

УДК 004.932.72

Шифр специальности ВАК: 2.3.8

С.И. Бабаев, Б.В. Костров, Е.Р. Муратов, А.С. Тарасов, В.Ю. Тарасова

Алгоритм уточнения совмещения разнородных изображений на основе ключевых точек

Актуальность. Рассматривается алгоритм уточнения совмещения разнородных изображений. Совмещение изображений осуществляется на основе нейросетевого отбора кандидатных областей. Совмещение разнородных изображений может служить основой для систем определения положения беспилотной воздушной техники, что является актуальной задачей в условиях отсутствия или вызванного помехами искажения сигнала глобальных систем навигации. **Цель исследования.** Разработка и экспериментальная оценка алгоритма уточнения совмещения изображений, сочетающего быстрый нейросетевой отбор кандидатных областей с точным позиционированием на основе сопоставления и геометрической верификации ключевых точек. Особое внимание уделяется работе в условиях, когда кадр, полученный с сенсора ЛА, пересекает границы нескольких спутниковых тайлов. **Научная новизна** заключается в определении положения ЛА при отсутствии сигналов глобальных навигационных систем на основе совмещения разнородных изображений подстилающей поверхности для заданного района полёта. На первом этапе производится отбор кандидатных областей с помощью сиамской свёрточной нейронной сети. Сеть анализирует пары, состоящие из текущего бортового изображения и изображений подстилающей поверхности, вычисляя для каждой пары численную оценку визуального сходства, что позволяет сократить зону поиска. На втором этапе выполняется уточнение координат на основе сопоставления ключевых точек. **Методы.** Рассмотрены методы обнаружения, описания и сопоставления локальных признаков. В результате применения алгоритма уточнения на тестовой выборке средняя ошибка совмещения снизилась с 300 пикселей после нейросетевого отбора до 20 пикселей. Наиболее точные результаты получены для случаев, когда кадр ЛА полностью располагается в пределах одного спутникового тайла. **Практическая значимость.** Разработанный алгоритм обеспечивает определение местоположения ЛА с точностью 10–20 м при времени обработки ≤ 207 мс на основе разнородных изображений подстилающей поверхности.

Ключевые слова: разнородные изображения, совмещение изображений, сиамская архитектура нейронной сети, ключевые точки, визуальное сходство.

DOI: 10.21293/1818-0442-2026-29-2-107-115

Введение

В данной работе рассматривается уточнение совмещения разнородных изображений (на основе небольшого фрагмента, соответствующего кадру, снятому строго вертикально) после первичного, возможно неточного, нейросетевого совмещения. Задача совмещения разнородных изображений, полученных с сенсора ЛА и из спутниковых данных, является частным случаем визуальной навигации в условиях отсутствия сигналов глобальных навигационных систем. Данная задача решается большим количеством исследователей как в Российской Федерации, так и за рубежом. Совмещением разнородных снимков в отечественной научной школе занимались, в частности, А.И. Новиков и А.И. Ефимов; их работы охватывают вопросы сопоставления разнородных изображений, отличающихся по ракурсу и разрешению [1–3].

В работах Ю.В. Визильтера и С.Ю. Желтова предложены методы корреляционно-экстремального совмещения изображений, применимые к данной задаче [4–6].

Для снижения погрешности совмещения пары разнородных изображений предлагается использовать нейросетевые технологии, в частности, декомпозированную сиамскую архитектуру свёрточной нейронной сети.

Так как спутниковые изображения имеют высокое разрешение, их необходимо предварительно обработать и разделить на части. Тайл – это небольшой

фрагмент (участок) большой карты или изображения подстилающей поверхности. Каждый тайл записывается в базу данных изображений подстилающей поверхности (БД) и имеет уникальный идентификатор и метаданные.

Задача совмещения разнородных изображений осуществляется на основе пары изображений: одно получено в текущий момент времени с ЛА, второе – из БД изображений. Также известен предполагаемый маршрут ЛА. На предварительном этапе с использованием нейросетевого алгоритма осуществляется отбор тайла-кандидата для совмещения.

Алгоритм уточнения совмещения разнородных изображений на основе ключевых точек состоит из двух этапов: предварительный отбор тайлов-кандидатов в окрестности заданного маршрута на основе сиамской нейросетевой архитектуры и непосредственно уточнение, где с использованием ключевых точек осуществляется визуальная локализация внутри этих областей [7–9].

Компанией SandboxAQ разработана система AQNav для геомагнитной навигации в условиях отсутствия или ненадежности сигналов систем глобальной навигации [10]. Данная система использует дорогостоящее оборудование (квантованные сенсоры) и позволяет осуществлять навигацию с погрешностью до ~1,8 км [11].

Астрономическая навигация для БПЛА рассматривается в работе [12], погрешность навигации

составляет до 4 км. Для астронавигации [12] необходимы специализированное оборудование (оптика), а также чистое звёздное небо. Существенным недостатком данной системы является невозможность работы в условиях слабой видимости (туман, облачность, городская засветка), а также днем.

С помощью программного комплекса «Навигация» решается задача уточнения положения БПЛА вблизи известной точки [13]. Для экспериментальных исследований [13] проводились имитационное моделирование видеосъемки, полученной с сенсора БПЛА, подготовка исходных данных и полетных заданий. Предложенный подход ориентирован на решение более общей задачи – определение визуальной локализации «с нуля» после полной потери ориентировки в окрестностях заданного маршрута. На первом этапе осуществляется отсеивание зоны неопределенности («разброс»), на втором этапе – позиционирование.

Таким образом, целью данной работы является разработка и исследование алгоритма связывания ключевых точек между изображениями пары «спутник–ЛА» для уточнения позиции ЛА. Известны ориентировочная высота и углы ориентации (крен, тангаж, азимут), что позволяет прогнозировать и компенсировать ожидаемые геометрические искажения. В отличие от подходов, ориентированных на непосредственное совмещение изображений либо на уточнение положения вблизи заранее известной точки [4–6, 13–16], предлагаемый алгоритм использует двухэтапную схему: нейросетевое сужение области поиска и последующее локальное уточнение по ключевым точкам с геометрической верификацией. Отличительной особенностью является выполнение уточнения в кандидатной области, включающей несколько соседних тайлов, что позволяет обрабатывать случаи пересечения кадром ЛА границ тайлов. Новизна работы состоит прежде всего в указанной схеме совместного использования нейросетевого отбора и локального геометрического уточнения, а не в отдельных базовых операциях ORB, SIFT или RANSAC.

