективности фильтрации приводятся некоторые объективные числовые характеристики (максимальное отношение сигнал/шум (S/N), оценка дисперсии шума (ND), среднеквадратичное отклонение от оригинала (MSD)), подтверждающие, что MVF и ROMVF гораздо более эффективны, чем ранговый фильтр (ROF).

Возможность решения задачи коррекции частот при помощи линейной фильтрации в пространственной области была обоснована в [15]. Реализация соответствующих фильтров на CCNN исследована в [13, 16]. Коррекция (усиление) частот очень эффективно решает задачу выделения деталей, размеры которых равны апертуре фильтра, либо меньше. Следующий линейный фильтр был предложен для решения задачи в [15]:

$$\hat{B}(i, j) = G_1 B(i, j) + G_2 (B(i, j) - B_m) + G_3 B_m + c. \quad (10)$$

Таблица 1. Сравнение MVF, ROMVF и ROF на тестовом изображении "Lena", искаженном шумом (Рис. 4б).

Изобр.	Lena	Lena+ шум	ROF	ROMVF	MV F
S/N	57.0	35.3	47.9	48.8	48.4
ND (σ)	62.1	491.8	15.1	8.2	12.0
MSD	0	24.6	16.7	13.9	14.3

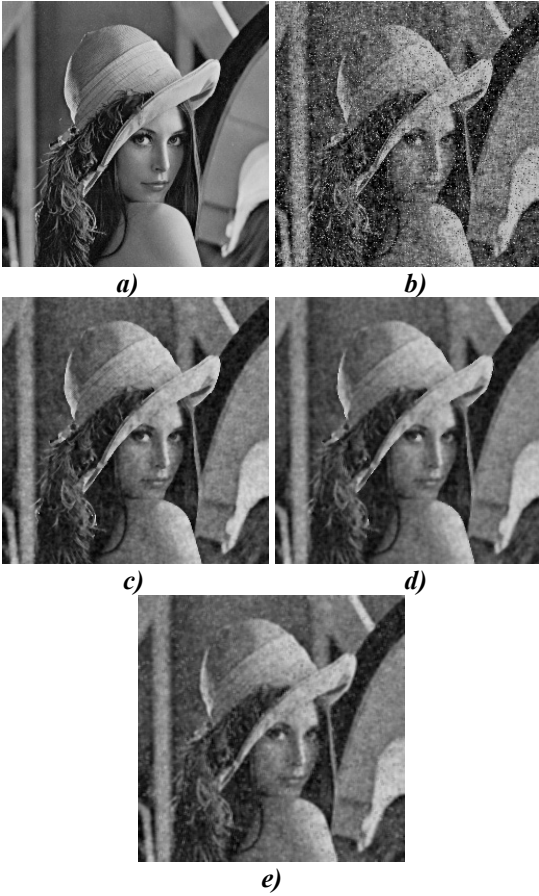


Рис. 4. Многозначная фильтрация: (а) исходное изображение; (б) исходное изображение, искаженное гауссовским шумом с нулевым средним и импульсным шумом с искажением 10%; (с) многозначная фильтрация по набору весов (9); (д) рангово-многозначная фильтрация; (е) ранговая фильтрация (для сравнения).

Этот фильтр осуществляет коррекцию средних и высоких частот, что ведет к выделению на изображении деталей приблизительно средних размеров. Здесь v_m - локальное среднее по окну, окружающему отсчет $B(i,j)$; $\hat{B}(i,j)$ значения яркости (ij)-того пиксела до и после обработки соответственно, G_1, G_3 - коэффициенты, определяющие коррекцию низких частот, G_2 определяет коррекцию высоких частот, c - константа, влияющая на яркость фона после обработки. Апертура фильтра (10) должна быть приблизительно равна размерам деталей, которые следует выделить. Как было показано выше, MVF, определяемые формулами (6) - (7), можно рассматривать как своеобразное обобщение линейных низкочастотных фильтров. Учитывая, что фильтр (10) - линейный, его естественно также несложно обобщить на случай MVF. Действительно, из (10), принимая во внимание (6) и (7), мы получаем следующий фильтр для глобальной коррекции частот:

$$\hat{Y}(i,j) = P \left[c + (G_1 + G_2)Y(i,j) + (G_3 - G_2) \sum_{Y(k,l) \in R_{ij}} Y(k,l) \right], \quad (11)$$

