

ՀԱՅԱՍՏԱՆԻ ՀԱՆՐԱՊԵՏՈՒԹՅԱՆ ԿՐԹՈՒԹՅԱՆ, ԳԻՏՈՒԹՅԱՆ,
ՄՇԱԿՈՒՅԹԻ ԵՎ ՍՊՈՐՏԻ ՆԱԽԱՐԱՐՈՒԹՅՈՒՆ

ՀԱՅԱՍՏԱՆԻ ԱԶԳԱՅԻՆ ՊՈԼԻՏԵԽՆԻԿԱԿԱՆ ՀԱՄԱԼՍԱՐԱՆ

Խաչատրյան Տիգրան Բագրատի

ԱՆՕԴԱԶՈՒ ԹՈՉՈՂ ՄԱՐՔԵՐԻ ՆԱՎԱՐԿՄԱՆ ԱՎՏՈՄԱՏԱՑՄԱՆ
ՀԱՄԱԿԱՐԳԻ ՄՇԱԿՈՒՄԸ ԵՎ ՀԵՏԱԶՈՏՈՒՄԸ

Ե.13.02 «Ավտոմատացման համակարգեր» մասնագիտությամբ
տեխնիկական գիտությունների թեկնածուի գիտական աստիճանի
հայցման ատենախոսության

ՍԵՂՄԱԳԻՐ

Երևան 2025

МИНИСТЕРСТВО ОБРАЗОВАНИЯ, НАУКИ, КУЛЬТУРЫ И СПОРТА
РЕСПУБЛИКИ АРМЕНИЯ

НАЦИОНАЛЬНЫЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ АРМЕНИИ

Хачатрян Тигран Багратович

РАЗРАБОТКА И ИССЛЕДОВАНИЕ АВТОМАТИЗИРОВАННОЙ
СИСТЕМЫ НАВИГАЦИИ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ
АППАРАТОВ

АВТОРЕФЕРАТ

диссертации на соискание ученой степени кандидата
технических наук по специальности 05.13.02-
“Системы автоматизации”

Ереван 2025

Ատենախոսության թեման հաստատվել է Հայաստանի ազգային պոլիտեխնիկական համալսարանում (ՀԱՊՀ):

Գիտական ղեկավար՝	տ.գ.դ. Վազգեն Շավարշի Մելիքյան
Պաշտոնական ընդդիմախոսներ՝	տ.գ.դ. Աշոտ Գևորգի Հարությունյան Ֆ.-մ.գ.թ. Գարեգին Վրեժի Սարգսյան
Առաջատար կազմակերպություն՝	ՀՀ ԳԱԱ Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտ

Ատենախոսության պաշտպանությունը կայանալու է 2025թ. հունիսի 15-ին, ժամը 10⁰⁰-ին, ՀԱՊՀ-ում գործող «Կառավարման և ավտոմատացման» 032 մասնագիտական խորհրդի նիստում (հասցեն 0009, Երևան, Տերյան փ., 105, 17 մասնաշենք):

Ատենախոսությանը կարելի է ծանոթանալ ՀԱՊՀ-ի գրադարանում:

Սեղմագիրն արաքված է 2025թ. հունիսի 13-ին:

032 Մասնագիտական խորհրդի
գիտական քարտուղար, տ.գ.թ.



Անուշ Վազգենի Մելիքյան

Тема диссертации утверждена в Национальном политехническом университете Армении (НПУА)

Научный руководитель: д.т.н. Вазген Шаваршович Меликян

Официальные оппоненты: д.т.н. Ашот Геворкович Арутюнян
к.ф.-м.н. Гарегин Врежович Саргсян

Ведущая организация: Институт проблем информатики и
автоматизации НАН РА

Защита диссертации состоится 15-го июля 2025г. в 10⁰⁰ ч. на заседании Специализированного совета 032 — "Управления и автоматизации", действующего при Национальном политехническом университете Армении, по адресу: 0009, г. Ереван, ул. Теряна, 105, корпус 17.

С диссертацией можно ознакомиться в библиотеке НПУА.

Автореферат разослан 13-го июня 2025 г.

Ученый секретарь

Специализированного совета 032, к.т.н.



Ануш Вазгеновна Меликян

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Беспилотные летательные аппараты (БПЛА) - это летательные аппараты без экипажа. Основными преимуществами БПЛА являются доступность их создания и эксплуатации, а также гибкость, позволяющая применять их для различных задач.

БПЛА могут иметь различные степени автоматизации управления. Недостатками БПЛА считаются уязвимость системы дистанционного управления и сложность системы автоматического управления.

БПЛА в процессе работы должны уметь навигировать к указанному месту, ориентироваться на местности, избегать препятствий и т.д. В настоящее время БПЛА в основном используют глобальные навигационные спутниковые системы (ГНСС). Такие навигационные системы не могут применяться в местности, где нет доступа к ним.

Одним из решений вышеуказанной проблемы является разработка методов локализации, основанных только на визуальной информации, таких как визуальная одометрия (ВО). Наиболее перспективными из методов ВО являются подходы, основанные на обучении с применением нейронных сетей, которые способны на основе большого количества данных изучать особенности динамических сред, характерных для БПЛА. Несмотря на их большой потенциал для навигации БПЛА, последние имеют ряд ограничений, связанных с данными, необходимыми для обучения, точностью моделей, а также требуемыми вычислительными ресурсами.

Диссертация посвящена решению проблем и разработке средств, связанных с автоматизацией процесса навигации БПЛА, которые позволят с более высокой эффективностью и меньшими вычислительными ресурсами в реальном времени (РВ) определять траекторию полета БПЛА.

Объект исследования. Методы совершенствования автоматизированных навигационных систем БПЛА с внедрением искусственного интеллекта (ИИ). Разработка эффективных решений в средах с ограниченными ресурсами.

Цель работы. Разработка способов и средств повышения эффективности методов визуальной одометрии на основе обучения для БПЛА.

Методы исследования. В ходе выполнения диссертации были использованы современные архитектуры нейронных сетей, передовые стратегии их обучения, тестирования и оптимизации, а также другие подходы ИИ и соответствующие пакеты программного обеспечения.

