

**А.О. Касьянов, М.В. Потипак**

### **ПРИМЕНЕНИЕ ГРАФИЧЕСКОГО ПРОЦЕССОРА ДЛЯ ЦИФРОВОЙ ОБРАБОТКИ СИГНАЛОВ В РЕАЛЬНОМ МАСШТАБЕ ВРЕМЕНИ**

*Настоящая статья посвящена разработке энергоэффективных реализаций алгоритмов цифровой обработки сигналов в многоканальном радиолокаторе с пространственным разнесением передающих и приемных групп антенных элементов для оценки параметров целей на вычислителях с различной архитектурой. В соответствии общемировой тенденцией рассматривается возможность применения вычислителей с параллельной архитектурой для цифровой обработки широкополосных радиолокационных сигналов. Авторами предложена реализация процедуры обработки отраженного сигнала многоканального радиолокатора с пространственным разнесением передающих и приемных групп антенных элементов с использованием технологии общих вычислений на графических картах (GPGPU). Проведена оценка производительности разработанного решения на различных GPU с разной микроархитектурой. Предложен критерий оценки производительности алгоритма обработки в виде отношения пропускной способности алгоритма к пиковой пропускной способности памяти вычислителя. Проведена численная оценка эффективности использования пропускной способности памяти вычислителя разработанного алгоритма в сравнении с известными реализациями на GPU. Целью работы является обнаружение и оценка параметров целей в реальном масштабе времени с помощью многоканального радиолокатора с пространственным разнесением передающих и приемных групп антенных элементов, используя доступный на рынке вычислитель с минимально возможными массогабаритными характеристиками. Для достижения поставленной цели исследований решены задачи: – выбора и адаптации алгоритмов, позволяющих проводить оценку параметров целей в многоканальном радиолокаторе с пространственным разнесением передающих и приемных групп антенных элементов; – реализации выбранных алгоритмов с учетом архитектуры вычислителя, позволяющих проводить оценку фоновых элементов в реальном масштабе времени; – оценки производительности полученного решения. В процессе разработки алгоритма цифровой обработки сигнала многоканальным радиолокатором с пространственным разнесением передающих и приемных групп антенных элементов проведен анализ нескольких вариантов реализации алгоритма с учетом архитектуры параллельного вычислителя, что позволило выполнить обработку кадра радиоизображения, состоящего из 8 млн. комплексных отсчетов, за время менее 50 мс. на графическом процессоре NVIDIA Jetson AGXXavier. Показана обратная зависимость времени обработки кадра от величины пиковой пропускной способности памяти GPU. Предложен критерий оценки производительности алгоритма обработки. Проведена численная оценка эффективности использования пропускной способности памяти вычислителя разработанного алгоритма в сравнении с известными реализациями на GPU. Выигрыш разработанного алгоритма составляет в среднем 5 раз по сравнению с результатами, полученными другими авторами. По сравнению с ПЛИС реализация двумерного БПФ на GPU оказывается в 17 раз быстрее. Практическая значимость разработанного авторами функционального программного обеспечения не налагает никаких ограничений на количество приемных и передающих каналов и может быть использовано для обработки сигналов в многоканальных радиолокаторах с пространственным разнесением передающих и приемных групп антенных элементов с большим количеством каналов.*

*Параллельные вычисления; реальный масштаб времени; радиолокатор; технология радиолокационных измерений на основе применения множественных излучающих и множественных приемных каналов; технология общих вычислений на графических картах.*

**A.O. Kasyanov, M.V. Potipak**

### **USING A GPU FOR REAL-TIME DIGITAL SIGNAL PROCESSING**

*This paper is devoted to the development of energy-efficient implementations of digital signal processing algorithms in MIMO radar for estimating target parameters on computers with different architectures. In accordance with the global trend, the possibility of using computers with parallel architecture for digital processing of broadband radar signals is being considered. The authors proposed an implementation of the procedure for processing the reflected signal of MIMO radar using the technology of general computing on graphics cards (GPGPU). The performance of the developed solution was assessed on various GPUs*

*with different microarchitectures. A criterion for evaluating the performance of a processing algorithm is proposed in the form of the ratio of the algorithm's throughput to the peak throughput of the computer's memory. A numerical assessment of the efficiency of using the computer's memory bandwidth of the developed algorithm was carried out in comparison with known implementations on the GPU. The purpose of this work is to detect and evaluate target parameters in real time using MIMO radar, using a commercially available computer with the minimum possible weight and size characteristics. To achieve the set research purpose, the following problems were solved: – selection and adaptation of algorithms that allow the assessment of target parameters in MIMO radar; – implementation of selected algorithms taking into account the architecture of the computer, allowing for an assessment of the target background situation in real time; – assessment of the performance of the resulting solution. In the process of developing an algorithm for digital processing of a MIMO radar signal, several options for implementing the algorithm were analyzed taking into account the architecture of a parallel computer, which made it possible to process a radio image frame consisting of 8 million complex samples in less than 50 ms. by NVIDIA Jetson AGXXavier GPU. The inverse relationship between frame processing time and the peak GPU memory bandwidth is shown. A criterion for evaluating the performance of the processing algorithm is proposed. A numerical assessment of the efficiency of using the computer's memory bandwidth of the developed algorithm was carried out in comparison with known implementations on the GPU. The gain of the developed algorithm is on average 5 times compared to the results obtained by other authors. Compared to an FPGA, implementing 2D FFT on a GPU is 17 times faster. The practical significance of the functional software developed by the authors does not impose any restrictions on the number of receiving and transmitting channels and can be used for signal processing in MIMO radars with a large number of channels.*

*Radar; algorithm; parallel computing technology; real-time; Multi Input Multi Output (MIMO); digital signal processing; general-purpose computing on graphics processing units (GPGPU).*

**Введение.** В настоящее время вычислительные средства, способные выполнять цифровую обработку широкополосных сигналов в реальном масштабе времени, находятся на этапе активного развития. Из года в год повышается производительность вычислительных средств при снижении их энергопотребления и стоимости.

Повышение производительности обработки информации достигается двумя путями: экстенсивный – повышение тактовых частот, снижение латентности передачи данных и т.п., и интенсивный – за счёт архитектурных решений при построении вычислительных средств. Первый путь близок к исчерпанию своих возможностей ввиду приближения технологических процессов изготовления электронных компонентов к своим физическим пределам.

