



**МИНОБРНАУКИ РОССИИ**  
федеральное государственное бюджетное образовательное  
учреждение высшего образования  
**«Национальный исследовательский университет «МЭИ»**

Институт РЭ  
Кафедра Физики им.  
В.А. Фабриканта

**ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА**  
**(МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ)**

**Направление** 11.04.04 Электроника и наноэлектроника  
(код и наименование)

**Образовательная программа** Лазерная и оптическая измерительная  
электроника

**Форма обучения** очная  
(очная/очно-заочная/заочная)

**Тема:** Повышение точности и пространственного разрешения алгоритмов  
обработки для ближней фотограмметрии

**Студент** ЭР-03м-20 Пинчуков В.В.  
группа подпись фамилия и инициалы

**Руководитель**  
**ВКР** к.т.н. доцент Поройков А.Ю.  
уч. степень должность подпись фамилия и инициалы

**Консультант** уч. степень должность подпись фамилия и инициалы

**Внешний консультант** уч. степень должность подпись фамилия и инициалы

организация

**«Работа допущена к защите»**

**Заведующий кафедрой** д.т.н. доцент Скорнякова Н.М.  
уч. степень звание подпись фамилия и инициалы

**Дата**

**Москва, 2022**



**МИНОБРНАУКИ РОССИИ**  
федеральное государственное бюджетное образовательное  
учреждение высшего образования  
**«Национальный исследовательский университет «МЭИ»**

Институт РЭ  
Кафедра Физики им.  
В.А. Фабриканта

**ЗАДАНИЕ  
НА ВЫПУСКНУЮ КВАЛИФИКАЦИОННУЮ РАБОТУ  
(МАГИСТЕРСКУЮ ДИССЕРТАЦИЮ)**

<b>Направление</b>	11.04.04 Электроника и наноэлектроника (код и наименование)		
<b>Образовательная программа</b>	Лазерная и оптическая измерительная электроника		
<b>Форма обучения</b>	очная (очная/очно-заочная/заочная)		
<b>Тема:</b>	Повышение точности и пространственного разрешения алгоритмов обработки для ближней фотограмметрии		
<b>Студент</b>	ЭР-03м-20	подпись	Пинчуков В.В. фамилия и инициалы
<b>Руководитель ВКР</b>	к.т.н.	доцент	Поройков А.Ю. фамилия и инициалы
<b>Консультант</b>	уч. степень	должность	подпись фамилия и инициалы
<b>Внешний консультант</b>	уч. степень	должность	подпись фамилия и инициалы
<b>Заведующий кафедрой</b>	д.т.н.	доцент	Скорнякова Н.М. фамилия и инициалы
<b>Место выполнения работы</b>	ФГБОУ ВО «НИУ «МЭИ»		

## **1.Обоснование выбора темы выпускной квалификационной работы**

В настоящее время растёт интерес к методам измерения, которые могут предоставить глобальную информацию о динамике конструкций. Эта информация используется для фундаментальной физики, изучения особенностей конструкций и слежением за их динамикой, а также для проверки и обновления аналитических моделей. Для решения этой задачи используются оптические технологии, как методы бесконтактных измерений, который существенно развились за последние несколько десятилетий, благодаря исключительным достижениям в области компьютерных мощностей, объёмов памяти и фотодатчиков.

Оптические методы измерения пользуются спросом, особенно в строительстве, промышленности и авиастроении. Одним из широко используемых методов является метод ближней фотограмметрии. И как для любого метода измерения, для него всегда актуальным является вопрос повышения точности и пространственного разрешения.

**Научный руководитель**

**дата**

**Студент**

**дата**

## **2.Консультации по разделу**

---

---

---

---

**Подпись консультанта**

**дата**

## **3.Консультации по разделу**

---

---

---

---

**Подпись консультанта**

**дата**

#### **4. План выполнения выпускной квалификационной работы**

<b>№ п\п</b>	<b>Содержание разделов</b>	<b>Срок выполн- ния</b>	<b>Трудоём- кость, %</b>
I.	<p>Теоретическая часть</p> <p>1. Обоснование выбора метода оптических измерений формы поверхности.</p> <p>2. Разработка алгоритма рекалибровки на основе оптимизации параметров стереопары. Выбор метода оптимизации.</p> <p>3. Разработка алгоритма повышения точности и пространственного разрешения с помощью методов машинного обучения.</p>	09.02.22– 26.04.22	35
II.	<p>Экспериментальная часть</p> <p>1. Проверка работоспособности реализованного алгоритма оптимизации путем компьютерного моделирования процесса декалибровки.</p> <p>2. Проведение экспериментальных исследований работоспособности реализованного алгоритма оптимизации.</p> <p>3. Проведение экспериментальных исследований работоспособности нейросети.</p>	15.03.22– 08.05.22	30
III.	<p>Публикации</p> <p>Подготовка двух публикаций.</p>	22.03.22– 28.05.22	20
IV.	<p>Оформление диссертации</p> <p>Составление, оформление и редактирование диссертации.</p>	10.05.22– 15.06.22	15

## **5. Рекомендуемая литература**

1. Kirmse T. Recalibration of a stereoscopic camera system for in-flight wing deformation measurements // Measurement Science and Technology, vol. 27(5), 2016. 11 p.
2. Raffel M. et al. Particle image velocimetry // Handbuch Optische Messtechnik, 2018, P. 145-196.
3. Методы многомерной оптимизации: методические указания и задания к выполнению лабораторных работ по дисциплине «Методы оптимизации» для студентов направления «Прикладная математика»/ сост. Т. М. Попова. – Хабаровск : Изд-во Тихоокеан. гос. ун-та, 2012. – 44 с.
4. T. Hui, X. Tang, and C. Loy, “LiteFlowNet: A lightweight convolutional neural network for optical flow estimation,” in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 8981–8989.
5. Shengze C., et al., "Particle Image Velocimetry Based on a Deep Learning Motion Estimation", August 2019 IEEE Transactions on Instrumentation and Measurement pp (99):1-15.

## **АННОТАЦИЯ**

Отчёт содержит 113 страниц, 56 рисунков, 8 таблиц, 83 источника.  
**СТЕРЕОФОТОГРАММЕТРИЯ, ЗАДАЧА ОПТИМИЗАЦИИ,  
АЛГОРИТМЫ МАШИННОГО ОБУЧЕНИЯ, СВЁРТОЧНЫЕ НЕЙРОННЫЕ  
СЕТИ**

Целью работы является разработка алгоритмов повышения точности и пространственного разрешения для ближней фотограмметрии, которая используется для восстановления формы деформируемой поверхности. В нестационарных условиях эксплуатации камер под воздействием вибрационных и инерционных нагрузок возникают относительные смещения камер, что приводит к увеличению погрешности восстановления формы поверхности. Решить эту проблему предлагается методами оптимизации обобщённых целевых функций. Вторая проблема – это векторное поле низкой плотности, которое является следствием использования в этой области кросскорреляционного анализа. Для решения проблемы предлагается использовать существующие методы машинного обучения.

## **ABSTRACT**

The report contains 113 pages, 56 figures, 8 tables, 83 sources.  
**STEREOPHOTOGRAMMETRY, OPTIMIZATION PROBLEM, MACHINE  
LEARNING ALGORITHMS, CONVOLUTIONAL NEURAL NETWORKS**

The aim of the work is to develop algorithms for improving accuracy and spatial resolution for near photogrammetry, which is used to restore the shape of the deformable surface. In non-stationary operating conditions of the chambers, under the influence of vibration and inertial loads, relative displacements of the chambers occur, which leads to an increase in the error of restoring the shape of the surface. It is proposed to solve this problem by methods of optimization of generalized objective functions. The second problem is a low-density vector field, which is a consequence of the use of cross-correlation analysis in this field. To solve the problem, it is proposed to use existing machine learning methods.

## СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	9
1 Теоретическая часть.....	10
1.1 Обоснование выбора метода оптических измерений формы поверхности.....	10
1.1.1 Электронная интерферометрия спекл-структуры и шеарография .....	11
1.1.2 Лазерная доплеровская виброметрия.....	13
1.1.3 Фотограмметрия.....	14
1.1.4 Выбор метода оптических измерений формы поверхности.....	20
1.2 Разработка алгоритма рекалибровки на основе оптимизации параметров стереопары. Кросскорреляционный анализ .....	22
1.2.1 Модель и калибровка камер.....	22
1.2.2 Триангуляция, декалибровка и выбор метода оптимизации.....	27
1.2.3 Метод кроскорреляционного анализа .....	33
1.3 Разработка алгоритма повышения точности и пространственного разрешения с помощью методов машинного обучения.....	41
1.3.1 Обзор методов машинного обучения.....	41
1.3.2 Анализ и выбор алгоритма машинного обучения .....	52
2 Экспериментальная часть.....	69
2.1 Проверка работоспособности реализованного алгоритма оптимизации путем компьютерного моделирования декалибровки.....	69
2.1.1 Моделирование декалибровки.....	69
2.1.2 Оптимизация.....	72
2.2 Проведение экспериментальных исследований работоспособности реализованного алгоритма оптимизации.....	73
2.2.1 Лабораторная установка.....	73
2.2.2 Эксперимент .....	75

2.3 Проведение экспериментальных исследований работоспособности нейросети .....	79
2.3.1 Результаты применения нейросети .....	79
2.3.2 Замеченные особенности .....	81
2.4 Экспериментальное сравнение нейросетей с кросскорреляционным анализом .....	82
2.4.1 Сравнения в смежной области – PIV .....	82
2.4.2 Поиск равнозначных условий сравнения .....	88
2.4.3 Результаты сравнения .....	95
2.4.4 Предположения по ограничению сетей .....	99
2.4.5 Предложения по решению возникших проблем .....	101
ЗАКЛЮЧЕНИЕ .....	104
БИБЛИОГРАФИЧЕСКИЙ СПИСОК .....	106

## **ВВЕДЕНИЕ**

Работа посвящена повышению точности и пространственного разрешения в задаче восстановления трёхмерной формы поверхности. Рассматриваемым методом измерения является ближняя стереофотограмметрия, использующая как минимум две цифровые камеры для измерений.

Рассмотрены аналогичные методы измерения поверхности. Выполнено сравнение этих методов с фотограмметрией и обоснован выбор фотограмметрического метода. Затем обозначены проблемы точности этого метода. Предложены и реализованы пути решения рассматриваемых проблем.

Первая решаемая проблема – это декалибровка стереопары при внешних воздействиях. Эта проблема актуальна в рамках измерения формы крыла опытного летательного аппарата. Т.к. в полёте на камеры действуют вибрации и перегрузки, то могут возникать смещения одной камеры относительно другой, что приводит к увеличению погрешности восстановления поверхности крыла. Поэтому в работе рассмотрены критерии обнаружения декалибровки и предложен алгоритм уточнения внешних параметров камер.

Вторая проблема – это разреженное векторное поле, связанное с ограничениями метода на основе кросскорреляции. Кросскорреляционные алгоритмы хорошо зарекомендовали себя для поиска соответствий на изображениях во многих приложениях благодаря своей универсальности. Однако такие алгоритмы для достижения удовлетворительной скорости работы рассчитывают векторное поле смещения с некоторым шагом в пикселях изображения, что снижает пространственное разрешение восстанавливаемой трёхмерной поверхности в сравнении с возможным разрешением вектор на каждый пиксель, т.е. в полном разрешении. В последнее время стали интенсивно развиваться методы машинного обучения. Произведён обзор литературы по алгоритмам, которые начали применяться в задачах анализа оптического потока и по сообщениям исследователей имеют преимущества над уже известными алгоритмами.

## **1 Теоретическая часть**

### **1.1 Обоснование выбора метода оптических измерений формы поверхности**

В последнее время растёт интерес к методам измерения, которые могут предоставить глобальную информацию о динамике конструкций. Эта информация используется для фундаментальной физики, изучения особенностей конструкций и слежением за их динамикой, а также для проверки и обновления аналитических моделей. Подход к компьютерному моделированию, такой как метод конечных элементов (МКЭ), предоставляет наиболее полную информацию о всей конструкции. С другой стороны, обычные методы измерения (например, акселерометры или тензодатчики) для всех практических целей могут обеспечивать измерения только в нескольких локальных местах. Таким образом, существует серьёзное несоответствие между несколькими степенями свободы измерений для эмпирических данных по сравнению с потенциальными миллионами степенями свободы в модели МКЭ, что препятствует значимой и эффективной корреляции. Кроме того, обычные датчики могут вызывать массовую нагрузку и зачастую требуют проводки для передачи данных или питания. Таким образом, существует значительная потребность в бесконтактных измерениях, которые обеспечивали бы распределённую чувствительность, не добавляя при этом массы или жёсткости конструкции и не влияя на истинное динамическое движение путём добавления внешних датчиков.

Оптические технологии как методы бесконтактных измерений существенно развились за последние несколько десятилетий, благодаря исключительным достижениям в области компьютерных мощностей, объёмов памяти и фотодатчиков. Как правило, методы оптических измерений можно разделить на две группы: первая, подходы, использующие когерентное излучение, и вторая, подходы, использующие некогерентный свет. Лазерная доплеровская виброметрия (ЛДВ), электронная интерферометрия спекл-структуры (английский вариант – electronic speckle pattern interferometry – ESPI) и шеарография (английский вариант – digital speckle shearography – DSS)

используют когерентное излучение для отслеживания динамики структур. Вторая группа оптических методов, использующая некогерентный свет, называется системами на основе изображений. Фотограмметрия относится ко второй группе и использует изображение исследуемой конструкции, получаемые с помощью матричных фотоприёмников.

### **1.1.1 Электронная интерферометрия спекл-структуры и шеарография**

Спекл-структура создаётся интерференцией множества сферических фронтов, которые отражаются, когда шероховатая поверхность освещается когерентным светом. Разница в оптических путях, пройденных фронтами волн, приводит к интерференции, создавая таким образом зернистую пространственную картину (см. рисунок 1). Это явление наблюдается, когда шероховатость поверхности имеет отклонения, равные или превышающие длину волны освещдающего света. В ESPI спекл-эффект обычно нежелателен, так как ограничивает разрешение и точность измерений. Однако, информация которую несёт спекл, используется для измерения различных физических величин.

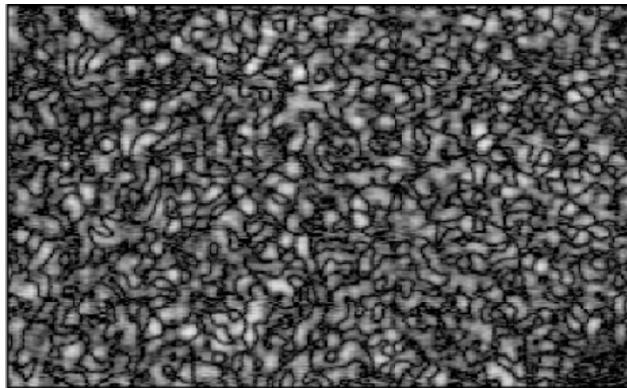


Рисунок 1 – Интерференционная картина, полученная с помощью интерферометра Маха-Цендера [1]

В настоящее время метод ESPI используется в измерениях деформации образца [2], обнаружения утечек в кегах из стали [3], в медицине [4]. Этот метод позволяет измерять поля смещений по всей поверхности объекта измерения.

Угол между источниками интерференционных волн определят вектор чувствительности. В зависимости от этого вектора интерферометры можно

разделить на два типа: первый, интерферометр с плоскостной чувствительностью и второй, интерферометр с внеплоскостной чувствительностью. В первом типе интерференционная картина создаётся двойным освещение объекта с целью измерения смещения в плоскости поверхности. Во втором типе интерферометров интерференционная картина создаётся путём объединения двух волн, одна из которых является опорной, а вторая исходит от поверхности объекта (объектный пучок). При наличии деформаций опорный луч является эталоном для волнового фронта, измененного объектом, и, таким образом, результатом являются интерференционные картины. Корреляция между этими картинами позволяет измерять смещения вне плоскости.

При внеплоскостных измерениях с помощью ESPI существует значительная разница в путях, пройденных опорными и объектными волнами. Это требует использования лазеров с высокой когерентностью, которые очень дороги. Кроме того, при статическом анализе необходима изоляция от внешних возмущений. Обычно это выполняется путём установки образца в среде с контролируемой вибрацией.

Некоторые из этих проблем можно преодолеть с помощью метода DSS. Это метод позволяет измерять градиенты смещений в выбранном направлении. DSS используется для определения повреждений в конструкциях [5], для измерения поверхности зеркально отражающих объектов [6]. До сих пор производятся исследования по улучшению качества и обработки получаемых данных с помощью DSS [7-9]. Метод DSS основан на тех же принципах, что и в ESPI, но теперь два волновых фронта исходят непосредственно от объекта. Эти два фронта смещаются в поперечном направлении друг относительно друга, образуя интерференционную картину.

ESPI метод позволяет измерять поля смещений, тогда как DSS позволяет измерять производные или градиенты этих полей. Поскольку они являются бесконтактными методами полного поля и высокого разрешения, то они подходят для измерения формы поверхности объекта.

### **1.1.2 Лазерная доплеровская виброметрия**

В основе метода лежит принцип оптического гетеродинного детектирования доплеровского сдвига волны, рассеянной от небольшой площади колеблющейся поверхности. Рассеянный свет собирается и коллимируется системой линз и смешивается с опорным пучком, частота которого сдвигается ячейкой Брэгга. Таким образом, система выдаёт аналоговое напряжение, пропорциональное мгновенной скорости вибрирующего объекта. Данный метод имеет широкий спектр применений для измерения частот и амплитуд колебаний. Диапазон частот простирается от статических отклонений до 100 кГц. Измеряемые смещения от 10 нм до 10 см [10].

Подобно другим методам интерферометрии, лазерный доплеровский виброметр сравнивает опорный луч и луч, отраженный от объекта. Тем не менее, ЛДВ измеряет частотный сдвиг, в то время как ESPI и DSS измеряют фазовый сдвиг между опорным и отраженными волнами пучка. Использование лазерной доплеровской виброметрии – один из широко используемых методов бесконтактных измерений вибрации, который нашёл применение в измерениях динамических напряжения твёрдых тел [11], анализе вибраций и воздействия водной среды на гребной винт [12,13] и медицине [14,15]. Этот метод измеряет одну точку на структуре за один раз. Сканирующий лазерный доплеровский виброметр использует зеркала для сканирования всей поверхности и обеспечения измерений по всему полю.

ЛДВ – это очень чувствительная система измерения с широким частотным диапазоном, которая измеряет высокочастотные колебания и может применяться к очень малым объектам. Однако существуют проблемы с использованием лазерного виброметра в полевых условиях, и эти измерения обычно требуют очень много времени. Для выполнения трехмерного измерения необходимы три лазерных виброметра. При сканировании нескольких точек ЛДВ последовательно сканирует каждую точку; поэтому, по сравнению с фотограмметрией и ESPI, ЛДВ может не иметь согласованности в измерениях,

поскольку физическое воздействие или поверхность могут изменяться во время измерения.

### 1.1.3 Фотограмметрия

Фотограмметрия – это метод измерения, который используется для определения геометрии, смещения и деформации конструкций с использованием фотографий или цифровых изображений. На заре своего существования фотограмметрия использовалась для воздушных и наземных целей, в значительной степени мотивированных военной разведкой. С развитием цифровых фотоаппаратов и возможностью покупать недорогие камеры с высоким разрешением приложения расширились и стали включать другие области науки, такие как гражданское строительство и машиностроение. Двумерная (2D) фотограмметрия использует одну фиксированную камеру и ограничивается измерениями деформаций и смещений в плоскости. Если тестовый объект изогнут, то возникают трехмерные (3D) деформации, двумерная фотограмметрия, как правило, больше не может обеспечить приемлемые результаты, поэтому появилась концепция стереокамер для трёхмерных измерений. В этом методе используют пару или серию изображений, записанных с помощью датчиков, таких как устройства с поверхностной зарядовой связью (ПЗС) и матрицы на комплементарном металл-оксидном полупроводнике (КМОП) для определения координат точек, паттернов и особенностей на изображении или для обнаружения границ объектов. Одна камера может использоваться для измерения смещения объектов при плоском движении. Однако 3D измерение должно выполняться с использованием стереокамер, относительное положение которых фиксируется во время теста, в то время как изображения записываются одновременно. Для трёхмерных измерений метод триангуляции используется для определения координат точек поверхности измеряемого объекта (см. рисунок 2).

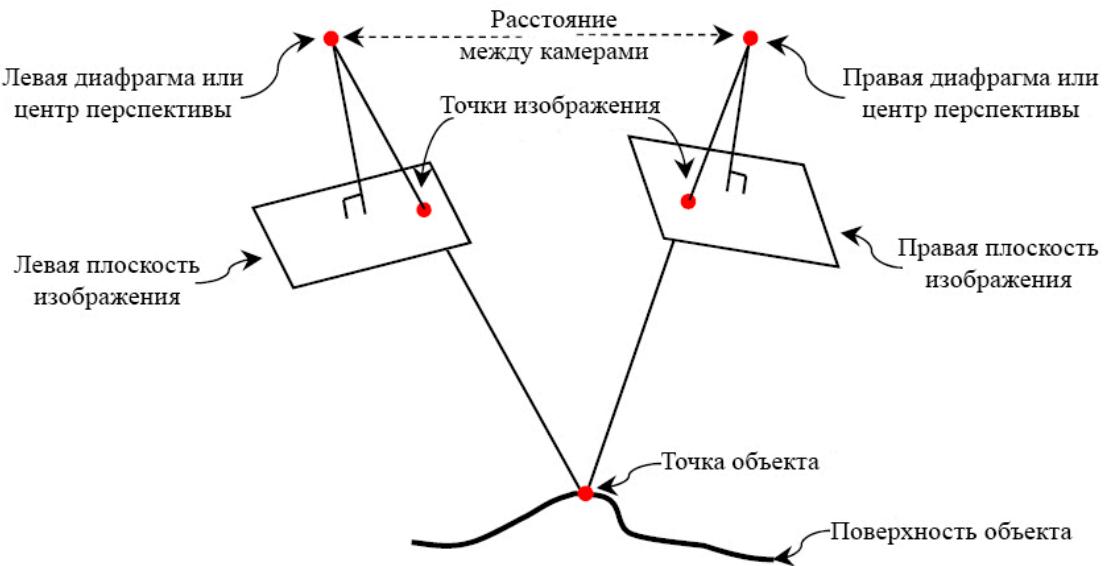


Рисунок 2 – Схема метода стереофотограмметрии, показывающая, как местоположение точки определяется с помощью триангуляции [16]

В общих чертах, метод фотограмметрии можно разделить на категории в зависимости от типа целей: отслеживание одиночных точек (английский вариант – Point Tracking – PT), корреляция цифровых изображений (английский вариант – Digital Image Correlation – DIC или Image Pattern Correlation Technique – IPCT, что является аналогом) или безмаркерный подход (английский вариант – Target-less approaches – TLA).

Фотограмметрия и корреляция цифровых изображений значительно развились за последние три десятилетия и были применены в измерениях деформаций крыльев самолётов, в условиях вибрационных нагрузок [17-21], вращающихся винтов самолёта [22] и вертолёта [23], деформаций поверхностей металлических конструкций [24-27], в деформации и моделировании гибких и мягких тел [28,29]. Совершенствование технологии камер и производство новых высокоскоростных камер в последнее время подтолкнуло исследователей начать использовать оптические подходы для измерения вибраций.

Немалую роль в развитии теоретической и математической базы сыграл метод анемометрии по изображению частиц (английский вариант – Particle Image Velocimetry – PIV). Этот метод позволяет по изображению одиночных или групп частиц определить их скорость, а, следовательно, скорость среды в которой они

находятся. В PIV также используется корреляция цифровых изображений [30] для определения смещений на фотографии.

В методике РТ используются камеры для определения координат отдельных точек, установленных на тестируемой конструкции. Для этого метода на конструкцию устанавливают группу оптических мишеней (обычно высококонтрастные круглые точки или световозвращающие мишени). Центры оптических целей находятся с использованием алгоритма поиска по эллипсу, а координаты точек в пространстве (для 3D) рассчитываются с помощью метода триангуляции. Трёхмерное отслеживание точек (3D РТ) определяет смещения целей путём отслеживания оптических мишеней на разных временных этапах и сравнения их координат с эталонным этапом. Обычно смещения объекта только в нескольких точках измеряется с помощью метода 3D РТ (см. рисунок 3).



Рисунок 3 – Фотография, показывающая оптические цели, установленные на лопасти ветровых турбин [16]

Корреляция цифровых изображений – это метод оптического измерения, который работает на основе вариации серого непрерывных изображений. Перед измерением необходимо подготовить область тестируемой конструкции. Подготовка поверхности обычно осуществляется путём нанесения на структуру высококонтрастного спекл-рисунка. Размер точек в этом методе обычно меньше, чем размер оптических мишеней, используемых в 3D РТ. После подготовки поверхности с помощью камер выполняется серия снимков. Метод ИРСТ работает при помощи корреляции перекрывающихся подмножеств (окон опроса) на всей интересующей области (см. рисунок 4). Размер окна опроса обычно составляет 15-30 пикселей, таким образом, оно включает в себя несколько точек узора (каждая точка обычно составляет 3-7 пикселей). Поскольку спекл-рисунок создаётся случайным образом, каждая область имеет уникальное значение интенсивности света, и программное обеспечение может распознать движение каждой уникальной подобласти узора на разных этапах. Смещение и деформации в интересующей области рассчитывается путём сопоставления каждого окна опроса с соответствующим окном на предыдущем этапе или восстановлением трёхмерной формы поверхности в каждый момент времени. Этот подход может обеспечить относительно непрерывные измерения на узорчатой области структуры и может называться методом измерения полного поля.

У безмаркерного подходов отсутствует оптическая мишень или шаблон для определения деформации структуры. Вместо этого используют внутренние особенности или края структуры, чтобы идентифицировать объект или область объекта, которую необходимо отслеживать. В этом методе используются алгоритмы обнаружения границ и сопоставления с эталоном. Алгоритм обнаружения краёв работает на основе того факта, что динамика движущегося объекта может быть обнаружена с использованием границ этого объекта. Метод сопоставления с эталоном позволяет обнаруживать измеряемый объект на основе эталонного изображения этого объекта.

Безмаркерный подход может быть не такими точными, как IPCT или РТ, но у него есть свои преимущества, когда нет возможности прикрепить оптические мишени или спектр-рисунок на поверхность объекта.

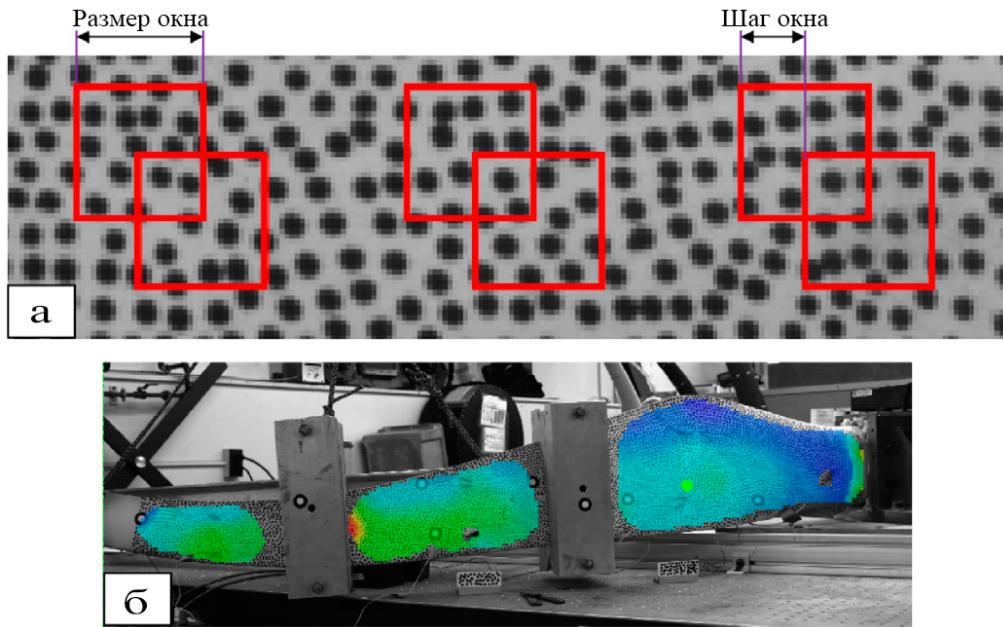


Рисунок 4 – (а) Фотография, показывающая размер окна опроса (подмножеств) и размер перекрытия; (б) фотография, показывающая узорчатую область лопасти ветряной турбины, на которую наложена измеренная деформация с использованием IPCT [16]

В таблице 1 [16] приведено сравнение различных методов оптических измерений и для полноты картины указан контактный метод измерения деформаций поверхности с помощью точечных датчиков, таких как акселерометры, тензодатчики и дифференциальный трансформатор для измерения линейных перемещений (английский вариант – Linear Variable Differential Transformer – LVDT).

Уровень шума фотограмметрических систем изменяется в зависимости от разрешения камер, поля зрения, тестируемого набора, угла наклона камеры и т.д. По данным работ [31, 32], внеплоскостная точность IPCT составляет приблизительно от 0,01 до 0,03 пикселя. Эти значения меняются в зависимости от размера окна опроса, способов фильтрации и тестирования.

Таблица 1 – Сравнение различных подходов к динамическим измерениям [16]

Характеристика	Точечные датчики (акселерометры, тензодатчики и LVDT)	Методы интерферометрии (ESPI, DSS)	Лазерный доплеровский виброметр	Фотограмметрия
Контактность измерений	Необходима проводка и может вызвать массовую нагрузку или изменение жесткости	Бесконтактный	Бесконтактный	Бесконтактный
Трудоёмкость получения передаточной функции	Легко получить передаточные функции	Требуется постобработка для получения передаточных функций	Легко получить передаточные функции	Требуется постобработка для получения передаточных функций
Размер области собираемых данных	Данные только в месте нахождения датчиков	Данные обо всей видимой поверхности объекта	Данные в заранее определенных точках на видимой поверхности объекта	Данные обо всей видимой поверхности объекта
Чувствительность метода	Чувствительность для каждой точки зависит от преобразователя	Чувствительность снижается по мере увеличения поля зрения	Чувствительность зависит от длины волны лазерного излучения.	Чувствительность снижается по мере увеличения поля зрения
Частота измеряемых смещений	Широкий частотный диапазон	Лучше для низкочастотных измерений	Широчайший частотный диапазон	Лучше для низкочастотных измерений
Чувствительность к калибровке и смещением всего объекта измерений	Не чувствителен к движениям объекта	Калибровка очень чувствительна (для ESPI) и чувствительна (для DSS) к изменениям условий установки и настройки.	Калибровка очень чувствительна к изменениям условий установки и движению объекта	Калибровка менее чувствительна к изменениям условий установки и движению объекта
Дороговизна	Недорого	Дорого	Очень дорогой	От умеренно до дорого в зависимости от камер
Сложность получения видов деформаций	Из-за проблем с чувствительностью очень сложно определить виды деформаций твердого тела и сложные деформации с использованием одного типа преобразователя.	Трудно извлекаемые виды деформаций. Применимо для более простых схем пространственной деформации и не применимо для очень гибких конструкций	Трудно извлекаемые виды деформаций. Применимо для более простых схем пространственной деформации и не применимо для очень гибких конструкций	Легко извлекаемые виды деформаций. Подходит для использования в очень гибкой конструкции. Очень точный для пространственно сложных деформаций
Скорость измерений	Быстрое измерение	Очень быстрое измерение	Очень трудоемкое измерение	Очень быстрое измерение
Пространственное разрешение	Низкое пространственное разрешение	Очень высокое пространственное разрешение	Высокое пространственное разрешение	Очень высокое пространственное разрешение (IPCT)

Продолжение таблицы 1

Характеристика	Точечные датчики (акселерометры, тензодатчики и LVDT)	Методы интерферометрии (ESPI, DSS)	Лазерный доплеровский виброметр	Фотограмметрия
Возможные типы измерений	В зависимости от типа датчика измеряет только смещение, ускорение или деформацию	Измеряет смещение и деформацию	Измеряет скорость; можно измерить напряжение, добавив дополнительное оборудование	Измеряет смещение и деформацию (IPCT)
Скорость обработки данных	Очень быстрая обработка данных	Быстрая обработка данных	Быстрая обработка данных	Длительная обработка данных (IPCT); быстрая обработка данных (PT)
Возможность измерения в реальном времени	Может быть в реальном времени	Может быть в реальном времени	Может быть в реальном времени	Обычно работает оффлайн
Применимость для вращающихся конструкций	Трудно использовать для вращающихся конструкций из-за проводки	Применимо для вращающихся конструкций	Трудно использовать для вращающихся конструкций	Подходит для вращающихся конструкций

Исследователи изучили влияние различных параметров на точность фотограмметрии в динамических условиях [33, 34]. В работе [33] показали, что при измерении IPCT спекл-рисунок оказывает меньшее влияние на точность системы по сравнению с углом наклона камеры, выдержкой и величиной измеренного смещения. В статье [34] провели систематическую оценку погрешности IPCT в общих динамических приложениях с помощью метода обработки изображений, который позволил оценить эффект движения в каждом изображении и, следовательно, неопределенность, связанную с IPCT на этих изображениях. Они также предложили метод предварительной обработки изображений [35] для улучшения IPCT в динамических приложениях. Тем самым показали, что предлагаемый метод может значительно снизить как ошибки смещения, так и неопределенность измерений.

#### 1.1.4 Выбор метода оптических измерений формы поверхности

В результате проведённого обзора можно заключить, что методы ESPI и DSS обеспечивают высокую точность измерений деформаций при вибрационных нагрузках, но имеют высокую чувствительность к юстировке

системы, которая дороже по сравнению с системой фотограмметрии. Также для этих методов сложно определить виды деформаций тела.

Метод ЛДВ позволяет производить измерения в очень широком частотном диапазоне вибраций. Но область измерений мала, что не позволяет единовременно получать информацию о всей конструкции, а, следовательно, приводит к несоответствию результатов измерений в разных областях объекта.

Метод фотограмметрии не обладает этим недостатком, хоть не подходит для измерений при высокочастотных вибрациях, но применим при низкочастотных. Этот метод позволяет легко определять виды деформаций тела и обладает точностью измерений сопоставимой с методами ESPI и DSS.

