УДК 004.021-004.8
DOI: 10.25559/SITITO.17.202103.541-549

Научная статья

# Многомерно-матричное определение операции свертки

**Е. И. Гончаров**
ФГБОУ ВО «Смоленский государственный университет», г. Смоленск, Российская Федерация
214000, Российская Федерация, г. Смоленск, ул. Пржевальского, д. 4
drbenvey1996@mail.ru

## Аннотация

Свертка – незаменимая операция для решения задачи из самых разных предметных областей: задач машинного обучения, анализа данных, обработки сигналов, фильтров обработки изображений. Целесообразно рассмотреть разные подходы и имеющиеся методы реализации операции свертки. Из-за сложности реализующих их алгоритмов на практике даже трехмерные свертки используются значительно реже, чем одно- и двумерные. Основная причина этого кроется в отсутствии единого строгого определения операции и перегруженности в математике термина "свертки". Алгебра многомерных матриц имеет в себе операции, семантически схожие со свертками и легко распараллеливается с помощью естественных обобщений операций на плоских матрицах на многомерный случай. Она уже доказала свою эффективность в тензорной алгебре. Поэтому целесообразно сконструировать определение свертки через операции на алгебре многомерных матриц. В статье рассмотрен подход к повышению эффективности алгоритмов свертки в программных системах, основанных на ней, в том числе и сверточных нейронных сетях. Автором предлагается многомерно-матричное определение операции свертки. На ее основе строится многомерно-матричная модель вычислений, которая позволят эффективно формализовать задачи, решение которых использует операции многомерной свертки, а также реализовать эффективное решение этих задач благодаря естественному параллелизму, присущему операциям алгебры многомерных матриц. В результате получена математическая модель операций свертки на основе алгебры многомерных матриц с операцией $(0, \mu)$-свернутого произведения. На практике в решении прикладных задач предложенная математическая модель операций свертки служит основой для разработки библиотек программ, эффективно реализующих эти операции за счет распараллеливания операции $(0, \mu)$-свернутого произведения.

**Ключевые слова:** операция свертки, компьютерное зрение, алгебра многомерных матриц, параллельные вычисления, искусственный интеллект

*Автор заявляет об отсутствии конфликта интересов.*

**Для цитирования:** Гончаров, Е. И. Многомерно-матричное определение операции свертки / Е. И. Гончаров. – DOI 10.25559/SITITO.17.202103.541-549 // Современные информационные технологии и ИТ-образование. – 2021. – Т. 17, № 3. – С. 541-549.

Original article

# Multi-Dimensional Definition of Convolution

**E. I. Goncharov**
Smolensk State University, Smolensk, Russian Federation
4 Przhevalsky St., Smolensk 214000, Russian Federation
drbenvey1996@mail.ru

## Abstract

Convolution is an indispensable operation for solving problems from a variety of subject areas: machine learning problems, data analysis, signal processing, image processing filters. However, due to the complexity of the algorithms that implement them, in practice, even three-dimensional convolutions are used much less often than one-and two-dimensional ones. The main reason for this lies in the absence of a single strict definition of the operation and the overload of the term "convolution" in mathematics. The multidimensional matrix algebra includes operations that are semantically similar to convolutions, and it is easily parallelized using natural generalizations of operations on flat matrices to the multidimensional case and has already proved its effectiveness in tensor algebra. Therefore, it is advisable to construct the definition of convolution through operations on the algebra of multidimensional matrices. The article considers an approach to improving the efficiency of convolution algorithms in software systems based on it, including convolutional neural networks. The author proposes a multidimensional matrix definition of the convolution operation. On its basis, a multidimensional matrix model of calculations is built, which will allow us to effectively formalize problems whose solution uses multidimensional convolution operations, as well as to implement an effective solution to these problems due to the natural parallelism inherent in the operations of the algebra of multidimensional matrices. As a result, a mathematical model of convolution operations based on the algebra of multidimensional matrices with the operation the $(0, \mu)$-convolution product is obtained. In practice, in solving applied problems, the proposed mathematical model of convolution operations serves as the basis for developing libraries of programs that effectively implement these operations by parallelizing the operation the $(0, \mu)$-convolution product.

**Keywords:** convolution, computer vision, multidimensional matrix algebra, parallel computing, artificial intelligence

*The author declares no conflict of interest.*

**For citation:** Goncharov E.I. Multi-Dimensional Definition of Convolution. *Sovremennye informacionnye tehnologii i IT-obrazovanie = Modern Information Technologies and IT-Education.* 2021; 17(3):541-549. DOI: https://doi.org/10.25559/SITITO.17.202103.541-549

## Введение

Термин свертка в математике сильно перегружен. В разных предметных областях есть операции называемые “сверткой” или имеющие это слово в названии: свертка на группах, свертка распределений, ($\lambda$, $\mu$)-свернутое произведения, $\mu$ – свертывание матрицы, свёртка тензора [1]. Свертки, используемые в задачах компьютерного зрения, возможно, наиболее популярны. С их помощью можно решать задачи анализа и прогнозирования временных рядов, обработки естественного языка, анализа сигналов на определенном временном промежутке, обработки изображений (фильтры для выделения краев, фрагментов, размытия и других операций преобразования для дальнейшего использования изображения в задаче распознавания).

## Цель исследования

Автор ставить целью исследования анализ имеющихся определений операции свертки и формирование определения этой операции в терминах алгебры многомерных матриц. Оно должно обеспечивает строгую формализацию этой операции на базе алгебры многомерных матриц и одновременно упрощать и снижать стоимость разработки подобных информационных систем, а также сокращать время выполнения запросов.