Второй, основанный на распараллеливании процессов обработки данных, напротив, набирает все большую популярность и успешно развивается. Настоящая популярность параллельных вычислений обусловлена в том числе и наличием специализированных средств разработки программного обеспечения, облегчающих реализацию алгоритмов на вычислителях с параллельной архитектурой. Немаловажным фактором является доступность специализированных библиотек, максимально эффективно реализующих различные математические функции в зависимости от архитектуры вычислителя. Из имеющихся на рынке коммерческих продуктов следует отметить экосистему NVIDIA, имеющую законченные программно-аппаратные решения на базе технологии реализации общих вычислений на графических картах.

Алгоритмы цифровой обработки сигналов в РЛС для оценки параметров цели достаточно хорошо проработаны. Тем не менее, наибольшую сложность представляет поиск энергоэффективных решений для обработки сигналов в многоканальных РЛС с пространственным разнесением передающих и приемных групп антенных элементов, позволяющих использовать широкую номенклатуру недорогих вычислителей.

**Обзор текущего состояния способов обработки сигналов в многоканальной РЛС.** В работе [1] предложен подход быстрой совместной оценки пеленга и дальности цели на основе нескольких глубоких нейронных сетей для обнаружения целей в сложных условиях. Предложенный алгоритм использует глубокое обучение для достижения точной локализации цели при низком отношении сигнал/шум. Авторы отмечают, что предложенный подход демонстрирует высокую эффективность и точность, однако это достигается при высокой вычислительной нагрузке.

В работах [2, 3] предложен оригинальный способ последовательного формирования матрицы каналов в многоканальной РЛС с пространственным разнесением передающих и приемных групп антенных элементов, позволяющий оценивать дальность до каждой из целей, их скорости и угловые положения целей за одну зондирующую посылку. К достоинствам предложенного подхода можно отнести простоту аппаратной реализации многоканальной РЛС с пространственным разнесением передающих и приемных групп антенных элементов, поскольку используется коммутация одного приемного и одного канала передачи к разным антенным элементам. Недостатком способа является высокая вычислительная нагрузка при оценке параметров целей.

В работе [4] предложен алгоритм оценки угловых координат целей для многоканальной РЛС с непрерывным излучением и пространственным разнесением передающих и приемных групп антенных элементов. Разработанный алгоритм является производным от MUSIC и использует рандомизированную матричную аппроксимацию для снижения вычислительной сложности при сохранении приемлемой точности. Авторы отмечают, что предложенный подход имеет точность, сравнимую с традиционным алгоритмом MUSIC, только при высоких отношениях сигнал-шум (ОСШ).

В работе [5] предлагается новый алгоритм локализации нескольких целей в азимутальной области для многоканальной РЛС с пространственным разнесением передающих и приемных групп антенных элементов, работающей в режиме непрерывного излучения. Предложенный метод позволяет снизить аппаратную и вычислительную нагрузку за счет использования разреженного массива случайно расположенных передающих и приемных антенных элементов, однако при этом снижается потенциальная разрешающая способность по угловой координате.

В работе [6] проведен обзор подходов для оценки параметров целей в многоканальной РЛС с пространственным разнесением передающих и приемных групп антенных элементов. В работе проводится экспериментальное сравнение различных алгоритмов оценки угловых координат целей, таких как пространственное быстрое преобразование Фурье (БПФ), Capon и MUSIC. Авторы отмечают высокую точность алгоритма MUSIC при высоких отношениях сигнал/шум (ОСШ), однако он требует предварительной оценки количества обнаруженных целей. Анализ приведенных в работе дальностно-азимутальных портретов целей показывает, что максимальная контрастность изображений достигается при использовании классического алгоритма пространственного БПФ, что также подтверждается экспериментальными результатами, приведенными в работе [7].

В работе [8] сообщается о реализации модуля обработки сигналов на программируемой логической интегральной схеме (ПЛИС) для РЛС с непрерывным излучением, которая включает в себя вычисление нескольких БПФ для определения дальности, скорости и пеленга целей. Для снижения вычислительной сложности обработки предложены модифицированные алгоритмы оценки скорости и порогового обнаружения.

В работе [9] рассматривается оптимизация стандартных алгоритмов обработки сигналов РЛС с непрерывным излучением с использованием библиотеки Cadence Tensilica для вычислений как с фиксированной, так и с плавающей точкой. Библиотека Cadence Tensilica предназначена для проектирования собственных специализированных вычислителей с последующим изготовлением в виде свербольших интегральных схем (СБИС).

В работах [10, 11] описываются методы обработки 3D данных радиолокационного наблюдения для многоканальной РЛС с пространственным разнесением передающих и приемных групп антенных элементов на графическом процессоре (GPU), реализующие ускорение вычисления трехмерного БПФ. В работе [11] используется высокоуровневая библиотека для ускорения расчетов общего вида на графических ускорителях NVIDIA *ArrayFire*, в то время как в работе [10] использовались библиотека *cuFFT*, входящая в состав NVIDIA CUDA SDK. Предложенные решения позволяют осуществлять обработку отраженных сигналов РЛС в реальном масштабе времени с использованием графической карты NVIDIA RTX A6000 [10] и NVIDIA GeForce RTX 3090 [11].

В работе [12] предложена эффективная реализация БПФ на ПЛИС Xilinx Zynq Ultra Scale+ для обработки радиолокационных сигналов в РЛС с непрерывным излучением. Решение позволяет использовать переменную длину входных данных для расчета БПФ, имеется возможность проводить оконную обработку, накопление, а также расчет амплитуды и фазы отраженного сигнала. Предлагаемая реализация БПФ в 7,32 раза быстрее, чем стандартное IP ядро БПФ Xilinx.

Работа [13] посвящена разработке вычислительного модуля с использованием классического двумерного БПФ для расчета дальностно-скоростного и дальностно-азимутального портретов на ПЛИС (Xilinx).

Проведенный в настоящее время обзор современного состояния подходов к обработке отраженных сигналов в многоканальной РЛС с пространственным разнесением передающих и приемных групп антенных элементов показал, что применяются в основном традиционные алгоритмы: расчет двумерного и трехмерного БПФ для оценки дальности-скорости и дальности-скорости-азимута целей. Это можно объяснить тем, что такие алгоритмы имеют известную временную сложность, и, как правило, готовую реализацию в виде библиотечных модулей, IP-ядер и т.п.

