

ИССЛЕДОВАНИЕ АЛГОРИТМОВ БЫСТРОГО ПРЕОБРАЗОВАНИЯ ФУРЬЕ ДВУМЕРНЫХ СИГНАЛОВ ДЛЯ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

Дубровина О.В., Башков Е.А.

Донецкий национальный технический университет

Дискретное преобразование Фурье (ДПФ) играет важную роль при анализе, синтезе и разработке алгоритмов и программно-аппаратных средств цифровой обработки сигналов [1]. Последние находят широкое применение в различных областях науки и техники: радиолокации, передачи и сжатии аудио и видео информации, геологии и т.п.

Большая распространенность ДПФ базируется на возможности построения быстрых алгоритмов – так называемых алгоритмов быстрого преобразования Фурье (БПФ). Однако и они не всегда обеспечивают выполнение преобразования в реальном времени, что, как правило, необходимо для практических приложений. Например, анализ в реальном времени сигналов большого объема (таких как снимки со спутника) стандартными средствами вычислительной техники, невозможен.

Над этой проблемой работают многие ученые мира, такие как Брендон Блойд, Юрий Доценко, Бартон Смит, Джон Манферделли и другие (см. [2]).

Одно из направлений повышения быстродействия обработки сигналов с помощью БПФ – распараллеливание вычислительного процесса преобразования и его реализация на некоторой универсальной или специализированной параллельной вычислительной системе.

Существует большое количество разработок связанных с реализацией алгоритмов БПФ на различных параллельных архитектурах: суперкомпьютерах [2, 3, 4], кластерах [2], Cell процессорах [2].

Параллельные вычислительные системы относятся к классу дорогостоящих средств ВТ и не доступны большинству пользователей. Но сейчас появилась возможность использовать новую технологию CUDA параллельных вычислений на широко распространенных графических процессорах семейства NVIDIA.

Это открывает перспективу параллельной реализации БПФ на широко распространенных низкостоймых компьютерах. Для

этого необходимо исследовать возможности создания и оценить характеристики реализации алгоритмов быстрого двумерного преобразования Фурье с использованием технологии NVIDIA CUDA.

CUDA (Compute Unified Device Architecture) - это технология от компании NVidia, предназначенная для разработки приложений для массивно-параллельных вычислительных устройств для GPU (графических процессоров), ориентированных в первую очередь, на графические приложения[2]. Технология CUDA работает на видеокартах NVIDIA начиная с 8400GS и выше. Разные видеокарты имеют разные возможности. В целом, если в видеокарте например 128 SP(Streaming Processor) — это значит что она содержит 8 SIMD MP (multiprocessor), каждый из которых делает одновременно 16 операций. Характеристики различных моделей видеокарт приведены в таблице 1.1.

Таблица 1.1 – Характеристики видеокарт NVIDIA [5]

Модель видеокарты	Кол-во мульти-процессоров
GeForce GTX 280	30
GeForce GTX 260	24
GeForce 9800 GX2	2x16
GeForce 9800 GTX, 9800 GTX+, 8800 GTS 512	16
GeForce 8800 Ultra, 8800 GTX	16
GeForce 9800 GT, 8800 GT, 9800M GTX	14
GeForce 9600 GSO, 8800 GTS, 8800M GTX, 9800 GT	12
GeForce 8800 GTS	12
GeForce 9600 GT, 8800M GTS, 9800M GTS	8
GeForce 9700M GT	6
GeForce 9500 GT, 8600 GTS, 8600 GT, 9700M GT, 9650M GS, 9600M GT, 9600M GS, 9500M GS, 8700M GT, 8600M GT, 8600M GS	4
GeForce 8500 GT, 8400 GS, 8400M GT, 9500M G, 9300M G, 8400M GS, 9400 mGPU, 9300 mGPU, 8300 mGPU, 8200 mGPU, 8100 Mgpu	2
GeForce 9300M GS, 9200M GS, 9100M G, 8400M G	1

Видеокарты разных серий имеет различные вычислительные возможности. Для создания программ с использованием технологии CUDA необходима одна из вышеперечисленных видеокарт, а так же

библиотека стандартных функций соответствующей версии, доступная на сайте разработчика <http://www.nvidia.ru>.

