Poster Session: WED-AM-045

BIOMECHANICS-GUIDED FACIAL ACTION UNIT DETECTION THROUGH FORCE MODELING

JUNE 18-22, 2023

CVPR

VANCOUVER, CANADA

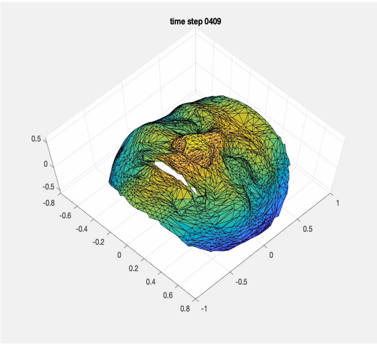
Zijun Cui, Chenyi Kuang, Tian Gao, Kartik Talamadupula, Qiang Ji

This work is supported by the Rensselaer-IBM AI Research Collaboration (http://airc.rpi.edu)

PREVIEW

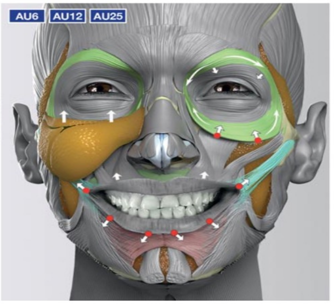
- Action Unit (AU) detection is to detect if the corresponding facial muscle is activated or not given 2D observations
- Biomechanics studies the dynamics of 3D geometric deformation caused by facial muscle contractions
- Leverage the facial biomechanics for AU detection given observable 2D videos

2D Video



3D Deformation

Biomechanics



[1]

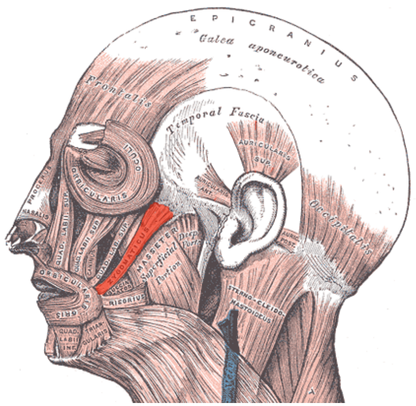
Muscle Activation Forces

For each frame,
AU is On or Off?

[1] https://anatomy4sculptors.com/blogs/articles/expressions?utm_source=pinterest&utm_medium=social

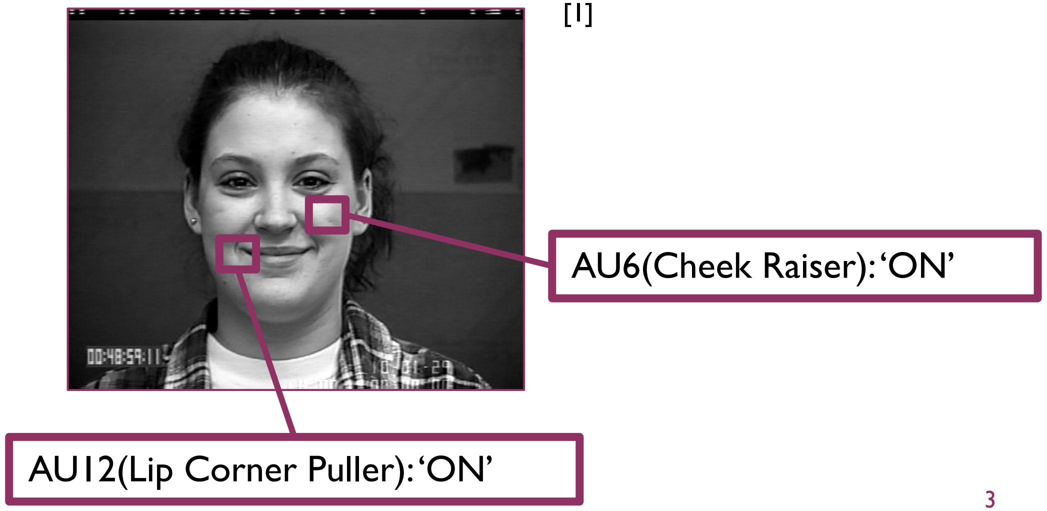
ACTION UNIT (AU) DETECTION

- Action Unit (AU) detection:
  - AU describes a local facial muscle activation. For example, *AU12 (lip corner puller)* is corresponding to the muscle *zygomatic major*
  - AU detection is to detect if the corresponding muscle is activated or not from 2D observations