Предложенный Texas Instruments в [14] подход к ЦОС, широко используемый в автомобильных многоканальных РЛС с пространственным разнесением передающих и приемных групп антенных элементов, работающих в режиме непрерывного излучения, с одной стороны, позволяет значительно снизить вычислительную сложность процедуры оценки угловых координат целей, а с другой – имеет существенный недостаток – некогерентное суммирование дальностно-скоростных портретов каждого канала перед процедурой обнаружения, что снижает ОСШ сигналов, отраженных от целей по сравнению с когерентной обработкой.

Большинство современных решений в качестве вычислительной платформы используют ПЛИС, основным недостатком которой является сложность в реализации комплексной процедуры обработки отраженного сигнала ввиду трудоемкости отладки.

Решения, использующие графические карты для обработки сигналов многоканальной РЛС с пространственным разнесением передающих и приемных групп антенных элементов, представлены мощными настольными системами, не всегда подходящими для создания компактных устройств.

Несмотря на достаточно хорошо проработанные подходы обработки отраженных сигналов в многоканальных РЛС с пространственным разнесением передающих и приемных групп антенных элементов и разнообразие алгоритмов, задача разработки эффективных процедур обработки, способных работать в реальном масштабе времени, является актуальной. Разработка подобных процедур сопряжена с необходимостью учитывать особенности архитектуры вычислителя, что в ряде случаев является значимым ограничением в выборе тех или иных алгоритмов обработки.

Целью работы является обнаружение и оценка параметров целей многоканальных РЛС с пространственным разнесением передающих и приемных групп антенных элементов в реальном масштабе времени, используя доступный на рынке вычислитель с минимально возможными массогабаритными характеристиками.

В ходе работы решались следующие задачи:

1. Выбор и адаптация алгоритмов, позволяющих проводить оценку параметров целей в многоканальных РЛС с пространственным разнесением передающих и приемных групп антенных элементов;
2. Реализация выбранных алгоритмов с учетом архитектуры вычислителя, позволяющая проводить оценку фоновых целей в реальном масштабе времени;
3. Оценка производительности полученного решения.

**Постановка задачи.** Рассмотрим задачу в следующей постановке. Пусть в многоканальной РЛС с пространственным разнесением передающих и приемных групп антенных элементов имеется  $N_{Tx}$  передающих и  $N_{Rx}$  приемных каналов. Приемные антенные элементы располагаются на расстоянии  $\lambda/2$  друг от друга, а излучающие – на расстоянии  $N_{Rx} \cdot \lambda/2$  [14]. Антенные элементы могут формировать группы приемных и передающих

элементов в двух плоскостях. Имеется четырехканальный трансивер с двумя парами независимых приемопередающих каналов. На базе трансивера строится малогабаритная РЛС ближнего радиуса действия с непрерывным излучением.

На рис. 1 приведено расположение приемных и передающих антенных элементов для  $N_{Tx} = 2$ ,  $N_{Rx} = 2$ , а также результирующая топология виртуальных антенных элементов.

Такая конфигурация антенной системы позволяет проводить оценку угловых координат цели в азимутальной и угломестной плоскостях.

Пусть имеется вычислитель на графическом процессоре NVIDIA Jetson AGX Xavier 32GB, построенный на микроархитектуре Volta.

Требуется обеспечить выдачу информации о дальности, скорости, а также азимуте и угле места каждой из обнаруженных РЛС целей в реальном масштабе времени.

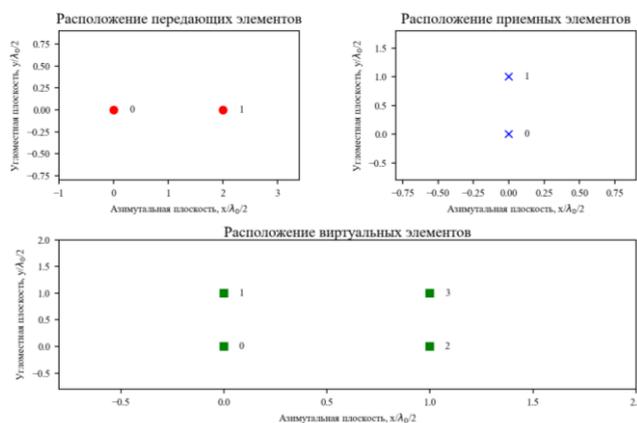


Рис. 1. Топология антенных элементов рассматриваемой РЛС

Каждый из  $N_{Rx}$  приемных каналов формирует массив данных, состоящий из  $N_t = 512$  отсчетов отраженного сигнала,  $L_T = 8192$  периодов зондирования. Каждый кадр, выдаваемый трансивером, содержит  $N_t \times L_T \times N_{Rx} = 512 \times 8192 \times 2 \sim 8,4$  млн. комплексных отсчетов с плавающей запятой одинарной точности. Реальный масштаб времени обеспечивается при обработке одного входного кадра данных, поступающего с трансивера, за время не более 50 мс. Таким образом, необходимо обеспечить пропускную способность процедуры обработки отраженного сигнала не менее 1,35 Гбайт/с.

**Выбор и адаптация алгоритмов для оценки параметров целей в РЛС.** В общем случае, обработку отраженного сигнала в многоканальных РЛС с пространственным разнесением передающих и приемных групп антенных элементов разделяют на три этапа [14]: корреляционная обработка (сжатие по дальности) и получение дальностно-временного портрета; межпериодная обработка сжатых по дальности отсчетов и получение дальностно-скоростного портрета; формирование пространственной характеристики направленности виртуальной антенной решетки по амплитудно-фазовому распределению тока на её элементах. На рис. 2 представлена последовательность обработки отсчетов отраженного сигнала, а на рис. 3 – укрупненный алгоритм обработки отраженного сигнала. Ключевой особенностью РЛС данного класса является необходимость ортогонального разделения зондирующих сигналов, излучаемых каждым из передающих элементов. Известны различные способы разделения сигналов, по частоте, времени, кодовой последовательности, разделение по поляризации и т.п. [6, 14]. При разработке данной РЛС было использовано разделение по частоте Доплера зондирующих сигналов её передающих элементов [15].

