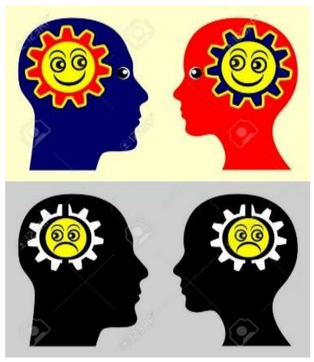
PERSON-SPECIFIC JOY EXPRESSION SYNTHESIS WITH GEOMETRIC METHOD

Sarra Zaied, Catherine Soladié, Pierre-Yves Richard
FAST Team, CentraleSupélec Rennes FRANCE, IETR -- Sarra.zaied@supelec.fr

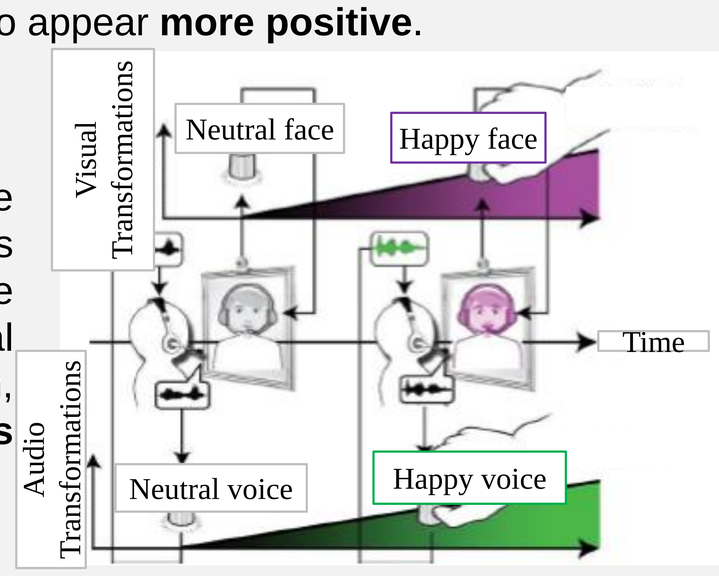


Introduction The theory of **peripheral emotional feedback** states that our emotional experiences are under the retroactive influence of our own expressions [1]. Actually, **facial expressions become a tool for psychological analysis**. The fact that putting on a smile or frown may have an **implicit, automatic effect in one's emotional experience** holds tremendous potential for clinical remediation in psychiatric disorders. We address this situation head-on, proposing a framework able to channel the psychological mechanism of facial feedback for clinical application to **post-traumatic stress disorders (PTSD).**

Objective

❖ Via our framework implemented in **a mirror-like device**, the observers can see themselves in a gradually **more positive way**; without their knowing, their reflected face is **algorithmically Transformed in real-time** to appear **more positive**.

❖ Using our system, we expect that observers believe the emotional tone of their transformed facial reflection as their own, and **align their feelings with the transformation.**