Анализ снимков с ЛА

Ключевой задачей является определение координат ЛА в условиях отсутствия сигналов спутниковой навигации. Решение данной задачи сводится к сопоставлению данных, полученных с сенсора ЛА, и спутникового изображения. Сложность представляет сопоставление разнородных изображений: текущего кадра с борта ЛА и заранее полученного снимка подстилающей поверхности, что обусловлено следующими факторами:

- изменение условий съемки (сезон, время суток, погода);
- различия в ракурсе и проекции;
- наличие шумов и искажений;
- ограниченная вычислительная мощность бортового оборудования.

Для повышения устойчивости алгоритма к изменяющимся условиям съемки предлагается использо-

вать подавление низкочастотных составляющих с применением спектрального преобразования Уолша, которое вычисляется следующим образом:

$$\text{Wal}(k, n) = \sum_{j=0}^{m-1} (\text{Rad}_j(n))^{k_j(r)},$$

где $\text{Wal}(k, n)$ – функция Уолша порядка k в точке n ; $N = 2^m$ – длина сигнала; $k_j(r)$ – j -й бит в коде Грея числа r , который вычисляется как

$$k_j(r) = k_j \oplus k_{j+1},$$

где k_j – j -й бит двоичного представления k ; $\text{Rad}_j(n)$ – функция Радемахера порядка j :

$$\text{Rad}_j(n) = (-1)^{n/2^j}.$$

Такая обработка направлена на уменьшение влияния плавных изменений яркости и крупномасштабных различий освещения между бортовым и спутниковым изображениями. Фильтрация применяется к обоим изображениям до извлечения локальных признаков. Её вклад оценивался отдельным экспериментом на той же тестовой выборке: без фильтрации средняя ошибка (Validation Accuracy) составила 0,7300; с фильтрацией – 0,8081. Количество удаляемых коэффициентов преобразования Уолша составляет $> 50\%$.

Для уточнения совмещения применяются методы, основанные на локальных признаках и ключевых точках, вычислительная сложность которых допускает реализацию на бортовом оборудовании с ограниченными вычислительными ресурсами.

Совмещение разнородных изображений внутри кандидатной области

Предложенный алгоритм осуществляет уточнение внутри кандидатной области, которая может включать несколько тайлов. Уточнение осуществляется с использованием ключевых точек.

На рис. 1 приведены различные варианты расположения снимка с ЛА относительно тайлов местности в окрестности заданного маршрута. На рис. 1 в рамке, выделенной пунктиром (затемнённый участок), отмечен кадр с ЛА, полученный в текущий момент времени. Точка-треугольник в центре этой прямоугольной рамки (FRAME UAV) – центр кадра с ЛА. Также выделены границы тайлов, на которые разбит спутниковый снимок, и отмечены номера тайлов.

Возможны следующие варианты расположения снимка с ЛА относительно тайлов местности: кадр с ЛА полностью перекрывает тайл спутникового снимка (см. рис. 1, а), перекрытие нескольких тайлов спутникового снимка (см. рис. 1, б–г).

Ключевые точки

Для уточнения совмещения предлагается использовать ключевые точки. Они, как правило, располагаются на участках с высокой информативностью. Для нахождения ключевых точек используются детекторы, а для их описания – дескрипторы. Дескриптор формирует для точки вектор фиксированной размерности.

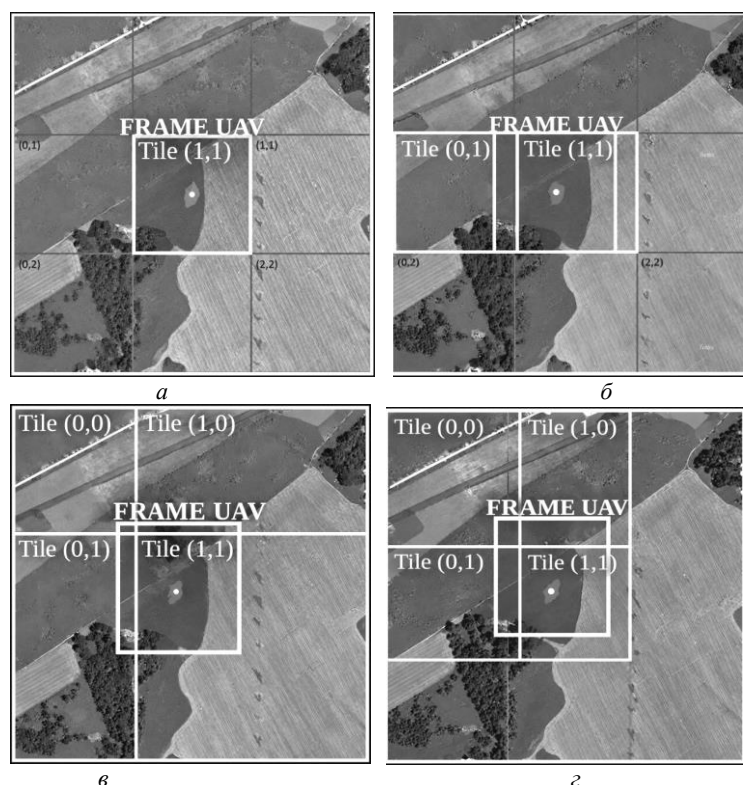


Рис. 1. Визуализация расположения ЛА относительно тайлов местности

Благодаря дескрипторам становится возможным идентифицировать и сравнивать одни и те же локальные особенности на разных изображениях, несмотря на изменения масштаба, условий освещения и перспективы.

В данной работе рассматривались методы обнаружения, описания и сопоставления локальных признаков, приведённые в табл. 1 [17–22]. Следует отметить, что LoFTR относится к детектор-свободным методам сопоставления, тогда как Deep Features и TIMM Features используют признаки, извлекаемые предобученными свёрточными нейронными сетями. Поэтому далее эти подходы сопоставляются как альтернативные способы получения соответствий между парой изображений, а не только как детекторы ключевых точек [23, 24].

В табл. 1 приведены характеристики найденных соответствий.

