Mesh registration via geometric feature homogenization and offset cross-attention: application to 3D photogrammetry

Inés A. Cruz-Guerrero[1], Connor Elkhill[1], Jiawei Liu[1], Phuong Nguyen[2], Brooke French[2], and Antonio R. Porras[1, 2, 3, 4]

[1] Department of Biostatistics and Informatics, Colorado School of Public Health, University of Colorado Anschutz Medical Campus
[2] Department of Pediatric Plastic and Reconstructive Surgery, Children's Hospital Colorado
[3] Department of Pediatric Neurosurgery, Children's Hospital Colorado
[4] Departments of Pediatrics, Surgery and Biomedical Informatics, School of Medicine, University of Colorado Anschutz Medical Campus

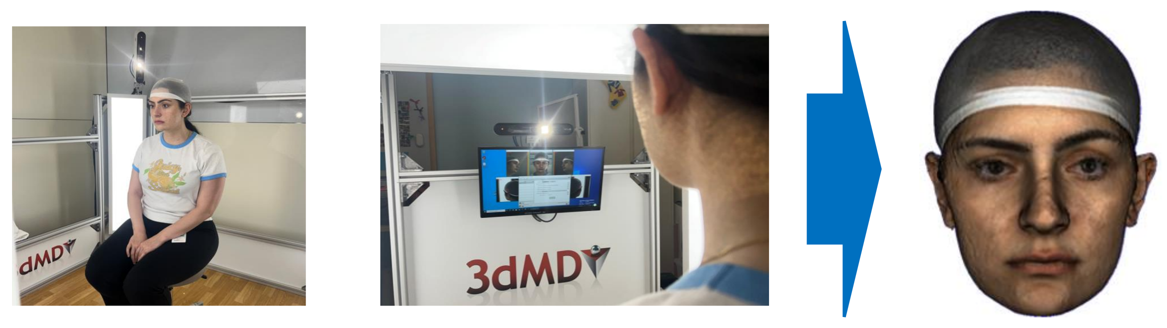
colorado school of public health

University of Colorado Anschutz Medical Campus

Children's Hospital Colorado

Background

3D photogrammetry



Cost efficient | Non-invasive | Accurate | Fast

State-of-the-art challenges

- ❑ High spatial resolutions with large number of vertices
- ❑ Variable number or vertices between patients
- ❑ Require sampling and/or interpolation
- ❑ Disregard local mesh structure information

Motivation

3D photogram registration is necessary to enable:

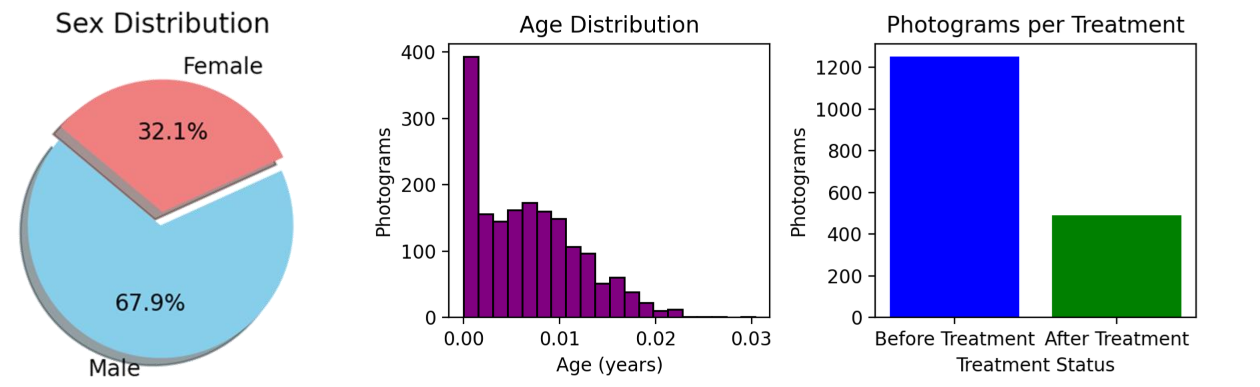
- ❑ Quantitative population studies (e.g., anatomical modeling)
- ❑ Longitudinal pediatric evaluation

