

## Высокопроизводительные архитектуры цифровой фильтрации изображений в системе остаточных классов на основе метода Винограда

М.В. Валуева<sup>1</sup>, П.А. Ляхов<sup>1</sup>, Н.Н. Нагорнов<sup>1</sup>, Г.В. Валуев<sup>1</sup>

<sup>1</sup> Северо-Кавказский федеральный университет,  
355017, г. Ставрополь, ул. Пушкина, д. 1

### Аннотация

Непрерывное совершенствование методов регистрации, обработки и хранения визуальной информации приводит к необходимости улучшения технических характеристик систем цифровой обработки изображений. В данной работе предлагаются новые высокопроизводительные архитектуры цифровых фильтров для обработки изображений по методу Винограда с вычислениями в системе остаточных классов с модулями специального вида. Для оценки производительности и аппаратных затрат предложенных архитектур проведено аппаратное моделирование с использованием программируемых логических интегральных схем в среде автоматизированного проектирования Xilinx Vivado 2018.3 для целевой платы Artix-7 xc7a200tffg1156-3. Результаты аппаратного моделирования показали, что предлагаемые архитектуры цифровых фильтров повышают производительность до 5,42 раза по сравнению с известными методами за счет увеличения аппаратных затрат. Результаты данного исследования могут быть использованы при проектировании сложных систем обработки и анализа изображений, с целью увеличения их производительности.

**Ключевые слова:** цифровые фильтры, обработка изображений, система остаточных классов, фильтрация по методу Винограда, Field-Programmable Gate Array.

**Цитирование:** Валуева, М.В. Высокопроизводительные архитектуры цифровой фильтрации изображений в системе остаточных классов на основе метода Винограда / М.В. Валуева, П.А. Ляхов, Н.Н. Нагорнов, Г.В. Валуев // Компьютерная оптика. – 2022. – Т. 46, № 5. – С. 752-762. – DOI: 10.18287/2412-6179-CO-933.

**Citation:** Valueva MV, Lyakhov PA, Nagornov NN, Valuev GV. High-performance digital image filtering architectures in the residue number system based on the Winograd method. Computer Optics 2022; 46(5): 752-762. DOI: 10.18287/2412-6179-CO-933.

### Введение

Цифровая обработка изображений широко востребована в различных областях человеческой деятельности, таких как медицинская визуализация [1], геолокация [2], системы видеонаблюдения [3], контроль качества продукции на производстве [4] и многих других. Цифровая фильтрация используется в сложных системах обработки изображений с целью очистки от шума [5], сегментации [6], детектирования [7] или распознавания [8]. Линейные пространственные фильтры составляют основу многих систем обработки изображений, например, систем нейросетевого анализа изображений [9, 10]. В данной работе будет рассматриваться именно этот вид фильтров. Они делятся на сглаживающие фильтры, фильтры выделения контуров и фильтры повышения резкости [11].

Фильтрация изображения в оттенках серого  $I$  размера  $R \times C$  производится следующим образом. Изображение обрабатывается фильтром  $W$  размером  $k \times k$ , при этом размер обработанного изображения меньше исходного на  $k-1$ , если  $k$  нечетное, и на  $k$ , если  $k$  четное. Чтобы размер изображения не уменьшался после фильтрации, к исходному изображению добавляют рамку из столбцов и строк соответствующего размера [12]. Далее будем рас-

сматривать только случай, когда  $k$  – нечетное, так как он является наиболее распространенным на практике. Таким образом, на вход подается изображение  $I_p$  размером  $(R+k-1) \times (C+k-1)$ . После фильтрации изображения  $I_p$  фильтром размером  $k \times k$  на выходе формируется изображение  $I_f$  размером  $R \times C$  следующим образом:

$$I_f(x, y) = \sum_{i=-t}^t \sum_{j=-t}^t I_p(x+i+t, y+j+t) \cdot W(i, j), \quad (1)$$

где  $0 \leq x < R$ ,  $0 \leq y < C$ ,  $t = \lfloor k/2 \rfloor$ . Схема процесса фильтрации изображения представлена на рис. 1.

Как видно из формулы (1), операции умножения и сложения лежат в основе цифровой фильтрации. Основная вычислительная нагрузка при линейной фильтрации заключается в многократном выполнении операции умножения. Аппаратная реализация процесса фильтрации позволяет увеличить скорость обработки изображений [13]. В связи с этим исследователи ведут разработки в области проектирования эффективных с точки зрения скорости энергопотребления и площади сумматоров и умножителей. В работе [14] авторы предлагают новое устройство умножения с накоплением. Авторы статьи [15] представили новое усеченное устройство умножения с накоплением

для реализации цифровых фильтров с конечной импульсной характеристикой.

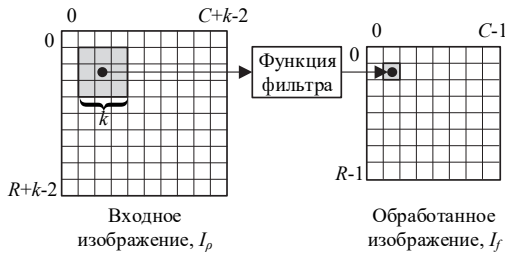


Рис. 1. Фильтрация изображения

Одним из подходов к увеличению быстродействия цифрового фильтра является уменьшение количества операций умножения. В работе [16] предложен метод фильтрации Винограда, который позволяет сократить количество умножений в процессе фильтрации за счет увеличения количества сложений. В статье [17] авторы предлагают использовать алгоритм Винограда для реализации операции свертки в сверточной нейронной сети при выполнении вычислений на графическом процессоре. Авторы статьи [18] предлагают аппаратный ускоритель сверточной нейронной сети, реализованный на FPGA (Field-Programmable Gate Array) с применением алгоритма Винограда.

Еще одним подходом к увеличению производительности, то есть количества обработанных кадров в единицу времени, цифровой фильтрации является использование непозиционной системы счисления – системы остаточных классов (СОК) [19]. Она позволяет представлять числа большой разрядности в виде небольших остатков от деления по модулю и выполнять вычисления параллельно по каждому модулю. Применение такого подхода позволяет увеличить скорость выполнения модульных операций сложения и умножения, являющихся основой для выполнения цифровой фильтрации. В работе [20] авторы предлагают архитектуру усеченного блока умножения с накоплением, использующего арифметику СОК. В статье [21] представлен метод сглаживающей фильтрации изображений с использованием СОК. Авторами предлагается метод квантования коэффициентов маски фильтра. В работе [22] предлагается архитектура устройства для сглаживания изображений и выделения контуров, которая использует трехмодульный СОК. Авторы статьи [23] предлагают методологию построения цифровых фильтров в СОК для автоматизации процесса проектирования и обеспечения эффективного соотношения скорости и энергоэффективности устройств. В работе [24] представлена аппаратная реализация фильтра с конечной импульсной характеристикой на FPGA в СОК с набором модулей специального вида.

В данной работе предлагаются архитектуры устройств для фильтрации изображений методом Винограда с использованием СОК с модулями специального вида  $2^\alpha$  и  $2^\alpha \pm 1$ ,  $k \in N$ , где  $N$  – множество

натуральных чисел. Для демонстрации их высокой производительности в экспериментальной части статьи представлены результаты аппаратного моделирования предложенной архитектуры, и проведено сравнение с другими известными методами.

Оставшаяся часть статьи организована следующим образом. В параграфе 1 кратко описано представление чисел в СОК и особенности применения СОК в цифровой фильтрации изображений. В параграфе 2 представлен метод Винограда для цифровой фильтрации. В параграфе 3 предложен метод построения цифровых фильтров на основе метода Винограда и арифметики СОК. Архитектуры предлагаемых устройств для фильтрации изображений описаны в параграфе 4. Параграф 5 содержит результаты аппаратного моделирования. Обсуждение результатов моделирования представлено в параграфе 6. В заключении изложены основные выводы по проведенному исследованию.

### 1. Применение системы остаточных классов для фильтрации изображений

В СОК целое число  $A$  представляется как набор остатков от деления  $\{a_1, a_2, \dots, a_n\}$  на соответствующие попарно взаимно простые модули  $\{p_1, p_2, \dots, p_n\}$ , так что  $a_i = |A|_{p_i}$  [25]. Произведение всех модулей

$$P = \prod_{i=1}^n p_i$$

называется динамическим диапазоном СОК. Арифметические операции над числами, представленными в СОК, производятся параллельно по каждому модулю:

$$A * B = (|a_1 * b_1|_{p_1}, |a_2 * b_2|_{p_2}, \dots, |a_n * b_n|_{p_n}), \quad (2)$$

где  $*$  означает операцию сложения, вычитания или умножения.

