Toward a procedural fruit tree rendering framework for image analysis

Thomas Duboudin\textsuperscript{1}, Maxime Petit\textsuperscript{1}, and Liming Chen\textsuperscript{1}

\textsuperscript{1}LIRIS, CNRS UMR 5205, Ecole Centrale de Lyon, France

Abstract

We propose a procedural fruit tree rendering framework, based on Blender and Python scripts allowing to generate quickly labeled dataset (i.e. including ground truth semantic segmentation). It is designed to train image analysis deep learning methods (e.g. in a robotic fruit harvesting context), where real labeled training datasets are usually scarce and existing synthetic ones are too specialized. Moreover, the framework includes the possibility to introduce parametrized variations in the model (e.g. lightning conditions, background), producing a dataset with embedded Domain Randomization aspect.

Keywords: Synthetic Fruit Dataset, Harvesting Robotics, Procedural Model, Domain Randomization

1 Introduction and Previous Work

State-of-the-art methods for object recognition and grasping are currently mostly based on deep neural network. Despite tremendous results, these methods require a huge amount of labeled data in order to be trained. This is a major drawback in the field of robotics fruit-harvesting, where existing dataset are usually too small and/or dedicated to a single specie (e.g. \cite{1, 2}).

We previously tackled this issue for indoor object grasping robots using a simulated environment and rendering engine \cite{3, 4}. We apply here a similar strategy for fruits harvesting problems by defining a framework capable of generating scenes of fruit trees coupled with procedural scripts controlling parameters (e.g. position of the fruits, type of background, lightning condition) to introduce realistic variations for outdoor data\cite{1}.

One of the most photo-realistic synthetic dataset of fruit trees is the work of Barth et al. \cite{2} for sweet pepper. This come at the price of a huge computational cost (10 min/frame with a 16 core processor) preventing it to be easily extendable in order to produce large dataset for other fruits. In fact, such high degree of photo-realism does not seem to be needed for deep simulated learning \cite{5}. That is why we aim at only an adequate photo-realism with quick rendering and easy-generation method, allowing scientists to create their own fruits or tree dataset according to their precise research interests.

2 Material and Methods

We chose the open-source Blender as the 3D-modelling and rendering software. Every options accessible through the Blender GUI can be reached and modified through a Python API, enabling us to entirely control the simulation with scripts.

The framework can be decomposed in two steps : first the generation of the tree or plant models, then a rendering script.

\footnotesize{Source code available at https://github.com/tduboudi/IAMPS2019-Procedural-Fruit-Tree-Rendering-Framework}
Figure 2: Schema of the camera point of view generation. Camera origin $O$ and target point $T$ are uniformly picked from respectively the hollow external cylinder $C_e$ and the internal cylinder $C_i$ (centered around the model).

to be continuous, we can also create robotic-like continuous trajectories toward branches on a tree. In addition, depth of field and rendering-engine parameters (e.g. number of rays for the ray-tracing engine) can be randomly modified within a certain range every time-step.

Background and lightning are also controlled in the rendering script, and can be easily changed every few time-steps to generate diverse images. This is a much needed feature allowing Domain Randomization, known to be efficient in reducing the Reality Gap encountered when transferring deep neural networks from simulation to the real world [7]. In order to have complex backgrounds we use freely available spherical HDRIs (high-dynamic-range 360° images), which control the lightning and the background objects. The tree models are placed in the center of the HDRIs.

3 Results and Discussion

Overall, the rendering of one pair of images (512x512 pixels, raw and ground truth for semantic segmentation) is fairly quick with ~[10s:30s] with GPU-rendering (NVIDIA GTX 1080). Fig[1] illustrates different possible variations for the scene generation (e.g. fruits, light) and the quality of the rendering.

4 Conclusion and Future Work

We designed a framework for semantic segmentation and object recognition field related to robotics fruit harvesting problems. It allows a quick and efficient fruit trees scene generation with parametrized variations, thus producing labelled images embedding a Domain Randomization aspect.

We aim to extend our framework to reproduce the variations occurring during the lifetime of the fruits and tree. For instance, an implementation of the fruit senescence and decay [8] unlocks the training of mature fruit picking robot and disease detection systems. We also plan to improve the photo-realism of the simulated image by using a Generative Adversarial Network as in [9], allowing to keep the rendering time needed quite low.

Acknowledgment

This work is supported by the french National Research Agency (ANR), through the ARES labcom (grant ANR 16-LCV2-0012-01) and by the CHIST-ERA EU project "Learn-Real".

References


