GPU Acceleration of non-iterative and iterative algorithms in Fluorescence Lifetime Imaging Microscopy

Gang Wu,1,2 Thomas Nowotny,1 Yu Chen,3 and David Day-Uei Li2
1University of Sussex, School of Engineering and Informatics, Brighton, UK
2University of Strathclyde, Centre for Biophotonics, Glasgow, UK
3University of Strathclyde, Physics, Glasgow, UK

1. Summary
Graphics Processing Unit (GPU) enhanced Fluorescence Lifetime Imaging Microscopy (FLIM) algorithms are presented, and their results are compared with the latest research results. The GPU based approaches are suitable for highly parallelized sensor systems and promising for high-speed FLIM applications.

2. FLIM System
FLIM Analysis System is used to extract the lifetime of fluorescent samples in biological research and medical diagnosis. It contains a light source (e.g. laser), a photon detector, a time-correlated single-photon counting (TCSPC) camera, lifetime analysis software, and a PC with graphical user interface (GUI).

3. FLIM Algorithms
FLIM generates images by analyzing the exponential decay (fluorescence lifetime) of fluorescence intensity (from fluorescent proteins tagged on biological samples) at each camera pixel. Lifetime can be extracted from an exponential histogram, as shown in the figure, by the following algorithms.

A. Iterative algorithms

Algorithm Function
LSM[1] \( A = \sum_{j=0}^{N-1} \left( \frac{N_j}{N} \right) \times \ln \left( \frac{N_j}{N} \right) \)

where \( N_j \) is the photon number of the jth bin, \( N \) is the photon number of the first bin, and \( A \) is the number of lifetime components.

B. Non-iterative algorithms

A. Thread-based non-iterative algorithms

The histogram of each pixel is analyzed by an independent CUDA thread, as shown in figure below, and each block contains 512 threads. This configuration allows analyzing a large number of pixels simultaneously, the exact number being determined by the number of streaming multi-processors.

4. GPU Implementation

A. Block-based iterative algorithms

To realize parallel FLIM analysis in a GPU, the histogram for each pixel is analyzed by a separate block of CUDA, as shown in the Figure and each such block contains 256 threads that roughly correspond to the 256 time bins.

B. Thread-based non-iterative algorithms

The histogram of each pixel is analyzed by an independent CUDA thread, as shown in figure below, and each block contains 512 threads. This configuration allows analyzing a large number of pixels simultaneously, the exact number being determined by the number of streaming multi-processors.

5. Simulations and FLIM data analysis

A. Simulation

2000 photons were collected, the number of time bin was 256, the width of each time bin was 100ps and the size of the image was 512 by 512 pixels.

B. Experiment

We demonstrate the performances of the GPU based BCMM on two-photon FLIM images of gold nanorods (GNRs)-Cy5 labelled A375 cells. GNRs were conjugated with Cy5 labelled oligonucleotide (GNR-Cy5) and fixed with paraformaldehyde. FLIM was performed using a confocal microscope (LSM 510, Carl Zeiss) equipped with a time-correlated single photon counting (TCSPC) module (SPC-830, Becker & Hickl GmbH).

6. Discussion

FLIM analysis is well-suited for GPU acceleration because it is highly parallelizable. Each pixel in a FLIM frame can be processed independently of any other pixel, and, depending on the details of the algorithm, there is a lot of room for parallelization even within the processing of an individual pixel.

7. References