CMA-ES and its variants applied on car shape optimisation by exploring differentiability of MeshSDF

Huajian Qiu

***Abstract*—The original CMA-ES can find smaller drag value than gradient descent optimiser. CMA-ES with SGD or line search doesn't give clear advantage than before mixing two methods. Adam has noisy and vibrant optimisation trail. The SGD gives continuous and smooth shape changing progress. All of these optimisers can converge to different final optimum if starting from different position.**

I. Definitions

$$\mathcal{L}_{phy} = \oint_S p(v)\, d\vec{s}_{x_0}(v)$$

$$v = v(\vec{z}), \vec{z} \in \mathbb{R}^{256}$$

$$v\, is\, vertex, p\, is\, pressure, s\, is\, surface\, area$$

I use physical loss to represent the loss caused by air pressure drag. Notice that I never observed the loss caused by the bounding space(reserving space for engine and driver) so far. Physical loss is what we finally wish to optimise if the predicted pressure is reliable. To prevent the optimiser from exploring the latent space that doesn't represent meaningful car shapes, it is critical to add a regularisation term to the full loss. A large regularisation loss usually indicates the current latent is far from all known car latents and then the shape is no longer a car.

$$\mathcal{L}_{full} = \mathcal{L}_{phy}(\vec{z}) + \mathcal{L}_{regular}(\vec{z})$$

$$\mathcal{L}_{regular}(\vec{z}) = distance(\vec{z}, top\, 10\, closest\, car\, latent\, to\, \vec{z})$$

II. Comparisons of different optimisers

I compared five different optimisers starting from same initial car latent code and with same constant learning rate. As shown in Figure 1, as to the lines of CMA-ES and its variants, I plot the loss of all its candidates in each generation by the sequence of function evaluation. So they appears more vibrant than the others. The group size of candidates is 5 and each candidate is optimised by 15 runs of inner optimiser. Overall, the running progress of SGD is most smooth and it converges to a certain loss and no longer decreases. CMA finds a smaller value (even negative drag loss) but the regularisation loss is very large in this case.

I also tried starting from different initial car latent code with SGD and CMA. They tended to end up with another different and reasonable car shape.

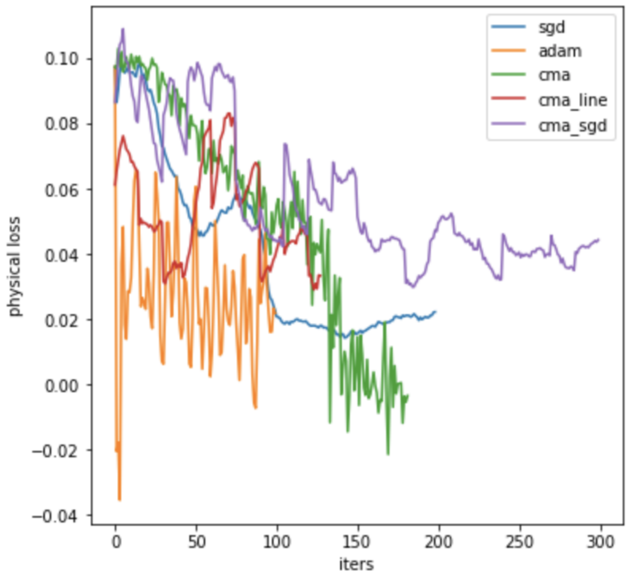

Fig. 1. Physical loss: optimisation process on the car shape latent space

III. Structure of latent space

The physical loss value doesn't necessary reflect the goodness of optimisation result, especially given the fact that a small physical loss sometimes doesn't represent a car. Therefore I begin another part of project to explore the structure of latent space.

The latent space is 256 dimension. I project all latents vector into 2 dimensional space by principle component analysis(PCA). It only reserves the most two significant dimension but still provides meaningful insights for trails of different optimisers. By applying PCA, 1205 cars each with 256 dimensional latent vectors are projected into a new 256D coordinates and I only reserve the most significant 2D for visualisation. The singular values describe the explained components along each axis. Here I report the 256 singulars in Figure 2 which show most of variance are explained by first few dimensions. Combined by the regularisation requirement that new explored latents should be as close to known latents as possible, we can conclude that every optimisers should explore these most important dimensions.