В течение зондирующей посылки комплексные отсчеты отраженного сигнала накапливаются в буфере, находящемся в ОЗУ вычислителя, формируя кадр (шаг 1 алгоритма, приведенного на рис. 3). Поскольку отсчеты сигнала на выходе трансивера имеют тип

данных *int16*, а дальнейшие вычисления используют числа с плавающей запятой одинарной точности *float32*, необходимо провести преобразование типов входных данных. Для согласования порядка следования отсчетов каналов приемника и порядка, принятого в процедуре обработки, необходимо выполнить их упорядочивание (шаг 2 алгоритма обработки, приведенного на рис. 3).

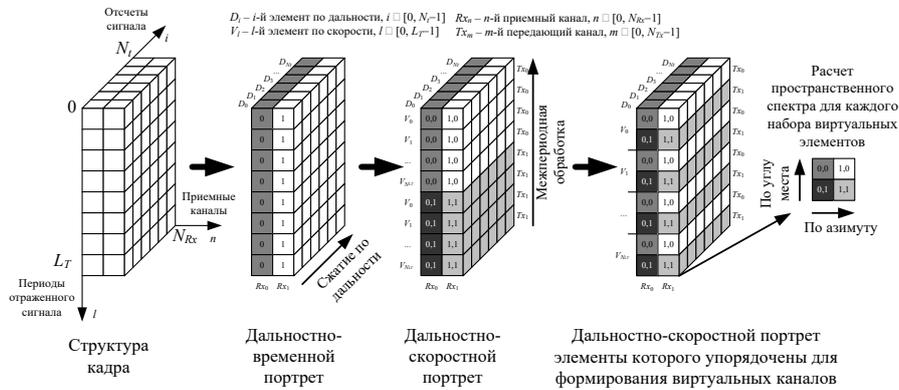


Рис. 2. Обработка отраженного сигнала в рассматриваемой РЛС



Рис. 3. Укрупненная блок-схема алгоритма обработки отраженного сигнала в рассматриваемой РЛС

Такое упорядочивание данных необходимо для эффективной работы контроллера памяти вычислителя и позволяет использовать линейный доступ к данным на последующих этапах обработки [16]. Упорядочивание данных выполнялась с помощью собственной функции ядра, написанной на CUDA C. После процедуры упорядочивания, буфер в ОЗУ вычислителя содержит данные, подготовленные для выполнения процедуры оптимальной обработки (сжатия по дальности).

Следующим этапом обработки является получение дальностно-временного портрета (см. рис. 2) посредством выполнения процедуры корреляционной обработки (сжатие по дальности) накопленных периодов отраженного сигнала. Каждая  $l$ -я строка дальностно-временного портрета  $\dot{J}_l$  представляет собой свертку  $l$ -го периода отраженного сигнала  $\dot{S}_l$  с коэффициентами согласованного фильтра в частотной области

$$\dot{J}_{l,n} = F_{\langle N_t \rangle}^{-1} \left\{ F_{\langle N_t \rangle} \left\{ \dot{S}_{l,n} \right\} \times C_{MF} \right\}, \quad (1)$$

где  $l = [0, L_T - 1]$  – индекс периода принятого сигнала,  $n = [0, N_{Rx} - 1]$  – индекс приемного канала,  $F_{\langle m \rangle}$ ,  $F_{\langle m \rangle}^{-1}$  – символ прямого и обратного преобразования Фурье вдоль размерности кадра  $m$ ,  $C_{MF}$  – дискретизированная передаточная функция согласованного фильтра (СФ), или частотные коэффициенты СФ.

Для вычисления преобразований Фурье использовалась библиотека *cuFFT*, входящая в состав NVIDIA CUDA SDK. Особенностью шага 3 алгоритма обработки, приведенного на рис. 3, является возможность интерполяции получаемого дальностного портрета. Интерполяция производится в частотной области, непосредственно перед обратным преобразованием Фурье. Далее выполняется межпериодное сжатие по скорости (шаг 4 алгоритма обработки, приведенного на рис. 3) для расчета дальностно-скоростного портрета (см. рис. 2). Каждый  $i$ -й столбец дальностно-скоростного портрета  $\dot{G}_i$  представляет собой прямое преобразование Фурье  $i$ -х столбцов дальностно-временного портрета по всем периодам зондирования

$$\dot{G}_{i,n} = F_{\langle L_T \rangle} \left\{ \dot{J}_{i,n} \right\}, \quad (2)$$

где  $i = [0, N_t - 1]$  – индекс, соответствующий номеру отсчета принятого сигнала.

Поскольку используется разделение каналов передачи по частоте Доплера, дальностно-скоростной портрет каждого приемного канала разделяется на  $N_{Tx}$  равных участков по строкам (периодам зондирования). Таким образом, первые  $\lfloor L_T / N_{Tx} \rfloor$  строк дальностно-скоростного портрета первого приемного канала, где  $\lfloor arg \rfloor$  – означает округление аргумента  $arg$  до ближайшего целого в меньшую сторону, будут иметь периоды отраженного сигнала, полученные в ходе зондирования первым передающим каналом, вторые  $\lfloor L_T / N_{Tx} \rfloor$  строки будут иметь периоды отраженного сигнала, полученные в ходе зондирования вторым передающим каналом и т.д. Возвращаясь к рис. 2 следует отметить, что на нём участки дальностно-скоростных портретов, принадлежащие различным комбинациям приемных и передающих каналов, выделены различными оттенками серого цвета.

Для вычисления диаграмм отражения целей в азимутальной и угломестной плоскостях необходимо произвести предварительную перестановку элементов дальностно-скоростного портрета  $\dot{G}$  с учетом координат мест расположения виртуальных каналов, как это показано на рис. 2. Упорядочивание данных выполняется собственной функцией ядра, написанной на CUDA C (шаг 5 алгоритма обработки, приведенного на рис. 3). В результате перестановок формируется матрица  $\dot{G}'$ , содержащая амплитудно-фазовые распределения виртуальных каналов рассматриваемой РЛС. Для формирования характеристик направленности итоговой виртуальной антенной решетки в двух плоскостях выполняется прямое двумерное преобразование Фурье (шаг 6 алгоритма обработки, приведенного на рис. 3).

