# Исследование возможности применения свёрточных нейронных сетей для обработки фотограмметрических изображений

В.В. Пинчуков <sup>1</sup>, А.Ю. Поройков <sup>1</sup>, Е.В. Шматко <sup>1</sup>

<sup>1</sup> Национальный исследовательский университет «МЭИ», Красноказарменная улица, д. 14, Москва, 111250, Россия

#### Аннотация

Ближняя фотограмметрия широко используется для измерения формы поверхности различных объектов и ее деформаций. Обычно для этого используют стереопару изображений исследуемого объекта, полученных с разных ракурсов с помощью нескольких цифровых видеокамер. Форма поверхности измеряется путем триангуляции набора соответствующих двухмерных точек с этих изображений с помощью заранее известного расположения камер относительно друг друга. Для поиска таких точек применяют фотограмметрических методов различные алгоритмы. Несколько используют кросскорреляцию для этой цели. В данной работе рассматривается возможность замены корреляционного алгоритма с помощью нейросетей для определения смещений на изображениях. Они позволяют увеличить скорость расчета и пространственное разрешение результатов измерений. Для проверки возможности их применения была получена серия экспериментальных изображений поверхностей с различной деформацией. Были проведены вычислительные эксперименты по обработке этих изображений с помощью выбранных нейросетей и классического алгоритма кросскорреляции. Определены ограничения по использованию сравниваемых алгоритмов и оценена их погрешность при восстановлении трехмерной формы поверхности. Проведенное физическое моделирование для проверки выбранных нейросетей к обработке изображений для задачи фотограмметрии показали их работоспособность и эффективность.

#### Ключевые слова

Ближняя фотограмметрия, кросскорреляция, метод корреляции фоновых изображений, сверточные сети.

# Exploring the Application of Convolutional Neural Networks for Photogrammetric Image Processing

V.V. Pinchukov<sup>1</sup>, A.Yu. Poroykov<sup>1</sup>, E.V. Shmatko<sup>1</sup>

<sup>1</sup> National Research University "Moscow Power Engineering Institute", Krasnokazarmennaya 14, Moscow, 111250, Russia

#### Abstract

Close-range photogrammetry is widely used to measure the surface shape of various objects and its deformations. Usually, a stereo pair of images of the object under study, obtained from different angles by means of several digital video cameras, is used for this purpose. The surface shape is measured by triangulating a set of corresponding two-dimensional points from these images using a predetermined location of cameras relative to each other. Various algorithms are used to find these points. Several photogrammetric methods use cross-correlation for this purpose. This paper discusses the possibility of replacing the correlation algorithm with neural networks to determine

EMAIL: vpinchukov@list.ru (В.В. Пинчуков); poroykovay@gmail.com (А.Ю. Поройков); shmatko.97@bk.ru (Е.В. Шматко) ORCID: 0000-0002-5816-0424 (В.В. Пинчуков); 0000-0002-9284-1397 (А.Ю. Поройков); 0000-0003-3565-9502 (Е.В. Шматко)



<sup>© 2022</sup> Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

ГрафиКон 2022: 32-я Международная конференция по компьютерной графике и машинному зрению, 19-22 сентября 2022 г., Рязанский государственный радиотехнический университет им. В.Ф. Уткина, Рязань, Россия

offsets in the images. They allow to increase the calculation speed and the spatial resolution of the measurement results. To verify the possibility of their application, a series of experimental images of surfaces with different deformations were obtained. Computational experiments were performed to process these images using selected neural networks and a classical cross-correlation algorithm. The limitations on the use of the compared algorithms were determined and their error in restoring the three-dimensional shape of the surface was estimated. The physical simulation to verify the selected neural networks for image processing for the task of photogrammetry showed their performance and efficiency.

#### Keywords

Close-range photogrammetry, cross correlation, image pattern correlation technique, convolution networks.

## 1. Введение

Один из методов для восстановления трехмерной поверхности основан на кросскорреляционной обработке стереопар изображений. Такой метод называется методом корреляции фоновых изображений (МКФИ, в англоязычных источниках IPCT) [1-4]. Он основан на алгоритмах обработки другого метода – анемометрии по изображению частиц (PIV) и является другим вариантом метода корреляции цифровых изображений (DIC). В работе описана попытка применения методов машинного обучения для замены кросскорреляционного расчета в методе IPCT.