## Теоретический анализ формальных определений

В функциональном анализе операция свертки для двух функций $f$ и $g$ интегрируемых относительно меры Лебега определяется следующим образом: $(f * g)(x) \stackrel{\text{def}}{=} \int_{\mathbb{R}^n} f(y)g(x-y)dy$. Результат свертки соответствует взаимно-корреляционной функции для этих двух функций[1]. Здесь операцию свёртки можно интерпретировать как «схожесть» одной функции с отражённым и сдвинутым образом другой.

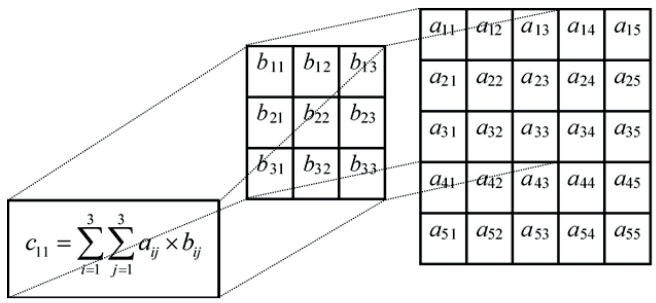
В силу того, что наш мир дискретен, в прикладных задачах, как правило, фигурируют дискретные модели и функции. Так цифровой обработке сигналов широко используется $n$-мерная дискретная свертка сигналов. Ее также называют цифровой фильтрацией[2] [2], [3]. Эта операция обозначается $Y = H\underbrace{**\ldots*}_{n}U$ и определяется следующим образом: пусть $U(u_{l_1},\ldots,u_{l_n})$ и $H(h_{p_1},\ldots,h_{p_n})$ два дискретных $n$-мерных сигнала (отсчета исходного сигнала и отсчета импульсной характеристики фильтра). Результатом их свертки будет выходной сигнал[3] $Y(y_{m_1},\ldots,y_{m_n})$, где $Y(y_{m_1},\ldots,y_{m_n}) = \underbrace{\sum_{k_1}\ldots\sum_{k_n}}_{n} U(u_{k_1},\ldots,u_{k_n})H(h_{m_1-k_1},\ldots,h_{m_n-k_n})$ [4].

В задачах компьютерного зрения под одноканальной многомерной сверткой по ядру понимают операцию, определенную следующим образом: пусть $A$ – входной многомерный массив размерности $\underbrace{n\times\ldots\times n}_{N}$, который называется сигналом, а $B$ – многомерный массив размерности $\underbrace{m\times\ldots\times m}_{N}$, который называется ядром. Свертка $A$ и $B$ определяется следующим образом: $(A\times B)(i_1,\ldots,i_N) = \sum_{j_1=1}^{m}\ldots\sum_{j_N=1}^{m} A(i_1+j_1-1,\ldots,i_N+j_N-1)\times B(j_1,\ldots,j_N)$, где $i_k = 1,\ldots,n+1-m$; $k = 1,\ldots,N$ (*) [5].

Вообще говоря, в задачах компьютерного зрения в операциях свертки в смысле определения присуще (*) участвуют много параметров таких как: размерность, отступы, шаги, линейное смещение и так далее [6]. Помимо свертки по ядру производятся свертки по фильтрам. Тем не менее, смысл операции и особенности реализующих ее алгоритмов от этих параметров меняются незначительно. В дальнейшем, без ограничения общности, предполагается, что свертка выполняется без нулевого заполнения (Zero Padding), с шагом равным 1 и без линейного смещения.

## Теоретический анализ интуитивного “определения”

Часто можно встретить неформальное описание операции свертки. Ниже приводится пример такого описания. Ядро “скользит” по входному сигналу с заданным шагом [7]. Размерность свертки определяет число измерений для “скольжения”. На каждом шаге поэлементно умножается содержимое ядра на “окно”, ему соответствующего. Результат суммируется и заносится в соответствующую позицию матрицы результата. На рисунке 1 приведен пример одноканальная двумерня свертка матрицы $A$ по ядру $B$.




Р и с. 1. Одноканальная двумерня свертка матрицы $A$ по ядру $B$

F i g. 1. Single-channel two-dimensional convolution of matrix $A$ by kernel $B$

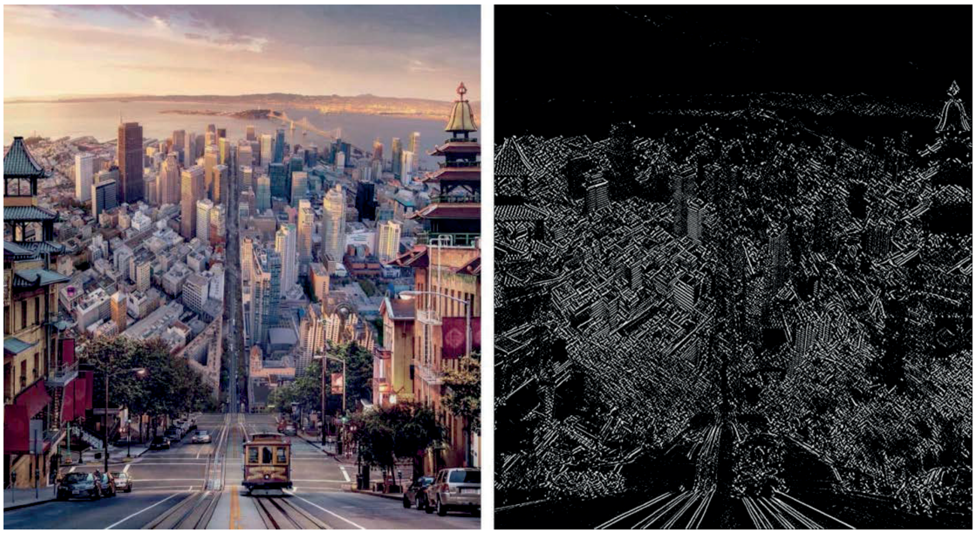
Такие неформальные определения дают понятную визуализацию операции, из которой можно понять смысл работы некоторых фильтров. Для примера, рассмотрим результат применения горизонтального и вертикального ядра Собеля: $\begin{pmatrix}-1 & 0 & 1\\ -2 & 0 & 2\\ -1 & 0 & 1\end{pmatrix}, \begin{pmatrix}-1 & -2 & -1\\ 0 & 0 & 0\\ 1 & 2 & 1\end{pmatrix}$ — в задаче выделения контура изображений [8], [9]. На рисунке 2 приведен оригинал изображения и результат его обработки фильтром Собеля. Для участка