Научная новизна:

- Предложены новые подходы к решению задач автоматизации навигации беспилотных летательных аппаратов, включая гибридный метод создания набора фотореалистичных данных, внедрение трансформерной нейронной сети для оценки глубины, сжатие трансформерной сети методом дистилляции знаний и применение дополнительных методов оптимизации модели, которые позволяют преодолеть проблему накопления ошибки масштаба, повысить точность и устойчивость системы, одновременно обеспечивая их практическую применимость в условиях ограниченных вычислительных ресурсов.
- Разработан гибридный подход к созданию набора искусственных данных,

благодаря которому с использованием метода эффективного объединения реальных изображений, полученных с беспилотного летательного аппарата, и симуляционной среды, лучше охватывая реалистичные рабочие условия беспилотных летательных аппаратов, включая модели полета, разнообразие ракурсов камеры и изменения окружающей среды, и затрачивая на 50% больше времени на подготовку набора данных, стало возможным повысить точность модели визуальной одометрии для реальных изображений, полученных с беспилотного аппарата, примерно на 36%.

- Предложена интеграция трансформерной нейронной сети для определения глубины в монокулярную процедуру визуальной одометрии, которая в различных экспериментальных сценариях обеспечила уменьшение ошибок масштабного отклонения на 57...64% за счет увеличения использования памяти графического процессора в три раза и снижения скорости обработки на 72%. Данный метод, учитывая значительное улучшение точности и существенное снижение кумулятивного эффекта ошибки масштаба, обеспечивает его применимость в практических навигационных системах беспилотных летательных аппаратов.
- Спроектирована ресурсоэффективная модель оценки глубины, которая благодаря дистилляции знаний и оптимизации с помощью библиотеки TensorRT при работе с параметрами точности FP16 увеличивает скорость в среднем на 63% при потере точности всего на 1...2%, обеспечивая скорость обработки, превышающую 60 кадров в секунду, что полностью удовлетворяет требованиям навигации беспилотных летательных аппаратов в реальном времени.

Практическая ценность работы. Предложенные в диссертации средства автоматизации навигации беспилотных летательных аппаратов были реализованы в программном средстве (ПС) “Neural Visual Odometry“, которое внедрено в ООО “ТЕСВАН“. Благодаря удобному пользовательскому интерфейсу, валидации данных, работе в реальном времени и возможности гибкого изменения программных и аппаратных настроек, оно позволило сократить время, требуемое для процесса навигации беспилотных летательных аппаратов, в 2...3 раза. Сравнение результатов с использованием ПС и без него применением показало отклонение точности до 5%.

На защиту выносятся:

- гибридный метод создания фотореалистичного искусственного набора данных для визуальной одометрии;
- метод визуальной одометрии с интегрированной трансформерной нейронной сетью для определения глубины;
- метод сжатия нейронной сети для определения глубины посредством дистилляции знаний и оптимизации модели визуальной одометрии;
- программное средство “Neural Visual Odometry“, способствующее процессу автоматизированной навигации беспилотных летательных аппаратов;

Достоверность научных положений. Достоверность научных результатов

подтверждена программной реализацией представленных в диссертации процедур, результатами моделирования и экспериментов, математическими обоснованиями.

Внедрение. Разработанное ПС “Neural Visual Odometry” внедрено в ООО “ТЕСВАН” и успешно используется для реализации визуальной одометрии в реальном времени с помощью различных видеоматериалов, а также для анализа и исследования результатов.

Апробация работы. Основные научные и практические результаты диссертации докладывались на:

- Международном симпозиуме "IEEE East-West Design & Test Symposium (EWDTS)" (Батуми, Грузия, 2023 г.);
- Международном симпозиуме "IEEE East-West Design & Test Symposium (EWDTS)" (Ереван, Армения, 2024 г.);
- научных семинарах кафедры "Информационные технологии и автоматизация" ННУА (Ереван, Армения, 2022–2024 гг.);
- научных семинарах кафедры "Микроэлектронные схемы и системы" ННУА (Ереван, Армения, 2022 - 2024 гг.).

Публикации. Основные положения, представленные в диссертации, обобщены в десяти научных статьях, список которых представлен в конце автореферата.

Структура и объем диссертации. Работа состоит из введения, трех глав, основных выводов, списка литературы из 127 наименований и трех приложений. В первом приложении представлен акт внедрения диссертации, во втором - фрагменты функциональной структуры реализованного программного средства “Neural Visual Odometry”, в третьем - списки использованных рисунков, таблиц и сокращений. Объем диссертации составляет 128 страниц, а вместе с приложениями - 141 страница. Диссертация написана на армянском языке.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы диссертации, сформулированы цель и основные задачи исследования, представлены разработанные методы, научная новизна, практическое значение и основные научные положения, выносимые на защиту.

В первой главе представлены основные методы автоматизации навигации беспилотных летательных аппаратов (БПЛА) на основе визуальной одометрии (ВО), рассмотрены различные подходы к анализу визуальной информации, изучены существующие методы извлечения и обработки данных.

Современные методы ВО включают три основных направления: классические геометрические методы, методы на основе обучения и гибридные подходы. Классические геометрические методы, в свою очередь, делятся на две группы: признаковые методы и прямые методы.

Признаковые методы основаны на выделении и отслеживании визуальных элементов в последовательности изображений. Структура системы (рис. 1) включает несколько взаимосвязанных компонентов: предварительную обработку изображения, обнаружение особенностей, их описание и оценку движения. Алгоритмы обнаружения особенностей сочетают достижения классического

компьютерного зрения и глубокого обучения.

Основная задача оптимизации в признаковых методах формулируется следующим образом:

$$\min_{\xi} \sum_{i=1}^n \rho(\|e_i(\xi)\|_{(\Sigma_i)}), \quad (1)$$

где ξ выражает параметры движения; e_i представляет ошибку проецирования; $\rho(\cdot, \cdot)$ является устойчивой функцией потерь, а Σ_i - ковариационная матрица измерений.

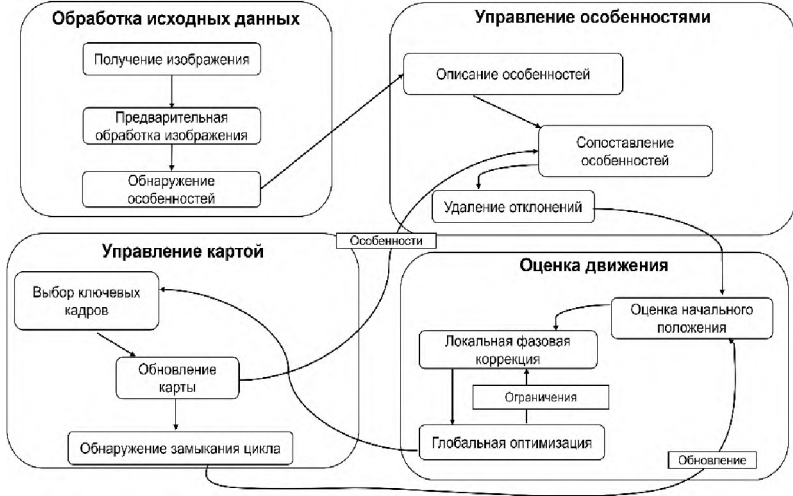


Рис. 1. Основная структура и поток данных системы признакового метода ВО

Прямые методы ВО (рис. 2) работают непосредственно с интенсивностью пикселей изображения, минимизируя ошибку фотометрических измерений (ОФИ). Математическая формулировка минимизации ОФИ между изображениями имеет вид:

$$E(T) = \sum_{x \in \Omega} \rho(\|I_2(\pi(T \cdot \pi^{-1}(x, d(x)))) - I_1(x)\|_Y), \quad (2)$$

где Ω представляет область изображения; π - функция проекции; π^{-1} - функция обратной проекции; $d(x)$ - глубина в пикселе x ; ρ - устойчивая весовая функция, а Y - параметр нормы Хубера.

