

## Применение глубокого обучения в задаче интерполяции кадров

В. И. Самсонов<sup>1,2</sup>

<sup>1</sup>Московский физико-технический институт (государственный университет)

<sup>2</sup>Яндекс, Кафедра анализа данных

Задача интерполяции кадров является одной из задач компьютерного зрения, в которой требуется увеличить число кадров видео для улучшения визуального восприятия человеком. В анимационных фильмах (мультфильмах) увеличение частоты кадров значительно улучшает качество видео. Многие анимационные фильмы рисуются художниками — интерполяция позволяет значительно сократить затраты на отрисовку и облегчить задачу художника, поскольку отрисовка каждого отдельного кадра занимает значительное время.

Правильная интерполяция кадров должна обрабатывать сложные нелинейные трансформации и учитывать структуру движения. Для этого требуется понимать семантическую составляющую кадров. Большинство простых методов не способны обеспечить достаточное качество интерполяции, особенно при больших смещениях объектов между кадрами.

В задаче интерполяции кадров большинство предложенных методов являются методами обучения без учителя (unsupervised). В данной работе представлен метод обучения с учителем с использованием глубоких сверточных сетей. Сверточные сети широко применяются в задачах семантической сегментации и текстовой разметки изображений, что показывает их способность к распознаванию объектов и пониманию семантической структуры изображений. Одним из главных достоинств метода обучения с учителем является возможность адаптации к характеру движений и визуального стиля конкретного видео.

В результате данной работы предложена нейросетевая модель, которая качественно решает задачу интерполяции кадров без использования дополнительной априорной информации. Также в данной работе показано, как использовать оптический поток в качестве априорной информации для обучения, что значительно сокращает время обучения, а также позволяет улучшить качество интерполяции.

### Литература

1. P. Fischer, A. Dosovitskiy, E. Ilg, P. Häusser, C. Hazirbas, V. Golkov, P. van der Smagt, D. Cremers, and T. Brox, “FlowNet: Learning optical flow with convolutional networks,” *CoRR*, vol. abs/1504.06852, 2015.
2. N. Mayer, E. Ilg, P. Häusser, P. Fischer, D. Cremers, A. Dosovitskiy, and T. Brox, “A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation,” *CoRR*, vol. abs/1512.02134, 2015.
3. G. Long, L. Kneip, J. M. Alvarez, and H. Li, “Learning image matching by simply watching video,” *CoRR*, vol. abs/1603.06041, 2016.
4. J. Johnson, A. Alahi, and F. Li, “Perceptual losses for real-time style transfer and super-resolution,” *CoRR*, vol. abs/1603.08155, 2016.
5. I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. C. Courville, and Y. Bengio, “Generative adversarial networks,” *CoRR*, vol. abs/1406.2661, 2014.