Electromyography-Informed Facial Expression Reconstruction for Physiological-Based Synthesis and Analysis

Tim Büchner[1], Christoph Anders[2], Orlando Guntinas-Lichius[2], Joachim Denzler[1]

[1]Computer Vision Group [2]Jena University Hospital

Friedrich Schiller University Jena, Germany

FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

Computer Vision Group

UNIVERSITÄTSKLINIKUM jena

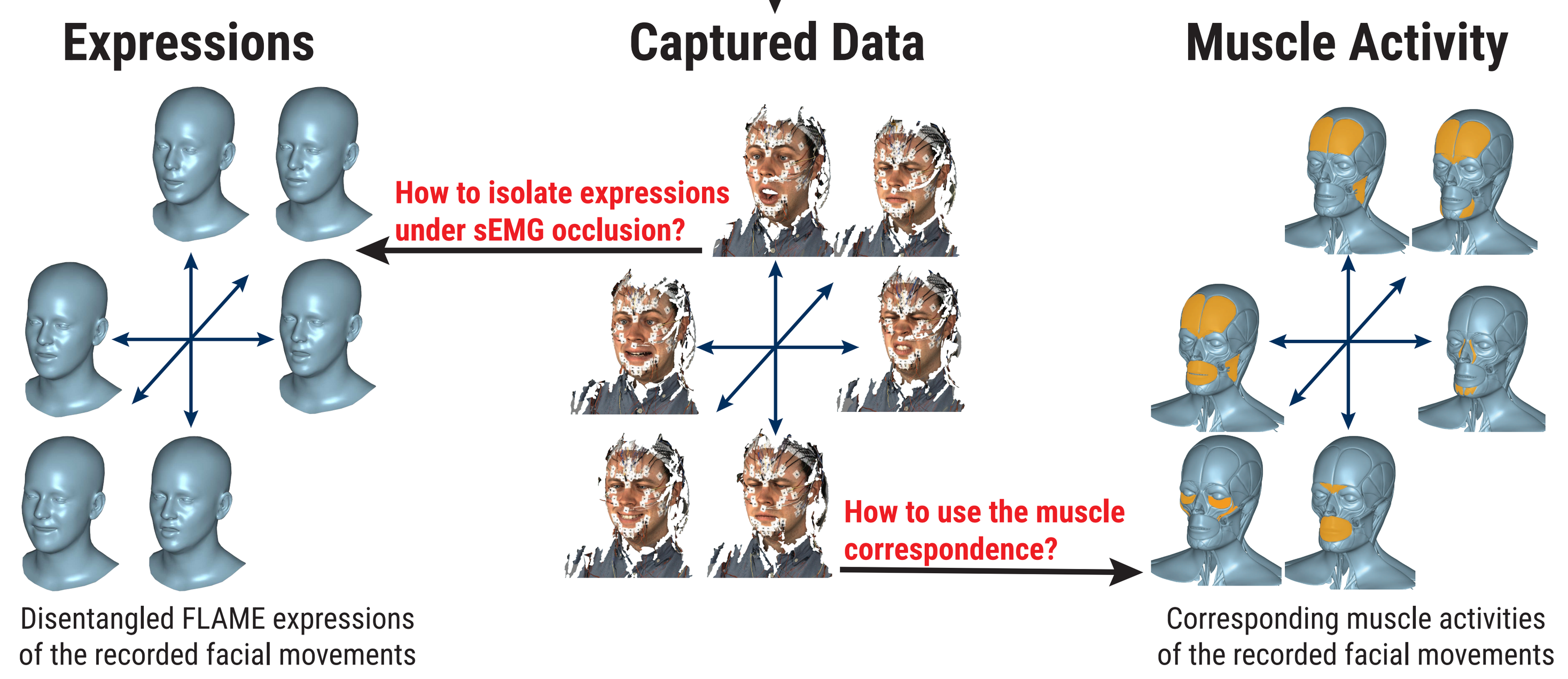
CVPR Nashville JUNE 11-15, 2025

Highlight Paper

Mimics and Muscles

We want to **learn the non-linear bidirectional correspondence** between facial movements and muscle activity for **more natural expression synthesis** and **camera-based muscle analysis**