Goal

To build an efficient and automated method for 3D photogram registration

Data

- ❑ ModelNet40 dataset [1]
  - 12,311 3D point clouds with 1024 points
  - 40 object categories
  - Uniformly sampled from original meshes
- ❑ 3D photograms dataset [2]
  - 1,744 3D photograms of children with craniosynostosis
  - Average age: 2.16 ± 2.32 years (range birth -18 years)
  - Average number of nodes: 23,686±915



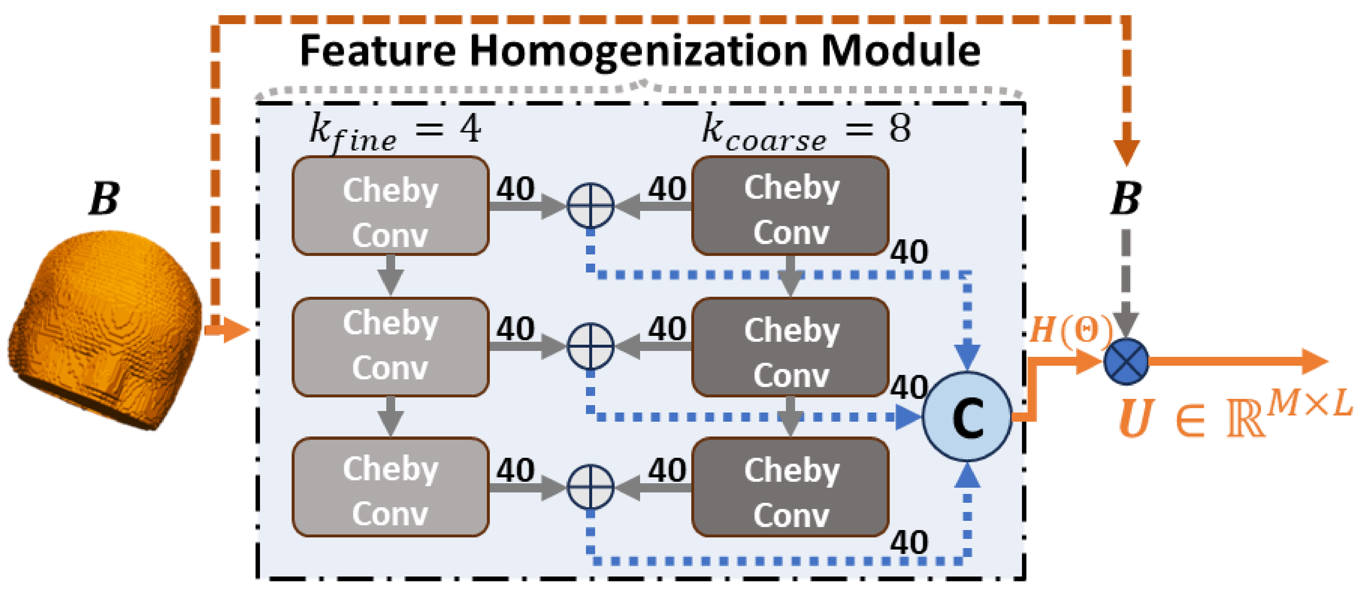
Methods

Feature homogenization network

Dimensionality reduction via feature homogenization:

$$\boldsymbol{U}(\boldsymbol{B},\boldsymbol{\Theta}) = \boldsymbol{B}^{T}\boldsymbol{H}(\boldsymbol{\Theta})$$

$\boldsymbol{B} \in \mathbb{R}^{NxM}$: variable dimensionality input data
$\boldsymbol{H}(\boldsymbol{\Theta}) \in \mathbb{R}^{NxL}$: learned feature transformation
$\boldsymbol{U}(\boldsymbol{B},\boldsymbol{\Theta}) \in \mathbb{R}^{MxL}$: uniform dimensionality output features
$\boldsymbol{\Theta}$: trainable coefficients



