2nd RSEEME Engineering Projects Showcase
ENGN4200 Individual Research Project
May 2019

Australian National University

An Automated System for 3D Modelling and Feature Extraction of Small-scale Objects; Combining Computer Vision and Deep Learning

Student: Hengjia Li
Email: u5629478@anu.edu.au

Supervisor: Dr. Chuong Nguyen, Dr. Marnie Shaw

Examiner:

Scope

The project aims to produce:

- an image-based 3D modelling system for small-scale objects, for example insects
- a CNN-based framework that implements visual attribute grounding i.e. for indicating important features of target images

Previous Work

A 3D nature-color capturing device [1] and DISC3D [3] are existing solutions for performing the digitization of small-scale objects.

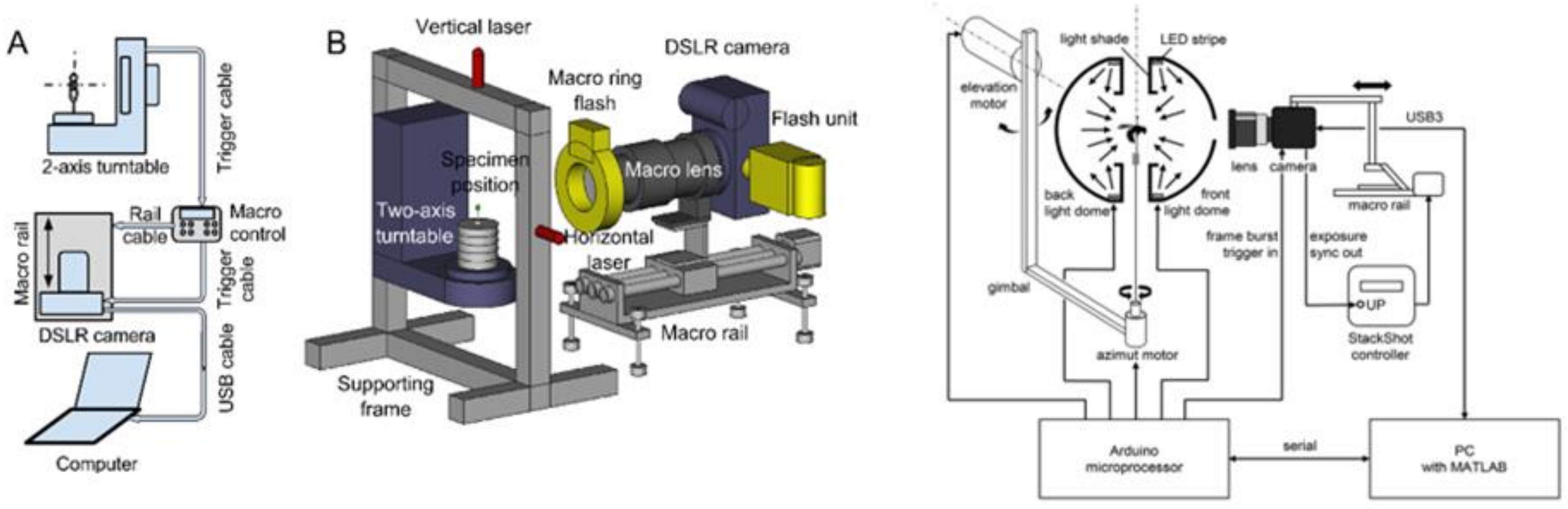

Figure 1: Previous work of image based 3D modelling systems. Left: 3D Nature-Colour Capturing Device [1]. Right: DISC3D [3]

CAM and Grad-CAM [2] are two example applications that localize important regions of an image, reflecting the network's decision.

