

**Санкт-Петербургский политехнический университет
Петра Великого
Институт физики, нанотехнологий и телекоммуникаций**

На правах рукописи

Павлов Виталий Александрович

**Обработка динамических изображений в телевизионной системе селекции
объектов с адаптацией нейросетевого вычислителя и фотоприёмника**

Направление подготовки 11.06.01 «Электроника, радиотехника и системы связи»

Код и наименование

Направленность 11.06.01_01 «Радиотехника, в том числе системы и устройства телевидения»

Код и наименование

НАУЧНЫЙ ДОКЛАД

об основных результатах научно-квалификационной работы (диссертации)

Автор работы: Павлов Виталий Александрович
Научный руководитель: д.т.н., профессор,
Макаров Сергей Борисович

Санкт Петербург – 2020

Научно-квалификационная работа выполнена в ВШ Института физики, нанотехнологий и телекоммуникаций федерального государственного автономного образовательного учреждения высшего образования «Санкт-Петербургский политехнический университет Петра Великого»

Директор Высшей школы:

Величко Елена Николаевна,
к. т. н., доцент

Научный руководитель:

Макаров Сергей Борисович,
д. т. н., профессор

Рецензент:

Чесноков Михаил Николаевич
д. т. н., профессор,
ЗАО «Институт телекоммуникаций», г. Санкт-Петербург, ведущий сотрудник

С научным докладом можно ознакомиться в библиотеке ФГАОУ ВО «Санкт-Петербургский политехнический университет Петра Великого» и на сайте Электронной библиотеки СПбПУ по адресу: <http://elib.spbstu.ru>

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность работы. Настоящая работа посвящена решению части крупной научной проблемы разработки методов синтеза телевизионных систем с учётом совместной адаптации фотоприёмных матриц и бортового нейросетевого вычислителя.

Работа состоит в разработке и исследовании нейросетевых методов, алгоритмов и устройств обнаружения, распознавания, сопровождения и адаптивного кодирования для бортовых телевизионных систем видеоконтроля. В работе производится:

- синтез и исследование алгоритма обнаружения и распознавания подвижных искусственных космических объектов (ИКО) на сложном фоне для адаптации кадровой частоты телевизионных бортовых систем;
- разработка метода адаптации телевизионных систем к динамике сюжета на дальней дистанции;
- разработка и исследование метода адаптивного кодирования видеопоследовательностей делимых на доминантный и фоновые объекты;
- разработка метода обнаружения, распознавания и сопровождения группы наземных объектов в условиях подвижной камеры и фона.

В теоретическом плане актуальность работы связана с необходимостью выбора оптимальных параметров и синтеза на их основе бортовых телевизионных систем наблюдения.

В методологическом плане исследование методов обнаружения, распознавания, сопровождения и адаптивного кодирования изображений объектов в видеопоследовательности является актуальным в силу возросшей потребности в автоматизации и автономности работы бортовых телевизионных систем.

С практической точки зрения, актуальность работы заключается в потребности создания бортовых телевизионных систем для обнаружения, распознавания и сопровождения объектов на звёздном небе и земной поверхности.

Ключевой особенностью предлагаемой в данной работе системы на основе нейросетевых методов является устойчивость к изменению сюжета и движению доминантных и фоновых объектов.

Цель и задачи исследования. Цель работы заключается в разработке нейросетевых методов обработки динамических изображений искусственных космических и наземных объектов бортовыми телевизионными системами для повышения точности распознавания, снижения передаваемой фоновой информации и устойчивости к сложной фоно-целевой обстановке при сопровождении.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Разработка и исследование нейросетевого метода обнаружения ИКО на звёздном фоне.
2. Разработка метода адаптации кадровой частоты фотоприёмника.
3. Разработка и исследование нейросетевого метода адаптивного кодирования изображений в задачах автоматического обнаружения ИКО.
4. Разработка и исследование метода обнаружения, распознавания и сопровождения группы наземных объектов с борта летательного аппарата.

Научная новизна работы заключается в следующем:

1. Синтезирована структура адаптивной бортовой телевизионной системы с автоматической селекцией ИКО по динамическим признакам с помощью свёрточной нейронной сети. Получены оценки быстродействия и аппаратных затрат, приемлемых для бортовой реализации.
2. Синтезирована структура бортовой телевизионной системы адаптации кадровой частоты фотоприёмника.
3. Синтезирована структура бортовой телевизионной системы для адаптивного кодирования видеопоследовательностей, делимых на доминантный и фоновые объекты.
4. Выявлена зависимость сложности структур свёрточных нейронных сетей семейства *YOLO* от сложности модели сцены, включающей сложный фон и наземные обнаруживаемые объекты при наблюдении с борта летательного аппарата.

Теоретическая и практическая значимость. В теоретическом плане в работе осуществляется попытка объединения методов теории решений и теории информации для определения параметров системы видеонаблюдения. В результате было

найден оптимальное значение времени накопления кадра, при котором значение информационного риска минимально. Установлено, что:

- При превышении пороговых значений амплитуды и смаза сигнала, информационный риск сводится к нулю и достигается максимальное значение качества передаваемой информации.
- Для задачи селекции космических аппаратов (КА) на звёздном фоне при широком диапазоне изменения дальности и, как следствие, амплитуды и формы сигналов применение нейросетевых методов обеспечивает высокую надёжность обнаружения, распознавания и сопровождения сигналов объектов.

Апробация работы. Работа в целом обсуждена на секции НТС АО «НИИ телевидения» и отдельные вопросы обсуждены на конференциях:

- Международная конференция «*Emerging Trends in Applied and Computational Physics*», 2019 г.
- Международная конференция молодых ученых «*International youth conference on electronics, telecommunications and information technologies*» (YETI), 2019 г.;
- Международная конференция «*IEEE EExPolytech-2019: Electrical Engineering and Photonics*», 2019 г.;
- 15-я международной конференции «Телевидение: передача и обработка изображений» 2018 г.
- Всероссийская астрометрическая конференция «Пулково-2018», 2018 г.
- 18-я международная конференция «Цифровая обработка сигналов и ее применение - *DSPA-2016*», 2016 г.
- XLIV «Неделя науки» СПбПУ, 2015 г.