Исходя из полученной информации и возможностей лабораторного стенда выбор метода оптических измерений для определения деформаций поверхности падает на фотограмметрический метод на основе IPCT. В отличии от принципиально других подходов он не требует сложного и чувствительного к юстировке оптического устройства – интерферометра. При этом он обеспечивает высокую точность измерений, которая в основном ограничена характеристиками используемых камер и алгоритмами обработки изображений. Именно IPCT среди других фотограмметрических методов выбран потому, что этот метод позволяет измерять всё поле объекта в отличии от РТ. При этом по сравнению с TLA имеет более высокую точность, но требует спекл-рисунка на измеряемой конструкции.

Постоянно разрабатываются новые камеры с улучшенными техническими характеристиками, которые позволяют производить измерения с более высокой скоростью, лучшим разрешением, большей чувствительностью и меньшими затратами. Это делает фотограмметрию многообещающей альтернативой другим методам измерения.

## **1.2 Разработка алгоритма рекалибровки на основе оптимизации параметров стереопары. Кросскорреляционный анализ**

Проблема декалибровки параметров камер известна давно, но большинстве приложений, либо не требуются параметры камер, либо нет требования к высокой точности измерений. Поэтому погрешность переотображения объекта на изображении в 1 или 2 пикселя является для таких областей приемлемой. Но для достижения высокой точности восстановления трёхмерной формы поверхности объекта необходима субпиксельная точность в 0,1 пикселя и менее, что может быть достигнуто в лабораторных экспериментах и “тепличных” условиях, когда нет внешних воздействий на камеры.

В главе рассматривается проблема, связанная с использованием системы камер (стереопары) для измерения поверхности крыла испытуемого летательного аппарата прямо в небе [36-38]. Эта проблема – декалибровка внутренних и внешних параметров стереопары в результате вибрационных нагрузок. Для начала остановимся на декалибровки внешних параметров.

Первое, что будет необходимо сделать, это определится с моделью камеры и количеством изменяемых параметров. Затем выбрать метод первоначальной калибровки камер. Определить каким образом чувствуют внешние параметры в процессе восстановления трёхмерной формы поверхности, называемой триангуляцией. Далее выбрать критерий определения изменения внешних параметров и методику исправления (оптимизации) этих параметров. Также в конце этой главы будет описан алгоритм кросскорреляционной обработки изображений, который участвует в экспериментальной проверке разработанного алгоритма рекалибровки. Все описанные выше этапы необходимы для наиболее полного понимания проблемы и её последовательного анализа.

### **1.2.1 Модель и калибровка камер**

Камера выполняет линейные преобразования из трехмерного проективного пространства  $P^3$  в двумерное проективное пространство  $P^2$ , как показано на рисунке 5 [39]. Модель камеры — это математическое соотношение,

описывающее, что происходит, когда сцена преобразуется в изображение. Проективное преобразование в общем случае можно разложить на три более простых преобразования, что соответствует трем переходам между четырьмя системами координат.

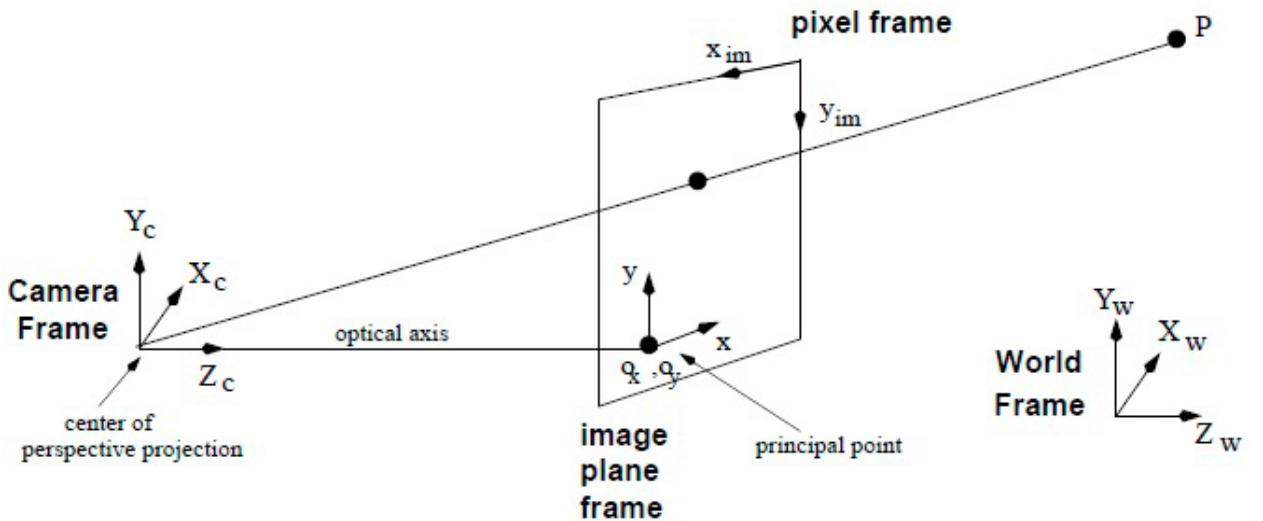


Рисунок 5 – Координаты камеры и мировые координаты [39]

Точка  $M = [x \ y \ z]^T$  в 3D пространстве представлена своими однородными координатами  $M = [x \ y \ z \ 1]^T$ . Точно так же её изображение на матрице камеры (или плоскости изображения),  $m = [x \ y]^T$ , представлено однородными координатами,  $m = [x \ y \ 1]^T$ . Основное свойство модели камеры состоит в том, что связь между мировыми координатами и координатами пикселя является линейно-проективной. Эта зависимость остается линейной независимо от выбора обеих координат [40].

Таким образом,  $M$  и  $m$  связаны соотношением

$$sm = \mathbf{P}M, \quad (1)$$

где  $s$ , масштабный коэффициент, называется глубиной;

$\mathbf{P}$  — это матрица  $3 \times 4$ , называемая матрицей перспективной проекции.

Поскольку  $s$  — это произвольная шкала, которая может принимать любое значение, мы будем просто обозначать уравнение (1) как

$$m \cong \mathbf{P}M. \quad (2)$$

Камера моделируется матрицей перспективной проекции (perspective projection matrix – PPM)  $\mathbf{P}$ , которую можно разложить с помощью QR-факторизации на произведение

$$\mathbf{P} = \mathbf{A}[\mathbf{R} \ \mathbf{T}], \quad (3)$$

где  $\mathbf{A}$  называется матрицей калибровки камеры, которая отображает нормализованные координаты изображения в координаты изображения на матрице камеры. Эта матрица имеет ряд элементов, называемых внутренними параметрами (камеры). Матрица  $\mathbf{A}$  определяется как

$$\mathbf{A} = \begin{bmatrix} f \cdot k_u & f \cdot k_u \cos \theta & u_0 \\ 0 & \frac{f \cdot k_v}{\sin \theta} & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (4)$$

где внутренние параметры:

- 1)  $f$  — фокусное расстояние камеры;
- 2)  $k_u$  и  $k_v$  — коэффициенты масштабирования по горизонтали и вертикали. Их инверсии определяют размер пикселя в мировой единице координат;
- 3)  $u_0$  и  $v_0$  — точка пересечения оптической оси с плоскостью матрицы камеры (главная точка камеры);
- 4)  $\theta$ , угол между осями матрицы камеры. Возможно использование не ортогональной пиксельной сетки, т. е. в общем случае может не равняться  $90^\circ$ .

Когда  $\theta = 90^\circ$ , матрица  $\mathbf{A}$  определяется как

$$\mathbf{A} = \begin{bmatrix} f \cdot k_u & 0 & u_0 \\ 0 & f \cdot k_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (5)$$

Поскольку мы не можем отделить фокусное расстояние камеры от коэффициентов масштабирования, у нас есть пять внутренних параметров  $a_u = f \cdot k_u$ ,  $a_v = f \cdot k_v$ ,  $u_0$ ,  $v_0$  и  $\theta$ , которые вносят вклад в косой элемент  $\gamma$  ( $\gamma = f \cdot k_u \cos \theta$ ).

$$\mathbf{A} = \begin{bmatrix} a_u & \gamma & u_0 \\ 0 & a_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (6)$$

Обратите внимание, что отношение  $r = a_u/a_v$  является соотношением сторон.

**D** представляет собой матрицу  $4 \times 4$ , определяемую как:

$$\mathbf{D} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix}, \quad (7)$$

где

- 1) **R** представляет собой матрицу вращения  $3 \times 3$  (имеющую 3 параметра). Определяемую как

$$\mathbf{R} = \begin{bmatrix} r_1^T \\ r_2^T \\ r_3^T \end{bmatrix}, \quad (8)$$

- 2) **T** представляет собой трехмерный вектор переноса (который имеет 3 параметра). Определяемый как,

$$\mathbf{T} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}. \quad (9)$$

Система координат правой камеры может быть преобразована в систему координат левой камеры посредством поворота **R** и переноса **T**. Матрица **D**

описывает изменение положения камеры (мировая система координат) и имеет 6 независимых параметров (3 параметра для  $\mathbf{R}$  и еще 3 параметра для  $\mathbf{T}$ ), называемых внешними параметрами.

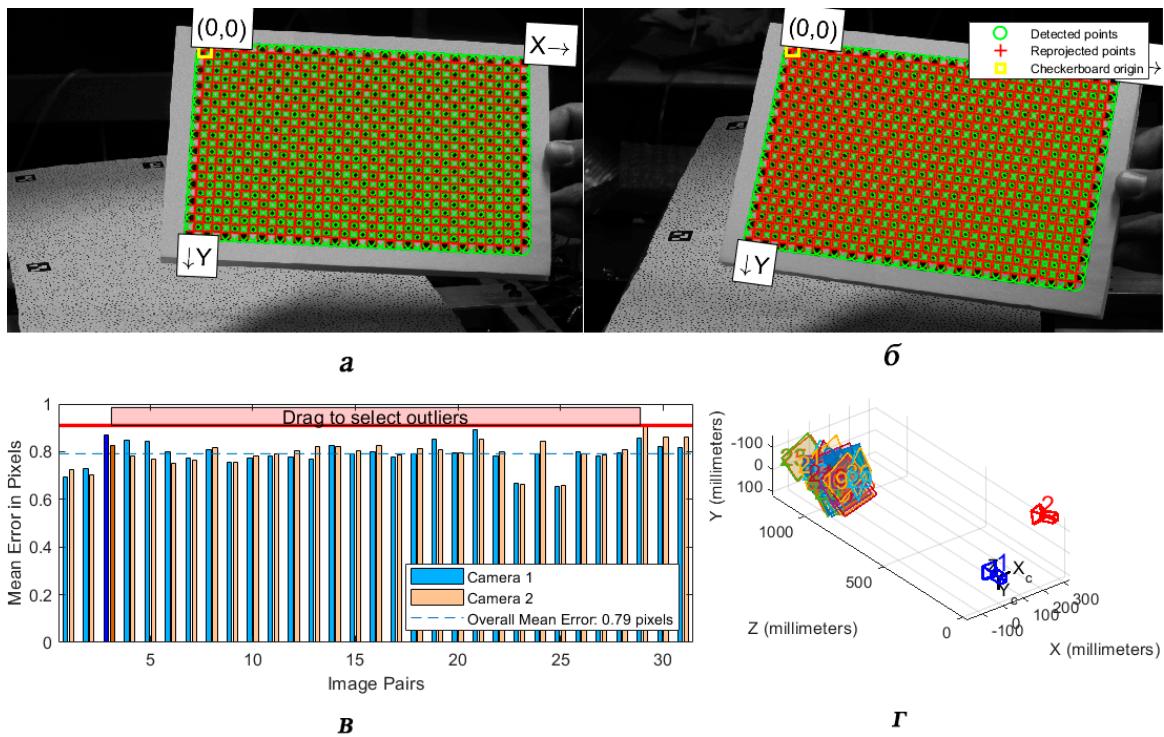
Основным методом калибровки является калибровка по шаблону [41]. Когда полностью контролируем процесс формирования изображения, лучший способ выполнить калибровку – это сделать несколько изображений объекта или шаблона известных размеров с разных точек зрения. Метод шахматной доски относится именно к этой категории.

Искомые 3D-точки – это углы клеток шахматной доски. Любой угол доски может быть выбран в качестве начала мировой системы координат. Оси  $x$  и  $y$  расположены вдоль доски, ось  $z$  перпендикуляра доске. Поэтому все точки на шахматной доске находятся в плоскости  $xy$ .

Клетки шахматной доски легко различимы на изображении и их легко обнаружить [42]. Более того, углы клеток шахматной доски идеально подходят для их локализации, поскольку имеют резкие градиенты в двух направлениях. Кроме того, углы находятся на пересечении линий шахматной доски. Все эти факторы определяют надёжность поиска углов клеток, расположенных на шахматной доске.

На рисунке 6 представлен пример работы калибровки стереосистемы. На рисунках 6(а) и 6(б) представлены результаты поиска узлов шахматной доски для левой и правой камеры соответственно, под 6(в) показана средняя ошибка репроекции для каждой пары изображений, под 6(г) визуализация взаимного расположения камер и измеренных шахматных досок.

В результате калибровки стереосистемы получаются матрицы внутренних параметром камер, матрица поворота и вектор смещения одной камеры относительно другой.



- а) – определение узлов шахматной доски с левой камеры; б) – определение узлов шахматной доски с правой камеры; в) – ошибки репроекции для каждой пары изображений; г) – визуализация взаимного расположения камер и измеренных калибровочных шахматных досок

Рисунок 6 – Этапы калибровки

### 1.2.2 Триангуляция, декалибровка и выбор метода оптимизации

Триангуляция в стереоанализе — это задача вычисления трехмерного положения точек на изображениях с учетом карты диспарантности и геометрии стереоустановки. Рассмотрим восстановление положения точки  $P(x, y, z)$  в реальном мире, которая на изображениях представляется точками  $O_l$  и  $O_r$  (сопряженная пара). Обозначим координаты через  $(x_l, y_l)$  и  $(x_r, y_r)$  для левого и правого изображений соответственно. Тогда учитывая подобие треугольников на рисунке 7 [39,40] можно записать

$$\frac{z}{f} = \frac{x}{x_l}, \frac{z}{f} = \frac{x-T}{x_r}, \frac{z}{f} = \frac{y}{y_l} = \frac{y}{y_r}. \quad (10)$$

Уравнения модели описывают взаимосвязь точек в трехмерной системе координат камеры с точками на двумерной действительной системе координат

изображения. После ряда подстановок и преобразований из уравнений (10) можно выразить неизвестные координаты  $x, y, z$  точки  $P$ .

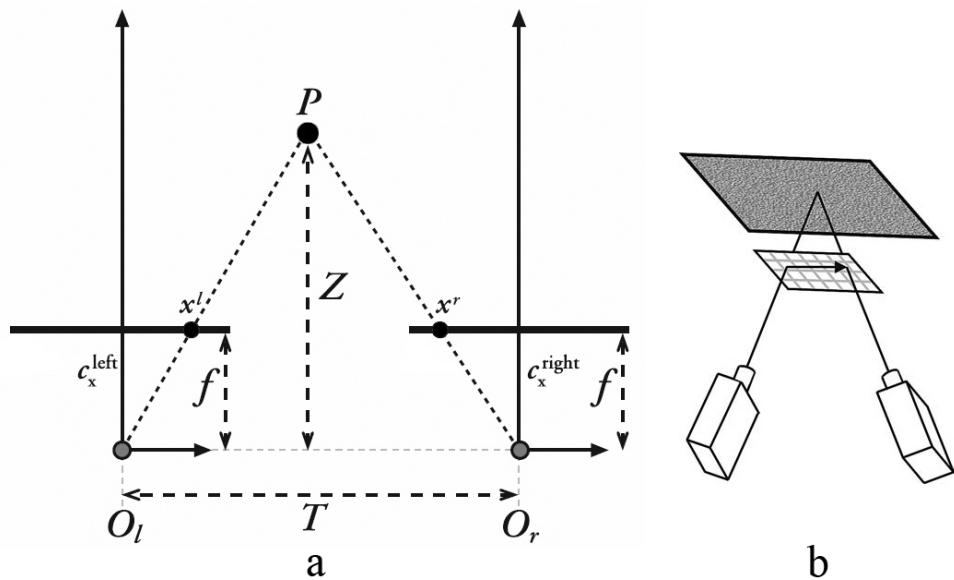


Рисунок 7 – Определение несоответствий точек и глубины [39,40]: а – схема наблюдения проекции одной точки на две камеры; б – визуальная схема вектора диспаранности

Необходимо учитывать переход от пикселей к линейным единицам. Таким образом, итоговые координаты опорных точек в трехмерном пространстве запишутся в таком виде [43]

$$\begin{aligned} X &= \frac{x_l T}{x_l - x_r}, \\ Y &= \frac{y_l T}{x_l - x_r}, \\ Z &= \frac{f \cdot T}{x_l - x_r}. \end{aligned} \quad (11)$$

Расстояние до объекта  $Z$  обратно пропорционально диспаритности ( $x_l - x_r$ ), причем расстояние до близких объектов можно измерить точнее, чем до удаленных. Заметим также, что диспаритность пропорциональна расстоянию между центрами линз  $T$ . Поэтому при фиксированной погрешности измерения диспаритности точность определения глубины будет возрастать с увеличением базы. К сожалению, чем дальше разнесены камеры, тем меньше похожи

изображения. Например, некоторые объекты, снятые одной камерой, даже могут быть невидимыми другой. Диспаратность пропорциональна также эффективному фокусному расстоянию  $f$ , поскольку изображение увеличивается с увеличением фокусного расстояния.

На практике две камеры, с помощью которых получаются стереопары, невозможно точно выставить, как полагалось при упрощенном рассмотрении [44]. Поэтому трудно достичь того, чтобы оптические оси сходились в одну точку на поверхности объекта. Поэтому соотношения и расчеты, приведённые в формулах (10) и (11), значительно усложняются и для выполнения триангуляции в общем виде необходимо знать, как вектор смещения, так и вектор поворота одной камеры относительно другой.

Стоит понимать, что декалибровка параметров камер в результате внешних воздействий является проблемой не для каждого приложения. Так, например, если камеры закреплены в помещении на статичной конструкции, то о них каких внешних воздействиях, кроме как усталости крепёжных элементов речи идти не может. Но существуют задачи, в которых камеры находятся на открытой местности [45], поэтому погодные условия влияют на параметры камер, или камеры могут быть закреплены на конструкции, которая испытывает физические нагрузки и вибрации. В случае испытания экспериментальный летательных аппаратов воздействует как минимум один из факторов – это не статичность конструкции, к которой крепятся камеры.

Обычно в работах по перекалибровки внешних параметров камер для расчёта новых параметров опираются на особые точки на изображениях [46, 47], калибровочную мишень [48] или на статичный фон [49] по смещениям камер относительно которых можно определить, как внешние параметры камер, так и оптимизировать внутренние параметры. Но для задачи, рассматриваемой в этой работе, подобный подход невозможен, в силу того, что, когда самолёт находится в небе нет ни статичного фона изображения из-за движения, ни статичных точек из-за деформаций крыла и фюзеляжа самолёта. Поэтому для критерия оптимизации внешних параметров камер используется ошибка репроекции

[20, 50], которая позволяет оценить точность внешних параметров камер в динамических условиях. Для лучшего понимания того, что из себя представляет ошибка репроекции приведён рисунок 8, на котором  $A_{Left}$  и  $A_{Right}$  матрицы внутренних параметров левой и правой камеры соответственно (см. формулу (6)), а  $R$  и  $T$  поворот и смещение одной камеры относительно другой.

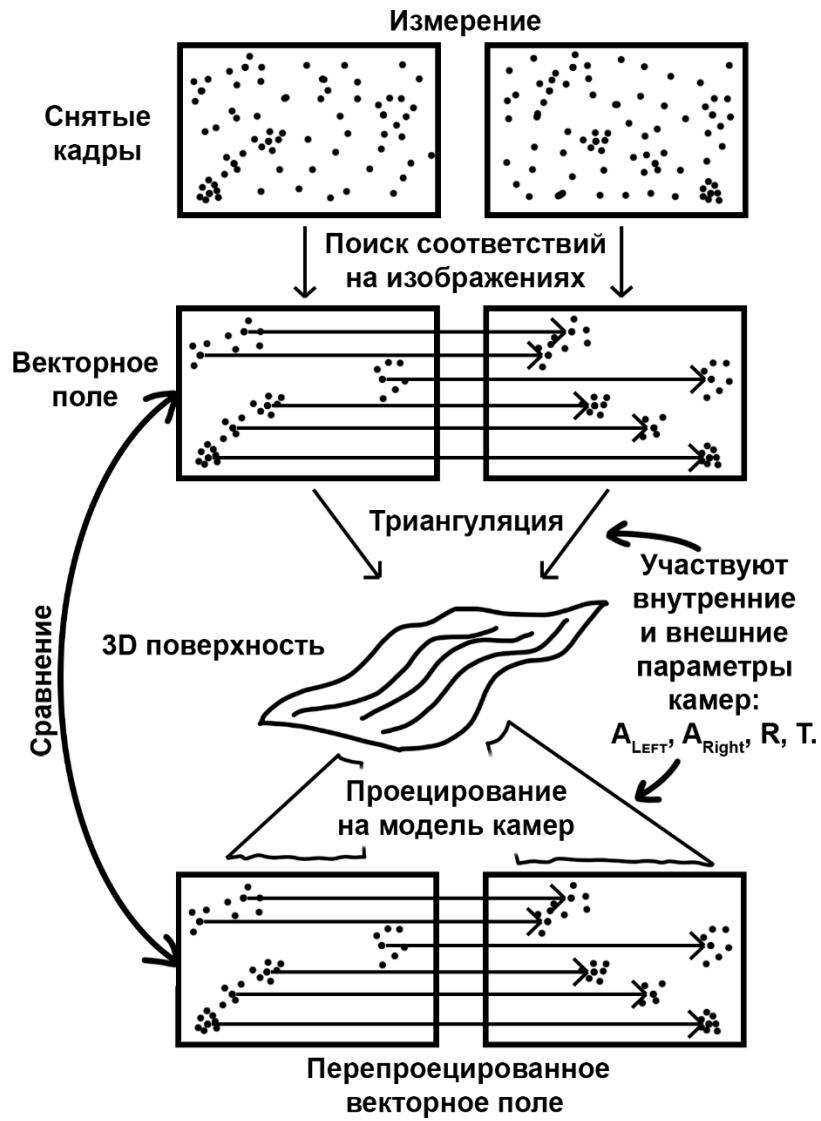


Рисунок 8 – Схема расчёта ошибки репроекции

Задача оптимизации множества параметров – это математическая задача по поиску глобального минимума функции многих переменных. Существуют множество методов использующие только значения функции в точках или первую производную функции или первую и вторую производную. Чем больше используется информации о минимизируем функции, тем быстрее сходимость

метода, но далеко не всегда тоже самое можно сказать о надёжности сходимости. Поэтому методы нулевого порядка, использующие только информацию о значении функции в точках, хоть и медленнее, но у них выше вероятность сходимости к глобальному минимуму.

В результате изучения литературы и компьютерного моделирования процесса оптимизации выбор пал на алгоритм безусловной оптимизации Нелдера-Мида совместно с методом внешних штрафных функций.

Алгоритм Нелдера-Мида основан на построении симплекса, который деформируется на каждой итерации поиска. Вычисляя значения целевой функции в вершинах симплекса, получаем информацию о характере изменения этой функции в области расположения симплекса. Поиск точки минимума целевой функции с помощью симплексов производится следующим образом:

- 1) На каждой итерации происходит вычисление целевой функции во всех точках симплекса и их упорядочивание по возрастанию значений.
- 2) Затем осуществляется последовательная попытка построить новые симплексы с лучшими значениями целевых функций, путём отражения точек с худшими значениями.
- 3) Если последовательная попытка отражения двух худших вершин оканчивается неудачей, то производится сжатие симплекса к точке с наименьшим значением и осуществляется новая итерация.
- 4) Поиск завершается, как только разность между значениями функции в точках симплекса становится достаточно малой.

Достоинства метода: простота, малое количество заранее установленных параметров, простая стратегия поиска, вычисления только значений функции, небольшой объём требуемой памяти.

Недостатки метода: метод работает эффективно при размерности целевой функции  $N \leq 6$ .

Условие  $N \leq 6$  соответствует данной задаче, т.к. в процессе оптимизации целевая функция зависит от вектора смещения с тремя элементами

соответствующим осям  $x$ ,  $y$ ,  $z$  и 3-х углов поворота одной камеры относительно другой. Метод Нелдера-Мида является одним из самых эффективных при  $N \leq 6$ .

Метод штрафных функций необходим для ограничения области поиска, путём резкого увеличения значения минимизируемой целевой функции за пределами допустимой области. Этот метод применяется для решения задач оптимизации в общей постановке, т.е. при наличии как ограничений-неравенств, так и ограничений-равенств. В рассматриваемом методе штрафные функции выбирают такими, что их значения равны нулю внутри и на границе допустимой области, а вне её – положительны и возрастают тем больше, чем больше нарушается ограничения. Таким образом «штрафуется» удаление от допустимой области.

Подобная комбинация методов показывает лучшую и более устойчивую сходимость, чем алгоритмы условной оптимизации, например, как комплексный метод Бокса и т.п.

Точностную характеристику алгоритма напрямую определить невозможно, т.к. она зависит от множества факторов, но возможно оценить по достигаемым значениям целевой функции в результате оптимизации. Например, на этапе калибровки средняя ошибка репроекции равна 0,75 пикселя, это означает, что среднее отклонение исходных 2-х мерных координат от перепроектированные по двум осям  $x$  и  $y$  составляет 0,75 пикселя. В результате оптимизации вероятно достижение значений ошибки репроекции таких же как на этапе калибровке и даже меньших значений, но не всегда.

Пределы применимости алгоритма ограничиваются только максимальным заданным числом итераций в одной попытке и числом попыток для поиска лучшего решения. Различные попытки поиска решения отличаются друг от друга только центральной точной построения симплекса, т.е. область целевой функции из которой начинается поиск. Это необходимо для того, чтобы исключить неудачные попытки поиска, когда получаемое решение удалено от глобального минимума целевой функции.

## 1.2.3 Метод кросскорреляционного анализа

### 1.2.3.1 Однопроходная кросскорреляция

Сигнал для обработки состоит из пары изображений, на каждом из которых имеется набор измеряемых точек (см. рисунок 9).

Локальный поиск наилучшего совпадения между изображениями достигается путём использования дискретной функции взаимной корреляции [38]

$$R_{II}(x, y) = \sum_{i=-M}^M \sum_{j=-N}^N I(i, j)I'(x + i, y + j). \quad (11)$$

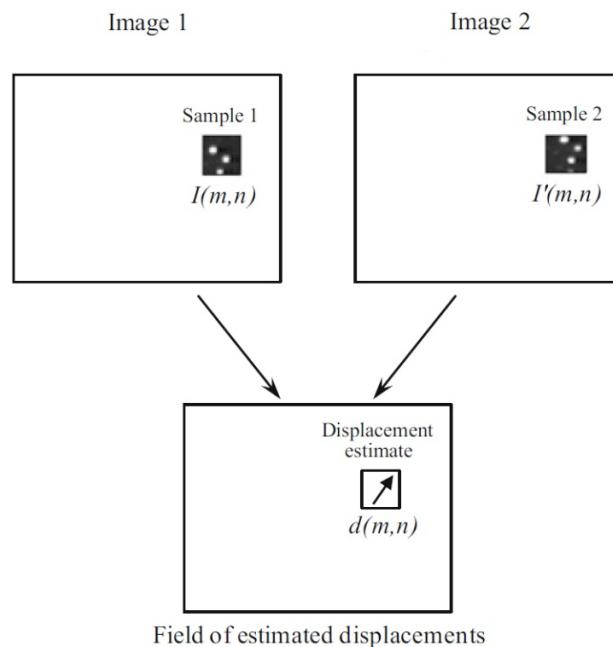


Рисунок 9 – Концептуальная схема окон опроса и определения смещения [43]

Переменные  $I$  и  $I'$  являются выборками (например, значениями интенсивности), извлеченными из изображений. Для каждого сдвига выборки  $(x, y)$  сумма произведений всех перекрывающихся интенсивностей пикселей дает одно значение взаимной корреляции  $R_{II}(x, y)$ . Применяя эту операцию для диапазона сдвигов ( $-M \leq x \leq +M, -N \leq y \leq +N$ ), формируется плоскость корреляции размером  $(2M + 1) \times (2N + 1)$ . Для значений сдвига, при которых изображения частиц образцов совпадают друг с другом, сумма произведений интенсивностей

пикселей будет больше, чем где-либо еще, что приведет к высокому значению взаимной корреляции.

При изучении этой прямой реализации функции взаимной корреляции становятся очевидными две вещи: во-первых, количество умножений на значение корреляции увеличивается пропорционально размеру окна опроса (или выборки), а во-вторых, метод взаимной корреляции по своей сути восстанавливает только линейные сдвиги. Этот метод первого порядка не может измерить никаких поворотов или деформаций. Следовательно, взаимная корреляция между двумя образцами изображений частиц даст только вектор смещения первого порядка, то есть средний линейный сдвиг частиц в пределах окна запроса. Это означает, что размер окна опроса следует выбирать достаточно малым, чтобы можно было пренебречь эффектами высшего порядка.

Альтернатива вычислению взаимной корреляции напрямую с использованием уравнения (12) заключается в использовании теоремы корреляции, которая утверждает, что взаимная корреляция двух функций эквивалентна комплексно-сопряженному умножению их преобразований Фурье (см. рисунок 10)

$$R_{II} = \hat{I} \cdot \hat{I}'^*, \quad (12)$$

где  $\hat{I}$  и  $\hat{I}'$  - преобразования Фурье функций  $I$  и  $I'$  соответственно. На практике преобразование Фурье эффективно реализуется для дискретных данных с использованием быстрого преобразования Фурье или БПФ, что сокращает количество вычислений, требуемых  $O[N^2]$  операций до  $O[N \log_2 N]$  операций.

Использование двумерного БПФ для вычисления плоскости взаимной корреляции имеет ряд свойств, эффекты которых необходимо учитывать.

Фиксированные размеры выборки: наиболее распространенная реализация БПФ требует, чтобы входные данные имели размерность по основанию 2 (то есть выборки  $16 \times 16$  пикселей или  $32 \times 32$  пикселей). По причинам, объясненным ниже,

обычно невозможно просто заполнить образец нулями, чтобы сделать его образцом размера по основанию 2.

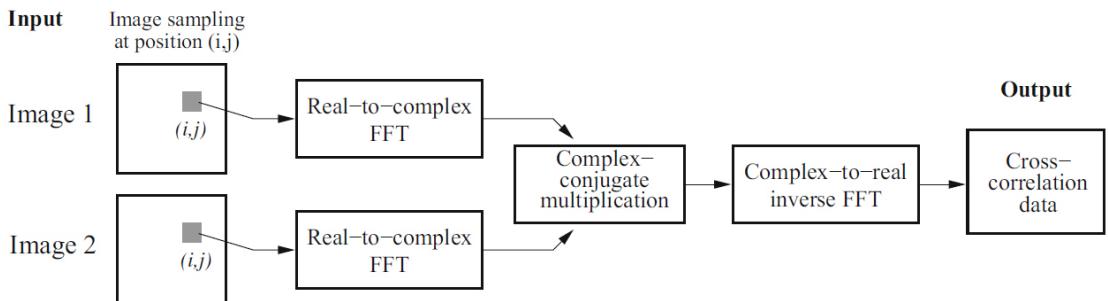


Рисунок 10 – Реализация взаимной корреляции с использованием быстрых преобразований Фурье [44]

Периодичность данных: по определению преобразование Фурье представляет собой интеграл (или сумму) по области от отрицательной бесконечности до положительной бесконечности. Однако на практике интегралы (или суммы) вычисляются по конечным областям, что оправдано предположением, что данные являются периодическими, то есть сигнал (то есть выборка изображения) постоянно повторяется во всех направлениях. Хотя для спектральной оценки существует множество методов для работы с ассоциированными артефактами, например, использование окон, их использование при вычислении взаимной корреляции внесет систематические ошибки или даже скроет сигнал корреляции в шуме.

Один из этих методов, заполнение нулями, который влечет за собой увеличение размера выборки по сравнению с исходным размером путем заполнения нулями, будет работать плохо, потому что данные (то есть образец изображения) обычно состоят из ненулевого (шумового) фона, на котором сигнал (т.е. изображения частиц) накладывается. Неоднородность границ, возникающая в процессе заполнения нулями, загрязняет спектры данных высокочастотным шумом, который, в свою очередь, ухудшает сигнал взаимной корреляции. Немного более продвинутая технология обработки данных с помощью окна БПФ устраняет эффекты неоднородности края, но приводит к

неравномерному взвешиванию данных в плоскости корреляции и смещению восстановленного вектора смещения.

Сглаживание: поскольку предполагается, что входные наборы данных для алгоритма корреляции на основе БПФ являются периодическими, сами данные корреляции также являются периодическими. Если данные длины  $N$  содержат сигнал (т.е. смещения), превышающий половину размера выборки  $N/2$ , то пик корреляции будет свернут обратно в план корреляции, чтобы появиться на противоположной стороне. Для смещения  $d_{x,true} > N/2$ , измеренное значение будет  $d_{x,Meas.} = d_{x,true} - N$ .

Ограничение диапазона смещения: Как упоминалось ранее, размер образца  $N$  ограничивает диапазон максимального извлекаемого смещения до  $\pm N/2$ . Однако на практике мощность сигнала корреляционного пика будет уменьшаться с увеличением смещения из-за пропорционального уменьшения возможных совпадений частиц. Ранее в литературе сообщалось, что значение  $N/3$  является адекватным пределом возможности восстановления вектора смещения. Более консервативным, но широко используемым пределом является  $N/4$ , иногда называемое правилом одной четверти. Однако при использовании методов итеративной оценки с методами сдвига окна эти требования являются существенными только для первого прохода и устаревают для всех других проходов.

