Vector-Valued Infinite Task Learning in Style Transfer

Zoltán Szabó
Department of Statistics, LSE
(joint work with Alex Lambert, Sanjeel Parekh, Florence d'Alché-Buc)

Advanced Statistical Methods for High Dimensional Data, CMStatistics
Dec. 19, 2021

Style transfer

- Goal: transfer an object according to a target style.

Numerous applications

- Computer vision [Ulyanov et al., 2016, Choi et al., 2018, Puy and Pérez, 2019, Yao et al., 2020], NLP [Fu et al., 2018], audio signal processing [Grinstein et al., 2018].
- Graphics: animating digital characters & avatars $\rightarrow$ body MOCAP [Aristidou et al., 2017, Aberman et al., 2020].
- Health & industry: digital twinning [Tao et al., 2019, Barricelli et al., 2019, Lim et al., 2020].

Style transfer

- Goal: transfer an object according to a target style.

Numerous applications

- Computer vision [Ulyanov et al., 2016, Choi et al., 2018, Puy and Pérez, 2019, Yao et al., 2020], NLP [Fu et al., 2018], audio signal processing [Grinstein et al., 2018].
- Graphics: animating digital characters & avatars → body MOCAP [Aristidou et al., 2017, Aberman et al., 2020].
- Health & industry: digital twinning [Tao et al., 2019, Barricelli et al., 2019, Lim et al., 2020].

Our aim

general principled style transfer: object → continuum of styles

Emotion transfer

- Given: set of emotions.
- Goal: transform object representations of
  - faces [Choi et al., 2018], hands [Irimia et al., 2019], body movement [Aristidou et al., 2017], . . .
  - repr: 2D images, 3D meshes, body skeletons, MOCAP sequences.

Emotion transfer

- Given: set of emotions.
- Goal: transform object representations of
  - faces [Choi et al., 2018], hands [Irimia et al., 2019], body movement [Aristidou et al., 2017], . . .
  - repr: 2D images, 3D meshes, body skeletons, MOCAP sequences.

Our example

- style := emotion,
- object representation := facial landmark locations.

Emotion transfer

- Given: set of emotions.
- Goal: transform object representations of
  - faces [Choi et al., 2018], hands [Irimia et al., 2019], body movement [Aristidou et al., 2017], . . .
  - repr: 2D images, 3D meshes, body skeletons, MOCAP sequences.

Our example

- style := emotion,
- object representation := facial landmark locations.

Why facial landmarks?

- Useful for facial identification [Mane and Shah, 2019], expression analysis [Devries et al., 2014], medical diagnosis [Balaei et al., 2017].

Problem formulation

- Object space: $\mathcal{X}$. Style space: Θ.
- Goal: (object, style) $\mapsto$ object, i.e. an

$$h : \mathcal{X} \times \Theta \mapsto \mathcal{X}, \text{ or } h : \mathcal{X} \mapsto (\Theta \mapsto \mathcal{X}).$$

Problem formulation

- Object space: $\mathcal{X}$. Style space: Θ.
- Goal: (object, style) $\mapsto$ object, i.e. an

$$h:\mathcal{X}\times\Theta\mapsto\mathcal{X},\text{ or } h:\mathcal{X}\mapsto(\Theta\mapsto\mathcal{X}).$$

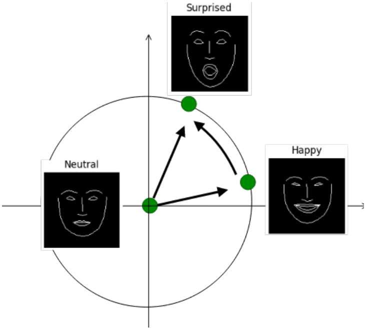
- In our case: landmarks $\mapsto$ $\underbrace{(\text{emotion}\mapsto\text{landmarks})}_{\text{function-valued regression}}$.

Problem formulation

- Object space: $\mathcal{X}$. Style space: Θ.
- Goal: (object, style) $\mapsto$ object, i.e. an

$$h : \mathcal{X} \times \Theta \mapsto \mathcal{X}, \text{ or } h : \mathcal{X} \mapsto (\Theta \mapsto \mathcal{X}).$$

- In our case: landmarks $\mapsto$ $\underbrace{(\text{emotion} \mapsto \text{landmarks})}_{\text{function-valued regression}}$.