Таблица 1
Сравнение методов локальных признаков

Метод	Время обработки, мс	Найдено ключевых точек	Количество совпадающих точек	Доля совпадений, %
BRISK	583,93	100	7	7,0
LoFTR	2 800,15	228	22	9,6
ORB	48,76	100	10	10,0
AKAZE	154,41	100	11	11,0
Deep Features	1 491,8	20	6	30,0
ORB+SIFT	1 104,18	21	8	38,1
SIFT	254,02	11	5	45,5
TIMM Features	1 112,4	14	9	64,3

Метод Deep Features использует активации свёрточных слоёв предобученной CNN (например, ResNet50) в качестве дескрипторов [18]. TIMM Features использует модели из библиотеки timm, содержащей широкий набор предобученных архитектур, что позволяет выбирать глубину и тип извлекаемых признаков.

В рамках данной задачи требуется высокая надёжность каждого соответствия при приемлемой скорости сопоставления пар изображений. Рассмотрим нейросетевые методы Deep/TIMM Features, SIFT и комбинацию ORB+SIFT. В табл. 2 приведено сравнение их вычислительных характеристик.

Исходя из данных табл. 1 и 2, можно сделать вывод, что нейросетевые методы TIMM и Deep Features демонстрируют высокую долю совпадений (до 64,3%), однако требуют дополнительного времени и памяти на инициализацию модели. LoFTR формирует большое число соответствий, но в проведённом эксперименте имел низкую долю принятых совпадений (9,6%) и наибольшее время обработки. Связка ORB+SIFT выбрана как компромиссный вариант для бортового применения: в измерениях она обеспечила 29,1 пары/с при объёме памяти 38 МБ, т.е. несколько более высокую скорость и меньший объём памяти, чем SIFT (27,6 пары/с и 45 МБ).

Таблица 2
Сравнение методов локальных признаков

Метод	Скорость, пар/с	Память, МБ	Инициализация, мс
Deep Features	14,0	450	1 200
SIFT	27,6	45	0
ORB+SIFT	29,1	38	0
TIMM Features	79,5	280	860

При этом окончательное решение о соответствии принимается после фильтра Lowe и геометрической верификации RANSAC, поэтому доля исходных совпадений из табл. 1 не является единственным критерием выбора.

Алгоритм уточнения совмещения разнородных изображений

Алгоритм уточнения совмещения разнородных изображений состоит из следующих этапов:

Этап 1. Предварительная подготовка.

1. Фильтрация низких частот.
2. Подготовка БД спутниковых снимков. Разделение изображений на тайлы.

3. Нейросетевой отбор тайлов. Осуществляется анализ тайлов в окрестности заданного маршрута. Тайл-кандидат выбирается по значению косинусного сходства признаков, рассчитываемого по следующей формуле:

$$\text{cosine}_{\text{similar}}(A, B) = \frac{(A, B)}{\|A\| \cdot \|B\|},$$

где (A, B) – скалярное произведение векторов-признаков A и B ; $\|A\|$ и $\|B\|$ – их нормы; i – индекс суммирования; m – размерность вектора признаков. Здесь обозначение m относится только к размерности вектора признаков и не связано с параметром m , используемым выше при описании преобразования Уолша.

На рис. 2 приведена зависимость средней ошибки совмещения пары изображений (в пикселях) от размера кандидатной области тайлов.

На рис. 3 приведена зависимость времени поиска и средней ошибки в пикселях от размера области поиска.

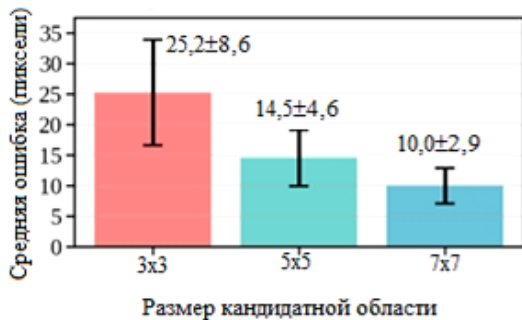


Рис. 2. Зависимость средней ошибки совмещения пары изображений (в пикселях) от размера кандидатной области тайлов

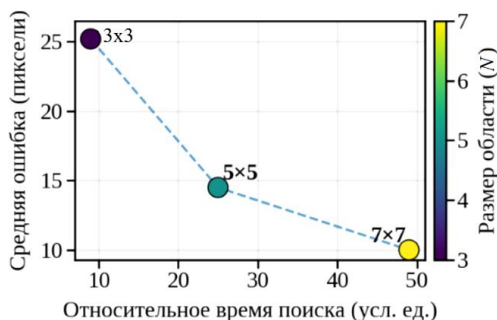


Рис. 3. Зависимость времени поиска и средней ошибки в пикселях от размера области поиска

Этап 2. Детектирование и описание ключевых точек.

1. Детектирование точек ORB-детектором. Для каждого изображения с ЛА и каждого кандидатного тайла применяется ORB-детектор, позволяющий устойчиво находить ключевые точки на разных масштабах.

2. Вычисление дескрипторов SIFT. Для ключевых точек, обнаруженных детектором ORB, вычисляются SIFT-дескрипторы.

Этап 3. Сопоставление и верификация.

1. Поиск соответствий. Для поиска соответствий используется алгоритм FLANN. Для отсека ложных совпадений применяется критерий отношения Лоу (Lowe ratio test): сохраняются только пары точек, для которых отношение расстояния до ближайшего соседа к расстоянию до второго ближайшего соседа меньше 0,8 [20].

2. Геометрическая верификация. Для каждого кандидатного тайла с достаточным количеством совпадений необходимо вычислить гомографию с помощью RANSAC [25]. Сохраняются только внутренние точки, соответствующие найденной модели.

Этап 4. Уточнение совмещения изображений.

1. Производится комбинирование метрик: количества совпадений, доли внутренних точек после RANSAC и оценки нейросетевого алгоритма. Перед объединением показатели приводятся к сопоставимому диапазону. Выбирается тайл с максимальной итоговой оценкой.

2. Центр кадра с ЛА проектируется на выбранный тайл. Локальные координаты тайла преобразуются в глобальные координаты. В результате определяются глобальные координаты точки, соответствующей центру кадра; при принятии модели съёмки и известной ориентации ЛА эта точка используется для уточнения оценки его положения.