Ошибка смещения: еще один побочный эффект периодичности корреляционных данных состоит в смещении оценок корреляции. При увеличении сдвигов меньше данных фактически коррелирует друг с другом, поскольку периодически продолжающиеся данные шаблона корреляции не вносят вклада в фактическое значение корреляции. Значения на краю плоскости корреляции вычисляются только из перекрывающейся половины данных и должны быть соответственно взвешены. Если значения корреляции не взвешены соответствующим образом, оценка смещения будет смещена к более низкому значению. Эта ошибка уменьшается с увеличением размера выборки. Более

крупные изображения частиц и вместе с ними более широкие корреляционные пики связаны с большими ошибками смещения.

Если все вышеперечисленные точки обработаны должным образом, алгоритм запроса на основе БПФ, показанный на рисунке 10, надежно предоставит необходимые данные корреляции, из которых можно будет извлечь карту смещений.

### 1.2.3.2 Субпиксельная точность

Одной из наиболее важных, но не обязательно легко понимаемых особенностей цифровой оценки является то, что положение пика корреляции можно оценить с точностью до субпикселей. Точность оценки порядка  $1/10 - 1/20$ -го пикселя реалистична для выборок размером  $32 \times 32$  пикселя из 8-битных цифровых изображений. Поскольку сами входные данные дискретизированы, значения корреляции существуют только для интегральных сдвигов. Тогда наивысшее значение корреляции позволит определить смещение только с погрешностью  $\pm 1/2$  пикселя. Однако, поскольку функция взаимной корреляции является статистической мерой наилучшего совпадения, сами значения корреляции также содержат полезную информацию.

Надежным методом является описание данных корреляции некоторой функцией. Особенno для узких пиков корреляции широкое распространение получил подход с использованием только трех соседних значений для оценки компонента смещения. Причем наиболее часто применяется аппроксимация пикового значения Гаусса. Разумным объяснением этого является то, что сами изображения частиц, если они правильно сфокусированы, описывают функции интенсивности Эйри, которые очень хорошо аппроксимируются гауссовым распределением интенсивности. Можно показать, что корреляция между двумя гауссовскими функциями также приводит к гауссовой функции. Трехточечные оценки обычно лучше всего подходят для довольно узких корреляционных пиков, сформированных из изображений частиц в диапазоне диаметров 2–3 пикселя.

Координаты пика с субпиксельной точностью  $(x_0, y_0)$  по гауссу определяются по формуле (14)

$$x_0 = i + \frac{\ln R_{(i-1,j)} - \ln R_{(i+1,j)}}{2 \ln R_{(i-1,j)} - 4 \ln R_{(i,j)} + 2 \ln R_{(i+1,j)}};$$

$$y_0 = j + \frac{\ln R_{(i,j-1)} - \ln R_{(i,j+1)}}{2 \ln R_{(i,j-1)} - 4 \ln R_{(i,j)} + 2 \ln R_{(i,j+1)}}, \quad (14)$$

где  $R$  – кросскорреляционная матрица;

$R_{(i,j)}$  – максимум этой матрицы.

### 1.2.3.3 Поиск и интерполяция выбросов

При поиске вектором смещения очень часто возникают, так называемые ошибочные вектора – выбросы, которое появляются из-за статистических флюктуаций. Поэтому крайне важно, чтобы эти векторы были удалены и заменены до того, как будут выполнены дальнейшие шаги в анализе. Типичными особенностями ошибочных векторов смещения, некоторые из которых можно наблюдать на рисунке 11(а), являются следующие:

- их величина и направление значительно отличаются от окружающих их соседей;
- во многих случаях они выглядят как изолированные ложные векторы;
- они также могут быть сгруппированы в группы, иногда указывающие в одном направлении;
- они очень часто появляются на краях поля данных (около поверхности модели, на краях выпадающих областей, на краях освещенной области, в областях повышенного движения вне плоскости или там, где присутствует сильный градиент).

Основным методом поиска выбросов является нормализованный медианный тест. Небольшая модификация медианного теста приводит к очень мощной схеме проверки ложных векторов скорости. Вестервил и Скарано [51] продемонстрировали, что нормализация стандартного медианного теста дает

довольно универсальную функцию плотности вероятности для невязки, так что одно пороговое значение может применяться для эффективного обнаружения паразитных векторов. Нормализация требует, чтобы невязка  $r_i$ , определяемая как:  $r_i = |U_i - U_{med}|$  сначала определяется для каждого окружающего вектора  $\{U_i|i = 1, \dots, 8\}$ . Затем определяется медиана этих восьми остатков, среднеквадратичная, и используется для нормализации стандартного теста медианы следующим образом

$$\frac{|U_{med} - U_{i,j}|}{r_{med} + \varepsilon_0} < \varepsilon_{thresh} \quad (13)$$

Дополнительный член  $\varepsilon_0$  требуется для учета оставшихся колебаний, полученных в результате корреляционного анализа. На практике это значение должно быть установлено около 0,1-0,2 пикселя, что соответствует среднему уровню шума данных [52].

После поиска ложных векторов необходимо их заменить на основе значений ближайших соседей. Для этой задачи лучше всего подходит метод билинейной интерполяции, который выполняется в несколько итераций. Итеративный подход необходим для того, что значение интерполируемого вектора приближалось к среднему значению некоторой области, а не к среднему двух соседей, что позволяет с высоким качеством интерполировать целые области неверный векторов.

#### **1.2.3.4 Многопроходная кросскорреляция**

Объем данных в процессе запроса может быть значительно увеличен путем использования смещения окна, равного локальному целочисленному смещению во втором проходе запроса. Путем смещения окон опроса в соответствии со средним смещением доля согласованных изображений частиц по сравнению с несогласованными изображениями частиц увеличивается, тем самым увеличивая отношение сигнал/шум пика корреляции. Кроме того, шум

измерения или погрешность в смещении значительно уменьшается, когда смещение изображения частицы меньше половины пикселя, где оно масштабируется пропорционально смещению. Смещение окна запроса может быть относительно легко реализовано. Алгоритм выглядит следующим образом:

Шаг 1: Выполните стандартный цифровой запрос со смещением окна запроса, близким к среднему смещению данных.

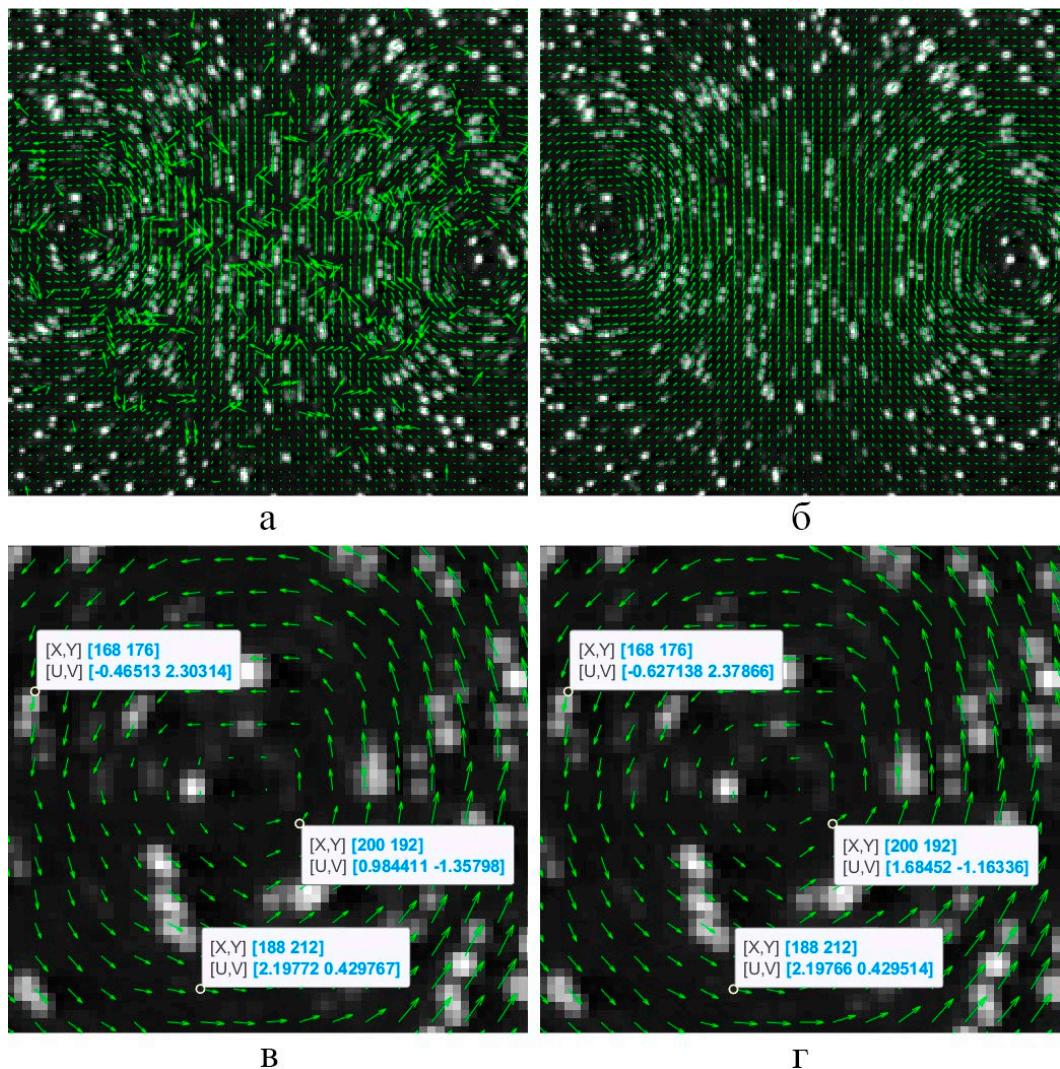
Шаг 2: Просканируйте данные на наличие выбросов, используя предопределенный критерий проверки, как описано выше. Замените данные выброса интерполяцией из допустимых соседей.

Шаг 3: Используйте оценки смещения для локальной корректировки смещения окна запроса до ближайшего целого числа.

Шаг 4: Повторяйте опрос до тех пор, пока целочисленные векторы смещения не сойдутся до  $\pm 1$  пикселя. Обычно требуется три прохода.

Скорость этого многопроходного запроса может быть значительно увеличена путем сравнения нового целочисленного смещения окна с предыдущим значением, что позволяет пропустить ненужные вычисления корреляции. Выход данных может быть дополнительно увеличен за счет ограничения области поиска пика корреляции на последнем проходе запроса.

На рисунке 11 представлены результаты работы, описанных выше алгоритмов. На рисунке 11(а) видно, что встречается множество векторов, который значительно отличаются от соседей, и это наблюдается на всей области корреляции. Уже на рисунке 11(б) произведён поиск и замена выбросов нормализованным средним с билинейной интерполяцией, что заметно улучшило картину структуры векторного поля. На рисунках 11(в) и 11(г) приведена разница между однопроходным и многопроходным запросом. Не стоит забывать про правило одной четверти, описанного выше, которое накладывает ограничение на размер окна запроса, поэтому для многопроходного опроса количество итераций обычно около трёх.



- а) – однопроходный запрос без поиска выбросов; б) – однопроходный запрос с поиском и интерполяцией выбросов; в) – однопроходный запрос с субпиксельной точностью;  
г) – многопроходный запрос с субпиксельной точностью

Рисунок 11 – Результаты работы реализованной программы кросскорреляции

Есть и множество других примеров, поэтому для задачи повышения точности фотограмметрических измерений необходимо провести обзор методов машинного обучения и проанализировать их возможности в исследуемой области.

### **1.3 Разработка алгоритма повышения точности и пространственного разрешения с помощью методов машинного обучения**

#### **1.3.1 Обзор методов машинного обучения**

Машинное обучение является классом методов такой области, как искусственный интеллект (см. рисунок 12). Главным отличием этих методов от

других является не прямое решение задачи на основе строго заданных уравнений или соотношений, а обучение за счёт нахождения решения множества однотипных задач. Машинное обучение активно развивается в настоящее время, т.к. позволяет решать с большей скоростью уже известные проблемы, для которых есть решение, так и справляться с теми задачами, которые считалось невозможно решить.

Уже сейчас методы машинного обучения используются в различных областях жизни. Например, в поисковых системах для быстрого поиска и выдачи запроса используются ансамблевые методы, т.к. они отличаются высокой точностью и скоростью работы. Или системы распознавания лиц, в частности в программном обеспечении камер смартфонов. Ещё интересным применением является в области обработки видео для компьютерной индустрии, где используют нейросеть для того, чтобы на основе нескольких последовательных кадров низкого разрешения получить кадры более высокого разрешения, что снижает нагрузку на систему и позволяет играть с повышенным разрешением картинки при снижении требований к компьютеру.



Рисунок 12 – Иллюстрация областей искусственного интеллекта [53]

Машинное обучение – это методы, позволяющие получить результат по входному набору данных без явного программирования логики обработки. Машинное обучение является часть такой области, как искусственный интеллект, что проиллюстрировано на рисунке 12. Нейросети всего лишь

область машинного обучения, а глубокое обучение – архитектура нейросетей, один из подходов к их построению и обучению.

Методы машинного обучения делятся по типам решаемых задач:

- Классификация – определение категории объекта на основе его признаков.
- Кластеризация – разделение объектов на категории по объединяющему признаку.
- Регрессия – прогнозирование на основе выборки объектов с различными признаками.
- Уменьшение размерности – является частным случаем кластеризации, направленным на уменьшение числа признаков, описывающих объект.
- Идентификация – исторически возникла из задачи классификации, когда вместо определения класса объекта потребовалось уметь определять, обладает ли объект требуемым свойством или нет.
- Поиск выбросов – нахождение объектов, которые отличаются по своим свойствам от объектов обучающей выборки.
- Поиск правил (ассоциация) – это анализ потока данных с целью выявления закономерностей (правил).

Это основные типы решаемых задач, но существуют и другие. Также машинное обучение разделяется по типу обучения:

- Обучение с учителем – в этом случае подготовленная выборка содержит правильный ответ, а алгоритму необходимо выявить зависимость, то есть определить, как получить верный ответ.
- Обучение без учителя – здесь уже отсутствуют верные ответы, имеется только массив данных, который необходимо разделить на группы по каким-либо признакам или обработать так, чтобы минимизировать какой-либо параметр, который может косвенно связан с правильностью ответа.
- Обучение с подкреплением – состоит приспособлении (адаптации) алгоритма к среде, таким образом пытаясь найти оптимальные действия, имеющие последствия, находясь в наборе различных сценариев.

Обучения с учителем и без учителя относятся к классическому обучению.

В отдельный класс выделяют ансамблевые методы, так как эти методы используют одновременно два и более алгоритмов машинного обучения для построения более точных результатов.

На рисунке 13 показаны основные направления машинного обучения. Рассмотрим области применения этих направлений:

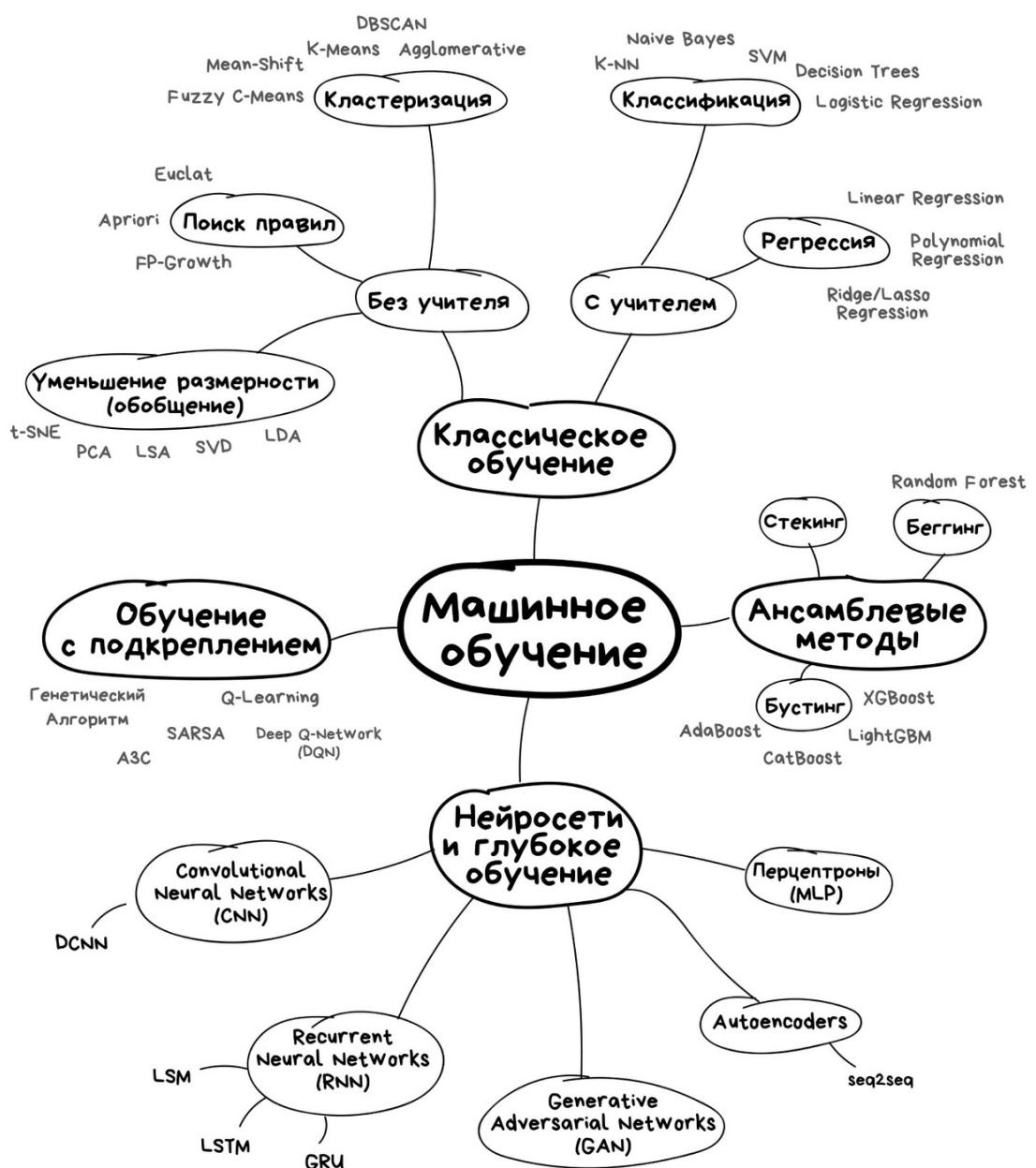


Рисунок 13 – Основные направления машинного обучения [53]

### Регрессия

- анализ спроса, объёма продаж,
- прогноз стоимости товаров,
- любые зависимости анализируемого параметра от времени;

### Классификация

- поиск схожих файлов,
- фильтрация данных,
- распознавание рукописных символов;

### Кластеризация

- сжатие файлов и изображений,
- анализ данных,
- определение аномалий,
- сегментация данных;

### Поиск правил

- анализ поведения пользователей на сайтах,
- расстановка объектов,
- прогноз акций;

### Уменьшение размерности

- визуализация многомерных данных,
- поиск аномалий на изображении,
- системы рекомендаций,
- поиск схожих файлов;

### Обучение с подкреплением

- беспилотные автомобили,
- самоуправляемые роботы,
- управление ресурсами компаний,
- игры;

### Ансамблевые методы

- поисковые системы,

- распознавание объектов,
- компьютерное зрение;

### Нейросети

- определение объектов на фото и видео,
- распознавание и синтез речи,
- обработка изображений,
- машинный перевод.

Из вышеперечисленных областей применения видно, что для задачи определения деформаций объекта на изображении и видео, что является частью области анализа оптического потока, подходят ансамблевые методы и нейронные сети.

Ансамблевые методы имеют лучшую точность среди методов машинного обучения, но на данный момент нет целенаправленных исследований применения этих методов к анализу оптического потока. Ближайшим по идеям к ансамблевым методам является генеративно-состязательные сети (Generative adversarial network – GAN), которые в основном применяются для синтеза изображений по их описанию или другим параметрам и классификации изображений. В работах [54, 55] анализируются возможности GAN для оптического потока, но по данному направлению мало информации.

Проанализировав направления машинного обучения, можно сказать, что на данный момент достаточно развитыми для применения в области анализа оптического потока и перспективными являются нейронные сети с глубоким обучением. Точного определения глубокого обучения нет, но обычно ссылаются на то, что сеть имеет хотя бы один скрытый слой и обучается методом обратного распространения ошибки. На практике сети отличают по названию конкретной реализации без внимания на то, есть ли глубокое обучение или нет.

Теперь рассмотрим некоторые виды нейронных сетей (см. рисунок 14) и их более детальное применение.

Вариационный автоэнкодер (Variational Autoencoder – VAE) [56] подобен сети прямого распространения. Основная идея автоэнкодера – автоматическое

# A mostly complete chart of Neural Networks

©2019 Fjodor van Veen & Stefan Leijnen asimovinstitute.org

- Input Cell
- Backfed Input Cell
- △ Noisy Input Cell
- Hidden Cell
- Probabilistic Hidden Cell
- △ Spiking Hidden Cell
- Capsule Cell
- Output Cell
- Match Input Output Cell
- Recurrent Cell
- Memory Cell
- △ Gated Memory Cell
- Kernel
- Convolution or Pool

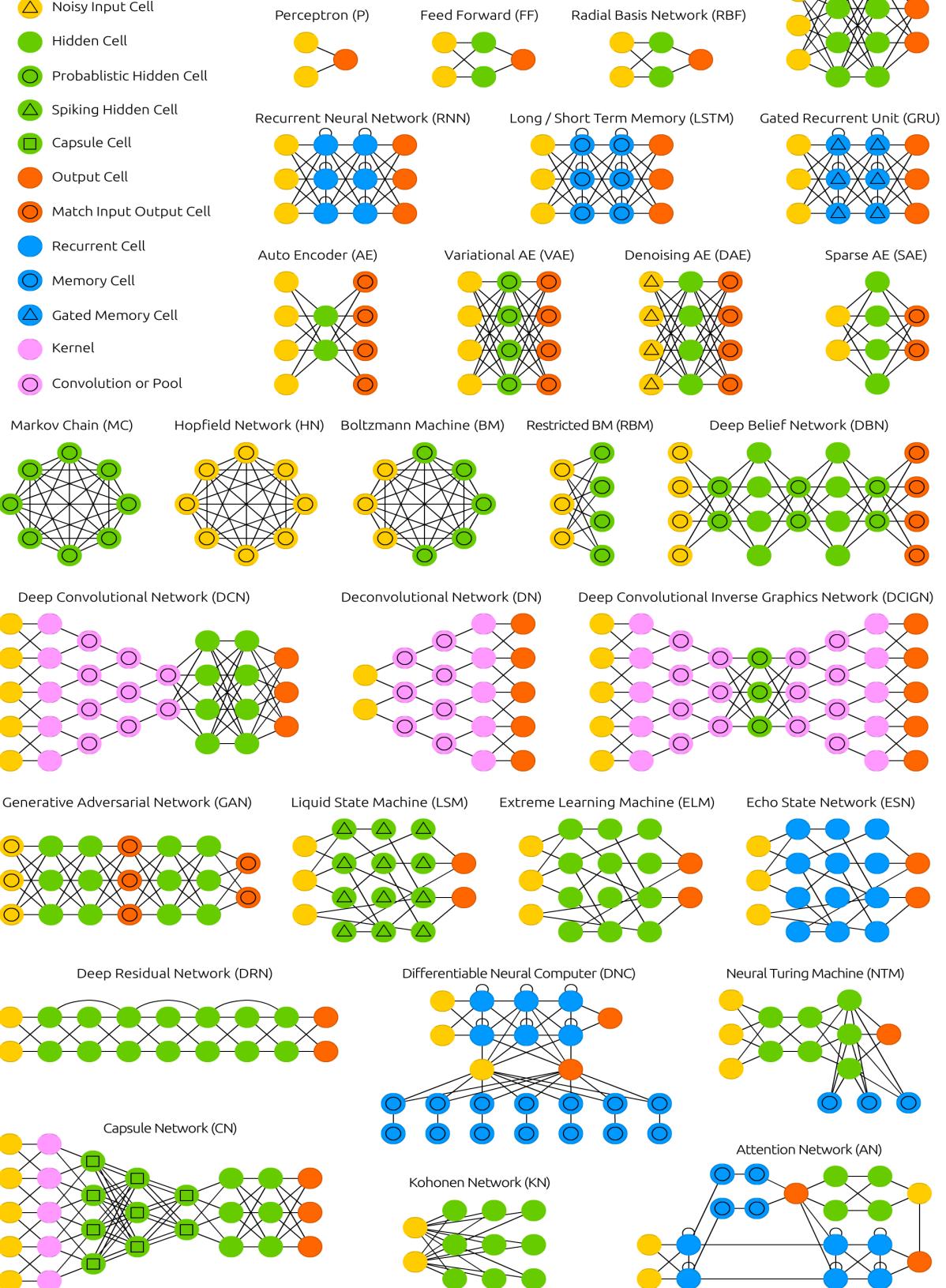


Рисунок 14 – Виды нейронных сетей [62]

кодирование информации как при сжатии. У вариационного автоэнокодера та же архитектура, но другой способ обучения.

Применения:

- генерация изображений, например, человеческих лиц,
- синтез музыки и речи,
- поиск аномалий на изображении.

В генеративно-состязательной сети (GAN) [57] суть заключается в том, что используются две сети, где одна из сетей генерирует данные (“генератор”), а вторая – анализирует (“дискриминатор”). Дискриминатор получает на вход или обучающие данные, или сгенерированные первой сетью. То, насколько точно дискриминатор сможет определить источник данных, служит потом для оценки ошибок генератора. Таким образом, происходит своего рода соревнование, где дискриминатор учится лучше отличать реальные данные от сгенерированных, а генератор стремится стать менее предсказуемым для дискриминатора. Это работает отчасти потому, что даже сложные изображения с большим количеством шума в конце концов становятся предсказуемыми, но сгенерированные данные, мало отличающиеся от реальных, сложнее научиться отличать. GAN достаточно сложно обучить, так как задача здесь – не просто обучить две сети, но и соблюдать необходимый баланс между ними. Если одна из частей (генератор или дискриминатор) станет намного лучше другой, то GAN никогда не сойдётся.

Применения:

- получение фотoreалистичных изображений,
- восстановление 3D модели объекта по его фрагментным изображениям,
- улучшение качества изображений.

Рекуррентные нейронные сети (Recurrent neural networks – RNN) [58] – это сети прямого распространения, но со смещением во времени: нейроны получают информацию не только от предыдущего слоя, но и от самих себя в результате предыдущего прохода. Следовательно, здесь важен порядок, в котором мы подаем информацию и обучаем сеть.

Применение [59, 60]:

- обработка естественного языка,
- сегментация кадров видео,
- распознавание движений,
- формирование словесного описание по изображению.

Свёрточные нейронные сети (Convolutional neural networks – CNN) [61] – эти сети кардинально отличаются от других тем, что в основном работают с изображениями. Обработка осуществляется не за раз, а путём сканирования некоторой подобласти и последующего сдвига (выполнение свёртки), что позволяет уменьшать размер данных, передаваемых на следующий слой. Данное направление пользуется большим интересом и поэтому активно разрабатываются различные реализации сетей и находят новые применения этой архитектуре.

Применения:

- Любая обработка изображений (классификация, нахождения объектов, сравнение, улучшение качества)
- обработка естественного языка

Также приведём примеры нейронных сетей, являющихся производными от выше упомянутых, предложенных относительно недавно.

Глубокие свёрточные обратные глубинные сети (Deep convolutional inverse graphics networks – DCIGN) [63] являются вариационным автоэнкодером со сверточной и разверточной сетями в качестве кодирующей и декодирующей частей соответственно. Такие сети представляют черты изображения в виде вероятностей и могут научиться строить изображение объектов вместе, взглянув лишь на каждый объект по отдельности.

Применения:

- моделирование сложных преобразований на изображениях
- генерация и синтез изображений
- определение объектов на изображении

Глубокие остаточные сети (Deep residual networks – DRN) [64] – это глубокие сети прямого распространения с дополнительными связями между слоями, соединяющими не только соседние слои, но и более отдаленные. Вместо того, чтобы искать способ находить соответствующие исходным данным входные данные через, скажем, пять слоев, сеть обучена ставить в соответствие входному блоку пару “выходной блок + входной блок”. Таким образом входные данные проходят через все слои нейросети и подаются почти полностью обработанными последним слоям.

Применения:

- генерация музыкальных композиций
- распознавание объектов
- аналогичны сетям RNN

Управляемые рекуррентные блоки (Gated recurrent units – GRU) [65] – это разновидность LSTM (Long short term memory – долгая краткосрочная память). У них на один фильтр меньше, и они немного иначе соединены: вместо входного, выходного фильтров и фильтра забывания здесь используется фильтр обновления. Этот фильтр определяет и сколько информации сохранить от последнего состояния, и сколько информации получить от предыдущего слоя. Фильтр сброса состояния работает почти так же, как фильтр обновления, но расположен немного иначе. На следующие слои отправляется полная информация о состоянии — выходного фильтра здесь нет. В большинстве случаев GRU работают так же, как LSTM, самое значимое отличие в том, что GRU немного быстрее и проще в эксплуатации (однако обладает немного меньшими выразительными возможностями).

Применения:

- анализ временных рядов,
- распознавание речи и рукописного ввода,
- аналогичны RNN.

Ниже приведены конкурентные реализации сетей, между которыми будет производиться выбор. В целом если рассматривать работы по нейронным сетям,

то больший успех и популярность в области обработки оптического потока имеют сети на основе архитектуры CNN [66].

Одной из первых сетей на основе CNN в области оптического потока является FlowNet (Досовицкий А. и др. [67] 2015 год), которую применяют для определения движения на видео. В работе [68] 2017 года представлено развитие этой сети FlowNet2. В статье [69] 2017 года предложена сеть PIV-DCNN для задач PIV, что напрямую связано с областью измерения деформаций, но данная реализация отличается низкой производительностью в сравнении с методами кросскорреляции. В работе [70] 2018 года от инженеров NVIDIA представлена сеть PWC-Net, также на основе CNN, для определения движения на видео. Данная реализация превосходит по точности и времени выполнения сеть FlowNet2, только нет исследований адаптации этой архитектуры к задачам PIV. В работе [71] 2019 года представлена адаптация сети FlowNet к задачам PIV, под названием PIV-NetS. Авторы этой статьи добились точности сопоставимой с методами кросскорреляции и классическим методом анализа оптического потока HS (Horn and Schunck), при этом время обработки изображений значительно ниже этих методов, что позволяет анализировать структуры в режиме реального времени. Работа [72] 2019 года аналогична работе [71], но сеть основана на облегчённой и улучшенной версии FlowNet2 под названием LiteFlowNet. Полученная в этой работе сеть названа PIV-LiteFlowNet.

Работы [54, 55] 2019 года демонстрируют сети на основе GAN. В [54] показано применение GAN для определения движения на изображения аналогично задачам, решаемым сетью FlowNet. В [55] GAN используются для задач PIV с целью получения сверхвысокого разрешения. Но эти работы не предоставляют не сравнения с существующими сетями, не характерные времена работы этих реализаций, что важно при выборе сетей.

В работе [73] 2020 года предложены сети SCNN и MTPC на основе CNN для исследования потока за цилиндром, что является частной задачей PIV. Время работы этих сетей сопоставимо с PIV-NetS.

В [74] 2020 года авторы предлагают дополнение к классическому методу HS для шумоподавления и обнаружения движения на видео в виде сети RBM (Restricted Boltzmann machine – ограниченная машина Больцмана).

В работе [75] 2020 года используют сеть SNN (Shallow neural network – малая нейронная сеть) для отслеживания отдельных частиц и определения 3D траекторий этих частиц.

Следующий шаг – детальный разбор и сравнение реализаций машинного обучения в задачах анализа оптического потока. По результатам которого будет сделан выбор конкретной сети и платформы для её реализации.

### 1.3.2 Анализ и выбор алгоритма машинного обучения

Начнём рассмотрение и сравнение с нейронных сетей в области анализа оптического потока не приспособленных к задачам PIV [67, 68, 70]. В работе [67] предлагаются сети FlowNetS и FlowNetC, архитектуры которых представлены на рисунках 15 и 16.

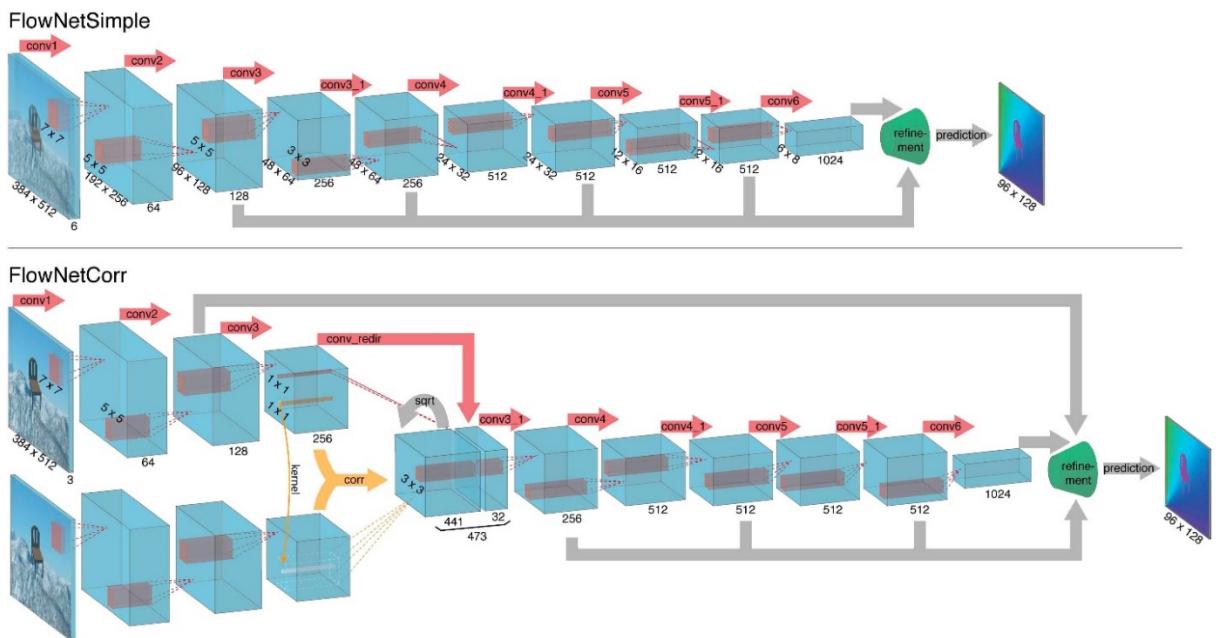


Рисунок 15 – Две сетевые архитектуры: FlowNetSimple (вверху) и FlowNetCorr (внизу). Зеленая воронка является заполнителем для расширяющейся части уточнения, показанной на рисунке 16. Сети, включая часть уточнения, обучаются от начала до конца [67]

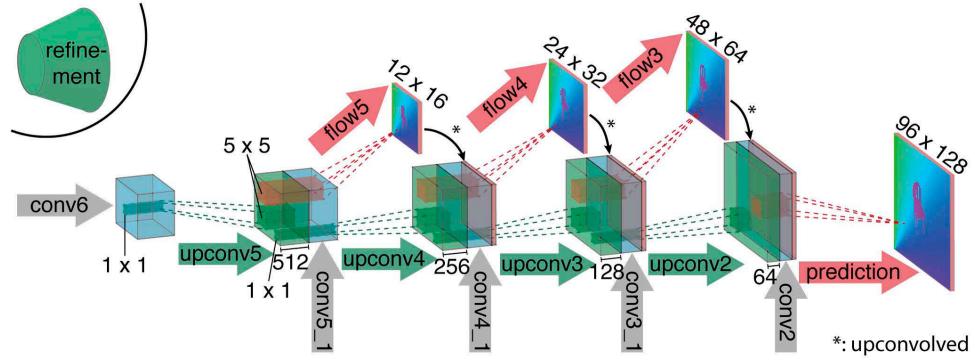


Рисунок 16 – Уточнение грубых карт признаков для прогноза с высоким разрешением [67]

Обе сети имеют схожую структуру. Имеются свёрточные слои, которые анализируя входную информацию уменьшают разрешение для грубой (общей) оценки, и уточняющие слои, которые повышают разрешение.