Offset cross-attention (OCA) network

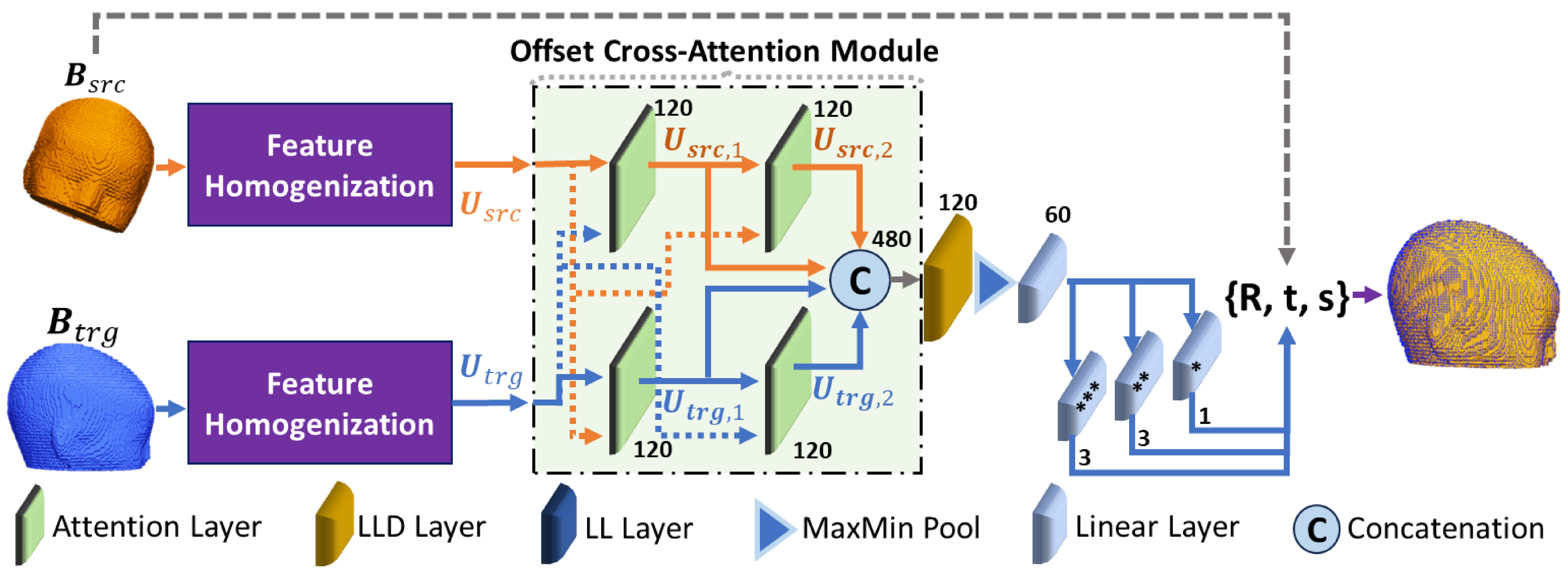
Efficient comparisons between latent feature representations:

- ❑ Offset Cross-Attention operation

$$OCA(\boldsymbol{Z}_1, \boldsymbol{Z}_2) = LBL(\boldsymbol{Z}_1 - softmax(\boldsymbol{QK}^{\top})\boldsymbol{V}) + \boldsymbol{Z}_1$$

- ❑ Four OCA layers

$$\boldsymbol{U}_{src,1} = OCA(\boldsymbol{U}_{src}, \boldsymbol{U}_{trg}) \qquad \boldsymbol{U}_{trg,1} = OCA(\boldsymbol{U}_{trg}, \boldsymbol{U}_{src})$$
$$\boldsymbol{U}_{src,2} = OCA(\boldsymbol{U}_{src,1}, \boldsymbol{U}_{src}) \qquad \boldsymbol{U}_{trg,1} = OCA(\boldsymbol{U}_{trg,1}, \boldsymbol{U}_{trg})$$