Число  $A = \{a_1, a_2, \dots, a_j\}$  может быть преобразовано из СОК в позиционную систему счисления (ПСС) с использованием Китайской теоремы об остатках [25]

$$A = \left| \sum_{i=1}^j \left( |P_i^{-1}|_{m_i} a_i |P_i| \right) \right|_P, \quad (3)$$

где  $P_i = P / p_i$ , а  $|P_i^{-1}|_{p_i}$  – мультипликативный обратный элемент для  $P_i$ .

Фильтрация изображений по формуле (1) в СОК выполняется в несколько этапов (рис. 2). Сперва необходимо перевести данные из ПСС в СОК. Затем производится фильтрация параллельно по нескольким вычислительным каналам, которые соответствуют модулям СОК. Далее производится обратное преобразование из СОК в ПСС.

Вид модулей СОК влияет на производительность вычислений, поэтому их выбор является важной задачей при проектировании прикладных систем, использующих арифметику СОК. Во-первых, набор мо-

дулей должен обеспечивать достаточный динамический диапазон системы для однозначного представления чисел в СОК. Во-вторых, модули должны быть сбалансированы таким образом, чтобы время выполнения операций по каждому каналу было примерно одинаково и не было длительного простоя системы по какому-либо вычислительному каналу. Наконец, модули СОК специального вида  $2^\alpha$  и  $2^\alpha \pm 1$ ,  $k \in N$ , где  $N$  – множество натуральных чисел, позволяют избежать затратной по ресурсам операции деления по модулю. Проектирование арифметических устройств по модулю вида  $2^\alpha$  сводится к реализации  $\alpha$ -битных устройств в двоичной системе счисления. Но требование попарной взаимной простоты модулей СОК позволяет нам использовать лишь один модуль вида  $2^\alpha$ , так как остальные модули должны быть нечетными. Использование модулей вида  $2^\alpha$  и  $2^\alpha \pm 1$  позволяет использовать эффективные техники суммирования по модулю, аналогичные используемым в традиционной двоичной системе счисления [26]. Но практическая реализация вычислительного канала по модулю  $2^\alpha + 1$ ,  $k \in N$ , требует введения дополнительной логики по методу «diminished-1» для отслеживания нулевой кодовой комбинации, что является нежелательным явлением при разработке системы с минимальными аппаратными и временными затратами [27, 28]. Таким образом, модули вида  $2^\alpha - 1$  предпочтительнее для использования в качестве нечетных модулей системы.



Рис. 2. Фильтрация изображения в СОК

Фильтрация по методу Винограда также может быть реализована с использованием арифметики СОК. В следующих параграфах представлен метод фильтрации Винограда и предложены архитектуры фильтров на основе данного метода и СОК с модулями специального вида.

### 2. Фильтрация по методу Винограда

Одномерная фильтрация по методу Винограда в матричной форме имеет вид:

$$\vec{z} = \mathbf{A}^T \left( (\mathbf{G}\vec{w}) \odot (\mathbf{B}^T \vec{d}) \right), \quad (4)$$

где оператор  $\odot$  обозначает поэлементное умножение матриц,  $\mathbf{A}$ ,  $\mathbf{G}$  и  $\mathbf{B}$  – матрицы преобразования,  $\vec{w}$  – маска одномерного фильтра,  $\vec{d}$  – вектор данных,  $\vec{z}$  – результат фильтрации [17]. Алгоритм одномерной фильтрации по методу Винограда принято обозначать  $F(n, k)$ , где  $n$  – размер вектора  $\vec{z}$ , а  $k$  – размер маски фильтра  $\vec{w}$ .

Двумерная фильтрация по методу Винограда в матричной форме имеет вид [17]:

$$\mathbf{Z} = \mathbf{A}^T \left( (\mathbf{G}\mathbf{W}\mathbf{G}^T) \odot (\mathbf{B}^T \mathbf{D}\mathbf{B}) \right) \mathbf{A}, \quad (5)$$

где  $\mathbf{W}$  – двумерная маска фильтра,  $\mathbf{D}$  – двумерный массив данных и  $\mathbf{Z}$  – двумерный массив результата фильтрации. Алгоритм двумерной фильтрации по методу Винограда принято обозначать  $F(n \times n, k \times k)$ .

Рассмотрим пример одномерной фильтрации по методу Винограда  $F(2, 3)$ . Пусть имеется вектор данных  $\vec{d} = \{d_i\}$ ,  $0 \leq i \leq 3$ , и фильтр с маской  $\vec{w} = \{w_j\}$ ,  $0 \leq j \leq 2$ , в результате фильтрации формируется вектор  $\vec{z} = \{z_l\}$ ,  $0 \leq l \leq 1$ . Представим  $\vec{z}$  и  $\vec{w}$  в виде полиномов, тогда  $z(x) = z_1x + z_0$ ,  $w(x) = w_2x^2 + w_1x + w_0$ . Произведение  $z(x)w(x)$  может быть записано в виде

$$\begin{aligned} z(x)w(x) &= (z_1x + z_0)(w_2x^2 + w_1x + w_0) = \\ &= z_1w_2x^3 + (z_1w_1 + z_0w_2)x^2 + (z_1w_0 + z_0w_1)x + z_0w_0. \end{aligned} \quad (6)$$

Представим слагаемые из выражения (6) в виде:

$$\begin{aligned} z_0w_0 &= z_0w_0, \\ z_1w_0 + z_0w_1 &= (z_0 + z_1)(w_0 + w_1) - z_0w_0 - z_1w_1, \\ z_1w_1 + z_0w_2 &= (z_0 + z_1)(w_1 + w_2) - z_0w_1 - z_1w_2, \\ z_1w_2 &= z_1w_2. \end{aligned} \quad (7)$$

Умножим обе части равенств на соответствующие элементы вектора  $\vec{d}$ .

$$\begin{aligned} d_0z_0w_0 + d_1z_1w_0 + d_1z_0w_1 + d_2z_1w_1 + d_2z_0w_2 + \\ + d_3z_1w_2 &= d_0z_0w_0 + d_1(z_0 + z_1)(w_0 + w_1) - \\ - d_1z_0w_0 - d_1z_1w_1 + d_2(z_0 + z_1)(w_1 + w_2) - \\ - d_2z_0w_1 - d_2z_1w_2 + d_3z_1w_2. \end{aligned} \quad (8)$$

Приравнявая соответствующие множители при  $z_0$ , получим:

$$\begin{aligned} d_0w_0 + d_1w_1 + d_2w_2 &= (d_0 - d_2)w_0 + \\ + \frac{1}{2}(d_1 + d_2)(w_0 + w_1 + w_2) + (d_2 - d_1)(w_0 - w_1 + w_2). \end{aligned} \quad (9)$$

Приравнявая соответствующие множители при  $z_1$ , получим:

$$\begin{aligned} d_1w_0 + d_2w_1 + d_3w_2 &= \\ = \frac{1}{2}(d_1 + d_2)(w_0 + w_1 + w_2) - \\ - \frac{1}{2}(d_2 - d_1)(w_0 - w_1 + w_2) - (d_1 - d_3)w_2. \end{aligned} \quad (10)$$

Данные выражения (8–10) можно записать в виде

$$\begin{bmatrix} z_0 \\ z_1 \end{bmatrix} = \begin{bmatrix} d_0 & d_1 & d_2 \\ d_1 & d_2 & d_3 \end{bmatrix} \cdot \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} m_1 + m_2 + m_3 \\ m_2 - m_3 - m_4 \end{bmatrix}, \quad (11)$$

где

$$\begin{aligned} m_1 &= (d_0 - d_2)w_0, \\ m_2 &= 1/2(d_1 + d_2)(w_0 + w_1 + w_2), \\ m_3 &= 1/2(d_2 - d_1)(w_0 - w_1 + w_2), \\ m_4 &= (d_1 - d_3)w_2. \end{aligned}$$

Пусть

$$\bar{\mathbf{w}} = [w_0 \ w_1 \ w_2]^T, \quad \bar{\mathbf{d}} = [d_0 \ d_1 \ d_2 \ d_3]^T.$$

Тогда матрицы преобразования для  $F(2 \times 2, 3 \times 3)$  имеют вид

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}, \quad (12)$$

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & -1 \\ 0 & -1 \end{bmatrix}.$$

В случае двумерной фильтрации  $F(2 \times 2, 3 \times 3)$  необходимо воспользоваться формулой (5). На рис. 3 представлена схема фильтрации изображения методом Винограда  $F(2 \times 2, 3 \times 3)$ .

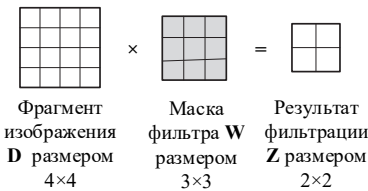


Рис. 3. Схема фильтрации фрагмента изображения по методу Винограда  $F(2 \times 2, 3 \times 3)$

В следующем параграфе предложена архитектура устройства двумерной фильтрации методом Винограда  $F(2 \times 2, 3 \times 3)$  с использованием арифметики СОК.