и следующий фильтр для коррекции только высоких частот.

$$\hat{Y}(i,j) = P \left[c + (1+G)Y(i,j) - G \sum_{Y(k,l) \in R_{ij}} Y(k,l) \right], \quad (12)$$

где R_{ij} - локальное окно вокруг пиксела $Y(i,j)$; $Y(i,j)$ и $\hat{Y}(i,j)$ являются значениями яркости в (i,j)-том пикселе до и после обработки соответственно, коэффициенты G_1, G_3 определяют коррекцию низких частот, G_2 определяет коррекцию высоких частот, G - корректирующий коэффициент (в (12)), c - константа. Из (11) и (12) легко получить наборы весов для их реализации на CNN-MVN. Фильтр, определяемый соотношением (11) может быть реализован на CNN-MVN при помощи следующего набора весов:

$$W_0 = C; W = \begin{pmatrix} G_3 - G_2 & \dots & \dots & G_3 - G_2 \\ \cdot & & & \cdot \\ \cdot & \dots & G_1 + G_3 & \dots \\ \cdot & & & \cdot \\ G_3 - G_2 & \dots & & G_3 - G_2 \end{pmatrix} \quad (13)$$

Чтобы усилить средние пространственные частоты и соответственно выделить детали приблизительно размера $n \times m$, апертура фильтра и здесь должна быть в точности равна $n \times m$. Таким образом, матрица W должна быть в точности той же размерности. Рекомендуемые значения весовых параметров следующие: $0 \leq G_1, G_3 \leq 1$, $5 \leq G_2 \leq 14$. Фильтр, определяемый соотношением (12), может быть реализован следующим набором весов размерности (3 x 3) (апертура (3 x 3) - наиболее эффективная для данного фильтра):

$$W_0 = C; W = \begin{pmatrix} -0.5 & -0.5 & -0.5 \\ -0.5 & w_{22} & -0.5 \\ -0.5 & -0.5 & -0.5 \end{pmatrix}, \quad (14)$$

где C - константа (может быть равна нулю), w_{22} - параметр. Рекомендуемые значения для веса w_{22} следующие: $6 \leq w_{22} \leq 14$. $w_{22}=6$ соответствует более жесткому и $w_{22}=14$ соответствует более мягкому усилению высоких частот. Замечание, касающееся реализации. Функция (2) не определена в нуле, однако взвешенная сумма в (7) и (8), хотя и с вероятностью близкой к нулю, но может быть равной нулю. В этом случае целесообразно использовать вес w_0 для избежания подобной ситуации. Отметим также, что во избежание инверсии "белое" - "черное" (0-255) k в (2а-б) можно взять чуть больше, чем 256, например 264-280, однако подобная ситуация также - редкий случай. Кроме того, для избежания той же инверсии можно пользоваться дополнительной проверкой: если $|B_{ij} - \hat{B}_{ij}| \geq k/2$, то оставляем $\hat{B}_{ij} = B_{ij}$, т.е., в этом случае яркость в рассматриваемой точке не изменится. Но принципиаль-

ной необходимости в этом нет, поскольку это также редчайший случай. Преимущества коррекции частот при помощи MVF иллюстрируются на Рис. 5.

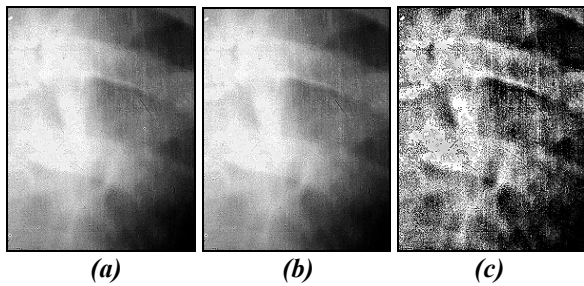


Рис. 5 *Коррекция частот (выделение деталей): (a) оригинал; (b) коррекция при помощи MVF (фильтр (12), набор весов (14)); (c) - коррекция при помощи линейного фильтра (10).*