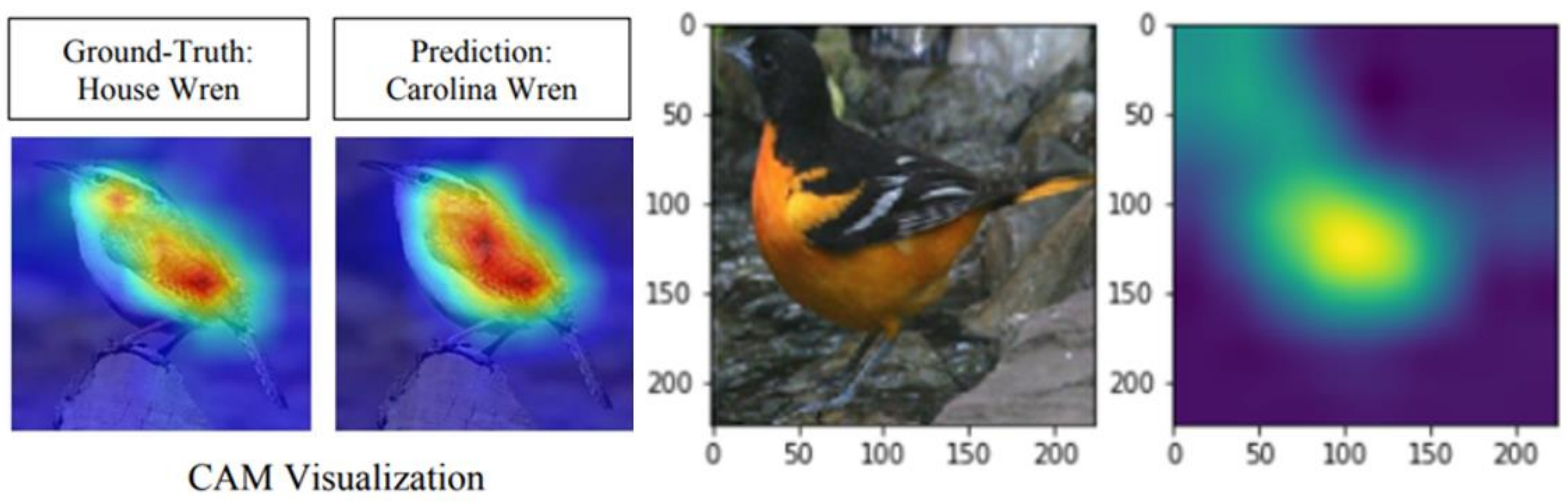

Figure 2: CAM [2] and visual attribute grounding visualizations.

Methodology

We implemented our own 3D modelling setup in the lab.

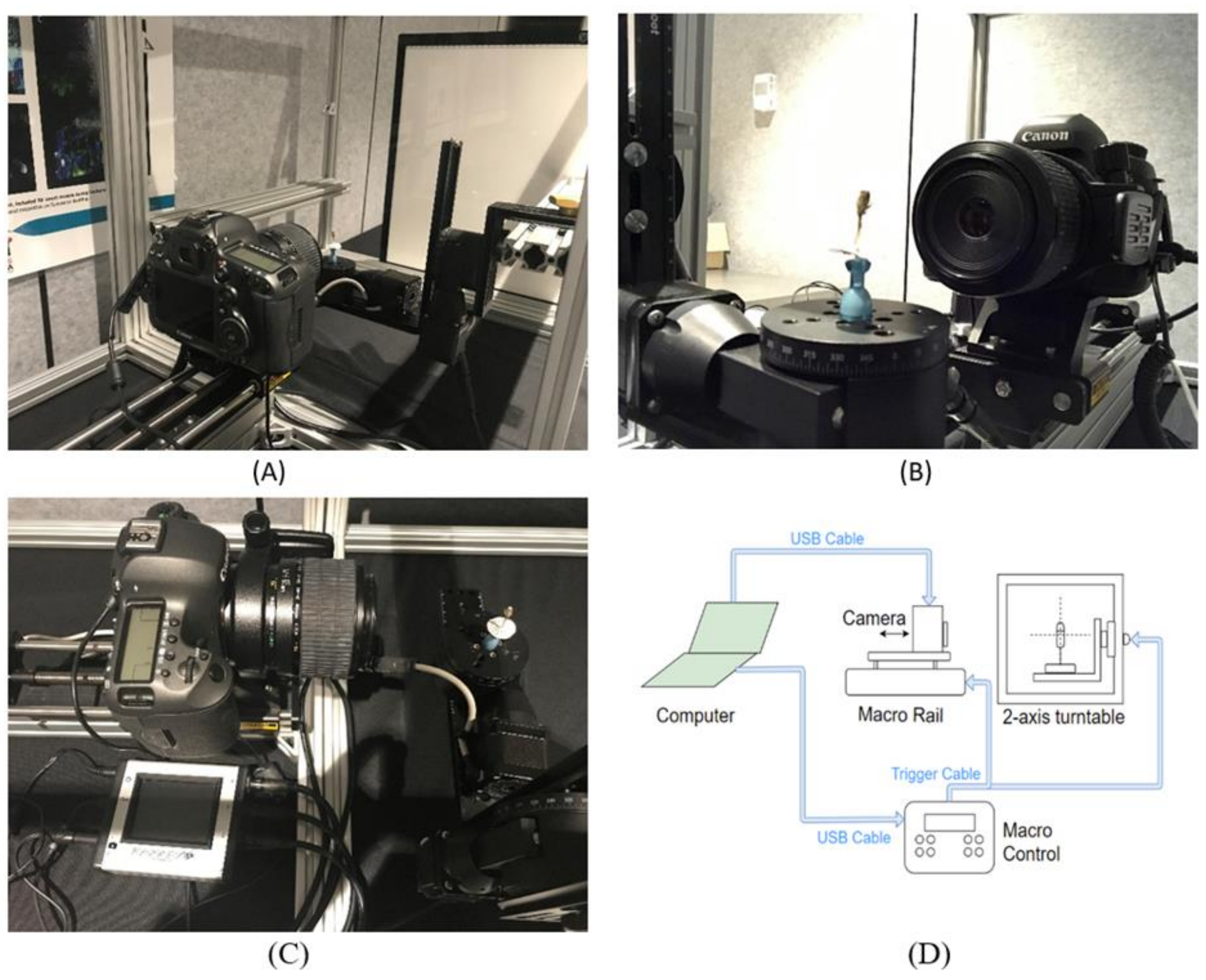

Figure 3: **Experiment setup. Figure (A), (B) and (C): pictures of the 3D modelling system. (D): schematic setup**