### 3. Метод проектирования цифровых фильтров на основе метода Винограда и арифметики системы остаточных классов

В работе предлагается новый метод построения цифрового фильтра размером  $k \times k$  для обработки изображений на основе метода Винограда и арифметики СОК. Разделим изображение на фрагменты меньшего размера  $m \times m$ , так что каждый фрагмент будет обрабатываться с помощью  $F(n \times n, k \times k)$ , где  $k < m$  и  $n < m$ . То есть для обработки изображения требуется выбирать фрагменты  $\mathbf{D}$  размером  $m \times m$  с шагом  $k$  по строкам и столбцам. Пусть фрагмент изображения обрабатывается фильтром с маской  $\mathbf{W}$  раз-

мером  $k \times k$ , тогда после обработки формируется фрагмент  $\mathbf{Z}$  размером  $n \times n$ . Например, для случая  $F(2 \times 2, 3 \times 3)$ , изображенного на рис. 3, выбираются фрагменты изображения  $\mathbf{D}$  размером  $4 \times 4$  с шагом 2, результатом обработки фрагмента фильтром  $\mathbf{W}$  с маской размером  $3 \times 3$  по формуле (5) является фрагмент обработанного изображения  $\mathbf{Z}$  размером  $2 \times 2$ . То есть при фильтрации методом Винограда  $F(2 \times 2, 3 \times 3)$  формируются сразу четыре пикселя обработанного изображения.

В выражении (5) преобразование маски фильтра  $\mathbf{GWG}^T$  может быть выполнено предварительно один раз и, следовательно, не несет существенной вычислительной нагрузки. Введем обозначения  $\mathbf{U} = \mathbf{GWG}^T$ ,  $\mathbf{V} = \mathbf{B}'\mathbf{D}\mathbf{B}$  и  $\mathbf{M} = \mathbf{U}\mathbf{V}$ , тогда  $\mathbf{Z} = \mathbf{A}'\mathbf{M}\mathbf{A}$ . Цифровой фильтр, выполняющий вычисления по методу Винограда, изображен на рис. 4 и состоит из:

- 1) блока памяти, в котором хранится преобразованная маска фильтра  $\mathbf{U}$ ;
- 2) устройства преобразования входных данных  $\mathbf{D}$ , выполняющего вычисления матрицы  $\mathbf{V}$ ;
- 3) устройства поэлементного умножения матриц  $\mathbf{U}$  и  $\mathbf{V}$ , на выходе которого формируется матрица  $\mathbf{M}$ ;
- 4) устройства итогового преобразования матрицы  $\mathbf{M}$ , выходом которого является фрагмент обработанного изображения  $\mathbf{Z}$ .

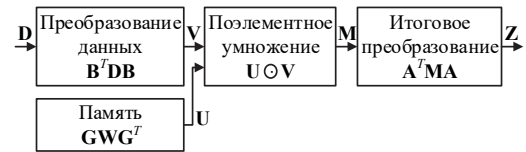


Рис. 4. Схема устройства фильтрации фрагмента изображения по методу Винограда  $F(n \times n, k \times k)$

Коэффициенты маски фильтра обычно являются вещественными числами. Для представления в цифровом устройстве их необходимо привести к целочисленному виду. В работе [21] представлен метод масштабирования коэффициентов для фильтров Гаусса. В соответствии с данным методом исходные коэффициенты фильтра умножаются на  $2^\gamma$ , где  $\gamma = 10 + t$  – параметр масштабирования. Далее коэффициенты округляются к большему. Обозначим маску фильтра с масштабированными коэффициентами как  $\tilde{\mathbf{W}}$ . После наложения цифрового фильтра на изображение выполняется масштабирование полученных результатов на  $2^{-\gamma}$  и округление к меньшему.

При использовании СОК вычисления производятся параллельно по каждому модулю системы (рис. 2). Как упоминалось выше, выбор модулей СОК является важным этапом при проектировании цифровых устройств. В работе предлагается использовать модули специального вида  $2^\alpha$  и  $2^\alpha \pm 1$ . Динамический диапазон СОК для устройства фильтрации изображения с 8-битным представлением пикселей должен удовлетворять неравенству [29]

$$P \geq 255 \cdot \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} \widetilde{W}(i, j). \quad (13)$$

На рис. 5 представлена схема фильтра, выполняющего преобразование фрагмента изображения  $\mathbf{D}$  по методу Винограда  $F(n \times n, k \times k)$  в СОК. На вход устройства поступает фрагмент изображения, представленный в СОК с набором модулей  $\{2^{\alpha_1}, 2^{\alpha_2}-1, \dots, 2^{\alpha_\beta}-1\}$ , где  $\beta$  – количество модулей, таким образом,

$$\mathbf{D} = \{|\mathbf{D}|_{2^{\alpha_1}}, |\mathbf{D}|_{2^{\alpha_2}-1}, \dots, |\mathbf{D}|_{2^{\alpha_\beta}-1}\}.$$

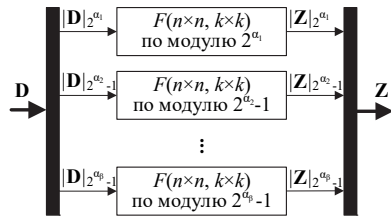


Рис. 5. Схема устройства фильтрации фрагмента изображения по методу Винограда  $F(n \times n, k \times k)$  с вычислениями в СОК

Вычисления производятся параллельно по  $\beta$  каналам, которые соответствуют набору модулей СОК. На вы-

ход устройства поступает фрагмент обработанного изображения  $\mathbf{Z}$ , представленный в СОК.

В следующем параграфе описаны архитектуры устройств для фильтрации изображений по методу Винограда  $F(2 \times 2, 3 \times 3)$  с использованием СОК, построенные на основе предлагаемого метода.

#### 4. Предлагаемые архитектуры устройств фильтрации изображений на основе метода Винограда

Далее представлено описание предлагаемых архитектур устройств фильтрации изображений на основе метода Винограда  $F(2 \times 2, 3 \times 3)$  с вычислениями в СОК. В качестве модулей СОК выбраны модули специального вида  $2^\alpha$  и  $2^\alpha \pm 1$ , так что для каждого вида модулей предлагаются разные устройства. Предлагаемую архитектуру устройства цифровой фильтрации по модулю  $2^\alpha$  в СОК можно также использовать для выполнения вычислений в ПСС с числами разрядности  $\alpha$  бит.

Для случая  $F(2 \times 2, 3 \times 3)$  размер матриц  $\mathbf{U}$ ,  $\mathbf{V}$  и  $\mathbf{M}$  составляет  $4 \times 4$ . Вычисление матриц  $\mathbf{V}$  и  $\mathbf{Z}$  может быть представлено в виде:

$$\mathbf{V} = \begin{bmatrix} d_{0,0} - d_{2,0} - (d_{0,2} - d_{2,2}) & d_{0,1} - d_{2,1} + d_{0,2} - d_{2,2} & -(d_{0,1} - d_{2,1}) + d_{0,2} - d_{2,2} & d_{0,1} - d_{2,1} - d_{0,3} + d_{2,3} \\ d_{1,0} + d_{2,0} - (d_{1,2} + d_{2,2}) & d_{1,1} + d_{2,1} + d_{1,2} + d_{2,2} & -(d_{1,1} + d_{2,1}) + d_{1,2} + d_{2,2} & d_{1,1} + d_{2,1} - d_{1,3} - d_{2,3} \\ d_{2,0} - d_{1,0} - (d_{2,2} - d_{1,2}) & d_{2,1} - d_{1,1} - d_{1,2} + d_{2,2} & -(d_{2,1} - d_{1,1}) - d_{1,2} + d_{2,2} & d_{2,1} - d_{1,1} + d_{1,3} - d_{2,3} \\ d_{1,0} - d_{3,0} - (d_{1,2} - d_{3,2}) & d_{1,1} - d_{3,1} + d_{1,2} - d_{3,2} & -(d_{1,1} - d_{3,1}) + d_{1,2} - d_{3,2} & d_{1,1} - d_{3,1} - d_{1,3} + d_{3,3} \end{bmatrix}, \quad (14)$$

$$\mathbf{Z}_{0,j} = \begin{bmatrix} M_{0,0} + M_{1,0} + M_{2,0} + M_{0,1} + M_{1,1} + M_{2,1} + M_{0,2} + M_{1,2} + M_{2,2} \\ M_{0,1} + M_{1,1} + M_{2,1} - (M_{0,2} + M_{1,2} + M_{2,2}) - M_{0,3} - M_{1,3} - M_{2,3} \end{bmatrix}^T, \quad (15)$$

$$\mathbf{Z}_{1,j} = \begin{bmatrix} M_{1,0} - M_{2,0} - M_{3,0} + M_{1,1} - M_{2,1} - M_{3,1} + M_{1,2} - M_{2,2} - M_{3,2} \\ M_{1,1} - M_{2,1} - M_{3,1} - (M_{1,2} - M_{2,2} - M_{3,2}) - M_{1,3} + M_{2,3} + M_{3,3} \end{bmatrix}^T,$$

где  $0 \leq j \leq 1$ .