Zygomaticus major

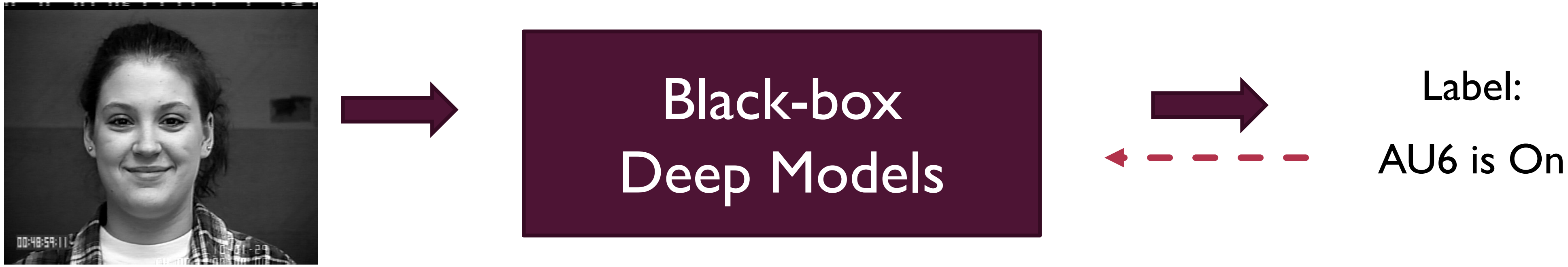
[1]



[1] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. 2010.

EXISTING AU MODELS

- Existing deep models for AU detection are mainly data-driven:

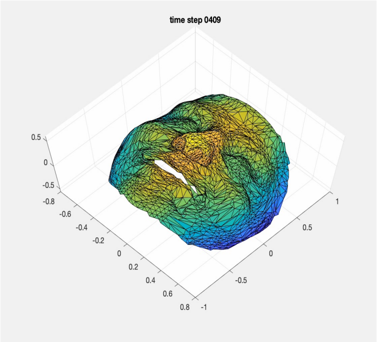


  - They reply on *sufficient annotated training data* to perform well
  - They may *not generalize* well to unseen subjects

BIOMECHANICS-GUIDED AU DETECTION

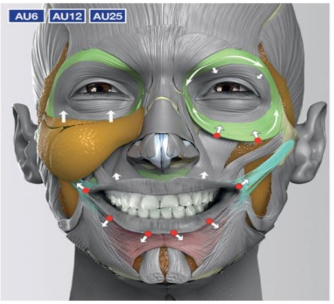
- Leverage the facial biomechanics for AU detection given observable 2D videos
- Biomechanics studies the dynamics of 3D geometric deformation caused by facial muscle contractions
- Biomechanics-guided AU detection through facial muscle activation force modeling
  - Step 1: Estimate physically plausible muscle activation forces using an Encoder-Decoder Network
  - Step 2: Perform AU detection with estimated muscle activation forces

2D Video (from BP4D)