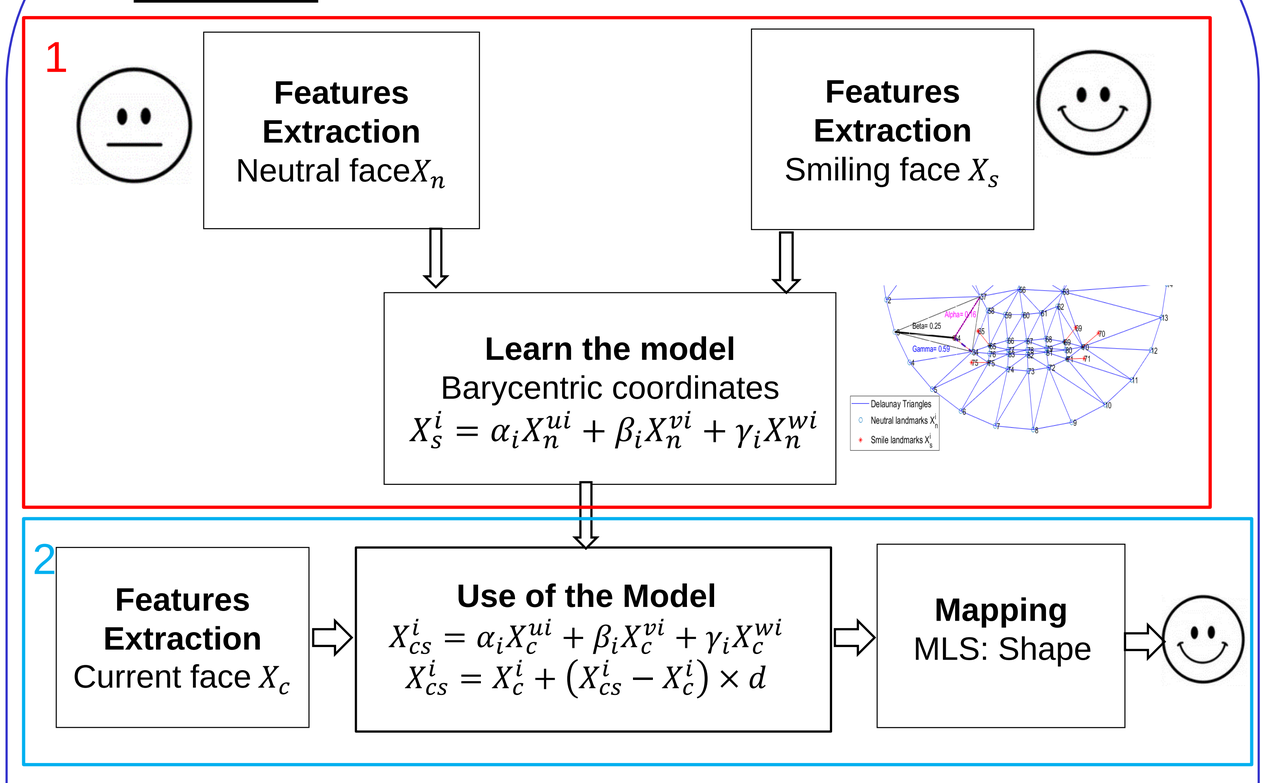


❖ The expressions are highly personal and each person smiles differently. That's why, to act positively on the emotional state of a subject while keeping the credibility, we generate **a person-specific joy expression** based on previous knowledge of her own way of smiling.

State of the art

- The geometric methods [2] ➔ relevant shape deformation but it lacks local details in the generated images. Use the same model for all the subjects.
- The generative models [3] ➔ add texture details (wrinkles and teeth) but the generated smiles are not those of the person.

❖ Method:



Overview of our framework

➢ Statistical study on CK, Oulu-CASIA and MMI databases ➔ the **smile is personal, straight and it occurs in a different way for each person** ➔ which justifies the necessity of a prior knowledge of the subject to learn a person specific model.

- ***Learning*** ➔ learns **a unique and specific parametric model** for each person thanks to a prior knowledge of the subject's way of smiling.
- ***Test*** ➔ synthesizes real-time smile expression with **different intensities** from a **single image** while preserving the user's identity.

