

Comparing and stress testing ICON and PIFu

Andrew Morgan

February 9, 2023

Abstract

ICON stands for Implicit Clothed humans Obtained from Normals and is a modern way of generating 3D representations of humans from a single RGB image. PIFu stands for Pixel-Aligned Implicit Function and has the same functionality.

Links to the papers: [PIFu](#) [ICON](#).

1 Comparison

We start by comparing the run-times of both methods, ICON roughly averaged a 3 minute run-time and PIFu roughly averaged a 2 minute 30 second run-time (using an Amazon EC2 T4g Instance). Now lets look at the different cases and how each algorithm performed:

1.1 Simple Pose



Figure 1: The left is a picture of Joel, the right is of Dan. Both with a similar pose from a similar angle

Number of vertices produced by each model (in thousands):

	ICON	PIFu
Joel	31	29
Dan	29	29

There is little difference in the two models generated in terms of vertices and overall shape, the two main differences between the outputted models are the representation of Joel's feet and Dan's head.

Figure 2 shows, ICON generates a strange jagged geometry, whereas PIFu is closer with a discernible difference between shin and foot.

Figure 3 shows, ICON puts the ear to far forward and ends up squishing the head to be slightly to short. PIFU seems to have the opposite problem placing the ear to far back making the head to long.

Figure 4 shows, ICON does a better job at texturing Joel's face with less areas that are left blank.

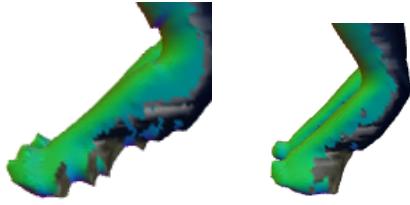


Figure 2: The left is from ICON, the right from PIFu

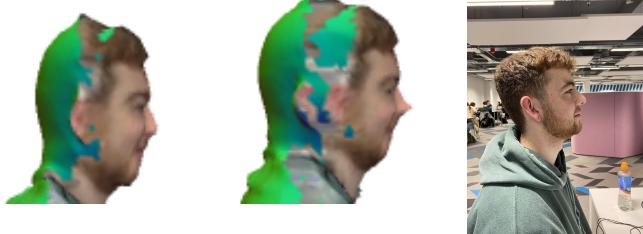


Figure 3: The left to right, this is ICON, PIFu and base truth



Figure 4: The left is from ICON, the right from PIFu



Figure 5: Dan in various poses, from left to right I'll refer to these as leg up, one eye and squat

1.2 Other Pose

Number of vertices produced by each model (in thousands):

	ICON	PIFu
Leg up	26	25
One eye	30	30
Squat	49	46

In figure 6 we see ICON separate Dan's right arm from his body and in figure 7 we see ICON separates Dan's hand from his head, in both of these cases this separation isn't present in the ground truth. It appears that ICON favours poses where the subject isn't in contact with themselves. In figures 6 and 7 we see both method struggling to identify shoes. Figure 8 shows that neither method could reproduce a the squat pose, ICON does outperform PIFu here as ICON gets the general shape

whereas PIFu has the image collapsing toward the middle. Interestingly the position outputted by PIFu is impossible for a human to do.



Figure 6: The leg up pose: left is from ICON, the right from PIFu



Figure 7: The one eye pose: left is from ICON, the right from PIFu



Figure 8: The squat pose: left is from ICON, the right from PIFu

1.3 Small Note on Occlusion

Both methods cannot handle occlusion of the body, producing random models. ICON does well with the leg occlusion producing a kneeling model, whereas PIFu struggles to capture the and produce a model.



Figure 9: On the left we have leg occlusion and on the right we have body occlusion



Figure 10: The leg occlusion: left is from ICON, the right from PIFu