As an application of PCA dimension reduction, Figure 3 shows the runnings of SGD from three different starting car latent code. This figure indicates the coarse position of three

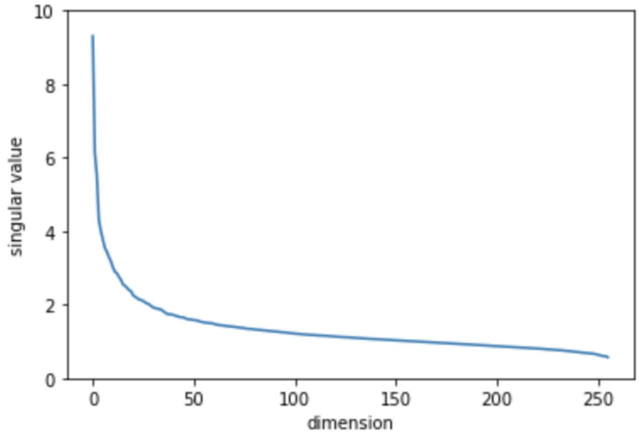

Fig. 2. Explained components by each dimension

local minimum. Figure 4 and Figure 5 shows the pattern of

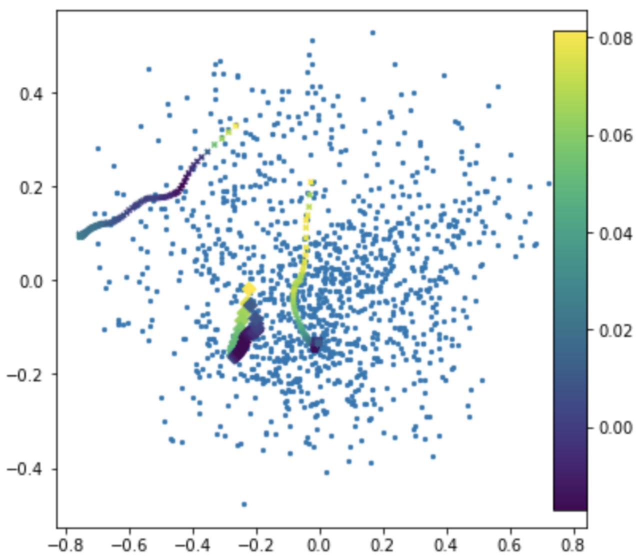

Fig. 3. latent trails from 3 different initial car latent, optimised by SGD. The blue points represent 1025 car latents used for training encoder and decoder, the gradually varied points represent the latent code after each iteration. The color of three trails represents the physical loss of latent.

CMA-ES optimisation. These two figures are generated by projecting same set of 256D latents into two different 2D space. Reading these two figures and the Table I, we can infer that after long enough optimisation, CMA optimiser only explores a few parts of original 1025 car latent space and it explore in a direction relatively perpendicular to original 1025 car latent space. This can be supported by the fact that at the end of optimisation, the most closet known latent is still the starting latent. This is a dangerous sign that the optimiser didn't explore the proper latent space enough and got close to the dangerous zone where latent codes don't represent meaningful car shape. Figure 6 shows the pattern of CMA-ES with SGD, which is a mixture of CMA and gradient method and an analog of CMA-line-search. CMA-line-search

TABLE I
CMA OPTIMISATION RESULTS STARTING FROM 424TH CAR LATENT.

closest latents	most closest	2nd	3rd	4th	5th
latent index	424	732	178	612	22
penalty	1.160	1.169	1.182	1.202	1.217

shows promising potential on benchmark functions and is the starting motivation of this project. But in the context of the car shape optimising problem, this method fails because it shows no advantages over CMA or SGD. Figure 6 shows locally this optimiser has desired property of smoothing running but globally it has no much difference with original CMA. What's worse, it takes much more time than original CMA (if inner optimiser SGD runs 15 steps, then the mixed method takes 15-fold times than CMA), which seems not worthwhile.

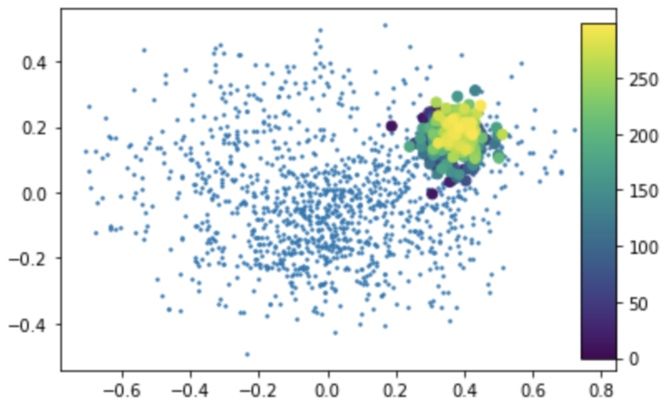

Fig. 4. CMA optimisation results starting from 424th car latent. All latents are projected to maximize the variance of original 1205 car latents. The color of three trail represents the order of latent generated by CMA optimisers

