

Распределенная обработка изображений на основе однотипных IP-ядер в архитектуре ПЛИС/FPGA

В.М. Захаров¹, С.В. Шалагин¹, Б.Ф. Эминов¹

¹Казанский национальный исследовательский технический университет имени А.Н. Туполева, Карла Маркса 10, Казань, Россия, 420111

Аннотация. Решается задача обработки изображений, двумерных массивов данных, за счет реализации двумерного быстрого преобразования Фурье (БПФ) при использовании однотипных аппаратных модулей – IP-ядер в архитектуре ПЛИС семейства Virtex-6. Показана возможность параллельной реализации каждого из этапов двумерного БПФ на основе четырех преобразований вида «бабочка» (ПрБ) над четырьмя элементами обрабатываемого массива данных. Получены оценки временной и аппаратной сложности IP-ядра, реализующего ПрБ, которое применяется и при реализации одномерного БПФ. Результаты могут быть использованы для оценки аппаратных и временных затрат при выполнении двумерного БПФ над массивом заданной размерности при использовании распределенных вычислительных систем с программируемой архитектурой, как существующих, так и перспективных.

1. Введение

В настоящее время актуальна задача обработки изображений в реальном масштабе времени. Возможности программной реализации алгоритмов, реализующих данную задачу на ЭВМ общего назначения, ограничены возможностями фон-неймановской архитектуры. Выходом из сложившейся ситуации является применение специализированных ЭВМ. В частности – ЭВМ, включающих в свой состав аппаратные ускорители как ASIC, так и ПЛИС класса FPGA.

В работе рассмотрена распределенная реализация двумерного быстрого преобразования Фурье (БПФ) на основе однотипных IP-ядер, выполненных в архитектуре ПЛИС/FPGA. На основе оценок аппаратной сложности и времени задержки функционирования IP-ядер получены оценки аппаратной и временной сложности выполнения двумерного БПФ для изображения заданной размерности.

Известен алгоритм вычисления двумерного БПФ на основе процедуры одномерного БПФ [1, 2]. Недостатком данного алгоритма является то, что он выполняется в два этапа. На первом этапе вычисляются образы для строк, затем, на втором этапе, – для столбцов. Или наоборот, на первом этапе – для столбцов, на втором – для строк. В любом случае, пока не выполнены все операции для первого этапа, невозможно перейти к выполнению операций на втором этапе. Предлагаемый в работе алгоритм лишен данного недостатка: операции выполняются над элементами двумерного массива данных параллельно, без разделения на этапы.

Полученные результаты применимы для решения задач распределенной обработки изображений в режиме реального времени с применением специализированных графических ускорителей на основе ПЛИС/FPGA, как существующих, так и перспективных.

2. Двумерное быстрое преобразование Фурье

Рассмотрим двумерное дискретное преобразование Фурье (ДПФ) для массива чисел $\{x_{mn}\} = X$ размерности $N \times N$, $N = 2^a$. Оно представимо следующим образом [1, 2]:

$$G_{uv} = N^{-2} \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} x_{mn} \cdot W_N^{mu+nv}, \quad (1)$$

где $W_N^{mu+nv} = \exp\left(-2\pi j \frac{mu+nv}{N}\right)$, $u = \overline{0, N-1}$, $v = \overline{0, N-1}$, $\{G_{uv}\}_{N \times N} = G$ – образ для X . По аналогии выполняется и обратное двумерное ДПФ:

$$x_{mn} = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} G_{uv} \cdot W_N^{mu+nv}, \quad m = \overline{0, N-1}, \quad n = \overline{0, N-1}.$$

Система вида (1) представима, по аналогии с одномерным БПФ, в виде:

$$G_{uv} = N^{-2} \left[S_{2n_1, 2m_1} + W_N^u \cdot S_{2n_1+1, 2m_1} + W_N^v \cdot S_{2n_1, 2m_1+1} + W_N^{u+v} \cdot S_{2n_1+1, 2m_1+1} \right], \quad (2)$$

где $u = \overline{0, N-1}$, $v = \overline{0, N-1}$, $S_{2n_1, 2m_1} = \sum_{n_1=0}^{N/2-1} \sum_{m_1=0}^{N/2-1} x_{2n_1, 2m_1} \cdot W_N^{2(n_1u+m_1v)}$.