Главное преимущество MVF-коррекции (по сравнению с линейным фильтром, определяемым (10)) является сохранение общего динамического диапазона наряду с одновременным усилением локальных контрастов. Линейный фильтр (10), являясь отличным аппроксиматором соответствующего фильтра в частотной области, и, имея вычислительную простоту, тем не менее не лишен недостатка: линейность предопределяет недостаточную коррекцию, либо (и чаще всего) наоборот, слишком сильную, в результате чего происходит смещение динамического диапазона изображения к краю (белому, либо черному) и какие-то важные мелкие детали могут остаться невидимыми. На Рис. 5a представлен фрагмент рентгеновского медицинского изображения (опухоль легкого) крайне низкого качества. Чтобы попытаться рассмотреть структуру опухоли (является ли она гомогенной или нет), а также выделить ее детали, мы использовали набор весов (13) (апертура 25×25 , $G_1 = 0.25$; $G_2 = 10.0$; $G_3 = 0.4$). Результат обработки представлен на Рис. 5b (неоднородная структура опухоли, ее детали достаточно отчетливо видны). Рис. 5c содержит для сравнения результат обработки исходного изображения с Рис. 5a. Этот результат - лучшее, чего можно было добиться линейным фильтром (10), однако детали на данном изображении не выделены настолько отчетливо, как на изображении с Рис. 5b, являющегося результатом MVF-коррекции.

3.3. Алгоритмы обработки, описываемые не пороговыми булевыми функциями и их CNN-UBN реализация

В [9, 13] были рассмотрены алгоритмы выделения контуров и фильтрации импульсных помех, а также их реализация на CNN-UBN. Здесь хотелось бы развить идеи, представленные в указанных работах.

Булева непороговая функция, определяемая соотношением (16a) описывает процедуру выделения “восходящих” перепадов яркости и соответствующих им контуров. Алгоритм обучения (3), примененный к этой функции ($m=4$ в (2c)) дает набор весов (16b) для реализации функции на CNN-UBN.

Обработка полутонового изображения производится путем отдельной обработки битовых плоскостей и их последующего объединения в результирующее изображение. На Рис. 6b представлен результат выделения контуров при помощи функции (16a) на исходном изображении (Рис. 6a).

Для сравнения на Рис. 6c приводится результат выделения контуров на том же исходном изображении при помощи оператора Лапласа. Булева непороговая функция (17a) выделяет на бинарном изображении “нисходящие” перепады яркостей и соответствующие им контуры.

