Hyperparameter-Free Losses for Model-Based Monocular Reconstruction


Eduard Ramon


Guillermo Ruiz


Thomas Batard


Xavier Giró-i-Nieto

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

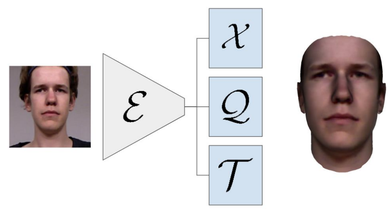
ICCV 2019 Seoul, Korea

Abstract

This work proposes novel hyperparameter-free losses for single view 3D reconstruction with morphable models (3DMM). We dispense with the hyperparameters used in other works by exploiting geometry, so that the shape of the object and the camera pose are jointly optimized in a sole term expression. This simplification reduces the optimization time and its complexity. Moreover, we propose a novel implicit regularization technique based on random virtual projections that does not require additional 2D or 3D annotations. Our experiments suggest that minimizing a shape reprojection error together with the proposed implicit regularization is especially suitable for applications that require precise alignment between geometry and image spaces, such as augmented reality. We evaluate our losses on a large scale dataset with 3D ground truth and publish our implementations to facilitate reproducibility and public benchmarking in this field.

Motivation

Learning monocular 3D reconstruction



- X: Geometry
- Q: Camera rotation
- T: Camera translation

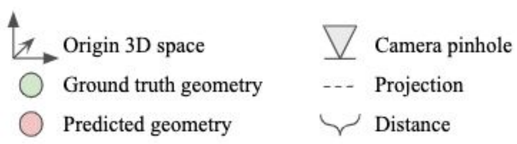
Traditional multiterm losses

$$\mathcal{L}_{Coarse} = ||\boldsymbol{x} - \hat{\boldsymbol{x}}||_2^2 + \alpha||[\boldsymbol{q}, \boldsymbol{t}] - [\hat{\boldsymbol{q}}, \hat{\boldsymbol{t}}]||_2^2$$

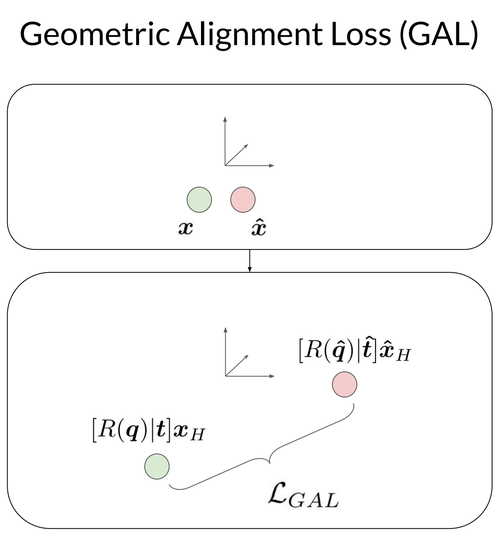
$$\mathcal{L}_{XQT} = ||\boldsymbol{x} - \hat{\boldsymbol{x}}||_2^2 + \beta||\boldsymbol{q} - \hat{\boldsymbol{q}}||_2^2 + \gamma||\boldsymbol{t} - \hat{\boldsymbol{t}}||_2^2$$

How to avoid the hyperparameters α, β and γ ?

Method: Hyperparameter-free losses

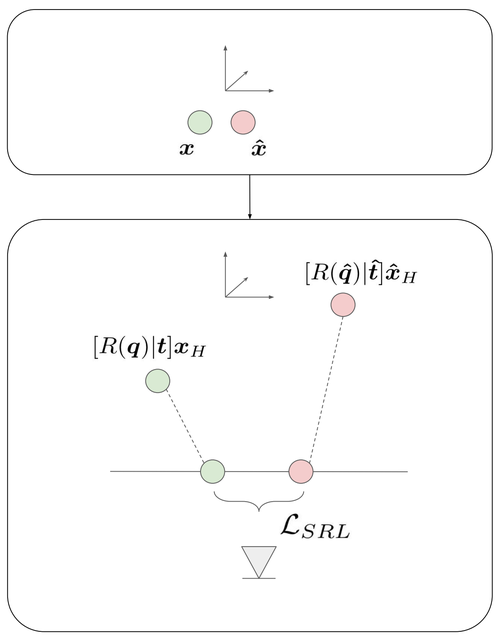


Geometric Alignment Loss (GAL)



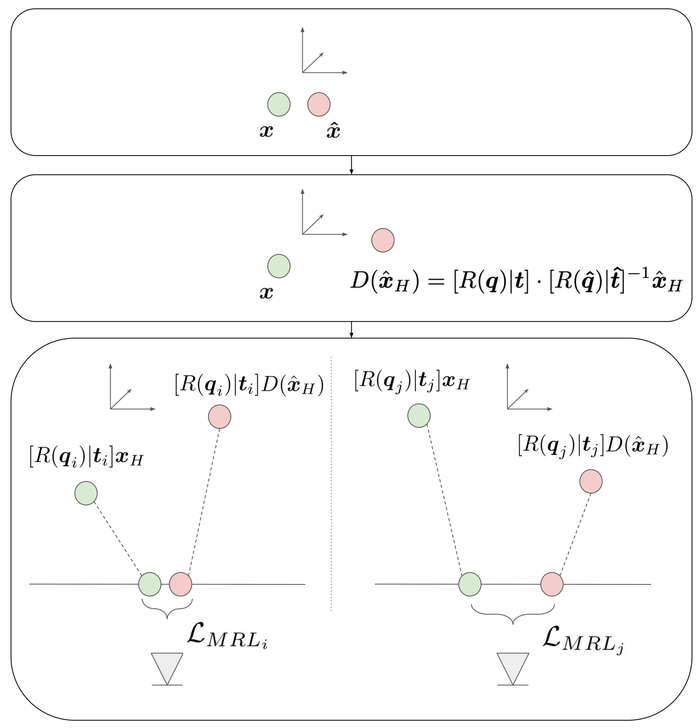
$$\mathcal{L}_{GAL} = ||[R(\boldsymbol{q})|\boldsymbol{t}]\boldsymbol{x}_H - [R(\hat{\boldsymbol{q}})|\hat{\boldsymbol{t}}]\hat{\boldsymbol{x}}_H||_1$$