В результате, двумерное БПФ реализуемо на основе однотипных операций, выполняемых над элементами матриц D размерности $2^d \times 2^d$, подматриц $\{x_{nm}\}$, $m = \overline{0, N-1}$, $n = \overline{0, N-1}$, $d = \overline{1, a}$, $a = \log_2 N$:

$$\begin{aligned} D_{u,v} &= \left[D_{u,v} + W_N^u \cdot D_{u+2^{d-1}, v} + W_N^v \cdot D_{u, v+2^{d-1}} + W_N^{u+v} \cdot D_{u+2^{d-1}, v+2^{d-1}} \right], \\ D_{u+2^{d-1}, v} &= \left[D_{u,v} - W_N^u \cdot D_{u+2^{d-1}, v} + W_N^v \cdot D_{u, v+2^{d-1}} - W_N^{u+v} \cdot D_{u+2^{d-1}, v+2^{d-1}} \right], \\ D_{u, v+2^{d-1}} &= \left[D_{u,v} + W_N^u \cdot D_{u+2^{d-1}, v} - W_N^v \cdot D_{u, v+2^{d-1}} - W_N^{u+v} \cdot D_{u+2^{d-1}, v+2^{d-1}} \right], \\ D_{u+2^{d-1}, v+2^{d-1}} &= \left[D_{u,v} - W_N^u \cdot D_{u+2^{d-1}, v} - W_N^v \cdot D_{u, v+2^{d-1}} + W_N^{u+v} \cdot D_{u+2^{d-1}, v+2^{d-1}} \right], \end{aligned} \quad (3)$$

где $u = \overline{0, 2^{d-1}-1}$, $v = \overline{0, 2^{d-1}-1}$.

Замечание. Количество матриц D размерности $2^d \times 2^d$, определено как 4^{a-d} , $d = \overline{1, a}$, $a = \log_2 N$.

Систему вида (3) представим в матричном виде как комплекс однотипных преобразований «бабочка» (ПрБ), применяемых при вычислении одномерного БПФ:

$$\begin{aligned} \begin{pmatrix} D_{u,v} \\ D_{u, v+2^{d-1}} \end{pmatrix} &= \begin{pmatrix} 1 & W_N^v \\ 1 & -W_N^v \end{pmatrix} \cdot \begin{pmatrix} A_1 \\ B_1 \end{pmatrix}, & \begin{pmatrix} D_{u+2^{d-1}, v} \\ D_{u+2^{d-1}, v+2^{d-1}} \end{pmatrix} &= \begin{pmatrix} 1 & W_N^v \\ 1 & -W_N^v \end{pmatrix} \cdot \begin{pmatrix} A_2 \\ B_2 \end{pmatrix}, \\ \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} &= \begin{pmatrix} 1 & W_N^u \\ 1 & -W_N^u \end{pmatrix} \cdot \begin{pmatrix} D_{u,v} \\ D_{u+2^{d-1}, v} \end{pmatrix}, & \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} &= \begin{pmatrix} 1 & W_N^u \\ 1 & -W_N^u \end{pmatrix} \cdot \begin{pmatrix} D_{u, v+2^{d-1}} \\ D_{u+2^{d-1}, v+2^{d-1}} \end{pmatrix}, \end{aligned} \quad (4)$$

В результате, для выполнения одной операции вида (3) над четырьмя элементами массива чисел X , требуется, согласно (4), четыре ПрБ. Указанные операции требуется выполнить над массивом из N^2 элементов $\log_2 N$ раз. Согласно (3) и (4) справедливы

Утверждение 1. Для реализации двумерного БПФ над массивом чисел X размерности $N \times N$ требуется выполнить $a \cdot N^2/4$ операций вида (3) или $a \cdot N^2$ ПрБ.

Утверждение 2. Однотипные операции вида (4) над массивом чисел X размерности $N \times N$ возможно выполнить параллельно.

Утверждения 1 и 2 обосновывают возможность применения специализированных вычислителей, реализующих однотипную операцию, представленную в виде (3) или (4).

3. Сложность реализации на ПЛИС однопипного преобразования «бабочка»

Рассмотрим реализацию аппаратного модуля, реализующего ПрБ (далее – Модуля) как IP-ядра в архитектуре ПЛИС класса FPGA. Модуль позволяет оперировать с 18-разрядными числами. Функционально-логическая схема Модуля, реализованного в схематехническом редакторе САПР ISE Design Suite 14,7 (Xilinx, Inc.) включает в себя:

- блок умножения комплексного числа на константу (ComplexNumMultiplier),
- четыре сумматора, два из которых выполняют операцию сложения, остальные – операцию вычитания в обратном коде,
- регистры для обеспечения конвейерной обработки данных.

Дополнительный разряд регистра служит для хранения знакового разряда.

Схема ComplexNumMultiplier включает четыре умножителя (MULT18x18), два сумматора на 16 разрядов, один из которых реализует операцию вычитания, а также шесть регистров на 16 разрядов для хранения промежуточных результатов.