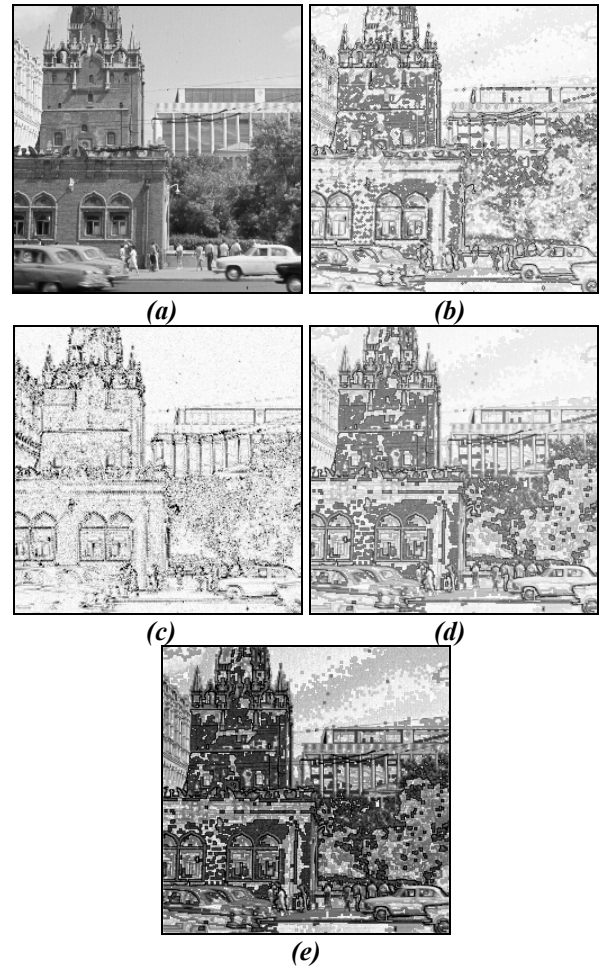


Рис. 6 *Выделение контуров: (a) оригинал; (b) обработка по функции (16a); (c) оператор Лапласа; (d) обработка по функции (17a); (e) контурное изображение (d) с усиленным контрастом*

(a)

$$Y \begin{bmatrix} x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 \\ x_7 & x_8 & x_9 \end{bmatrix} = x_5 \& (\bar{x}_1 \vee \bar{x}_2 \vee \bar{x}_3 \vee \bar{x}_4 \vee \bar{x}_6 \vee \bar{x}_7 \vee \bar{x}_8 \vee \bar{x}_9);$$

(b) $W = (6.0, 17.0); \begin{bmatrix} (-3.23, 3.64) & (-7.83, -0.02) & (-2.50, -0.13) \\ (-6.94, 3.34) & (35.06, 0.66) & (-6.50, 2.13) \\ (-1.83, 0.02) & (-7.23, -3.64) & (-1.00, 1.00) \end{bmatrix}$ (16)

Алгоритм обучения (3) дает набор весов (17b) для реализации функции (17a) на клеточной сети из универсальных элементов.

$$(a) \quad Y \begin{bmatrix} x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 \\ x_7 & x_8 & x_9 \end{bmatrix} = \bar{x}_5 \& (x_1 \vee x_2 \vee x_3 \vee x_4 \vee x_6 \vee x_7 \vee x_8 \vee x_9);$$

$$(b) \quad W = (-1.7, -15); \begin{bmatrix} (-0.91, 0.74) & (0.78, -1.0) & (-0.1, -0.04) \\ (-0.06, -0.08) & (1.4, 1.6) & (0.86, -0.96) \\ (-0.15, -0.07) & (-0.06, -0.05) & (-0.14, -0.03) \end{bmatrix} \quad (17)$$

4. Распознавание и интерполяция изображений путем анализа ортогонального спектра на сети из многозначных нейронов

4.1. Распознавание изображений

Решение задачи распознавания изображений на нейросетях становится все более популярным в последние годы. В работах [8-10] были предложены различные сети из многозначных нейронов для решения подобной задачи. Общий недостаток предложенных в данных работах сетей и подходов к решению проблемы состоит в слишком тесной зависимости качества распознавания от обучающего множества, а также значительными трудностями при распознавании сдвинутых и повернутых изображений, а также изображений с измененным динамическим диапазоном яркостей. Эти недостатки вызваны тем, что обучение и распознавание происходит путем анализа изображений непосредственно в пространственной области. В то же самое время хорошо известен подход к распознаванию, основанный на анализе коэффициентов низкочастотной части ортогонального спектра [17]. Мы хотим предложить здесь новую нейросеть из MVN, ориентированную на решение задачи распознавания изображений путем анализа коэффициентов низкочастотной части ортогонального спектра. В качестве ортогонального преобразования использовалось преобразование Уолша, упорядоченное по Уолшу. Рассмотрим проблему распознавания портретов сорока людей (мы использовали в точности ту же базу данных, сто и в предшествующих работах [8-9]). Обучающее множество возьмем намеренно минимальной мощности - по одному представителю для каждого класса (лица). На Рис. 7 представлен фрагмент обучающего множества.

Для распознавания следует найти подходящее число спектральных коэффициентов (входов нейрона) для проведения обучения с последующим предъявлением сети для распознавания объектов, не входивших в обучающую выборку. Учитывая свойства многозначных нейронов, следующая структура нейросети является наиболее подходящей для решения нашей задачи. Это - одноуровневая нейросеть, содержащая столько же нейронов, сколько классов объектов (в данном случае, сколько разных лиц) мы собираемся распознавать. Каждый нейрон должен распознавать объекты из соответствующего

класса как "свой", объекты же из всех остальных классов - как "чужие".