Одним из первых предложений по применению нейросетей к задачам PIV предложено в работе [5]. Так как это были первые попытки в новом направлении, то предложенные идеи сильно пересекались с кросскорреляционными алгоритмами. Так же загружались в сеть два окна опроса 32×32 пикселя с двух изображений и сеть предсказывала вектор смещения, соответствующий этим окнам опроса. Первые попытки хоть и не показали более лучших результатов по сравнению с уже известными методами, но они показали работоспособность идеи, которую в дальнейшем стали развивать. Подобное исследование проводилось и в работе [6], где изучалась сеть с архитектурой, предложенной в [7]. Более подробно методы машинного обучения для задач диагностики в гидродинамике приведены в работах [8-9].

В результате обзора литературы по методам машинного обучения для задачи измерения формы поверхности была выбрана нейросеть PIV-LiteFlowNet-en [10], которая основана на сети LiteFlowNet [11]. Обе эти сети являются потомками сети FlowNet [12], предназначенной для приема двух изображений на вход и оценки смещения оптического потока на выходе. По данным [12], сеть обучается на большом синтетическом наборе данных и обеспечивает приемлемую точность оценки жесткого движения. Однако исходная FlowNet не может быть напрямую применена к задачам PIV, что было показано в работе [13].

Большое количество нейросетей, которые были рассмотрены в обзоре, были реализованы с на основе библиотеки caffe. На данный момент их запуск связан с большим количеством технических трудностей из-за устаревшей кодовой базы. Поэтому произведён повторный поиск реализаций этих сетей на библиотеке pytorch. Для дальнейшего исследования и сравнения нейросети с кросскорреляционным алгоритмом будут использоваться две сети как PIV-LiteFlowNet-en, так и LiteFlowNet. Вторая сеть позволяет оценивать большие смещения, что является важным фактором, что будет показано позже. PIV-LiteFlowNet-en в отличии от LiteFlowNet на выходе сети даёт разрешение изображения равным входному изображению без использования билинейной интерполяции, что повышает точность определения малых смещений.

Обе сети требуют использования GPU с поддержкой CUDA для запуска. Важным является объём видеопамяти карты, так как о него зависит какого разрешения изображения возможно обработать на видеокарте. Благодаря сервису Google Colab, который выделяет на пользователя около 11 GiB памяти GPU, возможно обработать изображения размером 1900×2000 пикселей с глубиной цвета 24 бит формата jpg. Этого достаточно для обработки экспериментальных изображений, используемых в данной работе, при полном разрешении в рабочей области.

Для оценки результатов обработки сетями были получены 150 пар экспериментальных изображений поверхности с различными деформациями. Для этого был использован имитатор деформируемой поверхности, описанный в [14]. Имитатор позволяет произвольно устанавливать форму гибкой поверхности с помощью цифровых сервомашинок.

Для восстановления формы поверхности был использован метод IPCT по алгоритму, описанному в [4]. Первый этап алгоритма – это поиск кодовых маркеров на изображении для проведения перспективного преобразования изображений стереопары. После преобразования измеряемая поверхность будет ориентирована перпендикулярно оптической оси камеры. Размер каждой стереопары будет индивидуален, но обычно не превышает 1700×1500 пикселей. Второй этап расчета – это кросскорреляционный анализ, в результате которого получаются векторные поля смещений точек поверхности между изображениями стереопары. Третий этап – это расчет триангуляции для определения трехмерных координат поверхности. На рисунке 1 приведен пример обработки стереопары с помощью сети PIV-LiteFlowNet-en. Результаты перспективного преобразования 1(в-г).



**Рисунок 1** – Пример обработки стереопары, все отсчёты измеряются в пикселях: а, б – исходные изображения; в, г – результаты перспективного преобразования; д – визуализация разреженного векторного поля; е – визуализация векторного поля в полном разрешении; ж – представление амплитуды деформаций при помощи цветовой карты

## 2. Ограничения и поиск равнозначных условий

Прежде чем сравнить две выбранных сети и кросскорреляционный расчет, необходимо определить равнозначные условия для этих алгоритмов. На рисунке 2 приведены векторные поля для одной и той же стереопары, но при разных разрешениях этих изображений. Из рисунка 2 видно, что сеть PIV-LiteFlowNet-en не справляется со смещениями больше ~12-13 писклей. Для сети LiteFlowNet этот показатель ~80-90 пикселей.