Публикации. По материалам диссертации опубликовано 18 научных работ, из них 9 статей в журналах из перечня ВАК, в том числе 1 статья без соавторов; 8 докладов на научно-технических конференциях; 1 свидетельство о государственной регистрации программы для ЭВМ.

Положения, выносимые на защиту. В ходе решения поставленных задач были сформулированы следующие положения:

1. Для повышения дальности автоматического обнаружения ИКО в условиях больших вариаций дистанций и формы сигналов, объектов и фонов при динамическом смазе изображений следует применять адаптацию кадровой частоты и обнаружитель на базе свёрточных нейронных сетей.

2. Для повышения сжатия сигналов изображений с сохранением информации о доминантном объекте следует использовать разделение видеопоследовательностей на доминантный объект и фон с помощью нейросетевого метода, сегментацию изображения и различные значения коэффициента сжатия для объекта и фона.

3. Сложность структуры и обучающей выборки свёрточных нейронных сетей семейства *YOLO* при наблюдении с борта летательного аппарата определяется сложностью модели сцены, включающей сложный фон и наземные обнаруживаемые объекты.

СОДЕРЖАНИЕ РАБОТЫ

Работа состоит из введения, четырёх глав, заключения, списка опубликованных работ, используемой литературы и приложений.

Во введении обоснована актуальность диссертационной работы, сформулирована цель и задачи исследования, описана научная новизна, показаны теоретическая и практическая значимость, изложены основные положения, выносимые на защиту.

В первой главе рассмотрены проблемы обнаружения объектов, производится обзор существующих методов обнаружения, распознавания и сопровождения мало-размерных искусственных космических и наземных объектов на изображениях и видеопоследовательностях; рассматриваются проблема сжатия изображений и сегментация. Показано, что для обнаружения и распознавания малоразмерных искусственных космических и наземных объектов следует применять подходы на основе свёрточных нейронных сетей. Для адаптивного кодирования изображений следует опираться на принцип доминантной информации.

Вторая глава посвящена решению задачи обнаружения и распознавания ИКО на фоне звёзд для случая подвижной камеры и изменяющегося фона по динамическим признакам (смаз) с помощью нейросетевого обнаружителя. Проведённые исследования позволили определить: оптимальный порог, при котором значение риска стремится к нулю; оптимальное значение времени накопления кадра, которое определяет значение скоростного смаза; необходимую сложность свёрточной нейронной сети. Полученные результаты позволили синтезировать структурную схему бортовой телевизионной системы наблюдения с адаптацией кадровой частоты.

Третья глава посвящена адаптивному кодированию изображений в задачах автоматического сопровождения ИКО на ближней дистанции. Предложен алгоритм адаптивного выделения доминантного объекта и подавления фона, в рамках которого синтезирована бортовая телевизионная система адаптивного кодирования видеопоследовательностей, разделяющая изображения на доминантный объект и фон. Данная система позволяет осуществлять сжатие информации с разными коэффициентами качества для объекта и фона.

Четвёртая глава посвящена обнаружению, распознаванию и сопровождению группы наземных объектов на последовательности аэрофотоснимков, которые характеризуются многообразием классов объектов и изменяющимся фоном. По результатам исследований нейросетевых подходов сделан вывод, что с ростом сложности модели сцены растёт сложность архитектуры нейронной сети (число слоёв). На основе обученного нейросетевого обнаружителя получен алгоритм сопровождения нескольких объектов в условиях подвижной камеры и изменяющегося фона.

Объектом исследования является телевизионная система, включающая нейросетевую вычислитель, используемый для задач обнаружения, распознавания, сопровождения и адаптивного кодирования изображений объектов.

Предметом исследования в данной работе являются методы адаптации вычислителей и фотоприёмников телевизионных систем к динамическим свойствам сюжета при наблюдении за подвижными объектами на сложном фоне.

Основные методы исследования, применяемые в работе:

1. методы теории решений и теории информации;
2. методы компьютерного моделирования;
3. методы экспериментальной оценки.

Результаты

Для обнаружения и распознавания точечных ИКО на звёздном фоне, примеры которых приведены на рисунке 1, в работе предложено применять нейросетевой обнаружитель. Длина подвижного объекта (смаза), изображённого на рисунке 1 слева, зависит от времени накопления кадра и является доминантным признаком различия сигналов при обнаружении ИКО.

Для формализации задачи обнаружения сигнала объекта на звёздном фоне вводятся три гипотезы: H_0 – наличие шума; H_1 – наличие фона и шума; H_2 – наличие сигнала объекта, фона и шума. Количество информации о координате объекта вычисляется по формуле Вудворда:

$$I_{д\epsilon} = \frac{1}{2} \log \frac{T^2}{2\pi e \epsilon_0} = \frac{1}{2} \log \frac{T^2 \Psi^2 \beta^2}{2\pi e} = \log \left(\frac{N\Psi}{\sqrt{2\pi e}} \right), \quad (1)$$

где T – интервал наблюдения, Ψ – отношение сигнал/шум, β – количество пикселей.



Рисунок 1 – Изображения объектов на расстоянии более 1000 км. Левый рисунок – подвижный объект (негатив). Правый рисунок – неподвижный объект (негатив).

Потеря доминантной информации при обнаружении вычисляется, как:

$$\Delta I_{\text{д}} = [p(H_0|H_2) + p(H_1|H_2)] \log \left(\frac{N\Psi}{\sqrt{2\pi e}} \right), \quad (2)$$

где $p(H_0|H_2) + p(H_1|H_2)$ – вероятность пропуска объекта, $N = T\beta$ – число пикселей.