[1] Колмогоров А. Н., Фомин С. В. Элементы теории функций и функционального анализа. 7-е изд. М.: Физматлит, 2004. 572 с.

[2] Рабинер Л. Теория линейных дискретных систем / Рабинер Л., Гоулд Б. // Теория и применение цифровой обработки сигналов; под ред. Ю. Н. Александрова. М.: Мир, 1978. С. 18-88.

[3] Гутников В. С. Фильтрация измерительных сигналов. Л.: Энергоатомиздат, Ленингр. отд-ние, 1990. 192 с.

изображения, не содержащего в себе границ объектов, пиксели имеют одинаковое значение, поэтому общий вывод ядра в этой точке равен 0. На фрагментах с вертикальными (горизонтальными) гранями существует разница между пикселями слева и справа (сверху и снизу) от грани, и ядро вычисляет эту ненулевую разницу, находя ребра [10]. Таким образом, другое определение позволяет изучить эту операцию с другой стороны, сделав очевидными некоторые ее особенности. Тем не менее, это "определение" не обладает достаточной математической строгостью, а при переходе к большему числу измерений теряется визуализация.



Р и с. 2. Оригинальная картинка и пример применения оператора Собеля
F i g. 2. The original picture and an example of the application of the Sobel operator

## Теоритический анализ многомерно-матричного определения

Это определение строится на алгебре многомерных матриц, введенной математиком Н. П. Соколовым[4]. Ниже приведены необходимые определения базовых элементов алгебры многомерных матриц по Соколову.

Многомерная матрица определяется как система элементов $A_{i_1 i_2 \ldots i_p}$, расположенных в точках p-мерного пространства и определяемых координатами $(i_1, i_2, \ldots, i_p)$, где $i_a = 1, 2, \ldots, N_{i_\alpha}$; $a = 1, 2, \ldots p$; $N_{i_\alpha} \in \mathbb{N}$ [11]. Многомерные матрицы принято записывать с помощью двухмерных сечений в виде квадратной или прямоугольной таблицы. Например, представление четырехмерной матрицы A порядка 2 имеет такой вид

$$A = \left\| \begin{array}{cc|cc} a_{1111} & a_{1112} & a_{1211} & a_{1212} \\ a_{1121} & a_{1122} & a_{1221} & a_{1222} \\ \hline a_{2111} & a_{2112} & a_{2211} & a_{2212} \\ a_{2121} & a_{2122} & a_{2221} & a_{2222} \end{array} \right\| .$$

Пусть матрицы $A = \left\| A_{i_1 \ldots i_p} \right\|$ и $B = \left\| B_{i_1 \ldots i_q} \right\|$ p- и q-мерные соответственно. Совокупности индексов этих матриц $i_1, \ldots, i_p$ и $i_1, \ldots, i_q$ разбиваются на три группы, содержащие соответственно. Совокупности индексов этих матриц $i_1, \ldots, i_p$ и $i_1, \ldots, i_q$ разбиваются на три группы, содержащие соответственно к, $\mu$ и $\nu$ индексов (к, $\mu$, $\nu \geq 0$). Причем к+$\mu$=p, а $\mu$+$\nu$=q. Для полученных групп индексов используются обозначения: $l = (l_1, \ldots, l_\kappa)$, $c = (c_1, \ldots, c_\mu)$ и $m = (m_1, \ldots, m_\nu)$. Тогда матрицы A и B можно представить в виде $A = \left\| A_{lc} \right\|$ и $B = \left\| B_{cm} \right\|$. Индексы разбиения c называются кэлиевыми, а индексы разбиения m и l — свободными. Матрица $C = \left\| C_{lm} \right\|$, элементы которой вычисляются по формуле $C_{lm} = \sum_{(c)} A_{lc} \times B_{cm}$, называется (0, μ)-свернутым произведением матриц A и B и обозначается ${}^{0,\mu}(A \times B) = C$ (**). Из его определения следует, что при изменении значения μ результатом (0, μ)-свернутого произведения многомерных матриц A и B могут получаться многомерные матрицы разных размерностей. Вычислительная сложность такой операции $O(N^{\mu+\kappa+\nu})$.

Для определения свертки средствами алгебры многомерных матриц авторы предлагает строить вспомогательную $A'$ [12]. Она состоит из элементов $A$, расположенных в дополнительных размерностях так, что бы ее двухмерные сечения содержали элементы, которые сворачиваются ядром при каждом шаге свертки. Матрица $A'$ имеет вдвое больше измерений, чем матрица A. Это следует из того, что в матрице A' каждый элемент расположен в p-мерном пространстве и сам представляет собой p-мерный блок матрицы A. Авторы доказывают утверждение:

«Пусть $A = \left\| a_{i_1 \ldots i_p} \right\|$ $(i_1, \ldots, i_p = 1, \ldots, N_A)$ — свертываемая матрица. $B = \left\| b_{j_1 \ldots j_p} \right\|$ $(j_1, \ldots, j_p = 1, \ldots, N_B)$ — ядро свертки. Тогда одноканальная p-мерная свертка матрицы A по фильтру B это ${}^{0,p}(A' \times B) = \sum_{(c)} a'_{lc} \times b_c = \left\| c_l \right\|$, где $A' = \left\| a'_{lc} \right\|$, $l_1, \ldots, l_p = 1, \ldots, N_{A'}$ и

[4] Соколов Н. П. Введение в теорию многомерных матриц. Киев: Наукова Думка, 1972. 177 с.