Optimization

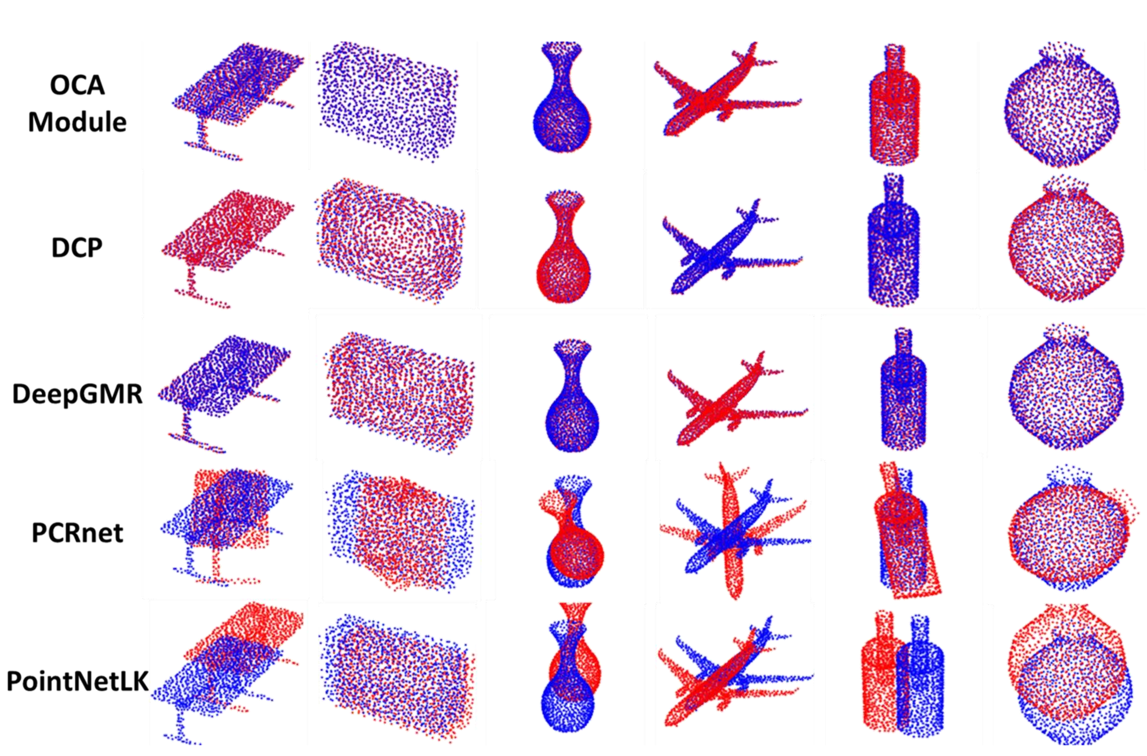
$$\mathcal{L} = \left\|R(\alpha,\beta,\gamma)^{T}R(\hat{\alpha},\hat{\beta},\hat{\gamma}) - I\right\|_F$$
$$+\|t-\hat{t}\|^2 + \|s-\hat{s}\|^2$$

$\{\alpha,\beta,\gamma\}$: rotation angles
t : translation vector
s: scaler factor