Therefore, we record both **muscle activity and facial movement simultaneously!**



Expressions
Captured Data
Muscle Activity
How to isolate expressions under sEMG occlusion?
How to use the muscle correspondence?
Disentangled FLAME expressions of the recorded facial movements
Corresponding muscle activities of the recorded facial movements

A Novel Dataset

Synchronized Multi-Modal Face Data

- Multi-channel muscle activity with two sEMG schemes
- Frontal video for expression with Intel RealSense camera

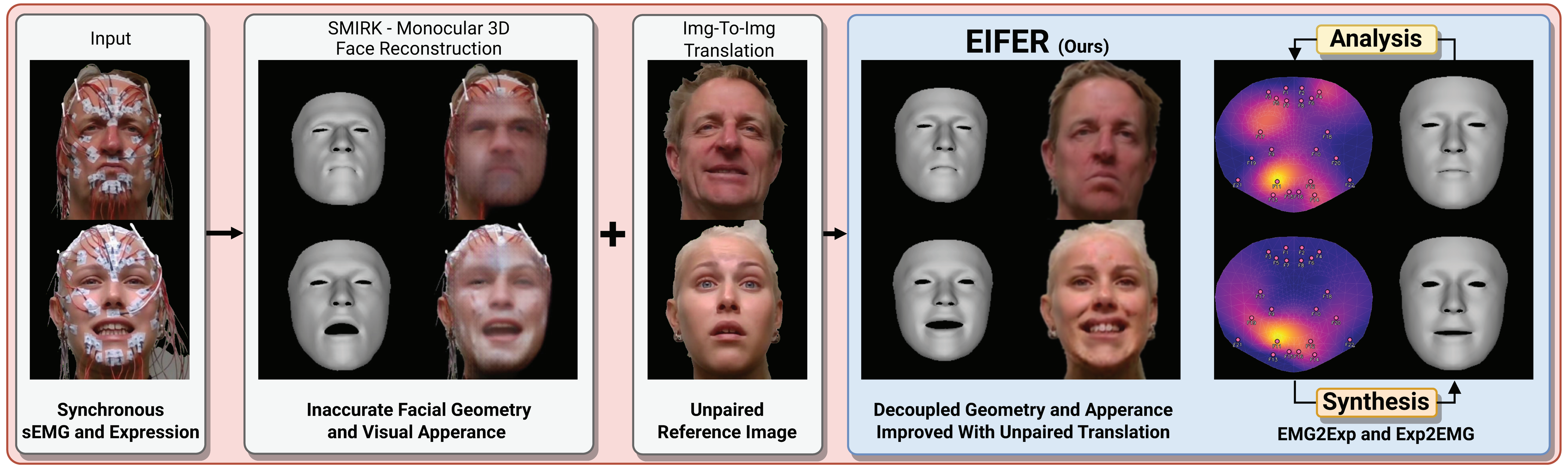
Rich in Variety

- 36 + 1 (20 + 1 public) healthy participants recorded
- 11 functional (individual) movements
- 6x4 emotional (complex) movements
- sEMG electrode free reference recordings
- Recording repeated after two weeks

A Unique Resource

- First dataset for data-driven muscle-expression-link
- **Many Tasks Await! Improve & Explore!**