Шумовая информация вычисляется по следующей формуле:

$$I_{\text{ш}} = (N - M - 1) p(H_2 | H_0) \log \left(\frac{N\Psi}{\sqrt{2\pi e}} \right), \quad (3)$$

где $p(H_2|H_0)$ – вероятность ложной тревоги, M – число фоновых объектов.

Фоновая информация вычисляется, как:

$$I_{\text{ф}} = Mp(H_2 | H_1) \log \left(\frac{N\Psi}{\sqrt{2\pi e}} \right). \quad (4)$$

Информационный риск вычисляется по формуле:

$$R = [(N - M - 1) p(H_2 | H_0) + Mp(H_2 | H_1) + p(H_0 | H_2) + p(H_1 | H_2)] \log \left(\frac{N\Psi}{\sqrt{2\pi e}} \right). \quad (5)$$

Решение задачи обнаружения и классификации ИКО с помощью принципа доминантной информации осложняется следующими особенностями:

- форма разностного сигнала зависит от значения смаза и амплитуд сигналов;
- наличие априорной неопределённости амплитуды и формы доминантного и фоновых сигналов.

В работе показано, что для обнаружения объектов по динамическим признакам перспективно применять нейросетевые подходы, а именно свёрточную нейронную

сеть *TinyYOLOv2*. В таблице 1 приведены значения вероятностей правильного обнаружения неподвижных (P_1) и подвижных (P_2) объектов; вероятностей ложного обнаружения неподвижных (P_{EF1}) и подвижных объектов (P_{EF2}); абсолютная (Δ) и относительная (δ) погрешность локализации ограничивающего прямоугольника для разного числа слоёв нейросетевого обнаружителя. Результаты получены на наборе данных, который содержал 1000 изображений объектов класса 1 и 2. Обучающий набор данных содержал 2500 изображений подвижных (смаз) и неподвижных (точечный объект) объектов.

Таблица 1 – Оценки работы сети *TinyYOLOv2* для разного количества слоёв.

Число сверточных слоев	P_1	P_2	$P_{\text{среднее}}$	P_{EF1}	P_{EF2}	Δ , пиксель	δ	среднее время, с
9	0,90	0,90	0,90	$6,0 \times 10^{-6}$	$2,6 \times 10^{-6}$	3,0	0,1	0,25
8	0,88	0,89	0,89	$6,0 \times 10^{-6}$	$3,0 \times 10^{-6}$	3,1	0,1	0,20
7	0,84	0,85	0,85	$6,5 \times 10^{-6}$	$4,5 \times 10^{-6}$	3,1	0,1	0,15
6	0,80	0,80	0,80	$7,0 \times 10^{-6}$	$5,0 \times 10^{-6}$	5,3	0,2	0,10

Пример работы *TinyYOLOv2*, обученной на реальных изображениях космических объектов, представлен на рисунке 2.

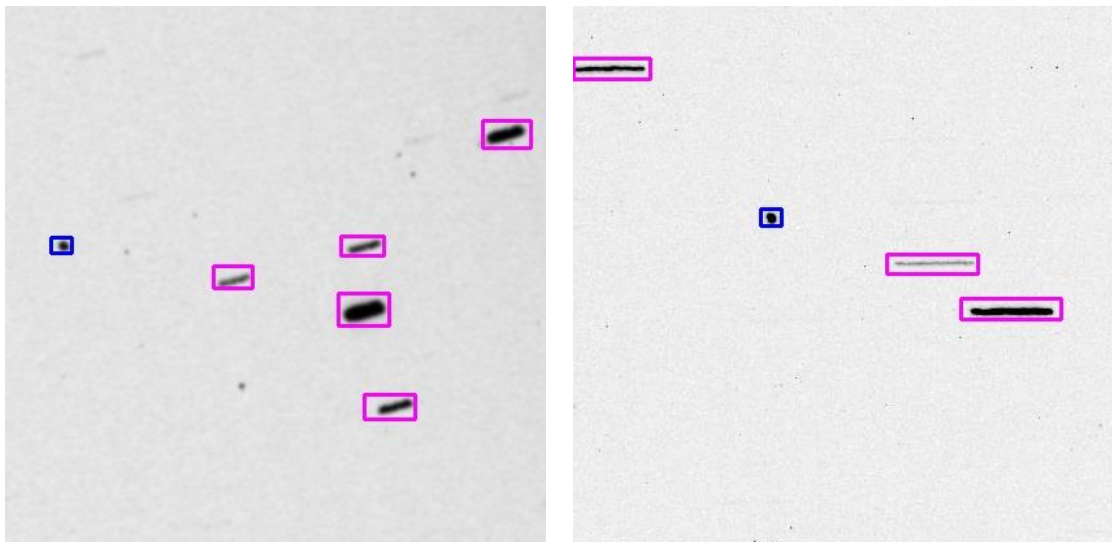


Рисунок 2 – Пример работы *TinyYOLOv2* на двух фрагментах (негатив) с разрешением 416×416 . Синий цвет – неподвижный точечный объект, фиолетовый цвет – подвижный объект.

На рисунке 3 представлены теоретическая и экспериментальная зависимости нормированного информационного риска от времени накопления кадра для разного количества слоёв *TinyYOLOv2*.