Схема устройства для вычисления одного элемента матрицы  $\mathbf{V}$  по модулю  $2^\alpha$  представлена на рис. 6а. Данное устройство производит преобразование данных из матрицы  $\mathbf{D}$  и генерирует элемент  $V_{i,j}$ ,  $0 \leq i \leq 3$ ,  $0 \leq j \leq 3$ . Далее будем обозначать данное устройство как DT (data transformation), на входы которого совместно подаются вектор входных данных

$$\bar{\mathbf{h}}^{i,j} = \{h_0^{i,j}, h_1^{i,j}, h_2^{i,j}, h_3^{i,j}\},$$

сформированный из матрицы  $\mathbf{D}$ , и корректирующая константа  $c_{DT}^{i,j}$ , которая равна количеству операций вычитания. Корректирующие константы необходимы для представления отрицательных чисел в дополнительном коде, который соответствует инверсии числа в двоичном коде с добавлением 1. Инверсию двоичного числа будем обозначать чертой сверху. Для расчета элемента  $V_{0,0}$  необходимо учитывать, что

$$\bar{\mathbf{h}}^{0,0} = \{d_{0,0}, d_{2,2}, \overline{d_{2,0}}, d_{0,2}\},$$

а  $c_{DT}^{0,0} = 2$ . Аналогично формируются данные для вычисления других элементов матрицы  $\mathbf{V}$ . Особенностью расчета элемента  $V_{1,1}$  является то, что на вход поступают инвертированные числа

$$\bar{\mathbf{h}}^{1,1} = \{d_{1,1}, d_{2,1}, \overline{d_{2,0}}, \overline{d_{0,2}}\}$$

и корректирующая константа  $c_{DT}^{1,1} = 0$ . Сложение чисел производится с помощью дерева сумматоров с сохранением переноса (Carry-Save Adder, CSA) и сумматора (Kogge-Stone Adder, KSA) по модулю  $2^\alpha$ , то есть без учета старшего бита переноса [30, 31]. Таким образом, для вычисления матрицы  $\mathbf{V}$  по модулю  $2^\alpha$  используются 16 устройств DT, изображенных на рис. 6а, работающих параллельно. Устройство DT для вычисления одного элемента матрицы  $\mathbf{V}$  по модулю  $2^\alpha - 1$  имеет схожий принцип работы (рис. 6б). Отличие заключается в использовании техники циклического переноса старших бит (End-Around-Carry, EAC) [26]. Кроме того, для выполнения вычислений по модулю  $2^\alpha - 1$  отрицательные числа представляют-

ся в обратном коде, следовательно, отсутствуют корректирующие константы.

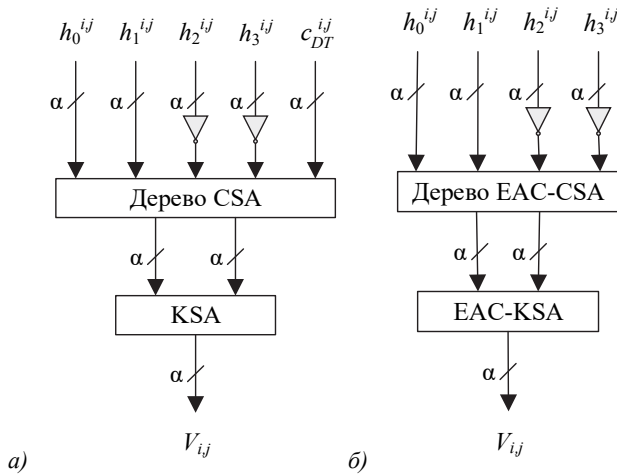


Рис. 6. Схема устройства DT для вычисления одного элемента матрицы V: а) по модулю  $2^\alpha$ ; б) по модулю  $2^{\alpha-1}$

Поэлементное умножение матриц **U** и **V** производится с помощью 16 параллельных умножителей двух чисел, изображенных на рис. 7, обозначенных MUL (multiplication). На вход устройства поступают элементы матриц **U** и **V**. На рис. 7а представлена схема устройства умножения двух чисел по модулю  $2^\alpha$ . Сперва с помощью массива вентилей AND формируются частичные произведения [30]. Так как расчеты производятся по модулю  $2^\alpha$ , то биты старше  $\alpha$  не участвуют в вычислениях. Затем производится сложение по модулю  $2^\alpha$  с помощью дерева сумматоров CSA и сумматора KSA. На рис. 7б представлена схема устройства умножения двух чисел по модулю  $2^{\alpha-1}$ . Для выполнения вычислений по модулю  $2^{\alpha-1}$  используется техника EAC. Таким образом формируется матрица **M** размером  $4 \times 4$ .

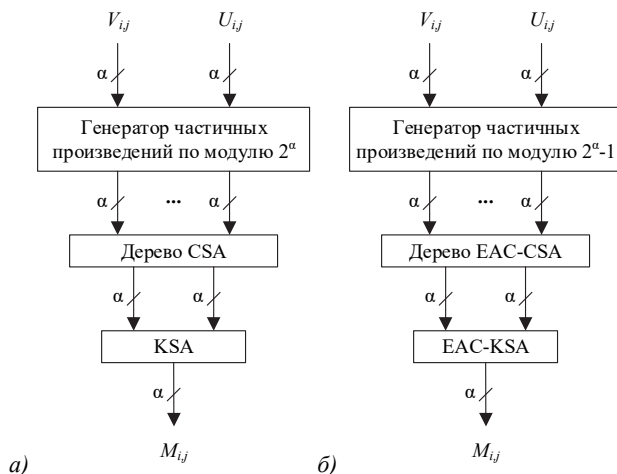


Рис. 7. Схема устройства MUL для умножения двух чисел: а) по модулю  $2^\alpha$ ; б) по модулю  $2^{\alpha-1}$

Схема устройства для вычисления *i*-й строки матрицы **Z** представлена на рис. 8. Введем для данного устройства обозначение FTR (final transform for

row). Для вычисления матрицы **Z** используются два устройства FTR, изображенных на рис. 8, работающих параллельно. На вход поступает вектор  $\vec{r}^i = \{r_0^i, r_1^i, \dots, r_{11}^i\}$ , сформированный из элементов матрицы **M**. Для выполнения расчетов по модулю  $2^\alpha$  на вход устройства также подаются корректирующие константы  $\vec{c}_{FTR}^i = \{c_0^i, c_1^i\}$  (рис. 8а). Для расчета элементов первой строки

$$Z_{0,j}, j = 0,1, \\ \vec{r}^{-0} = \{M_{0,0}, M_{1,0}, M_{2,0}, M_{0,1}, M_{1,1}, M_{2,1}, M_{0,2}, M_{1,2}, M_{2,2}, \\ \overline{M_{0,3}}, \overline{M_{1,3}}, \overline{M_{2,3}}\} \text{ и } \vec{c}_{FTR}^{-0} = \{0,5\}.$$

Для расчета элементов второй строки

$$Z_{1,j}, j = 0,1, \\ \vec{r}^{-1} = \{M_{1,0}, \overline{M_{2,0}}, \overline{M_{3,0}}, M_{1,1}, \overline{M_{2,1}}, \overline{M_{3,1}}, M_{1,2}, \overline{M_{2,2}}, \overline{M_{3,2}}, \\ \overline{M_{1,3}}, \overline{M_{2,3}}, M_{3,3}\} \text{ и } \vec{c}_{FTR}^{-1} = \{6,7\}.$$

Сложение чисел по модулю  $2^\alpha$  производится с помощью дерева сумматоров CSA и сумматора KSA. Для выполнения вычислений по модулю  $2^{\alpha-1}$  используется техника EAC, а отрицательные числа представляются в обратном коде (рис. 8б).

Предлагаемая схема устройства для двумерной фильтрации с использованием модифицированного метода Винограда на основе СОК представлена на рис. 9. Вычисления производятся по формуле (5). Преобразование маски фильтра  $\mathbf{U} = \mathbf{G}\mathbf{W}\mathbf{G}^T$  производится предварительно, и результат хранится в памяти устройства. На рис. 9а представлена схема предлагаемого устройства фильтрации по модулю  $2^\alpha$ . Так как для выполнения операций с отрицательными числами требуется их представление в дополнительном коде, то в памяти устройства также хранятся корректирующие константы. Блок преобразования данных состоит из 16 устройств DT, представленных на рис. 6а. Блок поэлементного умножения матриц состоит из 16 устройств MUL, изображенных на рис. 7а. Блок итогового преобразования состоит из двух устройств FTR, представленных на рис. 8а. На рис. 9б изображена схема предлагаемого устройства фильтрации по модулю  $2^{\alpha-1}$ . В памяти данного устройства хранится только преобразованная маска фильтра. Блоки преобразования данных, поэлементного умножения матриц и итогового преобразования состоят из соответствующих устройств, представленных на рис. 6б, рис. 7б и рис. 8б.