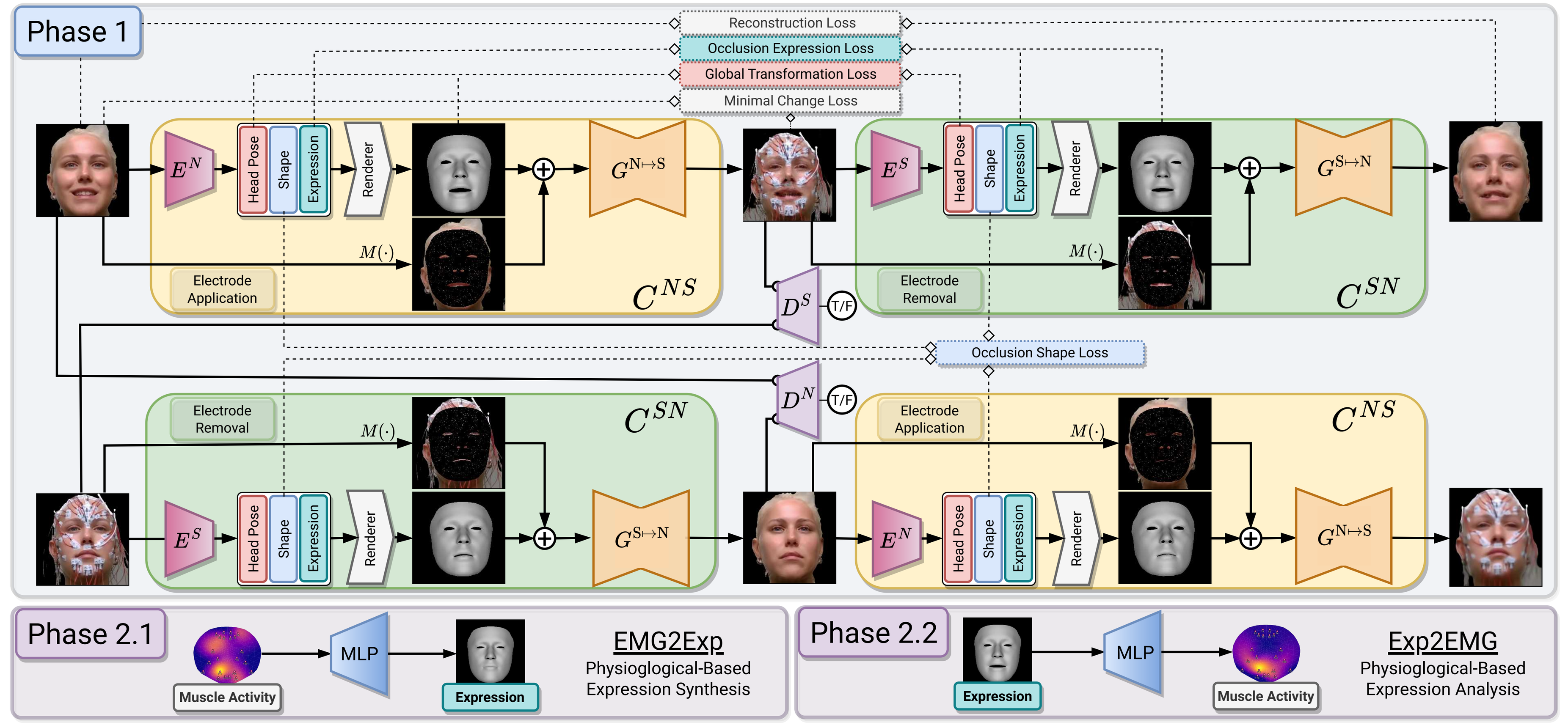
https://eifer-mam.github.io



Input
Synchronous sEMG and Expression
SMIRK - Monocular 3D Face Reconstruction
Inaccurate Facial Geometry and Visual Apperance
Img-To-Img Translation
Unpaired Reference Image
EIFER (Ours)
Decoupled Geometry and Apperance Improved With Unpaired Translation
Analysis
Synthesis
EMG2Exp and Exp2EMG

Muscles Fire, Faces Express – Two Sides of One Event! Yet muscle activity electrodes conceal the expressions we aim to capture.

How can we bridge the gap between **Mimics and Muscles** for 3D reconstruction, synthesis, and analysis?