Рис. 7. Фрагмент обучающей выборки.

Наилучшие результаты были получены, когда первые $k/4$ (из k) секторов на комплексной плоскости (см. (2a), (2b) и Рис. 1) зарезервированы для "своих" для данного нейрона объектов, а остальные $3/4$ секторов - для "чужих" объектов (Рис. 8). Таким образом, выходные значения $0 \dots k/4 - 1$ для i -го нейрона соответствуют классификации объекта, как принадлежащем i -му классу. Выходные значения $k/4 \dots k - 1$ соответствуют классификации объекта как "чужого" для данного нейрона. Для проверки эффективности предложенной сети и подхода к решению задачи мы намеренно создали экстремальную ситуацию, включив в обучающую выборку лишь по одному представителю каждого класса. Обучающее множество состояло соответственно из спектров Уолша изображений из обучающей выборки. Для поиска оптимального числа спектральных коэффициентов (входа нейрона), мы брали для анализа коэффициенты с первых нескольких диагоналей двумерной матрицы спектра (ибо на данной диагонали все коэффициенты соответствуют одинаковой частоте, каждая следующая диагональ соответствует более высокой частоте) за исключением нулевой диагонали (нулевого коэффициента), слишком тесно связанной с данным конкретным объектом. Оптимальные результаты были получены при $k=512$ в (2a-b) (соответственно при нормализации спектральных коэффициентов к диапазону $[0 \dots 511]$). Для обучения сети каждый нейрон обучался в соответствии с алгоритмом (4). Секторы $0 \dots 127$ были зарезервированы для "своих", а секторы $128 \dots 511$ - для "чужих" для данного нейрона образов. Сходимость обучения для каждого нейрона наступала примерно после 1000 итераций, что занимает несколько секунд на компьютере класса Pentium-133.

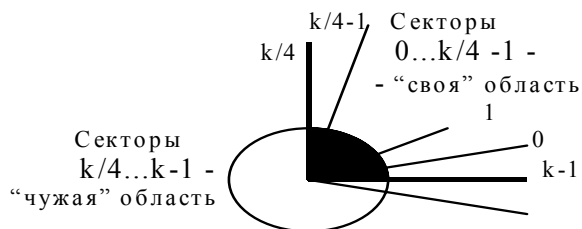
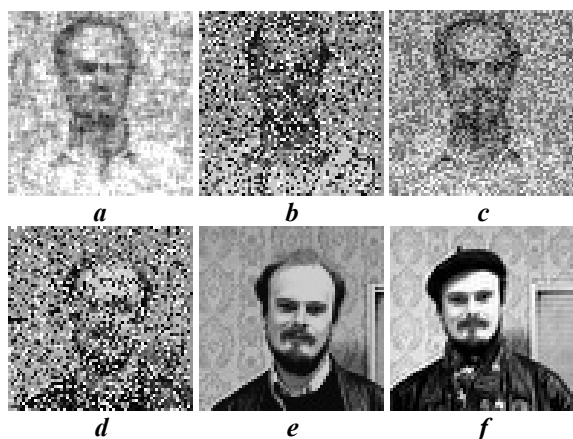


Рис. 8. Распределение секторов для обучения и распознавания.

Для распознавания серия изображений, не представленных в обучающей выборке, но являющихся другими, либо искаженными портретами тех же людей, была предъявлена для распознавания. Результаты более, чем положительные. На Рис. 9 представлены 12 изображений из класса "Игорь", которые не были представлены в обучающей выборке, но были успешно распознаны. Таблица 2 показывает, что при распознавании имелся достаточный запас прочности, страхующий от ошибок.

