

Збільшення кількості кадрів в потоковому відео за інтерполяцією нейронної мережі

Вінницький національний технічний університет

Анотація

Запропоновано новий метод штучного збільшення кадрів у потоковому відео, який дозволяє відтворювати уповільнене відеофрагменту без розмитості та різкого переходу між кадрами

Ключові слова: проміжний кадр, нейронна мережа, інтерполяція

Abstract

A new method of artificially enlarging frames in streaming video is proposed, which allows you to play slow motion video without blurring and sharp transition between frames

Keywords: intermediate frame, neural network, interpolation

Вступ

За статистикою, більшість відеокамер знімають з відео з частотою близько 30 кадрів в секунду. Хоч і в наш час, все частіше зустрічаються камери, які мають змогу знімати відео з частотою навіть 240 кадрів в секунду, але тим не менш, це вимагає великої кількості пам'яті та не є економним варіантом у випадку мобільних пристроїв.

Суттєво зменшити використання ресурсів при зйомці відео з високою частотою кадрів, дозволяють нові системи, що генерують кілька проміжних кадрів, які є просторово і часово когерентними.

Принцип роботи збільшення кількості кадрів за допомогою згорткової нейронної мережі (CNN)

Більшість існуючих методів фокусуються на однокадровій інтерполяції DVF (яскравим прикладом є популярний плагін TwixtorPro, який використовується у відеоредакторах) [1]. Згорткова нейронна мережа (Convolutional Neural Network (CNN)) оцінює відеопотік - структуру руху, об'єктів, поверхонь і ребер в сцені. Робить вона це і вперед по часовій шкалі, і в зворотній послідовності для двох вхідних кадрів. Потім система передбачає, як пікселі будуть переміщатися з одного кадру в наступний, створюючи 2D-вектори цих переміщень[2]. Потім працює друга нейромережа, яка передбачає карту видимості - виключає ті пікселі, які повинні бути перекриті об'єктами, щоб прибрати артефакти. І система за допомогою всіх отриманих даних спотворює нові кадри між двома вихідними, щоб забезпечити плавність переходу.

В якості об'єкта використання для досягнення мети, було взято бібліотеку глибинного навчання PyTorch і графічні процесори Nvidia Tesla V100 GPUs [3]. Систему тренували за допомогою 11 000 відеороликів повсякденній і спортивній активності, знятих з частотою в 240 кадрів в секунду. Завдяки цьому вона стала прогнозувати проміжні кадри.

Простий спосіб досягти цього, це тренувати її безпосередньо виведенням RGB-пікселів [4]. Для цього, однак, мережа повинна навчитися інтерпретувати не тільки патенти руху, але і їх появу у вигляді двох вхідних зображень шляхом запозичення оптичного потоку з пікселів (Рисунок 1). У RGB важко

генерувати високоякісне проміжне зображення таким чином, через що виникають дефекти у зображенні. Щоб компенсувати це, використовується функція зворотної деформації (Рисунок 2), яка може бути реалізована за допомогою білінійної інтерполяції і є диференційованою [5].

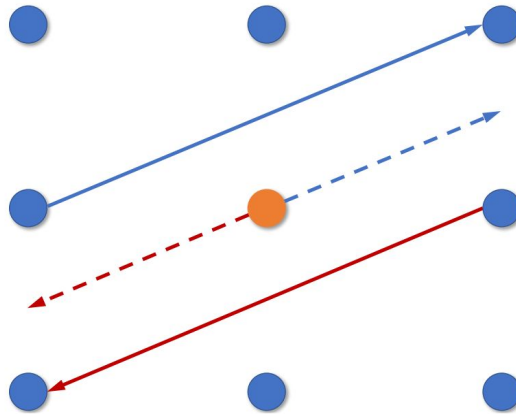


Рисунок 1 - Ілюстрація наближення проміжного оптичного потоку. Помаранчевий піксель запозичує оптичний потік з пікселів на першому та другому зображеннях.

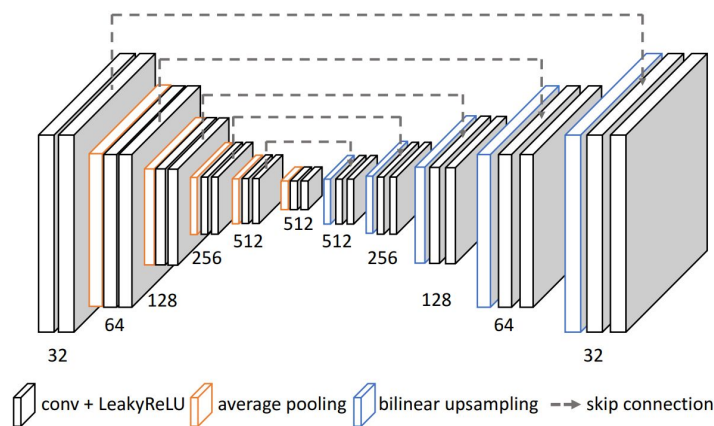


Рисунок 2 - Ілюстрація архітектури обчислень потоку та інтерполяції потоків CNN.

Висновок

Завдяки інтерполяції нейронної мережі, ми можемо створювати стільки проміжних відеокадрів, скільки потрібно між двома вхідними зображеннями. Технологія дозволяє робити відео набагато більш плавними і менш розмитими, ніж у випадку звичайного уповільнення швидкості. Звичайно, є відмінності між штучно створеними уповільненими відео і оригіналом, відзнятим з високою частотою кадрів, але за рахунок економії використаних ресурсів, даний метод стає більш доступним для звичайних користувачів.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black, and R. Szeliski. A database and evaluation methodology for optical flow. IJCV, 92(1):1–31, 2011. 2, 3, 6, 7, 8
2. J. Barron, D. Fleet, and S. Beauchemin. Performance of optical flow techniques. IJCV, 12(1):43–77, 1994. 2
3. T. Brox and J. Malik. Large displacement optical flow: Descriptor matching in variational motion estimation. PAMI, 33(3):500–513, 2011. 2, 8

4. D. J. Butler, J. Wulff, G. B. Stanley, and M. J. Black. A naturalistic open source movie for optical flow evaluation. In ECCV, 2012. 2, 6
5. A. Dosovitskiy, P. Fischery, E. Ilg, C. Hazirbas, V. Golkov, P. van der Smagt, D. Cremers, T. Brox, et al. FlowNet: Learning optical flow with convolutional networks. In ICCV, 2015.

Петрук Вадим Вікторович — студент групи ЛОТ-19м, факультет комп'ютерних систем та автоматики, Вінницький національний технічний університет, Вінниця, e-mail: wadimosss@gmail.com.

Науковий керівник: **Кожем`яко Андрій Вікторович** — к.т.н, доцент кафедри лазерної та оптоелектронної техніки Вінницького національного технічного університету, м. Вінниця.

Vadim Petruk Viktorovich — student of Lot-19m group, Faculty of Computer Systems and Automation, Vinnytsia National Technical University, Vinnytsia, email: wadimosss@gmail.com.

Supervisor: **Kozemyako Andriy Viktorovich** — Cand. Sc. (Eng.), Docent of Laser and Optoelectronic Technology, Vinnitsa National Technical University, Vinnitsa.