$c_1, \dots, c_p = 1, \dots, N_B$. Пусть в самом общем случае $A = \left\| a_{i_1 \dots i_p i_{p+1}} \right\|$ — свертываемая матрица, где $i_1, \dots i_p = 1, \dots, N_A$ — индексы матриц каналов, а значения индекса $i_{p+1} = 1, \dots, n$ — номера каналов. $B = \left\| b_{j_1 \dots j_p j_{p+1}} \right\|$ — фильтр, где $j_1 = 1, \dots, n$ — номер ядра свертки, а $j_2, \dots, j_{p+1} = 1, \dots, N_B$ — индексы ядер свертки.

Таким образом, доказывается, что посредством операции (0, μ)-свернутого произведения могут быть выражены всевозможные свертки $A$ по ядру (фильтру) B» [12].

Например, рассмотрим одноканальную двумерую свертку матрицы A по ядру B (см. рисунок 1). Пусть даны входная матрица $A = \begin{Vmatrix} a_{11} & \dots & a_{15} \\ & \dots & \\ a_{51} & \dots & a_{55} \end{Vmatrix}$ и ядро свертки $B = \begin{Vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{Vmatrix}$. В этом случае матрица $A'$ представляет собой четырехмерную матрицу

$$A' = \left\| \begin{array}{ccc|c|ccc} a'_{1111} & \dots & a'_{1113} & & a'_{1311} & \dots & a'_{1311} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a'_{1131} & \dots & a'_{1133} & & a'_{1331} & \dots & a'_{1333} \\ \hline & \dots & & \dots & & \dots & \\ \hline a'_{3111} & \dots & a'_{3113} & & a'_{3311} & \dots & a'_{3313} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a'_{3131} & \dots & a'_{3133} & & a'_{3331} & \dots & a'_{3333} \end{array} \right\| = \left\| \begin{array}{ccc|c|ccc} a_{11} & \dots & a_{13} & & a_{13} & \dots & a_{15} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{33} & \dots & a_{33} & & a_{33} & \dots & a_{35} \\ \hline & \dots & & \dots & & \dots & \\ \hline a_{31} & \dots & a_{33} & & a_{33} & \dots & a_{35} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{51} & \dots & a_{53} & & a_{53} & \dots & a_{55} \end{array} \right\| = \left\| A'_{lc} \right\|,$$

$l_1, l_2, c_1, c_2 = 1, \dots, 3$.

Тогда результатом выполнения свертки A по ядру B с шагом 1 будет совпадать с ${}^{0,2}(A' \times B)$, где ${}^{0,2}(A' \times B) = \sum_{(c)} a'_{lc} \times b_c = \left\| c_l \right\|$.

И как показано на рисунке 1

$$c_{11} = \sum_{(c)} a'_{11c} \times b_c = \sum_{c_1=1}^{3} \sum_{c_2=1}^{3} a'_{11c_1c_2} \times b_{A_1A_2} = \sum_{i=1}^{3} \sum_{j=1}^{3} a_{ij} \times b_{ij}.$$

Из утверждения прямо следует, что посредством операции (0, μ)-свернутое произведение могут быть выражены всевозможные свертки по ядру (фильтру) и свертки сигналов. Все виды сверток имеют однотипное формальное представление, легко адаптируемое к каждому конкретному случаю за счет построения $A'$. Таким образом, можно дать определение свертки через (0, μ)-свернутое произведение. Отметим, что в таком подходе ядро не «скользит» по направлениям, а «прыгает» по двухмерным сечениям. Такое определение в равной степени как позволяет визуально представить операцию, так и имеет достаточную строгость.

## Имеющиеся методы выполнения сверток

Ввиду популярности операции свертки на сегодняшний день существует огромное количество библиотек с методами, его реализующими: cv2.filter2D, tensorflow. compat. v1.conv2d, scipy. sepfir2d, numpy. convolve [13, 14, 15, 16]. В таблице 1 приведена зависимость времени выполнения операции двухмерной одноканальной свертки по ядру $5 \times 5$ (средствами tensorflow. compat. v1.conv2d в среде Colab) от числа пикселей в изображении. В таблице 2 приведена зависимость времени выполнения операции двухмерной трехканальной свертки по фильтру $3 \times 5 \times 5$ (средствами cv2.filter2D в среде Colab) от числа пикселей в изображении.

Т а б л и ц а 1. **Время для свертки с tensorflow. compat. v1.conv2d**

T a b l e 1. **Time to convolve with tensorflow.compat.v1.conv2d**

| Число пикселей, млн. | Время, с. |
|---|---|
| | CPU |
| 0,92 | 0,471 |
| 2,07 | 0,911 |
| 3,69 | 1,483 |
| 8,29 | 2,909 |
| 16,38 | 6,445 |
| 55,06 | 15,506 |

Т а б л и ц а 2. **Время для свертки с cv2.filter2D**

T a b l e 2. **Time to convolve with cv2.filter2D**

| Число пикселей, млн. | Время, мс. |
|---|---|
| | CPU |
| 0,92 | 71 |
| 2,07 | 159 |
| 3,69 | 282 |
| 8,29 | 672 |
| 16,38 | 1303 |
| 55,06 | 2551 |

Из результатов видно, что, не смотря на кажущуюся простоту операции, одна свертка может занимать секунды на картинках с большим разрешением (Ultra HD и выше) [17], [18].

