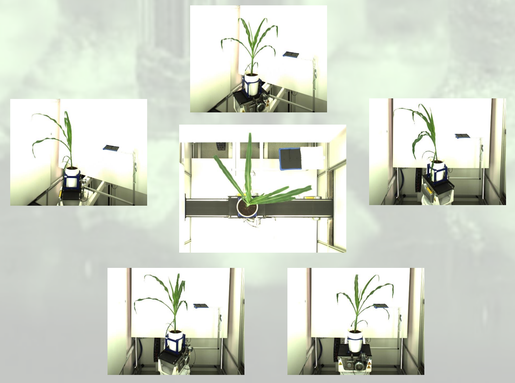
GPU-based maize plant analysis: accelerating CNN segmentation and voxel carving

Simon Donné, Bart Goossens, Stijn Dhondt, Nathalie Wuyts, Hiep Luong, Dirk Inzé, Wilfried Philips

UNIVERSITEIT GENT IPI

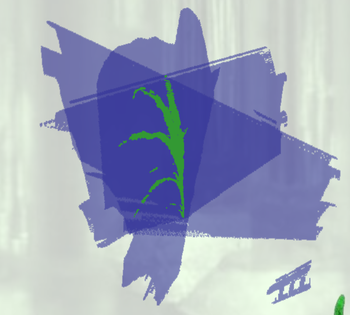
Plant phenotyping

PHENOVISION is a high-throughput plant phenotyping system for crop plants in greenhouse conditions. A conveyor belt transports plants between automated irrigation stations and imaging cabins.
The aim is to phenotype maize varieties grown under different conditions. To this end we model the plants in 3D and automate the measuring of the plants.



Voxel carving

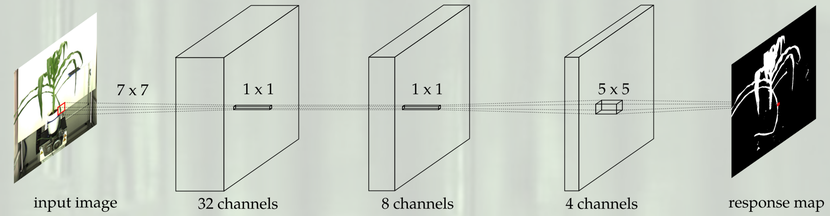
In order acquire a 3D model of the maize plants, we perform voxel carving to find a visual hull of the plant. This allows us to fit a parametric model in a later step, at which point we can easily perform automated measurements and even evolution through time.



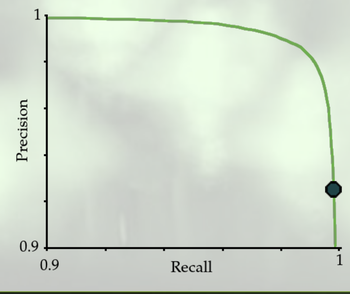
Voxel carving works by testing, for each candidate in the voxel cube, whether it corresponds to a plant pixel in all cameras. A perfect target for parallelization!

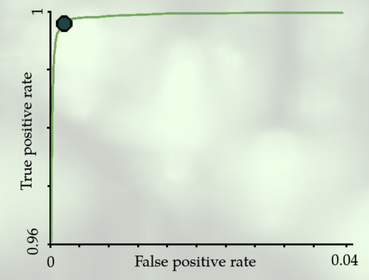
Neural net segmentation

Voxel carving requires binary input images, so we need to segment the images into *plant* and *not plant*. This is done using a pre-trained **convolutional neural net** (4 layers) which is able to discern between the maize plant and the background.



After training on manually annotated ground truth, we achieve high accuracy, with the precision-recall tuneable through the threshold.



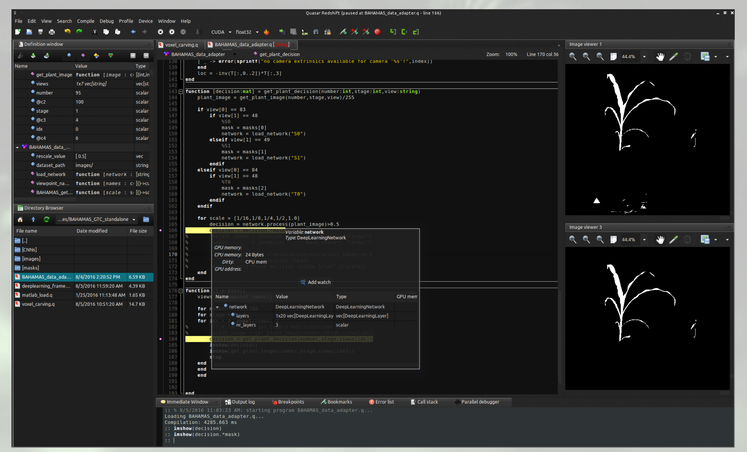


Quasar

Quasar is a new **programming language** (compiler, runtime optimization,…) for fast development on heterogeneous hardware:

- compact code
- automatic parallelization (OpenCL, OpenMP, CUDA, …)
- high-level programming like MATLAB with low-level GPU tweaks as required for optimization
- extensive debugging and profiling support using the Redshift **IDE**

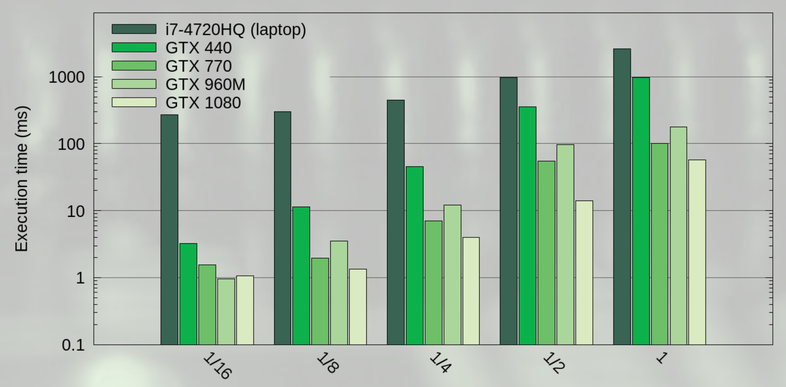
More info: **http://gepura.io**



Results

Because of the high throughput of the PHENOVISION system, the execution time of the processing is of the utmost importance. As the input images are of very high quality and resolution (4008x2672), the 3D model can be very accurate. But at full speed, the system images on the order of one plant per minute, which means that our preprocessing should occur in a fraction of this – let's say 30 seconds.
Both for the forward pass of the segmentation neural net and the voxel carving, the NVIDIA cards boast an impressive performance gain over the CPU: the GTX 1080 achieves a speed-up factor of 15-40x!
As the very large voxel cubes can pose a memory issue for the weaker cards, we have implemented block-wise processing of the voxel cube: yet the size of these blocks does not impact the execution time much.
The GTX 1080 is able to perform segmentation on 7 input images at full scale and voxel carving on a 1024^3 cube within 30 seconds – well within the margins of the PHENOVISION imaging speed. To put this into perspective: on the CPU we need to process the images in half resolution and carve a 256^3 cube to achieve the same throughput!

Full kernel convolution acceleration using the GPU



Voxel carving acceleration using the GPU