Abstract

Objective Lipid Data Analyzer identifies lipids and their structure based on mass spectrometry. A bottle neck of the analysis was the Savitzky-Golay Filter which is used for smoothening the raw data of the chromatogram. This implementation is needed to detect the begin and the end of a peak in the chromatogram without changing the shape and position of the peak significantly. The aim was to speed-up the calculation by filtering in parallel.

Methods To reach the goal, the GPU (Graphic Processing Unit) was invoked to perform the calculation by using the programming model CUDA (Compute Unified Device Architecture). The existing Java code was ported to the C-like programing model. A native library was created which was included into the existing project and tested on three different platforms. If a CUDA capable device is installed, the library is called which uses the GPU.

Results The GPU version of the algorithm led to a speed-up between 7-170 depending on the input data and available graphics card. When smoothing a whole chromatogram file the CUDA version led to a speed-up of 6.5-43 compared to the Java version. Due to different floating point representations there are numerical deviations compared to the Java implementation. From 8084 lipids 28 differ in their result by more than 0.1 %. In average the deviation of the peak area is 0.05 % and the maximum 1.56 %.

Conclusion Enabling the GPU using CUDA leads to a noticeable speed-up of the analysis. The degree of performance enhancement is although very depending on the type of data which is processed and also on the used hardware.

Key Words: Savitzky-Golay; CUDA; Lipid Data Analyzer; Bioinformatics