По сути сети имеют сжимающую часть и расширяющуюся часть и обучаются как единое целое с использованием обратного распространения ошибки. Сеть FlowNetS получает на вход пару изображений, где сама сеть решает каким образом необходимо обработать изображения. В FlowNetC подаются два изображения отдельно и только после нескольких свёрточных слоёв выполняется кросскорреляция. Результат кросскорреляции передаётся на последующие свёрточные слои.

В работе [68] представлено развитие FlowNet. Основные изменения заключаются в том, что авторы показали улучшение точности за счёт комбинации различных сетей (см. рисунок 17).

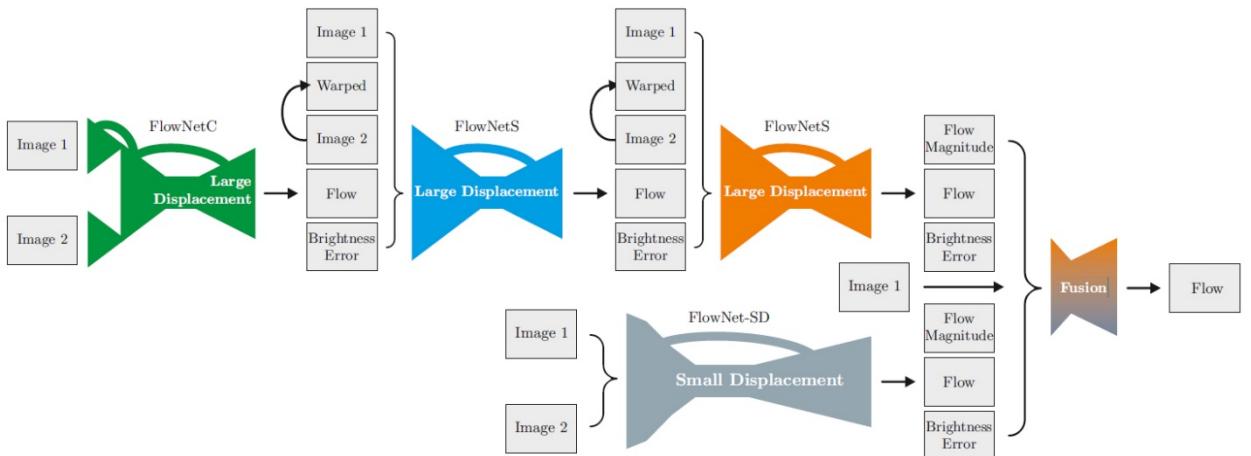


Рисунок 17 – Схематическое изображение полной архитектуры FlowNet2 [68]

Новая архитектура состоит из одной FlowNetC, двух FlowNetS и FlowNet-SD. FlowNet-SD – это сеть, обученная на данных имеющих малые смещения единицы пикселей и субпиксельные смещения. Полученная таким образом сеть названа FlowNet2. Она имеет значительное улучшение точности по сравнению с её предшественницами, но время её выполнения выросло. На рисунке 18 представлены качественные результаты работы рассмотренных сетей и других методов.

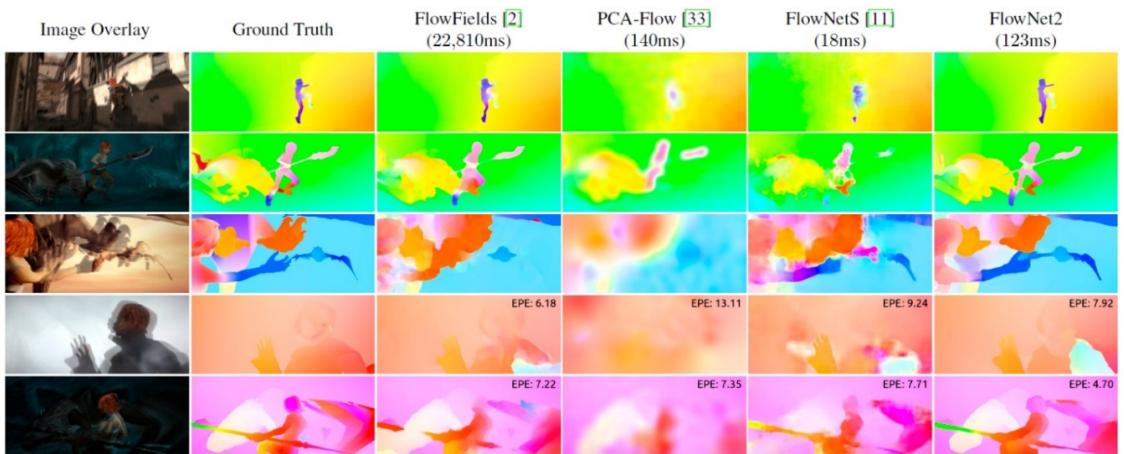


Рисунок 18 – Примеры оценок полей течения разными методами на открытой библиотеке Sintel. FlowNet2 работает аналогично FlowFields и может извлекать мелкие детали, в то время как методы, работающие на сопоставимых скоростях, работают намного хуже [68]

В работе [70] представлена сеть PWC-Net, которая точнее и быстрее FlowNet2. Архитектура этой сети в сравнении с традиционным подходом изображена на рисунке 19.

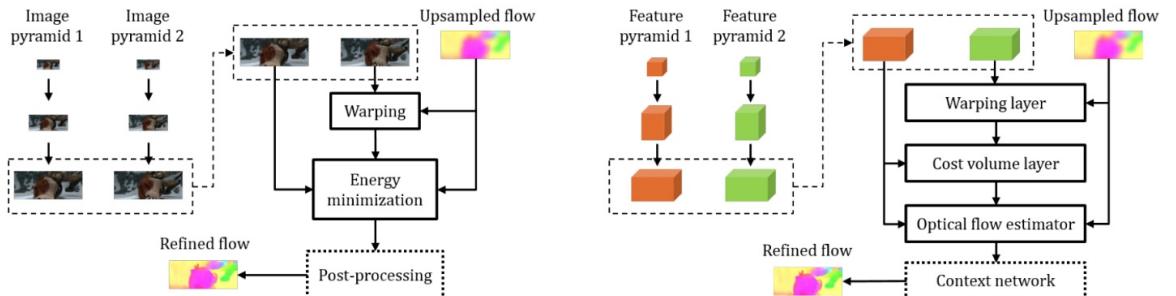


Рисунок 19 – Традиционный подход от грубого к мелкому по сравнению с PWC-Net. Слева: пирамида изображения и уточнение на одном уровне пирамиды с помощью подхода минимизации энергии. Справа: пирамида характеристик и уточнение на одном уровне пирамиды по PWC-Net. PWC-Net искажает характеристики второго изображения с помощью потока с повышенной дискретизацией, вычисляет объем затрат и обрабатывает объем затрат с помощью CNN [70]

Название PWC-Net состоит из отличительных методов этой сети. Первый: использование пирамид (Pyramid), а точнее пирамид свёрток. Второй: деформация (Warping) одного из изображений для уточнения признаков на следующем уровне. Третий: корреляция между пирамидами свёрток изображений называемой в этой работе объём затрат (Cost value). PWC-Net примерно в 17 раз меньше по размеру (по кол-ву параметров), в 2 раза быстрее в выводе и легче обучается, чем FlowNet2. PWC-Net работает со скоростью около 35 кадров в секунду на изображениях библиотеки Sintel с разрешением (1024×436).

Тестирование и сравнение сетей производится на открытых библиотеках:

Sintel. Бенчмарк, созданный на основе открытого и нарисованного в Blender мультика Sintel в двух версиях, которые обозначаются как clean и final. Второй намного сложнее, т.к. содержит множество атмосферных эффектов, шумов, блюра и прочих сложностей для алгоритмов вычисления оптического потока.

KITTI. Это набор изображений, размеченный под приложения для self-driving автомобилей и собранный с помощью технологии LIDAR. Он широко используется для валидации алгоритмов вычисления оптического потока и содержит множество довольно сложных случаев с резкими переходами между кадрами.

Middlebury. Малый набор из 8 пар изображений с небольшими смещениями, который, тем не менее, иногда используется при валидации алгоритмов вычисления оптического потока.

Таблица 2 – Сравнение нейронных сетей, значения под библиотеками — это среднеквадратичная (Евклидова) ошибка (EPE).

Архитектура	Sintel clean		Sintel final		KITTI		Middlebury		Время, мс (GPU)
	train	test	train	test	train	test	train	test	
FlowNetS	4,50	6,96	5,45	7,52	8,26	-	1,09	-	18
FlowNetC	4,31	6,85	5,87	8,51	9,35	-	1,15	-	32
FlowNet2	2,02	3,96	3,14	6,02	1,28	1,8	0,35	0,52	123
PWC-Net	2,55	4,39	3,93	5,04	1,45	1,7	-	-	30

На рисунке 20 представлено наглядное сравнение современных методов анализа оптического потока.

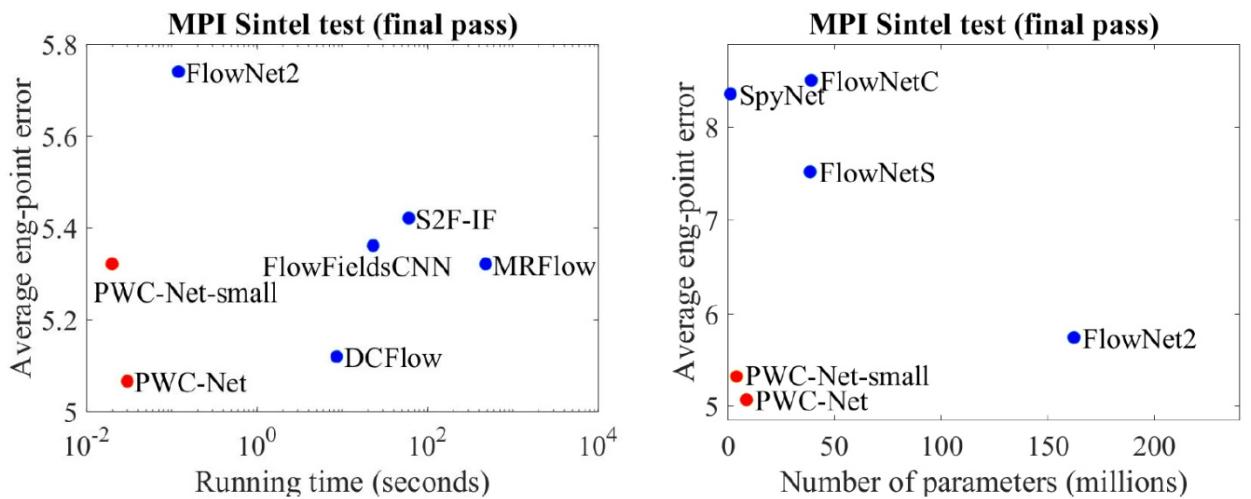


Рисунок 20 – Слева: PWC-Net превосходит все опубликованные методы в teste финального прохода MPI Sintel как по точности, так и по времени работы. Справа: PWC-Net достигает наилучшего баланса между размером и точностью среди существующих сквозных моделей CNN [70]

Рассмотрим существующие нейронные сети, адаптированные под задачи PIV [69, 71, 72]. Первыми кто попытались применить сети CNN к задачам PIV были авторы работы [69]. Авторы разработали сеть PIV-DCNN (глубокая свёрточная нейронная сеть) на основе библиотеки MatConvNet2. PIV-DCNN использует архитектуру каскадной регрессии для достижения грубой (больших смещений) и точной (малых смещений) оценки потока частиц. Подобно многопроходному методу кросскорреляции сеть ищет совпадения на частях изображения при каждом смещении окна опроса, (что значительно ухудшает производительность, т.к. сеть многократно опрашивается) при этом не использует деформацию изображения. На рисунке 21 представлена архитектура этой сети.

Несмотря на то, что PIV-DCNN требует больших вычислительных затрат она имеет преимущества над методом WIDIM (window deformation iterative multi-grid – итеративная многосеточная деформация окна, являющиеся модификацией кросскорреляционной обработки) в виде лучшего

пространственного разрешения и меньшего количества выбросов, но также, как и WIDIM извлекает разреженные карты векторов смещения.

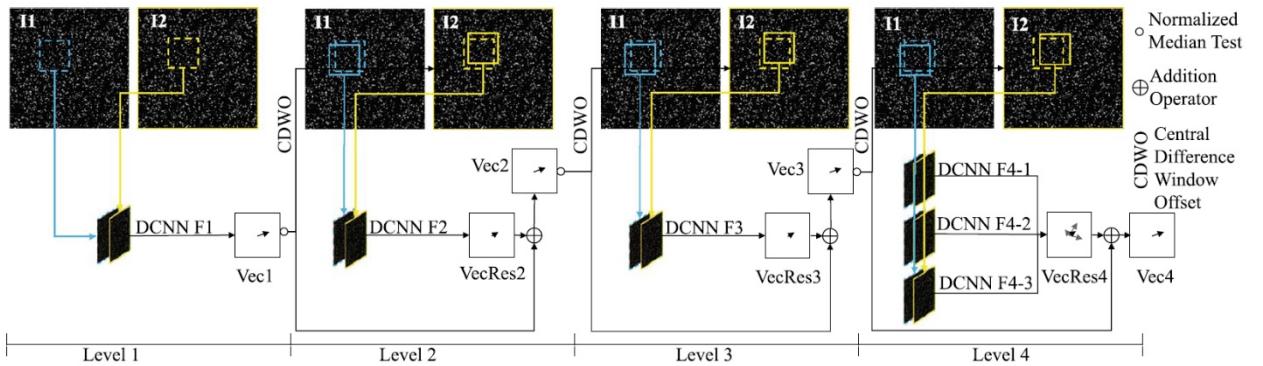


Рисунок 21 – Каскадные сверточные сети для PIV. Входными данными являются образцы изображения двух частиц ( $I_1$  и  $I_2$ ). Глубокие CNN обозначаются как  $F_1$ ,  $F_2$ ,  $F_3$ ,  $F_{4-1}$ ,  $F_{4-2}$  и  $F_{4-3}$  от уровня 1 до уровня 4. Каждая сеть может предсказать вектор смещения из двух участков (цветных квадратов).  $F_1$  на уровне 1 задуман как экстрактор большого смещения, а сети на уровне 2, 3 и 4 могут уточнять вектор, оценивая остаточный вектор ( $VecRes$ ) после смещения окна центральной разности (два окна компенсируют половину прогнозируемого смещения, сплайн-интерполяция для значения интенсивности изображения) [69]

В работе [71] представлена адаптация FlowNetS к задачам PIV (рисунок 22). Основные изменения полученной сети состоят в следующем:

- добавление деконволюционных слоёв до полного разрешения изображения, в отличии от FlowNetS, где выход получается в 4 раза меньше

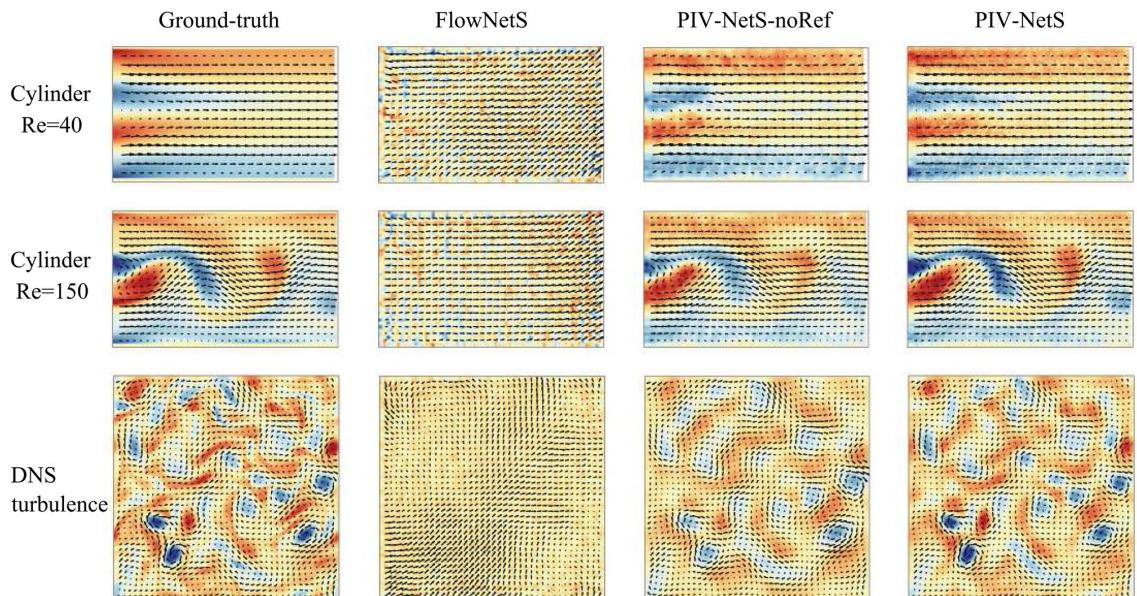


Рисунок 22 – Истинное поле и оценочные поля, извлеченные из изображений частиц с помощью различных сетевых моделей. Цветная карта демонстрирует завихренность потоков [71]

входа и расширяется до входного разрешения методами интерполяции. Это связано с тем, что в оценки PIV необходимо исследовать мелкомасштабные структуры по сравнению с задачами оценки жёстких движений;

- нормировка интенсивности между изображениями на каждом слое сети;
- и главное отличие в наборе обучающих данных, что будет продемонстрировано на рисунке 22.

Полученная сеть названа PIV-NetS. И для сравнения сеть без добавления деконволюционных слоёв названа PIV-NetS-noRef. На рисунке 22 можно увидеть, как справляются представленные сети с оценкой потока частиц.

На рисунке 23 можно увидеть сравнение сетей с классическим методом оптического потока HS и корреляционным методом WIDIM. Время работы и число получаемых векторов представлены в таблице 3.

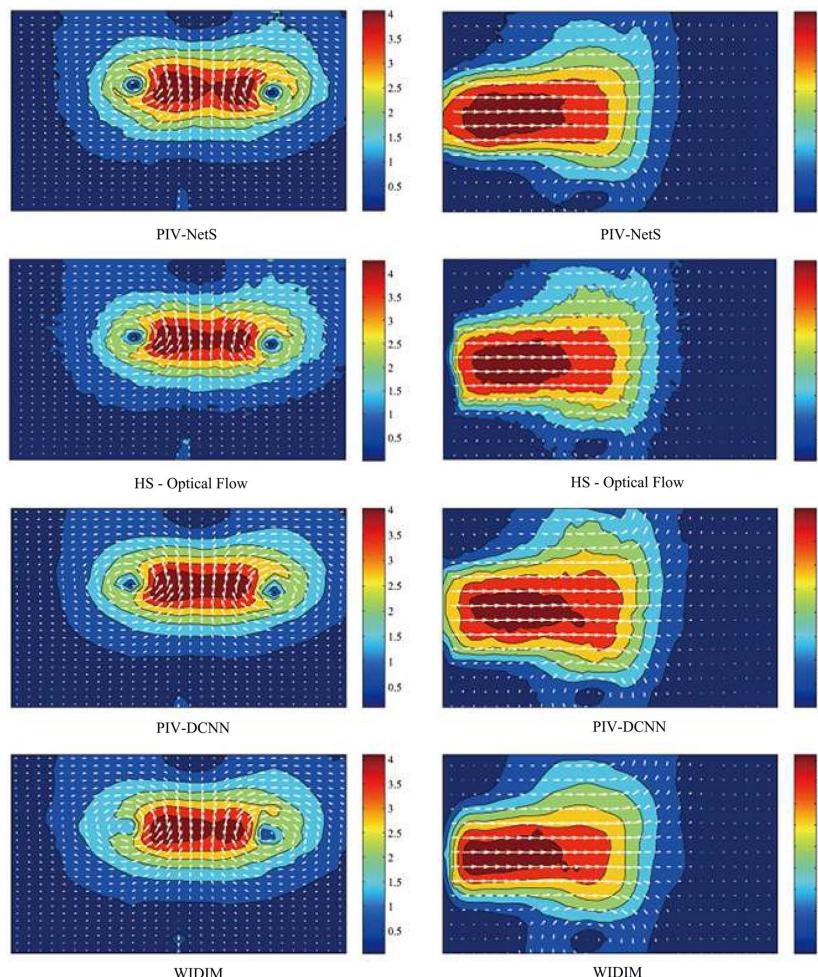


Рисунок 23 – Поля скорости, оцененные по лабораторным PIV-изображениям: поток вихревой пары (слева) и струйный поток (справа). Цветной фон с контурами обозначает соответствующую величину скорости [71]

Таблица 3 – Время выполнения (единицы: мс) различных методов лабораторной оценки PIV [71]

	PIV-NetS	HS	WIDIM	PIV-DCNN	
	GPU	CPU	CPU	CPU	GPU
Поток вихревой пары	51 (480×512)	3171 (480×512)	916 (27×29)	6970 (27×29)	2388 (27×29)
Струйный поток	79 (512×1024)	8140 (512×1024)	1713 (29×61)	15328 (29×61)	4721 (29×61)

Работа [72] это продолжение и развитие работы [71]. В основе сети авторы используют архитектуру LiteFlowNet [76] – улучшение сети FlowNet2, архитектура которой представлена на рисунке 24. LiteFlowNet состоит из двух подсетей: кодировщика NetC и декодера NetE. NetC специализируется на извлечении признаков, а NetE выполняет грубую оценку оптического потока.

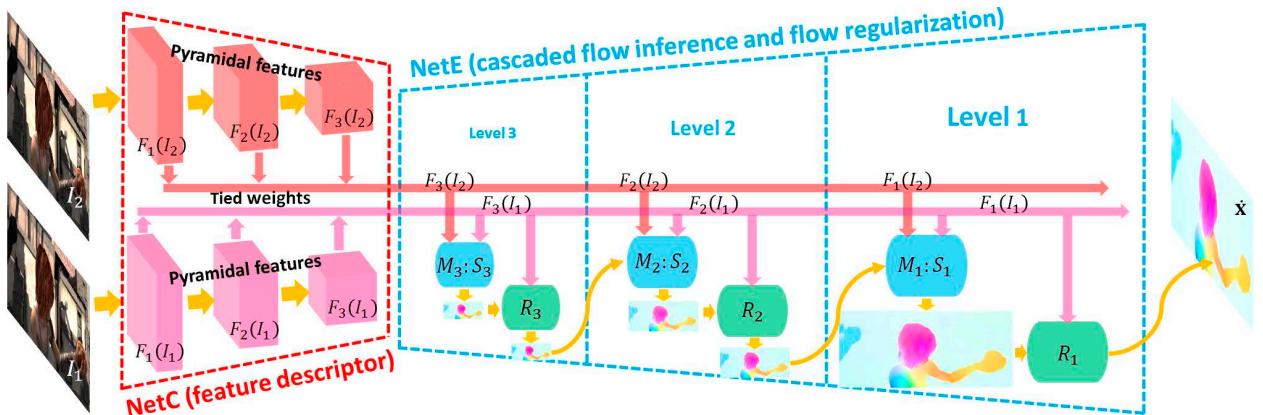


Рисунок 24 – Сетевая структура LiteFlowNet. Для простоты представления показан только трехуровневый дизайн. Учитывая пару изображений ( $I_1$  и  $I_2$ ), NetC генерирует две пирамиды высокоуровневых функций ( $F_k(I_1)$  розовым и  $F_k(I_2)$  красным,  $k \in [1; 3]$ ). NetE дает многомасштабные поля потока, каждое из которых генерируется модулем каскадного вывода потока M: S (синим цветом, включая блок сопоставления дескрипторов M и блок уточнения субпикселей S) и модуль регуляризации R (зеленый цвет). Модули вывода потока и регуляризации соответствуют условиям точности данных и регуляризации в традиционных методах минимизации энергии соответственно [76]

Доработка сети LiteFlowNet для задач PIV соответствует изменениям, внесённым в FlowNetS для создания PIV-NetS (см. рисунок 25). Полученная сеть названа PIV-LiteFlowNet-en. В этой работе представлено сравнение с PIV-NetS, которую называют PIV-FlowNetS-en.

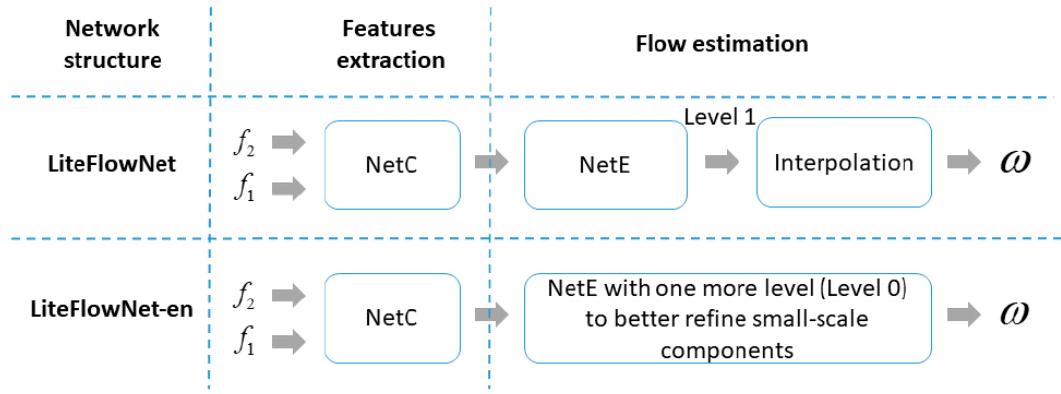


Рисунок 25 – Сравнение исходной структуры и модифицированной структуры LiteFlowNet [72]

На рисунке 26 представлено сравнение ранее упомянутых сетей и хорошо зарекомендованных методов.

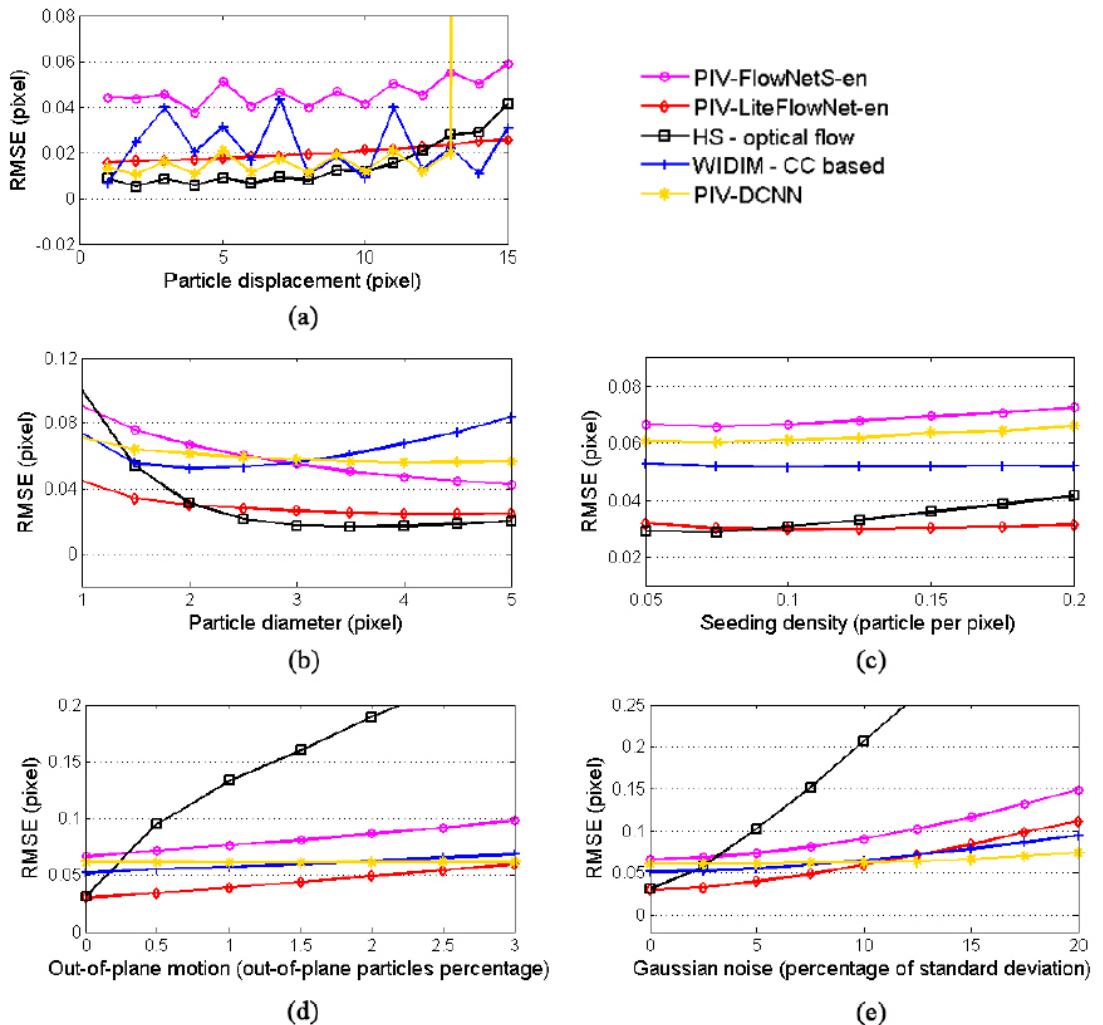


Рисунок 26 – RMSE (средняя квадратическая ошибка) относительно (а) смещения частицы в однородном потоке, (б) диаметра изображения частицы, (в) плотности затравки частицы, (г) внеплоскостного движения частицы, (д) гауссовского шума изображений потока с выходом вихрей над круговой цилиндр [72]

Также на рисунке 27 можно увидеть разницу в пространственном разрешении на синтетическом тесте. В таблице 4 представлено сравнение между различными методами оценки PIV на 500 изображениях с разрешением  $256 \times 256$  пикселей. Загрузка сетевой модели производится одни раз перед началом обработки набора изображений. Обработка пары изображений для модели PIV-LiteFlowNet-en без пост-обработки составляет 44 мс.

В работе [54] впервые пытаются применить сеть GAN к расчёту оптического потока на стереоизображениях. Предлагаемая авторами сеть SceneFlowGAN в работе обучена контролируемым способом (обучение с учителем), но сообщается о возможности неконтролируемого обучения и создания набора данных с натуралистичной сценой при помощи этой сети.

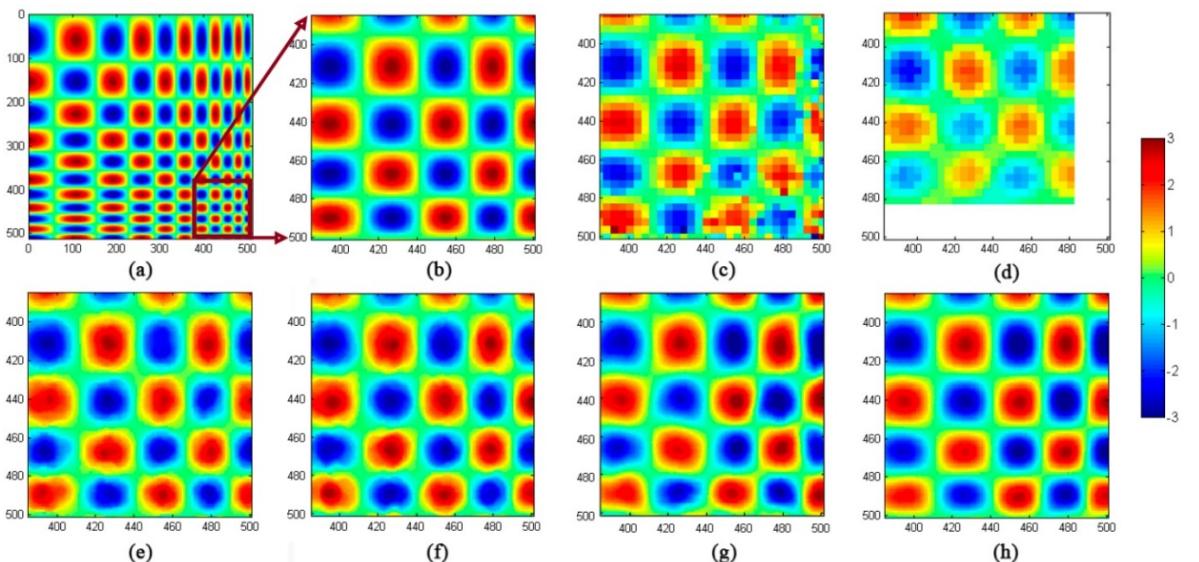


Рисунок 27 – Тест на пространственное разрешение. Карты цвета фона представляют величины скорости  $v$ -компоненты. (а) и (б) отображают полное разрешение и увеличенную область достоверных данных, соответственно. Остальные цифры представляют собой оценочные результаты (с) метода WIDIM, (д) модели PIV-DCNN, (е) метода оптического потока HS, (ф) метода OF + STE, (г) модели PIV-FlowNetS-en, (х) модели PIV-LiteFlowNet-en, в увеличенном окне. Модель PIV-DCNN не может оценить скорость вблизи границы изображения, поэтому на рисунке (д) отсутствует часть [72]

SceneFlowGAN применяется для стереопар изображений. Генератор этой сети основан на архитектуре SceneEDNet [77]. Полученные данные кривых обучения авторами работы говорят о недостаточном наборе обучающих данных. Поэтому результат тестирования показывает значительные ошибки по

сравнению с существующими сетями. Также не сообщается о времени обработки этой сетью.