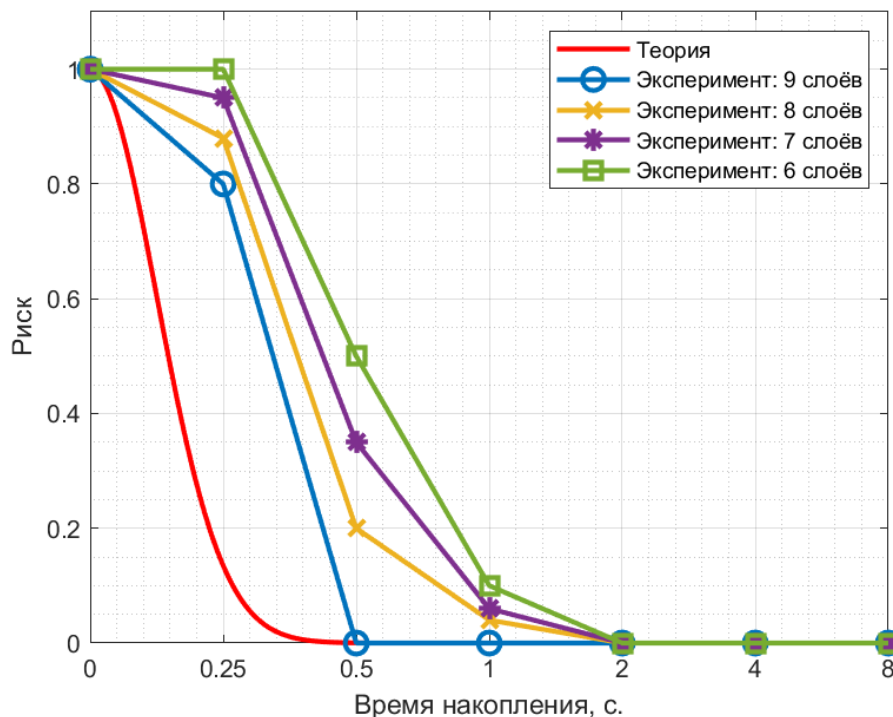


Рисунок 3 – Зависимость нормированного информационного риска от времени накопления кадра для разного количества слоёв *TinyYOLOv2*.

В таблице 2 приведена вычислительная сложность нейросетевого детектора для разного числа слоёв. Вычислительная сложность для одного свёрточного слоя вычисляется по формуле: $O = C \cdot M \cdot N \cdot H \cdot W \cdot S^2$, где C – количество каналов; $M = 2$ – количество типов арифметических операций; N – количество свёрточных фильтров; H и W – высота и ширина карты признаков после операции свёртки, соответственно; S – размер свёрточных фильтров.

Таблица 2 – Вычислительная сложность нейросетевого детектора для разного количества слоёв.

Количество слоёв	Вычислительная сложность (входное разрешение: 416×416 пикселей)
9	$6,946 \cdot 10^9$
8	$2,952 \cdot 10^9$
7	$2,559 \cdot 10^9$
6	$1,556 \cdot 10^9$

В результате исследований получено, что структура сети с 6 свёрточными слоями является перспективной для использования в условиях жестких ресурсных ограничений.

При параметрическом синтезе системы наблюдения за ИКО, необходимо определить оптимальное значение кадровой частоты (время накопления кадра) и метод её перестройки во время процесса наблюдения. Применение адаптации кадровой частоты позволяет увеличить дальность обнаружения при малых значениях скоростного смаза ИКО.

Оптимизация кадровой частоты рассматривается для случая, в котором фон является подвижным (скоростной смаз), а объект неподвижным (точечный объект). Оптимальное время накопления кадра вычисляется по следующей формуле:

$$T_{к\text{ опт}} = \operatorname{argmin}\{c_{\tau}T_{к} + c_{R}(R_{\phi} + R_{ш})\}, \quad (6)$$

где риск R включает в себя: R_{ϕ} – влияние фона и $R_{ш}$ – влияние шума; c_{τ} и c_{R} – весовые коэффициенты, τ – задержка принятия решения, $T_{к}$ – время накопления кадра. Таким образом, для минимизации τ требуется выбрать значение $T_{к}$, при котором значение R стремится к нулю.

Оптимальное значение $T_{к}$ определяется следующим выражением:

$$T_{\text{ опт}} = \max\{T_{\psi\text{ мин}}, T_{s\text{ мин}}\}, \quad (7)$$

где $T_{\psi\text{ мин}}$ – время накопления кадра, обеспечивающее ОСШ выше порогового; $T_{s\text{ мин}}$ – время накопления кадра, которое обеспечивает значение смаза выше порогового.

Для работы алгоритма адаптации времени накопления кадра вводятся два порога: верхний $\gamma_{в}$ и нижний $\gamma_{н}$. При адаптации времени накопления кадра значения порогов должны удовлетворять следующему условию: $\gamma_{в}/\gamma_{н} = 2 + \delta$, где δ – компонент, определяемый с помощью среднеквадратических значений ошибок оценивания смаза и ОСШ. Таким образом, выражение (7) принимает вид:

$$\max\{T_{\psi\text{ мин}}, T_{s\text{ мин}}\} < T_{\text{ опт}} < (2 + \delta)\max\{T_{\psi\text{ мин}}, T_{s\text{ мин}}\}. \quad (8)$$

Структурная схема бортовой телевизионной системы с адаптацией кадровой частоты представлена на рисунке 4. Для стабилизации работы системы в её структуру вводятся два контура отрицательной обратной связи. Эти контуры содержат фильтр

сглаживания оценки \hat{s} смаза изображений фоновых звёзд и оценки $\hat{\psi}$ ОСШ выделяемого КО. Уравнение управления выглядит следующим образом:

$$Y_{k+1} = \begin{cases} Y_{-k+1} = \text{sign}(\hat{s}_k - \gamma_{sb}) \wedge \text{sign}(\hat{\psi}_k - \gamma_{\psi b}) \\ Y_{+k+1} = \text{sign}(\gamma_{sh} - \hat{s}_k) \vee \text{sign}(\gamma_{\psi h} - \hat{\psi}_k) \end{cases}, \quad (9)$$

где Y_{k+1} – сигнал управления; γ_{sb} и γ_{sh} – верхний и нижний порог для смаза, соответственно; $\gamma_{\psi b}$ и $\gamma_{\psi h}$ – верхний и нижний порог для ОСШ, соответственно.