Авторы предлагают построение программно аппаратного комплекса, реализующего алгебру многомерных матриц [19]. ($\lambda$, $\mu$)-свернутое произведение многомерных матриц это обобщение умножения плоских матриц. Этот факт обеспечивает возможность без труда обобщить на многомерный случай хорошо известные, доказавшие свою эффективность и простые в понимании и реализации алгоритмы параллельного умножения матриц [20]. (0, $\mu$)-свернутое произведение является частным случаем ($\lambda$, $\mu$)-свернутого произведения, поэтому для него может быть применен тот же параллельных алгоритм. Ниже приводится описание параллельного алгоритма выполнения (0, $\mu$)-свернутого произведения многомерных матриц. За $C_k$ будем обозначать зафиксированный набор индексов $C_k = (c_0^k, \dots, c_{\mu-1}^k)$, всего таких наборов $N_\mu$ и так далее. Тогда для реализации (0, $\mu$)-свернутого произведения необходимо выполнить последовательность вложенных циклов, описанную на псевдокоде ниже (см. рисунок 3).

```
For (i=0; i< N^κ; i++) { // Multiplication of matrices
|   For (k=0; k< N^ν; k++) {
|   |   tmp=0;
|   |   For (t=0; t< N^μ; t++)
|   |   |   tmp=tmp+ A_{L_i C_t} × B_{C_t M_k} ;
|   |   C_{L_i M_k} =tmp;
|   }
}
```

Р и с. 3. Псевдокод (0, $\mu$)-свернутого произведения

F i g. 3. Pseudocode of the (0, $\mu$)-folded product

Таким образом $(0, \mu)$-свернутое произведение многомерных матриц $A$ и $B$ выражается через умножение плоских матриц $A^*$ и $B^*$, полученных из многомерных матриц $A$ и $B$ соответственно, где

$$A^* = \begin{pmatrix} a_{L_0C_0} & \dots & a_{L_0C_{N^\mu-1}} \\ \dots & \dots & \dots \\ a_{L_{N^\kappa-1}C_0} & \dots & a_{L_{N^\kappa-1}C_{N^\mu-1}} \end{pmatrix}, \quad B^* = \begin{pmatrix} b_{C_0M_0} & \dots & b_{C_0M_{N^\nu-1}} \\ \dots & \dots & \dots \\ b_{C_{N^\mu-1}M_0} & \dots & b_{C_{N^\mu-1}M_{N^\nu-1}} \end{pmatrix}.$$

Итак, для выполнения $(0, \mu)$-свернутого произведения матриц $A$ и $B$ необходим CPU, который отправляет соответствующие им плоские матрицы $A^*$ и $B^*$ в видеопамять GPU[5]. Результат их умножения на GPU может быть получен любым удобным параллельным способом. Он пересылается в оперативную память и CPU записывает его в ответ (см. рисунок 4). Для создания такого программно-аппаратного комплекса наиболее эффективны вычислительные системы, обладающие большим количеством GPU [21], [22].

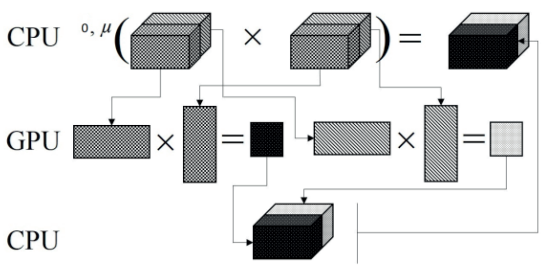



Рис. 4. Схема $(0, \mu)$-свернутого произведения
Fig. 4. Scheme of $(0, \mu)$-folded product

## Полученные результаты

С использованием алгебры многомерных матриц авторам удалось решать задачи из самых разных предметных областей [13, 23, 24, 25]. Был доказан гомоморфизм между алгеброй многомерных матриц и реляционной алгеброй. Это породило новый подход к распараллеливанию запросов в базах данных. С ее помощью были построены полиномиальные алгоритмы для вывода ассоциативных правил и маршрутизации. Так как она является естественным обобщением матричной алгебры, то есть возможность обобщить многие матричные алгоритмы шифрования, тем самым увеличив их криптостойкость. Были обобщены алгоритмы Хилла и Диффи-Хеллмана. Все операции из тензорной алгебры выражается в терминах алгебры многомерных матриц, что позволяет с ее помощью решать все задачи тензорной алгебры. Она удачно подходит для построения цепей Маркова и решения задач на графах. Спроектированный программно аппаратный комплекс помимо решения вышеперечисленных задач можно будет использовать и для построения сверточных нейронных сетей в виду наличия определения операции свертки в терминах алгебры многомерных матриц.

## Заключение

В статье был проведен анализ различных подходов к операции свертки в контексте задач компьютерного зрения. Было введено другое определение этой операции, которое совмещает в себе математическую строгость и интуитивную простоту уже имеющихся. Именно в силу изложенных причин, формальное представление сверток посредством $(0, \mu)$-свернутого произведения многомерных матриц имеет важное практическое значение. Были проанализированы имеющиеся методы реализации сверток. Было предложен аппаратно-программный комплекс, с помощью которого, в том числе, возможно построение сверточных нейронных сетей.

## Список использованных источников

[1] Кисиль, В. В. Об алгебре псевдодифференциальных операторов, порожденной свертками на группе Гейзенберга / В. В. Кисиль // Сибирский математический журнал. – 1993. – Т. 34, № 6. – С. 75-85. – URL: http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=smj&paperid=811&option_lang=rus (дата обращения: 17.03.2021).

