

个人工作

MINGRUIBO

- 阅读 TDAN、DUF、TSM 并分享。
 - DUF：首次将 DCN（最初命名为 Dynamic Upsampling Filters）应用到 Video Super-Resolution 问题上。
 - DUF: Dynamic Upsampling Filters
 - framework = DUF + Residual
 - TDAN：提出 EDVR 的 framework，也可以认为首次使用了可变卷积核（毕竟之前不叫 PCD）。
 - <https://github.com/YapengTian/TDAN-VSR-CVPR-2020>
 - TDAN: Temporally Deformable Alignment Network
 - framework = Feature Extraction + Deformable Alignment + Aligned Frame Reconstruction
 - TSM：是一套与 EDVR 系列完全不同的框架，使用 Temporal Shift Module 进行时序建模，并提供了两种版本：online(uni-direction)、offline(bi-direction)。
 - <https://github.com/mit-han-lab/temporal-shift-module>
 - TSM: Temporal Shift Module
 - 实质上是一种 pseudo 3D-CNN
 - 只对其中一部分通道进行 shift，节省了内存开销；由于 shift 操作本身并不会引入计算，计算复杂度与 2D-CNN 相同。
 - 进行了时序建模能力（3DCNN）和空间建模能力（2DCNN）之间的 trade-off

ZHENGYUNLONG

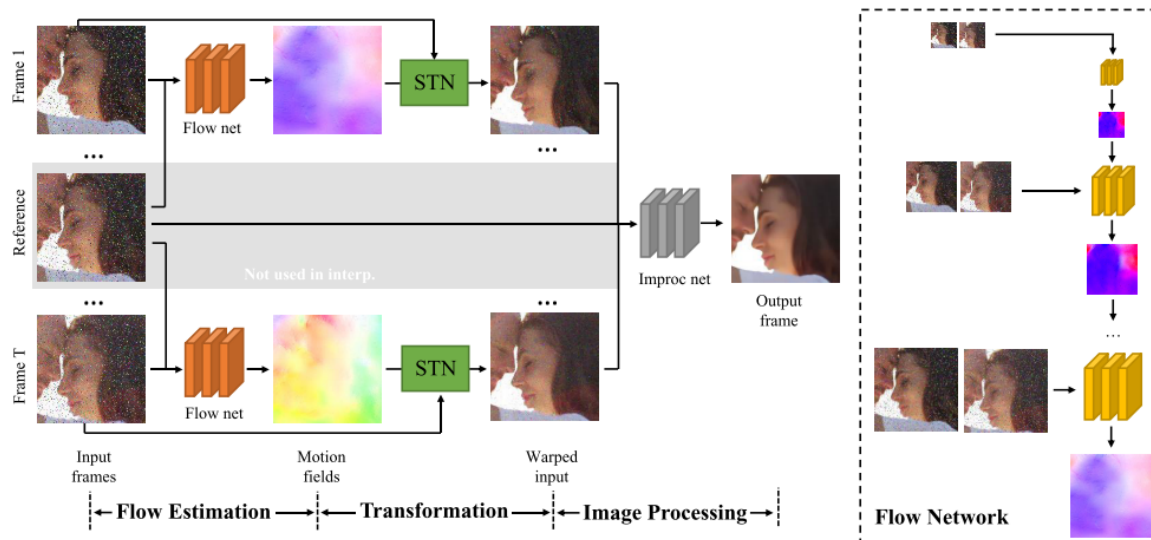
- 分享 Task Oriented Flow 模型。

介绍

Task-Oriented Flow 模型简称 TOFlow 模型，与传统的通过 CNN 训练光流相比，TOFlow 没有单独训练光流，而是整个网络联合训练，去学习某个任务中最适合表达特征的光流。

该模型适用于视频补帧、去噪、超分辨任务，不同任务模型结构及参数有所差异。

TOFlow 模型结构



如图所示，TOFlow模型分为三个模块：Flow Estimation(流估计)，Transformation(转换)，Image Processing(图像处理)。

Flow Estimation:

此模块的输入为N帧（补帧N=3，去噪和超分辨N=7），经过N-1个结构相同的光流网络估计运动场。

（师兄指出：N-1个光流网络实际上是光流估计由粗到精的过程，经过降采样的、尺度最小的帧能估计到粗略的光流，根据该光流进一步得到上一尺度，依次递进）。

显然，依照参考帧，一次任务需进行两次运动场的估计，分别为参考帧之前的运动以及参考帧之后的运动。

需要特别注意的是，对于补帧任务无参考帧。

Transformation

此模块依据得到的运动场进行帧对齐，TOFlow模型采用的是STN（Spatial Transfoemation Network）用于帧对齐。

Image Processing

此模块用于生成最终结果，对于超分辨任务，该模块有四个卷积层和RELU层。

数据集

该模型提出的论文采用了Vimeo-90K数据集，数据集形式为每个样例7张图片，数据集共82G。

参考

Video Enhancement with Task-Oriented Flow, Tianfan Xue Baian Chen Jiajun Wu Donglai Wei William T. Freeman