Для перевода в географические координаты используется модель билинейной интерполяции, рассчитываемой по формуле

$$\text{lat}_{\text{est}} = \text{lat}_{\text{TL}} - \frac{y}{H} \cdot (\text{lat}_{\text{TL}} - \text{lat}_{\text{BR}}),$$

$$\text{lon}_{\text{est}} = \text{lon}_{\text{TL}} + \frac{x}{W} \cdot (\text{lon}_{\text{BR}} - \text{lon}_{\text{TL}}),$$

где $(\text{lat}_{\text{TL}}, \text{lon}_{\text{TL}})$ – координаты верхнего левого угла тайла; $(\text{lat}_{\text{BR}}, \text{lon}_{\text{BR}})$ – координаты нижнего правого угла тайла; W, H – ширина и высота тайла в пикселях; $(\text{lat}_{\text{est}}, \text{lon}_{\text{est}})$ – расчётные географические координаты точки.

Этап 5. Оценка достоверности. Точность результата оценивается на основе количества геометрически согласованных совпадений. Для верификации выполняется визуализация итогового совмещения.

Подготовка эксперимента

Для тестирования механизма сопоставления ключевых точек был сформирован набор данных, состоящий из 2000 пар разнородных изображений, полученных с разных локаций: городская застройка, песчаная пустыня с застройкой, чистый лес и мегаполис. Все пары изображений были предварительно выровнены по высоте и углу съёмки и имели извест-

ное эталонное взаимное положение. Исследовалось 11 различных сценариев. Такая структура данных позволяет оценивать остаточную ошибку совмещения относительно известного эталонного положения.

Для каждой пары изображений ошибка совмещения определяется как евклидово расстояние между оценённым и эталонным положением центра кадра в общей пиксельной системе координат. В качестве основных агрегированных метрик используются средняя ошибка, медиана, RMSE и среднее квадратическое отклонение. Для оценки корректности отдельных соответствий введён порог 10 пикселей: соответствие считается допустимым, если расстояние между эталонной и найденной точками не превышает этого значения. Такое значение соответствует 10 м на местности, что является физически обоснованным допуском.

Экспериментальные результаты

В данной работе была проведена экспериментальная оценка работоспособности предложенного алгоритма уточнения и его сравнение с базовыми подходами. Для сравнения выбраны: прямое сопоставление SIFT как классический локальный метод; нейросетевой отбор тайла как результат первого этапа предлагаемого конвейера; VGG16+L2 как базовый вариант глобального CNN-дескриптора с расстоянием L2; CVM-Net как специализированная нейросетевая архитектура для сопоставления изображений разных ракурсов [26]. Сравнительные характеристики приведены в табл. 3.

Точность совмещения до применения алгоритма уточнения с использованием ключевых точек составляла 300 пикселей, после – до 20 пикселей. Предложенный алгоритм демонстрирует высокую точность на тестовой выборке на участках с выраженными текстурами (городская застройка, дороги, изгиб реки). Однако его эффективность снижается на однородных поверхностях (водная гладь, поля), где количество ключевых точек недостаточно для точного сопоставления.

Наилучший результат у сценария «Мегаполис», наихудший результат – «Обработываемое поле».

Результаты, представленные в табл. 3, получены на одном тестовом наборе данных путем повторных запусков исследуемых методов.

		Мегаполис, ОАЭ Accuracy: 0,919 Precision: 0,750				Город, Рязань Accuracy: 0,783 Precision: 0,727				Коттеджи, США Accuracy: 0,660 Precision: 0,333							
Истинный класс	Positive	TP: 3 (3,0%)		FP: 7 (7,1%)		Истинный класс	Positive	TP: 8 (8,7%)		FP: 17 (18,5%)		Истинный класс	Positive	TP: 15 (15,0%)		FP: 4 (4,0%)	
	True Negative	FN: 1 (1,0%)		TN: 88 (88,9%)			True Negative	FN: 3 (3,3%)		TN: 64 (69,6%)			True Negative	FN: 30 (30,0%)		TN: 51 (51,0%)	
		Pred Positive		Pred Negative				Pred Positive		Pred Negative				Pred Positive		Pred Negative	
		Предсказанное значение						Предсказанное значение						Предсказанное значение			

Рис. 5 (начало)

На рис. 4 приведен результат совмещения. Линией выделена траектория полета ЛА, квадратами выделены тайлы.

На рис. 5 приведены оценка ошибки первого и второго рода для различных наборов сценариев.

Таблица 3

Сравнительные характеристики методов

Метод	Средняя ошибка, пикс.	Время, мс
SIFT (прямое)	>500	180
Нейросетевой отбор тайла	300	106
VGG16 + L2	150	300
CVM-Net	80	450
Алгоритм уточнения на основе ключевых точек	20	207



Рис. 4. Траектория полета ЛА и выбранные тайлы для совмещения

Выводы

В рамках данной работы разработан и экспериментально исследован алгоритм уточнения совмещения изображений, полученных с разных сенсоров. Подход включает первоначальный нейросетевой отбор и последующее позиционирование за счёт сопоставления локальных признаков [27, 28].

Экспериментальная проверка показала, что применение связки ORB+SIFT с последующей фильтрацией соответствий по критерию Лоу и геометрической верификацией методом RANSAC позволяет снизить среднюю ошибку совмещения с 300 до 20 пикселей относительно результата после этапа нейросетевого поиска.