Методы на основе обучения (рис. 3) часто обходят необходимость поиска особенностей и геометрического моделирования, оценивая положение камеры непосредственно из последовательности изображений. Основная задача может быть сформулирована как определение функции отображения f_{θ} , которая оценивает относительное преобразование положения камеры $T_{t,t+1} \in SE(3)$ между последовательными кадрами I_t и I_{t+1} :

$$T_{t,t+1} = f_{\theta}(I_t, I_{t+1}), \quad (3)$$

где θ представляет параметры нейронной сети, подлежащие обучению.

Методы визуальной одометрии на основе обучения (ВООО) являются наиболее

перспективными для навигации БПЛА благодаря способности изучать сложные визуальные закономерности из больших объемов данных и экономической эффективности.

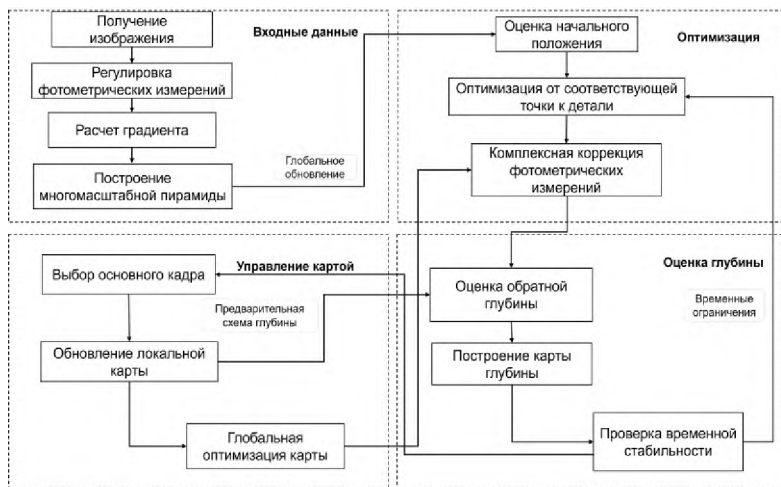


Рис. 2. Основная структура и поток данных системы прямого метода ВО

Однако указанные методы также сталкиваются с проблемами неопределенности масштаба, необходимостью обширных размеченных данных и высокими вычислительными ресурсами для обучения моделей.

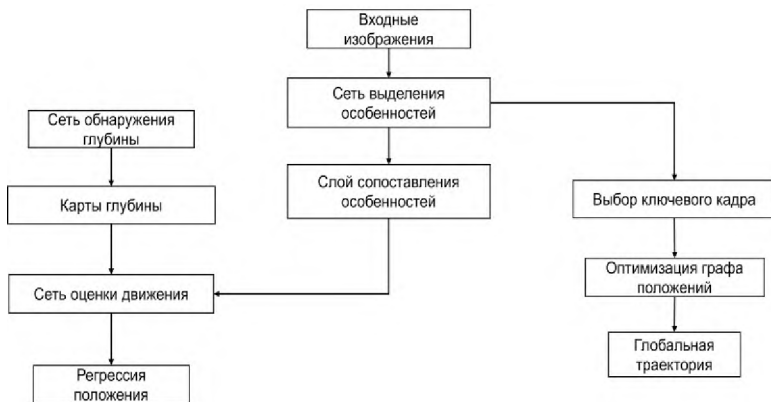


Рис. 3. Процедура систем BOOO

Гибридные методы визуальной одометрии (ГМВО) сочетают преимущества классических геометрических методов и методов BOOO. ГМВО, хотя и обладают

потенциалом объединения преимуществ обоих подходов, сталкиваются с проблемами эффективной синхронизации геометрических и обучающих компонентов. Рассогласование данных между модулями может привести к системе, которая не обеспечивает ни стабильной точности геометрии, ни гибкости изученных признаков.

Современные методы автоматизации навигации БПЛА, несмотря на значительный прогресс, сталкиваются с рядом существенных ограничений. Методы на основе анализа визуальных характеристик ограничиваются однородными структурами или изображениями с малой информативностью, где надежное обнаружение особенностей затруднено. Одной из наиболее проблемных особенностей является накопление ошибки масштаба, особенно характерное для монокулярных систем ВО.

Во второй главе представлены разработанные методы и даются решения проблем, описанных в первой главе.

Гибридный метод создания фотореалистичного искусственного набора данных для визуальной одометрии

Преимущество предлагаемого метода (рис. 4) заключается в его гибридном подходе. Гибридный метод одновременно использует точные структурные детали реальных изображений, а также гибкость и скорость симуляционных сред. Предлагаемый набор данных предназначен для обнаружения глубины на изображениях, полученных с БПЛА, и реализации ВО с их помощью. Создание искусственного набора данных начинается с этапа сбора и предварительной обработки данных.



Рис. 4. Процедура создания фотореалистичного набора данных

Для оптимизации качества получаемых изображений и полноты покрытия были разработаны процедуры планирования полета (табл. 1). Полеты осуществляются на контролируемых высотах, обычно в диапазоне от 30 до 120 м, как определено целями применения набора данных.

Скорости БПЛА регулируются от 4 до 7 м в секунду. Эти параметры позволяют получать высококачественные изображения, подходящие для реализации стабильной реконструкции структуры из движения (РСД).

Таблица 1

Параметры полёта и настройки оборудования

Параметр	Установленное значение
Модель БПЛА	DJI Mini 2 SE
Разрешение камеры	4000 × 3000 пикселей
Фокусное расстояние камеры	24 мм (эквивалент 35 мм)
Диапазон высот	30...120 м
Скорость полета	4...7 м/с
Лобовое перекрытие	70...80%
Боковое перекрытие	60...70%

Собранные с помощью БПЛА изображения затем подвергались процедуре трехмерной реконструкции с использованием техник РСД и многоракурсного стерео (MPC).

