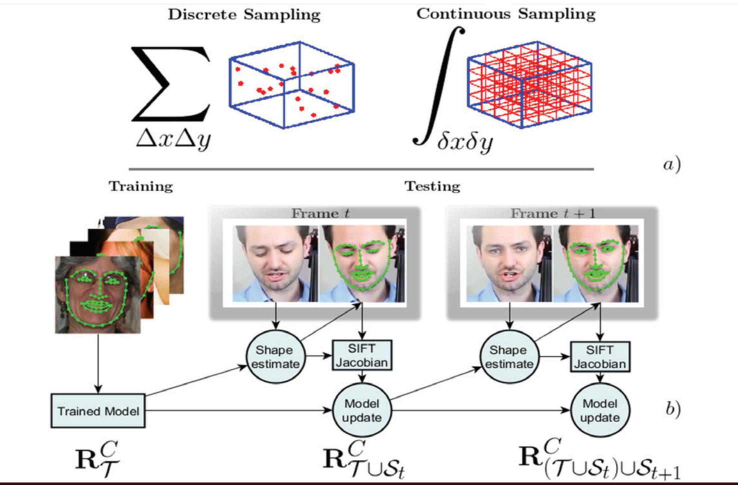
Computer Vision Laboratory
cvl.cs.nott.ac.uk
The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Cascaded Continuous Regression for Real-time Incremental Face Tracking

Enrique Sánchez-Lozano, Brais Martinez, Georgios Tzimiropoulos, Michel Valstar

Computer Vision Laboratory, School of Computer Science, University of Nottingham, UK



Code Available!!
https://continuousregression.wordpress.com

Linear Regression $\xrightarrow{\text{Functional Data Analysis}}$ Continuous Regression [10]

$$\sum_{j=1}^{M}\sum_{k=1}^{K}\|loss(\mathbf{R},\delta\mathbf{p})\|_2^2 \qquad \sum_{j=1}^{M}\int\|loss(\mathbf{R},\delta\mathbf{p})\|_2^2 d\delta\mathbf{p}$$

Ideal Loss Function $\xrightarrow{\text{Monte-Carlo Sampling}}$ Linear Regression

$$\sum_{j=1}^{M}\int_{\delta\mathbf{p}}\boxed{p(\delta\mathbf{p})}\|loss(\mathbf{R},\delta\mathbf{p})\|_2^2 d\delta\mathbf{p} \qquad \sum_{j=1}^{M}\sum_{k=1}^{K}\|loss(\mathbf{R},\delta\mathbf{p})\|_2^2$$

Our method

Problem

- This paper addresses real-time incremental Face Tracking

Motivation

- SDM [1] does not support Incremental Learning (IL)
- Chehra [2] shows that SDM can be trained in parallel, but IL is very slow (~4 fps)

Contributions

- Continuous Regression revisited (show how to incorporate a data term in CR)
- Cascaded Continuous Regression (show that it performs the same as SDM)
- Incremental Learning for CCR, **one order of magnitude faster** than SDM
- Fully automatic system state-of-the-art results on 300VW [3]

Linear Regression

Ground-truth

$$\sum_{j=1}^{M}\sum_{k=1}^{K}\|\delta\mathbf{p}_{j,k}-\mathbf{R}f(\mathbf{I}_j,\boxed{\mathbf{p}_j^*}+\delta\mathbf{p}_{j,k})\|_2^2$$

$$\boxed{\mathbf{R}=\mathbf{Y}\mathbf{X}^T\left(\mathbf{X}\mathbf{X}^T\right)^{-1}}$$

Linear Regression Closed-form Solution

Cascaded Regression (par-SDM)

Cascade Level — Statistics previous level

$$\mathbf{p}_{j,k}^{\boxed{(i+1)}}\sim\mathcal{N}(\mathbf{p}_j^*+\boxed{\boldsymbol{\mu}^{(i)},\boldsymbol{\Sigma}^{(i)}})$$

Incremental Learning for par-SDM

Training set — Updating set

$$\mathbf{R}_{\mathcal{T}\cup\mathcal{S}}=\mathbf{R}_{\boxed{\mathcal{T}}}-\mathbf{R}_{\mathcal{T}}\mathbf{Q}+\mathbf{Y}_{\boxed{\mathcal{S}}}\mathbf{X}_{\mathcal{S}}^T\mathbf{V}_{\mathcal{T}\cup\mathcal{S}}$$
$$\mathbf{Q}=\mathbf{X}_{\mathcal{S}}\mathbf{U}\mathbf{X}_{\mathcal{S}}^T\mathbf{V}_{\mathcal{T}}$$
$$\mathbf{U}=(\mathbb{I}_K+\mathbf{X}_{\mathcal{S}}\mathbf{V}_{\mathcal{T}}\mathbf{X}_{\mathcal{S}}^T)^{-1}$$
$$\mathbf{V}_{\mathcal{T}\cup\mathcal{S}}=\mathbf{V}_{\mathcal{T}}-\boxed{\mathbf{V}_{\mathcal{T}}\mathbf{Q}}$$

$$K=10\,,\,m=24\,,\,d=2000$$

$$\boxed{\mathcal{O}(K^3)}\sim\boxed{\mathcal{O}(m^3)} \qquad \boxed{\mathcal{O}(d^3)}\ll\boxed{\mathcal{O}(d^2)}$$