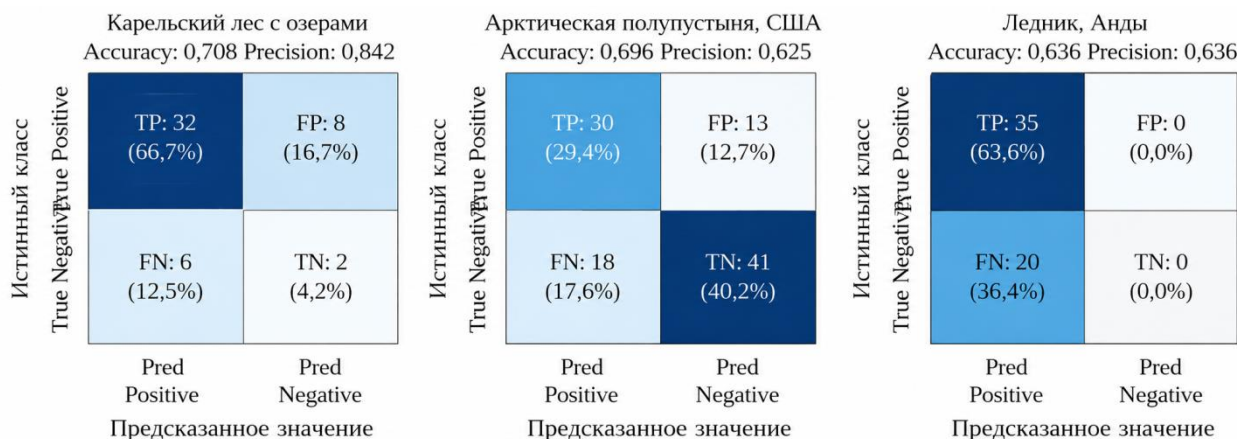


Рис. 5 (окончание). Ошибки первого и второго рода для различных сценариев

Среднее время выполнения полного уточнения составляет 207 мс, что соответствует частоте около 4,8 Гц. При скорости ЛА 50 м/с за один такой интервал аппарат перемещается примерно на 10,4 м.

Алгоритм сохраняет работоспособность в сценариях, при которых кадр ЛА пересекает границы нескольких тайлов спутниковых изображений.

Наиболее высокая точность достигается на участках с хорошо выраженными и устойчивыми текстурными деталями. На однородных поверхностях число геометрически согласованных соответствий уменьшается, вследствие чего возрастает ошибка и снижается доля успешных уточнений. Дальнейшие исследования целесообразно направить на повышение устойчивости алгоритма на слабоструктурированных участках и на сокращение времени обработки на бортовой вычислительной платформе.

Литература

- Новиков А.И. Алгоритмы автоматической идентификации объектов на разнородных изображениях и совмещения изображений / А.И. Новиков, А.И. Ефимов, Д.А. Колчаев // Сборник трудов IV Междунар. конф. и молодежной школы «Информационные технологии и нанотехнологии» (ИТНТ–2018). – Самара: Новая техника, 2018. – С. 650–663.
- Ефимов А.И. Комбинированный алгоритм совмещения облаков точек / А.И. Ефимов, Т.Н. Крючкова, А.И. Ярославцева // GraphiCon – 2023: труды 33-й Междунар. конф. по компьютерной графике и машинному зрению. – М.: Институт прикладной математики им. М.В. Келдыша РАН, 2023. – С. 539–548.
- Sablina V.A. Combined approach to object contour superimposition in heterogeneous images / V.A. Sablina, A.I. Efimov, A.I. Novikov // 2017 6th Mediterranean Conference on Embedded Computing (MECO). – Bar, Montenegro: IEEE, 2017. DOI: 10.1109/MECO.2017.7977175.
- Алгоритм автоматического совмещения сенсорной и синтезируемой видеoinформации для авиационной системы комбинированного видения / М.А. Лебедев, М.А. Бондаренко, Д.В. Комаров, Д.Г. Степаньянц, О.В. Выголов, Ю.В. Визильтер, С.Ю. Желтов // Вестник компьютерных и информационных технологий. – 2014. – № 7 (121). – С. 8–13.
- Сравнение изображений по форме с использованием диффузной морфологии и диффузной корреляции /

Ю.В. Визильтер, В.С. Горбачевич, А.Ю. Рубис, О.В. Выголов // Компьютерная оптика. – 2015. – Т. 39, № 2. – С. 265–274.

6. Визильтер Ю.В. Современные технологии искусственного интеллекта и их применение в авиационных комплексах / Ю.В. Визильтер, Б.В. Вишняков, С.Ю. Желтов // XVI Всерос. мультikonф. по проблемам управления (МКПУ–2023): матер. мультikonф. – Волгоград: Волгоград. гос. техн. уни-т, 2023. – С. 13–16.

7. Кочетуров В.В. Сиамские свёрточные нейронные сети для распознавания лиц / В.В. Кочетуров, О.И. Федяев // Программная инженерия: методы и технологии разработки информационно-вычислительных систем (ПИИВС-2024): сборник матер. и докл. V Междунар. науч.-практ. конф. – Донецк: Дон. нац. техн. ун-т, 2024. – С. 178–184.

8. Ладанова Е.О. Разработка методики и новой структуры сиамской нейронной сети для анализа геопространственных данных // Научно-технический вестник Поволжья. – 2024. – № 12. – С. 144–147.

9. Li Y. A Survey on Siamese Network: Methodologies, Applications, and Opportunities / Y. Li, C.L.Ph. Chen, T. Zhang // IEEE Transactions on Artificial Intelligence. – 2022. – Vol. 3, No. 6. – P. 994–1014. DOI: 10.1109/TAI.2022.3207112.

10. Broughton G. Robust Image Alignment for Outdoor Teach-and-Repeat Navigation / G. Broughton, P. Linder, T. Rouček, T. Vintr, T. Krajník // Proceedings of the 2021 European Conference on Mobile Robots (ECMR). – Bonn, Germany: IEEE, 2021. – P. 1–6. DOI: 10.1109/ECMR50962.2021.9568832.

11. Rozsypalek Z. Semi-Supervised Learning for Image Alignment in Teach and Repeat Navigation / Z. Rozsypalek, G. Broughton, P. Linder, T. Roucek, K. Kusumam, T. Krajník // Proceedings of 37th ACM/SIGAPP Symposium on Applied Computing (SAC '22). – New York, USA: ACM, 2022. DOI: 10.1145/3477314.3507061.

12. Teague S. An Algorithm for Affordable Vision-Based GNSS-Denied Strapdown Celestial Navigation / S. Teague, J. Chahl // Drones. – 2024. – Vol. 8, No. 11. – Art. 652. DOI: 10.3390/drones8110652.

13. Программный комплекс для автоматизации научных исследований в области создания перспективных систем навигации беспилотных летательных аппаратов по данным видеонаблюдения / Б.А. Алпатов, П.В. Бабаян, Ю.С. Коблов, В.С. Муравьев, В.В. Стротов, А.Б. Фельдман // Вестник Рязан. гос. радиотехнического ун-та. – 2012. – № 40. – С. 9–15.

14. Оптико-электронная корреляционно-экстремальная навигационная система малого беспилотного летательного аппарата / А.А. Шейников, В.А. Малкин, Э.А. Смоленский, Л.А. Иваницкий // Системный анализ и прикладная

информатика. – 2024. – № 2. – С. 30–40. DOI: 10.17212/2307-5981-2024-2-30-40.

