Diffeomorphic Temporal Alignment Nets

Ron Shapira Weber[1], Matan Eyal[1], Nicki Skafte Detlefsen[2], Oren Shriki[1] and Oren Freifeld[1]

1. Ben-Gurion University
2. Technical University of Denmark

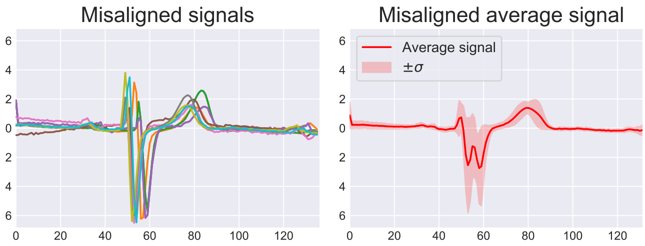
October 2019

Author's summary (NeurIPS 2019)

Outline

Problem Formulation - Time-Series Joint Alignment

- Time-series data often presents a significant amount of nonlinear misalignment.
- This may include simple phase-shift between similar instances, artifacts due to sampling methods and/or stretching and squeezing of different regions of the signal.
- Even trivial task such as computing the arithmetic mean (i.e. average signal) is hard under nonlinear misalignment .



Problem Formulation - Time-Series Joint Alignment

- Let $(\boldsymbol{U}_i)_{i=1}^{N}$ be a set of N time-series observations. The nonlinear misalignment could be written as:

$$(\boldsymbol{U}_i)_{i=1}^{N} = (\boldsymbol{V}_i \circ W_i)_{i=1}^{N} \tag{1}$$

- Where $\boldsymbol{U}_i$ is the i^{th} misaligned signal, $\boldsymbol{V}_i$ is the i^{th} latent aligned signal and W_i is a latent warp of the domain of $\boldsymbol{V}_i$.
- For technical reasons, the misalignment is usually viewed in terms of $T_i \triangleq W_i^{-1}$, the inverse warp of W_i, implicitly suggesting W_i is invertible.
- It is also typically assumed that $(T_i)_{i=1}^{N}$ belong to some nominal family of warps, parametrized by $\boldsymbol{\theta}$:

$$(\boldsymbol{V}_i)_{i=1}^{N} = (\boldsymbol{U}_i \circ T^{\boldsymbol{\theta}_i})_{i=1}^{N}\,, \quad T_i = T^{\boldsymbol{\theta}_i} \in \mathcal{T}\ \forall i \in (1, \ldots, N)\,. \tag{2}$$

Problem Formulation - Time-Series Joint Alignment

- We argue that this problem should be seen as a learning one, mostly due to the need for generalization, and define:

Definition (the joint-alignment problem)

Given the observations $(\boldsymbol{U}_i)_{i=1}^N$, infer the latent $(T^{\boldsymbol{\theta}_i})_{i=1}^N \subset \mathcal{T}$.

- To solve the joint alignment problem one must define a:
  - Warp family - $\mathcal{T}$
  - Model capable of generalizing the learned alignment.
  - Metric measuring the misalignment.
- We will define each in following slides.

Preliminaries - Spatial/Temporal Transformer Nets

- Introduced by Jaderberg et al. (2015) , the Spatial Transformer Nets (STN) is an end-to-end differential module that allows for certain invariances to spatial warps.
- Paramerized by any differential transformation family (*i.e.* Affine), the STN Learns input-depended spatial transformation to allow for robuts spatial invariance in deep neural networks.

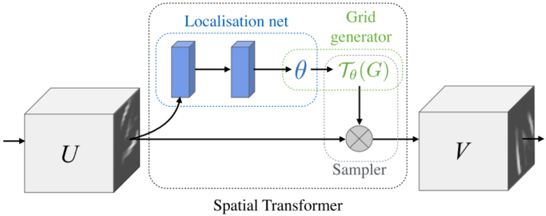


Figure taken from: Jaderberg et al. NIPS (2015)

Preliminaries - Spatial/Temporal Transformer Nets

- Temporal Transformer Nets (TTN), the time-series analog of STN.