Процедуры РСД и MPC, реализуемые через специализированное программное обеспечение 3DF Zephyr для трехмерной реконструкции сцены с помощью фотографий, преобразуют последовательности перекрывающихся изображений БПЛА в пространственные представления.

Затем трехмерные модели импортируются в движок рендеринга Blender, где с путём моделирования полета и камеры можно автоматически генерировать большое количество искусственных изображений.

Скорость генерации набора данных предлагаемым методом на различных графических процессорах (ГП) в 2...3 раза превышает скорость создания набора данных с реальными изображениями (табл. 2).

Таблица 2

Сравнение временных показателей рендеринга и производительности различных методов генерации изображений

Метод	Аппаратная конфигурация	Скорость рендеринга (кадр/с)	Генерация изображений в час (кадров/час)
Предлагаемый гибридный метод	ГП (RTX 3060)	25,3	144
	ГП (RTX 3090)	12,4	300
	ГП (RTX 4070)	10,3	360
Метод получения реальных изображений с БПЛА	Камера DJI Mini 2 SE	25,2	145
Симуляционный метод создания искусственных изображений	ГП (RTX 3060)	12,5	298
	ГП (RTX 3090)	8,1	415
	ГП (RTX 4070)	6,8	550

По сравнению с другими методами создания искусственных изображений, она выше примерно на 50%, так как процедура включает получение небольшого количества реальных изображений с БПЛА, а также дополнительные шаги, направленные на повышение качества.

Для проверки практической полезности и реалистичности созданного искусственного набора данных оценивается модель BOOO TartanVO. Критерии оценки включают абсолютную ошибку траектории (АОТ), среднеквадратичное отклонение (СКО) и среднее отклонение поворота (СОП). Сравнение вышеуказанных двух моделей показывает, что при обучении на предлагаемом наборе данных ошибка модели сокращается на 23,72...30,24% (табл. 3).

Таблица 3

Сравнительные данные производительности метода BOOO TartanVO при обучении на разных наборах данных

Набор данных	СКО (м)	СОП (°)	АОТ (м)
TartanVO (обученный на TartanAir)	2,15	2,62	1,62
TartanVO (обученный на предлагаемом наборе данных)	1,64	1,91	1,13

Для тестирования способности к обобщению модели, обученные на двух наборах данных, оцениваются для реальных траекторий БПЛА (табл. 4). Модель, обученная на искусственных данных, сокращает ошибку модели, обученной с помощью TartanAir, примерно на 36%.

Таким образом, представленные улучшения показывают, что предлагаемый искусственный набор данных лучше охватывает реалистичные рабочие условия БПЛА. Благодаря более детальному моделированию случаев применения БПЛА, а также использованию гибридного подхода, затрачивая на 50% больше времени на подготовку набора данных, стало возможным повысить точность модели TartanVO для набора изображений, полученных с реального БПЛА, примерно на 36%.

Таблица 4

Сравнительные данные обобщения производительности ВО в случае изображений реального мира

Набор данных	СКО (м)	СОП (°)	АОТ (м)
TartanVO (обученный на TartanAir)	3,87	3,93	2,27
TartanVO (обученный на предлагаемом наборе данных)	2,46	2,48	1,47

Метод визуальной одометрии с интегрированной трансформерной нейронной сетью для определения глубины

Предлагаемый метод (рис. 5) интегрирует модель оценки глубины на основе стереотрансформерной сети в процедуру модели ВО TartanVO. В начале процесса пара стереоизображений передается модели оценки глубины на основе трансформерной сети, которая создает карту абсолютной глубины.



Рис. 5. Интеграция нейронной сети для оценки глубины в модель BOOO

Эти данные о глубине вместе с монокулярными входными изображениями TartanVO используются в процессе оценки положения с помощью сверточной нейронной сети (CHC) и комплексной регулировки. В результате получается оптимизированная траектория с абсолютным масштабом, которая по точности превосходит результат только монокулярного подхода. Для обучения использовались 30000 пар стереоизображений. Общее время обучения составило примерно 16 часов с использованием ГП NVIDIA RTX 4070 (8 Гб). Динамика потерь обучения (рис. 6) и валидации (рис. 7) подтверждают, что модель успешно сходилась в течение циклов обучения и отсутствует переобучение.

Численные результаты (табл. 5) показывают значительное снижение ошибки масштаба: улучшение в различных сценариях примерно на 57...64%.

Таблица 5

Оценки накопления ошибки масштаба ВО для последовательностей искусственных и реальных изображений

Тип последовательности	Название последовательности	Количество кадров	Накопление отклонения TartanVO (%)	Накопление отклонения TartanVO+ сети глубины (%)	Накопление отклонения DeepVO (%)
Искусственная	Farm 01	2000	3,5	1,3	10,5
	Urban 02	2000	4,1	1,9	10,8
Реальная	UAV Real 01	1500	9,8	3,5	13,0
	UAV Real 02	1500	10,2	3,8	11,8

Для количественной оценки роста требуемых вычислительных ресурсов было проведено тестирование (табл. 6).



Рис. 6. Динамика потерь набора обучения модели ВО с интегрированной нейронной сетью для оценки глубины

Результаты показывают, что интеграция сети для оценки глубины увеличила использование памяти ГП на 2.1 Гб (примерно трехкратный рост). Скорость обработки снизилась с 77,6 кадр/с до 21,4 кадр/с (снижение примерно на 72%). Эти различия, хотя и заметны, сохраняют производительность в РВ или близкую к РВ, что подходит для типичных операций БПЛА.

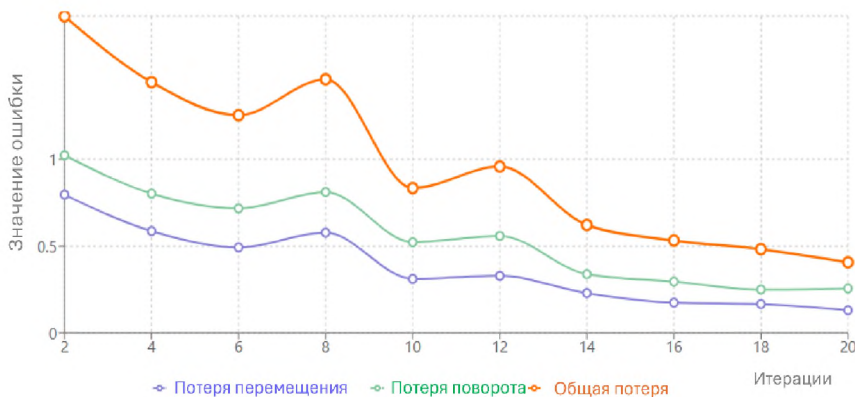


Рис. 7. Динамика потерь валидационного набора модели ВО с интегрированной нейронной сетью для оценки глубины

Таблица 6

Сравнительные данные использования вычислительных ресурсов и производительности

Метод	Использование памяти ГП (Гб)	Скорость обработки (кадр/с)
TartanVO (монокулярный)	1,2	77,6
TartanVO + стерео-трансформерная сеть глубины	3,3	21,4

Метод сжатия нейронной сети для определения глубины посредством дистилляции знаний и оптимизации модели визуальной одометрии.