Для сравнения сетей с различными диапазонами измеряемых смещений было решено провести расчет для изображений с разными исходными разрешениями. В этом случае смещения на изображениях будут изменяться пропорционально их размеру. Это позволит сравнить результаты работы алгоритмов на одних и тех же экспериментальных данных. Расчёт СКО ошибки репроекции был проведен для 10 разрешений каждой экспериментальной стереопары. Ряд разрешений, использованный в расчетах, был получен по следующей формуле

$$R_k = R_0 - \frac{R_0}{10}k,$$
 (1)

где k = 0, 1, 2...9;  $R_0$  – исходный размер стороны изображения. При этом соотношение сторон изображений сохраняется. Для алгоритма на основе кросскорреляции необходимо также определить параметры для обработки изображений. Универсальные параметры выбрать невозможно, т.к. смещения могут достигать более 100 пикселей. Поэтому размер окна опроса

необходимо определять индивидуально для каждого разрешения. Итоговые параметры обработки тремя тестируемыми алгоритмами представлены в таблице 1.



**Рисунок 2** – Результат обработки PIV-LiteFlowNet-en для разных исходных разрешений изображений

Разрешение, пикселей	Размер окна опроса, пикселей	Смещение окна опроса, пикселей
1700×1500	256	128
1530×1350	256	128
1360×1200	196	98
1190×1050	196	98
1020×900	128	64
850×750	128	64
680×600	64	32
510×450	64	32
340×300	32	16
170×150	32	16

Таблица 1 – Параметры обработки изображений с помощью тестируемых алгоритмов

На рисунке 3 представлен результат тестирования трех алгоритмов. Каждая кривая является усреднением по 150 стереопарам. Для проверки влияния общей интенсивности на изображении на основе каждой пары изображений создавалась инверсная версия этой пары. Это связано с тем, что для метода IPCT характерны черные точки на белом фоне, а для PIV белые точки на черном фоне.

Из результатов обработки можно сделать следующие выводы.

1. Алгоритм кросскорреляции показывает стабильную ошибку репроекции почти для всех разрешений. При этом на него не влияет инверсия интенсивности.

2. LiteFlowNet показывает лучшие результаты среди всех алгоритмов, а инверсия изображения ухудшает его работу.

3. PIV-LiteFlowNet-en показывает плохие результаты из-за больших смещений на изображениях. При разрешении 340×300 пикселей смещения становятся достаточно малыми, но из-за большого сжатия качество изображений не позволяет алгоритму добиться высокой точности. Инверсия изображений улучшает работу алгоритма.

4. При разрешении в 340×300 и 170×150 все алгоритмы показывают ухудшение точности из-за сильного сжатия изображения.



Разрешение, пикс.

**Рисунок 3** – Среднее значение СКО ошибки репроекции для разных разрешений по трём исследуемым алгоритмам

Для лучшего понимания поведения алгоритмов на рисунке 4 приведены графики СКО ошибки репроекции для двух стереопар: при малых и больших смещениях на изображениях. На рисунке 4(а) графики в одной плоскости для малых и больших смещений, 4(б) величина этого смещения для каждого разрешения, 4(в) увеличенная область 4(а), демонстрирующая поведение алгоритмов при малых смещениях. Сравнивая графики, можно ещё раз убедиться, что максимально оцениваемое смещения сетью PIV-LiteFlowNet-en составляет около 10 пикселей, а сетью LiteFlowNet около 80 пикселей.

Для кросскорреляции оценка максимального смещения зависит от размера окна опроса. Максимальное измеряемое смещение должно быть меньше 1/2 или 1/3 размера окна опроса. По графикам 4(б) и 4(в) видно, что как только смещения на изображениях опускаются до 10 и ниже пикселей, то сеть PIV-LiteFlowNet-en показывает лучшие результаты по сравнению с другими, что еще раз подтверждает наличие максимального оцениваемого смещения. Среднее СКО ошибки репроекции на рисунке 4 для всех алгоритмов довольно велик, т.е. близок или больше 1 пикселя. Это объясняется тем, что при больших смещениях, как рисунке 4, ошибка резко возрастает, поэтому средний уровень ошибки подымается.