Для оценки временных параметров реализации БПФ с помощью технологии NVIDIA CUDA на компьютерах класса IBM PC разработаны две экспериментальные программы.

Первая программа – базовая, реализует алгоритм на обычном персональном компьютере. Все вычисления выполняются последовательно – сначала вычисляется БПФ для каждой строки изображения, а потом для каждого столбца.

Вторая программа – реализует алгоритм двумерного БПФ для видеокарты GeForce 9300M G. Кратко, подход к распараллеливанию на GPU заключается в следующем: сначала одновременно выполняется БПФ над всеми строками изображения, а потом параллельно над всеми столбцами.

Для вычисления быстрых преобразований Фурье в технологии CUDA существует стандартная библиотека CUFFT. Недостатком этой библиотеки является ограничение на размер быстрого преобразования. Так же арифметика чисел с двойной точностью поддерживается только в последних версиях библиотеки.

CUDA вариант FFT поддерживает 1D, 2D, и 3D преобразования комплексных и действительных данных, пакетное исполнение для нескольких 1D трансформаций в параллели, размеры 2D и 3D трансформаций могут быть в пределах [2, 16384], для 1D поддерживается размер до 8 миллионов элементов [6].

Эксперименты выполнялись на компьютере класса IBM PC со следующими характеристиками:

- Процессор Intel(R) CPU T2390 @1,86 GHz;
 - Оперативная память 2 Гб.
- Характеристики видеокарты NVIDIA GeForce 9300M G:
- Глобальная память: 268107776 б;
 - Разделяемая память на каждый блок: 16384 б;
 - Количество регистров в каждом блоке: 8192 б;
 - Максимальное количество потоков на блок: 512;
 - Размер константной памяти: 65536 б.

В результате выполнения программы было получено время выполнения двумерного быстрого преобразования Фурье для различных размеров изображения. Полученные данные приведены в таблице 2. Для сравнения в таблице так же приведены результаты выполнения базовой программы.

Таблица 2 – Время выполнения двумерного БПФ

Высота	Ширина	CPU (сек)	GPU 9300 M (сек)	GPU 8600 GT (сек)	Выигрыш
4	4	0,0156250	0,000394 с	0.000195	39,65
8	8	0,046875	0,001040	0.000246	45,07
16	16	0,03125	0,000638	0.000356	48,98
64	64	0,09375	0,002686	0.001473	34,90
128	128	0,34375	0,007875	0.004216	43,65
256	256	0,7031	0,031599	0.015506	22,25
512	512	5,57	0,147223	0.040733	37,83
1024	1024	26,8125	0,376411	0.195710	71,2319,
2048				0.787242	

Полученные результаты (см. таблицу 2) показывают, что реализация алгоритма быстрого преобразования Фурье двумерного сигнала на видеокартах NVIDIA с помощью технологии CUDA по временным затратам в среднем более чем на порядок превосходит прямую реализацию алгоритма на одном процессоре.

Данные результаты не являются окончательными, так как БПФ вычисляется для чисел с одинарной точностью. При работе с числами двойной точности возможно снижение эффективности. Так же результат зависит от мощности видеокарты и количества ее процессоров.

Для получения более высокой производительности и дальнейшего уменьшения времени вычисления возможно использование видеокарт с большим количеством процессоров, или даже нескольких видеокарт [3, 4].

Литература

1. <http://sonder.ru/content/view/472/32/>
2. Говидараю Н.К., Ллойд Б., Доценко Ю., Смит Б., Манферделли Д. Высокопроизводительные дискретные преобразования Фурье на графических процессорах/ компания Microsoft.
<http://www.masters.donntu.edu.ua/2009/fvti/dubrovina/library/translate.htm>
3. Schivea H., Chiena C., Wonga S., Tsaia Y., Chiueha T. Graphic-Card Cluster for Astrophysics (GraCCA) - Performance Tests / Сборник статей по математике, физике, программированию и др.
<http://xxx.lanl.gov/ftp/arxiv/papers/0707/0707.2991.pdf>.
4. <http://uraldev.ru/news/id/1222>.
5. Nvidia Cuda Programming Guide 1.1.
6. <http://www.ixbt.com/video3/cuda-1.shtml>.

Получено 27.05.09