3D Deformation

Biomechanics

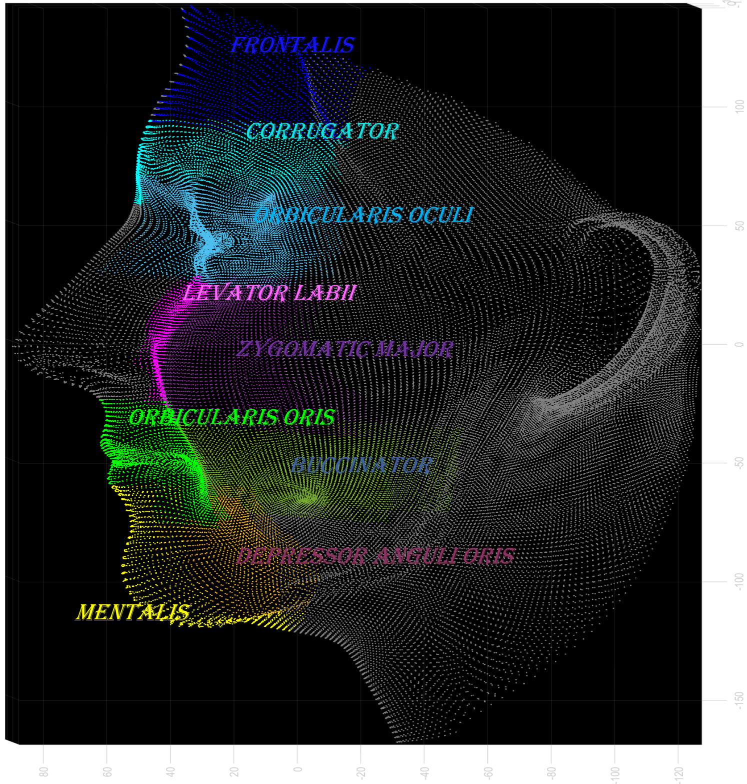


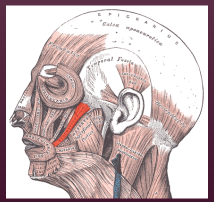
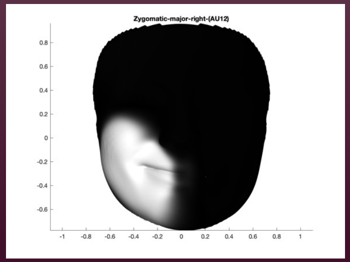
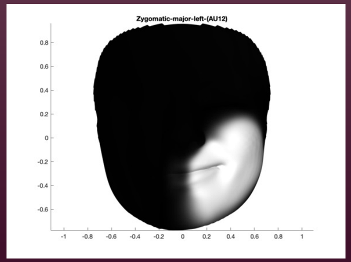
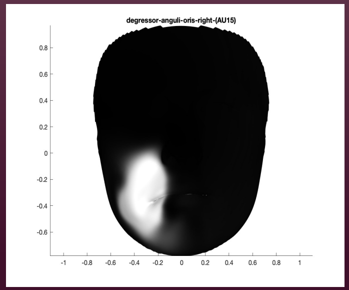
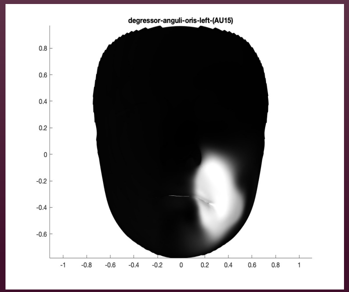
Muscle Activation Forces

For each frame, AU is On or Off?

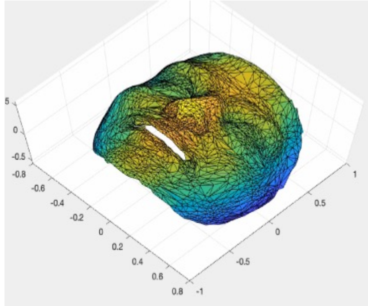
FACIAL MUSCLE ACTIVATION FORCE

- For each muscle, a subset of vertices are displaced as the muscle contracts
- We identify 9 major muscles and define their effective areas on the mesh



	Muscle	Effective Areas		Action Unit
Example 1	 *Zygomaticus major*			*AU12* *Lip Corner Puller*
Example 2	 *Depressor anguli oris*			*AU15* *Lip Corner Depressor*