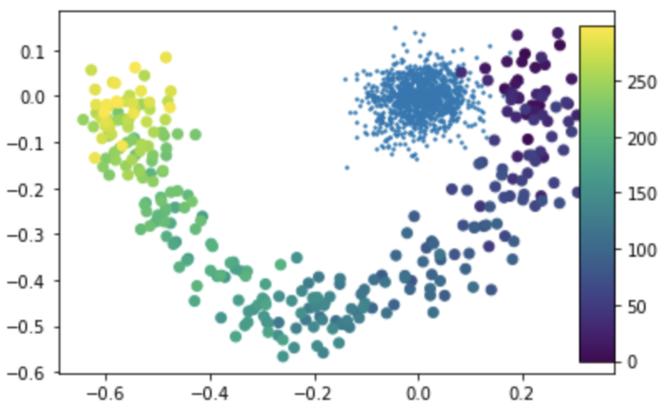

Fig. 5. CMA optimisation results starting from 424th car latent. All latents are projected to maximize the variance of latents generated by optimisers. The color of three trail represents the order of latent generated by CMA optimisers

IV. VISUALISATION OF SHAPING EVOLUTION

This section Figure 7 and Figure 8 shows again the importance of suitable regularisation term. As a general rule,

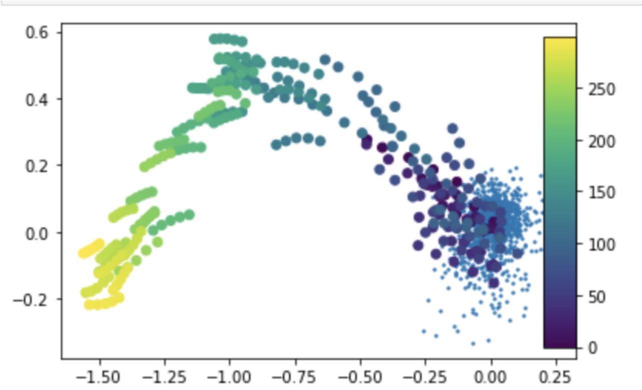

Fig. 6. CMA-SGD optimisation results starting from 32th car latent. All latents are projected to maximize the variance of latents generated by optimisers. The color of three trail represents the order of latent generated by CMA optimisers

CMA can easily be attempted to explore the latent space far from known car shapes, where a very small even negative drag loss is common to find. This can be seen as an undesired artifact of latent space structure. Gradient descent methods are rarely disturbed by this property for unknown reasons. But this artifact affect stochastic-search based method badly.

As an another remark, the shapes found by CMA are usually located at the edge of valid latent space. They are close to become shapes that is no longer a car as shown in Figure 9

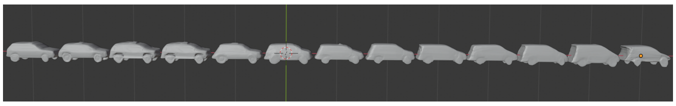

Fig. 7. one successful example of CMA optimisation, with reasonable regularisation

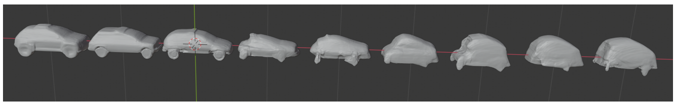

Fig. 8. one failed example of CMA optimisation, without any regularisation, it finds very small drag loss but no longer a car

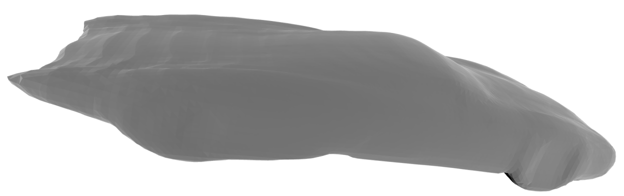

Fig. 9. an example shape after CMA optimisation, with very small drag loss but nearly losing the shape of car

V. Conclusion

In this project, the CMA doesn't display its advantage as an global optimiser. It still only explore parts of domain and tends to diverge from known car shapes. Choosing a larger initial variance can help only partly. Because if the initial variance is too big, then it is prone to throw exceptions because no mesh can be got from marching algorithm on SDF values.

CMA-ES with line search, SGD and Adam has similar problem as CMA-ES.

I infer the latent space is ill-defined and afflicted by trivial minimum, which means we need to add some carefully-designed regularisation to ensure the found minimum is really meaningful. Otherwise although we find a very small minimum, it just doesn't represent a car.