Experiments and results

1 OCA Evaluation with ModelNet40 dataset [1]

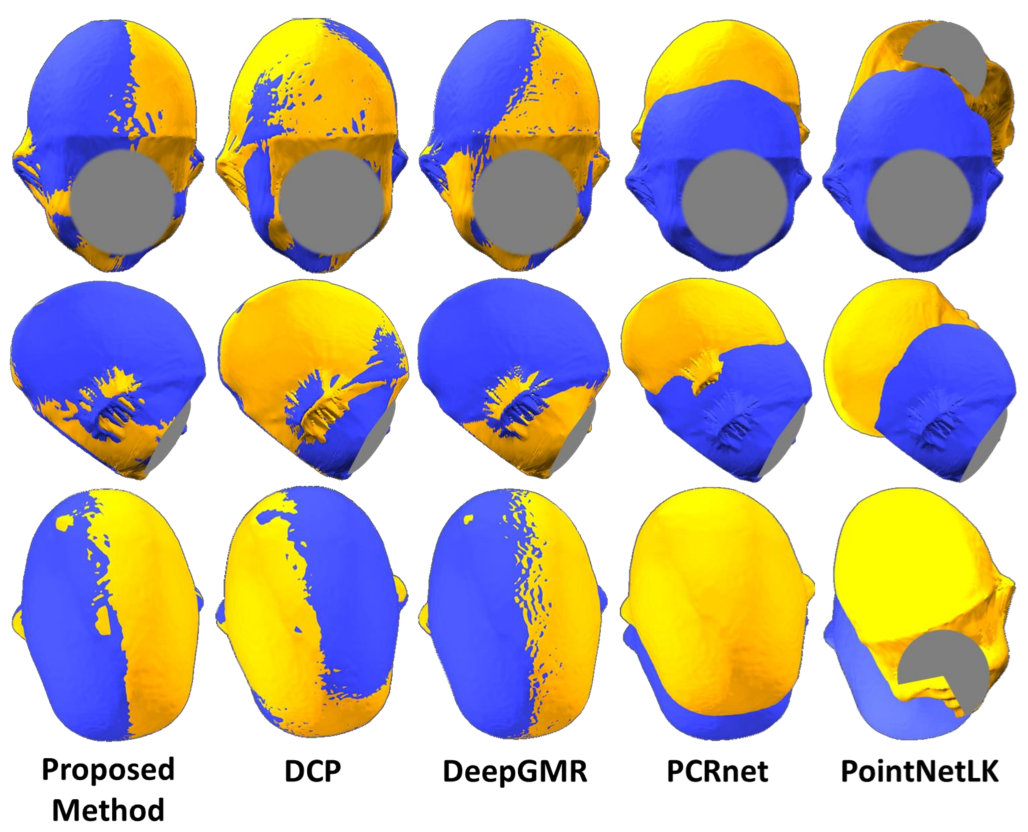
- ❑ Training size: 9,846
- ❑ Testing size: 2,468
- ❑ Rotation: ±45º
- ❑ Translation: ±0.5



Method	MSE		MAE		MSE	Chamfer
	{α, β, γ}	**t**	{α, β, γ}	**t**	(src,trg)	(src,trg)
DCP	5.43	**4E-4**	1.45	**0.01**	**4E-4**	0.46
PointNetLK	80.30	0.08	1.52	0.24	0.09	0.23
DeepGMR	32.67	1E-3	2.45	0.02	2E-3	**0.02**
PCRnet	683.18	0.08	22.63	0.24	0.05	0.13
OCA module	**3.42**	1E-3	**1.29**	0.02	1E-3	0.07