BIOMECHANICS REPRESENTATION
-- ORDINARY DIFFERENTIAL EQUATION



3D mesh vertex $\mathrm{u}(t)$

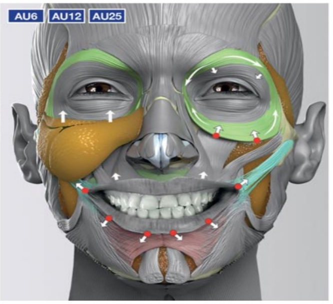
$\boldsymbol{u}_i(t) = \{\boldsymbol{u}_{i,x}(t), \boldsymbol{u}_{i,y}(t), \boldsymbol{u}_{i,z}(t)\}$

Ordinary Differential Equation

External force for each vertex

$\boldsymbol{F}_i^{ext} = \{F_{i,x}, F_{i,y}, F_{i,z}\}$

Projection



Muscle activation force in 3D

$\boldsymbol{F}_j^{mus}$

To reduce the dimension, a generalized coordinate defined by blendshape basis is considered

$$\boldsymbol{u} = c_1 B_1 + c_2 B_2 + \cdots + c_K B_K$$

- Blendshape basis $B = \{B_k\}$ *with* $\mathrm{B_k} \in \mathrm{R^{N\times3}}$ being k-th blendshape is known
- Linear coefficients $\boldsymbol{c} = \{c_k\}$ are corresponding to a mesh $\boldsymbol{u}$
- With the generalized coordinate, we only need to deal with $\mathbf{c} \in R^{K\times1}$. K is the dimension of the blendshape basis

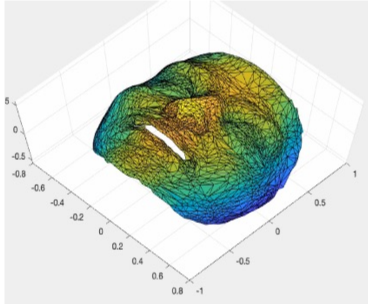
Applying Euler-Lagrangian equation, we derive the dynamic law in the generalized coordinate:

$$M^B \frac{d^2\boldsymbol{c}(t)}{dt^2} + \frac{dM^B}{dt}\dot{\boldsymbol{c}} - \frac{1}{2}\dot{\boldsymbol{c}}^T\left(\frac{\partial M^B}{\partial \boldsymbol{c}}\right)^T\dot{\boldsymbol{c}} = Q$$

$$\Rightarrow M^B \frac{d^2\boldsymbol{c}(t)}{dt^2} = Q$$

- $M^B \in R^{K\times K}$ is the generalized mass, and can be analytically computed with M and B
- $Q \in R^{K\times1}$ is the generalized force, and can be analytically computed with F^{ext} and B