В работе [55] предлагается использование уже разработанных сетей SRGAN и ESRGAN в механике жидкостей (задачи PIV). Сети на основе векторного поля низкого разрешения, полученного при помощи метода кросскорреляции, увеличивают разрешение в 4 или 8 раз. Данная работа исследует разрешение потока жидкости за цилиндрическим объектом. Полученные результаты показывают высокую точность реконструкции поля. Но нет данных по времени обработке изображений, хотя авторы ESRGAN заявляют о работе в режиме реального времени.

Таблица 4 – Сравнение различных методов оценки PIV. Времена выполнения указаны в секундах [72]

	PIV-LiteFlowNet-en	PIV-FlowNetS-en	HS	WIDIM	PIV-DCNN
Процессор	GPU	GPU	CPU	CPU	CPU
Число вектор	256×256	256×256	256×256	29×29	25×25
Загрузка сети	1,45	1,40	-	-	-
Выполнение	0,044	0,010	1,075	0,422	3,926
Медианная фильтрация	0,038	0,038	0,038	0,001	0,001
Общее время (500 изображений)	42,45	25,9	556,5	211,5	1963,5

В работе [73] разработаны сети SCNN (статическая CNN) и MTPC (нейронная сеть с множественными временными путями). Эти сети предназначены для повышения разрешения из грубых оценок векторов в задачах PIV в 4 и 8 раз, аналогично работе [55]. SCNN принимает на вход только одно изображение, а MTPC исследует серию кадров, таким образом имея больше информации для повышения разрешения. Время работы сетей для восстановления формы потока до  $1536 \times 2048 \times 3$  из данных  $384 \times 512 \times 3$  занимает 0,03 и 0,38 секунды для моделей SCNN и MTPC соответственно. Сеть SCNN показывает точность лучше бикубической интерполяции, а MTPC точность выше SCNN.

В работе [74] предлагается сеть на основе RBM (ограниченной машины Больцмана) для задач шумоподавления и обнаружения границ (краёв движущегося объекта) на видео. Предложенная сеть является лишь дополнением к классическому методу расчёта оптического потока HS.

В работе [75] изучается применение нейронных сетей для задач 3D-PTV (трёхмерной отслеживающей велосиметрии частиц). Авторы разработали SNN (мелкую нейронную сеть) для быстрого решения уравнений отображения частиц. Расчёт положения 1000 частиц занимает не более 0,5 секунды, для сравнения классический итерационный метод решения 3D-PTV занимает около 11 часов.

В процессе изучения актуальных сетей были обнаружены ещё несколько интересных решений для задач оценивания оптического потока [77-80]. Работы [77, 78] посвящены оценке оптического потока при стереосъёмке.

Первой из этих работ является работа [77], где представлена сеть SceneEDNet для оценки оптического потока по стереоизображениям. На рисунке 28 представлена идея оценки по стереопаре. Авторы SceneEDNet не предоставляют подробную архитектуру сети, но показывают реализацию аналогичную CNN сетям.

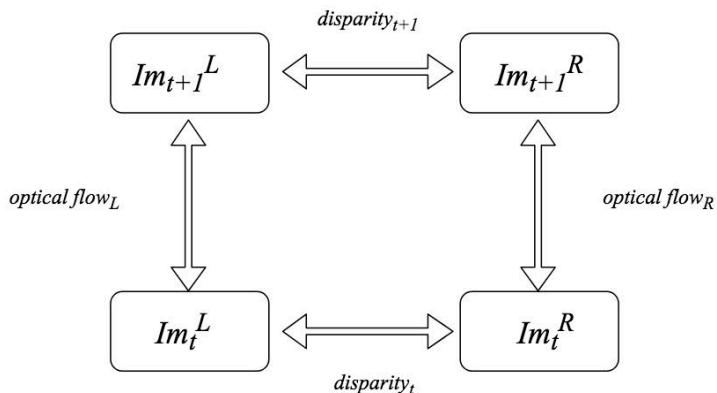
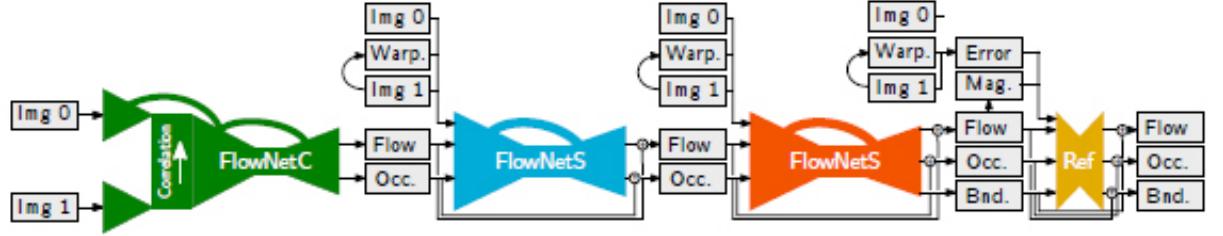
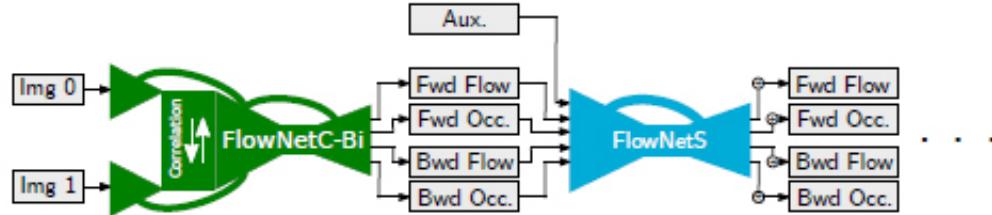


Рисунок 28 – Связь между парами стереоизображений. Изображения с последовательными временными шагами дают левый и правый оптический поток. Тогда как стереопары дают несоответствия [77]

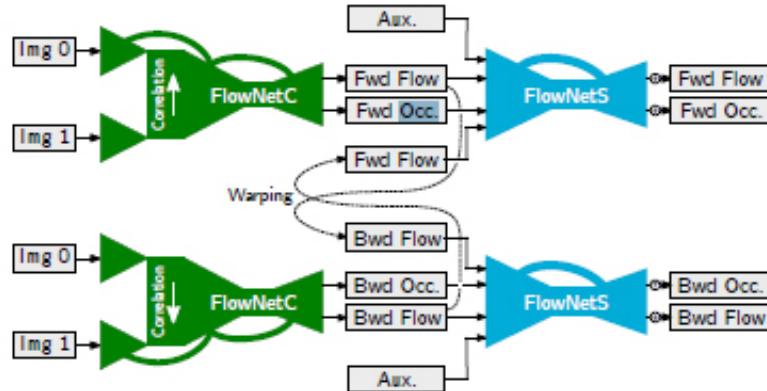
Работа [78] производит учёт окклузий, движений и границ глубины сетью для оценки диспаратности. Авторы предлагают комбинацию сетей DispNet и FlowNet для улучшения оценки смещения при стереосъёмке. Архитектура решения показана на рисунке 29.



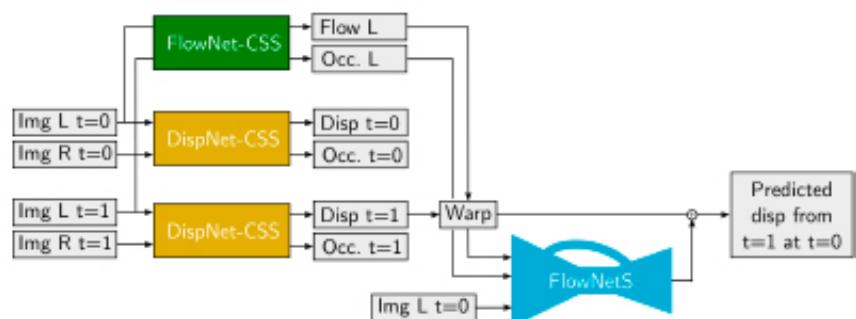
(a) Extension of FlowNet2 with occlusions and residual connections.



(b) Architecture for joint estimation of forward/backward flows and occlusions. See figure caption for symbol explanation.



(c) Dual forward and backward estimation architecture with mutual warping. See figure caption for symbol explanation.



(d) Extending FlowNet-CSS and DispNet-CSS to a full scene flow network.

Рисунок 29 – Обзор возможных стеков уточнения для границ потока, оклюзии и движения. Остаточные соединения показаны только на первом рисунке и обозначены знаком + в другом месте. Доп. относится к изображениям плюс искаженное изображение для каждого входного потока соответственно. Архитектуры для случая несоответствия аналогичны [78]

Работы [77] и [78] позволяют производить обработку стереопар, что даёт больше информации о сцене, но в данном направлении ещё мало работ и не производились попытки приспособления существующих решений к задачам PIV.

В работе [79] представлена сеть UnLiteFlowNet для PIV, которая обучалась без учителя. Данная модель позволяет решить проблему нехватки обучающих данных для обучения с учителем. В некоторых сложных структурах поля (см. рисунок 30) сеть UnLiteFlowNet превосходит по точности PIV-LiteFlowNet, но обычно имеет большие среднеквадратичные ошибки (см. таблицу 5).

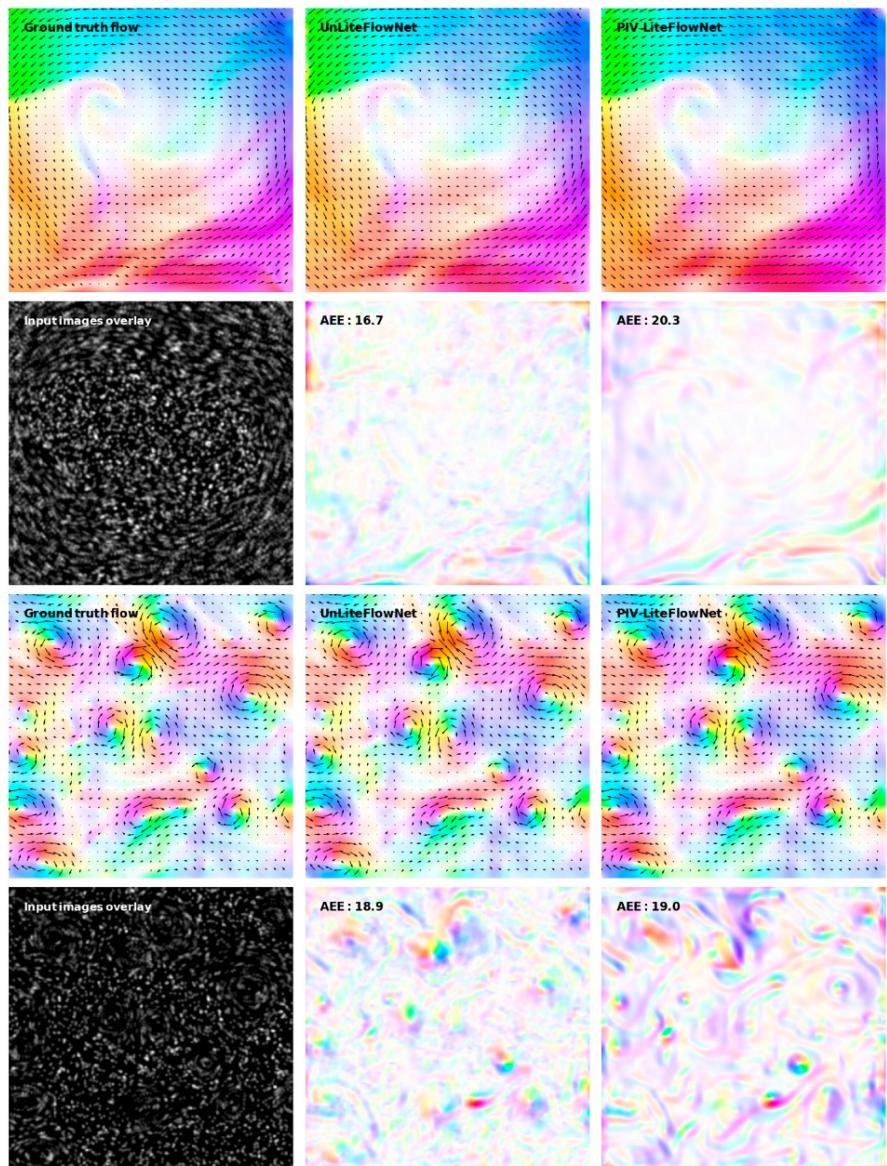


Рисунок 30 – Визуальные сравнения между данными истинного потока (слева), полностью неконтролируемой моделью UnLiteFlowNet-PIV (в центре) и PIV-LiteFlowNet (справа), показанными в первой, третьей и пятых строках для соответственно [79]

Таблица 5 – Усредненная ошибка конечной точки (AEE) для набора данных PIV (усредненная по всем парам изображений), единица ошибки установлена на пиксель на 100 пикселей для упрощения сравнения [79]

Методы	Back-Step		Cylinder		JHTDB channel		DNS turbulence		SQG	
	train	test	train	test	train	test	train	test	train	test
WIDIM	-	3,4	-	8,3	-	8,4	-	30,4	-	45,7
HS optical flow	-	4,5	-	7,0	-	6,9	-	13,5	-	15,6
PIV-NetS-noRef	13,6	13,9	19,8	19,4	24,6	24,7	50,6	52,5	51,9	52,5
PIV-NetS	5,8	5,9	6,9	7,2	16,3	15,5	27,1	28,2	28,9	29,4
PIV-LiteFlowNet	5,5	5,6	8,7	8,3	10,9	10,4	18,8	19,6	19,8	20,2
PIV-LiteFlowNet-en	3,2	3,3	5,2	4,9	7,9	7,5	11,6	12,2	12,4	12,6
UnLiteFlowNet-PIV	-	10,1	-	7,8	-	9,6	-	13,5	-	19,7

В работе [80] разработана сеть DCVNet, которая, по заявлениюм авторов, быстрее и точнее сети PWC-Net в задачах анализа оптического потока благодаря расширению расчёта объёма затрат (Cost value), чтобы одновременно фиксировать как малые, так и большие смещения на изображениях. Сравнение предложенной сети с существующими аналогами представлено на рисунке 31.

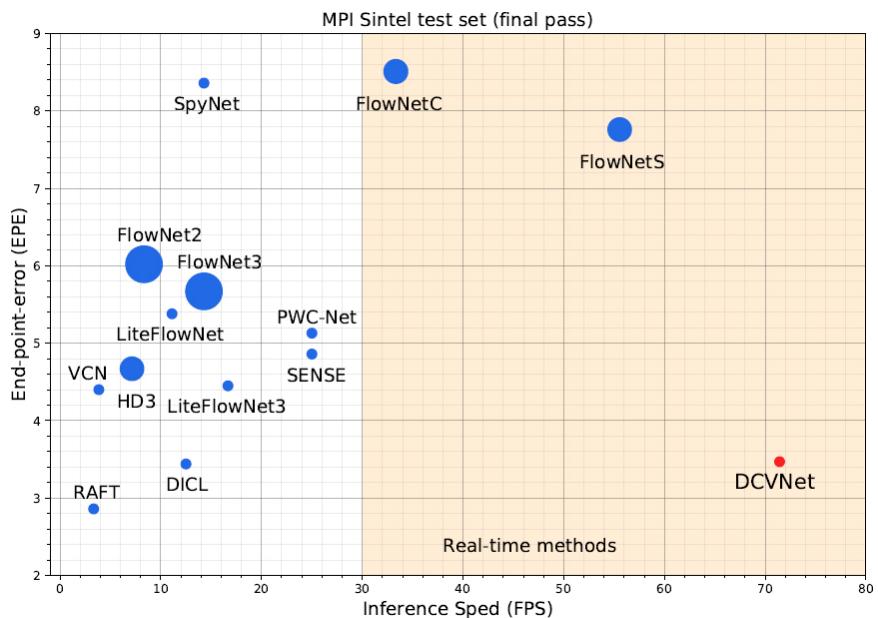


Рисунок 31 – Сравнение скорости вывода, размера модели и ошибки конечной точки (EPE) для разных моделей. Для каждой модели радиус круга указывает количество параметров (больший радиус означает больше параметров). Центр круга соответствует EPE модели. Предлагаемый DCVNet работает значительно быстрее, чем другие, обеспечивая скорость вывода в реальном времени. В то же время он компактен и обеспечивает сопоставимую частоту ошибок с существующими подходами [80]

Выбор архитектуры сети падает на PIV-LiteFlowNet-en. Рассмотрим, чем уступают другие модели. Сети для задач оптического потока такие, как: FlowNetS, FlowNetC, FlowNet2, LiteFlowNet, PWC-Net, DCVNet в исходном виде не приспособлены к задачам PIV, хотя вариации FlowNet имеют адаптацию, но в их архитектуры внесены изменения, поэтому можно считать их новым моделями, которые, обычно, в названии имеют аббревиатуру PIV.

Из сетей на основе CNN приспособленных к PIV: PIV-DCNN, PIV-NetS (или PIV-FlowNetS-en), PIV-LiteFlowNet-en лучший результат по точности показывает сеть PIV-LiteFlowNet-en, при этом немного уступая по времени обработки PIV-NetS.

По сетям на основе GAN мало работ и информации для сравнения их другими решениями.

Сети SCNN и MTPC, которые повышают разрешение, не подходят, т.к. уже PIV-NetS и PIV-LiteFlowNet-en дают разрешение на каждый пиксель, т.е. в исходном разрешении входного изображения. Вероятно, попробовать сети повышения разрешения для построения субпиксельных карт векторов смещения, если это возможно, но работ в данном направлении не было обнаружено.

Для сети на основе RBM необходим классический метод оптического потока HS, что говорит о заведомо худших результатах в сравнении с PIV-NetS.

Сеть SNN используется для быстрого решения уравнений отображения, поэтому не подходит по сути к задаче определения деформаций.

Сети SceneEDNet и сеть на основе DispNet и FlowNet [78], хоть не имеют адаптации к PIV, но работают со стереоизображениями, поэтому идеи и методы из этих работ могут пригодиться к исследуемой области.

Сеть UnLiteFlowNet имеет, обычно, большие ошибки, чем PIV-LiteFlowNet-en, но обучается неконтролируемым способом, поэтому может быть полезна в случае нехватки обучающих данных.

Подробнее о PIV-LiteFlowNet-en. Эта модель не применялась к стереоизображениям, но для того, чтобы извлекать максимум информации из

сцены можно воспользоваться идеями, предложенными на рисунках 28 и 29. Для начала возможно загружать не серию снимков во времени, а пары изображений с разных камер, тем самым каждый раз меняя опорный кадр и производить расчёт карт диспаратности для этой пары. Потом к этому можно добавить динамику деформации поверхности объекта с каждой из камер.

В случае заинтересованности в улучшении скорости обработки можно будет попытаться приспособить сети PWC-Net или DCVNet к исследуемой задаче. Если на результате точности будет сказываться нехватка обучающих данных, то возможно применение UnLiteFlowNet.

Большим плюсом PIV-LiteFlowNet-en, в сравнении с некоторыми рассмотренными сетями, является наличие открытого исходного кода. Данная сеть реализована на библиотеке Caffe и может запускаться как на языке Python, так и с помощью среды Matlab.

## **2 Экспериментальная часть**

### **2.1 Проверка работоспособности реализованного алгоритма оптимизации путем компьютерного моделирования декалибровки**

#### **2.1.1 Моделирование декалибровки**

Процесс моделирования декалибровки заключается в следующем:

1) Задаются параметры камер как внутренние (фокусные расстояния, размер пикселя и координаты центральной точки), так и внешние (вектор смещения и повороты камер друг относительно друга). Все эти параметры определяются в эксперименте из этапа калибровки камер.

2) Формируется объект измерения, как набор точек в 3-х мерном пространстве.

3) Без смещения (декалибровки), путём проецирования 3-х мерных точек на плоскость изображения камер, рассчитываются 2-х мерные точки, которые выступают в роли действительно получаемых изображений с камер.

4) Вносится известное смещение (декалибровки) в вектор смещения или поворота камер друг относительно друга.

5) С учётом смещённых внешних параметров камер и полученных в п. 3 2-х мерных точек рассчитываются, путём триангуляции, 3-х мерные точки объекта.

6) Потом полученные 3-х мерные точки снова проецируются на плоскость изображения камер, в результате чего получаем перепроецированные 2-х мерные точки.

7) Далее сравнивая исходные 2-х мерные точки и перепроецированные оценивается величина ошибки репроекции.

Если величина смещения равна нулю, то ошибка репроекции будет близка к нулю, не нулевое значение обусловлено только машинной погрешностью.

Приведём пример моделирования декалибровки. Зададим следующие параметры камер: отношение фокуса к размеру пикселя равно 7500; центральная точка камеры исходя из её разрешения [1250; 980]; смещение одной камеры относительно другой  $T_x = -315$  мм,  $T_y = 115$  мм,  $T_z = 40$  мм; поворот одной камеры

относительно другой вокруг оси  $x$  равен  $3,5^\circ$ , оси  $y$  равен  $12^\circ$ , оси  $z$  равен  $-1^\circ$ . Такие параметры выбраны потому, что близки к действительным калибровочным параметрам. Полученный объект и изображения объекта на камерах представлены на рисунке 32.

Параметры измеряемого объекта следующие: расстояние до объекта 1,5 м; размер объекта  $30 \times 20$  см; количество точек  $10 \times 10$  штук; угол наклона объекта  $45^\circ$ . Объект находится напротив левой камеры. Расчёт зависимости ошибки репроекции от изменения внешних параметров камер при нулевых начальных смещениях представлен на рисунке 33.

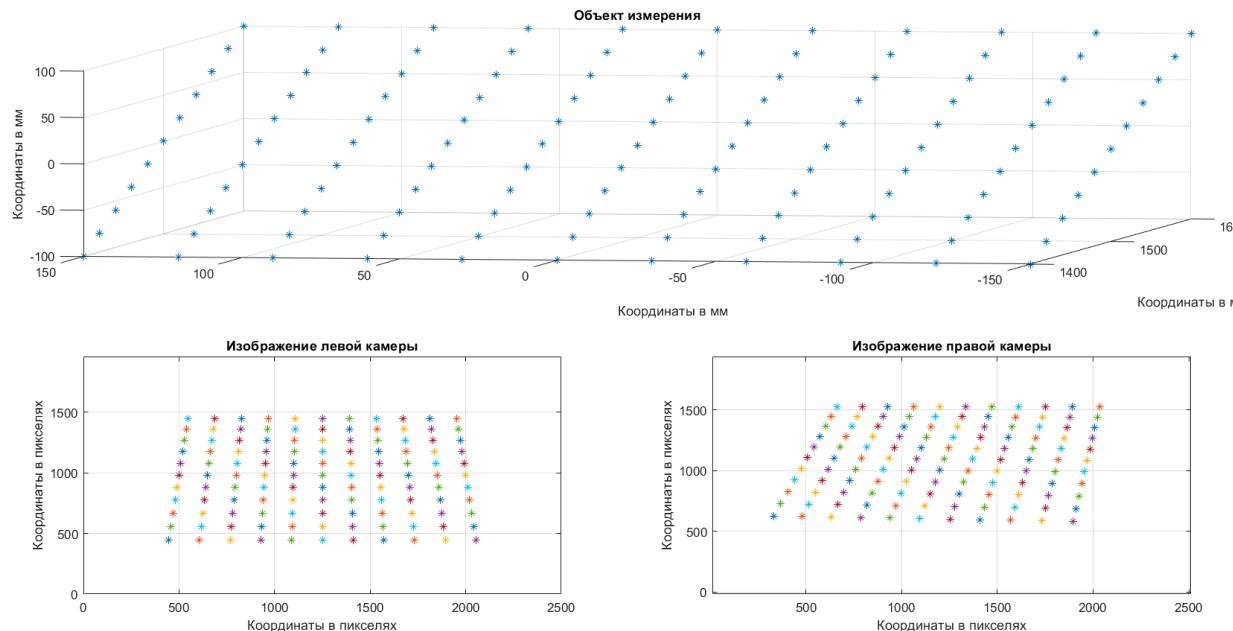


Рисунок 32 – Внешний вид объекта и изображения, получаемые на камерах  
начало системы отсчёта координат

Если внести начальное смещение, например, поворот вокруг оси  $y$  сместить на  $+1,5^\circ$ , то на некоторых графиках сместится минимум (см. рисунок 34), что говорит о том, что простая оптимизация по отдельным осям смещения и поворота не приведёт к глобальному минимуму целевой функции, т.е. не будет получено верного или близкого к верному решения.

В процессе оптимизации минимизируется целевая функция, роль которой выполняет ошибка репроекции в зависимости от 6-ти параметров:  $T_x$ ,  $T_y$ ,  $T_z$  – элементы вектора смещения и  $\alpha$ ,  $\beta$ ,  $\gamma$  – углы поворота одной камеры

относительно другой [20]. В п. 4 смещения выполняются методом Нелдера-Мида и затем повторяются п. 5, 6, 7 и т.д., до тех пор, пока не достигнем максимального значения итераций или значения ошибки репроекции меньше заданного.

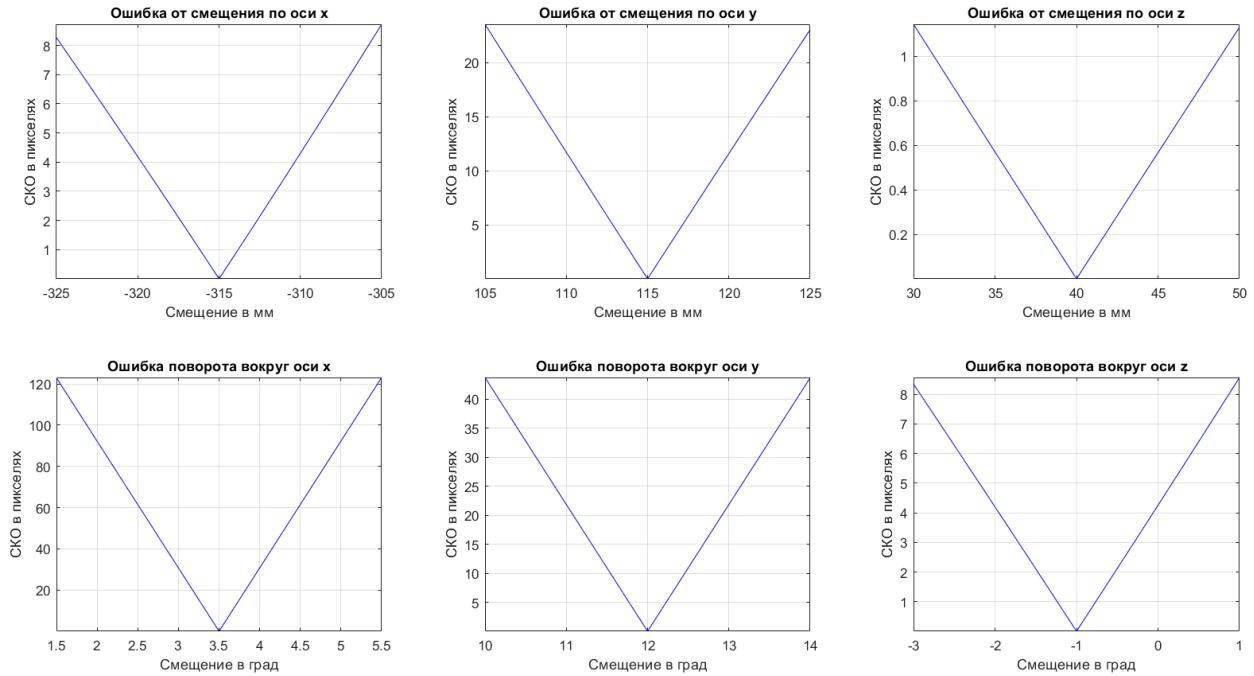


Рисунок 33 – Зависимость ошибки репроекции от изменения внешних параметров камер

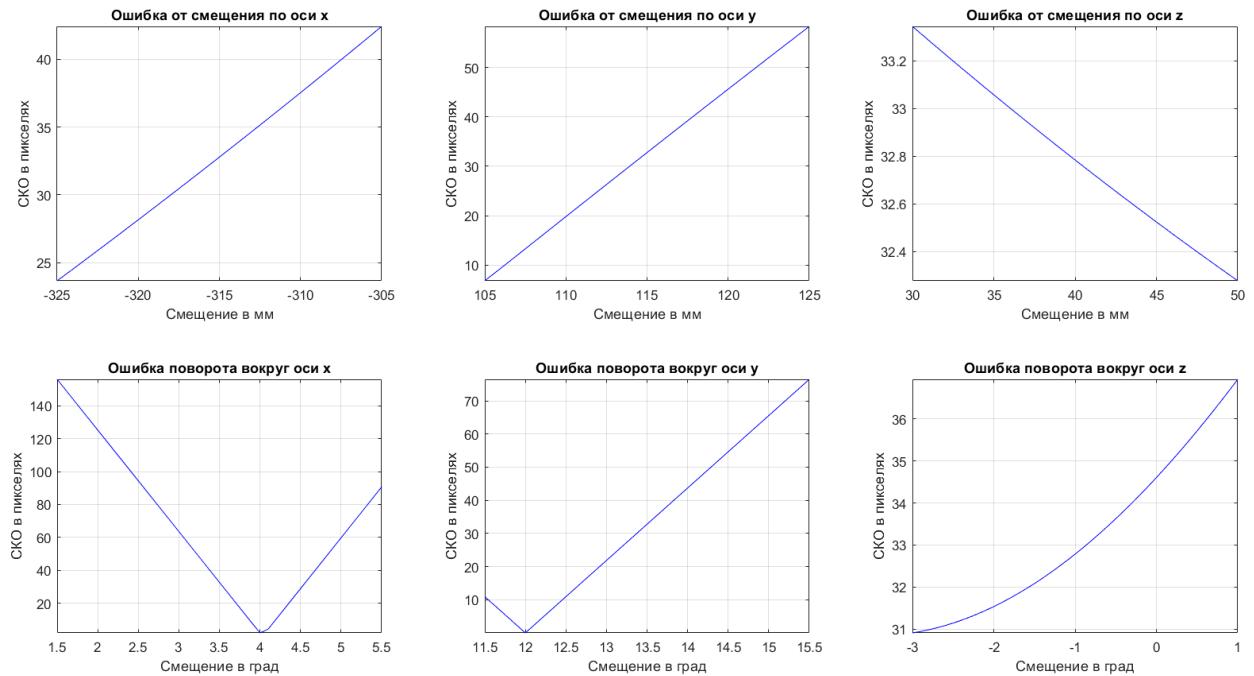


Рисунок 34 – Зависимость ошибки репроекции от изменения внешних параметров камер

## 2.1.2 Оптимизация

Приведём пример зависимостей ошибок репроекции от смещения по оси x и поворота вокруг оси x без оптимизации и с оптимизацией (см. рисунок 35). Вертикальная ось отображена в логарифмическом масштабе, красным результатом без оптимизации, а зелёным результатом с оптимизацией. Оптимизация выполнялась методом Нелдера-Мида совместно с методом штрафных функций с числом попыток для каждого из смещений равным 12 и максимальным количеством итераций в каждой попытке равным 500, шаги расчёта смещений равны 0,1 мм и 0,02°.

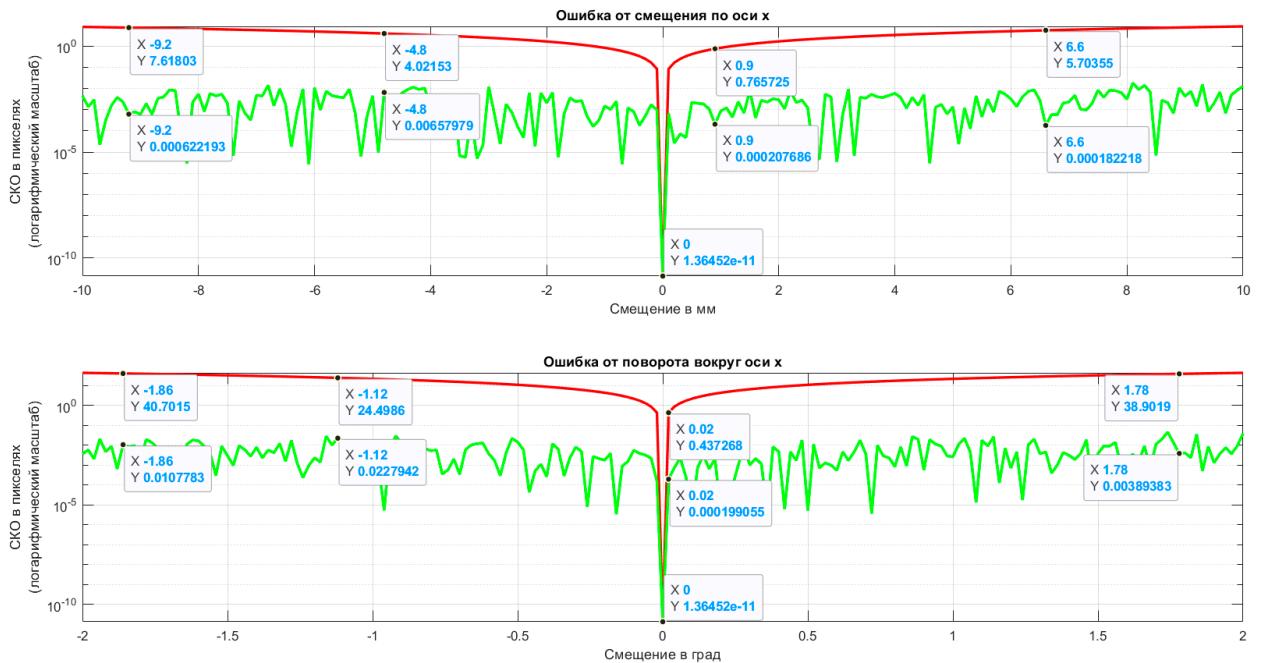


Рисунок 35 – Зависимость ошибки репроекции от смещения без оптимизации и с оптимизацией: вертикальная ось отображена в логарифмическом масштабе; красным – результат без оптимизации; зелёным – результат с оптимизацией

Из рисунка выше видно, что ошибка репроекции в результате оптимизации значительно уменьшилась, примерно на два порядка. Рассчитывать на такое же уменьшение ошибки при физическом моделировании не стоит, т.к. ошибка репроекции без смещений в реальных случаях значительно выше.