Cost function

- Training samples:
  - For each object $i \in [n]$: $|S_i|$ style transition pairs $\left\{(\theta^{\text{in}}_{i,j}, \theta^{\text{out}}_{i,j})\right\}_{j \in S_i}$.
  - $x_{i,j} \in \mathcal{X}$: object with input style $\theta^{\text{in}}_{i,j}$,
  - $y_{i,j} \in \mathcal{X}$: object with output style $\theta^{\text{out}}_{i,j}$.

Cost function

- Training samples:
  - For each object $i \in [n]$: $|S_i|$ style transition pairs $\left\{(\theta_{i,j}^{\text{in}}, \theta_{i,j}^{\text{out}})\right\}_{j \in S_i}$.
  - $x_{i,j} \in \mathcal{X}$: object with input style $\theta_{i,j}^{\text{in}}$,
  - $y_{i,j} \in \mathcal{X}$: object with output style $\theta_{i,j}^{\text{out}}$.
- Cost (quality of the reconstruction) of h:

$$\mathcal{R}_{\mathcal{S}}(h) := \frac{1}{n} \sum_{i \in [n]} \frac{1}{|S_i|} \sum_{j \in S_i} \ell\Big(\overbrace{h(\underbrace{x_{i,j}}_{\substack{\text{input}\\ \text{object}}})(\underbrace{\theta_{i,j}^{\text{out}}}_{\substack{\text{output}\\ \text{style}}})}^{\text{predicted output object}}, \underbrace{y_{i,j}}_{\substack{\text{output}\\ \text{object}}} \Big).$$

Cost function

- Training samples:
  - For each object $i \in [n]$: $|S_i|$ style transition pairs $\left\{\left(\theta_{i,j}^{\text{in}}, \theta_{i,j}^{\text{out}}\right)\right\}_{j \in S_i}$.
  - $x_{i,j} \in \mathcal{X}$: object with input style $\theta_{i,j}^{\text{in}}$,
  - $y_{i,j} \in \mathcal{X}$: object with output style $\theta_{i,j}^{\text{out}}$.
- Cost (quality of the reconstruction) of h:

$$\mathcal{R}_{\mathcal{S}}(h) := \frac{1}{n} \sum_{i \in [n]} \frac{1}{|S_i|} \sum_{j \in S_i} \ell\Bigg(\overbrace{h(\underbrace{x_{i,j}}_{\text{input object}})(\underbrace{\theta_{i,j}^{\text{out}}}_{\text{output style}})}^{\text{predicted output object}}, \underbrace{y_{i,j}}_{\text{output object}} \Bigg).$$

- Quadratic loss: $\ell = \frac{1}{2} \|\cdot\|_{\mathcal{X}}^2$.

Cost function

- Training samples:
  - For each object $i \in [n]$: $|S_i|$ style transition pairs $\left\{\left(\theta_{i,j}^{\text{in}}, \theta_{i,j}^{\text{out}}\right)\right\}_{j \in S_i}$.
  - $x_{i,j} \in \mathcal{X}$: object with input style $\theta_{i,j}^{\text{in}}$,
  - $y_{i,j} \in \mathcal{X}$: object with output style $\theta_{i,j}^{\text{out}}$.
- Cost (quality of the reconstruction) of h:

$$\mathcal{R}_{\mathcal{S}}(h) := \frac{1}{n} \sum_{i \in [n]} \frac{1}{|S_i|} \sum_{j \in S_i} \ell\Big(\overbrace{h\big(\underbrace{x_{i,j}}_{\substack{\text{input}\\\text{object}}}\big)\big(\underbrace{\theta_{i,j}^{\text{out}}}_{\substack{\text{output}\\\text{style}}}\big)}^{\text{predicted output object}}, \underbrace{y_{i,j}}_{\substack{\text{output}\\\text{object}}} \Big).$$

- Quadratic loss: $\ell = \frac{1}{2} \|\cdot\|_{\mathcal{X}}^2$.
- Hypothesis class (vv-RKHS): $h : \underbrace{\mathcal{X} \mapsto \underbrace{(\Theta \mapsto \mathcal{X})}_{\in \mathcal{F} := \mathcal{H}_G}}_{\in \mathcal{H} := \mathcal{H}_K}$.