[2] Liskov, B. Programming with abstract data types / B. Liskov, S. Zilles. – DOI 10.1145/942572.807045 // ACM SIGPLAN Notices. – 1974. – Vol. 9, issue 4. – Pp. 50-59.

[3] Baroody, A. J. An object-oriented approach to database system implementation / A. J. Baroody, D. J. DeWitt. – DOI 10.1145/319628.319645 // ACM Transactions on Database Systems. – 1981. – Vol. 6, No. 4. – Pp. 576-601.

[4] Barford, L. Filtering of Randomly Sampled, Time-Stamped Measurements / L. Barford. – DOI 10.1109/IMTC.2006.328336 // 2006 IEEE Instrumentation and Measurement Technology Conference Proceedings. – IEEE Press, Sorrento, Italy, 2006. – Pp. 1021-1026.

[5] Agarwal, R. C. Fast one-dimensional digital convolution by multi-dimensional techniques / R. C. Agarwal, C. S. Burrus. – DOI 10.1109/TASSP.1974.1162532 // IEEE Transactions on Acoustics, Speech, and Signal Processing. – 1974. – Vol. 22, No. 1. – Pp. 1-10.

[6] Dynamic Convolution: Attention Over Convolution Kernels / Chen Y. [и др.]. – DOI 10.1109/CVPR42600.2020.01104 // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – IEEE Press, Seattle, WA, USA, 2020. – Pp. 11027-11036.

[7] Маршалко, Д. А. Архитектура свёрточных нейронных сетей / Д. А. Маршалко, О. В. Кубанских // Ученые записки Брянского государственного университета. – 2019. – № 4(16). – С. 10-13. – URL: https://www.elibrary.ru/item.asp?id=42222091 (дата обращения: 19.06.2021).

[8] Kanopoulos, N. Design of an image edge detection filter using the Sobel operator / N. Kanopoulos, N. Vasanthavada, R. L. Baker. – DOI 10.1109/4.996 // IEEE Journal of Solid-State Circuits. – 1988. – Vol. 23, No. 2. – Pp. 358-367.

[9] An improved Sobel edge detection / W. Gao [и др.]. – DOI 10.1109/ICCSIT.2010.5563693 // 2010 3rd International

[5] Технология программирования CUDA / Д. Н. Тумаков, Д. Е. Чикрин, А. А. Егорчев, С. В. Голоусов. Казань: КГУ, 2017. 112 с.

Conference on Computer Science and Information Technology. – IEEE Press, Chengdu, 2010. – Pp. 67-71.

[10] Chaple, G. Design of Sobel operator based image edge detection algorithm on FPGA / G. Chaple, R. D. Daruwala. – DOI 10.1109/ICCSP.2014.6949951 // 2014 International Conference on Communication and Signal Processing. – Melmaruvathur, India, 2014. – Pp. 788-792.

[11] Ottaviani, G. Introduction to the Hyperdeterminant and to the Rank of Multidimensional Matrices / G. Ottaviani. – DOI 10.1007/978-1-4614-5292-8_20 // Commutative Algebra; ed. by I. Peeva. – Springer, New York, NY, 2013. – Pp. 609-638.

[12] Goncharov, E. Multidimensional Matrix Algebra Versus Tensor Algebra or μ > 0 / E. Goncharov, P. Iljin, V. Munerman. – DOI 10.1109/EIConRus49466.2020.9039478 // 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus). – St. Petersburg and Moscow, Russia, 2020. – Pp. 1949-1954.

[13] Afanasyeva, Z. S. Signature Detection and Identification Algorithm with CNN, Numpy and OpenCV / Z. S. Afanasyeva, A. D. Afanasyev. – DOI 10.1007/978-3-030-63319-6_43 // Software Engineering Perspectives in Intelligent Systems. CoMeSySo 2020. Advances in Intelligent Systems and Computing; ed. by R. Silhavy, P. Silhavy, Z. Prokopova. – Springer, Cham, 2020. – Vol. 1295. – Pp. 467-479.

[14] A convolutional neural network based on TensorFlow for face recognition / L. Yuan, Z. Qu, Y. Zhao, H. Zhang, Q. Nian. – DOI 10.1109/IAEAC.2017.8054070 // 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC). – IEEE Press, Chongqing, China, 2017. – Pp. 525-529.

[15] Jayalakshmi, G. S. Performance analysis of Convolutional Neural Network (CNN) based Cancerous Skin Lesion Detection System / G. S. Jayalakshmi, V. S. Kumar. – DOI 110.1109/ICCIDS.2019.8862143 // 2019 International Conference on Computational Intelligence in Data Science (ICCIDS). – IEEE Press, Chennai, India, 2019. – Pp. 1-6.

[16] Kumar, N. K. Detection and Recognition of Objects in Image Caption Generator System: A Deep Learning Approach / N. K. Kumar [и др.]. – DOI 10.1109/ICACCS.2019.8728516 // 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS). – IEEE Press, Coimbatore, India, 2019. – Pp. 107-109.

[17] Hung, K.-W. Real-Time Image Super-Resolution Using Recursive Depthwise Separable Convolution Network / K. -W. Hung, Z. Zhang, J. Jiang. – DOI 10.1109/ACCESS.2019.2929223 // IEEE Access. – 2019. – Vol. 7. – Pp. 99804-99816.

[18] Maggiori, E. High-resolution aerial image labeling with convolutional neural networks / E. Maggiori [и др.]. – DOI 10.1109/TGRS.2017.2740362 // IEEE Transactions on Geoscience and Remote Sensing. – 2017. – Vol. 55, No. 12. – Pp. 7092-7103.

