University of Zurich[UZH]

Department of Informatics

University of Zürich
Department of Informatics
Binzmühlestr. 14
CH-8050 Zürich
Phone. +41 44 635 43 11
Fax +41 44 635 68 09
www.ifi.uzh.ch/dbtg

UZH, Dept. of Informatics, Binzmühlestr. 14, CH-8050 Zürich

Prof. Dr. Michael Böhlen
Professor
Phone +41 44 635 43 33
Fax +41 44 635 68 09
boehlen@ifi.uzh.ch

Zürich, October 31, 2018

Bachelor's Thesis:
Implementation of Single-Point Discrete Fourier Transform on two dimensional data

The Australia Square Kilometer Array Pathfinder (ASKAP) is located in Australia and is one of the leading radio telescope facilities in the world. ASKAP roughly produces 2.5 GB/s, which is about 216 TB/day or 100 PB/year. One of the most basic and fundamental operations applied on radio astronomy data is the Fast Fourier Transform (FFT) to convert frequency data into dirty sky images. The goal of this Bachelor thesis is to design, implement and evaluate an incremental fourier transform algorithm for radio astronomy data.

What is the Discrete Fourier Transform?

Fourier Transform is a standard analysis technique for observing the spectral content of a signal or sequence. The Discrete Fourier Transform takes a time-domain finite length sequence as input and produces a frequency-domain finite length sequence as output. The equation 1 shows DFT of sequence f[k].

$$F[n] = \sum_{k=0}^{N-1} f[k] \exp^{-j\frac{2\pi}{N}nk} \quad where\ n = 0\ to\ N-1 \tag{1}$$

When the DFT is implemented as an efficient algorithm it is called the Fast Fourier Transform (FFT). A divide-and-conquer approach is taken in which DFT is decomposed into a cascade like connections of 2-point DFTs also known as Radix 2 FFT. With the decimation-in-time (DIT) algorithm the decomposition begins by decimating the input sequence f[k]. But with the decimation-in-frequency(DIF) algorithm the decomposition begins by decimating the output sequence F[n] and working backwards. As mentioned above that the radio astronomy data consists of frequency components so we would use decimation-in-frequency algorithm to get the dirty sky images.

For 1-D data the complexity of DFT is $O(N^2)$. The Fast Fourier Transform (FFT) computes the DFT and produces exactly the same result as evaluating the DFT definition directly. The most important difference is that FFT is much faster with time complexity $O(NlogN)$.

A two dimensional DFT is given as:

$$F(u,v) = \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} f(m,n)\exp^{-j2\pi\frac{um}{M}+\frac{vn}{N}} \quad where\ u,v = 0\ to\ N-1 \tag{2}$$

2D DFT is carried out by 1D transforming all rows of the 2D function $f[m,n]$ and then 1D transforming all columns of the resulting matrix. The order of the steps is not important [2].

Thesis

In a streaming environment data arrives continuously. In the scope of this thesis we consider a fixed-size 2D histogram that is updated continuously. The goal is to produce the real time Fourier Transform of 2D histogram with each update. Since the time complexity of 2D FFT is $O(N^2logN)$, it is not feasible to apply FFT for each new data point. Hence a new algorithm is needed that calculates the Fourier Transform of 2D histogram by only updating a single point but in an optimal way.

An optimization of Fourier Transform algorithm depends upon the total number of complex addition and complex multiplication operations. The Cooley-Tukey FFT algorithm has run-time complexity of $O(N^2logN)$ for NxN input with N^2logN complex multiplication operations and $2N^2logN$ complex addition operations. We have devised a new algorithm (Single-Point Fourier Transform) that has runtime complexity of $O(N^2)$ with N^2 complex additions but only $N/2$ complex multiplications. The Single-Point Fourier Transform shall be studied, implemented, progressed, evaluated and described.

Tasks

1. Literature study on Apache Flink [1], Discrete Fourier Transform [2][3].
2. Implement a streaming pipeline for Radio Astronomy Data.
3. Implement a parallel version of Single-Point Fourier Transform in Apache Flink.
4. Evaluate the run-time performance of Parallel streaming pipeline Single-Point Fourier Transform.
5. Write the Bachelor's thesis.
6. Present your thesis in a DBTG meeting

References

1. Carbone, P., Katsifodimos, A., Ewen, S., Markl, V., Haridi, S. et al. (2015): Apache flink: Stream and batch processing in a single engine
2. $www.fourier.eng.hmc.edu/e101/lectures/Image\_Processing/node6.html$
3. $www.eas.uccs.edu/\ mwickert/ece5655/lecture\_notes/ece5655\_chap10.pdf$

University of Zurich
Department of Informatics