Phase 1
Reconstruction Loss
Occlusion Expression Loss
Global Transformation Loss
Minimal Change Loss
Head Pose
Shape
Expression
Renderer
Electrode Application
Electrode Removal
Occlusion Shape Loss
Phase 2.1
Muscle Activity
MLP
Expression
EMG2Exp
Physioglogical-Based Expression Synthesis
Phase 2.2
Exp2EMG
Physioglogical-Based Expression Analysis

Phase 1: Reconstruction Under Occlusion (CycleGAN-Inspired)

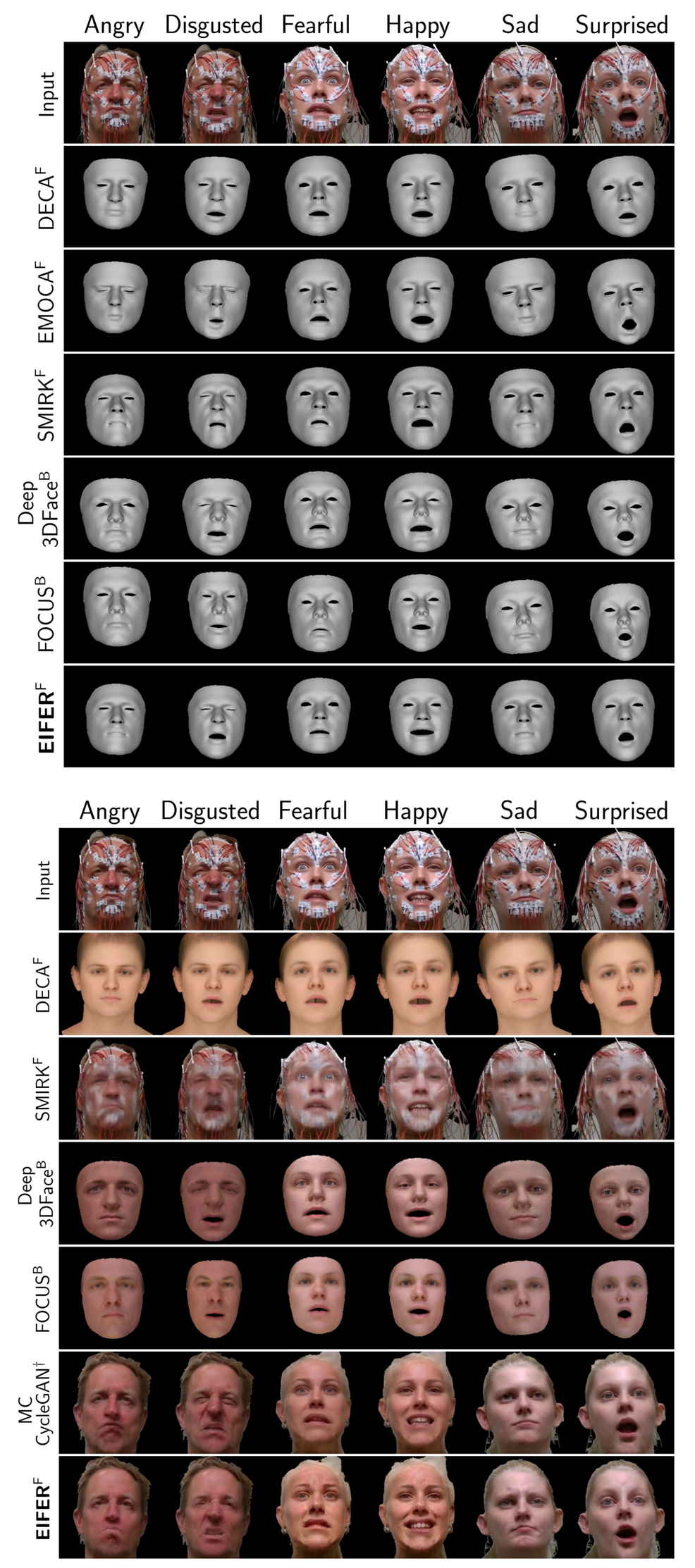
- **Framework:** Dual encoder-generator architecture (placing and removing of electrodes)
- **Input:** Occluded sEMG image + unpaired clean reference image
- **Encoder ($E_N + E_S$):** SMIRK[9]-based encoder architecture, estimate FLAME[6] parameters
- **Generator ($G_{N\to S} + G_{S\to N}$):** Combines rendered geometry with sparse color pixels (via masking) for photorealistic face generation as neural renderer
- **Multi-Stage Training (Simplify Convergence):**
  - **Generators Learn Electrode Locations**
    Focus on electrode application/removal (Encoders frozen, Generators get 50% pixels)
  - **Encoder Adapts to Occlusion**
    Adapt sensor encoder (E_S) on occluded faces (10% pixels), clean-side encoder (E_N) as teacher
  - **Final Decoupling of Geometry and Appearance**
    Force reliance on rendered geometry using 1% of appearance pixels
- **Training Principles:**
  - Cycle Consistency[11,14], Adversarial Training[15], and Minimal Change Loss[12]
  - Occlusion-Robust Regularizations (Occlusion Expression/Shape Loss, Global Transform Loss) ensuring consistent face geometry

