# 综述报告——基于光流的视频插帧

**摘要：**视频插帧是利用低帧率视频生成高帧率视频的技术，可以显著提高视频的流畅度和清晰度，光流法是一种常用的视频插帧方法，它通过计算相邻帧之间的像素运动来估计新的帧。本文分析总结了主要的光流插帧模型，展示了光流插帧技术的研究进展和现状，探讨了各种算法的原理，并在最后总结了光流插帧技术存在的主要矛盾。

**关键字：**视频插帧，光流，深度学习

视频插帧是一种[计算机视觉算法](https://so.csdn.net/so/search?q=%E8%AE%A1%E7%AE%97%E6%9C%BA%E8%A7%86%E8%A7%89%E7%AE%97%E6%B3%95&spm=1001.2101.3001.7020" \t "https://blog.csdn.net/q7w8e9r4/article/details/_blank)，用于在视频中插入额外的帧以提高视频的流畅度和观看体验。视频插帧技术可以应用于各种领域，比如电影特效、视频游戏、运动捕捉等，在近年来被广泛研究和应用。基于光流的视频插帧通过计算视频相邻帧之间的双向光流，分析相邻帧之间的像素变化，从而推测出像素点在下一帧中的位置，然后使用合适的扭曲算法生成插值帧。

近年来，深度学习方法在视频插帧领域取得了很大的突破，通过使用卷积神经网络（Convolutional Neural Networks，CNN）等深度学习框架，模型可以学习到视频中的时空特征，从而生成更真实的光流、更逼真的插值帧，如SuperSloMo[1]、DAIN[2]、BMBC[3]、ABME[4]、RIFE[5]等。

1. SuperSloMo

模型的目标是生成对应时刻的中间帧，SuperSlomMo通过融合利用光流扭曲的输入图像，生成目标插帧。因为无法获取中间帧，直接计算光流是困难的，为了解决这一问题，可以通过插值的方式计算，同样利用时序加权平衡正反向光流的贡献。

（二）DAIN

DAIN采用与SuperSloMo相同的策略生成中间帧，即先生成光流，然后根据光流对前后帧进行扭曲。DAIN的优势在于引入深度作为光流插值的权值，将深度值较小的前景物体作为主要的光流源，缓解了光流插值中遮挡问题。

（三）BMBC

该模型设计双边光流估计的策略，利用光流网络直接生成中间帧到前后帧的光流。给定输入，模型依次估计双边光流、双向光流，根据这些运动近似四个前向双边运动；同时提取上下文特征，使用六个双边光流扭曲输入帧、上下文特征，获得三组候选中间帧；最后利用帧合成网络将候选帧和输入帧融合，获得最终的插帧结果。

（四）ABME

BMBC遵循线性运动的假设，对于非对称运动的光流估计仍有局限性，ABME在此基础上更进一步，提出非对称双边光流估计方法。

ABME首先利用BMBC中提到的双边光流估计网络提取对称的双边光流，然后基于此光流生成锚定帧，即插值帧的临时中间结果，然后利用ABMR-Net网络从前后帧、锚定帧中估计新的非对称的双边光流，最终将结果输入帧合成网络，得到最终的输出结果。

（五）RIFE

视频插帧的关键是光流的生成，前人的研究都是对目标光流进行近似，但是这种线性关系在面对复杂运动时总是存在局限性，因此RIFE提出直接利用端到端的框架设计新的模型，使模型摆脱线性的约束，同时简化模型，提升模型的运行速度。

RIFE设计IFBlock模块，并将其进行堆叠形成端到端的结构，每层的输入为前后帧、融合权重、光流，输出为新的融合权重、光流。

（六）总结

对基于光流的视频插帧而言，光流的准确性决定了最终插帧结果的质量，因此提取出更准确的中间帧到前后帧的光流是至关重要的，但是由于我们需要生成中间帧，而生成光流又需要前后帧，因此存在一个“先有鸡还是先有蛋”的问题，如何解决这一问题是进一步提升插帧模型性能的关键。

本文总结的若干方法中，前期方法如SuperSloMo、DAIN等利用前后帧间的光流对目标光流进行近似，中期方法如BMBC、ABME等将中间帧的思想引入光流网络，直接生成中间帧，但仍不能摆脱线性运动的假设，后期方法如RIFE提出利用端到端模型规避这一问题，总之都在探索提取出准确的光流。

# 参考文献

[1] JIANG H, SUN D, JAMPANI V, et al. Super slomo: High quality estimation of multiple intermediate frames for video interpolation; proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, F, 2018 [C].

[2] BAO W, LAI W-S, MA C, et al. Depth-aware video frame interpolation; proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, F, 2019 [C].

[3] PARK J, KO K, LEE C, et al. Bmbc: Bilateral motion estimation with bilateral cost volume for video interpolation; proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XIV 16, F, 2020 [C]. Springer.

[4] PARK J, LEE C, KIM C-S. Asymmetric bilateral motion estimation for video frame interpolation; proceedings of the Proceedings of the IEEE/CVF International Conference on Computer Vision, F, 2021 [C].

[5] HUANG Z, ZHANG T, HENG W, et al. Real-time intermediate flow estimation for video frame interpolation; proceedings of the European Conference on Computer Vision, F, 2022 [C]. Springer.