

## Final Project: Histogram

=====

### 1. Notes:

- a. There are four attached files: Final Project CSC8007.docx (this document), histogram.cu, input\_test.txt, input\_data.txt
- b. Due date: May 11, 2015 mid-night
- c. Required Reading: Textbook. In particular please read chapter 9 histogram section.
- d. Submission:
  - i. You should submit your code and screenshot of your results for both input files input\_test.txt and input\_data.txt via blackboard.

### 2. Objectives

- a. To learn the GPU programming To learn how to program CUDA code to perform parallel computing using GPU blocks and threads.
- b. To learn how to apply GPU mechanisms and features in your applications. These mechanisms and features include memory allocations, memory copies, global and shared memory, atomic operations, and etc.

### 3. CPU Histogram Computation – Run and print out the result

The attached histogram.cu is simply a C code for the computation of a histogram of data from an input text files. The text file contains multiple lines of characters. The code generates the frequency histogram of characters from the input file.

The code first reads the text data from input\_test.txt and stores the data into a two dimensional array aTextData of char.

Then the method histogram is called to generate the frequency histogram of the characters in the array aTextData. The histogram data is stored in the array histo of unsigned int.

- a. Run the attached code and record the result. Please use the attached input file “input\_test.txt”. Note: your GPU implementation of the code should not change the output results!!!
- b. The output should look like this:

```
histogram frequency of characters a to z:
102 11 45 40 147 30 26 47 104 5 15 50 33 108 118 35 2 76 95 98 41 16 14 6 35 0
histogram frequency of characters in terms of percentiles:
0.0785 0.0085 0.0346 0.0308 0.1132 0.0231 0.0200 0.0362 0.0801 0.0038 0.0115 0.0
385 0.0254 0.0831 0.0908 0.0269 0.0015 0.0585 0.0731 0.0754 0.0316 0.0123 0.0108
0.0046 0.0269 0.0000 Press any key to continue . . .
```

#### 4. Programming Assignment

The attached histogram.cu is the CPU version of the code. You are required to modify the code to take advantage of the computing power of GPU. Please read the attached source code and make sure which parts of the code need to be changed. In particular, please read the comments very carefully. I put "TO DO:" in the comment areas if code changes are necessary. Here is list of details:

- a. Put your name and date in the comment area under "Modified by:"
- b. Re-implement the function `histogram()` to kernel function with new name `gpu_histogram()`.
- c. Modify the main method so that the code will work on GPU. The modifications include:
  - i. Declare some GPU pointers `dev_text_data` of `char` and `dev_histo` of `unsigned int`.
  - ii. Allocate cuda memory for the pointers
  - iii. Copy the array data from host to device
  - iv. Call the kernel function for computing the frequency histograms of characters
  - v. Add code to display the frequency histograms of all the alphabets (both low and up cases). Also add code to display the relative frequency histogram of all the alphabets in terms of percentiles.
- d. Modify the comments for the code you have changed.

#### 5. Grading: (Total of 110 points)

##### Section 3: CPU Histogram Computation (original code)

- a. (10 points) run and submit the results

##### Section 4: Programming Assignment

- b. (5 points) Put your name and date in the comment area under "Modified by:"
- c. (45 points) Re-implement the function `histogram()` to kernel function with new name `gpu_histogram()`.
- d. (45 points) Modify the main method so that the code will work on GPU. The modifications include. Here is the distribution of the points:
  - i. (5 points) Declare some GPU pointers `dev_text_data` of `char` and `dev_histo` of `unsigned int`.
  - ii. (5 points) Allocate cuda memory for the pointers
  - iii. (5 points) Copy the array data from host to device
  - iv. (10 points) Call the kernel function for computing the frequency histograms

of characters

- v. (20 points) Add code to display the frequency histograms of all the alphabets (both low and up cases). Also add code to display the relative frequency histogram of all the alphabets in terms of percentiles. Please run the code for input file `input_test.txt`.

As the results for `input_test.txt` are displayed in section 3 above, you can use the input data to verify that your GPU code is working. As I mentioned before, both CPU and GPU versions should generate the same results for the same test data.

- vi. Run the code again for `input_data.txt` and display the results.

You can submit your results by taking the screenshot of the display results.

- e. (5 points) Modify the comments for the code you have changed.