Phase 2: Physiological Mappings

- **EMG2Exp:** Synthesizes 3DMM expressions from sEMG signals (camera-free animations)
- **Exp2EMG:** Predicts sEMG from 3DMM expressions (facial electromyography)
- Monocular 3D face reconstruction methods are now usable the restored face images

Reconstruction

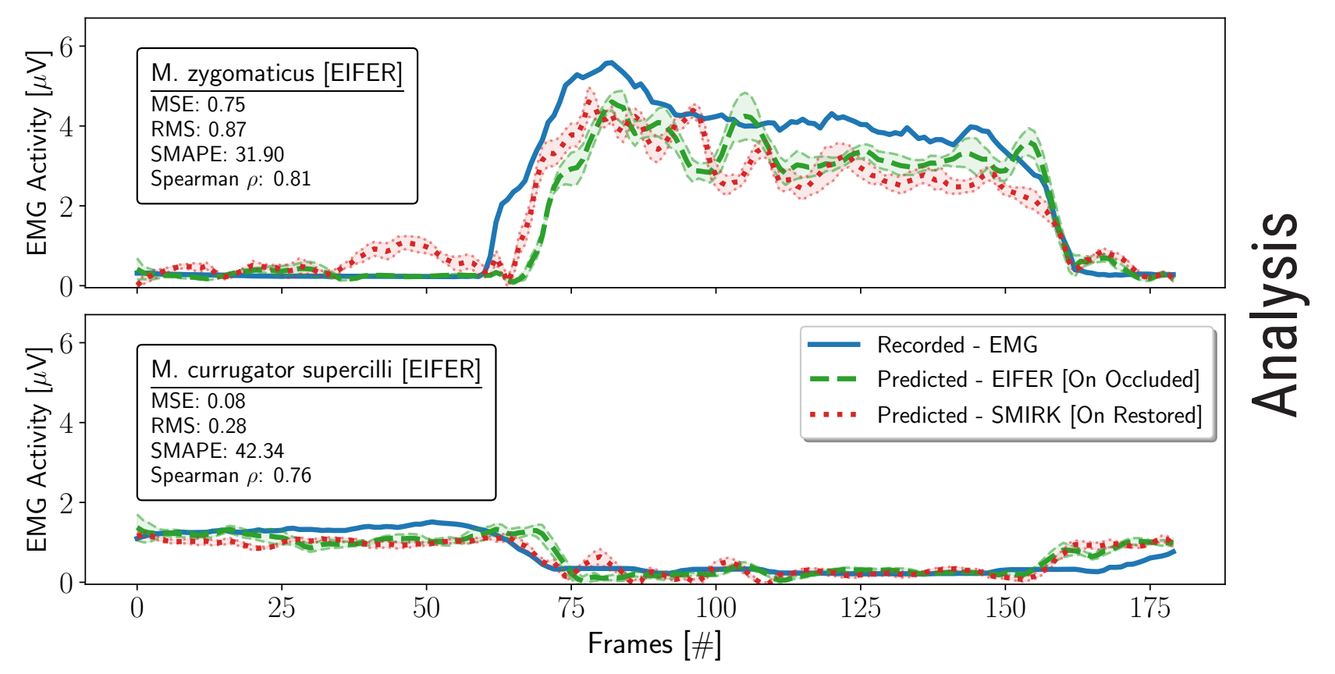
- Compare with existing reconstruction methods for FLAME[F] and BFM[B]: DECA[7], EMOCA[8], SMIRK[9], Deep3DFace[10], FOCUS[11], and MC-CycleGAN[12,13]
- Pixel-perfect ground truth does not exists for reconstruction evaluation
- Hence, we compute upper (**N**ormal-**N**ormal) and lower (**N**ormal-**S**ensor) performance bounds to evaluate the performance
- **Qualitative Evaluation**
  - Faithful geometry: consistent identity and expression intensity
    → Crucial for learning muscle activity and expression correspondence
  - Realistic appearance: neural generators create photorealistic faces
- **Quantitative Evaluation**
  - Matches and exceeds on photometric reconstruction metrics
  - Adversarial methods yield high FID due to powerful neural generators
  - Robust reconstruction possible with only 1% of input pixel
  - Performance similar to MC-CycleGAN but with face representation