Single View Reprojection Loss (SRL)



$$\mathcal{L}_{SRL} = ||\mathcal{P}(\boldsymbol{q}, \boldsymbol{t})(\boldsymbol{x}_H) - \mathcal{P}(\hat{\boldsymbol{q}}, \hat{\boldsymbol{t}})(\hat{\boldsymbol{x}}_H)||_1$$

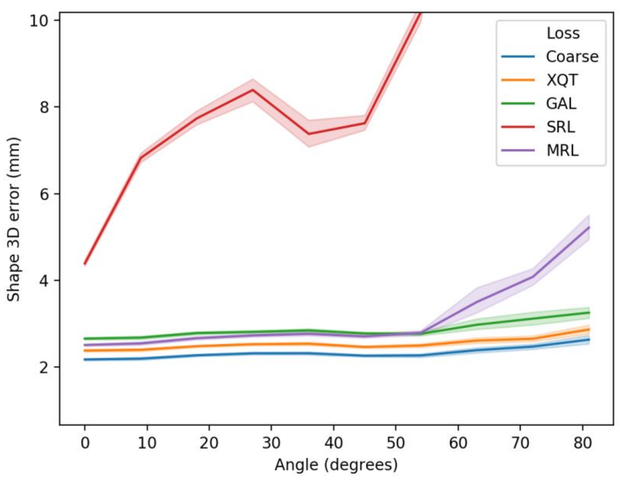
Multiview Reprojection Loss (MRL)

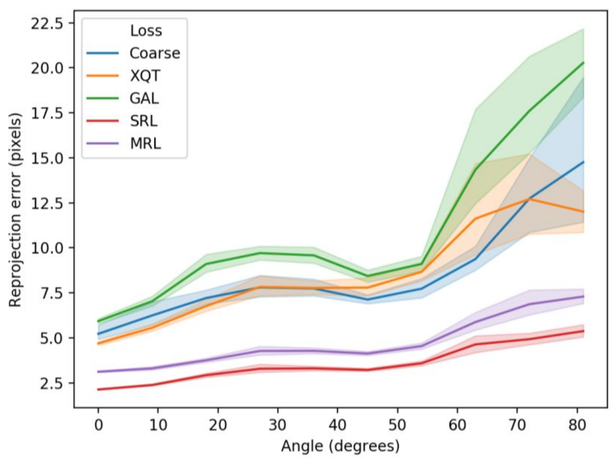


$$\mathcal{L}_{MRL} = \sum_{v=1}^{V} ||\mathcal{P}(\boldsymbol{q_v}, \boldsymbol{t_v})(\boldsymbol{x}_H) - \mathcal{P}(\boldsymbol{q_v}, \boldsymbol{t_v})(D(\hat{\boldsymbol{x}}_H))||_1$$

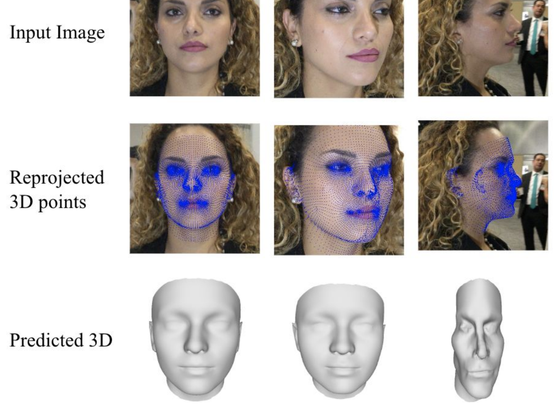
Results

Robustness against large poses

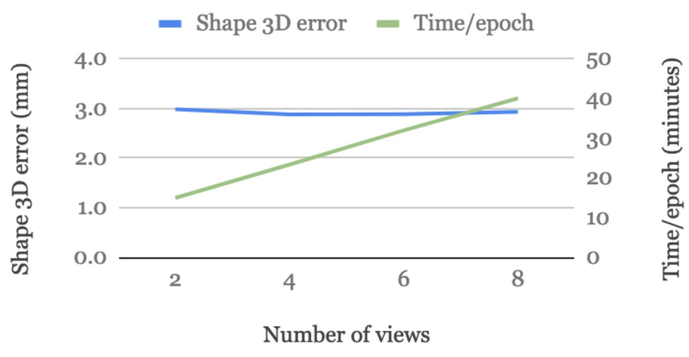




Effects of no regularization in 3D shape when using SRL



Shape accuracy and training time versus number of random views in MRL



Qualitative 3D shape error comparison