Trajectory interpretation: elements of $\mathcal{F}$

- Person $i \in [n]$: captured by a trajectory $z_i \in \mathcal{F}$.

  $$z_i: \text{ emotion } \theta \in \Theta \to \text{ landmarks } z_i(\theta) \in \mathcal{X}.$$

- z_i: observed at the style transition pairs $\left\{\left(\theta_{i,j}^{\text{in}}, \theta_{i,j}^{\text{out}}\right)\right\}_{j \in S_i}$.
- In other words,

$$x_{i,j} := z_i(\theta_{i,j}^{\text{in}}), \qquad y_{i,j} := z_i(\theta_{i,j}^{\text{out}}), \; i \in [n], j \in S_i.$$

Emotion representation: $\Theta \subset \mathbb{R}^p$

- Classical categorical description:

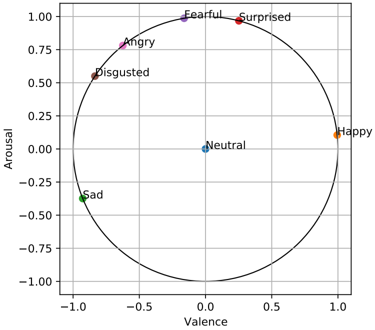
  'happy', 'sad', 'angry', 'surprised', 'disgusted', 'fearful'.

Emotion representation: $\Theta \subset \mathbb{R}^p$

- Classical categorical description:

  'happy', 'sad', 'angry', 'surprised', 'disgusted', 'fearful'.

- Valence-arousal model [Russell, 1980]: $\Theta \subset \mathbb{R}^2$,
  - valence: pleasure to displeasure,
  - arousal: high to low.

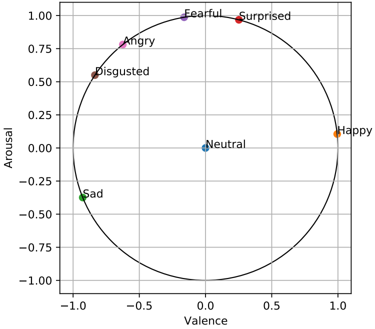


Emotion representation: $\Theta \subset \mathbb{R}^p$

- Classical categorical description:

  'happy', 'sad', 'angry', 'surprised', 'disgusted', 'fearful'.

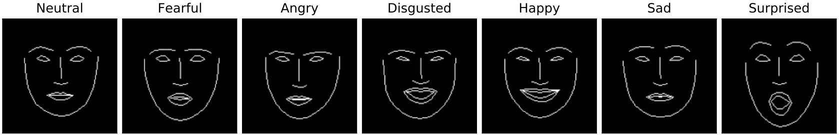
- Valence-arousal model [Russell, 1980]: $\Theta \subset \mathbb{R}^2$,
  - valence: pleasure to displeasure,
  - arousal: high to low.



- HigherD ($\Theta \subset \mathbb{R}^p$, $p \geq 2$) [Vemulapalli and Agarwala, 2019].

Object representation: $\mathcal{X} \subset \mathbb{R}^d$

- Face: landmarks points.
- Example: corners of the eyes, that of the mouth, ...



- Typically $M \approx 50 - 100 \Rightarrow \mathcal{X} \subset \mathbb{R}^{d:=2M}$.

Two problem families

- Single emotional input:
  - input emotion: identical & fixed for everyone (θ_0).
  - output emotion: same m number.

$\Rightarrow$ I-O emotion pairs: $\{(\theta_0, \theta_{i,j})\}_{j \in [m]}$, $|S_i| = m \; \forall i$.

Two problem families

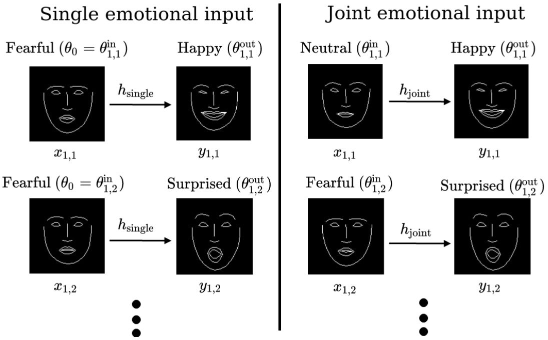
- Single emotional input:
  - input emotion: identical & fixed for everyone (θ_0).
  - output emotion: same m number.