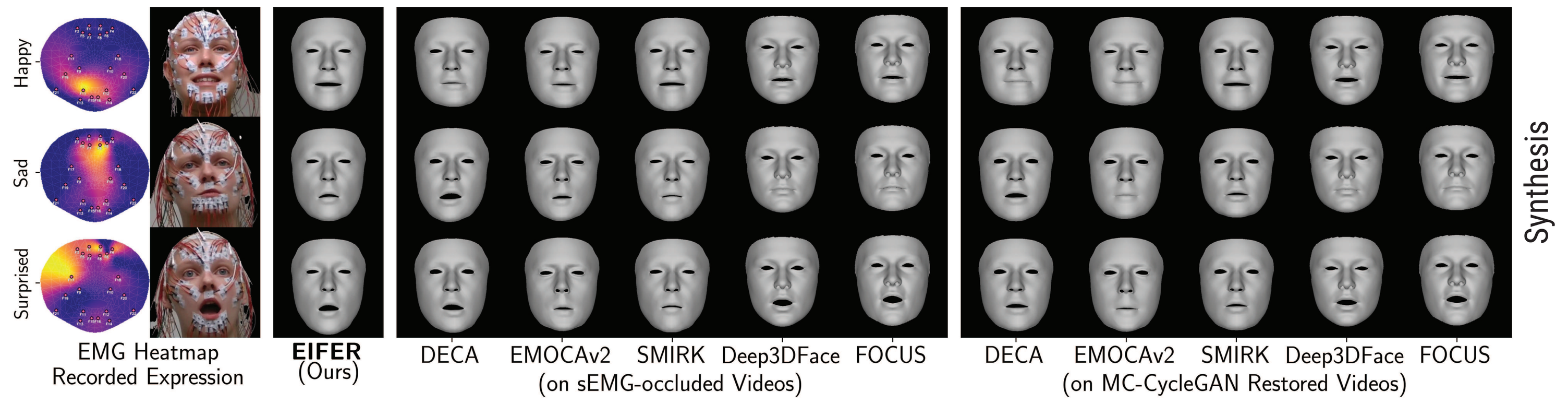


Method	SSIM (↑)	GMSD (↓)	PSNR (↑)	MDSI (↓)	FID (↓)
Baseline N-N	0.86±0.07	0.12±0.04	27.95±3.71	0.34±0.04	7.41± 3.72
Baseline N-S	0.39±0.05	0.33±0.01	13.69±1.27	0.62±0.02	285.42±38.18
DECA[F]	0.53±0.04	0.29±0.01	12.43±0.65	0.46±0.01	165.24±31.80
SMIRK[F]	0.47±0.06	0.31±0.02	14.45±1.41	0.58±0.02	275.80±46.38
Deep3DFace[B]	0.48±0.05	0.31±0.01	14.42±1.39	0.58±0.03	219.28±43.29
FOCUS[B]	0.46±0.05	0.32±0.02	13.95±1.35	0.58±0.03	227.71±50.21
MC-CycleGAN[†]	**0.66±0.08**	**0.24±0.03**	**19.38±2.39**	0.45±0.02	54.39±24.32
EIFER[F]	**0.66±0.09**	**0.24±0.03**	19.42±2.57	**0.44±0.03**	**52.56±27.75**