#### 3. Сравнение алгоритмов

Исходя из графиков на рисунке 4 минимум ошибки для сети PIV-LiteFlowNet-en достигается при разрешении 510:450, для сети LiteFlowNet при 680:600. На рисунке 5 представлены графики зависимости ошибки репроекции для каждой пары изображения из 150 снятых. Также для сравнения приставлено разрешение 850:750 и 1700:1500. Все графики для наглядности сортированы в порядке возрастания ошибки. На всех графиках, кроме случая 1700:1500, прослеживаются одинаковые закономерности. Сеть LiteFlowNet имеет меньшую ошибку в сравнении с кросскорреляционным алгоритмом при примерно одинаковой форме графика, лишь в нескольких случаях кросскорреляция превосходит по точности сеть. Сеть PIV-LiteFlowNet-en показывает лучшую точность среди всех алгоритмов примерно в 50 случаях. При этом она имеет лучше показатели при инверсии изображений. Исключением является случай при разрешении 510:450, где случай инверсии изображений не сильно отстаёт от исходных изображений.



**Рисунок 4** – Среднее значение СКО ошибки репроекции для различных амплитуд смещений: а – СКО ошибки репроекции для случая малых и больших смещений; б – величина смещения для двух случаев; в – увеличенная область графика (а) для демонстрации поведения алгоритмов при малых смещениях



**Рисунок 5** – Графики СКО ошибки репроекции для каждой снятой пары изображений и для трёх исследуемых алгоритмов при различных разрешениях

То, что сеть PIV-LiteFlowNet-en показывает лучшие результаты только ~50 случаях, связано с тем, что даже при разрешении 510:450 большинство пар изображений имеют смещения больше 10 пикселей. Поэтому лишь в 1/3 случаев эта сеть превосходит остальные алгоритмы.

Для демонстрации того, что не одни из алгоритмов не способен с высокой точностью обработать изображения в полном разрешении приведён случай 1700:1500, на котором видно, что приемлемая точность достигается кросскорреляцией примерно в 10 случаях и сетью LiteFlowNet примерно в 50 случаях, что даже не составляет половины всей выборки. Это связано с большими смещениями на изображениях.

# 4. Заключение

В работе рассмотрен вопрос применения нейросетей к восстановлению трёхмерной формы поверхности объекта методом стереофотограмметрии. Результаты их обработки сравнивались с уже зарекомендованным алгоритмом на основе кросскорреляции. Он с приемлемой скоростью позволяет оценить только разреженное векторное поле, по которому путём триангуляции рассчитываются трёхмерные точки. Для поиска решения этой проблемы произведён обзор методов машинного обучения, из которых выбраны две нейросети LiteFlowNet и PIV-LiteFlowNet-en. Эти сети позволяют оценить векторное поле в полном разрешении изображения при этом имеют большую скорость расчёта в сравнении с кросскорреляцией. Но полный выигрыш в быстродействии можно получить только с применением графического процессора.

Было установлено, что нейросети имеют ограничения по величине верно оцениваемого смещения. Для PIV-LiteFlowNet-еп это ограничение составило 12-13 пикселей, а для LiteFlowNet около 80 пикселей. Для первой сети это можно объяснить обучающей выборкой, а для второй – структурой сети. Также выявлена разница в обработке сетей исходных и инверсных изображений, что также является следствие обучающих выборок.

По результатам обработки LiteFlowNet превзошла алгоритм на основе кросскорреляции и PIV-LiteFlowNet-en в сумме по всех разрешениях изображений. Но если сравнивать в рамках ограничений алгоритмов, то PIV-LiteFlowNet-en имеет лучшую точность. При этом для обработки изображений характерных для фотограмметрии в полном разрешении ни один из методов не является удовлетворительным. Для полноценного применения таких нейросетей требуется их модификация под исследуемую задачу.

