Plug-and-Train Loss for Single View 3D Reconstruction

Thomas Batard

Xavier Giró-i-Nieto

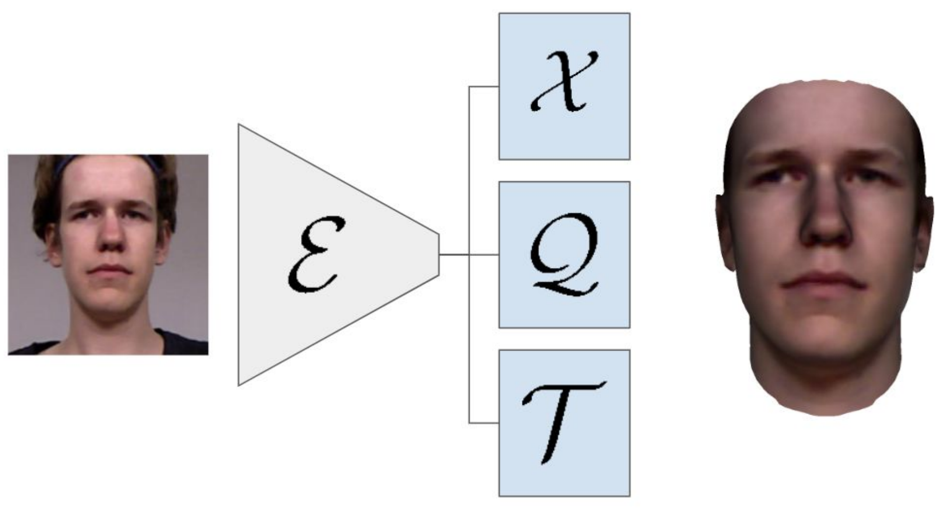
UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

crisalix VIRTUAL AESTHETICS

DOCTORATS INDUSTRIALS

Motivation

The single view reconstruction problem



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

Multi-term losses used for learning the mappings X, Q and T

$$\mathcal{L}_{data} = \mathcal{L}_x + \alpha\mathcal{L}_q + \beta\mathcal{L}_t$$