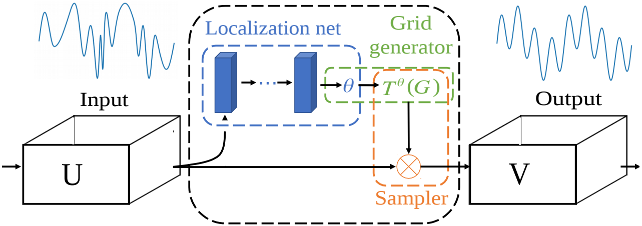


Figure taken from: Skafte et al. CVPR (2018)

Preliminaries - Diffeomorphisms

- As previously mentioned, $\mathcal{T}$ needs to be specified. In the context of time warping, diffeomorphisms are a natural choice (Mumford & Desolneux, 2010).

Definition (Diffeomorphisms)

A diffeomorphism is a differentiable invertible map with a differentiable inverse.

- Unfortunately, most representations of highly-expressive diffeomorphisms are complicated and computationally expensive.
- In our case, since the proposed method explicitly incorporates them in a DL architecture, it is even more important to drastically reduce the computational difficulties.

Preliminaries - CPAB

- We now define the warp family to be incorporated in our model.
- Following Skafte et al. (CVPR, 2018) incorporation of diffeomorphisms in STNs, the CPAB warps that had been proposed by Freifeld et al. (ICCV,2015; PAMI 2017) and are also used in this work.
- While any diffeomorphism could be integrated into DTAN, we chose CPAB warps since they:
  1. Are highly expressive and efficient.
  2. Offer an efficient and highly-accurate way to evaluate $x \mapsto \nabla_{\boldsymbol{\theta}} T^{\boldsymbol{\theta}}(x)$ (*i.e.* CPAB gradient).

Preliminaries - CPAB

- The name CPAB, short for CPA-Based, is due to the fact that these warps are based on Continuous Piecewise-Affine (CPA) velocity fields. The term "piecewise" is w.r.t. a partition, denoted by Ω, of the signal's domain into subintervals.
- Let $\mathcal{V}$ denote the linear space of CPA velocity fields w.r.t. such a fixed Ω, let $d = \dim(\mathcal{V})$, and let $v^{\boldsymbol{\theta}} : \Omega \to \mathbb{R}$, a velocity field parametrized by $\boldsymbol{\theta} \in \mathbb{R}^d$, denote the generic element of $\mathcal{V}$, where $\boldsymbol{\theta}$ stands for the coefficient w.r.t. some basis of $\mathcal{V}$ is

$$\mathcal{T} \triangleq \{T^{\boldsymbol{\theta}} : x \mapsto \phi^{\boldsymbol{\theta}}(x;1) \text{ s.t. } \phi^{\boldsymbol{\theta}}(x;t) = x + \int_0^t v^{\boldsymbol{\theta}}(\phi^{\boldsymbol{\theta}}(x;\tau))\, \mathrm{d}\tau \text{ where } v^{\boldsymbol{\theta}} \in \tag{3}$$

A. Abid and J. Zou. Autowarp: Learning a warping distance from unlabeled time series using sequence autoencoders. *arXiv preprint arXiv:1810.10107*, 2018.

O. Freifeld, S. Hauberg, K. Batmanghelich, and J. W. Fisher III. Highly-expressive spaces of well-behaved transformations: Keeping it simple. In *ICCV*, 2015.

O. Freifeld, S. Hauberg, K. Batmanghelich, and J. W. Fisher III. Transformations based on continuous piecewise-affine velocity fields. *IEEE TPAMI*, 2017.

M. Jaderberg, K. Simonyan, A. Zisserman, et al. Spatial transformer networks. In *Advances in neural information processing systems*, pages 2017–2025, 2015.

D. Mumford and A. Desolneux. *Pattern theory: the stochastic analysis of real-world signals*. AK Peters/CRC Press, 2010.

N. Skafte Detlefsen, O. Freifeld, and S. Hauberg. Deep diffeomorphic transformer networks. In *CVPR*, 2018.