FACIAL BIOMECHANICS INTEGRATION
-- CUSTOMIZE THE ARCHITECTURE OF ENCODER-DECODER NETWORK



3D mesh vertex $\mathrm{u}(t)$

Encoder

Hidden Features

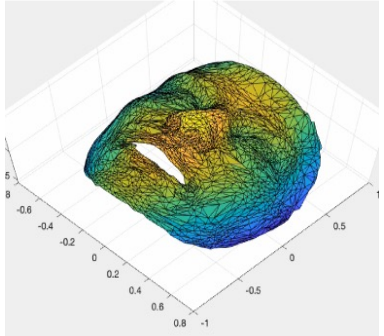
Physics parameters M^B and Q

Decoder

Decoder

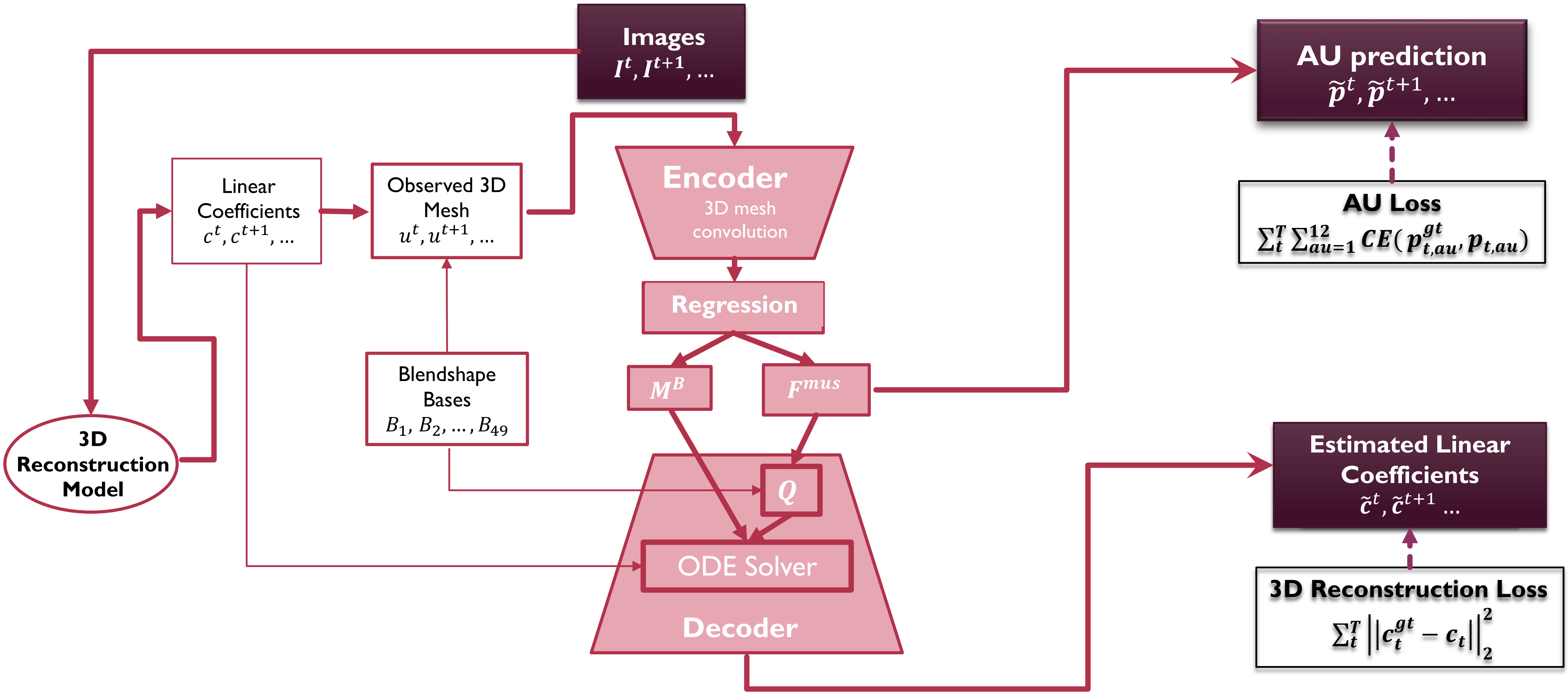
ODE solver:

$$\tilde{c} = \mathrm{ODEsolver}[\frac{d^2\boldsymbol{c}(t)}{dt^2} = (M^B)^{-1}\boldsymbol{Q}]$$



3D mesh vertex $\mathrm{u}(t+1)$

MODEL ARCHITECTURE



EXPERIMENTAL RESULTS
-- DATA EFFICIENCY EVALUATION

- We compare a baseline AU model trained with AU loss only and a baseline AU model trained with both AU loss and 3D reconstruction loss

Training Settings (BP4D)	AU Prediction (F1-score)	
	AU loss	AU loss + 3D reconstruction loss
100% Training Data	**.51**	.47
50% Training Data	.39	**.46**
20% Training Data	.35	**.44**
5% Training Data	.28	**.33**

- Using AU annotations only can't perform well on reduced training set
- Leveraging physics-based 3D reconstruction, AU detection performance significantly outperforms the one with AU loss only

THANK YOU!

Poster Session: WED-AM-045