LABWORK 3

HELLO CUDA

TRAN Quy Ban

Department of ICT

Introduction

Implement code for labwork 1 using CUDA.

How to implement the labwork

```
    __global__ void grayscale(uchar3 *input, uchar3 *output) {
int tid = threadIdx.x + blockIdx.x * blockDim.x;
output[tid].x = (input[tid].x + input[tid].y +input[tid].z) / 3;
output[tid].z = output[tid].y = output[tid].x;
}


void Labwork::labwork3_GPU() {
     uchar3 *devInput;
     uchar3 *devGray;
     uchar3 *hostGray;
    int pixelCount = inputImage->width * inputImage->height;
    char *hostInput = static_cast<char *>(malloc(pixelCount * 3));
    // Allocate CUDA memory
        cudaMalloc(&devInput, pixelCount *sizeof(uchar3));
        cudaMalloc(&devGray, pixelCount *sizeof(float));
    // Copy CUDA Memory from CPU to GPU
        cudaMemcpy(devInput, hostInput,pixelCount *sizeof(uchar3),cudaMemcpyHostToDev
    // Processing
        int blockSize = 64;
        int numBlock = pixelCount / blockSize;
        grayscale<<<numBlock, blockSize>>>(devInput, devGray);

    // Copy CUDA Memory from GPU to CPU
   cudaMemcpy(hostGray, devGray,pixelCount *sizeof(float),cudaMemcpyDeviceToHost);
    // Cleaning
        cudaFree(devInput);
}
```

The result of the labwork is shown as below:

What is the speed up?

Try experimenting with different block size values

Plot a graph of block size vs speedup

Discuss