Проведенное физическое моделирование для проверки выбранных подходов к обработке изображений для задачи фотограмметрии показали их работоспособность и эффективность. Но для их применения на практике необходимо решить несколько проблем. Выбранные нейросети не полностью подходят для исследуемой задачи из-за ограничения величины оцениваемого смещения и высокой сложности их запуска. Для успешного практического применения машинного обучения необходима модификация структуры выбранных нейросетей, или разработка собственной структуры, и их обучения на экспериментальных изображениях, характерных именно для фотограмметрии.

# 5. Благодарности

Работа выполнена в рамках проекта «Разработка фотограмметрической системы измерения формы поверхности для условий повышенных вибрационных нагрузок» при поддержке гранта НИУ «МЭИ» на реализацию программ научных исследований «Энергетика», «Электроника, радиотехника и IT» и «Технологии индустрии 4 для промышленности и робототехника» в 2020-2022 гг.

# 6. Список источников

Bakunowicz J., Święch Ł., Meyer R. Measuring structure deformations of a composite glider by optical means with on-ground and in-flight testing // Meas. Sci. Technol. 2016. Vol. 27, № 12. DOI: 10.1088/0957-0233/27/12/124013.

- [2] Kirmse T. Recalibration of a stereoscopic camera system for in-flight wing deformation measurements // Meas. Sci. Technol. IOP Publishing. 2016. Vol. 27, № 5. P. 54001. DOI:10.1088/0957-0233/27/5/054001.
- [3] Advanced In-Flight Measurement Techniques / F. Boden [et al.]. 2013. DOI: 10.1007/978-3-642-34738-2.
- [4] Crosscorrelation image processing for surface shape reconstruction using fiducial markers / E.V. Shmatko, V.V. Pinchukov, A.D. Bogachev, A.Yu. Poroykov // J. Physics: Conf. Ser. 2021. Vol. 2127. P. 012030. DOI: 10.1088/1742-6596/2127/1/012030.
- [5] Rabault J., Kolaas J., Jensen A. Performing particle image velocimetry using artificial neural networks: a proof-of-concept // Meas. Sci. Technol. 2017. Vol. 28, №12. P. 125301 DOI: 10.1088/1361-6501/aa8b87.
- [6] Lee Y., Yang H., Yin Z. PIV-DCNN: cascaded deep convolutional neural networks for particle image velocimetry // Experiments in Fluids. 2017. Vol. 58, № 12. P. 171. DOI:10.1007/s00348-017-2456-1.
- [7] Sun Y., Wang X., Tang X. Deep Convolutional Network Cascade for Facial Point Detection // 2013 IEEE Conference on Computer Vision and Pattern Recognition. 2013. P. 3476–3483. DOI: 10.1109/CVPR.2013.446.
- [8] Brunton S.L., Noack B.R., Koumoutsakos P. Machine learning for fluid mechanics // Annual review of fluid mechanics. 2020. Vol. 52. P. 477-508. DOI: 10.1146/annurev-fluid-010719-060214.
- [9] Znamenskaya I.A. Methods for Panoramic Visualization and Digital Analysis of Thermophysical Flow Fields. A Review. // Scientific Visualization. 2021. Vol. 13(3). P. 125–158. DOI:10.26583/sv.13.3.13.
- [10] Particle Image Velocimetry Based on a Deep Learning Motion Estimator / S. Cai [et al.] // IEEE Transactions on Instrumentation and Measurement. 2020. Vol. 69. P. 3538–3554. DOI:10.1109/TIM.2019.2932649.
- [11] Three-dimensional particle tracking velocimetry using shallow neural network for real-time analysis / Y. Gim [et al.] // Experiments in Fluids. 2020. Vol. 61. P. 1–8. DOI:10.1007/s00348-019-2861-8.
- [12] FlowNet: Learning Optical Flow with Convolutional Networks / A. Dosovitskiy [et al.] // 2015
  IEEE International Conference on Computer Vision (ICCV). 2015. P. 2758–2766.
  DOI:10.1109/ICCV.2015.316.
- [13] Dense motion estimation of particle images via a convolutional neural network / S. Cai [et al.] // Experiments in Fluids. 2019. Vol. 60. DOI:10.1007/s00348-019-2717-2.
- [14] Ivanova Y.V., Poroykov A.Y. Estimation of the measurement error of photogrammetric techniques by controlled flexible deformable surface // 2019 International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE). IEEE. 2019. 4 p. DOI:10.1109/REEPE.2019.8708779.