## **2.2 Проведение экспериментальных исследований работоспособности реализованного алгоритма оптимизации**

### **2.2.1 Лабораторная установка**

Экспериментальная установка реализована следующим образом (см. рисунок 36). В установке используется измерительная система LaVision StrainMaster, которая состоит из двух видео камер Imager SX, устройства синхронизации для одновременного получения изображений LaVision PTU и персонального компьютера с программным обеспечением DaVis 8.4.

Камеры Imager SX (Basler piA2400-17gm) основаны на монохромной ПЗС матрице размером 2/3" с разрешением  $2050 \times 2448$  пикселей, размером пикселей  $3,45 \times 3,45$  мкм, разрядностью 12 бит и максимальной скоростью съемки 17 к/с в максимальном разрешении. В качестве объективов использовались Fujinon HF25HA-1B с фокусным расстоянием 25 мм и относительным отверстие 1:1.4.

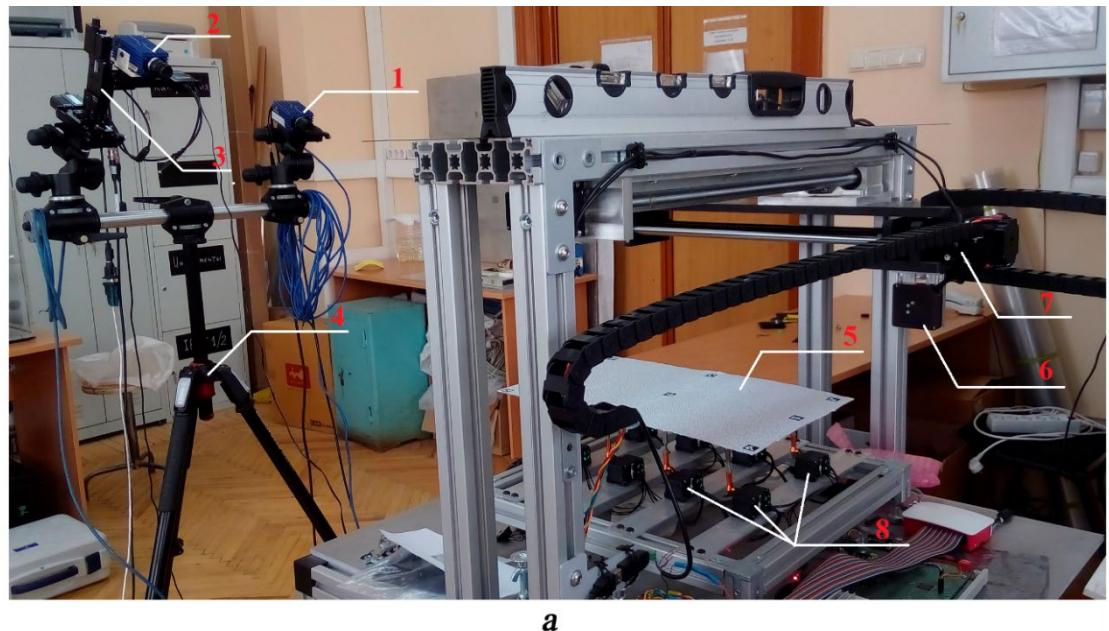
Подвижная камера крепится на 3-х координатный позиционер от компании Zaber. Он состоит из трёх отдельных линейных приводов X-LSM050A со следующими характеристиками: диапазон перемещения 50 мм, погрешность положения 20 мкм и шаг смещения меньше 0,1 мм.

Для сравнения результатов расчёта, полученных с помощью метода IPCT, и действительной формы поверхности объекта используется лазерная измерительная установка, характеристики которой приведены в таблице 6.

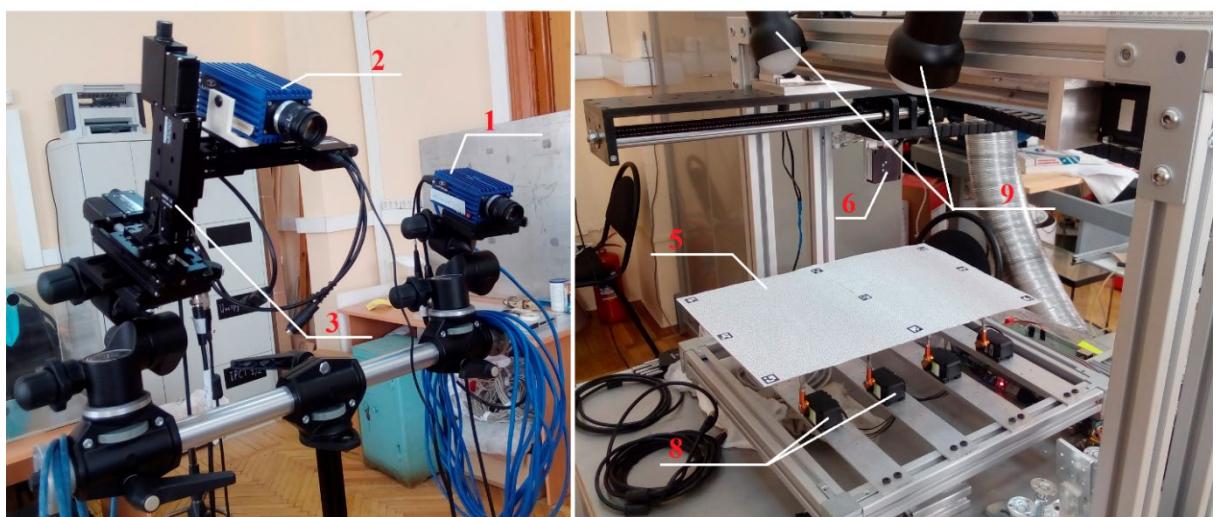
Ближняя граница и диапазон измерений лазерного триангуляционного датчика указан для направления вдоль оси распространения излучения. Для того, чтобы производить измерения в точках по всей поверхности объекта имеется 2-х координатный позиционер, осуществляющий перемещение лазерного датчика с минимальным шагом 0,1 мм. Для изменения формы поверхности объекта используются сервоприводы, закреплённые к обратной стороне. Они позволяют как локально изменять форму поверхности, так и смещать весь объект.

Таблица 6 – Характеристики лазерной измерительной установки

Ближняя граница	50 мм
Диапазон измерений	100 мм
Дискретность	0,01 мм
Погрешность не более	$\pm 0,15$ мм
Размер светового пятна	250 мкм



*a*



*b*

1 – неподвижная камера; 2 – подвижная камера; 3 – 3-х координатный позиционер;  
4 – штатив для камер; 5 – объект измерения; 6 – лазерная измерительная установка;  
7 – 2-х координатный позиционер лазера; 8 – сервоприводы для деформации поверхности  
объекта; 9 – дополнительное освещение

Рисунок 36 – Экспериментальная установка

## 2.2.2 Эксперимент

Для физического моделирования процесса оптимизации декалибровки необходимо решить две проблемы. Первая проблема – это наличие двух различных систем координат отсчёта для внешних параметров камер и величин смещения, задаваемых в 3-х координатном позиционере. Решение этой проблемы выполняется следующим образом:

1) Определяются внешние параметры камер при различных смещениях на позиционере.

2) Составляется система уравнений, где внешние параметры камер, в данном случае только вектор смещения камер, равен смещению на позиционере, умноженном на матрицу преобразования.

3) Выполняется поиск решения системы уравнений итерационным способом, т.к. из-за наличия погрешностей система не будет иметь точного решения. Поэтому минимизируется невязка уравнений методом наименьших квадратов.

В результате решения получается матрица перехода из одной системы координат (координат позиционера) в другую (координат камер).

Вторая проблема – это сопоставление измеренной с помощью камер формы поверхности объекта с измерениями лазерной установкой, т.к. получаемые координаты находятся в разных системах отсчёта. Решение этой проблемы осуществляется методом кросскалибровки, который заключается в следующем:

1) Выполняется съёмка объекта с камер вместе с измерением лазерного датчика. Тем самым получая 2-х мерные координаты в пикселях лазерного пятна на камерах вместе с 3-х мерными координатами с датчика.

2) Измеряется таким образом множество точек на поверхности объекта.

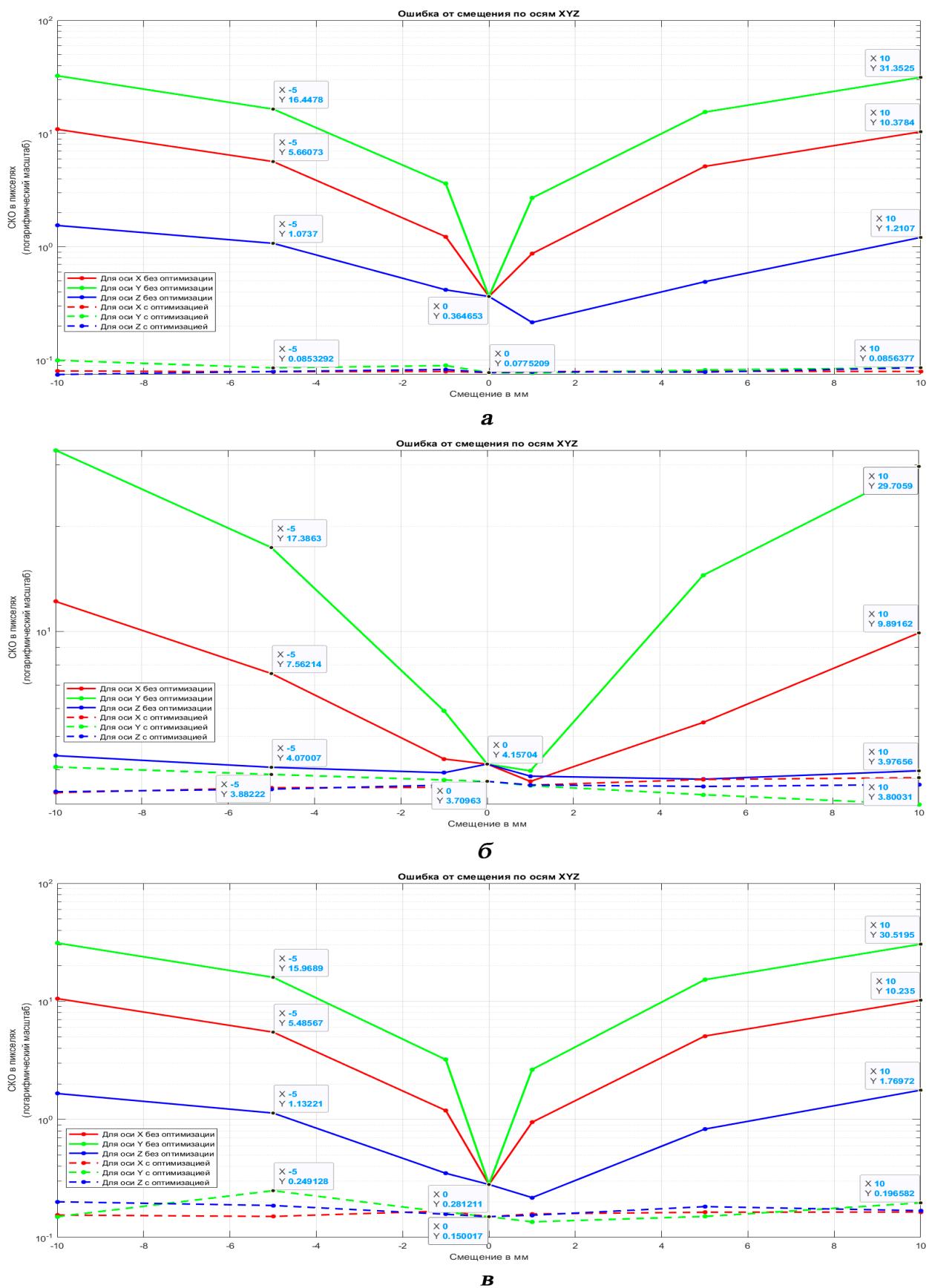
3) По полученным наборам 3-х и 2-х мерных координат, зная внутренние параметры камер, рассчитывается вектора смещения и матрицы поворота для каждой камеры, которые связывают системы отсчёта камер с системой отсчёта лазерного датчика.

4) В дальнейшем эти матрицы используется при триангуляции изображений с камер, чтобы перейти из системы отсчёта камер в систему отсчёта лазерного датчика.

В результате эксперимента построены графики зависимостей ошибки репроекции от смещения одной камеры относительно другой и выполнена оптимизация внешних параметров камер для уменьшения ошибки репроекции (см. рис. 37). На рисунке 37 сплошными линиями указаны ошибки репроекции без оптимизации, а пунктирными – с оптимизацией; красный цвет соответствует смещению вдоль оси  $x$ , зелёный –  $y$ , синий –  $z$ . Ось  $x$  направлена от неподвижной камеры в сторону подвижной камеры, ось  $z$  ориентирована в направлении объекта измерений, соответственно ось  $y$  перпендикулярна осям  $x$  и  $z$  и направлена вниз, т.е. в сторону основания штатива для камер. Измерения выполнены для различных форм поверхности объекта, соответственно 37(а), 37(б) и 37(в). Шаги смещения для всех осей в обе стороны от нуля равны 1, 5 и 10 мм в координатах позиционера. За ноль смещения взято положение на позиционере  $x = 25$  мм,  $y = 25$  мм и  $z = 25$  мм, т.к. диапазон смещений позиционера равен 50 мм.

Из полученных данных можно сделать вывод, что при всех трёх формах поверхности видна одна и та же зависимость ошибки репроекции без оптимизации для осей  $x$ ,  $y$  и  $z$ . Наибольшее влияние вносит смещение вдоль оси  $y$ , немногого меньшее влияние вдоль оси  $x$ , а смещение вдоль оси  $z$  незначительно влияет на ошибку.

Благодаря оптимизации для первой формы поверхности объекта получилось уменьшить ошибку при всех положениях камеры до значений меньших 0,1 пикселя, для второй формы меньше 4,0 пикселя, для третьей формы меньше 0,25 пикселя. Хоть для второй формы поверхности объекта явно не достигнут минимум ошибки, но её уменьшение значительно при смещениях 5 и 10 мм. Для первой и третьей формы поверхности объекта получилось опуститься ниже калибровочных значений ошибки и сохранять близкие к этим значениям ошибки даже при смещениях в 5 и 10 мм.



а) – первая форма поверхности объекта; б) – вторая форма поверхности объекта;  
в) – третья форма поверхности объекта

Рисунок 37 – Зависимости ошибки репроекции от смещений камеры

В приведённых выше результатах использовался, как и в компьютерном моделировании, метод оптимизации Нелдера-Мида совместно с методом штрафных функций, где ограничения по поиску смещений относительно начала поиска  $\pm 10$  мм, а поворота  $\pm 2^\circ$ . При максимальном количестве итераций в одной попытке 500 и количеством попыток равным 12. При оптимизации использовались все точки на изображении, который получались в результате кросскорреляции.

По рисункам 35 и 37 можно сказать, что смоделированные зависимости ошибки репроекции и экспериментальный зависимости схожи, что говорит о верной реализации модели. Оптимизация Нелдера-Мида совместно с методом штрафных функций успешно минимизирует ошибки репроекции, даже при значительных смещениях. Это свидетельствует об устойчивости и применимости реализованного алгоритма.

Из рисунка 36 можно заключить, что реализованный алгоритм успешно справляется с уменьшением ошибки репроекции. Также отмечу, что если будет необходимость ускорить работу алгоритма, то можно использовать не все точки изображения для оптимизации [50] и пробовать уменьшать количество попыток поиска решения.

## **2.3 Проведение экспериментальных исследований работоспособности нейросети**

### **2.3.1 Результаты применения нейросети**

Исследования о возможности восстановления формы трёхмерной поверхности на основе метода кросскорреляции с использованием стереопары уже проводились и называются методом корреляции фоновых изображений (IPCT), который является другим приложением метода PIV для измерения деформаций твёрдой поверхности.

Одним из первых предложений по применению нейросетей к задачам PIV предложено в работе [81]. 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  $32 \times 32$  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 [69], где изучалась сеть с архитектурой, предложенной в [82].

В разделе 2.3 посвящённой обзору и выбору методов машинного обучения для задачи измерения формы поверхности была выбрана нейросеть PIV-LiteFlowNet-en, которая основана на сети LiteFlowNet. Обе эти сети являются потомками сети FlowNet [67], предназначеннной для приема двух изображений на вход и оценки смещения оптического потока на выходе. По данным [67], 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 [71].

В изученных работах ссылались на реализации сетей на библиотеке caffe, с которой возникло много технических трудностей. Поэтому произведён повторный поиск реализаций этих сетей на библиотеке pytorch. Для дальнейшего

исследования и сравнения нейросети с кросскорреляционным методом будут использоваться две сети как PIV-LiteFlowNet-en, так и LiteFlowNet. Так как вторая сеть позволяет оценивать большие смещения, что является важным фактором. И это 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 11,17 GiB памяти GPU, возможно обработать изображения размером  $1900 \times 2000$  пикселей с глубиной цвета 24 бит формата jpg. Чего достаточно для обработки изображений в полном разрешении, если выделять рабочую область изображения. Пример выделения рабочей области показан на рисунке 38.

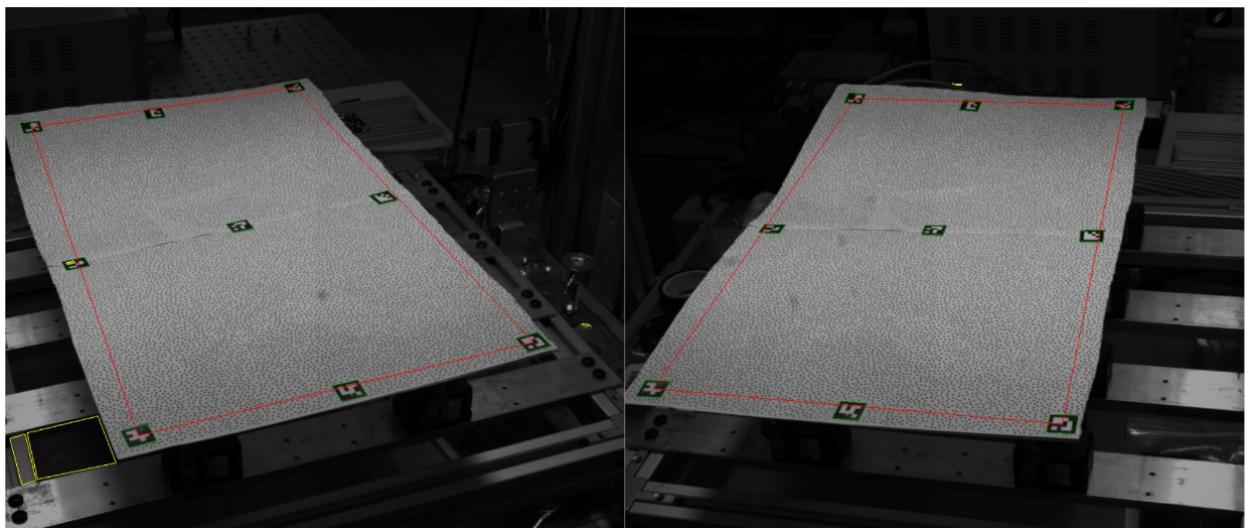
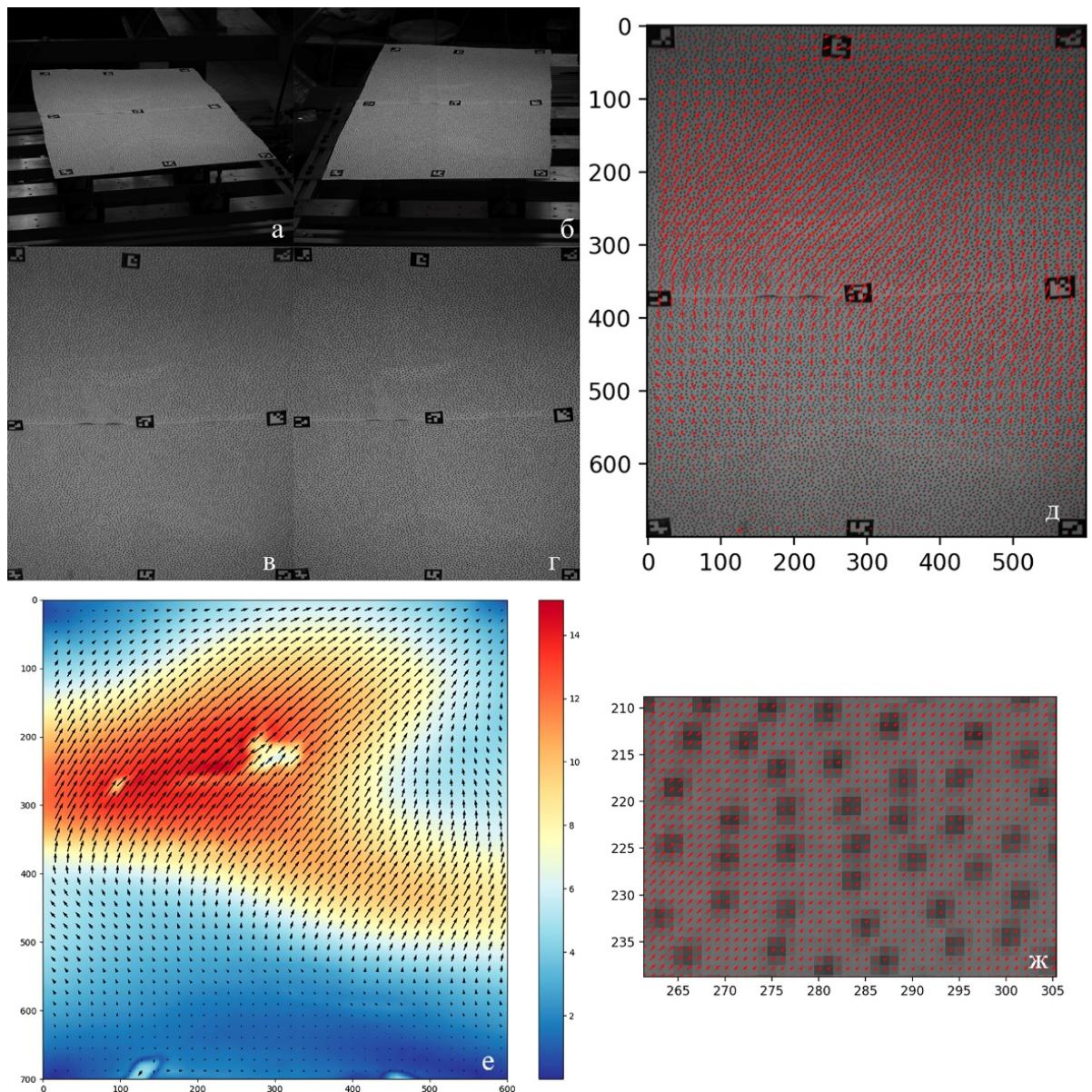


Рисунок 38 – Пример определения рабочей области по маркерам ArUco. Разрешения камер  $2050 \times 2448$  пикселей

На рисунке 39 приведена обработка стереопары сетью PIV-LiteFlowNet-en. Геометрические преобразования на рисунках 39(в) и 39(г) необходимы для выделения рабочей области и улучшения обработки. На рисунке 39(е) в области

больших смещений видны два светлых пятна. Это является следствием слишком больших смещений, для которых нейросеть не приспособлена.



а), б) – исходные изображения; в), г) – изображения с использованием геометрических преобразований; д) – визуализация разреженного векторного поля; е) – представление амплитуды деформаций при помощи цветовой карты; ж) – визуализация векторного поля в полном разрешении

Рисунок 39 – Пример обработки стереопары, все отсчёты измеряются в пикселях

### 2.3.2 Замеченные особенности

Прежде чем сравнивать сети и кросскорреляцию, необходимо определить равнозначные условия для этих методов. Как можно заметить на рисунке 39, уже видно, что сеть PIV-LiteFlowNet-en не способна оценивать большие смещения. Для наглядности на рисунке 40 приведены векторные поля для одной и той же

стереопары, но при разных разрешениях этих изображений. Из рисунка 40 видно, что сеть PIV-LiteFlowNet-en не справляется со смещениями больше  $\sim 12\text{-}13$  пикселей. Для сети LiteFlowNet этот показатель  $\sim 80\text{-}90$  пикселей.

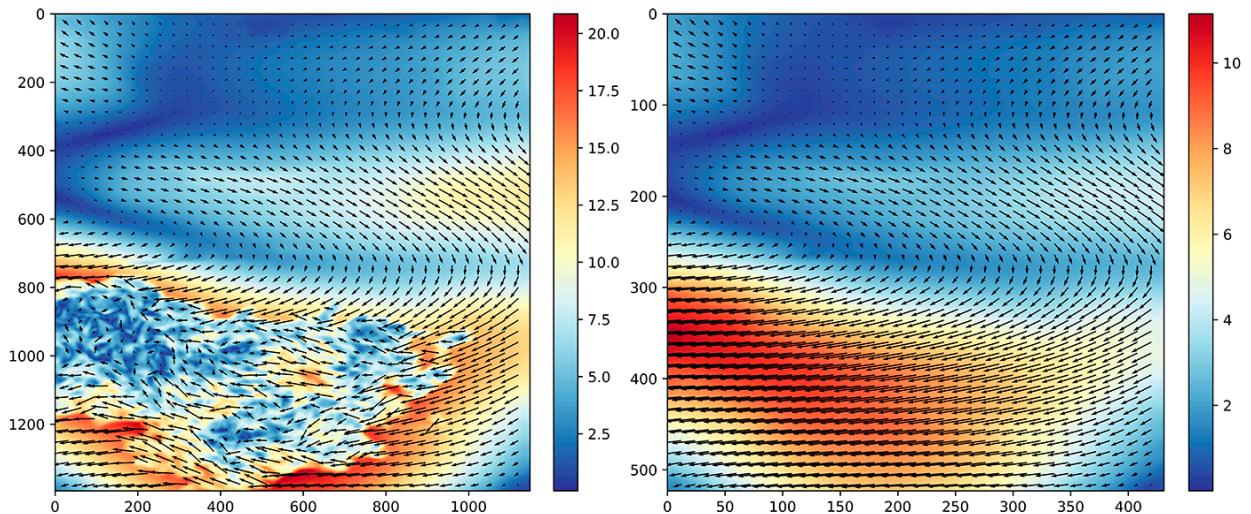


Рисунок 40 – Сравнение векторных полей для изображений разного разрешения; все отсчёты измеряются в пикселях

## 2.4 Экспериментальное сравнение нейросетей с кросскорреляционным анализом

### 2.4.1 Сравнения в смежной области – PIV

В разделе 2.3, которая посвящена обзору методов машинного обучения для задачи расчёта трёхмерной поверхности по стереопаре, выбран алгоритм машинного обучения под названием PIV-LiteFlowNet-en [72]. Также для сравнения и лучшего понимания разницы между различными сетями в текущей работе взята нейросеть LiteFlowNet [76], которая является прародителям выбранной сети. Разница между этими нейросетями уже описана ранее, но о некоторых особенностях необходимо будет ещё раз упомянуть.

В текущей главе будет произведено сравнение двух нейросетей и метода на основе кросскорреляции собственной разработки [83]. Первое, что необходимо сделать, это провести анализ точности всех трёх методов в смежной задаче – велосиметрии по изображениям частиц (PIV). Методы на основе кросскорреляции и PIV-LiteFlowNet-en изначально разрабатывались для этой задачи.

Второе, провести лабораторные измерения для получения набора данных по текущей задаче. И по полученному набору данных выполнить сравнение точности трёх методов между собой, выявить особенности и ограничения применимости этих методов на экспериментальных данных.

Прежде чем оценить максимальное рассчитываемое смещение нейросетям сравним их с кросскорреляцией на примере данных, предложенных авторами работы [71]. Этот набор является моделированием потоков с разными условиями и параметрами для задач PIV. Приблизительно по этим данным обучалась сеть PIV-LiteFlowNet-en. Но, несмотря на это, такое сравнение даст представление о точности сетей и программного комплекса собственной разработки, основанного на кросскорреляции, из работы [83]. В таблице 7 представлено описание и количество пар изображений, на которых выполнено тестирование. Все изображения имеют разрешение  $256 \times 256$  пикселей, представляют из себя белые точки на чёрном фоне, заданные гауссовым распределением (см. рисунок 41).

Таблица 7 – Набор данных для задач PIV [71]

Вид потока	Описание	Условия	Количество
Back-step	Поток со смещением	Re = 800	600
		Re = 1000	600
		Re = 1200	1000
		Re = 1500	1000
Cylinder	Потока вокруг цилиндра	Re = 40	50
		Re = 150	500
		Re = 200	500
		Re = 300	500
		Re = 400	500
DNS-turbulence	Однородный и изотропный турбулентный поток	/	2000
JHTDB-channel	Поток в канале от JHTDB	/	1900
JHTDB-channel hd	Изотропная турбулентность от JHTDB	/	600
JHTDB-isotropic 1024 hd	Изотропная турбулентность от JHTDB	/	2000
JHTDB-mhd 1024 hd	Турбулентность от JHTDB	/	800
SQG	Поверхность моря задаваемая моделью SQG	/	1500
Uniform	Равномерный поток	$ dx  \in [0, 5]$	1000

На рисунке 41 представлена визуализация различных видов потока при максимальном моделируемом смещении для каждого случая. В остальных случаях

амплитуда смещения потока меньше. Такое визуальное представление даёт качественную оценку о возможности сети рассчитывать различного вида деформации на изображениях. Хоть по рисунку 40 видно максимально смещение для каждого случая потока, но для лучшего представление о поведении потоков в наборе данных рассчитано максимальная амплитуда вектора для каждой пары изображений. Представлено это на рисунке 42. Из первого графика наглядно видно, что максимально смещение составляет  $\sim 11$  пикселей и только для случая JHTDB-isotropic 1024 hd, в остальных случаях потолок смещения ниже. Этот график является сортировкой по возрастанию смещения из остальных графиков. Все графики кроме первого будут нужны после сравнения точности трёх методов, так как максимум смещения на изображении коррелирует с точностью оценки.

Параметры обработки методом [83]: размер окна опроса  $24 \times 24$  пикселя, смещения окна опроса 4 пикселя, аппроксимация корреляционного пика гауссовым распределением. Из рисунка 43 видно, что нейросеть PIV-LiteFlowNet-en имеет лучшую точность и малый разброс ошибки для всех рассмотренных случаев потока. Сеть LiteFlowNet, которая не обучалась для задач PIV, хоть и имеет большую ошибку, но показывает хорошие результаты, что говорит о её универсальности для различных приложений. Кросскорреляционный метод в целом показал худшие результаты, за исключением случая “Cylinder”, при этом он имеет существенный разброс ошибки. Есть ещё один минус кросскорреляционной обработки – это векторное поле меньшей плотности. В то время как нейросети получают поле разрешением равным размеру входному изображению, то для кросскорреляции поле имеет в 4 раза меньшую плотность, не считая отступа от границ. Это связано с шагом между окнами опроса в 4 пикселя, а отступ от края изображения является ограничением метода. Но у кросскорреляции есть плюс в том, что она полностью выполняется на CPU, а нейросетям для достижения преимущества в скорости обработки необходимо использовать GPU.