  $\Rightarrow$ I-O emotion pairs: $\{(\theta_0, \theta_{i,j})\}_{j \in [m]}$, $|S_i| = m \; \forall i$.
- Joint emotional input:
  - m emotions for each person: $\{\theta_{i,a}\}_{a \in [m]}$, with all combinations,

  $\Rightarrow$ I-O emotion pairs: $\{(\theta_{i,a}, \theta_{i,b})\}_{a,b \in [m]}$, $|S_i| = m^2 \; \forall i$.

Two problem families

- Single emotional input:
  $\Rightarrow$ I-O emotion pairs: $\{(\theta_0, \theta_{i,j})\}_{j\in[m]}$, $|S_i| = m \ \forall i$.
- Joint emotional input:
  $\Rightarrow$ I-O emotion pairs: $\{(\theta_{i,a}, \theta_{i,b})\}_{a,b\in[m]}$, $|S_i| = m^2 \ \forall i$.



Towards the hypothesis class: $\mathcal{F}$

- Recall: $\underbrace{h : \mathcal{X} \mapsto \underbrace{(\boxed{\Theta \mapsto \mathcal{X}})}_{\in \mathcal{F} := \mathcal{H}_G}}_{\in \mathcal{H} := \mathcal{H}_K}$.

Towards the hypothesis class: $\mathcal{F}$

- Recall: $h : \underbrace{\mathcal{X} \mapsto \underbrace{(\boxed{\Theta \mapsto \mathcal{X}})}_{\in \mathcal{F} := \mathcal{H}_G}}_{\in \mathcal{H} := \mathcal{H}_K}$.
- $G : \Theta \times \Theta \to \underbrace{\mathcal{L}(\mathbb{R}^d)}_{\mathbb{R}^{d \times d}}$: matrix-valued kernel on Θ.
- Associated vv-RKHS: $\mathcal{H}_G = \overline{\operatorname{Span}} \left\{ G(\cdot, \theta)x : (\theta, x) \in \Theta \times \mathbb{R}^d \right\}$.

Towards the hypothesis class: $\mathcal{F}$

- Recall: $h : \underbrace{\mathcal{X} \mapsto \underbrace{(\boxed{\Theta \mapsto \mathcal{X}})}_{\in \mathcal{F} := \mathcal{H}_G}}_{\in \mathcal{H} := \mathcal{H}_K}$.
- $G : \Theta \times \Theta \to \underbrace{\mathcal{L}(\mathbb{R}^d)}_{\mathbb{R}^{d \times d}}$: matrix-valued kernel on Θ.
- Associated vv-RKHS: $\mathcal{H}_G = \overline{\operatorname{Span}}\left\{G(\cdot,\theta)x : (\theta, x) \in \Theta \times \mathbb{R}^d\right\}$.
- Simple & popular: separable kernel

$$G(\theta, \theta') = k_\Theta(\theta, \theta')\mathbf{A},$$

kernel $k_\Theta : \Theta \times \Theta \to \mathbb{R}$, $\mathbf{0} \preccurlyeq \mathbf{A} \in \mathbb{R}^{d \times d}$.

Towards the hypothesis class: $\mathcal{F}$

- Recall: $h : \mathcal{X} \mapsto \underbrace{(\underbrace{\boxed{\Theta \mapsto \mathcal{X}}}_{\in \mathcal{F} := \mathcal{H}_G})}_{\in \mathcal{H} := \mathcal{H}_K}$.
- $G : \Theta \times \Theta \to \underbrace{\mathcal{L}(\mathbb{R}^d)}_{\mathbb{R}^{d \times d}}$: matrix-valued kernel on Θ.
- Associated vv-RKHS: $\mathcal{H}_G = \overline{\operatorname{Span}} \left\{ G(\cdot, \theta)x : (\theta, x) \in \Theta \times \mathbb{R}^d \right\}$.
- Simple & popular: separable kernel

$$G(\theta, \theta') = k_\Theta(\theta, \theta')\mathbf{A},$$

kernel $k_\Theta : \Theta \times \Theta \to \mathbb{R}$, $\mathbf{0} \preccurlyeq \mathbf{A} \in \mathbb{R}^{d \times d}$.
- Intuition:
  - smoothness: Gaussian kernel $k_\Theta(\theta, \theta') = \exp\left(-\gamma \|\theta - \theta'\|_2^2\right)$, $\gamma > 0$,
  - dependency among output coordinates: $\mathbf{A}$.