15. TTC-ориентированная визуальная навигация систем БПЛА / Е.А. Конников, Д.А. Сергеев, Г.И. Голиков, Д.Г. Родионов, П.А. Поляков, А.Ф. Тихомиров // Программные системы и вычислительные методы. – 2025. – № 4. – С. 1–26. DOI: 10.7256/2454-0714.2025.4.75830.

16. Обнаружение объектов в аэронавигации с использованием вейвлет-преобразования и сверточных нейронных сетей: первый подход / Х. Фортуна-Сервантес, М. Рамирес-Торрес, Х. Мартинес-Карранса, Х. Мургуия-Ибарра, М. Мехия-Карлос // Труды Института системного программирования РАН. – 2021. – Т. 33, № 2. – С. 149–162. DOI: 10.15514/ISPRAS-2020-33(2)-9.

17. Lowe D.G. Distinctive Image Features from Scale-Invariant Keypoints // International Journal of Computer Vision. – 2004. – Vol. 60. – P. 91–110. DOI: 10.1023/B:VISI.0000029664.99615.94.

18. Денисов А.А. Анализ методов детектирования, описания и сопоставления ключевых точек изображений / А.А. Денисов, А.И. Новиков // Вестник Рязан. гос. радиотехн. ун-та. – 2024. – № 89. – С. 104–116.

19. Bay H. SURF: Speeded Up Robust Features / H. Bay, T. Tinne, V.G. Luc // Proc. of Computer Vision – ECCV 2006. – Lecture Notes in Computer Science. – 2006. – Vol. 3951. – Berlin, Heidelberg: Springer, 2006. https://doi.org/10.1007/11744023_3.

20. Mikolajczyk K. A performance evaluation of local descriptors / K. Mikolajczyk, C. Schmid // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2005. – P. 1615–1630. DOI: 10.1109/TPAMI.2005.188.

21. Дышлюк В.О. Исследование показателей качества и быстродействия поиска опорных точек на изображениях методом SURF // Молодой ученый. – 2018. – № 27. – С. 23–26.

22. DeTone D., Malisiewicz T., Rabinovich A. SuperPoint: Self-Supervised Interest Point Detection and Description // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). – Salt Lake City, UT, USA: IEEE. DOI: 10.1109/CVPRW.2018.00060.

23. Костров Б.В. Разработка гибридной архитектуры свёрточной нейронной сети с использованием спектральных преобразований / Б.В. Костров, С.И. Бабаев, А.И. Ефимов, В.Ю. Тарасова // Известия ЮФУ. Технические науки. – 2026. – № 1 (249). – С. 235–249. DOI: 10.18522/2311-3103-2026-1-235-249.

24. Костров Б.В. Подходы к формированию набора данных и его предварительной обработки в задачах обучения свёрточных сиамских нейронных сетей для коррекции положения летательного аппарата по аэрофотосъемке / Б.В. Костров, С.И. Бабаев, А.И. Ефимов, В.Ю. Тарасова // XXI век: итоги прошлого и проблемы настоящего плюс. – 2025. – Т. 14, № 3 (71). – С. 70–77.

25. Li J. Point Cloud Registration Based on One-Point RANSAC and Scale-Annealing Biweight Estimation / J. Li, Q. Hu, M. Ai // IEEE Transactions on Geoscience and Remote Sensing. – 2021. – Vol. 59, No. 11. – P. 9716–9729. DOI: 10.1109/TGRS.2020.3048026.

26. Hu S. CVM-Net: Cross-View Matching Network for Image-Based Ground-to-Aerial Geo-Localization / S. Hu, M. Feng, R.M.H. Nguyen, G.H. Lee // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018. Salt Lake City, UT, USA: IEEE, 2018. – P. 7258–7267. DOI: 10.1109/CVPR.2018.00758.

27. Св-во о гос. регистрации программы для ЭВМ № 2025691610 / Костров Б.В., Устюков Д.И., Тарасова В.Ю. // Свёрточная нейронная сеть с гибридными слоями для опре-

деления визуального сходства спутниковых снимков. Дата поступления 17 ноября 2025 г. Зарегистрировано в Реестре программ для ЭВМ 17 ноября 2025 г.

28. Никифоров М.Б. Метод определения местоположения съемки фотографии на основе визуального поиска / М.Б. Никифоров, В.Ю. Тарасова // XXI век: итоги прошлого и проблемы настоящего плюс. – 2023. – Т. 12, № 4 (64). – С. 103–110.

Бабаев Сергей Игоревич

Канд. техн. наук, доцент каф. электронных вычислительных машин (ЭВМ) Рязанского государственного радиотехнического университета им. В.Ф. Уткина» (РГРТУ) Гагарина ул., 59/1, г. Рязань, Россия, 390005
ORCID: 0000-0001-5829-8223
Тел.: +7 (491-2) 72-03-50
Эл. почта: babaev.s.i@gmail.com

Костров Борис Васильевич

Д-р техн. наук, проф., зав. каф. ЭВМ РГРТУ Гагарина ул., 59/1, г. Рязань, Россия, 390005
ORCID: 0000-0002-2984-4480
Тел.: +7 (491-2) 72-03-50
Эл. почта: kostrov.b.v@evm.rsreu.ru

Муратов Евгений Рашитович

Канд. техн. наук, доцент каф. ЭВМ РГРТУ Гагарина ул., 59/1, г. Рязань, Россия, 390005
ORCID: 0000-0002-1664-3954
Тел.: +7 (491-2) 72-03-50
Эл. почта: muratov_er@mail.ru

Тарасов Андрей Сергеевич

Канд. техн. наук, доцент каф. ЭВМ РГРТУ Гагарина ул., 59/1, г. Рязань, Россия, 390005
ORCID: 0000-0002-2812-3489
Тел.: +7 (491-2) 72-03-50
Эл. почта: vb2005@yandex.ru

Тарасова Валентина Юрьевна

Ст. преподаватель каф. ЭВМ РГРТУ Гагарина ул., 59/1, г. Рязань, Россия, 390005
ORCID: 0009-0007-4358-7736
Тел.: +7-903-837-37-69
Эл. почта: valentina2008.91@mail.ru

Поступила в редакцию: 15.04.2026.

Принята к публикации: 31.08.2026.