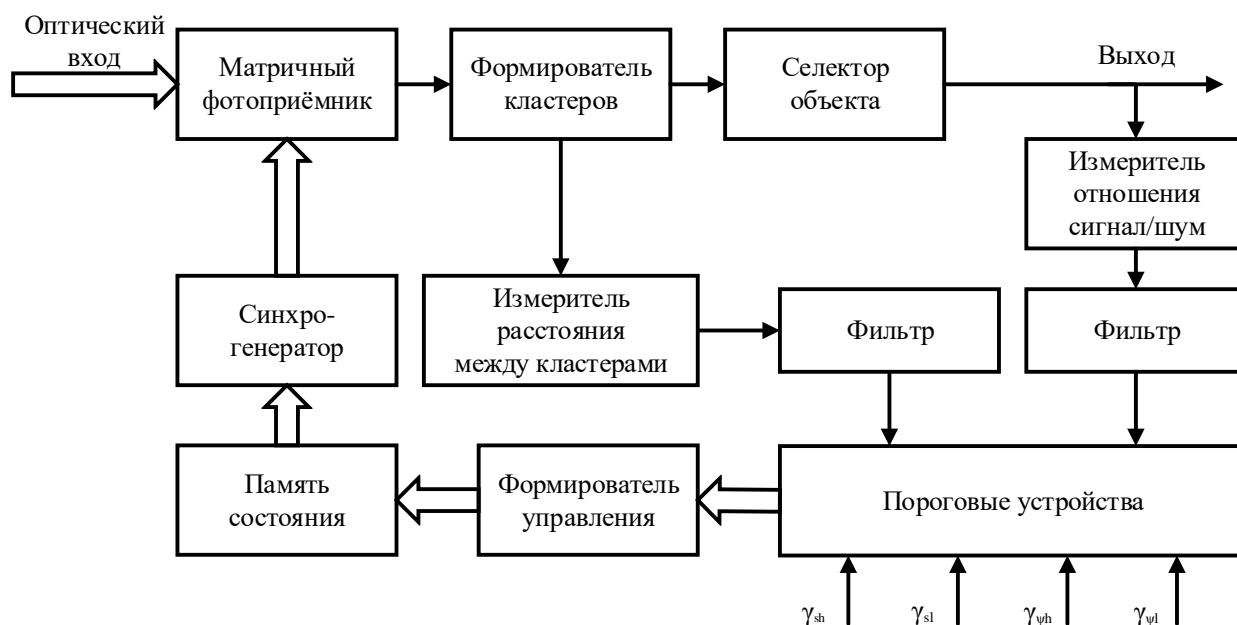


Рисунок 4 – Структурная схема бортовой телевизионной системы с адаптацией кадровой частоты.

Сигналы управления подаются на устройство памяти состояния. Для перевода системы в режим с большим T_k необходимо понизить значения смаза или ОСШ ниже порогов γ_{sh} и $\gamma_{\psi h}$, соответственно. Для перевода системы в режим с меньшим T_k необходимо повысить значения смаза и ОСШ выше γ_{sb} и $\gamma_{\psi b}$, соответственно.

Проведённые исследования, благодаря учёту информационной оценки шума и фона, позволили определить:

- оптимальный порог обнаружения смаза;
- оптимальное значение времени накопления кадра, которое определяет значение скоростного смаза (рисунок 3);
- минимальную необходимую сложность свёрточной нейронной сети.

При этом установлено:

- информационный риск стремится к нулю при достижении порогового значения смаза КА;
- для задачи селекции КА на звёздном фоне следует применять нейросетевые методы.

Для решения задачи адаптивного кодирования изображений применительно к задаче обнаружения ИКО предложено применять принцип доминантной информации, задачей которого является повышение качества передаваемой информации.

Блок схема разработанного алгоритма показана на рисунке 5. Данный алгоритм применяется на ближней и средней дистанции во время наблюдения за КА для оценки его координат, выделения доминантной информации и сжатия фоновой. В рамках предложенного алгоритма синтезирована телевизионная система адаптивного кодирования видеопоследовательностей, которая позволяет осуществлять сжатие с разными коэффициентами качества доминантной и фоновой информации.

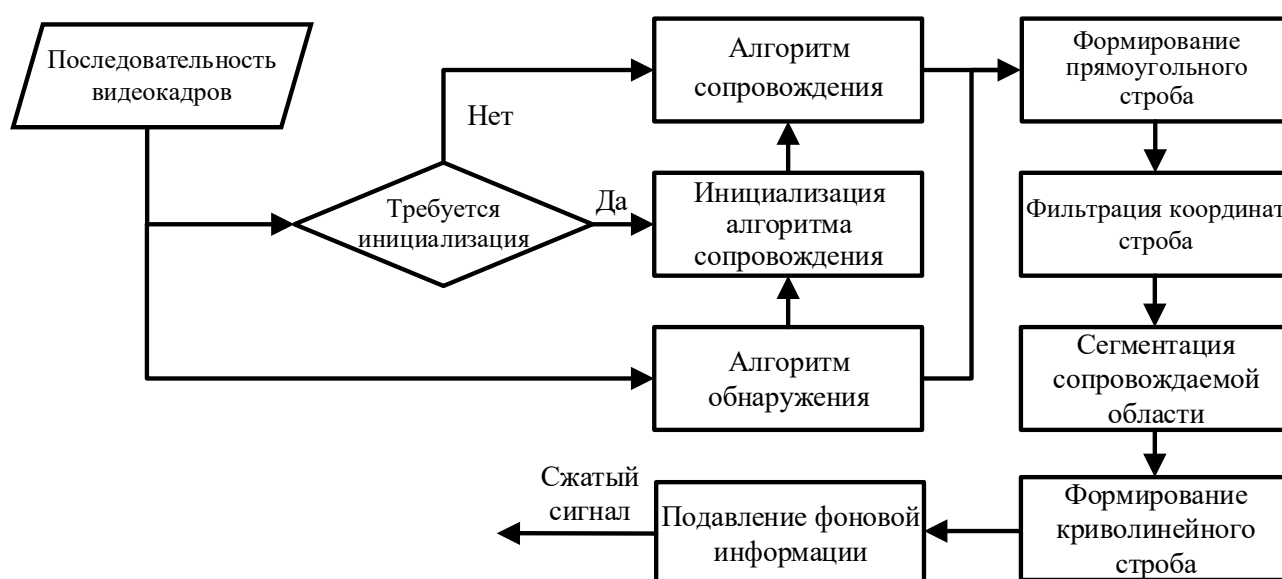


Рисунок 5 – Блок-схема алгоритма адаптивного выделения доминантной информации и сжатия фоновой.