Results

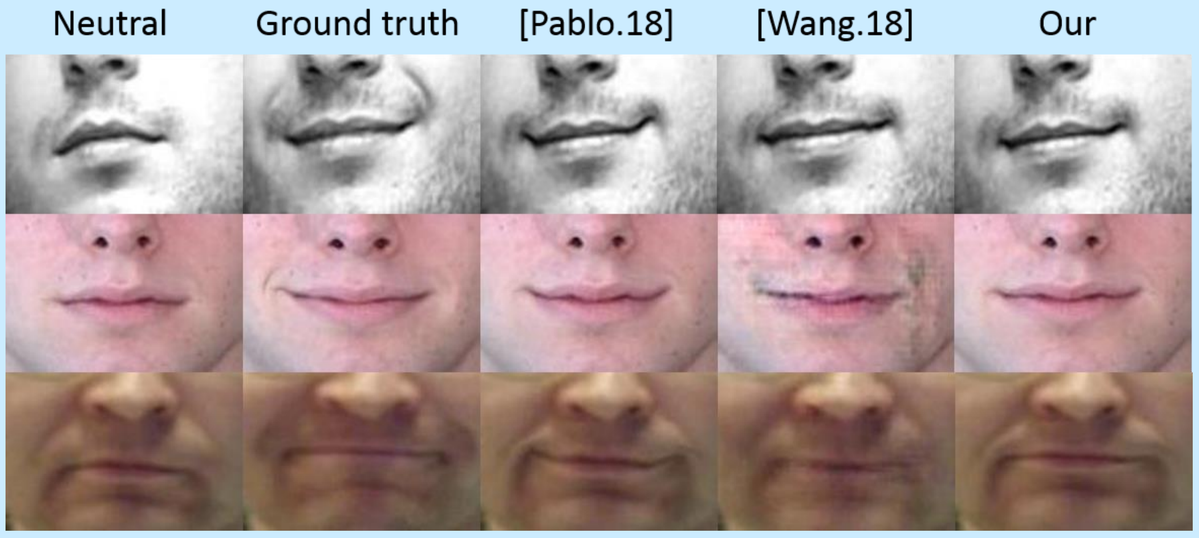
➢ Our model leads to preserve the morphology shape and the identity of the emotion by reproducing the specific way of smiling of each subject.

Photo-realistic images -- Personal smile -- Smiles with different intensities.

Evaluation

❖ Qualitative

➢ The visual fidelity shows that our method generates the closest smile shape to the ground truth for all the subjects.



Examples results with different methods [2], [3] and Our on the 3 databases CK, MMI and Oulu-CASIA respectively.

❖ Quantitative

➢ Statistical results on the generated smiles: $\theta = \left|\frac{a_{GT} - a}{1 + a_{GT}\, a}\right|$

	Database		
Method	CK	Oulu-CASIA	MMI
[Pablo.18] [2]	θ=12.10 σ=9.73	θ=17.81 σ=15.94	θ=12 σ=13.70
[Wang.18] [3]	θ=16.16 σ=12.24	θ=26 σ=18.21	θ=15.26 σ=12.66
Our Method	$\boldsymbol{\theta}$=6.65 $\boldsymbol{\sigma}$=7.57	$\boldsymbol{\theta}$=5.58 $\boldsymbol{\sigma}$=5.37	$\boldsymbol{\theta}$=4.38 $\boldsymbol{\sigma}$=5.27

Table 1. Results of the landmark of the left corner of the mouth. **Mean and standard deviation of angles** between GT trajectory and the synthesized trajectories are compared with the 3 methods on the 3 databases.

Conclusion

Our approach is a hybrid approach that :

- Learns a **person-specific model for each person** and transforms **in real-time** a captured face to appear more joyful.
- Generates for each subject a **photo-realistic joy expression** according to **her own expression**.
- Synthesizes smile expression with **different intensities** from a **single image**.

References:

[1]:Neal, David T., and Tanya L. Chartrand. "**Embodied emotion perception: amplifying and dampening facial feedback modulates emotion perception accuracy**." *Social Psychological and Personality Science* 2.6 (2011): 673-678.

[2]: Arias, P., Soladie, C., Bouafif, O., Robel, A., Seguier, R., & Aucouturier, J. J. (2017). Realistic transformation of facial and vocal smiles in real-time audiovisual streams. *IEEE Transactions on Affective Computing*.

[3]: Xueping Wang, Weixin Li, Guodong Mu, Di Huang, and Yunhong Wang, "Facial expression synthesis by u-net conditional GANs," in Proceedings ACM, 2018, pp. 283–290.