Synthesis and Analysis

- **Accurately Predicts Muscle Dynamics from Expressions**
  Camera-based electrode-free sEMG analysis by capturing dynamic, temporal muscle patterns and active/inactive muscle states
- **Faithfully Synthesizes Expressions from Muscle Signals**
  Generates realistic, complex facial movements from sEMG, for physiological animation & lifelike digital avatars
- **Unlocks Novel Physiological Insights & Applications**
  Integrates real muscle data into 3DMMs, advancing the study of expression-muscle links in health, disease, and digital humans





[1] P. F. Funk, R. Schneider, M. Schramm, G. F. Volk, C. Anders, and O. Guntinas-Lichius, "The complex facial muscle activation pattern in patients with postparalytic facial synkinesis: A prospective observational study using synchronous high-resolution surface electromyography on the synkinetic and the contralateral side," 2025

[2] P. Ekman and W. Friesen, "Facial Action Coding System: A Technique for the Measurement of Facial Movement," *Palo Alto: Consulting Psychologists Press*, 1978

[3] J. U. Blasberg, M. Gallistl, M. Degering, F. Baierlein, and V. Engert, "You look stressed: A pilot study on facial action unit activity in the context of psychosocial stress," *Comprehensive Psychoneuroendocrinology 2023*, vol. 15, p. 100187

[4] B. Egger et al., "3D Morphable Face Models-Past, Present, and Future," *ACM Transactions on Graphics 2020*, vol. 39, no. 5, p. 157:1-157:38

[5] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3D faces," *SIGGRAPH '99*, pp. 187–194.

[6] T. Li, T. Bolkart, M. J. Black, H. Li, and J. Romero, "Learning a model of facial shape and expression from 4D scans," *ACM Transactions on Graphics 2017*, vol. 36 pp. 1–17.

[7] Y. Feng, H. Feng, M. J. Black, and T. Bolkart, "Learning an animatable detailed 3D face model from in-the-wild images," *ACM Transactions on Graphics 2021*, vol. 40, pp. 1–13.

[8] R. Danecek, M. J. Black, and T. Bolkart, "EMOCA: Emotion Driven Monocular Face Capture and Animation," *CVPR 2022*

[9] G. Retsinas et al., "3D facial expressions through analysis-by-neural-synthesis." *CVPR 2024*, pp. 2490–2501

[10] Y. Deng, J. Yang, S. Xu, D. Chen, Y. Jia, and X. Tong, "Accurate 3d face reconstruction with weakly-supervised learning: From single image to image set," *CVPR 2019 Workshops*.

[11] C. Li, A. Morel-Forster, T. Vetter, B. Egger, and A. Kortylewski, "Robust model-based face reconstruction through weakly-supervised outlier segmentation," *CVPR 2023*, pp. 372–381.

[12] T. Büchner, S. Sickert, G. F. Volk, C. Anders, O. Guntinas-Lichius, and J. Denzler, "Let's get the FACS straight - reconstructing obstructed facial features," *VISAPP 2023*.

[13] T. Büchner, O. Guntinas-Lichius, and J. Denzler, "Improved Obstructed Facial Feature Reconstruction for Emotion Recognition with Minimal Change CycleGANs," *ACIVS 2023*, pp. 262–274.

[14] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," *CVPR 2017*, pp. 2223–2232.

[15] I. J. Goodfellow et al., "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.

Acknowledgments: Funded by **DFG project 427899908**
BRIDGING THE GAP: MIMICS AND MUSCLES (DE 735/15-1 and GU 463/12-1).

Contact: {tim.buechner, joachim.denzler}@uni-jena.de

Funded by DFG Deutsche Forschungsgemeinschaft German Research Foundation