[19] Goncharov, E. Software and Hardware Complex for Calculating Convolutions by Methods Multidimensional Matrix Algebra / E. Goncharov, V. Munerman, G. Yakovlev. – DOI 10.1109/ElConRus51938.2021.9396584 // 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus). – IEEE Press, St. Petersburg, Moscow, Russia, 2021. – Pp. 2176-2180.

[20] Захаров, В. Н. Параллельный алгоритм умножения многомерных матриц / В. Н. Захаров, В. И. Мунерман // Современные информационные технологии и ИТ-образование. – 2015. – Т. 11, № 2. – С. 384-390. – URL: https://www.elibrary.ru/item.asp?id=26167519 (дата обращения: 19.06.2021).

[21] Mohamed, K. S. Parallel Computing: OpenMP, MPI, and CUDA / K. S. Mohamed. – DOI 10.1007/978-3-030-37224-8_3 // Neuromorphic Computing and Beyond. – Springer, Cham, 2020. – Pp. 63-93.

[22] Vetter, J. S. Keeneland: Bringing heterogeneous GPU computing to the computational science community / J. S. Vetter [и др.]. – DOI 10.1109/MCSE.2011.83 // Computing in Science & Engineering. – 2011. – Vol. 13, No. 05. – Pp. 90-95.

[23] Гончаров, Е. И. Выбор параметров многомерных матриц для обобщенного алгоритма шифрования Хилла / Е. И. Гончаров, В. И. Мунерман, Т. А. Самойлова // Системы компьютерной математики и их приложения. – 2019. – № 20-1. – С. 111-116. – URL: https://www.elibrary.ru/item.asp?id=39103166 (дата обращения: 19.06.2021).

[24] Мунерман, В. И. Алгебраический подход к алгоритмизации задач маршрутизации / В. И. Мунерман, Т. А. Самойлова. – DOI 10.18127/j20729472-201805-08 // Системы высокой доступности. – 2018. – Т. 14, № 5. – С. 50-56.

[25] Мунерман, В. И. Построение архитектур программно-аппаратных комплексов для повышения эффективности массовой обработки данных / В. И. Мунерман // Системы высокой доступности. – 2014. – Т. 10, № 4. – С. 3-16. – URL: https://www.elibrary.ru/item.asp?id=22831892 (дата обращения: 19.06.2021).

*Поступила 19.06.2021; одобрена после рецензирования 26.08.2021; принята к публикации 06.09.2021.*

Об авторе:

**Гончаров Евгений Игоревич,** магистрант физико-математического факультета, ФГБОУ ВО «Смоленский государственный университет» (214000, Российская Федерация, г. Смоленск, ул. Пржевальского, д. 4), **ORCID: https://orcid.org/0000-0002-3393-5772,** drbenvey1996@mail.ru

*Автор прочитал и одобрил окончательный вариант рукописи.*

## References

[1] Kisil V.V. On the algebra of pseudodifferential operators which is generated by convolutions on the Heisenberg group. *Sibirskii Matematicheskii Zhurnal* = Siberian Mathematical Journal. 1993; 34(6):1066-1075. (In Eng.) DOI: https://doi.org/10.1007/BF00973470

[2] Liskov B., Zilles S. Programming with abstract data types. *ACM SIGPLAN Notices.* 1974; 9(4):50-59. (In Eng.) DOI: https://doi.org/10.1145/942572.807045

[3] Baroody A.J., DeWitt D.J. An object-oriented approach to database system implementation. *ACM Transactions on Da-*

*tabase Systems.* 1981; 6(4):576-601. (In Eng.) DOI: https://doi.org/10.1145/319628.319645

[4] Barford L. Filtering of Randomly Sampled, Time-Stamped Measurements. *2006 IEEE Instrumentation and Measurement Technology Conference Proceedings.* IEEE Press, Sorrento, Italy; 2006. p. 1021-1026. (In Eng.) DOI: https://doi.org/10.1109/IMTC.2006.328336

[5] Agarwal R., Burrus C. Fast one-dimensional digital convolution by multidimensional techniques. *IEEE Transactions on Acoustics, Speech, and Signal Processing*. 1974; 22(1):1-10. (In Eng.) DOI: https://doi.org/10.1109/TASSP.1974.1162532

[6] Chen Y., Dai X., Liu M., Chen D., Yuan L., Liu Z. Dynamic Convolution: Attention Over Convolution Kernels. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).* IEEE Press, Seattle, WA, USA; 2020. p. 11027-11036. (In Eng.) DOI: https://doi.org/10.1109/CVPR42600.2020.01104

[7] Marshalko D.A., Kubanskih O.V. Convolutional Neural Network Architecture. *Scientific notes of the Bryansk State University.* 2019; (4):10-13. Available at: https://www.elibrary.ru/item.asp?id=42222091 (accessed 19.06.2021). (In Russ., abstract in Eng.)