Babaev S.I., Kostrov B.V., Muratov E.R.,
Tarasov A.S., Tarasova V.Yu.

Keypoint-based algorithm for registration of heterogeneous images

This paper discusses an algorithm for refining heterogeneous images registration. Image registration is performed using neural network-based selection of candidate regions. **Relevance.** The registration of heterogeneous images can serve as a basis for unmanned aerial vehicle (UAV) positioning systems, which is a relevant task when signals from global navigation satellite systems are unavailable or distorted due to interference.

Purpose of the study. The purpose of this study is to develop and experimentally evaluate an image registration refinement algorithm that combines fast neural network-based selection of candidate regions with precise positioning based on matching and geometric verification of keypoints. Particular attention is paid to scenarios in which a frame acquired by an onboard UAV sensor crosses the boundaries of multiple satellite imagery tiles.

Scientific novelty. The scientific novelty lies in the position of a UAV in the absence of signals from global navigation satellite systems by registering heterogeneous images of the underlying surface within a specified flight area. At the first stage, candidate regions are selected using a Siamese convolutional neural network. The network analyzes pairs consisting of a current onboard image and images of the underlying surface, calculating a numerical visual similarity score for each pair, thereby reducing the search area. At the second stage, coordinates are refined based on keypoint matching.

Methods. Methods for detecting, describing, and matching local features are considered. As a result of applying the refinement algorithm to the test sample, the average alignment error decreased from 300 pixels after neural network selection to 20 pixels. The most accurate results were obtained for cases in which the UAV frame was entirely located within a single satellite imagery tile.

Practical significance. The developed algorithm enables UAV positioning with an accuracy of 10–20 m and a processing time of ≤ 207 ms based on heterogeneous images of the underlying surface.

Keywords: heterogeneous images, image registration, Siamese neural network architecture, keypoints, visual similarity.
DOI: 10.21293/1818-0442-2026-29-2-107-115

References

1. Novikov A.I., Efimov A.I., Kolchaev D.A. [Algorithms for automatic identification of objects on heterogeneous images and images superimposition]. *Informatsionnye tekhnologii i nanotekhnologii. [Information Technologies and Nanotechnologies]* Proc. of IV International Conference and Youth School ITNT–2018, Samara, Novaya tekhnika, 2018, pp. 650–663 (in Russ.).
2. Efimov A.I., Kryuchkova T.N., Yaroslavtseva A.I. [Combined Algorithm for Combining Point Clouds]. Proc. of 33rd International Conference on Computer Graphics and Vision GraphiCon 2023, 2023, no. 33, pp. 539–548 (in Russ.).
3. Sablina V.A., Efimov A.I., Novikov A.I. Combined approach to object contour superimposition in heterogeneous images. 2017 6th Mediterranean Conference on Embedded Computing (MECO), 2017, pp. 7977175.
4. Lebedev M.A., Bondarenko M.A., Komarov D.V., Stepaniants D.G., Vygodov O.V., Vizilter Yu.V., Zheltov S.Yu. [An algorithm for real and synthetic images matching with application to aviation combined vision system]. *Vestnik komp'yuternykh i informatsionnykh tekhnologii*, 2014, no. 7 (121), pp. 8–13 (in Russ.).
5. Vizilter Yu.V., Gorbatshevich V.S., Rubis A.Yu., Vygodov O.V. Image shape matching using diffusion morphology and diffusion correlation. *Computer optics*, 2015, vol. 39, no. 2, pp. 265–274 (in Russ.).
6. Vizilter Yu.V., Vishnyakov B.V., Zheltov S.Yu. *Sovremennyye tekhnologii iskusstvennogo intellekta i ikh primeneniye v aviatsionnykh kompleksakh* [Modern artificial intelligence technologies and their application in aviation systems]. XVI Vserossiyskaya mul'tikonferentsiya po problemam upravleniya (MKPU-2023) [Proc. of XVI All-Russian Multiconference on Control Problems], Volgograd, Volgograd State Technical University, 2023, pp. 13–16 (in Russ.).
7. Kocheturov V.V., Fedyaev O.I. *Siameskie svertochnyye neyronnyye seti dlya raspoznavaniya lits*. [Siamese convolutional neural networks for face recognition] Programmaya inzheneriya: metody i tekhnologii razrabotki informatsionno-vychislitel'nykh sistem (PIIVS-2024) [Software Engineering: Methods and Technologies for Developing Information and Computing Systems]. Proc. of the 5th International Scientific and Practical Conference, Donetsk, Donetsk National Technical University, 2024, pp. 178–184 (in Russ.).
8. Ladanova E.O. [Development of a methodology and a new structure of the Siamese neural network for analysis of geospatial data]. *Scientific and Technical Volga region Bulletin*, 2024, no. 12, pp. 144–147 (in Russ.).
9. Li Y., Chen C.L.Ph., Zhang T. A Survey on Siamese Network: Methodologies, Applications, and Opportunities. *IEEE Transactions on Artificial Intelligence*, 2022, vol. 3, no. 6, pp. 994–1014. DOI: 10.1109/TAI.2022.3207112.
10. Broughton G., Linder P., Rouček T., Vintř T., Krajník T. Robust Image Alignment for Outdoor Teach-and-Repeat Navigation. *Proceedings of the 2021 European Conference on Mobile Robots (ECMR)*, Bonn, Germany, IEEE, 2021, pp. 1–6. DOI: 10.1109/ECMR50962.2021.9568832.
11. Rozsypalek Z., Broughton G., Linder P., Rouček T., Kusumam K., Krajník T. Semi-Supervised Learning for Image Alignment in Teach and Repeat Navigation. *Proceedings of the 37th ACM/SIGAPP Symposium on Applied Computing (SAC '22)*, New York, USA, ACM, art. 4. DOI: 10.1145/3477314.3507045
12. Teague S., Chahl J. An Algorithm for Affordable Vision-Based GNSS-Denied Strapdown Celestial Navigation. *Drones*, 2024, vol. 8, no. 11, art. 652. DOI: 10.3390/drones 8110652.
13. Alpatov B.A., Babayan P.V., Koblov Yu.S., Muravyev V.S., Strotov V.V., Feldman A.B. [Software for automatization of scientific investigations aimed at development of perspective vision-based navigation systems for unmanned aerial vehicles]. *Vestnik of Ryazan state radio engineering university*, 2012, no. 40, pp. 9–15 (in Russ.).
14. Sheinikov A.A., Malkin V.A., Smolensky E.A., Ivanitsky L.A. [Optical-electronic correlation-extreme navigation system of small unmanned aerial vehicle]. *System analysis and applied information science*, 2024, no. 2, pp. 30–40. DOI: 10.17212/2307-5981-2024-2-30-40 (in Russ.).
15. Konnikov E.A., Sergeev D.A., Golikov G.I., Rodionov D.G., Polyakov P.A., Tikhomirov A.F. [TTC-oriented visual navigation systems for UAVS]. *Software systems and computational methods*, 2025, no. 4, pp. 1–26. DOI: 10.7256/2454-0714.2025.4.75830 (in Russ.).
16. Fortuna-Cervantes J., Ramírez-Torres M., Martínez-Carranza J., Murguía-Barra J., Mejía-Carlos M. Object Detection in Aerial Navigation using Wavelet Transform and Convolutional Neural Networks: A first Approach. *Proceedings of the Institute for System Programming of the RAS (Proceedings of ISP RAS)*, 2021, vol. 33, no. 2, pp. 149–162. DOI: 10.15514/IS-PRAS-2020-33(2)-9 (in Russ.).
17. Lowe D.G. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 2004, vol. 60, pp. 91–110. DOI: 10.1023/B:VISI.0000029664.99615.94.
18. Denisov A.A., Novikov A.I. [Analysis of image key point detection, description and matching methods]. *Vestnik of Ryazan state radio engineering university*, 2024, no. 89, pp. 104–116 (in Russ.).
19. Bay H., Tinne T., Luc V.G. SURF: Speeded Up Robust Features. Proc. of Computer Vision – ECCV 2006. ECCV 2006, *Lecture Notes in Computer Science*, vol. 3951. Springer, Berlin, Heidelberg. https://doi.org/10.1007/11744023_3.
20. Mikolajczyk K., Schmid C.A. Performance Evaluation of Local Descriptors. *IEEE Transactions on Pattern Ana-*