В следующем параграфе описано аппаратное моделирование на FPGA предложенных устройств для оценки их эффективности и сравнения с известными архитектурами устройств цифровой фильтрации.

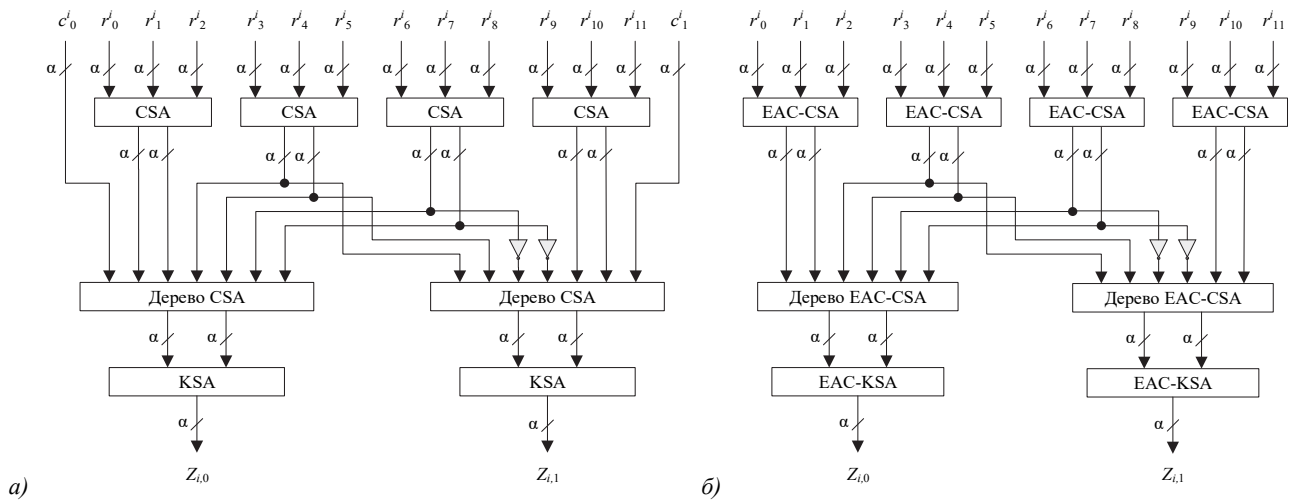


Рис. 8. Схема устройства FTR для вычисления  $i$ -й строки матрицы  $z$ : а) по модулю  $2^\alpha$ ; б) по модулю  $2^\alpha-1$

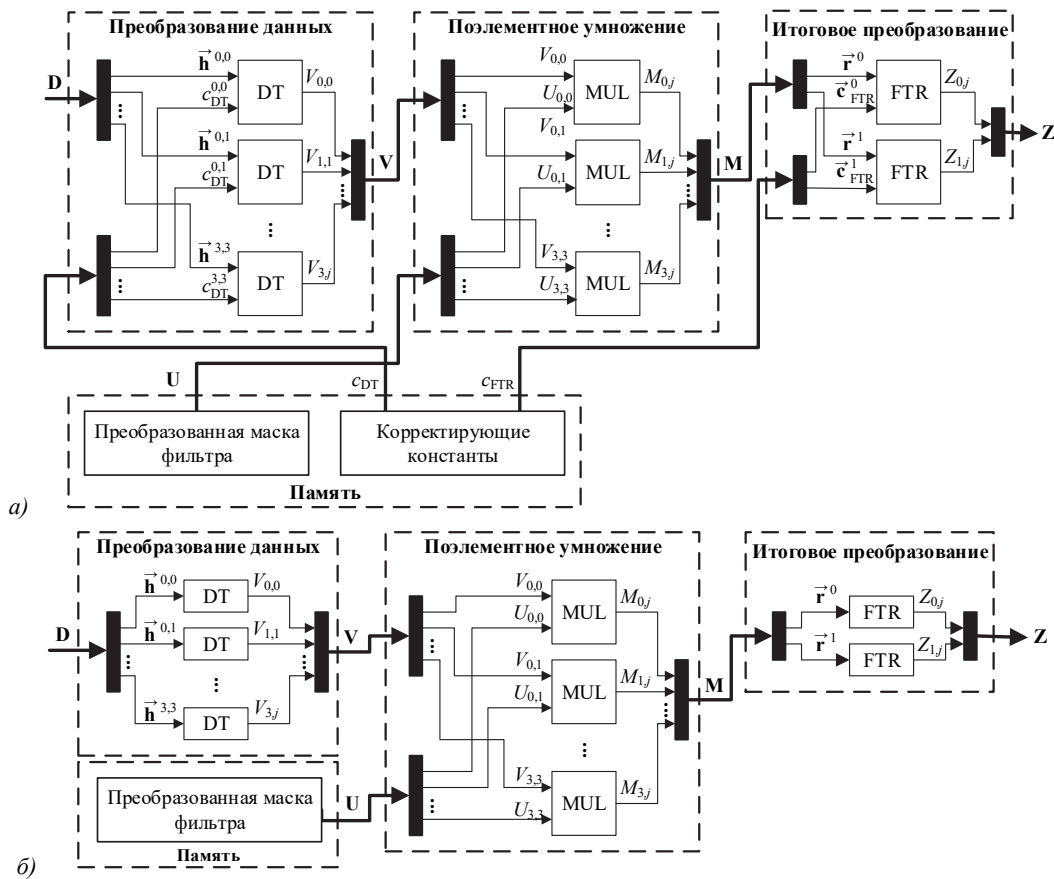


Рис. 9. Схема устройства для двумерной фильтрации с использованием модифицированного метода Винограда  $F(2 \times 2, 3 \times 3)$ : а) по модулю  $2^\alpha$ ; б) по модулю  $2^\alpha-1$

### 5. Аппаратное моделирование предлагаемых устройств фильтрации изображений на основе метода Винограда

Для проверки эффективности предлагаемых устройств фильтрации изображений по методу Винограда было проведено аппаратное моделирование. Для эксперимента использован фильтр Гаусса размером  $3 \times 3$ :

$$W = \frac{1}{15} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{bmatrix}. \tag{16}$$

Как видно из равенства (16), коэффициенты маски фильтра являются вещественными числами. В соответствии с методом масштабирования [21] исходные коэффициенты фильтра умножаются на  $2^l$ , где

$\gamma = 10 + t$ . В данном случае  $t = 1$  и  $\gamma = 11$ . Для фильтра Гаусса размером  $3 \times 3$  маска фильтра с масштабированными коэффициентами имеет вид:

$$\tilde{W} = \begin{bmatrix} 137 & 274 & 137 \\ 274 & 470 & 274 \\ 137 & 274 & 137 \end{bmatrix}. \quad (17)$$

Таким образом, для обработки изображения фильтром (17) динамический диапазон должен удовлетворять условию  $P \geq 539070$ , то есть разрядность устройства составляет 20 бит. Для реализации вычислений в СОК был выбран набор модулей специального вида  $\{2^7, 2^7-1, 2^6-1\}$ , динамический диапазон которого равен  $P = 1024128$  и удовлетворяет условию (13).

5.1. Теоретический анализ технических параметров предлагаемых устройств фильтрации изображений на основе метода Винограда

Для теоретической оценки задержки и площади цифрового устройства используем «unit-gate»-модель [32]. Основываясь на данной абстрактной модели, задержка  $U_{delay}$  и площадь  $U_{area}$  логических вентилей представляются следующим образом:

$$\begin{aligned} U_{delay}(NOT) &= 0, U_{area}(NOT) = 0; \\ U_{delay}(AND) &= 1, U_{area}(AND) = 1; \\ U_{delay}(OR) &= 1, U_{area}(OR) = 1; \\ U_{delay}(XOR) &= 2, U_{area}(XOR) = 2; \\ U_{delay}(XNOR) &= 2, U_{area}(XNOR) = 2. \end{aligned} \quad (18)$$

На основе параметров логических вентилей могут быть рассчитаны параметры более сложных устройств [15, 20]. Предлагаемые фильтры на основе метода Винограда  $F(2 \times 2, 3 \times 3)$  и СОК с модулями специального вида состоят из устройств DT, MUL и FTR, описанных выше. На основе параметров задержки и площади логических вентилей (18) рассчитаем параметры данных устройств.