Алгоритм состоит из следующих этапов:

- Обнаружение изображения КА с помощью алгоритма сопровождения и нейросетевого детектора *YOLOv2* (обучен на выборке из 1000 изображений КА разных ракурсов) и формирование прямоугольного строба.

- Фильтрация координат ограничивающего прямоугольника.
- Выделения доминантного объекта криволинейной областью с помощью алгоритма сегментации.
- Подавление фоновой информации.

В предложенном алгоритме для сжатия использовался алгоритм JPEG с 100% значением качества для доминантного объекта и 1% качества для фоновой информации. На рисунке 6 приведены примеры работы предложенного алгоритма.

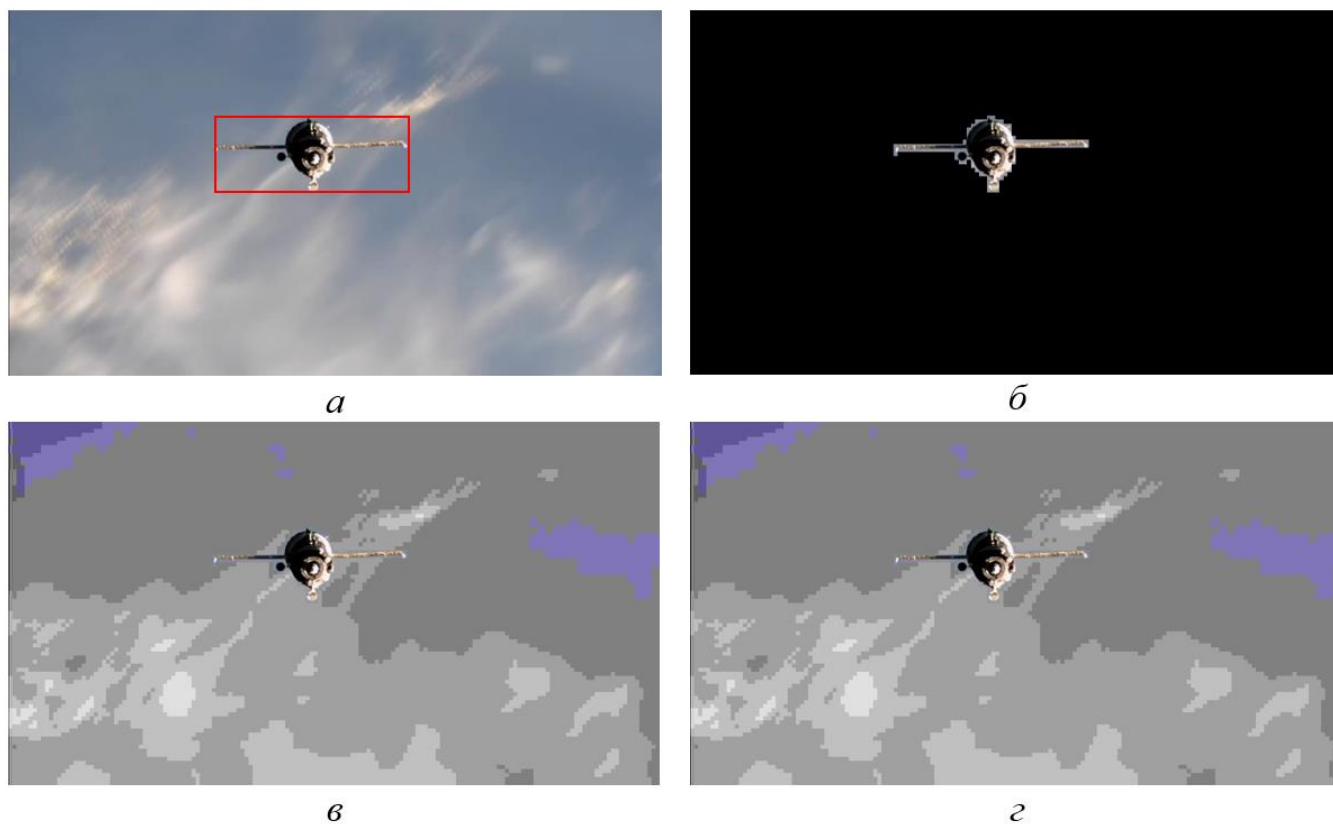


Рисунок 6 – Пример работы алгоритма: *а* – исходное изображение; *б* – сегментированное изображение; *в* – результат; *г* – сжатое изображение.

В таблице 3 представлены результаты работы предложенного алгоритма.

Таблица 3 – результаты работы алгоритма.

A_{Π}/A_{κ}	$A_{\kappa\sigma}/A_{\kappa}$	$S_a, \text{Кбит}$	$S_{\sigma}, \text{Кбит}$	$S_{\sigma}, \text{Кбит}$	$S_{\varepsilon}, \text{Кбит}$
0,06	0,02	22118,832	178,368	200,44	124,496

Обозначения в таблице: A_n/A_k – отношение площади ограничивающего прямоугольника A_n (рис. 6. а) к площади кадра A_k ; A_n/A_k – отношение площади криволинейного строга A_n (рис. 6. б) к A_k ; S_a, S_b, S_b, S_z – размеры файлов, содержащих рисунки ба, бб, бв и бг, соответственно.