lysis and Machine Intelligence, 2005, pp. 1615–1630. DOI: 10.1109/TPAMI.2005.188.

21. Dyshlyuk V.O. *Issledovanie pokazateley kachestva i bystrodeystviya poiska opornykh tochek na izobrazheniyakh metodom SURF* [Investigation of quality and speed metrics for keypoint detection in images using the SURF method]. *Molodoy uchenyi*, 2018, no. 27, pp. 23–26 (in Russ.).

22. DeTone D., Malisiewicz T., Rabinovich A. SuperPoint: Self-Supervised Interest Point Detection and Description. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, USA, IEEE. DOI: 10.1109/CVPRW.2018.00060.

23. Kostrov B.V., Babaev S.I., Efimov A.I., Tarasova V.Yu. [Convolutional neural network hybrid architecture development using spectral transformations]. *Izvestiya SFedU. Engineering Sciences*, 2026, no. 1 (249), pp. 235–249. DOI: 10.18522/2311-3103-2026-1-235-249 (in Russ.).

24. Kostrov B.V., Babaev S.I., Tarasova V.Yu. [Issues to the data set formation and preprocessing in the training convolutional Siamese neural networks tasks for an aircraft position correcting based on aerial photography]. *XXI century: resumes of the past and challenges of the present plus*, 2025, vol. 14, no. 3 (71), pp. 70–77 (in Russ.).

25. Li J., Hu Q., Ai M. Point Cloud Registration Based on One-Point RANSAC and Scale-Annealing Biweight Estimation. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, vol. 59, no. 11, pp. 9716–9729. DOI: 10.1109/TGRS.2020.3048026.

26. Hu S., Feng M., Nguyen R.M.H., Lee G.H. CVM-Net: Cross-View Matching Network for Image-Based Ground-to-Aerial Geo-Localization. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, Salt Lake City, UT, USA: IEEE, pp. 7258–7267. DOI: 10.1109/CVPR.2018.00758.

27. Kostrov B.V., Ustyukov D.I., Tarasova V.Yu. *Svertochnaya neyronnaya set' s gibridnymi sloyami dlya opredeleniya vizual'nogo skhodstva sputnikovyykh snimkov* [Convolutional neural network with hybrid layers for determining the visual similarity of satellite images]. Certificate of State Registration of a Computer Program No. 2025691610, 2025 (in Russ.).

28. Nikiforov M.B., Tarasova V.Yu. [Method for determining photo shooting location based on visual search]. *XXI century: resumes of the past and challenges of the present plus*, 2023, vol. 12, no. 4 (64), pp. 103–110 (in Russ.).

Sergey I. Babaev

Candidate of Technical Sciences, Associate Professor,
Department of Electronic Computers, Ryazan State Radio
Engineering University named after V.F. Utkin (RSREU)
59/1, Gagarin St., Ryazan, Russia, 390005
ORCID: 0000-0001-5829-8223
Phone: +7 (491-2) 72-03-50
Email: babaev.s.i@gmail.com

Boris V. Kostrov

Doctor of Engineering Sciences, Professor, Head
Department of Electronic Computers, RSREU
59/1, Gagarin St., Ryazan, Russia, 390005
ORCID: 0000-0002-2984-4480
Phone: +7 (491-2) 72-03-50
Email: kostrov.b.v@evm.rsreu.ru

Yevgeniy R. Muratov

Candidate of Technical Sciences, Associate Professor
Department of Electronic Computers, RSREU
59/1, Gagarin St., Ryazan, Russia, 390005
ORCID: 0000-0002-1664-3954
Phone: +7 (491-2) 72-03-50
Email: myratov_er@mail.ru

Andrey S. Tarasov

Candidate of Technical Sciences, Associate Professor,
Department of Electronic Computers, RSREU
59/1, Gagarin St., Ryazan, Russia, 390005
ORCID: 0000-0002-2812-3489
Phone: +7 (491-2) 72-03-50
Email: vb2005@yandex.ru

Valentina Yu. Tarasova

Senior Lecturer, Department of Electronic Computers, RSREU
59/1, Gagarin St., Ryazan, Russia, 390005
ORCID: 0009-0007-4358-7736
Phone: +7-903-837-37-69
Email: valentina2008.91@mail.ru

Received: 15.04.2026.

Accepted: 31.08.2026.