Hypothesis class: $\mathcal{H}$

- Recall: $h : \underbrace{\mathcal{X} \mapsto \underbrace{(\Theta \mapsto \mathcal{X})}_{\in \mathcal{F} := \mathcal{H}_G}}_{\in \boxed{\mathcal{H} := \mathcal{H}_K}}$.
- $K : \mathcal{X} \times \mathcal{X} \to \mathcal{L}(\mathcal{H}_G)$: operator-valued kernel.
- Associated vv-RKHS:

$$\mathcal{H}_K = \overline{\operatorname{Span}} \left\{ K(\cdot, x) f : (x, f) \in \mathcal{X} \times \mathcal{H}_G \right\}.$$

Hypothesis class: $\mathcal{H}$

- Recall: $h : \mathcal{X} \mapsto \underbrace{\underbrace{(\Theta \mapsto \mathcal{X})}_{\in \mathcal{F} := \mathcal{H}_G}}_{\in \boxed{\mathcal{H} := \mathcal{H}_K}}$.
- $K : \mathcal{X} \times \mathcal{X} \to \mathcal{L}(\mathcal{H}_G)$: operator-valued kernel.
- Associated vv-RKHS:

$$\mathcal{H}_K = \overline{\mathrm{Span}}\left\{K(\cdot, x)f : (x, f) \in \mathcal{X} \times \mathcal{H}_G\right\}.$$

- Separable kernel:

$$K(x, x') = k_{\mathcal{X}}(x, x')\mathrm{Id}_{\mathcal{H}_G},$$

with kernel $k_{\mathcal{X}} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$.

Hypothesis class: $\mathcal{H}$

- Recall: $h : \mathcal{X} \mapsto \underbrace{\underbrace{(\Theta \mapsto \mathcal{X})}_{\in \mathcal{F} := \mathcal{H}_G}}_{\in \boxed{\mathcal{H} := \mathcal{H}_K}}$.
- $K : \mathcal{X} \times \mathcal{X} \to \mathcal{L}(\mathcal{H}_G)$: operator-valued kernel.
- Associated vv-RKHS:

$$\mathcal{H}_K = \overline{\mathrm{Span}} \left\{ K(\cdot, x) f : (x, f) \in \mathcal{X} \times \mathcal{H}_G \right\}.$$

- Separable kernel:

$$K(x, x') = k_{\mathcal{X}}(x, x') \mathrm{Id}_{\mathcal{H}_G},$$

with kernel $k_{\mathcal{X}} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$.
- Intuition: smoothness $\Leftarrow k_{\mathcal{X}}$: Gaussian on $\mathcal{X}$.

Optimization

- vv-RKHS $\Rightarrow$ natural regularization.
- Task (vITL):

$$\min_{h \in \mathcal{H}_K} \mathcal{R}_\lambda(h) := \mathcal{R}_{\mathcal{S}}(h) + \frac{\lambda}{2} \|h\|^2_{\mathcal{H}_K}, \ \lambda > 0.$$

Optimization

- vv-RKHS $\Rightarrow$ natural regularization.
- Task (vITL):

$$\min_{h \in \mathcal{H}_K} \mathcal{R}_\lambda(h) := \mathcal{R}_\mathcal{S}(h) + \frac{\lambda}{2} \|h\|^2_{\mathcal{H}_K}, \ \lambda > 0.$$

- Representer lemma:

$$\hat{h}(x)(\theta) = \sum_{i=1}^{t} \sum_{j=1}^{m} k_\mathcal{X}(x, x_i) k_\Theta(\theta, \theta_{i,j}) \mathbf{A} \hat{\mathbf{c}}_{i,j}, \ \{\hat{\mathbf{c}}_{i,j}\}_{i \in [t], j \in [m]} \subset \mathbb{R}^d.$$

Optimization

- vv-RKHS $\Rightarrow$ natural regularization.
- Task (vITL):

$$\min_{h \in \mathcal{H}_K} \mathcal{R}_\lambda(h) := \mathcal{R}_\mathcal{S}(h) + \frac{\lambda}{2} \|h\|^2_{\mathcal{H}_K}, \ \lambda > 0.$$

