Discovering Geometric Biases in 3D Face Reconstruction: A Curvature-Aware Spectral Framework for Fairness Evaluation

Veronika Shilova[1,2], Emmanuel Malherbe[1], Giovanni Palma[3], Panagiotis-Alexandros Bokaris[3], Laurent Risser[2], Jean-Michel Loubes[2,4,5]

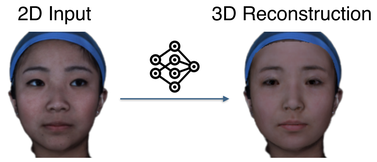
[1]Artefact Research Center, [2]Institut de Mathématique de Toulouse, [3]L'Oréal Recherche et Innovation, [4]Université de Toulouse, [5]INRIA

ECCV 2026 | MALMÖ | SEPT 8–12

Project page →

Problem Setting & Motivation

3D Face Reconstruction from a single 2D photo



3D Morphable Model (3DMM)

$$\mathscr{V} = \overrightarrow{\mu} + \sum_{j=1}^{n} \beta_j \overrightarrow{\psi_j},$$

with $\overrightarrow{\mu} \in \mathbb{R}^{3N_v}$ the mean face, $\overrightarrow{\psi_j}$ PCA basis and β_j the shape parameters, $j = 1,\ldots,N_{PCA}$.



Basel Face Model (BFM)
$N_v = 53{,}490$,
$N_{PCA} = 199$

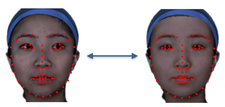


FLAME
$N_v = 5023$,
$N_{PCA} = 400$

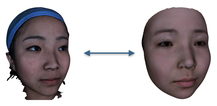
- Finite Training Data: BFM: 200 scans, FLAME: 3,800 scans
- Mostly young Caucasian adults

Reconstruction Errors

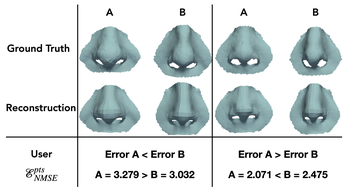
2D Landmarks Positions



Distance between Surfaces

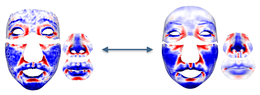


Comparison of Reconstruction Errors



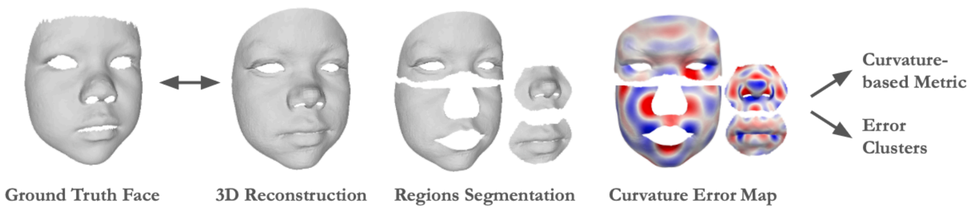
Contributions

Difference between Mean Curvatures



1. Curvature-Aware Evaluation Framework
2. Perceptual Validation
3. Systematic Bias Discovery

Methodology



What we had before: NMSE

$$\mathscr{E}^{pts}_{NMSE}(\mathscr{M},\mathscr{M}') = \frac{1}{|\mathscr{V}'|} \sum_{v \in \mathscr{M}'} ||v - T^{pts}_{\mathscr{M}'\to\mathscr{M}}(v)||^2,$$

where $T^{pts}_{\mathscr{M}'\to\mathscr{M}}$ is point-to-surface mapping.

- × Point-to-surface distance
- × Global fidelity only
- × Insensitive to ridges, creases, folds
- × Near-random alignment with human judgment (accuracy $\approx 42-54\,\%$)
- × Detects NO significant age or ethnicity bias

What we propose: Curvature Reconstruction Error (CRE)

$$\mathscr{E}_{CRE}(\mathscr{M},\mathscr{M}') = \frac{1}{Tr(M)} \sum_{v \in \mathscr{M}'} \left| H^{\mathscr{M}'}(v) - H^{\mathscr{M}}\left(T^{pts}_{\mathscr{M}'\to\mathscr{M}}(v)\right) \right| M_v,$$

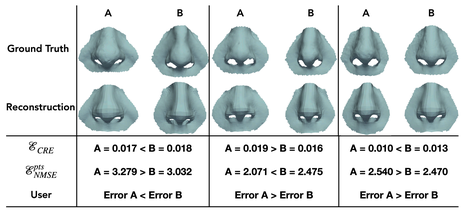
where $H^{(\cdot)}$ is mean curvature, M is the diagonal mass matrix whose entries M_v correspond to the barycentric area for each vertex $v \in \mathscr{V}'$.

- ✓ Curvature difference
- ✓ Sensitive to local surface geometry
- ✓ Captures ridges, folds, undulations
- ✓ Aligns with human judgment (accuracy $\geq 73\,\%$)
- ✓ Detects significant age bias and preliminary detects ethnicity bias

Experiments & Results

RQ1: Perceptual Validation — User study

Protocol
- 19 annotators, diverse backgrounds
- Pairwise comparisons of nose reconstructions
- 2,414 pairs → 1,085 labeled (≥75% agreement, ≥5 votes)
- BFM reconstructions of REALY faces

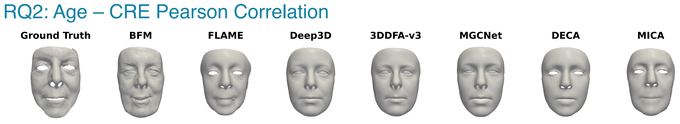


Curvature Reconstruction Error (CRE) better aligns with human perception of reconstruction quality