Continuous Regression Revisited

- Original continuous regression: i.i.d. uniform sampling distribution-> very limited in practice
- We propose to incorporate a data-term which helps us define the sampling volume, in which correlations are allowed:

$$\sum_{j=1}^{M}\int_{\delta\mathbf{p}}\boxed{p(\delta\mathbf{p})}\|\delta\mathbf{p}-\mathbf{R}f(\mathbf{I}_j,\mathbf{p}_j^*+\delta\mathbf{p})\|_2^2\,d\delta\mathbf{p}$$

Data term $\longrightarrow$ $\boldsymbol{\mu}\,,\,\boldsymbol{\Sigma}$

- We extend the image representation of [10] to use HOG features

$$\partial_x f(\mathbf{I},\mathbf{s})\approx\frac{f(\mathbf{I},[\mathbf{s}_x+\Delta x,\mathbf{s}_y])-f(\mathbf{I},[\mathbf{s}_x-\Delta x,\mathbf{s}_y])}{2\Delta x}$$

$$\mathbf{D}_j^*=[f(\mathbf{I}_j,\mathbf{p}_j^*)\,,\,\nabla_{\mathbf{p}}f(\mathbf{I}_j,\mathbf{p}_j^*)]$$

- Data term is parameterised by its first and second order statistics. **No prior assumption on the sampling distribution**!!!
- Sampling over images, not perturbations!!

$$\mathbf{A}=[\boldsymbol{\mu}\,,\,(\boldsymbol{\Sigma}+\boldsymbol{\mu}\boldsymbol{\mu}^T)] \qquad \mathbf{B}=\begin{pmatrix}1 & \boldsymbol{\mu}^T\\ \boldsymbol{\mu} & (\boldsymbol{\Sigma}+\boldsymbol{\mu}\boldsymbol{\mu}^T)\end{pmatrix}$$

$$\hat{\mathbf{B}}=\mathbf{B}\otimes\mathbb{I}_M$$

Continuous Regression Closed-form Solution

$$\boxed{\mathbf{R}=\mathbf{A}\left(\sum\mathbf{D}_j^*\right)^T\left(\mathbf{D}_{\mathcal{T}}^*\hat{\mathbf{B}}\mathbf{D}_{\mathcal{T}}^{*T}\right)^{-1}}$$

Incremental CCR (iCCR)

- Incremental updates are now:

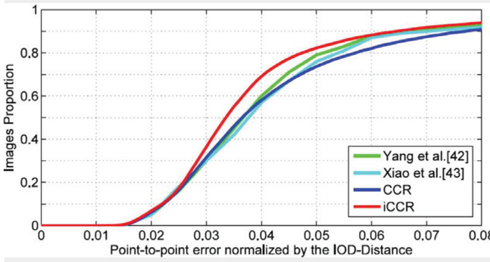
$$\mathbf{R}_{\mathcal{T}\cup\mathcal{S}}=\mathbf{A}\left(\sum\mathbf{D}_j^*+\mathbf{D}_{\mathcal{S}}\right)^T\mathbf{V}_{\mathcal{T}\cup\mathcal{S}}^{-1}$$

$$\mathbf{V}_{\mathcal{T}}^{-1}-\mathbf{V}_{\mathcal{T}}^{-1}\mathbf{D}_{\mathcal{S}}(\mathbf{B}^{-1}+\mathbf{D}_{\mathcal{S}}^*\mathbf{V}_{\mathcal{T}}^{-1}\mathbf{D}_{\mathcal{S}})^{-1}\mathbf{D}_{\mathcal{S}}^{*T}\mathbf{V}_{\mathcal{T}}^{-1}$$

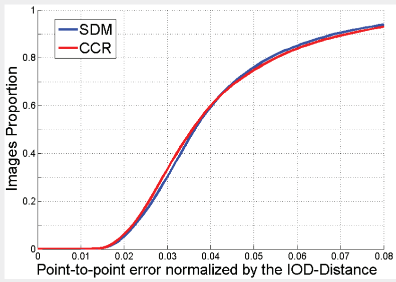
iCCR is an order of magnitude faster!!!!

CCR vs iCCR vs State of the Art

- Fully automatic system
- Category C of 300VW
- The Incremental Learning is crucial to attain State of the Art results!



CCR vs SDM (self-implementation)



- CCR is equivalent to SDM, but training is faster

Acknowledgments

- EU-Horizon 2020 Aria-Valuspa (http://aria-agent.eu)
- EPSRC Facial Deformable Models of Animals

References

[1] X. Xiong and F. De la Torre: Supervised Descent Method and its application to face alignment. CVPR 2013.
[2] A. Asthana, S. Zafeiriou, S. Cheng and M. Pantic: Incremental face alignment in the wild. In CVPR 2014.
[3] J. Shen, S. Zafeiriou, G. Chrysos, J. Kossaifi, G. Tzimiropoulos and M. Pantic: The first facial landmark tracking in-the-wild challenge: benchmark and resuts. In ICCV-W 2015
[10] E. Sánchez-Lozano, F. De la Torre and D. González-Jiménez: Continuous Regression for non-rigid image alignment. In ECCV 2012
[42] J. Yan, J. Deng, K. Zhang and Q. Liu: Facial shape tracking via spatio-temporal cascade shape regression. In ICCV-W 2015
[43] S. Xiao, S. Yan and A. Kassim: Facial landmark detection via progressive initialization. In ICCV-W 2015