- Representer lemma:

$$\hat{h}(x)(\theta) = \sum_{i=1}^{t} \sum_{j=1}^{m} k_\mathcal{X}(x, x_i) k_\Theta(\theta, \theta_{i,j}) \mathbf{A} \hat{\mathbf{c}}_{i,j}, \ \{\hat{\mathbf{c}}_{i,j}\}_{i \in [t], j \in [m]} \subset \mathbb{R}^d.$$

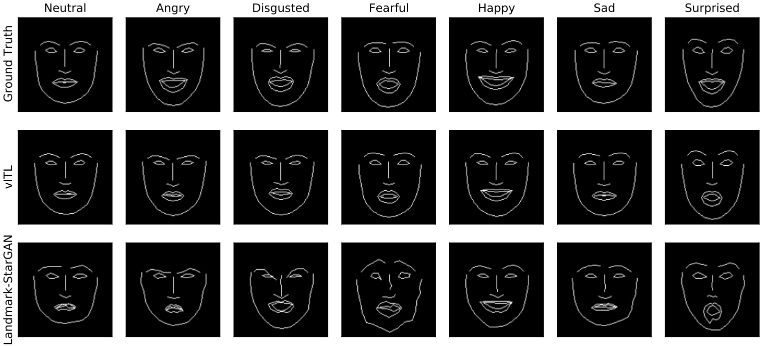
- With quadratic loss (ℓ): linear equation to $\hat{\mathbf{c}}_{i,j}$-s.

Quantitative illustration

Methods	MSE Error ↓		Emotion Classification Acc. ↑	
	KDEF frontal	RaFD frontal	KDEF frontal	RaFD frontal
vITL: θ_0 = neutral	0.010 ± 0.001	0.009 ± 0.004	76.12 ± 4.57	79.76 ± 7.88
vITL: θ_0 = fearful	0.010 ± 0.001	0.010 ± 0.005	76.22 ± 4.91	78.81 ± 8.36
vITL: θ_0 = angry	0.012 ± 0.002	0.010 ± 0.005	74.49 ± 2.31	78.10 ± 7.51
vITL: θ_0 = disgusted	0.012 ± 0.001	0.010 ± 0.004	74.18 ± 4.22	78.33 ± 4.12
vITL: θ_0 = happy	0.011 ± 0.001	0.010 ± 0.004	73.57 ± 2.74	80.48 ± 5.70
vITL: θ_0 = sad	0.011 ± 0.001	0.009 ± 0.004	75.82 ± 4.11	77.62 ± 5.17
vITL: θ_0 = surprised	0.010 ± 0.001	0.011 ± 0.006	74.69 ± 2.25	80.71 ± 5.99
vITL: Joint	$\mathbf{0.011} \pm 0.001$	$\mathbf{0.007} \pm 0.001$	$\mathbf{74.81} \pm 3.10$	$\mathbf{77.11} \pm 3.97$
Landmark-StarGAN	0.029 ± 0.003	0.024 ± 0.007	70.69 ± 8.46	65.88 ± 8.92

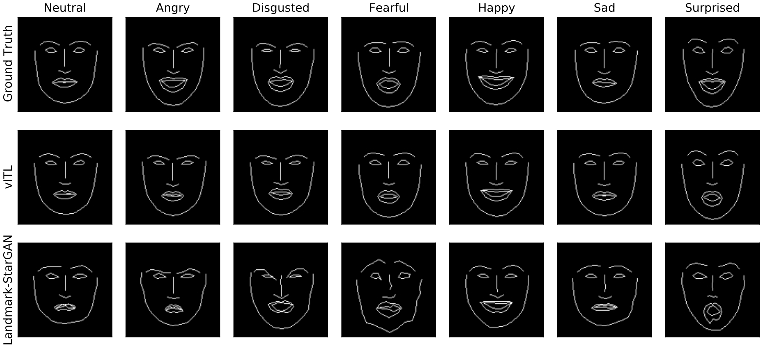
Qualitative illustration

Vs Landmark-StarGAN:

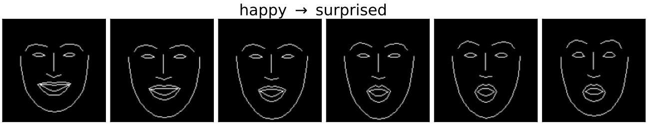


Qualitative illustration

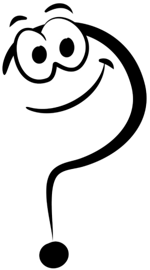
Vs Landmark-StarGAN:



Continuous traversal by vITL ($\hat{h}$):



Thank you for the attention!