2 Intra-subject registration of 3D photograms

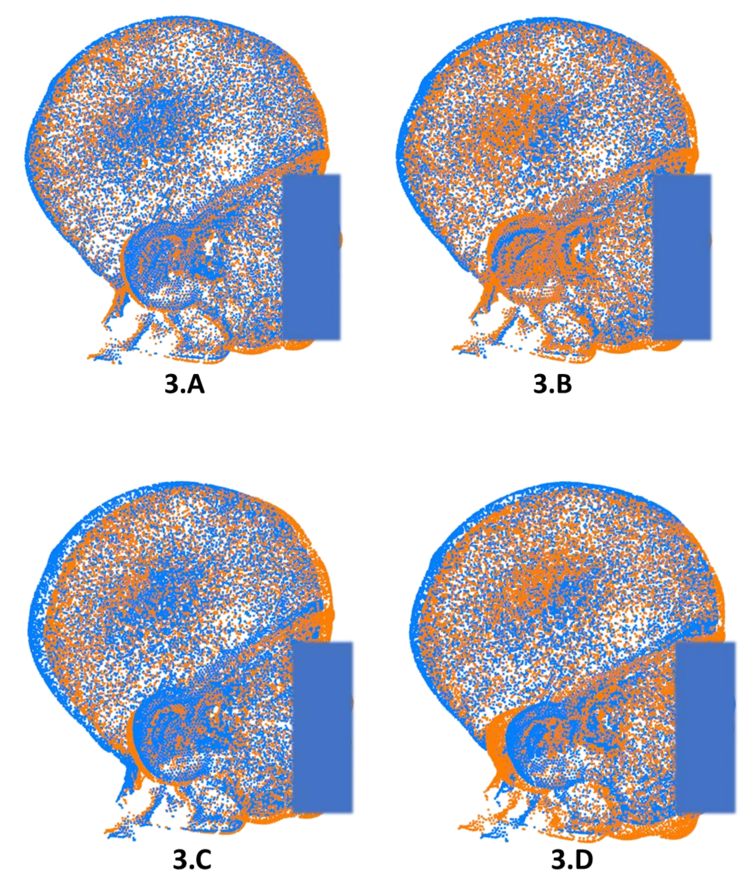
- ❑ Training size: 1,396
- ❑ Testing size: 174
- ❑ Rotation: ±45º
- ❑ Translation: ±0.5



Method	MSE		MAE		MSE	Chamfer
	{α, β, γ}	**t**	{α, β, γ}	**t**	(src,trg)	(src,trg)
DCP	60.09	0.01	5.01	0.06	0.03	4.81
PointNetLK	278.43	0.10	5.87	0.25	0.13	0.30
DeepGMR	31.53	0.01	**3.14**	0.06	0.02	0.189
PCRnet	494.64	0.01	15.12	0.25	0.08	0.19
Proposed method	**27.761**	**1E-3**	3.66	**0.03**	**0.01**	**0.02**

3 Intra/Inter-subject registration of 3D photograms with scaling

- ❑ Training size: 13,952
- ❑ Testing size: 1,744
- ❑ Rotation: ±45º
- ❑ Scaling: $[-0.5, 0.5]$
- ❑ Translation: ±50%



Metric		MSE			MAE			MSE	Chamfer
Experiment		{α, β, γ}	***t***	s	{α, β, γ}	***t***	s	(src,trg)	(src,trg)
3.A	Same subject	11.55	1E-3	-	2.03	0.02	-	0.01	0.02
3.B	Inter-subject	23.72	1E-3	-	3.51	0.02	-	0.01	0.02
3.C	Same subject	13.34	1E-3	4E-3	2.23	0.03	0.05	0.01	0.02
3.D	Inter-subject	30.42	1E-3	4E-3	3.94	0.03	0.05	0.01	0.02

Conclusions

- ❑ 3D surface registration method independent from the number of nodes and spatial resolution
- ❑ Feature homogenization module allows creating uniform representations independent from data dimensionality
- ❑ Offset cross-attention module improves registration performance by learning common spatial patterns between meshes
- ❑ State-of-the-art accuracy without data pre-processing or spatial sampling

References

[1] Wu, Z., et al. (2015). 3d shapenets: A deep representation for volumetric shapes. In Proceedings of the IEEE conference on computer vision and pattern recognition, 1912-1920.

[2] Elkhill, C., et al. (2023). Geometric learning and statistical modeling for surgical outcomes evaluation in craniosynostosis using 3D photogrammetry. Computer Methods and Programs in Biomedicine, 240, 107689.

[3] Defferrard, M., et al. (2016). Convolutional neural networks on graphs with fast localized spectral filtering. Advances in neural information processing systems, 29.

[4] Guo, M. H., et al. (2021). Pct: Point cloud transformer. Computational Visual Media, 7, 187-199.

NIH National Institute of Dental and Craniofacial Research R01DE032681

Ines Alejandro Cruz Guerrero, PhD
alejandro.cruzguerrero@cuanschutz.edu

Github code and more information: