1. When does it make sense to use constant memory?

When the data doesn’t change during the kernel execution.

1. What went well with this assignment?

Implementing the naïve algorithm was easy since I’m already familiar with allocating memory, copying elements from device to host and from host to device.

1. What was difficult?

The cache algo was hard. The sample code is partially provided, so I have to understand the entire code and then fill out the missing part. Also, it is designed for handling 2D matrix, which is different from naïve, which is 1D array.

1. How would you approach differently?

I would learn what pitch is prior to do this homework.

1. Anything else you want me to know?

The cache runs lot faster than naïve algo when dealing with smaller width, but the speed rises linearly. Naïve has a stable runtime throughout the size.

12/4Updates:

Discard the matrix part in cache, and focused on 1d array. No need to use pitch, use halo instead, and highly increases its time efficiency. It’s now far faster than naïve.