- **Acks**: A.L. and S.P. were funded by the research chair Data Science & Artificial Intelligence for Digitalized Industry and Services at Télécom Paris. Z.Sz. benefited from the support of the Europlace Institute of Finance and that of the Chair Stress Test, RISK Management and Financial Steering, led by the French École Polytechnique and its Foundation and sponsored by BNP Paribas.
- Emotion transfer by vITL (CtrlGen @ NeurIPS-2021).
- ITL @ AISTATS-2019.
- Code: `https://github.com/allambert/torch_itl`.

Kfir Aberman, Yijia Weng, Dani Lischinski, Daniel Cohen-Or, and Baoquan Chen. Unpaired motion style transfer from video to animation. *ACM Transactions on Graphics (TOG)*, 39(4):64, 2020.

Andreas Aristidou, Qiong Zeng, Efstathios Stavrakis, KangKang Yin, Daniel Cohen-Or, Yiorgos Chrysanthou, and Baoquan Chen. Emotion control of unstructured dance movements. In *ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pages 1–10, 2017.

Asghar Tabatabaei Balaei, Kate Sutherland, Peter A. Cistulli, and Philip de Chazal. Automatic detection of obstructive sleep apnea using facial images. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 215–218, 2017.

Barbara Rita Barricelli, Elena Casiraghi, and Daniela Fogli. A survey on digital twin: Definitions, characteristics, applications, and design implications. *IEEE Access*, 7:167653–167671, 2019.

Yunjey Choi, Minje Choi, Munyoung Kim, Jung-Woo Ha, Sunghun Kim, and Jaegul Choo. StarGAN: Unified generative adversarial networks for multi-domain image-to-image translation. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8789–8797, 2018.

Terrance Devries, Kumar Biswaranjan, and Graham W. Taylor. Multi-task learning of facial landmarks and expression. In *Canadian Conference on Computer and Robot Vision*, pages 98–103, 2014.

Zhenxin Fu, Xiaoye Tan, Nanyun Peng, Dongyan Zhao, and Rui Yan. Style transfer in text: Exploration and evaluation. In *Conference on Artificial Intelligence (AAAI)*, pages 663–670, 2018.

Eric Grinstein, Ngoc QK Duong, Alexey Ozerov, and Patrick Pérez. Audio style transfer. In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 586–590, 2018.

Ana-Sabina Irimia, Jacky CP Chan, Kamlesh Mistry, Wei Wei, and Edmond SL Ho. Emotion transfer for hand animation. In *Motion, Interaction and Games*, pages 1–2. 2019.

Kendrik Yan Hong Lim, Pai Zheng, and Chun-Hsien Che. A state-of-the-art survey of digital twin: techniques, engineering product lifecycle management and business innovation perspectives. *Journal of Intelligent Manufacturing*, 31: 1313–1337, 2020.

Shraddha Mane and Gauri Shah. Facial recognition, expression recognition, and gender identification. In *Data Management, Analytics and Innovation*, pages 275–290. Springer, 2019.

Gilles Puy and Patrick Pérez. A flexible convolutional solver for fast style transfers. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8963–8972, 2019.

James A Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6):1161–1178, 1980.

Fei Tao, He Zhang, Ang Liu, and A. Y. C. Nee. Digital twin in industry: State-of-the-art. *IEEE Transactions on Industrial Informatics*, 15(4):2405 – 2415, 2019.

Dmitry Ulyanov, Vadim Lebedev, Andrea Vedaldi, and Victor Lempitsky. Texture networks: Feed-forward synthesis of textures and stylized images. In *International Conference on Machine Learning (ICML)*, pages 1349–1357, 2016.

Raviteja Vemulapalli and Aseem Agarwala. A compact embedding for facial expression similarity. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5683–5692, 2019.

Xu Yao, Gilles Puy, Alasdair Newson, Yann Gousseau, and Pierre Hellier. High resolution face age editing. In *International Conference on Pattern Recognition (ICPR)*, 2020.