[8] Kanopoulos N., Vasanthavada N., Baker R.L. Design of an image edge detection filter using the Sobel operator. *IEEE Journal of Solid-State Circuits*. 1988; 23(2):358-367. (In Eng.) DOI: https://doi.org/10.1109/4.996

[9] Gao W., Zhang X., Yang L., Liu H. An improved Sobel edge detection. *2010 3rd International Conference on Computer Science and Information Technology.* IEEE Press, Chengdu; 2010. p. 67-71. (In Eng.) DOI: https://doi.org/10.1109/ICCSIT.2010.5563693

[10] Chaple G., Daruwala R.D. Design of Sobel operator based image edge detection algorithm on FPGA. *2014 International Conference on Communication and Signal Processing.* IEEE Press, Melmaruvathur, India; 2014. p. 788-792. (In Eng.) DOI: https://doi.org/10.1109/ICCSP.2014.6949951

[11] Ottaviani G. Introduction to the Hyperdeterminant and to the Rank of Multidimensional Matrices. In: Ed. by I. Peeva. Commutative Algebra. Springer, New York, NY; 2013. p. 609-638. (In Eng.) DOI: https://doi.org/10.1007/978-1-4614-5292-8_20

[12] Goncharov E., Iljin P., Munerman V. Multidimensional Matrix Algebra Versus Tensor Algebra or μ > 0. *2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*. IEEE Press, St. Petersburg and Moscow, Russia; 2020. p. 1949-1954. (In Eng.) DOI: https://doi.org/10.1109/EIConRus49466.2020.9039478

[13] Afanasyeva Z.S., Afanasyev A.D. Signature Detection and Identification Algorithm with CNN, Numpy and OpenCV. In: Ed. by R. Silhavy, P. Silhavy, Z. Prokopova. *Software Engineering Perspectives in Intelligent Systems. CoMeSySo 2020. Advances in Intelligent Systems and Computing*. 2020; 1295:467-479. Springer, Cham. (In Eng.) DOI: https://doi.org/10.1007/978-3-030-63319-6_43

[14] Yuan L., Qu Z., Zhao Y., Zhang H., Nian Q. A convolutional neural network based on TensorFlow for face recognition. *2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC).* IEEE Press, Chongqing, China; 2017. p. 525-529. (In Eng.) DOI: https://doi.org/10.1109/IAEAC.2017.8054070

[15] Jayalakshmi G.S., Kumar V.S. Performance analysis of Convolutional Neural Network (CNN) based Cancerous Skin Lesion Detection System. *2019 International Conference on Computational Intelligence in Data Science (ICCIDS).* IEEE Press, Chennai, India; 2019. p. 1-6. (In Eng.) DOI: https://doi.org/110.1109/ICCIDS.2019.8862143

[16] Kumar N.K., Vigneswari D., Mohan A., Laxman K., Yuvaraj J. Detection and Recognition of Objects in Image Caption Generator System: A Deep Learning Approach. *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS).* IEEE Press, Coimbatore, India; 2019. p. 107-109. (In Eng.) DOI: https://doi.org/10.1109/ICACCS.2019.8728516

[17] Hung K.-W., Zhang Z., Jiang J. Real-Time Image Super-Resolution Using Recursive Depthwise Separable Convolution Network. *IEEE Access*. 2019; 7:99804-99816. (In Eng.) DOI: https://doi.org/10.1109/ACCESS.2019.2929223

[18] Maggiori E. High-resolution aerial image labeling with convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*. 2017; 55(12):7092-7103. (In Eng.) DOI: https://doi.org/10.1109/TGRS.2017.2740362

[19] Goncharov E., Munerman V., Yakovlev G. Software and Hardware Complex for Calculating Convolutions by Methods Multidimensional Matrix Algebra. *2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus).* IEEE Press, St. Petersburg, Moscow, Russia; 2021. p. 2176-2180. (In Eng.) DOI: https://doi.org/10.1109/ElConRus51938.2021.9396584

[20] Zakharov V.N., Munerman V. I. *Parallel'nyj algoritm umnozhenija mnogomernyh matric* [Parallel Algorithm for Multidimensional Matrix Multiplication]. *Sovremennye informacionnye tehnologii i IT-obrazovanie* = Modern Information Technologies and IT-Education. 2015; 11(2):384-390. Available at: https://www.elibrary.ru/item.asp?id=26167519 (accessed 19.06.2021). (In Russ., abstract in Eng.)

[21] Mohamed K.S. Parallel Computing: OpenMP, MPI, and CUDA. *Neuromorphic Computing and Beyond.* Springer, Cham; 2020. p. 63-93. (In Eng.) DOI: https://doi.org/10.1007/978-3-030-37224-8_3

[22] Vetter J.S., et al. Keeneland: Bringing Heterogeneous GPU Computing to the Computational Science Community. *Computing in Science & Engineering*. 2011; 13(5):90-95. (In Eng.) DOI: https://doi.org/10.1109/MCSE.2011.83

[23] Goncharov E.I., Munerman V.I., Samoylova T.A. The Method of Selecting Parameters of Multidimensional Matrix for Hill Encryption Algorithm. *Sistemy komp'yuternoj matematiki i ih prilozheniya* = Computer Mathematics Systems and Their Applications. 2019; (20-1):111-116. Available at: https://www.elibrary.ru/item.asp?id=39103166 (accessed 19.06.2021). (In Russ., abstract in Eng.)

[24] Munerman V.I., Samoylova T.A. Algebraic Approach to Algorithmization of Routing Problems. *Highly Available Systems.* 2018; 14(5):50-56. (In Russ., abstract in Eng.) DOI: https://doi.org/10.18127/j20729472-201805-08

[25] Munerman V.I. Construction of hardware-software com-

plexes architecture to improve massively data processing. *Highly Available Systems.* 2014; 10(4):3-16. Available at: https://www.elibrary.ru/item.asp?id=22831892 (accessed 19.06.2021). (In Russ., abstract in Eng.)

*Submitted 19.06.2021; approved after reviewing 26.08.2021; accepted for publication 06.09.2021.*

About the author:

**Evgeniy I. Goncharov**, Master student of the Faculty of Physics and Mathematics, Smolensk State University (4 Przhevalsky St., Smolensk 214000, Russian Federation), **ORCID: https://orcid.org/0000-0002-3393-5772,** drbenvey1996@mail.ru

*The author has read and approved the final manuscript.*