Deep Video Denoising for Facial Signal Processing
Xiaoyan Li
University of Toronto

Motivation

- Spatial-temporal denoising. Temporal coherency.
- Video contains much more information.
- FastDVDnet (supervised) and UDVD (unsupervised).
- Apply two advanced video denoising methods to facial videos.
- Ground-truth video is not available.
- Image photoplethysmography (iPPG) reals the change of blood flow.
- Clean video could reflect blood flow better.

Proposed Method

- **UDVD**
  \[ L(\mu, \Sigma) = \frac{1}{2} \left[ (y - \mu)^T (\Sigma + \sigma^2 I) (y - \mu) \right] + \frac{1}{2} \log(\Sigma + \sigma^2 I) \]
  Where \( \mu \in \mathbb{R}^3 \) and \( \sigma \in \mathbb{R}^3 \) denote the mean vector and covariance matrix; \( y \) is observed noisy pixel

- **FastDVDnet**
  \[ L(\theta) = \frac{1}{2\tau} \sum_{i=1}^{m} \| P_i - \hat{P}_i \|^2, \]
  where \( \hat{P}_i = F(S_i^T, M_i^T; \theta) \) is the output of the network, and \( \theta \) is the set of all learnable parameters.

Related Work

- **FastDVDnet**
  1. Supervised learning
  2. Fast, a wide range of noise level.

- **UDVD**
  1. Unsupervised learning.

Experimental Results

- Results of frame denoising

- Results of iPPG extraction

References