Изображения, распознанные по 17 коэффициентам:



Изображения, распознанные по 35 коэффициентам:

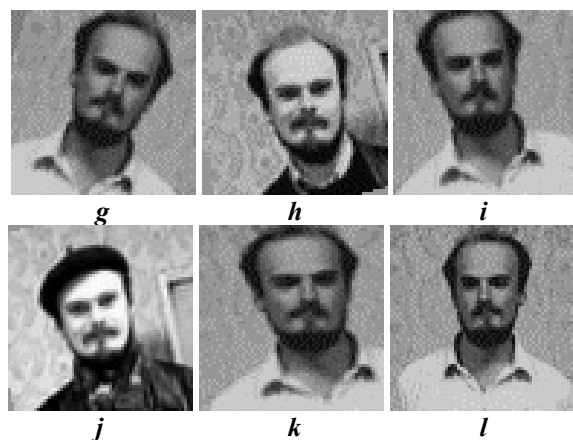


Рис. 9. Изображения, не представленные в обучающей выборке, но успешно распознанные.

Таблица 2.
Номера секторов, в которые попала взвешенная сумма при распознавании изображений (см. Рис. 9).

Image	a	b	c	d	e	f
Sector	60	62	62	64	65	75

Image	g	h	i	j	k	l
Sector	40	34	45	99	35	46

4.2. Решение задачи интерполяции изображения через экстраполяцию спектра с предсказанием спектральных коэффициентов при помощи многозначного нейроэлемента.

Пусть задан сигнал f и его дискретные значения $f_1, f_2, \dots, f_i, f_j, \dots, f_N$, взятые с некоторым шагом дискретизации. Задача состоит в вычислении значений f_{i+s} , $i = 1, \dots, N$; $i+s < j$; $s = 1/p, 2/p, \dots, (p-1)/p$; $p \in \{2, 4, 8, 16, \dots\}$. Например, $n=4$, и нужно получить значения $f_{i+0.25}, f_{i+0.5}, f_{i+0.75}$ между f_i и f_j . Традиционный подход к решению этой задачи - интерполяция в пространственно-временной области. Иной подход состоит в экстраполяции дискретного спектра сигнала в область более высоких частот с последующим вычислением обратного преобразования. Решение этой задачи обычно представляет весьма серьезную вычислительную проблему. Мы хотим предложить подход к ее решению, основанный на возможности экстраполировать временные ряды на единственном многозначном нейроне [8]. Таким образом, решение задачи сведется к следующему алгоритму: 1) Вычисление спектра (Фурье, косинусного, либо Уолша) сигнала f : $S_f = (s_1, \dots, s_N)$; 2) Предполагая, что имеет место зависимость $s_i = g(s_{i-q}, \dots, s_{i-1})$ (каждый спектральный коэффициент - есть функция от q ему предшествующих по частоте), можно обучить MVN реализации функции g ; 3) Экстраполяция спектра $\tilde{S}_f = (s_1, \dots, s_N, \tilde{s}_{N+1}, \dots, \tilde{s}_{pN})$ - предсказание с помощью MVN коэффициентов, соответствующих более высоким частотам; 4) Вычисление обратного преобразования размерности pN , что и позволяет получить необходимые значения сигнала f . Эффективность предложенного подхода иллюстрируется на Рис. 10. MVN был обучен по алгоритму (4) представлять каждый коэффициент косинусного спектра (нормализованного к диапазону $[0, 8191]$) изображения "Lena" как функцию от 63 предшествующих ему по частоте коэффициентов. Затем спектр был "увеличен" в 2 раза путем предсказания с помощью полученных весов значений коэффициентов в области более высоких частот. С помощью обратного преобразования было получено изображение, представленное на Рис. 10b. Для сравнения на Рис. 10c представлено изображение, полученное из того же оригинала путем билинейной интерпо-

ляции. Многие детали на нем “смазаны”, резкость частично потеряна по сравнению с изображением, полученным путем экстраполяции спектра.



Рис. 10 Экстраполяция спектра на MVN: (а) оригинал 256x256; (б) изображение 512x512, полученное из экстраполированного спектра оригинала; (с) изображение 512x512, полученное билинейной интерполяцией.

Благодарности

Данная работа частично поддержана грантом “Belgian program on interuniversity attraction poles (IUP-02) initiated by the Belgian State, Prime Minister’s Office for Science, Technology and Culture (DWTC).” Автор выражает также признательность Н.С. Мерзлякову за обсуждение вопросов, рассмотренных в статье.