Модуль реализован на ПЛИС XC6VLX240t-1FF1156 (семейство Virtex – 6, Xilinx, Inc.), которая включает в свой состав D-триггеры, генераторы булевых функций от шести переменных, (LUTs) и блоки ввода-вывода (БВВ).

При реализации Модуля задействовано:

- 210 из 301440 доступных D-триггеров (менее 0,1 %);
- 199 из 150720 LUTs (примерно 0,13%), из которых 161 использованы для реализации булевых функций, а 38 – как межсоединения;
- 83 из 37680 Slices (примерно 0,22%), включающих в себя по четыре LUTs, реализующих булевы функции от пяти или шести переменных и восемь D-триггеров;
- 130 из 600 БВВ (примерно 21,7%), из которых по 64 выделено на входы и на выходы, а два – под синхросигнал и сброс триггеров, соответственно.

Максимальное время задержки функционирования Модуля – 10,438 нс. Вычисление значений на выходах производится через три такта с момента поступления соответствующих значений на вход.

4. Обсуждение результатов

Согласно [3, 4], для приближенной к оптимальной реализации комбинационной схемы на ПЛИС/FPGA требуется задействование не более 0,5 ресурсов каждого из видов (D-триггеров, LUTs, Slices и блоков ввода-вывода).

Согласно полученным оценкам сложности, на одной ПЛИС семейства Virtex – 6 возможно разместить не более двух Модулей. Фактором ограничения является количество БВВ.

Для вычисления двумерного БПФ для массива чисел размерности $N \times N$ требуется $a \cdot N^2 / 4$ ПЛИС семейства Virtex – 6. Причем для вычисления операции, описываемой согласно (4), требуется одна ПЛИС указанного семейства, каждая из которых принимает на вход и выдает по четыре элемента ($D_{u,v}$, $D_{u+2^{d-1},v}$, $D_{u,v+2^{d-1}}$ и $D_{u+2^{d-1},v+2^{d-1}}$) и реализует четыре ПрБ. Оценка частоты работы распределенной вычислительной системы с программируемой архитектурой (РВС ПА) [5], включающей в себя данное количество ПЛИС, составляет не менее 95,8 МГц. Количество тактов задержки вычисления значений элементов $D_{u,v}$, $D_{u+2^{d-1},v}$, $D_{u,v+2^{d-1}}$ и $D_{u+2^{d-1},v+2^{d-1}}$ согласно (4) равняется ba с момента поступления на вход обрабатываемых элементов. В общем случае справедливо

Утверждение 3. Для обработки массива данных размера $N \times N$ на РВС ПА, включающей в свой состав P ПЛИС класса FPGA, принимающих на обработку за один такт q элементов и реализующие преобразование вида (4) за T тактов длительностью z единиц времени каждый, требуется не менее $\log N \cdot \left(\lceil N^2 / (q \cdot P) \rceil + T - 1 \right) \cdot z$ единиц времени.

Например, найдем нижнюю оценку времени, требуемого для обработки массива чисел размера $N \times N$, $N = 2^{10} = 1024$. Двумерное БПФ производим в 10 этапов, на каждом из которых

обрабатывается $N \times N = 2^{20}$ элементов. Каждая ПЛИС семейства Virtex – 6 принимает на обработку по четыре элемента массива, т.е. $q = 4$. Если в РВС ПА P указанных ПЛИС, то этап вычисляется за $\lceil N^2 / (4 \cdot P) \rceil + 5$ такта. Если принять количество ПЛИС/ Virtex – 6 равным 512, как в современных РВС ПА [5], то один этап двумерного БПФ реализуется за 517 тактов по 10,438 нс. каждый, а 10 этапов – за 5170 тактов. Что составляет примерно 5,40 мкс. на обработку одного этапа и примерно 54 мкс. на обработку всего массива. За одну секунду может быть обработано примерно 18,5 тыс. массивов размера 1024 на 1024.

В случае обработки массива 2048 на 2048 на РВС ПА, включающей в свой состав 512 ПЛИС/ Virtex – 6 нижняя оценка времени обработки, согласно утверждению 3, составляет 235 мкс., а за одну секунду может быть обработано примерно 4255 массивов указанного размера.

Сопоставим предложенную реализацию двумерного БПФ с известной его реализацией на основе «одномерных» БПФ [1, 2], которые обозначим как БПФ1, по аппаратным затратам – количеству ПЛИС семейства Virtex – 6. При выполнении N БПФ1 по столбцам массива чисел размерности $N \times N$ требуется $a \cdot N^2 / 2$ ПрБ. Столько же ПрБ требуется для реализации 1БПФ строкам массива чисел. Таким образом, необходимо выполнить $a \cdot N^2$ ПрБ, для чего требуется $a \cdot N^2 / 2$ ПЛИС/ Virtex – 6, каждая из которых реализует по два ПрБ. В результате, выполнение двумерного БПФ на основе предложенного алгоритма требует в два раза меньше ПЛИС семейства Virtex – 6 за счет того, что обрабатываемые данные на входе ПЛИС алгоритмически связаны с данными на выходе согласно (4), что позволяет реализовать на логических ресурсах одной ПЛИС указанного семейства (D -триггерах, LUTs и Slices) четыре ПрБ вместо двух.