Метод представляет собой сжатие сети для оценки глубины путем дистилляции знаний и оптимизации всей сети ВО с помощью библиотеки TensorRT. Дистилляция знаний передает представления, изученные большой сетью-учителем, более компактной сети-ученику. Таким образом, значительно снижаются вычислительные требования и требования к памяти. Благодаря этому подходу обеспечивается скорость обработки, соответствующая ограничениям РВ, за счет небольшой потери точности. Архитектура ученика включает удобные для мобильных устройств сверточные слои, выделяющие глубину.

Эффективность дистилляции знаний была оценена в различных сценариях с помощью экспериментов (табл. 7), которые включали обучение ученика без дистилляции, только с дистилляцией конечной карты глубины, и с комбинацией дистилляции глубины и сопоставления промежуточных признаков. Количество параметров и объем использования памяти были снижены с 344 миллионов до примерно 12 миллионов параметров, что позволило осуществить практическое развертывание для навигации БПЛА в РВ. Такое сокращение параметров стало возможным за счет снижения пороговой точности модели ученика на 2% по сравнению с моделью учителя.

Таблица 7

Основные показатели валидации сети для оценки глубины

Сценарий обучения	СКО (м)	СОП (м)	$\delta < 1,25$ (%)
Ученик (без дистилляции)	2,45	1,53	81,24
Ученик (только дистилляция глубины)	2,11	1,37	93,12
Ученик (глубина + сопоставление карты признаков)	2,01	1,25	95,3
Учитель	1,88	1,12	97,3

Основываясь на компактной модели ученика, полученной путем дистилляции знаний, следующий шаг включает оптимизацию с помощью библиотеки NVIDIA TensorRT. Предварительно обученная дистиллированная модель, созданная с использованием библиотеки PyTorch, проходит через процесс преобразования (рис.

8). Сначала модель PyTorch преобразуется в формат ONNX. Впоследствии модель ONNX подвергается оптимизации внутри TensorRT, что требует небольших корректировок, таких как упрощения слоев.

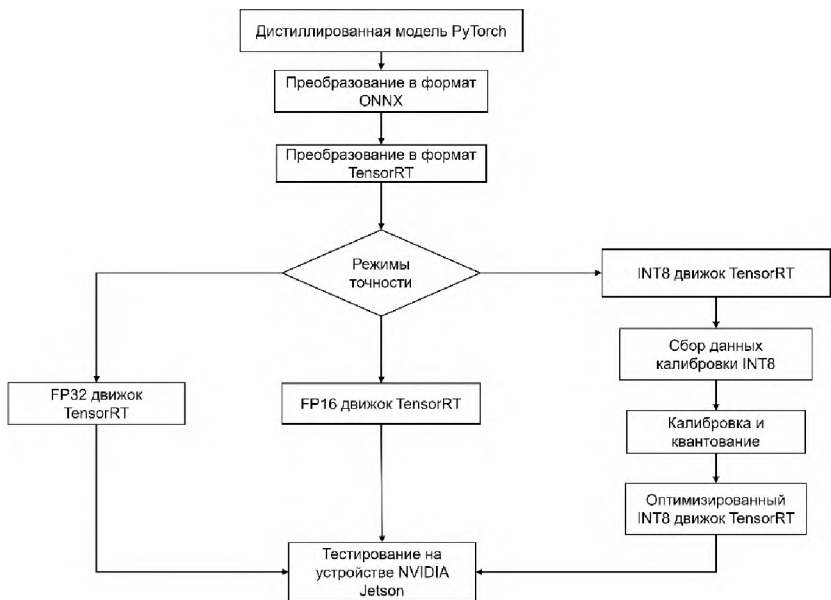


Рис. 8. Процесс оптимизации модели с помощью TensorRT

Для подробной оценки улучшений эффективности, достигнутых путем дистилляции знаний и оптимизации TensorRT, было проведено экспериментальное исследование (табл. 8).

Таблица 8

Результаты экспериментального исследования всего процесса

Модель	Скорость обработки (кадр/с)	СКО ошибка траектории (м)	Среднее отклонение (°)	АОТ (м)
TartanVO + сеть глубины ученика (без дистилляции, без TensorRT)	38	2,25	2,7	1,66
TartanVO + дистиллированная сеть глубины ученика (без TensorRT)	38	1,83	2,12	1,28
TartanVO + дистиллированная сеть глубины ученика (FP16 TensorRT)	62	1,86	2,13	1,3

Результаты исследования показывают, что при дистилляции знаний и оптимизации TensorRT скорость модели с точностью FP16 в среднем увеличилась на 63% за счет потери точности на 1...2%. Оценки подтверждают, что оптимизированная дистиллированная модель достигает скорости обработки в RV, превышающей 60 кадров/с.

В третьей главе представлена разработка программного средства “Neural Visual Odometry”, реализующего методы ВО с использованием нейронных сетей. Система автоматизирует весь процесс навигации от загрузки видеоматериалов до визуализации результатов, минимизируя человеческий фактор.

ПС облегчает работу с данными, используемыми в процессе навигации, устраняет необходимость промежуточных проверок и создает простой интерфейс пользователь-компьютер.

Процесс работы системы начинается с определения входных данных, где пользователь выбирает видеофайл или последовательность изображений. Затем настраиваются параметры моделей нейронных сетей, включая точность параметров сетей и тип модели оценки глубины. Программное средство также позволяет выбрать опцию удаленной обработки, когда вычисления выполняются на устройствах серии Jetson. Интерфейс программы разделен на несколько секций. В разделе "Setup" (рис. 9) пользователь может выбрать тип входного источника (видеофайл, последовательность изображений, поток с камеры или БПЛА), настроить точность модели TartanVO (FP32, FP16, INT8) и выбрать тип модели оценки глубины (стандартная или дистиллированная).