Параметры устройства DT по модулю  $2^\alpha$  равны

$$\begin{aligned} U_{delay}(DT) &= 2 \log_2 \alpha + 24, 4, \\ U_{area}(DT) &= 3\alpha \log_2 \alpha + 24\alpha + 1, \end{aligned} \quad (19)$$

а по модулю  $2^{\alpha-1}$

$$\begin{aligned} U_{delay}(DT) &= 2 \log_2 \alpha + 17, 6, \\ U_{area}(DT) &= 3\alpha \log_2 \alpha + 20\alpha. \end{aligned} \quad (20)$$

Параметры устройства MUL по модулю  $2^\alpha$  равны

$$\begin{aligned} U_{delay}(MUL) &= 8, 8 \log_2 \alpha + 0, 5\alpha^2 + 0, 5\alpha + 4, \\ U_{area}(MUL) &= 3\alpha \log_2 \alpha + 7, 5\alpha^2 - 10, 5\alpha + 1, \end{aligned} \quad (21)$$

а по модулю  $2^{\alpha-1}$

$$\begin{aligned} U_{delay}(MUL) &= 8, 8 \log_2 \alpha + \alpha^2 + 4, \\ U_{area}(MUL) &= 3\alpha \log_2 \alpha + 8\alpha^2 - 8\alpha. \end{aligned} \quad (22)$$

Параметры устройства FTR по модулю  $2^\alpha$

$$\begin{aligned} U_{delay}(FTR) &= 2 \log_2 \alpha + 31, 2, \\ U_{area}(FTR) &= 3\alpha \log_2 \alpha + 87\alpha + 1, \end{aligned} \quad (23)$$

а по модулю  $2^{\alpha-1}$

$$\begin{aligned} U_{delay}(FTR) &= 2 \log_2 \alpha + 31, 2, \\ U_{area}(FTR) &= 3\alpha \log_2 \alpha + 76\alpha. \end{aligned} \quad (24)$$

Предлагаемые устройства фильтрации на основе метода Винограда  $F(2 \times 2, 3 \times 3)$  состоят из 16 устройств DT и MUL и из 2 устройств FTR. Таким образом, параметры предлагаемого устройства по модулю  $2^\alpha$ , основываясь на (19), (21) и (23), равны

$$\begin{aligned} U_{delay}(F(2 \times 2, 3 \times 3)) &= \\ &= U_{delay}(DT) + U_{delay}(MUL) + U_{delay}(FTR) = \\ &= 12, 8 \log_2 \alpha + 0, 5\alpha^2 + 0, 5\alpha + 59, 6, \\ U_{area}(F(2 \times 2, 3 \times 3)) &= \\ &= 16U_{area}(DT) + 16U_{area}(MUL) + 2U_{area}(FTR) = \\ &= 102\alpha \log_2 \alpha + 120\alpha^2 + 390\alpha + 34, \end{aligned} \quad (25)$$

а по модулю  $2^{\alpha-1}$ , основываясь на (20), (22) и (24), равны

$$\begin{aligned} U_{delay}(F(2 \times 2, 3 \times 3)) &= \\ &= U_{delay}(DT) + U_{delay}(MUL) + U_{delay}(FTR) = \\ &= 12, 8 \log_2 \alpha + \alpha^2 + 52, 8, \\ U_{area}(F(2 \times 2, 3 \times 3)) &= \\ &= 16U_{area}(DT) + 16U_{area}(MUL) + 2U_{area}(FTR) = \\ &= 102\alpha \log_2 \alpha + 128\alpha^2 + 344\alpha. \end{aligned} \quad (26)$$

Было проведено сравнение предлагаемого устройства фильтрации изображений с устройством для фильтрации по методу Винограда без использования СОК [18]. Кроме того, проведено сравнение с устройством фильтрации, состоящим из блоков умножения с накоплением (MAC) [14], и устройством, состоящим из усеченных блоков умножения с накоплением (ТМАС) без арифметики СОК [15] и с использованием СОК [20].

Параметры устройства на основе метода Винограда без использования арифметики СОК [18] рассчитываются по формулам (25) при условии, что разрядность  $\alpha = 20$  бит.

Параметры фильтра, состоящего из MAC-блоков,  $FIR(MAC)$  [14] рассчитываются следующим образом:



$$\begin{aligned}
 U_{delay}(FIR(MAC)) &= 88 \log_2 \alpha + 50, \\
 U_{area}(FIR(MAC)) &= 30\alpha \log_2 \alpha + 80\alpha^2 - 40\alpha + 10.
 \end{aligned}
 \tag{27}$$

Параметры фильтра, состоящего из ТМАС-блоков,  $FIR(MAC)$  [15] вычисляются следующим образом

$$\begin{aligned}
 U_{delay}(FIR(TMAC)) &= 70 \log_2 \alpha + 14, \\
 U_{area}(FIR(TMAC)) &= 3\alpha \log_2 \alpha + 80\alpha^2 + 3\alpha + 1,
 \end{aligned}
 \tag{28}$$

для рассматриваемого случая разрядность  $\alpha = 20$  бит.

Параметры фильтра  $FIR(MAC)$  с вычислениями в СОК [20] по модулю  $2^\alpha$  вычисляются по формуле (29), а по модулю  $2^\alpha - 1$  равны

$$\begin{aligned}
 U_{delay}(FIR(TMAC)) &= 70 \log_2 \alpha + 14, \\
 U_{area}(FIR(TMAC)) &= 3\alpha \log_2 \alpha + 80\alpha^2 + 6\alpha.
 \end{aligned}
 \tag{29}$$

Для реализации вычислений в СОК был выбран набор модулей специального вида  $\{2^7, 2^7 - 1, 2^6 - 1\}$ .

В табл. 2 представлены результаты теоретической оценки параметров площади и задержки предложенного и известных фильтров на основе «unit-gate»-модели. Так же была произведена оценка времени обработки изображения размером  $256 \times 256$ .

Табл. 1. Теоретические значения технических параметров устройств на основе «unit-gate»-модели

Метод	Известные методы				Предлагаемый метод
	[14]	[15]	[20]	[18]	
Задержка	431	317	211	325	<b>138</b>
Площадь	33804	32321	<b>10985</b>	64651	29587
Время обработки изображения	$2,8 \cdot 10^7$	$2,1 \cdot 10^7$	$1,4 \cdot 10^7$	$5,3 \cdot 10^6$	$2,3 \cdot 10^6$

Анализ и обсуждение полученных теоретических результатов изложены в следующем параграфе.

### 5.2. Результаты аппаратного моделирования предлагаемых устройств фильтрации изображений на основе метода Винограда

Аппаратное моделирование на FPGA проводилось в среде автоматизированного проектирования Xilinx Vivado 2018.3 для целевой платы Artix-7 xc7a200tffg1156-3 со стратегией оптимизации Flow\_PerfOptimized\_high. Результаты аппаратного моделирования устройств для фильтрации изображений представлены в табл. 2. Качество обработки изображений всеми рассмотренными фильтрами одинаково. Для оценки устройств использовались такие параметры, как тактовая частота, количество занятых просмотрных таблиц (Look-up Table, LUT), энергопотребление и производительность, которые были получены в результате симуляции в среде проектирования. Под производительностью устройства подразумевается количество обработанных кадров размером  $256 \times 256$  пикселей в секунду. Для комплексной оценки временных и аппаратных затрат была введена метрика, учитывающая одновременно влияние производительности устройства и его площади:

$$\beta = \frac{P}{A},
 \tag{30}$$

где  $P$  – производительность устройства, а  $A$  – площадь устройства (количество занятых LUT) [33]. Чем больше значение данной метрики, тем более сбалансировано устройство, то есть обладает лучшим соотношением производительности и площади.

Табл. 2. Результаты аппаратного моделирования

Метод	Известные методы				Предлагаемый метод
	[14]	[15]	[20]	[18]	
Тактовая частота, МГц	31	36	<b>56</b>	37	42
Количество LUT	1872	1584	<b>738</b>	3633	3652
Энергопотребление, Вт	0,275	0,255	<b>0,251</b>	0,318	0,319
Производительность, кадр/с	473	549	854	2258	<b>2563</b>
$\beta$	0,253	0,347	<b>1,158</b>	0,622	0,702

Анализ полученных результатов аппаратного моделирования представлен в следующем параграфе.

### 6. Обсуждение

Теоретический анализ на основе «unit-gate»-модели (табл. 1) параметров предлагаемых устройств и известных аналогов показал, что задержка предлагаемого фильтра на основе метода Винограда и арифметики СОК в 1,53–3,12 раз меньше, а время обработки одного изображения размером  $256 \times 256$  в 2,30–12,17 раз меньше по сравнению с другими известными устройствами. Тем не менее фильтр на основе ТМАС-блоков с вычислениями в СОК [20] обладает в 2,69–5,89 раз меньшей площадью по сравнению с другими рассмотренными устройствами.