$$\boldsymbol{x} = \boldsymbol{m} + \Phi_{id}\boldsymbol{\alpha}_{id}$$

$$\mathcal{L}_{reg} = \gamma||\boldsymbol{\alpha}_{id}||_2^2$$

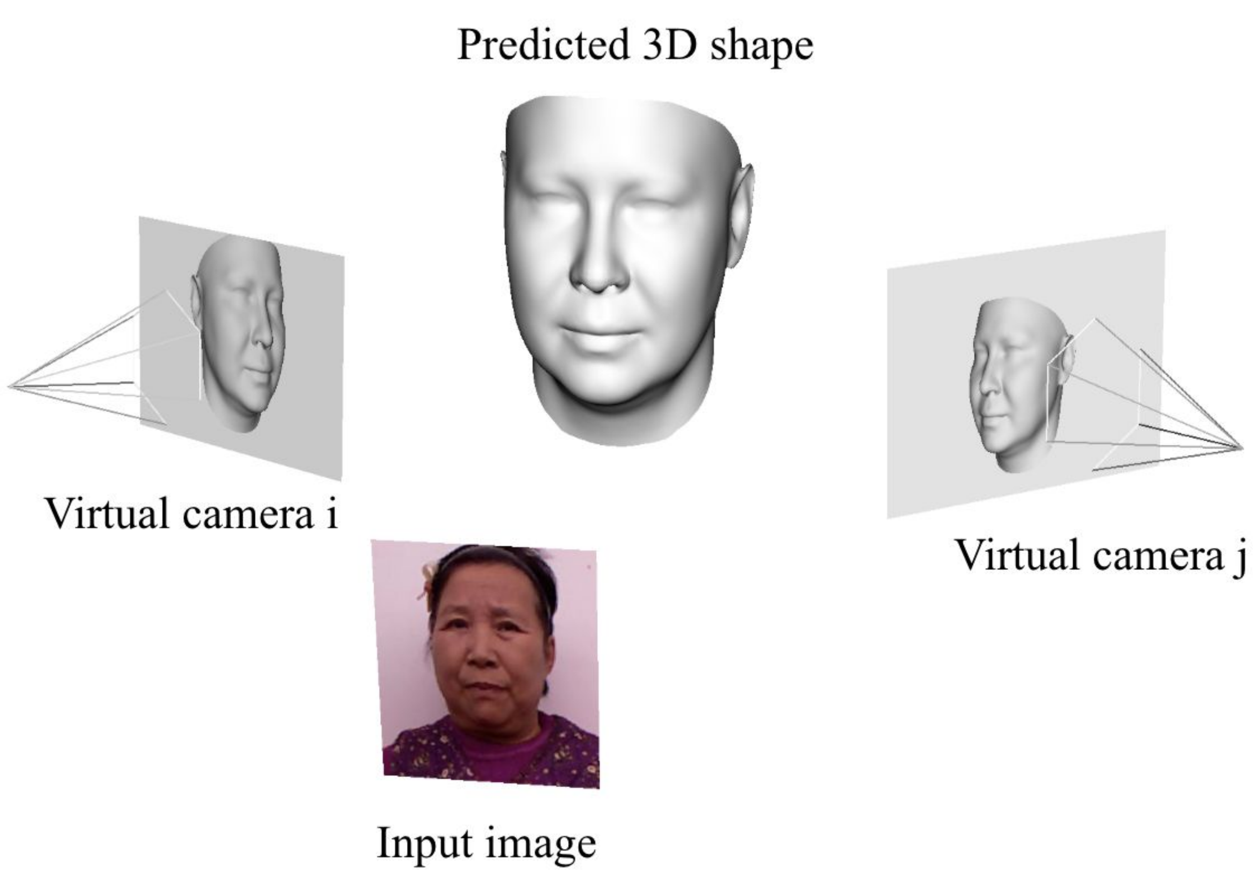
$$\mathcal{L} = \mathcal{L}_{data} + \mathcal{L}_{reg}$$

Which hyperparameters α, β and γ ?

Drawbacks:

- Problem dependent losses.
- Need of tuning a number of hyperparameters.
- Non optimal solutions.
- Waste of time and computational resources.

Multiview Reprojection Loss



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

$$\mathcal{P}_D(\boldsymbol{q}, \boldsymbol{t}) = K[R(\boldsymbol{q})|\boldsymbol{t}]D$$

$$D = [R(\boldsymbol{q})|\boldsymbol{t}] \cdot [R(\hat{\boldsymbol{q}})|\hat{\boldsymbol{t}}]^{-1}$$

Experiments

Implementation details:

- VGG16 encoder
- Training dataset with 6k facial scans and 50k images

Single term vs Multi-term losses

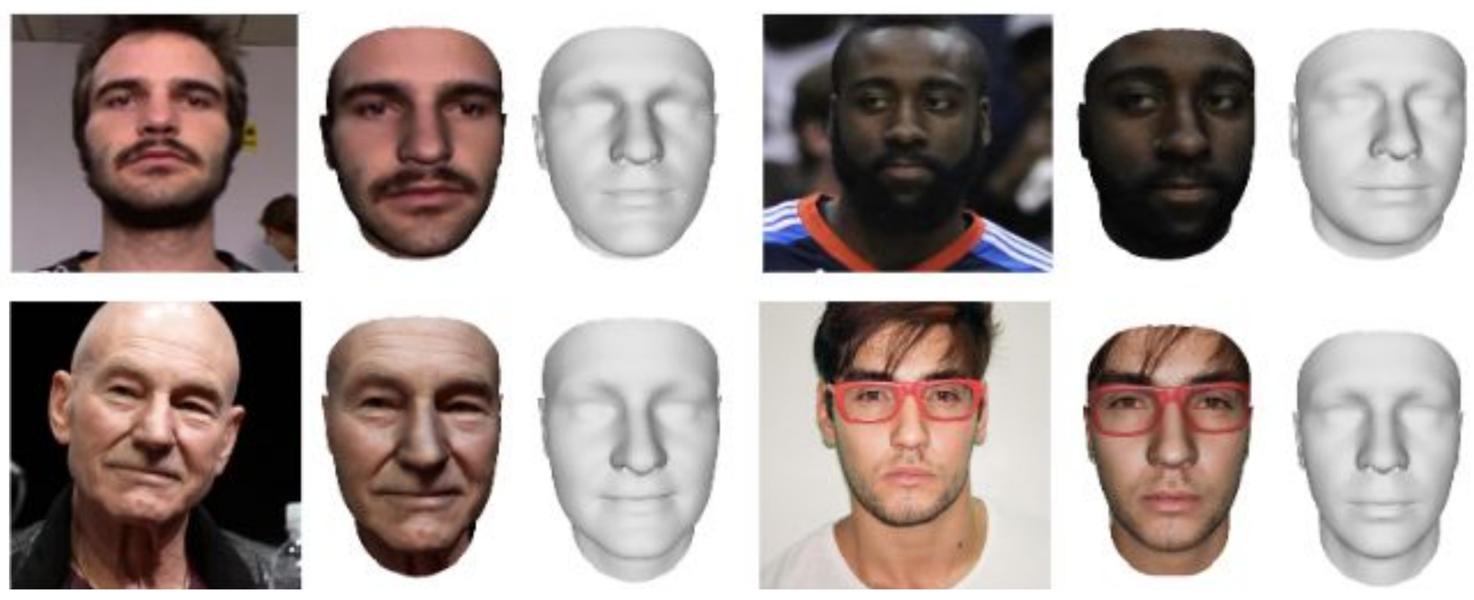
	Repro-jection (pixels)	Shape 3D (mm)	Camera position (mm)	Camera rotation (degrees)
Best pose	5.4	1.7	**2.39**	**2.51**
Best shape	8.58	**1.5**	2.50	2.62
MRL (ours)	**3.29**	1.7	2.86	3.04

Results on MICC dataset

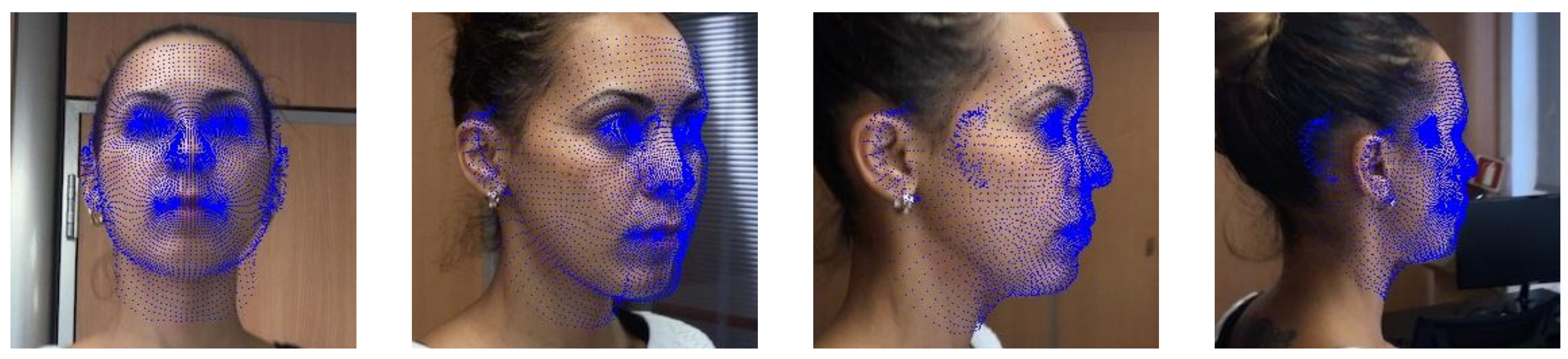
(WARNING: Our MRL results are over a subset of faces without expression)

Method	3DRMSE
3DMM [24]	$1.75 \pm .42$
Flow-based [11]	$1.83 \pm .39$
Discriminative [30]	$1.57 \pm .33$
MRL (ours)	$\mathbf{1.47 \pm .30}$

Robustness against face diversity



Robustness against large poses



Conclusions and future work

Conclusions:

- MRL can train single view reconstruction models without adding hyperparameters.
- Comparable performance against manually tuned losses.
- Robust models against input diversity.

Future work

- Evaluation of the method with non model-based geometries.
- Evaluation of single term losses in 3D space.