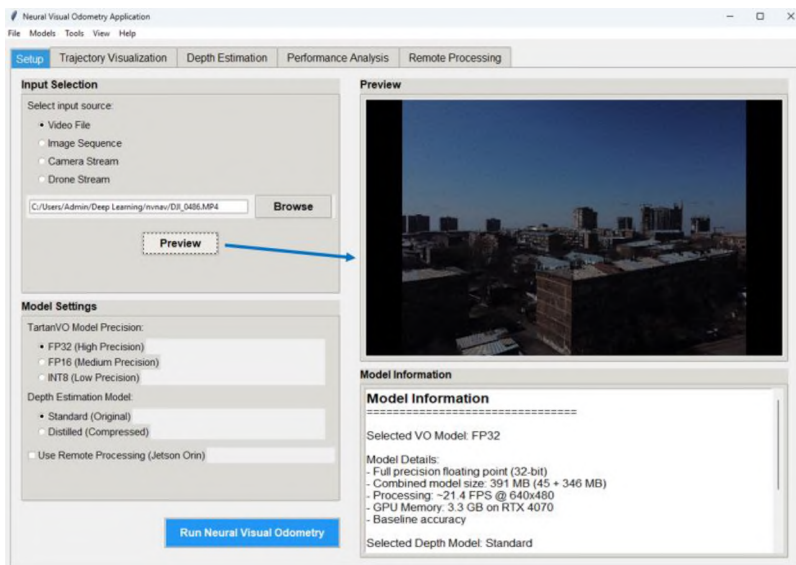


Рис. 9. Основной интерфейс "Setup" для настройки входных данных и моделей

В разделе "Trajectory Visualization" отображаются результаты ВО (рис. 10). Левая часть показывает графическое представление полученной траектории (вид сверху), а правая часть отображает текущий кадр. Пользователь может управлять воспроизведением, настраивать параметры отображения и переключаться между 2D и 3D режимами просмотра.

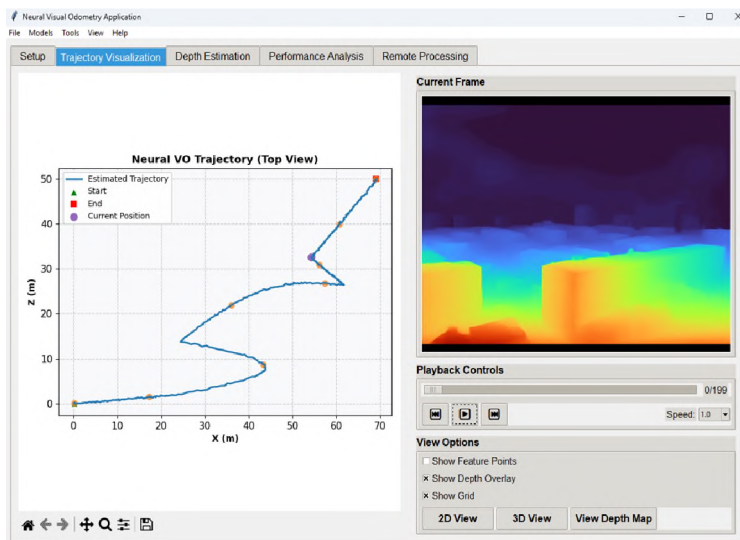


Рис. 10. Интерфейс отображения траектории БПЛА и карты глубины текущего кадра

Для оценки эффективности ПС проведено тестирование на наборе валидации. Сравнение результатов с применением ПС и без него (табл. 9) показало отклонение точности до 5%.

Таблица 9

Оценка программного средства

Метод	RMSE ошибка траектории (м)	Среднее отклонение (°/м)	ATE (м)
Применение предложенного метода	1,83	2,12	1,28
Применение программного средства	1,92	2,23	1,35

Это обусловлено различиями на этапе предварительной обработки. Программное средство позволяет быстро и удобно осуществлять ВО для различных входных данных, обеспечивая высокую точность и возможность работы в РВ.

ОСНОВНЫЕ ВЫВОДЫ ПО ДИССЕРТАЦИОННОЙ РАБОТЕ

1. Предложены новые подходы к решению задач автоматизации навигации беспилотных летательных аппаратов, включая гибридный метод создания набора фотореалистичных данных, внедрение трансформерной нейронной сети для оценки глубины, сжатие трансформерной сети методом дистилляции знаний и применение дополнительных методов оптимизации модели, которые позволяют преодолеть проблему накопления ошибки масштаба, повысить точность и устойчивость системы, одновременно обеспечивая их практическую применимость в условиях ограниченных вычислительных ресурсов [6].
2. Разработан гибридный подход к созданию набора искусственных данных, благодаря которому с использованием метода эффективного объединения реальных изображений, полученных с беспилотного летательного аппарата, и симуляционной среды, лучше охватывая реалистичные рабочие условия беспилотных летательных аппаратов - включая модели полета, разнообразие ракурсов камеры и изменения окружающей среды, и затрачивая на 50% больше времени на подготовку набора данных, стало возможным повысить точность модели визуальной одометрии для реальных изображений, полученных с беспилотного аппарата, примерно на 36% [5, 8].
3. Предложена интеграция трансформерной нейронной сети для определения глубины в монокулярную процедуру визуальной одометрии, которая в различных экспериментальных сценариях обеспечила уменьшение ошибок масштабного отклонения на 57...64% за счет увеличения использования памяти графического процессора в три раза и снижения скорости обработки на 72%. Данный метод, учитывая значительное улучшение точности и существенное снижение кумулятивного эффекта ошибки масштаба, обеспечивает его применимость в практических навигационных системах беспилотных летательных аппаратов [1, 2, 4].
4. Спроектирована ресурсоэффективная модель оценки глубины, которая благодаря дистилляции знаний и оптимизации с помощью библиотеки TensorRT при работе с параметрами точности FP16 увеличивает скорость в среднем на 63% при потере точности всего на 1...2%, обеспечивая скорость обработки, превышающую 60 кадров в секунду, что полностью удовлетворяет требованиям навигации беспилотных летательных аппаратов в реальном времени [3, 7, 10].
5. Предложенные в диссертации средства автоматизации навигации беспилотных летательных аппаратов были реализованы в программном средстве "Neural Visual Odometry", которое внедрено в ООО "ТЕСВАН". Благодаря удобному пользовательскому интерфейсу, валидации данных, работе в реальном времени и возможности гибкого изменения программных и аппаратных настроек, ПС позволило сократить время, требуемое для процесса навигации беспилотных летательных аппаратов, в 2...3 раза. Сравнение результатов с использованием ПС и без него

применением показало отклонение точности до 5% [9].

Основные результаты диссертации опубликованы в следующих работах:

1. **Խաչատրյան Տ.Բ.** Ինքնավար համակարգերում պատկերի խորության նույնականացման համար շարժումից պատկերի տեղափոխության հուշիչի ինտեգրումը // ՀՀ ԳԱԱ և ՀԱՊՀ տեղեկագիր. Տեխնիկական գիտությունների սերիա. - 2020. - Հատոր 73, համար 2. - էջ. 205-213:
2. **Խաչատրյան Տ.Բ.** Պատկերի խորության գնահատումը չվերահսկվող վարժեցմամբ ներդնային ցանցի միջոցով // ՀՀ ԳԱԱ և ՀԱՊՀ տեղեկագիր. Տեխնիկական գիտությունների սերիա. - 2020. - Հատոր 73, համար 3. - էջ. 328-333:
3. **Khachatryan T.** Depth Estimation AI Inferencing Benchmark on Jetson Xavier NX // Bulletin of RAU: Physics-Mathematics and Natural Sciences. - 2021. - no. 2. - P. 5-17.
4. **Khachatryan T.B., Davtyan D.F.** Depth Estimation AI Inferencing Comparison of Jetson Xavier NX And Coral Dev Board // Proceedings of the RA NAS and NPUA. Series of Technical Science. - 2022. - Vol. 75, no. 1. - P. 72-80, doi: 10.53297/0002306X-2022.v75.1-72
5. **Avetisyan A.A., Khachatryan T.B., Grigoryan M.T.** Photorealistic and Synthetic Stereo-Dataset Generation Method for Visual Odometry and Depth Estimation // Proceedings of the RA NAS and NPUA. Series of Technical Science. - 2023. - Vol. 76, no. 1. - P. 12-21, doi: 10.53297/0002306X-2023.v76.1-12
6. **Khachatryan T.B.** A Review of Visual Odometry for UAV Autonomous Navigation // Proceedings of NPUA: Information Technologies, Electronics, Radio Engineering. - 2023. - no. 1. - P. 9-17, doi: 10.53297/18293336-2023.1-9
7. **Nikoghosyan K.H., Khachatryan T.B., Harutyunyan E.A., Galstyan D.M.** Acceleration of transformer architectures on Jetson Xavier using TensorRT // Proceedings of NPUA: Information Technologies, Electronics, Radio Engineering. - 2023. - no. 2. - P. 30-40, doi: 10.53297/18293336-2023.2-30
8. **Khachatryan T., Galstyan D., Harutyunyan E.** A Comprehensive Approach for Enhancing Deep Learning Datasets Quality Using Combined SSIM Algorithm and FSRCNN // 2023 IEEE East-West Design & Test Symposium (EWDTS-2023). - 2023. - P. 1-4, doi: 10.1109/EWDTS59469.2023.10297040
9. **Melikyan V., Grigoryan M., Avetisyan A., Khachatryan T.** Accelerating CNN Models for Visual Odometry: Design and FPGA Implementation for Efficient Hardware Acceleration // 2023 IEEE East-West Design & Test Symposium (EWDTS-2023). - 2023. - P. 1-5, doi: 10.1109/EWDTS59469.2023.10297049
10. **Melikyan V., Khachatryan T., Galstyan D., Harutyunyan E.** Efficient Vision Transformer Deployment on Google Coral Through Knowledge Distillation // 2024 IEEE East-West Design & Test Symposium (EWDTS-2024). - 2024. - P. 1-3, doi: 10.1109/EWDTS63723.2024.10873747

ԱՄՓՈՓԱԳԻՐ

Անօղաչու թռչող սարքերը վերջին տասնամյակներում դարձել են ժամանակակից տեխնոլոգիաների կարևորագույն ձեռքբերումներից մեկը: Դրանք ներկայիս աշխարհում կիրառվում են բազմաթիվ ոլորտներում՝ սկսած արտակարգ իրավիճակների կառավարումից, մինչև գյուղատնտեսական մշտադիտարկումները: Դրանց առանցքային առավելությունը ինչպես արժեքային արդյունավետությունը, այնպես էլ մարդկային կյանքի ռիսկի նվազեցումն են:

Անօղաչու թռչող սարքերի արդյունավետ շահագործման համար էական է ունենալ հուսալի նավարկման համակարգ: Ներկայումս դրանց մեծ մասը հիմնվում է արբանյակային նավարկման համակարգերի վրա, որոնք, սակայն, ունեն էական սահմանափակումներ:

Այս խնդիրների հաղթահարման համար մշակվում են այլընտրանքային նավարկման մեթոդներ, որոնցից առավել հեռանկարայինը տեսողական տարածաչափությունն է: Վերջինս անօղաչու թռչող սարքին հնարավորություն է տալիս որոշելու իր դիրքն ու շարժման պարամետրերը, օգտագործելով միայն տեսախցիկից ստացված պատկերների հաջորդականությունը, ինչը հատկապես արժեքավոր է անհասանելի միջավայրերում:

Տեսողական տարածաչափության զարգացման խոստումնալից ուղղություն է ներդրանային ցանցերի կիրառումը, որոնք կարող են ավտոմատ կերպով ուսումնասիրել միջավայրի բնութագրիչները և իրականացնել ճշգրիտ դիրքորոշում: Այնուամենայնիվ, տեսողական տարածաչափության մեթոդները պահանջում են հզոր հաշվարկային ռեսուրսներ, որոնց օգտագործումը սահմանափակված է անօղաչու թռչող սարքերի էներգետիկ հնարավորություններով: Դրանք նաև թերանում են արագ փոփոխվող լուսավորության պայմաններում կամ միապաղաղ միջավայրերում, ինչպիսիք են ձյունածածկ դաշտերը կամ ջրային մակերևույթները:

Ատենախոսությունը նվիրված է անօղաչու թռչող սարքերի նավարկման գործընթացի ավտոմատացմանն առնչվող հիմնահարցերի լուծմանը և միջոցների մշակմանը, որոնք թույլ կտան, ավելի բարձր արդյունավետությամբ և քիչ հաշվողական ռեսուրսների կիրառմամբ, իրական ժամանակում որոշել ԱԹՄ-ի թռիչքի հետագիծը:

Առաջարկվել են անօղաչու թռչող սարքերի նավարկման ավտոմատացման խնդիրների լուծման մոտեցումներ, այդ թվում՝ ֆոտոիրական տվյալների հավաքածուի ստեղծման հիբրիդային մեթոդ, խորության գնահատման փոխակերպիչ ներդրանային ցանցի ներդրում, գիտելիքների գտման եղանակով փոխակերպիչ ցանցի սեղմում և մոդելի բարելավման լրացուցիչ մեթոդների կիրառում, որոնք թույլ են տալիս հաղթահարել մասշտաբի սխալի կուտակման խնդիրը, բարձրացնել

համակարգի ճշտությունը և կայունությունը, մինևույն ժամանակ ապահովելով դրանց գործնական կիրառելիությունը սահմանափակ հաշվողական ռեսուրսների պայմաններում:

Մշակվել է արհեստական տվյալների հավաքածուի ստեղծման հիբրիդային մոտեցում, որն իրական անօդաչու թռչող սարքից ստացված պատկերների և նմանակային միջավայրի արդյունավետ միավորման եղանակի շնորհիվ, ավելի լավ հաշվի առնելով անօդաչու թռչող սարքերի իրատեսական աշխատանքային պայմանները, ներառյալ թռիչքի մոդելները, տեսախցիկի տեսանկյունների բազմազանությունը և շրջակա միջավայրի փոփոխությունները, տվյալների հավաքածուի պատրաստման վրա ծախսելով 50%-ով ավելի ժամանակ, հնարավորություն է տվել տեսողական տարածաչափության մոդելի ճշտությունը իրական անօդաչու սարքից ստացված պատկերների դեպքում բարձրացնել մոտավորապես 36%-ով:

Առաջարկվել է խորությունների հայտնաբերման փոխակերպիչ նեյրոնային ցանցի ինտեգրում տեսողական տարածաչափության միակապիճային ընթացակարգում, որը տարբեր փորձարարական սցենարների դեպքում ապահովել է մասշտաբի շեղման սխալների նվազեցում 57-64%-ով ի հաշիվ գրաֆիկական պրոցեսորի հիշողության օգտագործման 3 անգամ աճի և մշակման արագության 72% նվազման, ինչը, հաշվի առնելով ճշտության զգալի բարելավումը և մասշտաբի սխալի կուտակային ազդեցության էական նվազումը, ապահովում է մեթոդի կիրառելիությունը գործնական անօդաչու թռչող սարքերի նավարկման համակարգերում:

Նախագծվել է ռեսուրսների առումով արդյունավետ խորության գնահատման մոդել, որը գիտելիքների գտման և TensorRT գրադարանի միջոցով բարելավման շնորհիվ FP16 կարգայնությամբ պարամետրերով գործելու պարագայում արագությունը միջին չափով մեծացնում է 63%-ով ի հաշիվ ճշտության ընդամենը 1-2% կորստի՝ ապահովելով 60 կադր/վայրկյանը գերազանցող առերեսման արագություն, ինչը լիովին բավարարում է անօդաչու թռչող սարքերի իրական ժամանակում նավարկման կիրառության պահանջները:

Մշակվել է անօդաչու թռչող սարքերի նավարկման ավտոմատացման «Neural Visual Odometry» ծրագրային միջոցը, որը ներդրված է «ՏԵՍՎԱՆ» ՍՊԸ-ում և հաջողությամբ օգտագործվում է տարաբնույթ տեսանյութերի միջոցով իրական ժամանակում տեսողական տարածաչափության իրականացման, ինչպես նաև արդյունքների վերլուծության և հետազոտության նպատակով՝ կրճատելով անօդաչու թռչող սարքերի նավարկման գործընթացի համար պահանջվող ժամանակը 2-3 անգամ: Առանց ծրագրային միջոցի և ծրագրային միջոցի կիրառմամբ արդյունքների համեմատությունը ցույց է տվել ճշտության մինչև 5% շեղում, որը պայմանավորված է նախամշակման փուլի տարբերություններով:

**DEVELOPMENT AND RESEARCH OF THE AUTOMATED NAVIGATION
SYSTEM FOR UNMANNED AERIAL VEHICLES**

SUMMARY

Unmanned aerial vehicles represent one of the most significant technological achievements in recent decades. They are currently deployed across numerous fields, from emergency management to agricultural monitoring. Their primary advantages lie in both cost-effectiveness and mitigating risks to human life.

For effective unmanned aerial vehicle operations, a reliable navigation system is essential. Currently, most unmanned aerial vehicles rely on satellite navigation systems, which have considerable limitations, particularly in enclosed spaces, deep canyons, or under electromagnetic interference conditions.

To overcome these limitations, alternative navigation methods are being developed, with visual odometry emerging as the most promising approach. This technique enables an unmanned aerial vehicle to determine its position and movement parameters using only a sequence of camera images, which is especially valuable in inaccessible environments.

A promising direction for visual odometry development is the application of neural networks, which can automatically analyze environmental characteristics and perform precise positioning. However, visual odometry methods demand substantial computational resources, the utilization of which is constrained by the energy limitations of unmanned aerial vehicles. These methods also underperform in rapidly changing lighting conditions or uniform environments, such as snow-covered fields or water surfaces.

This dissertation addresses issues related to automating unmanned aerial vehicle navigation processes and developing means to determine flight trajectories in real-time with greater efficiency and reduced computational demands.

The research proposes novel approaches for solving unmanned aerial vehicle navigation automation challenges, including: a hybrid method for creating visually realistic datasets, integration of transformer neural networks for depth estimation, transformer network compression using knowledge distillation, and application of additional model optimization techniques. These approaches help overcome scale error accumulation problems while enhancing system accuracy and stability within practical computational constraints.

A hybrid approach for artificial dataset creation has been developed that effectively combines real unmanned aerial vehicle-captured images with simulated environments. This method better encompasses realistic unmanned aerial vehicle operating conditions: including flight models, diverse camera angles, and environmental variations. While requiring 50% more time for dataset preparation, it has improved visual odometry model accuracy for real unmanned aerial vehicle imagery by approximately 36%.

The integration of transformer neural networks for depth detection in monocular visual odometry has been implemented. Across various experimental scenarios, this approach reduced scale deviation errors by 57-64%, albeit with a 3-fold increase in

graphics processing unit memory requirements and a 72% decrease in processing speed. Considering the significant accuracy improvements and substantial reduction in cumulative scale error effects, this method remains applicable for practical unmanned aerial vehicle navigation systems.

A resource-efficient depth estimation model has been designed using knowledge distillation and TensorRT library optimization with FP16 precision parameters. This model increases processing speed by an average of 63% while sacrificing only 1-2% in accuracy, achieving rendering speeds exceeding 60 frames per second, fully satisfying practical requirements for real-time unmanned aerial vehicle navigation.

The "Neural Visual Odometry" software tool for automating unmanned aerial vehicle navigation has been developed and implemented at "TESVAN" LLC. It is successfully employed for real-time visual odometry implementation through various video feeds, as well as for result analysis and research. This tool reduces unmanned aerial vehicle navigation process time by 2-3 times. Comparative analysis between navigation with and without the software tool shows accuracy deviations of up to 5%, attributable to differences in the preprocessing phase.