Core components:

- A Canon EOS DSLR camera with a macro lens
- A motorized macro-rail
- Two rotatory turntable
- A inbuilt controller
- A PC and MATLAB

Image Blending

1. Calibration by dot matrix
2A. Compare local saliency maps and then perform weighted-sum for fusion
2B. Run guided filter to optimize the local saliency maps, then perform weighted sum for fusion.

Visual Attribute Grounding

1. Finetune pre-trained Resnet-50 model.
2. Develop a Bayesian Inference framework that accumulates activations of last convolutional-layer filters, when the target attribute is fed in.

Results

Image Acquisition

Figure 4 : Image Blending Output

We compared different blending methods to visually and numerically test their performance:

Figure5: Comparison of blending output from simple saliency measurement (left), saliency measurement with dilation (middle) and guided filtering approach (right)

	Q_{MI}	rSFe	$N^{AB/F}$	Processing Time
Saliency measurement Method	1.0471	-0.1513	0.0765	49.916s
Saliency Measurement Method with dilation	1.0699	-0.3090	0.0446	55.163s
Guided Filtering Method	1.3150	-0.0717	0.0045	101.595s

Table 1: Numerical comparison between blending methods. Q_{MI}: indicate mutual information between source and blended images; rSFe: quality of blended image according to spatial frequency; $N^{AB/F}$: fusion artifacts.

Visual Attribute Grounding

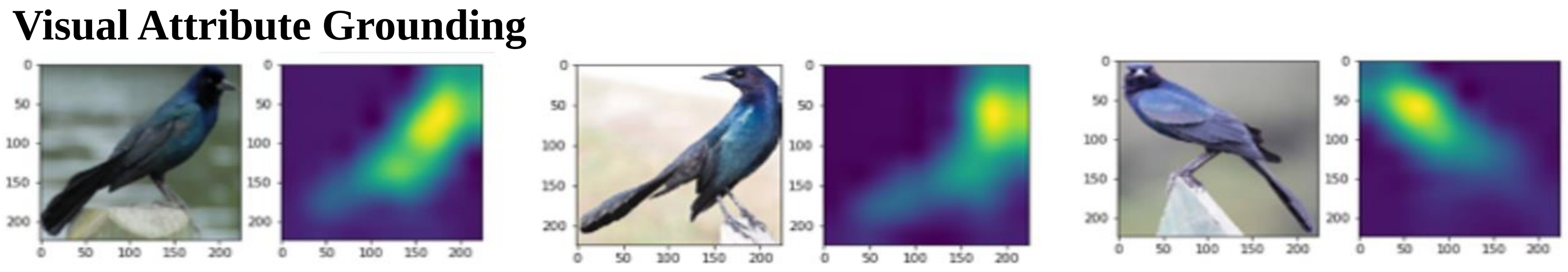

Figure 6: Grounding at blue chest

Factors that may effect detection performance:

- The natural shapes of wings/legs make these attributes harder to detect.
- The latent patterns that exist in training images. Human beings typically do not visually observe these patterns, but CNNs detect them. The vector angles between pattern-space activations were computed to prove the existence of latent patterns.

Potential solutions:

- Refine the CUB labels.
- Monitor the vector angles to detect latent patterns and the corresponding filters.

Conclusion

This research produces a proof of concept prototype for an image-based 3D digitization system, which is also capable of targeting important features of capture images. Further improvements include:

- Using a tilt-shift lens to reduce vibration during camera motion
- Investigating vector angles between pattern space activations to improve performance of feature detection

References

[1] Nguyen, C. V., Lovell, D. R., Adcock, M. & Salle, J. L., 2014. *Capturing Natural-Colour 3D Models of Insects for Species Discovery and Diagnostics.* [Online] Available at: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0094346 [Accessed 28 March 2019]

[2] Selvaraju, R. R. et al., 2016. *https://arxiv.org/abs/1512.04150.* [Online] Available at: https://arxiv.org/abs/1610.02391 [Accessed 26 March 2019]

[3] Ströbel, B., Schmelzle, S., Blüthgen, N. & Heethoff, M., 2018. *An automated device for the digitization and 3D modelling of insects, combining extended-depth-of-field and all-side multi-view imaging.* [Online] Available at: https://zookeys.pensoft.net/article/24584/element/4/430// [Accessed 28 March 2019]

Research School of Electrical, Energy and Materials Engineering
ANU College of Engineering & Computer Science