FML: Face Model Learning from Videos

Ayush Tewari1  Florian Bernard1  Pablo Garrido2  Gaurav Bharaj2  Mohamed Elgharib1  Hans-Peter Seidel1  Patrick Pérez3  Michael Zollhöfer4  Christian Theobalt1

1Max-Planck-Institute for Informatics  2Technicolor  3Valeo.ai  4Stanford University

1. Goal

- 3D human face model learning from monocular videos
- Joint learning of 3D face reconstruction from monocular images
- Real-time inference at ~200fps

2. Parametric Face Representation

- The parameters define the complete appearance of faces in 3D
- The geometry and reflectance models are learned from videos

3. Background / Challenges

- Construction of current 3D face models:
  - high-quality scans are required
  - not generalizable to in-the-wild faces
- Building a model from 2D data:
  - can capture more general faces
  - ill-posed problem
- Training on videos using multi-frame constraints

4. Pipeline

- Learned Models
  - Monocular vs. Multi-frame Reconstruction
  - Comparison to Tran18

5. Loss Function

\[ E(P) = w_{\text{land}} E_{\text{land}}(P) + w_{\text{photo}} E_{\text{photo}}(P) + w_{\text{reg}} E_{\text{reg}}(P) \]

- Landmark Alignment
- Photometric Alignment
- Laplacian smoothness
- Local sparsity
- Stabilization

6. Results and Evaluations

- Geometric reconstruction error on the BU-3DFE dataset
- FML: Face Model Learning from Videos

This work was supported by the ERC Consolidator Grant 4DReply (770784), the Max Planck Center for Visual Computing and Communication (MPC-VCC), and by Technicolor.

This image contains a diagram illustrating the FML: Face Model Learning from Videos system, including the pipeline, loss function, and various evaluation results. The diagram highlights key components such as the learned models, pipeline structure, and loss function equations.