За счет параллельно-последовательного ввода массива чисел X внутрь ПЛИС возможно существенно увеличить количество IP-ядер, реализующих Модуль, сконфигурированных на ПЛИС семейства Virtex – 6.

Для Модуля, реализованного на ПЛИС XC6VLX240t-1FF1156, требуется выделить дополнительно регистр на 64 разряда для хранения элементов массива чисел X . Фактором ограничения остается количество Slices, задействованных для реализации Модуля. В этом случае на указанной ПЛИС возможно разместить 227 IP-ядер, реализующих Модуль. Таким образом, на одной ПЛИС может быть реализовано $\lceil 227 / 4 \rceil = 56$ операций вида (4). Но при этом требуется дополнительно $224a$ тактов для последовательного ввода элементов массива X в ПЛИС. Согласно утверждению 3, когда $P = 512$, $q = 4$, $T = 230$ и $z = 10,438$ нс. нижняя оценка времени обработки массива чисел размера $N \times N$, $N = 2^{10} = 1024$ – примерно 77,4 мкс. За одну секунду может быть обработано примерно 12,9 тыс. массивов размера 1024 на 1024. В случае обработки массива 2048 на 2048 нижняя оценка времени данной операции – 238 мкс., а за одну секунду может быть обработано примерно 4202 массива указанного размера.

5. Заключение

На основе оценок временной и аппаратной сложности Модуля как однотипного IP-ядра в архитектуре ПЛИС семейства Virtex-6 вычислены оценки времени задержки функционирования и аппаратной сложности устройства, реализующего конвейерное вычисление двумерного БПФ с прореживанием по времени на РВС ПА. Соответствующие оценки для Модуля выполнены при использовании специализированной САПР ПЛИС/FPGA ISE Design Suite 14.7.

Результаты, полученные в работе, позволяют оценить потенциальные возможности по реализации двумерного БПФ на РВС ПА, как для существующих, так и для перспективных.

6. Литература

- [1] Гонсалес, Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. – М: Техносфера, 2006. – 1072 с.
- [2] Даджион, Д. Цифровая обработка многомерных сигналов / Д. Даджион, Р. Мерсеро. – М.: Мир, 1988. – 488 с.

- [3] Shalagin, S.V. Computer evaluation of a method for combinational-circuit synthesis in FPGAs / S.V. Shalagin // *Russian Microelectronics*. – 2004. – Vol. 33(1). – P. 46-54.
- [4] Шалагин, С.В. Реализация цифровых устройств в архитектуре ПЛИС/FPGA при использовании распределенных вычислений в полях Галуа / С.В. Шалагин. – Казань: изд-во КНИТУ-КАИ, 2016. – 228 с.
- [5] Дордопуло, А.И. Высокопроизводительные реконфигурируемые вычислительные системы / А.И. Дордопуло, И.А. Каляев, И.И. Левин // *Суперкомпьютеры*. – 2010. – Т. 3, № 3. – С. 44-48.

Благодарности

Работа выполнена при финансовой поддержке гранта РФФИ № 18-01-00120 «Специализированные устройства для генерирования и обработки массивов данных в архитектуре программируемых логических интегральных схем класса FPGA».

Авторы выражают благодарность студентам магистратуры КНИТУ-КАИ Бакиеву Л.Ю. и Песошину И.А. за помощь в получении экспериментальных данных.

Distributed image processing based on the same IP-cores in FPGA-architecture

V.M. Zakharov¹, S.V. Shalagin¹, B. Eminov¹

¹Kazan National Research Technical University name after A.N. Tupolev, Karl Marks 10, Kazan, Russia, 420111

Abstract. The problem of image processing, two-dimensional data arrays, to the implementation of two - dimensional fast Fourier transform (FFT) using the same type of hardware modules-IP-cores family Virtex-6 FPGA-architecture is solving. The possibility of parallel implementation of each stage of the two-dimensional FFT on the basis of four transformations of the «butterfly» (TrB) over four elements of the processed data array is shown. Estimates of the time and hardware complexity of the IP-core implementing TrB, which is used in the implementation of one-dimensional FFT, are obtained. The results can be used to estimate hardware and time complexity when performing a two-dimensional FFT over an array of a given dimension using distributed computing systems with programmable architecture, both existing and prospective.