Каждый элемент матрицы  $A$  представляет собой трехмерную комплекснозначную диаграмму направленности для каждой строки и столбца преобразованного дальностно-скоростного портрета  $\dot{G}'$

$$A_{i,l} = F_{\langle N_{Tx} \rangle} \left\{ F_{\langle N_{Rx} \rangle} \left\{ \dot{G}'_{i,l} \right\} \right\}. \quad (3)$$

Для вычисления двумерного преобразования Фурье используется библиотека *cuFFT*. При дальнейшей цифровой обработке принятого сигнала матрица  $A_{i,l}$ , найденная в результате преобразования Фурье, рассматривается как некий единый элемент, представляющий набор виртуальных излучателей антенной решетки для  $i$ -го отсчета по дальности и  $l$ -го отсчета по скорости (см. рис. 2).

Для диагностики состояния виртуальных каналов в процессе эксплуатации рассматриваемой РЛС выполняется поиск максимума амплитуды отраженного сигнала в каждом виртуальном канале (шаг 7 алгоритма обработки, приведенного на рис. 3). Реализация поиска использует модифицированную функцию ядра [5].

В работах [6–8, 10, 14] для обнаружения целей используется обнаружитель с постоянной вероятностью ложных тревог (CFAR). Авторы указанных работ отмечают высокую вычислительную сложность такого подхода, поэтому обнаружение целей в разработанной нами РЛС проводится на основе традиционной пороговой обработки сигналов, принимаемых рассматриваемой РЛС. Для корректного выбора порога обнаружения требуется знание уровня шума, который определяется по гистограмме распределения амплитуд отраженного сигнала, рассчитываемой индивидуально для каждого кадра (шаг 8 алгоритма обработки, приведенного на рис. 3). Для вычисления гистограммы используется модифицированный алгоритм приведения [18].

После определения порога обнаружения выполняется поиск элементов  $A_{i,l}$  (выражение (3)), амплитуды которых превышают заданный порог обнаружения. Поскольку реальные цели являются распределенными, то цель обычно содержит от нескольких десятков до сотен элементов  $A_{i,l}$ . Применение алгоритма кластеризации [19] позволяет группировать обнаруженные элементы  $A_{i,l}$  в кластеры и отождествлять обнаруженные кластеры целям в кадре (шаг 9 алгоритма обработки, приведенного на рис. 3).

Для повышения точности оценки угловых координат, по выражению (3) рассчитываются интерполированные характеристики направленности элементов  $A_{i,l}$ , превышающих порог обнаружения (шаг 10 алгоритма обработки, приведенного на рис. 3). Производится статистическая обработка каждого кластера, составленного из элементов  $A_{i,l}$ , для формирования дополнительной информации о цели, которая может быть использована для последующей классификации [20].

Результатом работы алгоритма обработки, приведенного на рис. 3, является список обнаруженных целей в кадре и их параметры (амплитуда отраженного сигнала, дальность, скорость, азимут и угол места). К дополнительным параметрам цели относится количество элементов  $A_{i,l}$ , амплитуды которых превысили порог обнаружения в каждом кластере, отождествленном с целью, среднее значение и СКО размеров цели по дальности, скорости, угловым координатам.

**Экспериментальная оценка вычислительной производительности.** В соответствии с алгоритмом обработки отраженного сигнала в рассматриваемой РЛС, приведенным на рис. 3, разработано функциональное программное обеспечение (ФПО) и проведена оценка производительности алгоритма обработки отраженного сигнала данной РЛС по каждому этапу обработки. В оценке производительности вычислительной системы рассматриваемой РЛС, помимо предлагаемого вычислителя, принимали участие несколько видеокарт, имеющих в распоряжении авторов. Участники тестирования приведены в табл. 1.

Таблица 1

**Наименования и параметры видеокарт, использованных для тестирования**

Модель вычислителя	Год начала выпуска	Микроархитектура	Пиковая пропускная способность памяти, Гбайт/с	Тактовая частота процессора, МГц
NVIDIA GeForce GTX 950	2015	Maxwell	105,76	1024
NVIDIA Jetson AGX Xavier 32GB	2018	Volta	88,128	854
NVIDIA GeForce GTX 1650 SUPER	2019	Turing	192,00	1485

В табл. 2 представлены усреднённые результаты замеров времени выполнения разработанной реализации алгоритма обработки отраженного сигнала рассматриваемой РЛС для одного и того же кадра данных радиолокационного наблюдения. Функциональное программное обеспечение компилировалось под конкретную вычислительную платформу компилятором *nvcc*, входящим в состав NVIDIA CUDA SDK.

Таблица 2

**Сравнение производительности видеокарт, использованных для тестирования**

№ п/п	Вычислительный этап	Время выполнения, мс		
		GTX 950	Jetson AGX Xavier	GTX 1650 SUPER
1.	Упорядочивание данных и преобразование типов (512×8192×2 комплексных отсчетов)	<b>7,509</b>	<b>6,109</b>	<b>12,403</b>
2.	Оптимальная обработка отраженного сигнала, сжатие по дальности	<b>14,594</b>	<b>16,155</b>	<b>8,193</b>
2.1	Прямое преобразование Фурье (512×8192×2 комплексных отсчетов)	1,652	2,252	0,842
2.2	Перемножение с комплексными отсчетами согласованного фильтра (512×8192×2 комплексных отсчетов)	1,545	1,567	0,812
2.3	Обратное преобразование Фурье с интерполяцией 8 (512×8×8192×2 комплексных отсчетов)	11,397	12,336	6,539
3.	Межпериодная обработка отраженного сигнала, сжатие по скорости	<b>5,66</b>	<b>8,031</b>	<b>2,625</b>
3.1	Транспонирование (256×8192×2 комплексных отсчетов)	1,952	4,307	0,742
3.2	Прямое преобразование Фурье (256×8192×2 комплексных отсчетов)	1,671	2,345	1,319
3.3	Транспонирование (256×8192×2 комплексных отсчетов)	2,037	1,379	0,564
4.	Упорядочивание данных для формирования амплитудно-фазового распределения (256×8192×2 комплексных отсчетов)	<b>1,078</b>	<b>0,881</b>	<b>0,410</b>
5.	Расчет характеристик направленностей виртуальной антенной решетки	<b>8,014</b>	<b>9,792</b>	<b>3,511</b>
5.1	Транспонирование (256×8192×2 комплексных отсчетов)	1,952	4,279	0,742
5.2	Прямое двумерное преобразование Фурье (256×8192×2 комплексных отсчетов)	4,021	4,365	2,207
5.3	Транспонирование (256×8192×2 комплексных отсчетов)	2,041	1,148	0,562
6.	Поиск максимума амплитуды в каждом виртуальном канале (256×8192×2 комплексных отсчетов)	<b>0,668</b>	<b>0,501</b>	<b>0,219</b>
7.	Вычисление гистограммы и определение порога обнаружения (256×8192×2 комплексных отсчетов)	<b>0,894</b>	<b>0,826</b>	<b>0,996</b>
8.	Обнаружение и кластеризация целей	<b>3,321</b>	<b>2,424</b>	<b>1,082</b>
8.1	Пороговая обработка 256×8192×2 комплексных отсчетов)	1,737	1,141	0,463

