Det Natur- og Biovidenskabelige Fakultet

Kontrakt: Projekt udenfor kursusregi

PERSONLIGE OPLYSNINGER

zfj900

Sarah Maria *

Hyatt *

zfj900@alumni.ku.dk *

53806111 *

Uddannelse: MSc in Computer Science *

Evt. specialisering:

OPLYSNINGER OM HOVEDANSVARLIG VEJLEDER

Cosmin Oancea *

cosmin.oancea@di.ku.dk *

Institut: Datalogisk Institut *

Institut: Vælg fra listen *

Institut: Vælg fra listen *

OPLYSNINGER OM PROJEKTPERIODEN

Projekt uden for kursusregi kan kun indgå som begrænset valgfrit, hvis dette fremgår af din studieordning, eller hvis du har fået en dispensation til, at det kan indgå som begrænset valgfrit.

Projektet tæller som: (•) Begrænset valgfrit ( ) Valgfrit *

( ) Både begrænset valgfrit og valgfrit (kun ved 15 ECTS-point)

Projektets ECTS-belastning: *

7.5 ECTS-point

Blok for påbegyndelse: *

Blok 3

Blok for afslutning: *

Blok 3

Dato for aflevering af projektrapport:

29/03/2020

Bemærk: Dato for aflevering kan tidligst ligge 2 uger før eksamensugen i projektets sidste blok.

*

Eksamensform:

Written report with no oral presentation.

Dit projekt uden for kursusregi skal bedømmes med 7-trinsskalaen og intern censur.

OPLYSNINGER OM PROJEKTETS FOKUS

☐ Jeg laver projekt udenfor kursusregi i en gruppe.

Studerende på en engelsksproget uddannelse skal skrive på engelsk. Studerende på en dansksproget uddannelse kan skrive på dansk eller engelsk. Det skal fremgå af projektkontrakten hvilket sprog, der er valgt.

Angiv hvilket sprog projektet udenfor kursusregi udarbejdes på *

○ Dansk ◉ Engelsk ○ Svensk ○ Norsk

Projektets titel/emne

Accelerating Approximate Nearest-Neighbours via Propagation-Assisted KD-Trees on GPUs *

Nærmere beskrivelse: *

Project Description

Computing nearest-neighbour fields (NNFs) between two images is useful for solving various computer vision problems. One common method is applying brute-force which has the complexity of O(n^2), while it is easy to implement it is also infeasible when n is large.
Another common solution is using KD-Trees which has an average complexity of O(n lg n). While this offers a cheaper traversal it still suffers from the curse of high dimensionality. Thus, the trade-off is accuracy for lower dimensionality and reduced algorithmic complexity.

Dimensionality and tree traversal can be further optimised by applying NNF in an approximate fashion that does not guarantee an exact solution, however the result has been found to be good enough in practice.

He and Sun have proposed one such method [1] which is claimed to be 10-20 times faster than other ANNF methods of comparable accuracy but it has only been evaluated on sequential CPU systems. The outline of the algorithm works as follows.

Assuming two similar images A and B:

1. Choose a fix number of dimensions (say 20), and pick a number (say 1000) of random patches from the images and perform Principal Component Analysis (PCA), i.e., reduce dimensionality by picking the most important 20 dimensions.

2. Build a KD-tree for the patches in B: at each recursive step, we choose and split the space by the median value.
The median split results in a binary balanced tree. The recursion continues until each leaf (but the last one) contains a statically chosen 'm' patches (say 50).

3. Approximate search through the KD-Tree:

a) Find an exact NNF solution for the first row of patches in A, by fully traversing through the KD-Tree and performing a brute-force search at each encountered leaf.

b) For the remaining patches we will get an initial estimate from the leaf to which they belong, requiring a simpler implementation, i.e., no backtracking. (At the leaf level we still perform a brute-force search).

4. Propagation Step: note that the NNF for the patches of the first row are highly accurate. We can use them to improve the accuracy of the other by a sequential-propagation procedure which updates each patch in row `i` in the following way:
a) lookup the leaf `l_{i-1}` of the top neighbour (i.e., row `i-1` computed in previous time step)
b) perform a brute-force w.r.t. that leaf (`l_{i-1}`), and choose the best between the resulted ANNF and the current result (from step 3.b).

Project Goal

The aim of the project is to implement a approach similar with He and Sun's, however utilising highly parallel hardware such as GPUs. To this extent we are going to develop a data parallel implementation of the main algorithmic steps in the Futhark language and/or perhaps CUDA. We are going to identify the performance bottlenecks and study techniques aimed at solving them.

References

[1] Computing Nearest-Neighbour fields Propagation-Assisted KD-Trees, Kaiming He and Jian Sun.

PROJEKT UDENFOR KURSUSREGI SOM GRUPPEPROJEKT

Tilføj de andre medlemmer af gruppen herunder

Navn * KU-brugernavn *

Oplysninger om projektopgaven og eksamen

Projektet udarbejdes som en: *

○ Fælles rapport ○ Rapport med individualiserede bidrag

Eksamen afholdes som en: *

○ Gruppeprøve ○ Individuel prøve

OPLYSNINGER OM VEJLEDNINGSFORLØB

Den hovedansvarlige vejleder og den studerende skal indbyrdes afklare deres forventninger til projektforløbet ved at besvare nedenstående spørgsmål.

Angiv hvor mange timer det forventes, at der vejledes: 20 *

Angiv hvor ofte det forventes, at I skal mødes (fastlæg evt. en mødeplan): *

Weekly.

Angiv hvad der forventes af vejleder til møderne: *

Advice and guidance for the project.

Angiv hvad for forventes af de(n) studerende til møderne: *

Showing up prepared for a project related discussion.

Andet:

UNDERSKRIFTER

Læs hvor du skal aflevere kontrakten under Studieinformation på KUnet > Bachelorprojekt og andre projekter > Tilmelding og kontrakt.

Studerende **Dato**

Med min underskrift på kontrakten står jeg inde for, at jeg har læst universitets vejledning til at undgå plagiering: Studieinformation > Kurser og undervisning > God videnskabelig praksis.

Hovedansvarlig vejleder **Dato**

Studieleder **Dato**