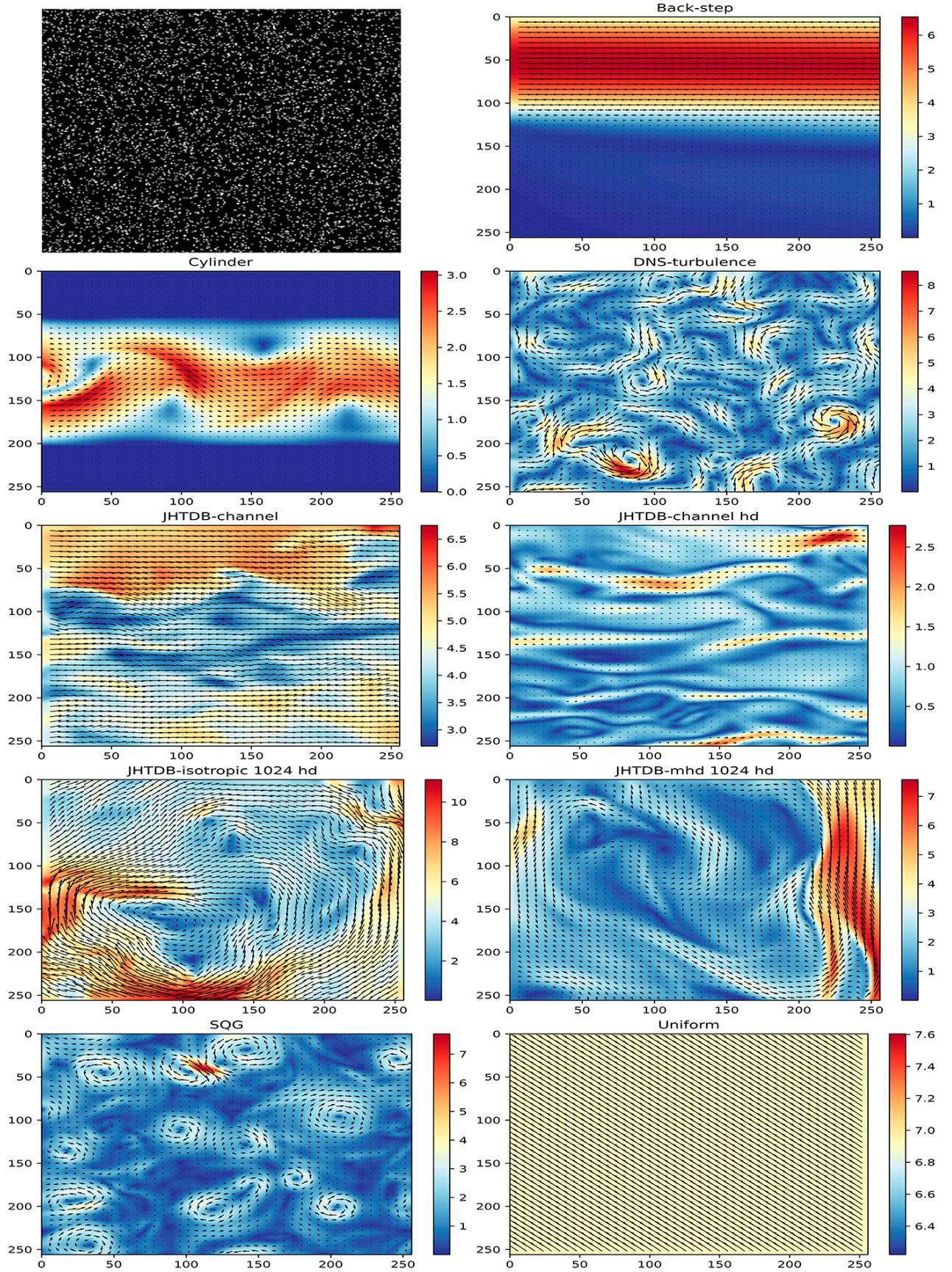


Рисунок 41 – Первая картинка – пример моделируемого потока; остальные картинки – визуализация видов потока из базы данных [71] при максимальном смещении для каждого случая; все отсчёты измеряются в пикселях

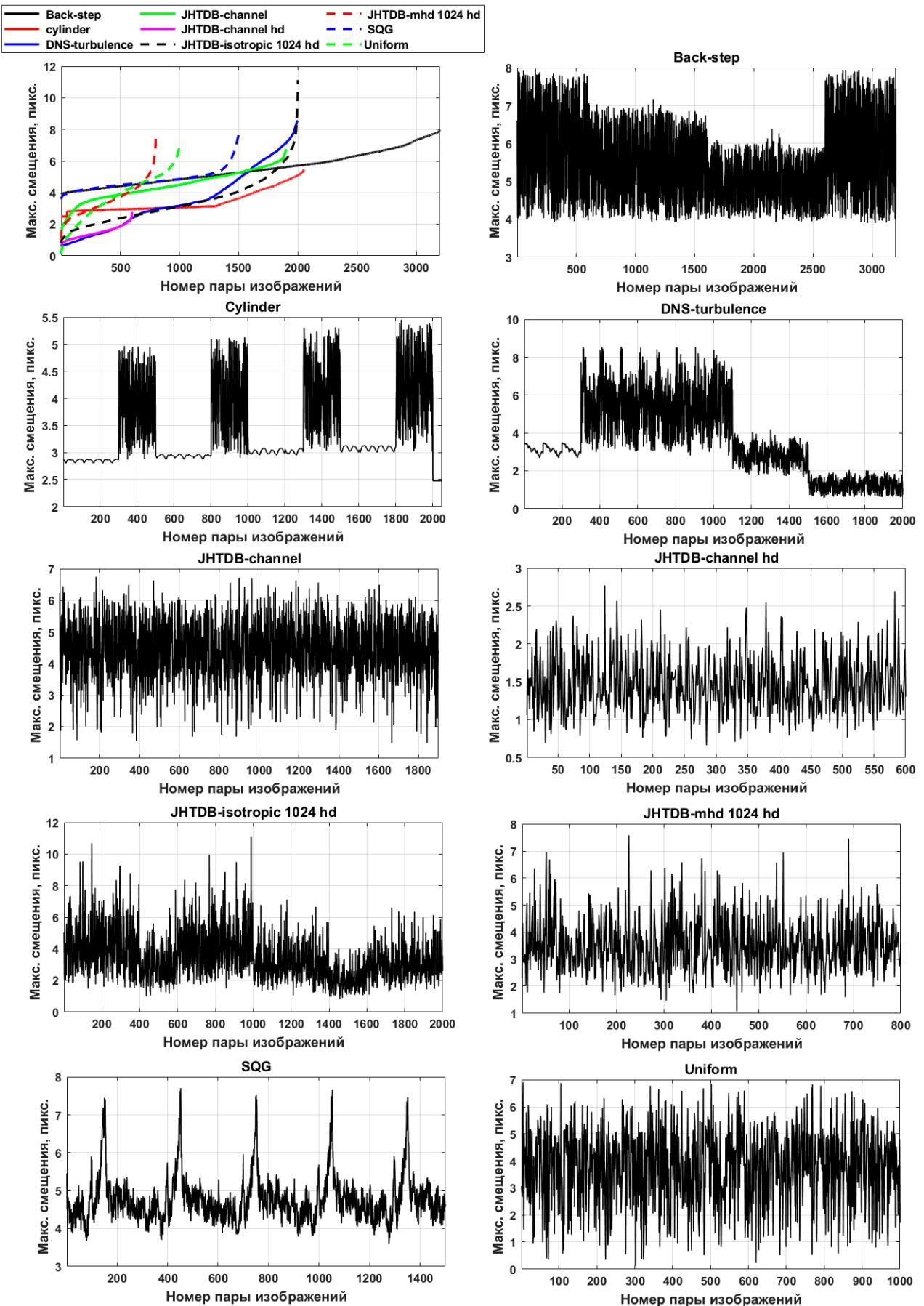


Рисунок 42 – Первый график – максимумы смещения, сортированные по возрастанию, остальные графики – максимумы смещений потока для каждой пары изображений

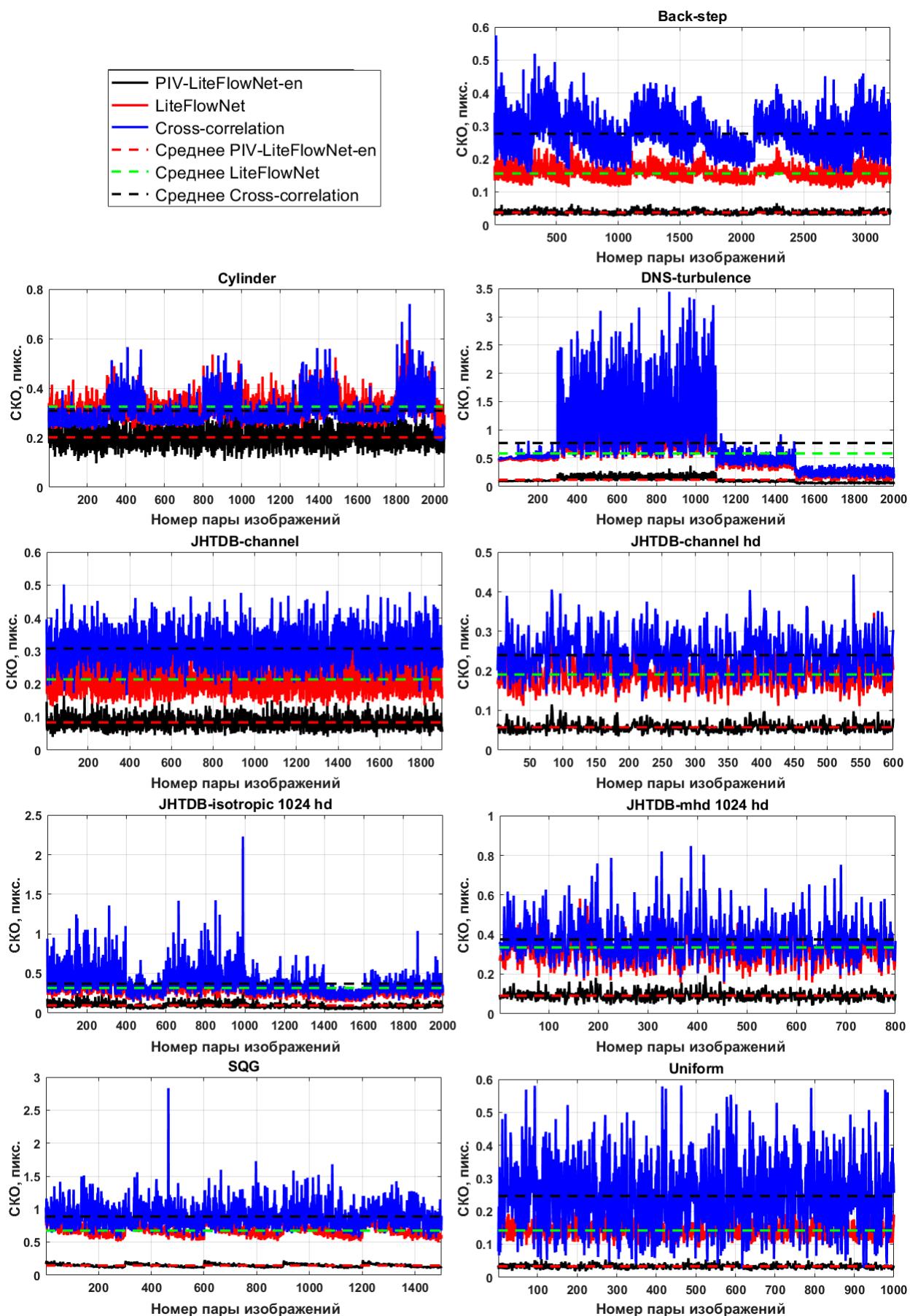


Рисунок 43 – Среднее квадратичное отклонение (СКО) между измеренным векторным полем и истинным для всех трёх исследуемых методов

Сравнивая графики на рисунках 42 и 43 можно заметить, что чем выше смещение потока по амплитуде, тем выше погрешность оценки этого смещения независимо от выбранного метода.

Произведено сравнение трёх методов определения смещений на изображениях, два из которых используют нейросети LiteFlowNet и PIV-LiteFlowNet-en, а третий метод основан на кросскорреляции. Критерием для оценки точности алгоритмов выступил набор данных для задач PIV [71]. Лучшие результаты по точности и разбросу ошибки показала сеть PIV-LiteFlowNet-en. Результат алгоритма, основанного на кросскорреляции, оказался худшим, но близким по точности к сети LiteFlowNet, хоть и имеет больший разброс ошибки, при этом из-за ограничений алгоритма плотность векторного поля меньше, чем у нейросетей.

#### **2.4.2 Поиск равнозначных условий сравнения**

Для проверки применимости нейросетей, в исследуемой области, сняты на лабораторной установке 150 пар изображений поверхности, эмитирующей форму поверхности крыла самолёта. Измерения производились при случайном положении сервоприводов, которые деформируют поверхность.

Для того, чтобы определить при каком разрешении изображений для измеренной выборки разумно обрабатывать изображения сетями решено провести расчёт СКО ошибки репроекции для 10 разрешений каждой стереопары из 150 и затем усреднить получение оценки по количеству стереопар.

Так сети не могут обрабатывать не прямоугольные изображения, то для каждой стереопары выполняются геометрические преобразование (см. рисунок 39), размер получаемого изображения определяется большей высотой и шириной из областей выделяемой по маркерам ArUco. В результате размер каждой стереопары индивидуален, но обычно не превышает  $1700 \times 1500$  пикселей.

Расчёт высоты и ширины сжатых изображений вёлся по следующей формуле

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

где  $k = 0,1,2\dots9$ ;

$R_0$  – исходный размер стороны изображения.

В результате, считая, что размеры изображений не превышают  $1700\times1500$  пикселей, получим следующий набор размеров

высота:  $1700 \rightarrow 1530 \rightarrow 1360 \rightarrow \dots 170$

ширина:  $1500 \rightarrow 1350 \rightarrow 1200 \rightarrow \dots 150$

при этом сохраняется соотношение сторон изображения.

На рисунке 44 представлен результат обработки набора данных. Каждый график является усреднением по 150 стереопарам. Для проверки влияния общей интенсивности на изображении на основе каждой пары изображений создавалась инверсная версия этой пары, то есть определялся максимум на изображении и из него вычиталось значение пикселя, пример на рисунке 45. В среднем максимум на изображениях составил 170, а минимум 10, обусловленный шумами, тогда динамический диапазон изображений составляет 160.

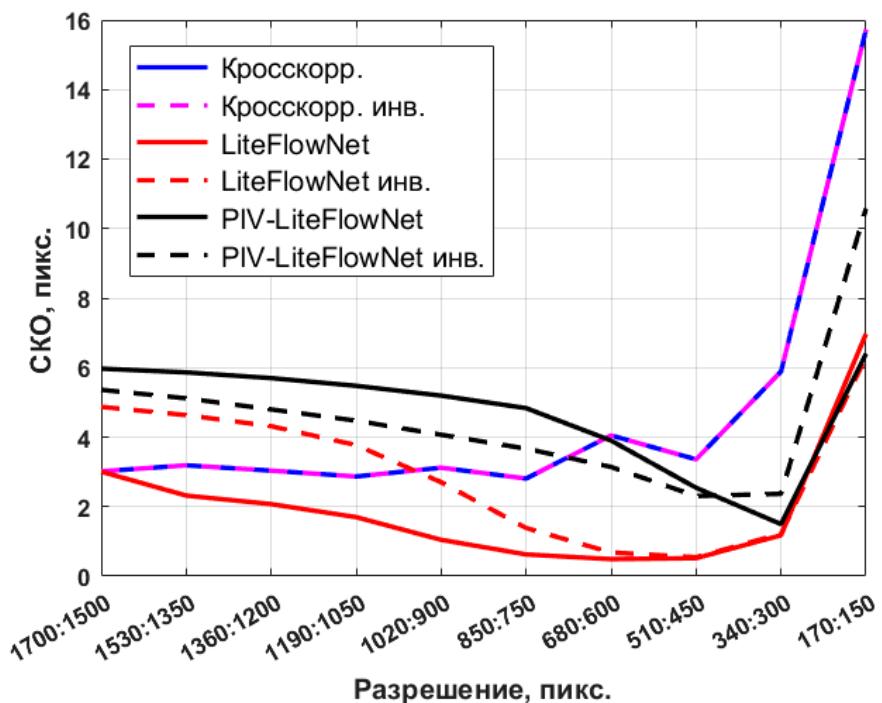
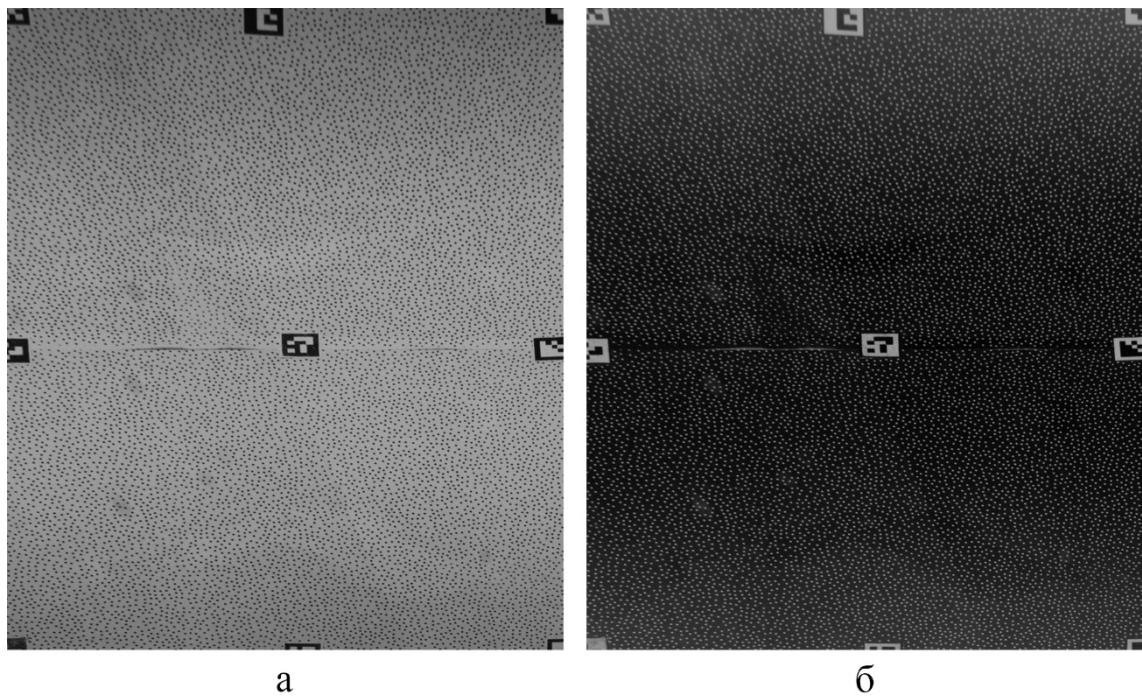


Рисунок 44 – Среднее значение СКО ошибки репроекции для разных разрешений по трём исследуемым методам



а) – исходное изображение; б) – инверсное изображение

Рисунок 45 – Пример инверсии изображения

Для метода на основе кросскорреляции необходимо определить параметры для обработки изображений. Универсальные параметры выбрать невозможно, т.к. смещения могут достигать 100 и больше пикселей, тогда разумно выбрать размер окна опроса 200-300 пикселей, но в этом случае для изображений малого разрешения не будет построено ни одного вектора. Поэтому размер окна опроса необходимо определять индивидуально для каждого разрешения.

Обработка методом на основе кроскорреляции производилась с параметрами, приведёнными в таблице 8.

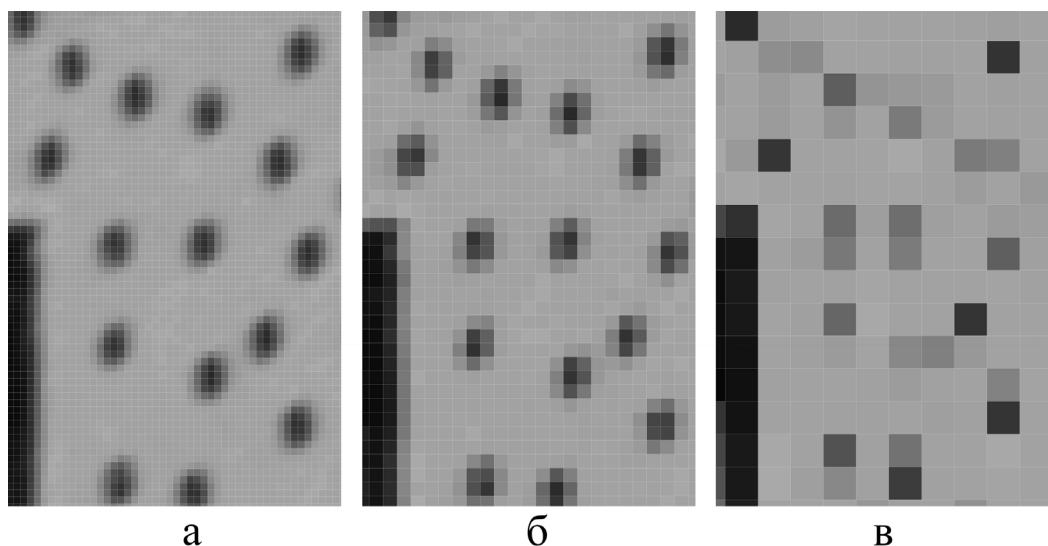
Таблица 8 – Параметры обработки изображений кроскорреляцией

Разрешение, пикс.	Размер окна опроса, пикс.	Смещение окна опроса, пикс.
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

Из рисунка 44 можно сделать следующие выводы:

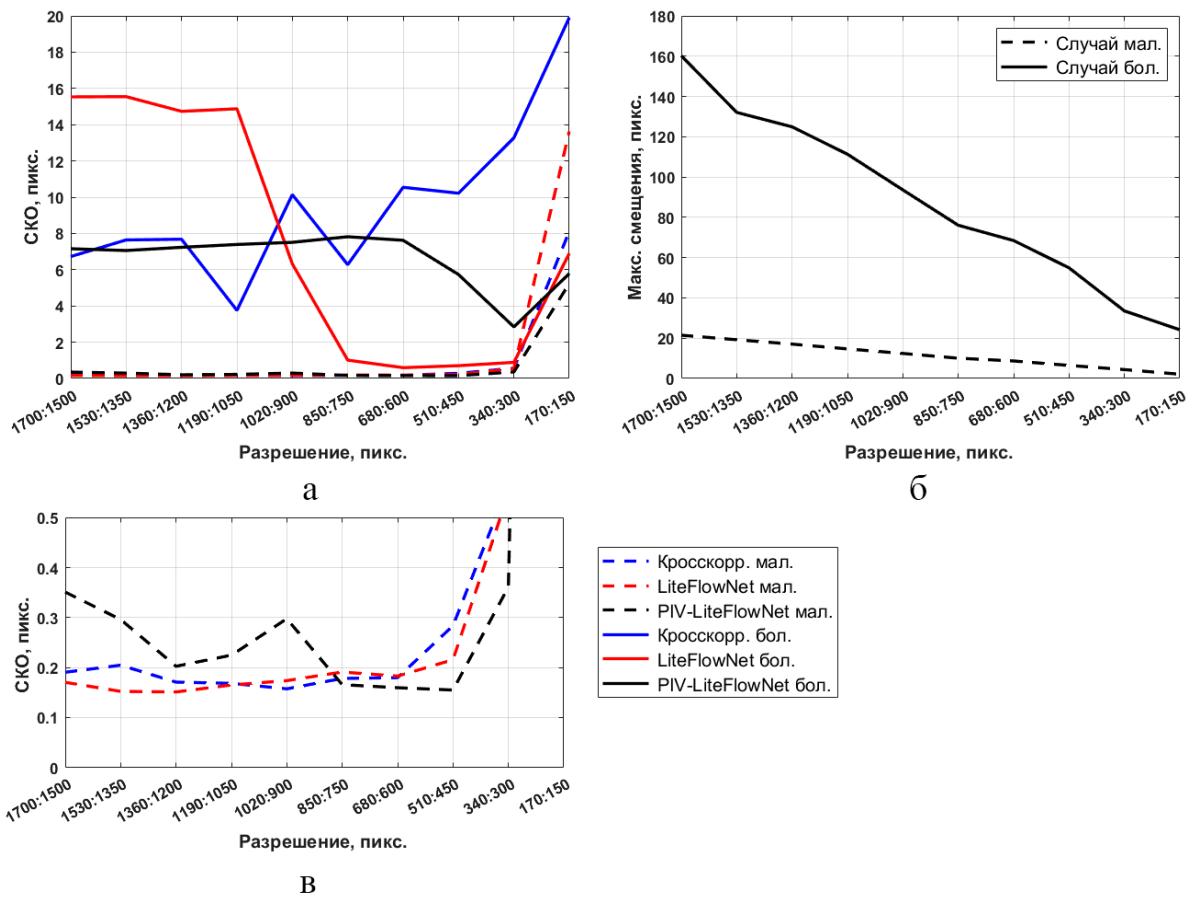
- 1) Метод кросскорреляции показывает стабильную ошибку репроекции почти для всех разрешений. При этом на метод не влияет инверсия интенсивности.
- 2) Метод LiteFlowNet показывает лучшие результаты среди всех методов, а инверсия изображения ухудшает его работу.
- 3) Метод PIV-LiteFlowNet-en явно не справляется из-за больших смещений на изображениях. Доходя до разрешения  $340 \times 300$ , видимо попадет на достаточно малые смещения, но здесь уже вступает погрешность из-за сильного сжатия изображения, что не позволяет методу добиваться высокой точности. Инверсия изображений улучшает работу метода.
- 4) При разрешении в  $340 \times 300$  и  $170 \times 150$  все методы показывают ухудшение точности из-за сильного сжатия изображения.

На рисунке 46 приведён пример области изображения при различных разрешениях. Видно, что при высоком сжатии, даже если кроскорреляция или сеть определяют смещения на изображениях с высокой точностью, то центры чёрных точек при высоком сжатии могут не совпадать с центрами этих точек без сжатия. Поэтому при таких условиях невозможно добиться низкой ошибки.



а) – высокое разрешение; б) – промежуточное разрешение; в) – низкое разрешение  
Рисунок 46 – Часть изображения при разных разрешениях

Для лучшего понимания поведения методов на рисунке 46 приведены графики СКО ошибки репроекции для двух стереопар: при малых и больших смещениях на изображениях. На рисунке 47(а) графики в одной плоскости для малых и больших смещений, 47(б) величина этого смещения для каждого разрешения, 47(в) увеличенная область 47(а), демонстрирующая поведение методов 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.



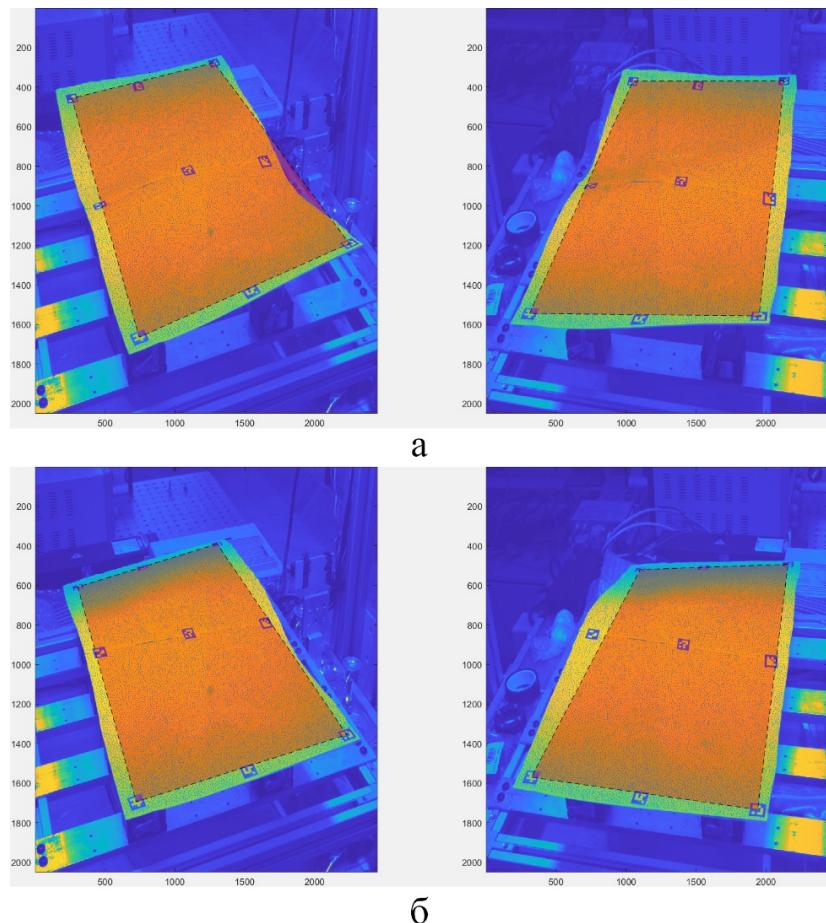
а) – СКО ошибки репроекции для случая малых и больших смещений; б) – величина смещения для двух случаев; в) – увеличенная область графика 47(а), для демонстрации поведения методов при малых смещениях

Рисунок 47 – Частные случаи обработки стереопар

Для метода кросскорреляции оценка максимального смещения зависит от размера окна опроса, если размер 256 пикселей, то максимум оцениваемого смещения на практике составляет 128 пикселей, но для перестраховки некоторые

исследователи рекомендуют размер окна опроса в 3 раза, превышающего максимум смещения. В таком случае максимум оцениваемого смещения будет составлять  $\sim 86$  пикселей.

По графикам 47(б) и 47(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 соответствует заключению из предыдущей главы, где производились сравнения на данных PIV. Среднее СКО ошибки репроекции на рисунке 44 для всех методов довольно велик, т.е. близок или больше 1 пикселя. Это объясняется тем, что при больших смещениях, как на рисунке 47, ошибка резко возрастает, поэтому средний уровень ошибки подымается. Но есть второй фактор, связанный с особенностями выделения рабочей области на изображениях (см. рис. 38). Пример подобной проблемы показан на рисунке 48. На левом фото рисунка 48(а) видно, что при



а) – 1-ый случай; б) – 2-ой случай; все отсчёты в пикселях  
Рисунок 48 – Частные случаи выделения рабочей области

выделении по обнаруженным маркерам ArUco область вышла за пределы измеряемой плоскости, что в результате поиска соответствий на изображениях приведёт к ошибочным векторам в этой области. Подобная ситуация характерна и для рисунка 48(б), только уже на правом фото в результате выделения не в область не попала существенная часть измеряемой плоскости.

Решить это проблему возможно выделением области по большему количеству маркеров, например, 6-ти для данного случая, получая в результате не четырёхугольных, шести угольник, по которому уже выполнять геометрические преобразования.

В итоге, после следующих операций:

- 1) Определения маркеров на изображениях.
- 2) Геометрического преобразования этих изображений.
- 3) Обработки изображений одним из трёх выше описанных методов.
- 4) Обратных геометрических преобразований.

Получаем векторное поле, представленное на рисунке 49, которое показывает какие области изображений соответствуют друг другу.

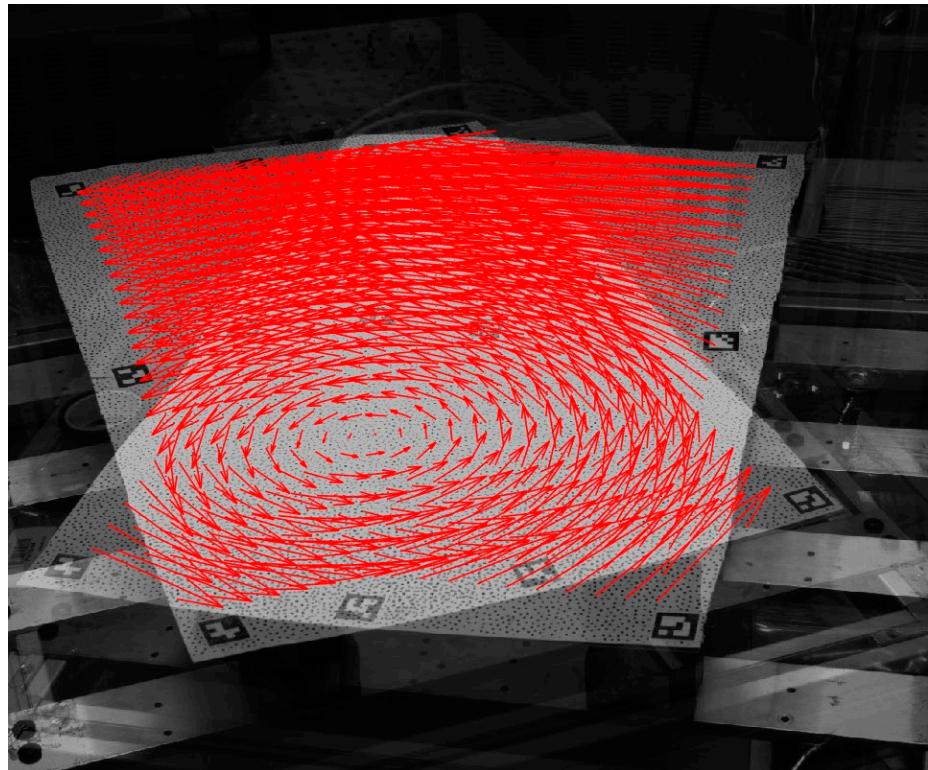


Рисунок 49 – Результирующее векторное поле (размер изображения 2050×2448 пикселей)

По рисунку 49 видно, что после обратных геометрических преобразований длина векторов может достигать значений в 800 пикселей и даже больше. Поэтому обрабатывать изображения без предварительной подготовки существующими методами невозможно для данной задачи.

#### 2.4.3 Результаты сравнения

Исходя из графиков на рисунке 44 минимум ошибки для сети PIV-LiteFlowNet-en достигается при разрешении 510:450, для сети LiteFlowNet при 680:600. На рисунках 50, 51 и 52 представлены графики зависимости ошибки репроекции для каждой пары изображения из 150 снятых. Также для сравнения

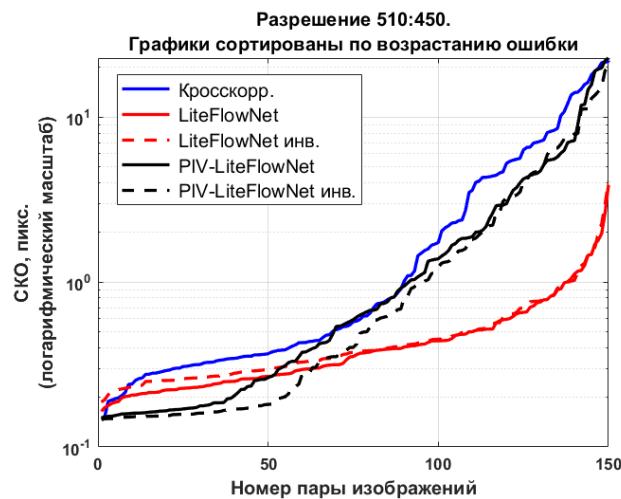


Рисунок 50 – Графики СКО ошибки репроекции для каждой снятой пары изображений и для трёх исследуемых методов при разрешении 510:450 пикселей

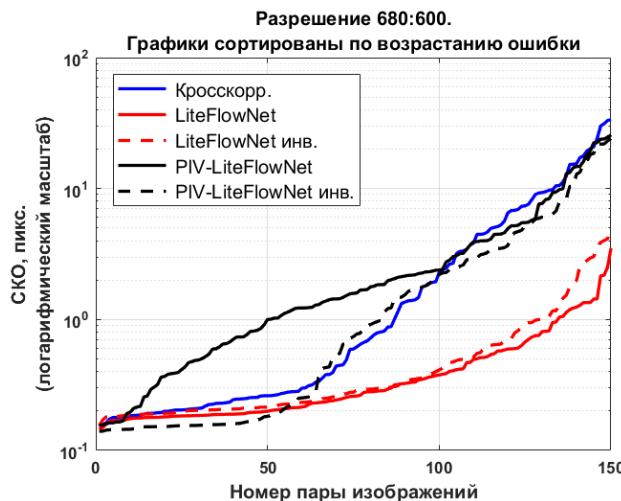


Рисунок 51 – Графики СКО ошибки репроекции для каждой снятой пары изображений и для трёх исследуемых методов при разрешении 680:600 пикселей

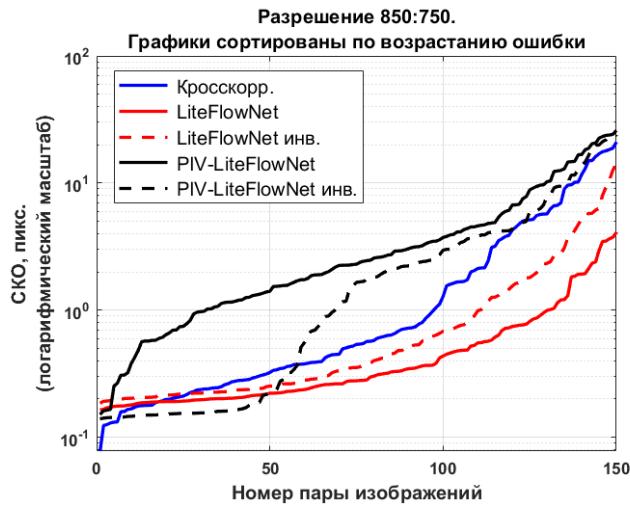


Рисунок 52 – Графики СКО ошибки репроекции для каждой снятой пары изображений и для трёх исследуемых методов при разрешении 850:750 пикселей

приставлено разрешение 850:750 на рисунке 52. Все графики для наглядности сортированы в порядке возрастания ошибки. На всех рисунках прослеживаются одинаковые закономерности. Сеть LiteFlowNet имеет меньшую ошибку в сравнении с кросскорреляционным методом при примерно одинаковой форме графика, лишь нескольких случаях не более 5-8 кросскорреляция превосходит по точности сеть. Сеть PIV-LiteFlowNet-en показывает лучшую точность среди всех методов примерно в 50 случаях, при этом имеет лучше показатели при инверсии изображений. Исключением является случай при разрешении 510:450, где случай инверсии изображений не сильно отстает от исходных изображений.