Результаты аппаратного моделирования предлагаемых устройств фильтрации и известных аналогов (табл. 2) показали, что тактовая частота устройства, разработанного с использованием метода [20], на 33,33–80,64% выше по сравнению с другими рассмотренными методами. Также данный метод требует на 53,41–79,79% меньше LUT и на 1,57–21,32% меньше энергозатрат по сравнению с другими рассмотренными методами. Предлагаемый в данной работе метод обладает в 1,13–5,42 раз большей производительностью по сравнению с другими рассмотренными известными методами. Тем не менее метрика  $\beta$  показывает, что фильтр на основе ТМАС-блоков с вычислениями в СОК [20] является в 1,65–4,58 раз более сбалансированным относительно производительности устройства и его площади.

Разница в теоретических параметрах устройств и полученных в результате аппаратного моделирования объясняется тем, что «unit-gate»-модель не учитывает нагрузочную способность устройств, а также задействованную память и время обращения к ней.

На основе результатов аппаратного моделирования можно сделать вывод о том, что предложенное устройство фильтрации изображений обладает высо-

кой производительностью по сравнению с рассмотренными известными методами построения устройств фильтрации за счет увеличения аппаратных ресурсов на реализацию. Высокая производительность устройств на основе метода Винограда объясняется тем, что на выходе формируется сразу четыре пикселя обработанного изображения. Использование СОК с модулями специального вида позволило увеличить скорость устройства фильтрации по сравнению с другими устройствами, использующими ПСС. Таким образом, предлагаемые архитектуры фильтрации изображений могут быть применены в системах, где ключевую роль играет производительность.

Предложенные архитектуры устройств могут быть использованы для фильтров с масками размером  $3 \times 3$  и коэффициентами произвольного вида. Поэтому при проведении эксперимента не учитывалась симметрия фильтра Гаусса. В дальнейших исследованиях планируется учесть особенности масок фильтров, например, симметрию или наличие нулевых коэффициентов, что позволит упростить архитектуру устройства фильтрации и еще больше повысить его производительность.

Перспективными направлениями дальнейших исследований являются: разработка устройств фильтрации изображений методом Винограда в СОК для других случаев  $F(n \times n, k \times k)$ , а также их внедрение в специальные процессоры обработки сигналов и системы нейросетевого анализа изображений с целью увеличения их производительности.

### Заключение

В работе представлены архитектуры устройств фильтрации изображений методом Винограда  $F(2 \times 2, 3 \times 3)$  с использованием СОК с модулями специального вида  $2^a$  и  $2^a - 1$ . Было проведено аппаратное моделирование устройства, спроектированного предлагаемым методом, и устройств, разработанных известными методами, для сглаживающего фильтра Гаусса с маской  $3 \times 3$ . Результаты моделирования показали, что предлагаемые архитектуры обладают высокой производительностью по сравнению с известными подходами за счет увеличения аппаратных затрат. Результаты данного исследования могут быть использованы в системах обработки и анализа изображений для увеличения производительности устройств. Применение предложенных устройств позволит проектировать сложные устройства обработки изображений, работающие в режиме реального времени.

### Благодарности

Работа выполнена в Северо-Кавказском центре математических исследований в рамках соглашения №. 075-02-2022-892 с Министерством науки и высшего образования Российской Федерации.

### References

- [1] Li H, Fan J, Yu K, Qi X, Wen Z, Hua Q, Zhang M, Zheng Q. Medical image coloring based on gabor filtering for internet of medical things. *IEEE Access* 2020; 8: 104016-104025. DOI: 10.1109/ACCESS.2020.2999454.
- [2] Kamirul K, Hasbi W, Hakim PR, Syafrudin AH. Automatic ship recognition chain on satellite multispectral imagery. *IEEE Access* 2020; 8: 221918-221931. DOI: 10.1109/ACCESS.2020.3042702.
- [3] Wang G, Li B, Zhang Y, Yang J. Background modeling and referencing for moving cameras-captured surveillance video coding in HEVC. *IEEE Trans Multimedia* 2018; 20(11): 2921-2934. DOI: 10.1109/TMM.2018.2829163.
- [4] Alonzo LMB, Chioson FB, Co HS, Bugtai NT, Baldovino RG. A machine learning approach for coconut sugar quality assessment and prediction. 2018 IEEE 10th Int Conf on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM) 2018; 1-4. DOI: 10.1109/HNICEM.2018.8666315.
- [5] Thanh DNH, Hien NN, Kalavathi P, Surya Prasath VB. Adaptive switching weight mean filter for salt and pepper image denoising. *Procedia Computer Science* 2020; 171: 292-301. DOI: 10.1016/j.procs.2020.04.031.
- [6] Minnema J, Wolff J, Koivisto J, Lucka F, Batenburg KJ, Forouzanfar T, Eijnatten M. Comparison of convolutional neural network training strategies for conebeam CT image segmentation. *Comput Methods Programs Biomed* 2021; 207: 106192. DOI: 10.1016/j.cmpb.2021.106192.
- [7] Hespeler SC, Hespeler SC, Nemati H, Dehghan-Niri E. Non-destructive thermal imaging for object detection via advanced deep learning for robotic inspection and harvesting of chili peppers. *Artificial Intelligence in Agriculture* 2021; 5: 102-117. DOI: 10.1016/j.aiaa.2021.05.003.
- [8] Chervyakov NI, Lyakhov PA, Nagornov NN, Valueva MV, Valuev GV. Hardware implementation of a convolutional neural network using calculations in the residue number system. *Computer Optics* 2019; 43(5): 857-868. DOI: 10.18287/2412-6179-2019-43-5-857-868.
- [9] Chriskos P, Frantzidis CA, Gkivogkli PT, Bamidis PD, Kourtidou-Papadeli C. Automatic sleep staging employing convolutional neural networks and cortical connectivity images. *IEEE Trans Neural Netw Learn Syst* 2020; 31(1): 113-123. DOI: 10.1109/TNNLS.2019.2899781.
- [10] Tian Y. Artificial intelligence image recognition method based on convolutional neural network algorithm. *IEEE Access* 2020; 8: 125731-125744. DOI: 10.1109/ACCESS.2020.3006097.
- [11] Gonzalez RC, Woods RE. *Digital image processing*. 3rd ed. Prentice Hall; 2008.
- [12] Bovik AC. *Handbook of image and video processing*, 2nd ed. San Diego: Elsevier Academic Press; 2005.
- [13] Nakasone S, Öfverstedt L, Wilken G, Skoglund U. An OpenCL implementation of an image filter on FPGA. 2019 IEEE 5th Int Conf on Computer and Communications (ICCC) 2019; 272-276. DOI: 10.1109/ICCC47050.2019.9064160.
- [14] Tung CA, Huang S. High-performance multiply-accumulate unit by integrating additions and accumulations into partial product reduction process. *IEEE Access* 2020; 8: 87367-87377. DOI: 10.1109/ACCESS.2020.2992286.
- [15] Lyakhov P, Valueva M, Valuev G, Nagornov N. A method of increasing digital filter performance based on truncated multiply-accumulate units. *Appl Sci* 2020; 10: 9052. DOI: 10.3390/app10249052.
- [16] Winograd S. Arithmetic complexity of computations. *SIAM*; 1980: 33. DOI: 10.1137/1.9781611970364.