8.2	Разметка отсчетов, превысивших порог обнаружение (256×8192×2 комплексных отсчетов)	1,028	0,868	0,384
8.3	Кластеризация отсчетов, превысивших порог обнаружение (256×8192×2 комплексных отсчетов)	0,345	0,277	0,155
8.4	Сбор дополнительной статистики по полученным кластерам (256×8192×2 комплексных отсчетов)	0,211	0,138	0,080
9.	Интерполяция угловых координат обнаруженных целей	<b>3,026</b>	<b>2,524</b>	<b>2,202</b>
9.1	Подготовка ММО-элементов (709×2×2 комплексных отсчетов)	0,009	0,009	0,006
9.2	Прямое двумерное преобразование Фурье с интерполяцией 32×16 (709×64×32 комплексных отсчетов)	3,017	2,515	2,196
<b>Итого:</b>		<b>44,764</b>	<b>47,244</b>	<b>31,640</b>

Как видно из табл. 2, на предлагаемом к применению вычислителе NVIDIA Jetson AGX Xavier 32GB удалось добиться обработки кадра радиолокационных данных, получаемых от данной РЛС, за время менее 50 мс.

Включение в тест доступных графических карт NVIDIA GeForce GTX 950 и NVIDIA GeForce GTX 1650 SUPER позволяет наглядно показать зависимость производительности разработанного алгоритма обработки сигнала на разных микроархитектурах GPU. Известно [21], что GPU имеет высокую латентность при обращениях в глобальную память, что в общем случае ограничивает скорость обработки данных на GPU. Итоговые результаты, приведенные в табл. 2, наглядно демонстрируют уменьшение времени расчета в зависимости от увеличения пропускной способности памяти GPU. Таким образом, пропускная способность памяти является удобной метрикой для прогнозирования производительности вычислительных алгоритмов на разных моделях GPU.

В работе [12] предложена реализация двумерного преобразования Фурье на ПЛИС. В тоже время в таблице 4, приведенной в работе [12], показано, что двумерное БПФ размерностью  $2048 \times 2048 = 4$  млн. комплексных отсчетов (плавающая запятая с одинарной точностью) на ПЛИС Xilinx Zynq Ultra Scale+ выполняется примерно за 168 мс. В тоже время, вычисление двумерного БПФ такой же размерности  $256 \times 8192 \times 2 = 4$  млн. комплексных отсчетов на NVIDIA Jetson AGX Xavier 32GB занимает всего около 9,8 мс (см. пятую строку табл. 2). Очевидно, что вычисление БПФ на предлагаемой к применению GPU NVIDIA Jetson AGX Xavier 32GB происходит примерно в 17 раз быстрее, нежели с помощью ПЛИС Xilinx Zynq Ultra Scale+. Транспонирование данных перед вычислением БПФ необходимо для организации эффективного доступа к данным, хранящимся в глобальной памяти GPU.

В работах [10, 11] приведены результаты замеров производительности разработанных алгоритмов обработки сигналов аналогичных РЛС на GPU, позволяющие провести сравнение их эффективности с алгоритмом, предложенным в настоящей работе.

В качестве критерия сравнения алгоритмов предлагается использовать отношение пропускной способности алгоритма  $B_a$ , отнесенной к пиковой пропускной способности памяти  $B$  вычислителя, на котором осуществлялся его запуск.

Расчетное соотношение для определения пропускной способности алгоритма  $B_a$  имеет вид

$$B_a = N_s C_s / T_p, \quad (4)$$

где  $N_s$  – количество комплексных отсчетов, шт.;  $C_s$  – объем одного комплексного отсчета, байт;  $T_p$  – время обработки одного кадра, состоящего из  $N_s$  отсчетов, мс.

В табл. 3 приведены значения пропускной способности различных вариантов реализации алгоритма цифровой обработки отраженного сигнала в рассматриваемой РЛС, рассчитанные с помощью соотношения (4).

Таблица 3

**Сравнение эффективности алгоритма цифровой обработки сигналов  
на тестируемых видеокάρтах в радиолокаторах с пространственным разнесением  
передающих и приемных групп антенных элементов**

Наименование параметра	RTX A6000	GeForce RTX 3090	Jetson AGX Xavier
Размерность кадра данных	8×8×256×1024	8000×32×4×64	512×8192×2
Количество комплексных отсчетов $N_s$ , шт	16 777 216	65 536 000	8 388 608
Время обработки одного кадра $T_p$ , мс	49,67	371,9	47,244
Объем одного комплексного отсчета $C_s$ , байт	8	8	8
Пропускная способность алгоритма обработки, $V_a$ Гбайт/с	2,7	1,41	1,42
Пиковая пропускная способность памяти $B$ , Гбайт/с	768	935,8	88,128
Эффективность использования пропускной способности памяти вычислителя алгоритмом обработки $V_a/B$ , %	0,35	0,15	1,61
Проигрыш в эффективности известных алгоритмов предложенному авторами, раз	4,6	10,73	1

Как видно из табл. 3, нами предложена наиболее эффективная реализация цифровой обработки сигналов РЛС данного класса с точки зрения максимально полного использования пропускной способности памяти вычислителя. При этом следует заметить, что полученная нами оценка эффективности использования пропускной способности памяти вычислителя алгоритма обработки отраженного сигнала рассматриваемой РЛС является приближенной. Тем не менее, предложенный критерий сравнения данного алгоритма ЦОС с известными алгоритмами является универсальной метрикой для прогнозирования производительности вычислительных алгоритмов цифровой обработки сигнала в многоканальных РЛС с пространственным разнесением передающих и приемных групп антенных элементов.