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 методы.

Для демонстрации того, что не одни из методов не способен с высокой точностью обработать изображения в полном разрешении приведён рисунок 53, 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. Предложения по решению этой проблемы описаны в разделе 2.4.5.

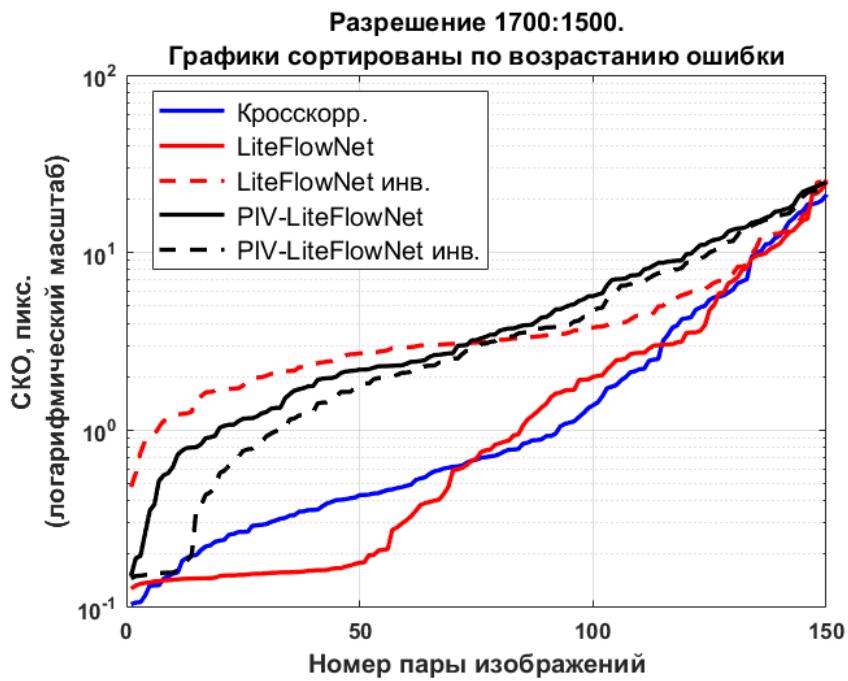
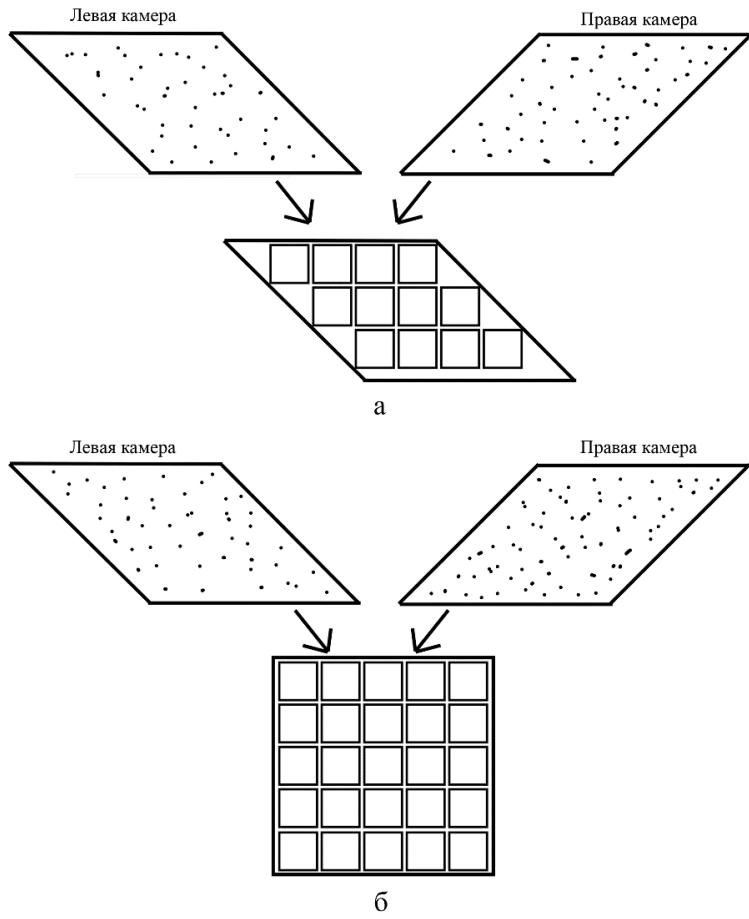


Рисунок 53 – Графики СКО ошибки репроекции для каждой снятой пары изображений и для трёх исследуемых методов при разрешении 1700:1500 пикселей

В используемом для кросскорреляции программном комплексе [83] применяется одно геометрическое преобразование (см. рисунок 54(а)), т.е. приводится изображение с правой камеры к левой по ArUco маркерам, а потом полученная произвольной формы четырёхугольная область разбивается на окна опроса, что приводит к прерывистому векторному полю на границе измеряемой поверхности. Чтобы исправить этот нежелательный эффект возможно выполнять два геометрических преобразования (см. рисунок 54(б)), т.е. приводить изображения с левой и правой камеры к среднему изображению, которое имеет прямоугольную форму (см. рисунок 39).

Так как выполнение геометрического преобразования вносит погрешность в результирующие векторное поле, то необходимо проверить не ухудшит ли точность выполнение двух геометрических преобразований вместо одного. Результат сравнения представлен на рисунке 55. Как видно по рисунку 55 выполнение двух геометрических преобразований не то, что увеличивает погрешность, а даже уменьшает ошибку репроекции векторного поля. Поэтому подобный метод улучшит точность программного комплекса [83] и упростит процесс разбиения изображения на окна опроса при кросскорреляции.



а) – одно геометрическое преобразование; б) – два геометрических преобразования

Рисунок 54 – Схема разбиения на окна опроса при различных геометрических преобразованиях

В результате проведённой работы выполнено сравнение трёх методов нейросетей PIV-LiteFlowNet-en, LiteFlowNet и метода на основе корреляции в задаче восстановления трёхмерной поверхности по стереопаре. Сначала проведена оценка этих методов в смежной области – PIV, где лучший результат показала сеть PIV-LiteFlowNet-en. Далее получен набор данных для исследуемой задачи, по которому, путём предварительной подготовки изображений, получены графики зависимости ошибки репроекции для всех трёх методов. Если не учитывать особенности и ограничения применимости каждого из методов, то лучшую оценку показал метод LiteFlowNet.

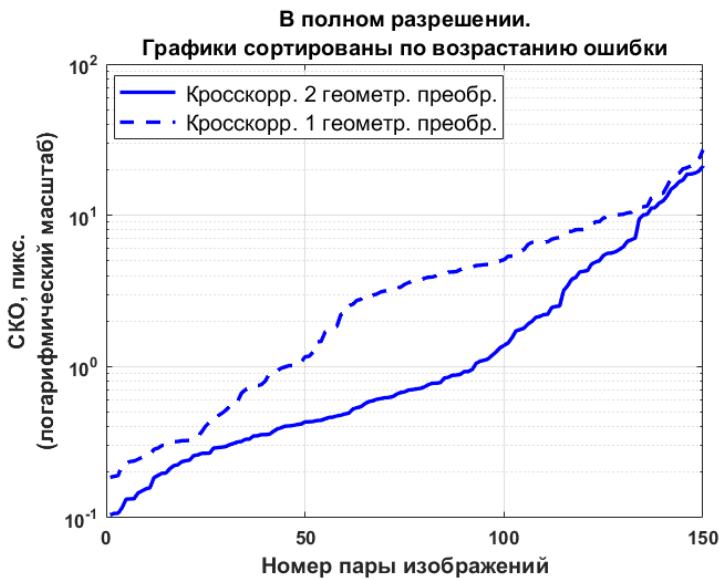


Рисунок 55 – Графики СКО ошибки репроекции для кросскорреляционного метода в случаях одного и двух геометрических преобразований

Стоит сказать о ограничениях применимости каждого из методов. Для PIV-LiteFlowNet-en максимум оцениваемого смещения составил ~12-13 пикселей, при этом сеть показывает лучшие результаты для инверсных изображений. LiteFlowNet верно оценила вектора длиной до ~80 пикселей и показала худшие результаты для инверсных изображений. Максимум оцениваемого смещения для метод на основе кросскорреляции зависит от выбранного размера окна опроса, но чем больше окно опроса, тем медленнее алгоритм.

Важной особенностью нейросетей в отличии от кросскорреляции является векторное поле более высокой плотности, при равной скорости обработки, что позволяет восстановить поверхность с большим разрешением. Поэтому использования сетей в исследуемой области является перспективным направлением.

#### 2.4.4 Предположения по ограничению сетей

Опишем предположений по поводу ограничения максимального оцениваемого смещения сетями. На рисунке 56 [67] приведена архитектура сетей FlowNetS, которая входит в сеть LiteFlowNet. На примере её архитектуры можно оценить максимально рассчитываемое смещение.

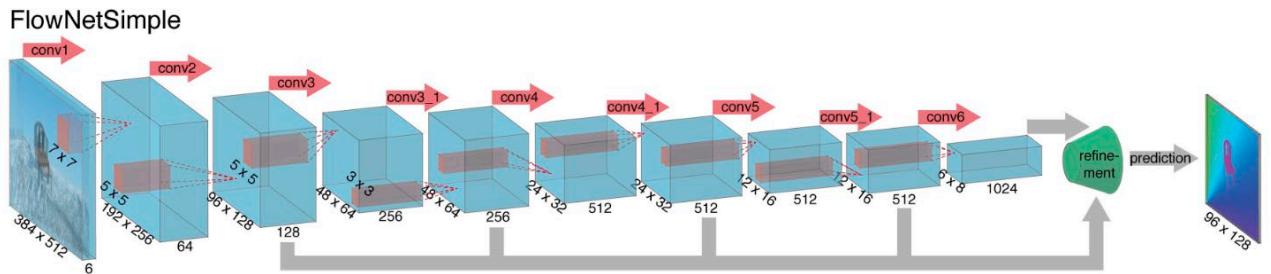


Рисунок 56 – Архитектура сети FlowNetS входящей в сеть LiteFlowNet [67]

На рисунке 56 видно, что исходное изображение разрешением  $384 \times 512$  в конечном итоге сжимается до размеров  $6 \times 8$ , но нам интересен предпоследний слой сети, на котором выполняется свёртка 6 (conv6). Его разрешение равно  $12 \times 18$ , что соответствует сжатию в 32 раза, при этом на этом слое сети выполняется свёртка с размером окна опроса  $3 \times 3$ . Если учесть сжатие изображения, то получается окно опроса соответствует области на исходном изображении  $96 \times 96$  пикселей, то есть максимально оцениваемое смещение на последнем слое сети со свёрткой, при идеальных условиях, составляет 96 пикселей. Так же обстоят дела и со слоем перед conv5, только сжатие изображения уже в 16 раз, максимум оцениваемого смещения на этом слое 48 пикселей. Сложим подобные оценки по всем слоям, где происходит сжатие:

$$|3 \times 3| + |3 \times 3| + |3 \times 3| + |5 \times 5| + |7 \times 7| \\ 96 + 48 + 24 + 20 + 7 = 195 \text{ пикселей.}$$

Казалось бы, что максимум оцениваемого смещения сетью равен 195 пикселей, но по факту получается не более 90 пикселей. Это объясняется тем, что самый большой вектор, который оценивается, на одном слое сети равен 96 пикселям. И если он был определен не верно, то все остальные слои сети не будут двигаться в верном направлении.

Тогда возникает вопрос – почему PIV-LiteFlowNet-en, которая основана на сети LiteFlowNet, имеет максимум определяемого смещения всего около 10 пикселей. В сеть PIV-LiteFlowNet-en вносились изменения, но незначительные, поэтому маловероятно, что они привели к такому сильному снижению максимального рассчитываемого вектора. Вероятнее всего это связано с обучающей выборкой сети, как можно было заметить на рисунке 42 максимум

смещения равен 11 пикселям, потому сеть просто не обучена распознавать большие смещения. Этим обстоятельством и объясняется разница поведения сетей при инверсии изображений (см. рисунок 44). Сеть PIV-LiteFlowNet-en обучалась по изображению белых точек на чёрном фоне (см. рисунок 41), поэтому она даёт лучшую оценку для инверсных изображений, а сеть LiteFlowNet обучалась на изображениях движущегося стула [76], для которого использовался светлый фон.

#### **2.4.5 Предложения по решению возникших проблем**

В этом разделе описаны предложения по решению основной проблемы, выявленной в этой работе – проблеме больших смещений на изображениях. Любой из трёх рассмотренных методов имеет ограничения по максимально оцениваемому смещению. Так для кросскорреляции максимум смещения ограничен размером окна опроса. Здесь возникает логичный вопрос: а почему бы не брать большое окно опроса и выполнять расчёт с ним? Но чем больше окно опроса мы выбираем, тем дольше расчёт. Поэтому использование больших окон опроса приведёт к значительным вычислительным затратам и ещё приведёт к большим отступам векторного поля от края изображения. Для нейросетей ограничения смещения обусловлено либо внутренней структурой слоёв, либо обучающей выборкой.

Для кросскорреляции эту проблему можно решить двумя способами. Первый способ использовать мультисеточный метод расчёта, который выполняет расчёт разреженного поля с большим окном опроса, а на следующей итерации уменьшает окно опроса (обычно в 2 раза) и выполняет расчёт смещений с учёт оценённого смещения на предыдущем этапе. Проблема заключается в том, что при значительных смещениях это приведёт к ненужным шагам, когда после первой итерации уже возможно уменьшить окно опроса в 4, 8 или 16 раз. Второй способ позволит избежать ненужных итераций. Он заключается в использовании начального разреженного векторного поля,

которое можно рассчитать, не задавая на прямую окно опроса, а, например, исходя из размеров обрабатываемого изображения.

Предложенные выше методы решения проблемы для кросскорреляции не решают ещё одну проблему – проблему погрешности геометрического преобразования. Эта проблема заключается в том, что в результате расчёта вектора смещения мы имеет величину смещения и погрешность расчёта смещения, например,  $5\pm0,1$  пикселя, а выполнение геометрического преобразования аналогично умножению и сложению с какими-то константами, в зависимости от положения на изображении. Таким образом если в результате преобразования смещение из 5 пикселей перейдёт в 500 пикселей, то и погрешность умножиться на величину не более 100, т.е. станет максимум 10 пикс. И уже получим в худшем случае вектор смещения  $500\pm10$  пикс., что является высокой абсолютной погрешностью.

Казалось, тогда почему бы не отказаться от геометрических преобразований, чтобы избежать дополнительной погрешности. Отказ от геометрических преобразований невозможен в силу того, что значительно увеличит вычислительные затраты на обработку изображений, на которые будут попадать области, не относящиеся к измеряемому изображению, и при этом усложнит предобработку изображений такую, как повышение контраста. Это ещё не учитывая того, что смещения на изображениях без геометрических преобразований будут достигать 800 и более пикселей (см. рисунок 49).

Единственный способ решения этой проблемы – введение дополнительного уточняющего шага расчёта после выполнения обратных геометрических преобразований. Например, к полученному вектору смещения  $500\pm10$  пикселей привязано окно проска размером  $32\times32$  пикселя и выполнив расчёт можно получить уточнение до смещения до исходных 0,1 пикселя. Такой подход немного замедлит вычисление из-за дополнительной итерации расчёта, но позволит компенсировать погрешность из-за геометрических преобразований.

Всё вышеописанное не относиться к нейросетям, т.к. в исходном виде они могут обрабатывать только прямоугольные изображения, размер смещения ограничен структурой сети и в существующие решения нельзя загрузить начальное векторное поле. Поэтому решение вышеописанных проблем для сетей PIV-LiteFlowNet и LiteFlowNet невозможно, здесь требуется глубокая модернизация структуры этих сетей под исследуемую задачу.

Также можно выделить два пути разработки новой сети. Первый путь создавать узконаправленную сеть, в которой будет заложен алгоритм поиска ArUco маркеров и выполняться уточнение векторного поля после обратных геометрических преобразований. Второй путь – это создание сети, которая может принимать на вход начальное разреженное векторное поле и на выходе будет выдавать уточнённое и более плотное векторное поле. Следом над этим полем будет выполняться обратное геометрическое преобразование. А для устранения ошибки обратного геометрического преобразования необходимо будет выполнять уточняющих шаг при помощи кросскорреляции.

## ЗАКЛЮЧЕНИЕ

В работе рассмотрены вопросы, связанные восстановлением трёхмерной формы поверхности конструкции методом стереофотограмметрии. Первая рассматриваемая проблема – это нарушение взаимного расположения камер во время измерений, которая актуальна для фотограмметрии крыла опытного летательного аппарата. Во время полёта на стереопару воздействуют вибрации и перегрузки, что приводит изменению внешних параметров стереосистемы. Решить эту проблему множеством известных методов для рекалибровки камер не представляется возможным из-за того, что они требуют в сцене наличия неподвижных точек или статичного фона, чего невозможно осуществить в полёте. Поэтому единственным критерием, по которому детектируется декалибровка и оптимизируются внешние параметры, является ошибка репроекции. Ошибка репроекции имеет положительную корреляцию с погрешностью восстановления трёхмерной формы поверхности и благодаря ей возможно оценить соответствие параметров камер действительности для любой сцены. Т.к. задача оптимизации внешних параметров камер сводится к математической задаче поиска глобального минимума функции множества переменных, то выбран алгоритм Нелдера-Мида, который использует только значения целевой функции. Этот алгоритм уступает по скорости методам использующие значения первой и второй производной, но имеет лучшую сходимость.

Вторая рассматриваемая проблема – увеличение пространственного разрешения восстанавливаемой трёхмерной поверхности. Метод на основе корреляции 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 и как множество методов машинного

обучения необходимо использование графического процессора. Также выявлены ограничения применимости сетей, связанные с максимальным оцениваемым смещением на изображениях.

Решение для первой проблемы является не самым лучшим, т.к. даже текущую реализацию возможно улучшить в скорости обработки подбором более подходящих параметров алгоритма. Поэтому в этой области есть ещё много работы, связанной с поиском оптимального алгоритма минимизации многомерной функции. Нейросети выбранные для решения проблемы пространственного разрешения имеют ограничения величины оцениваемого смещения. Из-за этого требуется модификация выбранных сетей. Т.к. алгоритмы машинного обучения являются интенсивно развивающейся областью, то они имеют перспективы в задаче восстановления трёхмерной поверхности.

## БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Araújo dos Santos J.V., Lopes H., "Application of Speckle Interferometry to Damage Identification", in B.H.V. Topping, (Editor), "Computational Methods for Engineering Science", Saxe-Coburg Publications, Stirlingshire, UK, Chapter 12, pp 299-330, 2012.
2. Pokharna H., Schajer G.S. Quasi single-frame measurements with phase-stepped ESPI // Opt. Lasers Eng. Elsevier Ltd, 2019. Vol. 121, № February. P. 181–188.
3. Pagliarulo V. et al. Leaks detection in stainless steel kegs via ESPI // Opt. Lasers Eng. Elsevier Ltd, 2018. Vol. 110, № June. P. 220–227.
4. Arun Kumar R. et al. Development of prototype of electronic speckle interferometry based spirometer // Opt. Lasers Eng. Elsevier Ltd, 2020. Vol. 136, № June.
5. Mininni M. et al. Damage identification in beams using speckle shearography and an optimal spatial sampling // Mech. Syst. Signal Process. Elsevier, 2016. Vol. 79. P. 47–64.
6. Yan P. et al. Shearography for non-destructive testing of specular reflecting objects using scattered light illumination // Opt. Laser Technol. 2019. Vol. 112, № November 2018. P. 452–457.
7. Lopes H., dos Santos J.V.A., Moreno-García P. Evaluation of noise in measurements with speckle shearography // Mech. Syst. Signal Process. Elsevier Ltd, 2019. Vol. 118. P. 259–276.
8. Arco A. et al. Application of spline interpolation to speckle shearography measurements for damage identification // Procedia Struct. Integr. Elsevier B.V., 2019, Vol. 17. P. 718–725.
9. Xu X. et al. Optimization of speckle pattern based on integer programming method // Opt. Lasers Eng. Elsevier Ltd, 2020. Vol. 133, № December 2019. P. 3–9.
10. Martens M.J.M. Laser-Doppler Vibrometer Measurements of Leaves. 1990. № Buchhave 1975. P. 1–22.

11. Zuo W. et al. LDV-based measurement of 2D dynamic stress fields in transparent solids // J. Sound Vib. Elsevier Ltd, 2020. Vol. 476.
12. Abbas S.H. et al. Underwater vibration analysis method for rotating propeller blades using laser Doppler vibrometer // Opt. Lasers Eng. Elsevier Ltd, 2020. Vol. 132, № May. P. 106133.
13. Jang J.K., Abbas S.H., Lee J.R. Investigation of underwater environmental effects in rotating propeller blade tracking laser vibrometric measurement // Opt. Laser Technol. Elsevier Ltd, 2020. Vol. 132, № June. P. 106460.
14. Aygün H., Apolskis A. The quality and reliability of the mechanical stethoscopes and Laser Doppler Vibrometer (LDV) to record tracheal sounds // Appl. Acoust. Elsevier Ltd, 2020. Vol. 161.
15. Kroschel K. Laser Doppler Vibrometry for Non-Contact Diagnostics 2019.
16. Baqersad J. et al. Photogrammetry and optical methods in structural dynamics – A review // Mech. Syst. Signal Process. Elsevier, 2017. Vol. 86. P. 17–34.
17. 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.
18. Gramola M., Bruce P.J.K., Santer M. Photogrammetry for accurate model deformation measurement in a supersonic wind tunnel // Exp. Fluids. Springer Berlin Heidelberg, 2019. Vol. 60, № 1. P. 1–11.
19. Кулеш В.П. Измерения деформации адаптивной носовой части крыла в аэродинамической трубе методом видеограмметрии // Учёные записки ЦАГИ, Том 45, №6, 2014. Р. 100–109.
20. 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.
21. Boden F. et al. Advanced In-Flight Measurement Techniques 2013.

22. Boden F. et al. In-flight measurements of propeller blade deformation on a VUT100 cobra aeroplane using a co-rotating camera system // Meas. Sci. Technol. 2016. Vol. 27, № 7.
23. Wei B. et al. A videogrammetric system for measuring the full-field blade deformation of a heavy-duty helicopter's rotating rotor // Rev. Sci. Instrum. 2018, Vol. 89, № 11.
24. Chen B. et al. Measurement of Weld-Induced Deformations in Three-Dimensional Structures Based on Photogrammetry Technique // Journal of Ship Production and Design, 2011. № May.
25. Khoo S. W. et al. A review of surface deformation and strain measurement using two-dimensional digital image correlation // Metrol. Meas. Syst., Vol. 23, 2016, No. 3, pp. 461–480.
26. Poroikov A.Y. Reconstruction of 3D profile of a deformed metallic plate by means of the image pattern correlation technique // Meas. Tech. 2014. Vol. 57, № 4. P. 390–395.
27. Jasiński R., Stebel K., Domin J. Application of the DIC technique to remote control of the hydraulic load system // Remote Sens. 2020. Vol. 12, № 21, P. 1–24.
28. Ivanova Y. V., Yu Poroykov A. Estimation of the measurement error of photogrammetric techniques by controlled flexible deformable surface // Proc. 1st IEEE 2019 Int. Youth Conf. Radio Electron. Electr. Power Eng. REEPE 2019. IEEE, 2019. P. 1–4.
29. Parker M.D. et al. Surface deformation tracking and modelling of soft materials // Biomech. Model. Mechanobiol. Springer Berlin Heidelberg, 2019. Vol. 18, № 4. P. 1031–1045.
30. Raffel M. et al. Particle image velocimetry // Handbuch Optische Messtechnik, 2018.
31. Schmidt T., Tyson J., Galanulis K., Full-field dynamic displacement and strain measurement using advanced 3d image correlation photogrammetry: part 1, Exp. Tech. 27 (2003) 47–50.

32. Schmidt T., Tyson J., Galanulis K., Full-field dynamic displacement and strain measurement-specific examples using advanced 3d image correlation, *Exp. Tech.* 27 (2003) 22–26.
33. Ehrhardt D.A., Beberniss T.J., Experimental investigation of dynamic out of plane displacement error in 3D digital image correlation, *Proceedings from 54th AIAA SDM*, (2013).
34. Zappa E., Mazzoleni P., Matinmanesh A., Uncertainty assessment of digital image correlation method in dynamic applications, *Opt. Lasers Eng.* 56 (2014) 140–151.
35. Zappa E., Matinmanesh A., Mazzoleni P., Evaluation and improvement of digital image correlation uncertainty in dynamic conditions, *Opt. Lasers Eng.* 59 (2014) 82–92.
36. Kirmse T. Messung der Ausbreitungsgeschwindigkeit eines Airbags mittels Image Pattern Correlation Technique (IPCT). In: Oldenburger 3D. Herbert Wichmann Verlag, 2009. Pp. 235–242.
37. Boden F., Stasicki B. Development of a rotating camera for in-flight measurements of aircraft propeller deformation by means of IPCT. In: New results in numerical and experimental fluid mechanics IX. Springer, 2014. Pp. 555–562.
38. Veerman H.P.J., Kannemans H., Jentink H.W. Highly accurate aircraft in-flight wing deformation measurements based on image correlation. In: Advanced in-flight measurement techniques. Springer, 2013. Pp. 15–32.
39. Induchoodan P. R. et al. Depth Recovery from Stereo Images // 2014 International Conference on Contemporary Computing and Informatics (IC3I), P. 745-750.
40. Wienke B. Stereo-PIV using self-calibration on particle images // *Experiments in fluids*, vol. 39(2), 2005. Pp. 267–280.
41. Шешин Е. П. и др., Калибровка системы стереоскопического компьютерного зрения // Труды МФТИ Том 10, №2, 2018, стр. 80-90.

42. Cvišic I., Markovic I. and Petrovic I., Recalibrating the KITTI Dataset Camera Setup for Improved Odometry Accuracy // arXiv:2109.03462v1 [cs.RO] 8 Sep 2021; 978-1-6654-1213-1/21 ©2021 IEEE.
43. Плотникова Е.М. Алгоритмы реконструкции трёхмерной модели сцены по двум цифровым изображениям // SWorld - 2-12 October 2012, Доклад/Технические науки – Информатика, вычислительная техника и автоматизация с. 2-12.
44. Тупицын И. В. Реконструкция трёхмерной модели объекта на основе стереопары при решении задач 3d-моделирования // Изв. Сарат. ун-та. Нов. сер. Сер. Математика. Механика. Информатика, 2011, 88-92.
45. Minagawa J. et. al., A Camera Recalibration Method for a Top-view Surveillance System Based on Relative Camera Pose and Structural Similarity // Conference: 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS).
46. Hermans C., Dumont M., Bekaert P. Extrinsic recalibration in camera networks // Fourth Canadian Conference on Computer and Robot Vision (CRV'07), IEEE, vol. 1, pp. 3–10, 2007.
47. Gaopan H. et. al., Self-recalibration of PTZ Cameras // Conference: 2010 International Conference on Machine Vision and Human-machine Interface, MVHI 2010, Kaifeng, China, April 24-25, 2010.
48. Malesa M. and Kujawinska M., Deformation measurements by digital image correlation with automatic merging of data distributed in time // © 2013 Optical Society of America.
49. Barrosa F. et. al., Robust reference system for Digital Image Correlation camera recalibration in fieldwork // January 2018 Procedia Structural Integrity 13:1993-1998.
50. Wieneke B. Volume self-calibration for 3d particle image velocimetry // Experiments in fluids, vol. 45(2), 2008. Pp. 549–556.
51. Westerweel, J., Scarano, F.: Universal outlier detection for PIV data. Exp. Fluids 39(6), 1096–1100 (2005).

52. Westerweel, J., Theoretical analysis of the measurement precision in particle image velocimetry. *Exp. Fluids* 29(1), S003–S012 (2000).
53. Сайт колледжа Доусона (Dawson college) URL: <https://www.dawsoncollege.qc.ca/ai/wp-content/uploads/sites/180/Algorithm-and-ethics.pdf>, дата обращения: 15.09.2021.
54. Ravi K., Snehasis M., “A Conditional Adversarial Network for Scene Flow Estimation” 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), 2019.
55. Zhiwen D., et al., “Super-resolution reconstruction of turbulent velocity fields using a generative adversarial network-based artificial intelligence framework”, *Phys. Fluids* 31, 125111 (2019).
56. Diederik P., Welling M., “Auto-encoding variational bayes” arXiv:1312.6114 (2013).
57. Goodfellow J., et al., “Generative adversarial nets”, *Advances in Neural Information Processing Systems*, 2014.
58. Jeffrey L., “Finding structure in time”, *Cognitive science* 14.2 (1990): 179-211.
59. VisinF., et al. ReSeg: A Recurrent Neural Network-based Model for Semantic Segmentation // In CVPR Deep Vision Workshop, 2016.
60. CheangT.K., el al. Segmentation-free Vehicle License Plate Recognition using ConvNet-RNN. – 2017.
61. LeCun, Yann, et al. “Gradient-based learning applied to document recognition.” *Proceedings of the IEEE* 86.11 (1998): 2278-2324.
62. Сайт института Азимова (The Asimov Institute) URL: <https://www.asimovinstitute.org/neural-network-zoo/>, дата обращения: 20.09.2021.
63. Kulkarni, Tejas D., et al. “Deep convolutional inverse graphics network.” *Advances in Neural Information Processing Systems*. 2015.
64. He, Kaiming, et al. “Deep residual learning for image recognition.” arXiv preprint arXiv:1512.03385 (2015).

65. Chung, Junyoung, et al. "Empirical evaluation of gated recurrent neural networks on sequence modeling." arXiv:1412.3555 (2014).
66. Piyush J., et al., "A review of machine learning applications in wildfire science and management", Environmental Reviews, 28 July 2020, doi: 10.1139/er-2020-0019.
67. Dosovitskiy A., et al., (2015) Flownet: Learning optical flow with convolutional networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp 2758–2766.
68. E. Ilg, N. Mayer, T. Saikia, M. Keuper, A. Dosovitskiy, and T. Brox. FlowNet 2.0: Evolution of optical flow estimation with deep networks. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
69. Lee Y, Yang H, Yin Z (2017) PIV-DCNN: cascaded deep convolutional neural networks for particle image velocimetry. *Exp Fluids* 58(12):171.
70. Deqing S., et al., "PWC-Net: CNNs for Optical Flow Using Pyramid, Warping, and Cost Volume", Conference: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
71. Shengze C., et al., "Dense motion estimation of particle images via a convolutional neural network", Springer Experiments in Fluids (2019) 60:73.
72. Shengze C., et al., "Particle Image Velocimetry Based on a Deep Learning Motion Estimation", August 2019 IEEE Transactions on Instrumentation and Measurement pp (99):1-15.
73. Bo Liu, et al., "Deep learning methods for super-resolution reconstruction of turbulent flows", *Phys. Fluids* 32, 025105 (2020).
74. Ang Li, et al., "Optical Flow Estimation and Denoising of Video Images Based on Deep Learning Models", August 2020 IEEE Access, Vol. 8, pp 122-134.
75. Yeonghyeon G., et al., "Three-dimensional particle tracking velocimetry using shallow neural network for real-time analysis", *Experiments in Fluids* (2020) 61:26.

76. T. Hui, X. Tang, and C. Loy, "LiteFlowNet: A lightweight convolutional neural network for optical flow estimation," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 8981–8989.
77. Ravi Kumar Thakur and Snehasis Mukherjee. SceneEDNet: A Deep Learning Approach for Scene Flow Estimation. In 2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV), pages 394–399. IEEE, 2018.
78. Eddy Ilg, et. al., "Occlusions, Motion and Depth Boundaries with a Generic Network for Disparity, Optical Flow or Scene Flow Estimation" Accepted to ECCV 2018, arXiv:1808.01838, pp. 1–26.
79. Mingrui Zhang and Matthew D. Piggott, "Unsupervised Learning of Particle Image Velocimetry", arXiv:2007.14487 (2020).
80. Huaizu Jiang Caltech and Erik Learned-Miller UMass Amherst, "DCVNet: Dilated Cost Volume Networks for Fast Optical Flow", arXiv:2103.17271 (2021).
81. J. Rabault, J. Kolaas, and A. Jensen, "Performing particle image velocimetry using artificial neural networks: a proof-of-concept," Measurement Science & Technology, 2017, vol. 28, no. 12, p. 125301.
82. Y. Sun, X. Wang, and X. Tang, "Deep convolutional network cascade for facial point detection," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 3476-3483.
83. E. V. Shmatko, V. V. Pinchukov, A. D. Bogachev and A. Yu. Poroykov, "Crosscorrelation image processing for surface shape reconstruction using fiducial markers," J. Physics: Conf. Ser., 2021, vol. 2127, p. 012030.