В работе показано, что применение принципа доминантной информации позволяет на порядок увеличить коэффициент сжатия передаваемых изображений без потери информации о сигнале доминантного объекта КА. В результате исследований разработана и синтезирована телевизионная система адаптивного кодирования видеопоследовательностей.

Для решения задачи обнаружения, распознавания и сопровождения группы наземных объектов на последовательности аэрофотоснимков в условиях изменяющегося фона предложено применять сочетание нейросетевого обнаружителя, алгоритма Куна–Манкреса и сравнение цветowych гистограмм.

Для исследования выбраны следующие классы объектов: легковые автомобили, крупногабаритный автотранспорт, здания, самолёты, вертолёты, корабли. Набор обучающих данных содержит 8000 изображений (86000 объектов), набор проверочных данных содержит 2000 изображений (21000 объектов), а набор тестовых данных содержит 1400 изображений (14000 объектов). Обучение проводилось на двух видеокартах NVIDIA Tesla K40X. Процесс обучения длился 200000 итераций. Примеры объектов представлены на рисунке 7. Для исследования в качестве нейросетевых обнаружителей выбраны подходы на основе свёрточных нейронных сетей *YOLOv2/v3*.



Рисунок 7 – Пример аэрофотоснимков наземных объектов.

Результаты работы трёх сетей семейства *YOLO* представлены в таблице 4. Все архитектуры показывают высокую скорость работы на графическом процессоре Nvidia GTX 1050TI (768 ядер CUDA) – от 11 до 50 кадров в секунду для разрешения 608×608.

В таблице 5 приведены значения индекса Жаккара и среднее значение времени обработки одного кадра для *TinyYOLOv2*, *YOLOv2* и *YOLOv3*.

Таблица 4. Результаты работы *TinyYOLOv2*, *YOLOv2* и *YOLOv3*.

Класс	<i>TinyYOLOv2</i>	<i>YOLOv2</i>	<i>YOLOv3</i>
	Вероятность правильного обнаружения		
Автомобиль	0,39	0,63	0,86
Крупногабаритный автотранспорт	0,42	0,64	0,76
Здание	0,63	0,80	0,77
Самолёт	0,67	0,88	0,87
Вертолёт	0,78	0,88	0,91
Корабль	0,60	0,82	0,84
Среднее значение	0,58	0,77	0,84

Таблица 5. Значение индекса Жаккара и среднее значение времени обработки одного кадра.

Параметр	<i>TinyYOLOv2</i>	<i>YOLOv2</i>	<i>YOLOv3</i>
Индекс Жаккара	0,35	0,50	0,60
Среднее время обработки на GPU, с	0,02	0,05	0,10

Примеры работы *TinyYOLOv2*, *YOLOv2* и *YOLOv3* приведены на рисунках 8, 9, 10 соответственно.

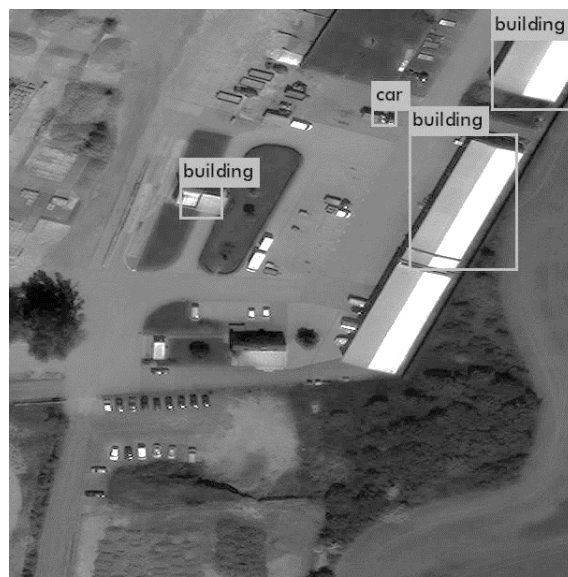


Рисунок 9 – Результаты *TinyYOLOv2*.

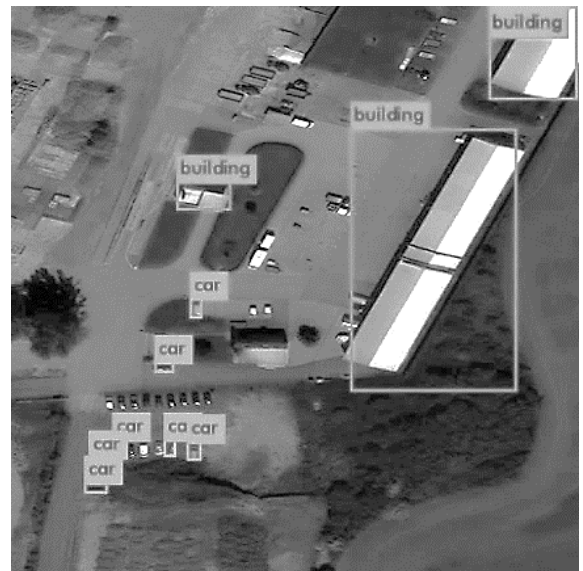


Рисунок 10 – Результаты *YOLOv2*.

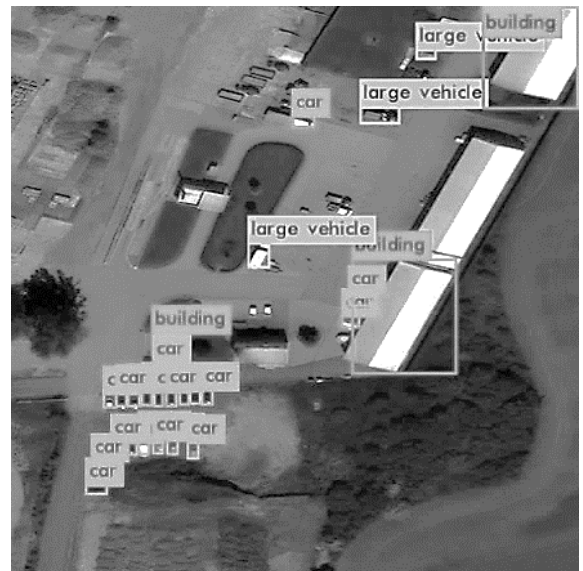


Рисунок 11 – Результаты *YOLOv3*.