Metric	Mapping	R	K	Norm	Pairwise-Accuracy, %	Pairwise-ROC-AUC, %
$\mathcal{E}^{REALY}_{NMSE}$	REALY	–	–	–	53.92	56.53
$\mathcal{E}^{pts}_{NMSE}$	Points-to-surface	–	–	–	49.32	55.02
$\mathcal{E}^{vtx}_{NMSE}$	Vertex-to-vertex	–	–	–	42.03	42.47
$\mathcal{E}_{CRE}$	Points-to-surface	**3**	**128**	L_1	**73.63**	**82.66**
$\mathcal{E}_{CRE}$	Points-to-surface	1	128	L_1	71.38	80.63
$\mathcal{E}_{CRE}$	Points-to-surface	1	512	L_1	68.32	78.17
$\mathcal{E}_{CRE}$	Points-to-surface	1	1024	L_1	66.52	74.35
$\mathcal{E}_{CRE}$	Points-to-surface	2	128	L_1	72.46	81.86
$\mathcal{E}_{CRE}$	Points-to-surface	2	512	L_1	68.77	77.46
$\mathcal{E}_{CRE}$	Points-to-surface	2	1024	L_1	66.79	77.46
$\mathcal{E}_{CRE}$	Points-to-surface	3	512	L_1	70.66	79.36
$\mathcal{E}_{CRE}$	Points-to-surface	3	1024	L_1	69.85	78.83
$\mathcal{E}_{CRE}$	Points-to-surface	4	128	L_1	73.00	81.85
$\mathcal{E}_{CRE}$	Points-to-surface	4	512	L_1	70.93	79.40
$\mathcal{E}_{CRE}$	Points-to-surface	4	1024	L_1	70.66	79.17
$\mathcal{E}_{CRE}$	Points-to-surface	3	128	L_2	72.28	81.17
$\mathcal{E}_{CRE}$	Points-to-surface	3	512	L_2	69.67	77.81
$\mathcal{E}_{CRE}$	Points-to-surface	3	1024	L_2	68.14	76.99
$\mathcal{E}_{CRE}$	Points-to-surface	3	128	*Huber*	67.60	75.82
$\mathcal{E}_{CRE}$	Points-to-surface	3	512	*Huber*	62.74	70.40
$\mathcal{E}_{CRE}$	Points-to-surface	3	1024	*Huber*	61.57	69.51

Fairness Insights

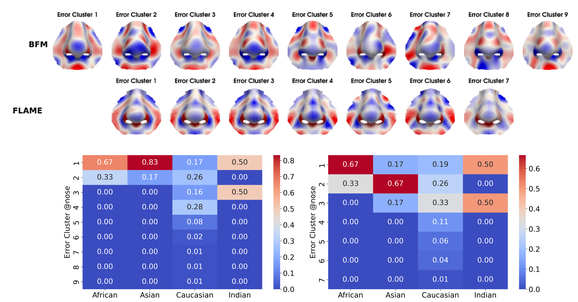
RQ2: Age – CRE Pearson Correlation



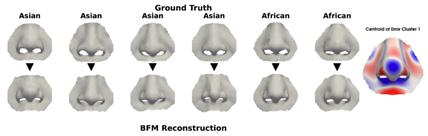
- CRE detects bias w.r.t. age, whereas NMSE-based metrics cannot capture it

		Pearson correlation: $\mathcal{E}_{CRE}$ vs. 3D Reconstruction											
		@nose			@mouth			@cheek			@forehead		
		All	Female	Male	All	Female	Male	All	Female	Male	All	Female	Male
3DMM	BFM	0.39	0.32	**0.59**	0.33	0.21	**0.67**	**0.77**	**0.72**	**0.88**	**0.52**	0.49	**0.59**
Repr.	FLAME	**0.78**	**0.79**	**0.78**	-0.03	-0.16	0.18	**0.82**	**0.84**	**0.82**	0.41	0.48	0.25
	Deep3D	0.24	0.18	0.41	-0.23	-0.19	-0.38	**0.69**	**0.81**	**0.62**	**0.58**	**0.54**	**0.66**
Model	3DDFA-v3	**0.67**	**0.65**	**0.73**	0.31	0.30	0.37	**0.79**	**0.78**	**0.88**	**0.53**	**0.55**	0.48
from	MGCNet	**0.54**	**0.50**	**0.67**	0.30	0.30	0.32	**0.80**	**0.79**	**0.89**	**0.51**	**0.54**	0.47
2D image	DECA	**0.73**	**0.75**	**0.70**	-0.32	-0.30	-0.35	**0.76**	**0.80**	**0.85**	0.40	0.46	0.28
	MICA	**0.69**	**0.72**	**0.65**	-0.31	-0.31	-0.30	**0.79**	**0.76**	**0.88**	0.43	0.48	0.31

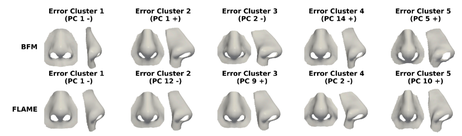
RQ3: Ethnicity with respect to Error Clusters



- Asian and African participants share a specific failure mode.



RQ4: Morphology Traits with respect to Error Clusters



- 3DMM failure modes are systematic rather than stochastic.

Conclusion

1. CRE aligns better with human perception of reconstruction quality. It achieves 74% of perceptual accuracy (vs 0.42–0.54 for NMSE)
2. CRE detects bias w.r.t. age which is strongest in cheek region (Pearson correlation up to 0.88).
3. CRE offers preliminary evidence of gender, ethnic, and morphological biases.

These findings suggest that 3DMM representational bottlenecks are inherent to their linear prediction subspaces rather than a failure of 2D-to-3D inference.