**Заключение.** При разработке алгоритма цифровой обработки сигнала в РЛС с пространственным разнесением передающих и приемных групп антенных элементов выполнен сравнительный анализ различных вариантов его практической реализации с учётом архитектуры применяемого параллельного вычислителя. Варианты реализации алгоритма ЦОС выбирались с учётом необходимости выполнения требования обеспечения линейного доступа к данным в глобальной памяти применяемого вычислителя. При рассмотрении возможных вариантов практической реализации вычислительного алгоритма ЦОС в РЛС данного класса предполагалось, что число используемых внутренних буферов памяти должно быть минимальным. Предпринятые меры позволили выполнить обработку кадра, состоящего из 8,4 млн. комплексных отсчетов, за время менее 50 мс на графическом процессоре NVIDIA Jetson AGX Xavier, что удовлетворяет требованию по функционированию подсистемы обработки в реальном масштабе времени.

Показана обратно пропорциональная зависимость времени обработки кадра от величины пиковой пропускной способности памяти GPU. Предложен критерий оценки производительности алгоритма обработки в виде отношения пропускной способности алгоритма  $V_a$  к пиковой пропускной способности памяти  $B$  вычислителя. Проведена численная оценка эффективности использования пропускной способности памяти вычислителя разработанного алгоритма в сравнении с известными реализациями на GPU. Выигрыш по времени обработки для разработанного алгоритма в среднем составляет в среднем до 5 раз по сравнению с результатами, полученными ранее для ЦОС на графических процессорах (см., например, работы [10] и [11]). При этом показано, что по сравнению с ПЛИС реализация двумерного БПФ на GPU оказывается в 17 раз быстрее.

Дальнейшее увеличение производительности разработанного алгоритма цифровой обработки сигнала в РЛС с пространственным разнесением передающих и приемных групп антенных элементов возможно за счет отказа от ряда операций транспонирования, которые были введены на этапе отладки для обеспечения линейного доступа к данным, хранящимся в глобальной памяти вычислителя.

## БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Su Y., et al.* Fast Target Localization in FMCW-MIMO Radar with Low SNR and Snapshot via Multi-Deep Net // *Remote Sensing*. – 2022. – Vol. 15, No. 1. – P. 1-21.
2. *Лобач В.Т., Касьянов А.О., Потупак М.В., Суматохин К.В.* Особенности цифрового формирования диаграммы направленности МИМО-радар // *Антенны*. – 2016. – № 7. – С. 47-52.
3. *Lobach V., Kas'yanov A., Potipak M.* Digital beam forming in MIMO-radar with frequency diversity // 2016 IEEE East-West Design & Test Symposium (EWDTS). IEEE. 14-17 Oct. 2016. – P. 1-4.
4. *Karim B.A., Ali H.K.* Computationally efficient MUSIC based DOA estimation algorithm for FMCW radar // *Journal of Electronic Science and Technology*. – 2023. – Vol. 21, No. 1. – P. 1-18.
5. *Singh H., Chattopadhyay A.* Multi-target Range and Angle detection for MIMO-FMCW radar with limited antennas // *arXiv preprint arXiv:2302.14327*. – 2023.
6. *Grove R.L.* MIMO radar Systems and Algorithms-Imperfections and Calibration. Copenhagen. – Technical University of Denmark, 2022. – 189 p.
7. *Patrick D., Huang Y., Brennan P.W.* FMCW based MIMO imaging radar // *ARMMS Conference*. 7-8 Apr. 2014. – Vol. 1. – P. 12-34.
8. *Eugin H., Lee J.* Hardware architecture design and implementation for FMCW radar signal processing algorithm // *Proceedings of the 2014 Conference on Design and Architectures for Signal and Image Processing*. IEEE. 8-10 Oct. 2014. – P. 1-6.
9. *Fixed-Point and Floating-Point FMCW Radar Signal Processing with Tensilica DSPs* // White paper. 2022. – URL: [https://www.cadence.com/content/dam/cadence-www/global/en\\_US/documents/tools/ip/tensilica-ip/tip-fmcw-radar-wp.pdf](https://www.cadence.com/content/dam/cadence-www/global/en_US/documents/tools/ip/tensilica-ip/tip-fmcw-radar-wp.pdf).
10. *Pitre E. et al.* Comparison of massively parallel algorithms on graphics processing unit for MIMO radar // *e-Prime-Advances in Electrical Engineering, Electronics and Energy*. – 2022. – Vol. 2. – P. 1-9.
11. *Liu G. et al.* A GPU based real time processing system for frequency division multiple input multiple output radar // *IET Radar, Sonar & Navigation*. – 2023. – Vol. 17, No. 10. – P. 1524-1537.
12. *Heo J. et al.* FPGA implementation of an efficient FFT processor for FMCW radar signal processing // *Sensors*. – 2021. – Vol. 21, No. 19. – P. 1-16.
13. *Damjanović V.D., Petrović M.L., Milovanović V. M.* On Hardware Implementations of Two-Dimensional Fast Fourier Transform for Radar Signal Processing // *IEEE EUROCON 2023-20th International Conference on Smart Technologies*. 6-8 Jul. 2023. – P. 400-405.
14. *Texas Instruments.* MIMO radar // *Application Report SWRA554*. – 2018.
15. *Doerry A.W.* SAR ambiguous range suppression // Report SAND2006-5332. Sandia National Laboratories (SNL). Albuquerque, NM. and Livermore, CA (United States). – 2006.
16. *Harris M.* How to Access Global Memory Efficiently in CUDA C/C++ Kernels // *Parallel Forall*. NVIDIA Corporation. 2013. – URL: <https://developer.nvidia.com/blog/how-access-global-memory-efficiently-cuda-c-kernels/>.
17. *Luitjens J.* Faster parallel reductions on Kepler // *Parallel Forall*. NVIDIA Corporation. – 2014. – URL: <https://devblogs.nvidia.com/faster-parallel-reductions-kepler>.
18. *Sakharnykh N.* GPUPro Tip: Fast histograms using shared atomics on Maxwell // *Parallel Forall*. NVIDIA Corporation. – 2015. – URL: <https://developer.nvidia.com/blog/gpu-pro-tip-fast-histograms-using-shared-atomics-maxwell/>.
19. *Kalentev O. et al.* Connected component labeling on a 2D grid using CUDA // *Journal of Parallel and Distributed Computing*. – 2011. – Vol. 71, No. 4. – P. 615-620.
20. *Касьянов А.О., Потупак М.В., Евтушенко А.Н.* Идентификация радиолокационной станции на основе анализа непреднамеренной внутриимпульсной модуляции // *Радиотехника*. – 2023. – Т. 87, № 2. – С. 73-83.
21. *Jia Z. et al.* Dissecting the NVidia Turing T4 GPU via microbench marking // *arXiv preprint arXiv:1903.07486*. – 2019.