Литература

1. L.O. Chua and L. Yang, "Cellular neural networks: Theory", IEEE Trans. on Circuits and Systems, vol. 35, pp.1257-1290, 1988.
2. S.Tan, J.Hao, J.Vandewalle “Cellular neural networks as a model of associative memories”, Proc. of the 1990 IEEE International Workshop on CNN and their applications (CNNA-90), Budapest, pp. 23-26, 1990.
3. H. Harrer and J.A. Nossek “Discrete-time cellular neural networks”, Int. J. Circuit Theory and Appl., vol.20, pp. 453-467, 1992.
4. N.N. Aizenberg, I.N. Aizenberg "Fast converged learning algorithms for multi-level and universal binary neurons and solving of the some image processing problems", Lect. Notes in Computer Science, Ed.-J. Mira, J. Cabestany, vol. 686, Shpringer, pp. 230-236, 1993.

5. N.N.Aizenberg, I.N.Aizenberg "CNN-like networks based on multi-valued and universal binary neurons: learning and application to image processing", Proc. of the Third IEEE International Workshop on Cellular Neural Networks and their Applications, Rome, pp.153-158, 1994
6. C. Rekeczky, T. Roska and A. Ushida “CNN Based Self-Adjusting Nonlinear Filters” Proc. of the Fourth IEEE International Workshop on Cellular Neural Networks and their Applications, Seville, Spain, pp. 309-314, 1996.
7. Aizenberg I.N. “Multi-valued non-linear filters and their implementation on cellular neural networks”, Frontiers in Artificial Intelligence and Applications, vol. 41 “Advances in Intelligent Systems” (F.C. Morabito - Ed.) IOS Press, Amsterdam-Berlin-Oxford-Tokyo-Washington DC, 1997, pp. 135-140.
8. N.N. Aizenberg, I.N. Aizenberg, G.A. Krivosheev "Multi-Valued Neurons: Mathematical model, Networks, Application to Pattern Recognition", Proc. of the 13 Int.Conf. on Pattern Recognition, Vienna, August 25-30, 1996, Track D, IEEE Computer Soc. Press, pp. 185-189, 1996.
9. Н.Н. Айзенберг, И.Н. Айзенберг, Г.А. Кривошеев. Нейросети на многозначных нейроэлементах: обучение и применение в обработке и распознавании изображений // Компьютерная оптика, 1995, Том 14-15, часть 1, стр. 179-186.
10. S.Jankowski, A.Loowski, M.Zurada “Complex-Valued Multistate Neural Associative Memory”, IEEE Trans. on Neural Networks, vol.7, No 6, November, pp.1491-1496, 1996.
11. Н.Н. Айзенберг, Ю.Л. Иваськив Многозначная пороговая логика, Киев: Наукова думка, 1977.
12. S.Haykin Neural networks. A comprehensive foundation. Macmillan College Publishing Company, New York, 1994.
13. I.N. Aizenberg “Processing of noisy and small-detailed gray-scale images using cellular neural networks” Journal of Electronic Imaging, 6(3), 1997, pp. 272- 285.
14. A.C. Bovik, T.S. Huang, D.C. Mudson A generalization of median filtering using linear combinations of order statistics’, IEEE Trans. Accoust., Speech, Signal Proc., vol.31, pp.1342-1349, 1983.
15. Т.П. Беликова. Моделирование линейных фильтров в задачах медицинской диагностики: в кн. "Цифровые методы обработки изображений и полей в экспериментальных исследованиях", Москва: Наука, 1990, стр. 130-152.
16. N.N. Aizenberg, I.N. Aizenberg, T.P. Belikova Extraction and localization of important features on gray-scale images: implementation on the CNN Proc. of the Third IEEE International Workshop on Cellular Neural Networks and their Applications (CNNA-94), Rome, pp. 207-212, 1994.
17. N. Ahmed, K.R. Rao Orthogonal Transforms for Digital Signal Processing, Springer 1975.