- [17] Lavin A, Gray S. Fast algorithms for convolutional neural networks. 2016 IEEE Conf on Computer Vision and Pattern Recognition (CVPR) 2016: 4013-4021. DOI: 10.1109/CVPR.2016.435.
- [18] Yepez J, Ko S. Stride 2 1-D, 2-D, and 3-D Winograd for convolutional neural networks. IEEE Trans Very Large Scale Integr VLSI Syst 2020; 28(4): 853-863. DOI: 10.1109/TVLSI.2019.2961602.
- [19] Molahosseini AS, De Sousa LS, Chang CH, eds. Embedded systems design with special arithmetic and number systems. Springer International Publishing; 2017.
- [20] Lyakhov P, Valueva M, Valuev G, Nagornov N. High-performance digital filtering on truncated multiply-accumulate units in the residue number system. IEEE Access 2020; 8: 209181-209190. DOI: 10.1109/ACCESS.2020.3038496.
- [21] Chervyakov NI, Lyakhov PA, Nagornov NN, Kaplun DI, Voznesenskiy AS, Bogayevskiy DV. Implementation of smoothing image filtering in the residue number system. 2019 8th Mediterranean Conference on Embedded Computing (MECO) 2019: 1-4. DOI: 10.1109/MECO.2019.8760190.
- [22] Vassalos E, Bakalis D, Vergos HT. RNS assisted image filtering and edge detection. 2013 18th Int Conf on Digital Signal Processing (DSP) 2013: 1-6. DOI: 10.1109/ICDSP.2013.6622821.
- [23] Cardarilli GC, Di Nunzio L, Fazzolari R, Nannarelli A, Petricca M, Re M. Design space exploration based methodology for residue number system digital filters implementation. IEEE Trans Emerg Topics Comput 2020; 10(1): 186-198. DOI: 10.1109/TETC.2020.2997067.
- [24] Belghadr A, Jaberipur G. FIR filter realization via deferred end-around carry modular addition. IEEE Trans Circuits Syst I Regul Pap 2018; 65(9): 2878-2888. DOI: 10.1109/TCSI.2018.2798595.
- [25] Omondi A, Premkumar B. Residue number systems: Theory and implementation. London: Imperial College Press; 2007.
- [26] Vergos HT, Dimitrakopoulos G. On Modulo  $2^n+1$  adder design. IEEE Trans Comput 2012; 61(2): 173-186. DOI: 10.1109/TC.2010.261.
- [27] Mohan PVA, Premkumar AB. RNS-to-Binary Converters for Two Four-Moduli Sets  $\{2^n-1, 2^n, 2^{n+1}, 2^{n+1}-1\}$  and  $\{2^n-1, 2^n, 2^{n+1}, 2^{n+1}+1\}$ . IEEE Trans Circuits Syst I Regul Pap 2007; 54(6): 1245-1254. DOI: 10.1109/TCSI.2007.895515.
- [28] Živaljević D, Stamenković N, Stojanović V. Digital filter implementation based on the RNS with diminished-1 encoded channel. 2012 35th Int Conf on Telecommunications and Signal Processing (TSP) 2012: 662-666. DOI: 10.1109/TSP.2012.6256380.
- [29] Chervyakov NI, Lyakhov PA, Kalita DI, Shulzhenko KS. Effect of RNS dynamic range on grayscale images filtering. 2016 XV Int Symp Problems of Redundancy in Information and Control Systems (REDUNDANCY) 2016: 33-37. DOI: 10.1109/RED.2016.7779323.
- [30] Parhami B. Computer arithmetic: algorithms and hardware designs. Oxford University Press; 2010.
- [31] Kogge PM, Stone HS. A parallel algorithm for the efficient solution of a general class of recurrence equations. IEEE Trans Comput 1973; C-22(8): 786-793. DOI: 10.1109/TC.1973.5009159.
- [32] Zimmermann R. Binary adder architectures for cell-based VLSI and their synthesis. Hartung-Gorre; 1998.
- [33] Valueva MV, Nagornov NN, Lyakhov PA, Valuev GV, Chervyakov NI. Application of the residue number system to reduce hardware costs of the convolutional neural network implementation. Math Comput Simul 2020; 177: 232-243. DOI: 10.1016/j.matcom.2020.04.031.

#### Сведения об авторах

**Валуева Мария Васильевна**, 1993 года рождения, окончила бакалавриат и магистратуру по специальности «Прикладная математика и информатика» в ФГАОУ ВО «Северо-Кавказский федеральный университет». Аспирант ФГАОУ ВО «Северо-Кавказский федеральный университет». Работает младшим научным сотрудником в отделе теоретико-числовых систем Северо-Кавказского центра математических исследований, ФГАОУ ВО «Северо-Кавказский федеральный университет». Область научных интересов: цифровая обработка изображений, искусственный интеллект, модулярная арифметика, цифровые устройства. E-mail: [mriya.valueva@mail.ru](mailto:mriya.valueva@mail.ru).

**Ляхов Павел Алексеевич**, 1988 года рождения, в 2009 году окончил Ставропольский государственный университет по специальности «Математика», кандидат физико-математических наук, доцент. Работает заведующим кафедрой математического моделирования ИМЕН ФГАОУ ВО «Северо-Кавказский федеральный университет». Область научных интересов: цифровая обработка сигналов и изображений, искусственный интеллект, модулярная арифметика, цифровые устройства, высокопроизводительные вычисления. E-mail: [ljahov@mail.ru](mailto:ljahov@mail.ru).

**Нагорнов Николай Николаевич**, 1992 года рождения, окончил бакалавриат и магистратуру по специальности «Прикладная математика и информатика» в ФГАОУ ВО «Северо-Кавказский федеральный университет», кандидат технических наук. Работает младшим научным сотрудником в ФГАОУ ВО «Северо-Кавказский федеральный университет». Область научных интересов: цифровая обработка изображений, искусственный интеллект, модулярная арифметика. E-mail: [sparta1392@mail.ru](mailto:sparta1392@mail.ru).

**Валуев Георгий Вячеславович**, 1992 года рождения, окончил магистратуру по специальности «Прикладная информатика» в ФГАОУ ВО «Северо-Кавказский федеральный университет». Аспирант ФГАОУ ВО «Северо-Кавказский федеральный университет». Работает младшим научным сотрудником в отделе теоретико-числовых систем Северо-Кавказского центра математических исследований, ФГАОУ ВО «Северо-Кавказский федеральный университет». Область научных интересов: цифровая обработка изображений, искусственный интеллект, высокопроизводительные вычисления. E-mail: [mail@gvvaluev.ru](mailto:mail@gvvaluev.ru).

ГРНТИ: 28.23.15

Поступила в редакцию 2 июня 2021 г. Окончательный вариант – 16 декабря 2021 г.

---

# High-performance digital image filtering architectures in the residue number system based on the Winograd method

M.V. Valueva<sup>1</sup>, P.A. Lyakhov<sup>1</sup>, N.N. Nagornov<sup>1</sup>, G.V. Valuev<sup>1</sup>  
<sup>1</sup> FSAEI HE "North-Caucasus Federal University",  
355017, Stavropol, Russia, Pushkina 1

## Abstract

Continuous improvement of methods for visual information registration, processing and storage leads to the need of improving technical characteristics of digital image processing systems. The paper proposes new high-performance digital filter architectures for image processing by the Winograd method with calculations performed in a residue number system with special-type moduli. To assess the performance and hardware costs of the proposed architectures, hardware simulation is carried out using a field-programmable gate array in a computer-aided design environment Xilinx Vivado 2018.3 for the target device Artix-7 xc7a200tffg1156-3. The results of hardware simulation show that the proposed filter architectures have 1.13–5.42 times higher performance, but require more hardware costs compared to the known methods. The results of this study can be used in the design of complex systems for image processing and analysis for their performance to be increased.

**Keywords:** digital filters, image processing, residue number system, Winograd filtering method, field-programmable gate array.

**Citation:** Valueva MV, Lyakhov PA, Nagornov NN, Valuev GV. High-performance digital image filtering architectures in the residue number system based on the Winograd method. *Computer Optics* 2022; 46(5): 752-762. DOI: 10.18287/2412-6179-CO-933.

**Acknowledgements:** The work is supported by the North-Caucasus Center for Mathematical Research under agreement No. 075-02-2021-1749 with the Ministry of Science and Higher Education of the Russian Federation.

---

## Authors' information

**Maria Vasilyevna Valueva** (b. 1993) graduated from bachelor's and master's degrees in "Applied Mathematics and Computer Science" at the FSAEI HE "North Caucasus Federal University". She is a postgraduate student in FSAEI HE "North Caucasus Federal University". Currently she works as the junior researcher of Number-Theoretic Systems department of the North-Caucasus Center for Mathematical Research, FSAEI HE "North Caucasus Federal University". Research interests are digital image processing, artificial intelligent, modular arithmetic, digital devices. E-mail: [mriva.valueva@mail.ru](mailto:mriva.valueva@mail.ru).

**Pavel Alekseyevich Lyakhov**, (b. 1988), graduated from Stavropol State University, specialty "Mathematics" in 2009. He is PhD of Physical and Mathematical Sciences and an associate professor. Currently he works as a head of Mathematical Modeling department in FSAEI HE "North Caucasus Federal University". Research interests: digital signal and image processing, artificial intelligent, modular arithmetic, digital devices, high-performance computing. E-mail: [ljahov@mail.ru](mailto:ljahov@mail.ru).

**Nikolay Nikolaevich Nagornov** (b. 1992) graduated from bachelor's and master's degrees in "Applied Mathematics and Computer Science" at the FSAEI HE "North Caucasus Federal University". He is PhD of Computer Sciences. Currently he works as the junior researcher at FSAEI HE "North Caucasus Federal University". Research interests are digital image processing, artificial intelligent, modular arithmetic. E-mail: [sparta1392@mail.ru](mailto:sparta1392@mail.ru).

**Georgii Vyacheslavovich Valuev** (b. 1992) graduated from master's degrees in "Applied Computer Science" at the FSAEI HE "North Caucasus Federal University". He is a postgraduate student in FSAEI HE "North Caucasus Federal University". He works as an junior researcher at Number-Theoretic Systems department of the North-Caucasus Center for Mathematical Research, FSAEI HE "North Caucasus Federal University". Research interests are digital image processing, artificial intelligent, high-performance computing. E-mail: [mail@gvvaluev.ru](mailto:mail@gvvaluev.ru).

---

*Received June 2, 2021. The final version – December 16, 2021.*

---