## REFERENCES

1. *Su Y., et al.* Fast Target Localization in FMCW-MIMO Radar with Low SNR and Snapshot via Multi-Deep Net, *Remote Sensing*, 2022, Vol. 15, No. 1, pp. 1-21.
2. *Lobach V.T., Kas'yanov A.O., Potipak M.V., Sumatokhin K.V.* Osobennosti tsifrovogo formirovaniya diagrammy napravlenosti MIMO-radara [Features of digital beamforming of MIMO radar], *Antenny* [Antennas], 2016, No. 7, pp. 47-52.

3. Lobach V., Kasyanov A., Potipak M. Digital beam forming in MIMO-radar with frequency diversity, *2016 IEEE East-West Design & Test Symposium (EWDTS)*. IEEE. 14-17 Oct. 2016, pp. 1-4.
4. Karim B.A., Ali H.K. Computationally efficient MUSIC based DOA estimation algorithm for FMCW radar, *Journal of Electronic Science and Technology*, 2023, Vol. 21, No. 1, pp. 1-18.
5. Singh H., Chattopadhyay A. Multi-target Range and Angle detection for MIMO-FMCW radar with limited antennas, *arXiv preprint arXiv:2302.14327*, 2023.
6. Grove R.L. MIMO radar Systems and Algorithms-Imperfections and Calibration. Copenhagen. Technical University of Denmark, 2022, 189 p.
7. Patrick D., Huang Y., Brennan P.W. FMCW based MIMO imaging radar, *ARMMS Conference*. 7-8 Apr. 2014, Vol. 1, pp. 12-34.
8. Eugin H., Lee J. Hardware architecture design and implementation for FMCW radar signal processing algorithm, *Proceedings of the 2014 Conference on Design and Architectures for Signal and Image Processing*. IEEE. 8-10 Oct. 2014, pp. 1-6.
9. Fixed-Point and Floating-Point FMCW Radar Signal Processing with Tensilica DSPs, *White paper*. – 2022. Available at: [https://www.cadence.com/content/dam/cadence-www/global/en\\_US/documents/tools/ip/tensilica-ip/tip-fmcw-radar-wp.pdf](https://www.cadence.com/content/dam/cadence-www/global/en_US/documents/tools/ip/tensilica-ip/tip-fmcw-radar-wp.pdf).
10. Pitre E. et al. Comparison of massively parallel algorithms on graphics processing unit for MIMO radar, *e-Prime-Advances in Electrical Engineering. Electronics and Energy*, 2022, Vol. 2, pp. 1-9.
11. Liu G. et al. A GPU based real time processing system for frequency division multiple input multiple output radar, *IET Radar. Sonar & Navigation*, 2023, Vol. 17, No. 10, pp. 1524-1537.
12. Heo J. et al. FPGA implementation of an efficient FFT processor for FMCW radar signal processing, *Sensors*, 2021, Vol. 21, No. 19, pp. 1-16.
13. Damjanović V.D., Petrović M.L., Milovanović V. M. On Hardware Implementations of Two-Dimensional Fast Fourier Transform for Radar Signal Processing, *IEEE EUROCON 2023-20th International Conference on Smart Technologies*. 6-8 Jul. 2023, pp. 400-405.
14. Texas Instruments. MIMO radar, *Application Report SWRA554*, 2018.
15. Doerry A.W. SAR ambiguous range suppression, *Report SAND2006-5332*. Sandia National Laboratories (SNL). Albuquerque. NM. and Livermore. CA (United States), 2006.
16. Harris M. How to Access Global Memory Efficiently in CUDA C/C++ Kernels, *Parallel Forall. NVIDIA Corporation*, 2013. Available at: <https://developer.nvidia.com/blog/how-access-global-memory-efficiently-cuda-c-kernels/>.
17. Luitjens J. Faster parallel reductions on Kepler, *Parallel Forall. NVIDIA Corporation*, 2014. Available at: <https://devblogs.nvidia.com/faster-parallel-reductions-kepler>.
18. Sakharnykh N. GPUPro Tip: Fast histograms using shared atomics on Maxwell, *Parallel Forall. NVIDIA Corporation*, 2015. Available at: <https://developer.nvidia.com/blog/gpu-pro-tip-fast-histograms-using-shared-atomics-maxwell/>.
19. Kalentev O. et al. Connected component labeling on a 2D grid using CUDA, *Journal of Parallel and Distributed Computing*, 2011, Vol. 71, No. 4, pp. 615-620.
20. Kas'yanov A.O. Potipak M.V., Evtushenko A.N. Identifikatsiya radiolokatsionnoy stantsii na osnove analiza neprednamerennoy vnutriimpul'snoy modulyatsii [Identification of a radar station based on the analysis of unintentional intrapulse modulation], *Radiotekhnika* [Radio engineering], 2023, Vol. 87, No. 2, pp. 73-83.
21. Jia Z. et al. Dissecting the NVidia Turing T4 GPU via microbench marking, *arXiv preprint arXiv:1903.07486*, 2019.

Статью рекомендовал к опубликованию д.т.н., профессор И.И. Турулин.

**Касьянов Александр Олегович** – Южный федеральный университет; e-mail: kasyanovao@sfedu.ru; г. Ростов-на-Дону, Россия; кафедра радиотехнических и телекоммуникационных систем; д.т.н.; профессор.

**Потипак Михаил Владимирович** – e-mail: potipak@sfedu.ru; г. Таганрог, Россия; Научно-исследовательский институт робототехники и процессов управления; к.т.н.; с.н.с.

**Kasyanov Alexander Olegovich** – Southern Federal University; e-mail: kasyanovao@sfedu.ru; Rostov-on-Don, Russia; the Department of Radio Engineering and Telecommunication Systems; dr. of eng. sc.; professor.

**Potipak Mikhail Vladimirovich** – e-mail: potipak@sfedu.ru; Taganrog, Russia; Scientific and technological laboratory “Designing on-board systems of robotic complexes”; cand. of eng.; senior researcher.