В таблице 6 представлены некоторые особенности модели сцены и эффективность нейросетевых детекторов.

Таблица 6. Особенности модели сцены и эффективность нейросетевых детекторов.

Особенность	<i>TinyYOLOv2</i>	<i>YOLOv2</i>	<i>YOLOv3</i>
Малые размеры объекта (пример: 20 × 30)	+	+	+
Большие размеры объекта (пример: 200×500)	–	+	+
Близкое расположение малоразмерных объектов	–	–	+

На рисунке 8 представлена зависимость количества слоёв от сложности модели сцены. По результатам исследований можно сделать вывод, что с ростом сложности модели сцены растёт сложность архитектуры нейронной сети (число слоёв).

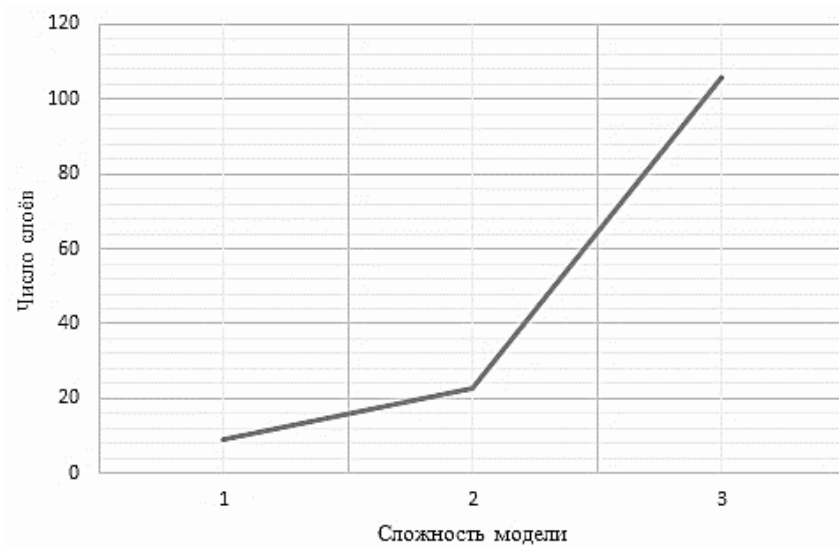


Рисунок 8 – Зависимость количества слоёв от сложности модели сцены.

В результате исследований было получено, что по критериям точности обнаружения, распознавания и скорости работы перспективным является нейросетевой детектор *YOLOv2*.

На основе обученной сети *YOLOv2* получен алгоритм сопровождения нескольких объектов в условиях подвижной камеры и изменяющегося фона, блок-схема которого представлена на рисунке 12.



Рисунок 12 – Схема алгоритма сопровождения нескольких объектов.

Принцип работы алгоритма сопровождения выглядит следующим образом:

- На первом кадре каждому обнаруженному объекту присваивается уникальный идентификатор. Текущий кадр пропускается в том случае, если объекты не были обнаружены.

- Если кадр не является первым и инициализация выполнена успешно, то выполняется поиск соответствий между обнаруженными и сопровождаемыми объектами.

Поиск соответствий выполняется с помощью алгоритма Куна–Манкреса, который работает с матрицей значений расстояния Евклида между ограничивающими прямоугольниками на текущем и предыдущем кадрах.

Поскольку в алгоритме Куна–Манкреса отсутствует механизм отслеживания момента исчезновения и появления того же самого объекта из поля зрения камеры, возможны два случая: количество объектов в текущем кадре уменьшилось по сравнению с предыдущим; количество объектов в текущем кадре увеличилось по сравнению с предыдущим. Для обработки обоих случаев предложено применять сравнение гистограмм изображений объектов. В результате экспериментов в рассматриваемом алгоритме для сравнения гистограмм была выбрана метрика – расстояния Бхаттачария. Значение этой метрики, близкое к 0, соответствует полному совпадению сравниваемых гистограмм, в случае полного несоответствия значение расстояния стремится к 1.

Для коррекции траектории группы объектов применяется рекурсивный фильтр, который является сочетанием линейного и нелинейного фильтров.

Для оценки производительности предложенного алгоритма выбраны следующие метрики: точность, полнота, индекс Жаккара и абсолютная погрешность локализации. Для экспериментов были выбраны 20 видеопоследовательностей, полученных с камеры летательного аппарата. Для сравнения производительности предложенного алгоритма были выбраны два алгоритма: «*IOU Tracker*» и «*TLD*».

Экспериментальные результаты предложенного алгоритма сопровождения группы объектов (АСГО) на 20 тестовых видеопоследовательностях приведены в таблицах 7 и 8.

Таблица 7. Среднее значение точности и полноты алгоритмов на 20 тестовых видеопоследовательностях.

Видео	Точность			Полнота		
	<i>АСГО</i>	<i>IOU Tr.</i>	<i>TLD</i>	<i>АСГО</i>	<i>IOU Tr.</i>	<i>TLD</i>
Среднее значение	0,95	0,94	0,67	0,94	0,46	0,74

Таблица 8. Среднее значение Индекса Жаккара и погрешности локализации (Δ), алгоритмов на 20 тестовых видеопоследовательностях.

Видео	Индекс Жаккара			Δ , пиксель		
	<i>АСГО</i>	<i>IOU Tr.</i>	<i>TLD</i>	<i>АСГО</i>	<i>IOU Tr.</i>	<i>TLD</i>
Среднее значение	0,7	0,7	0,56	4,9	5,6	25

На рисунке 13 представлены результаты работы предложенного алгоритма сопровождения.

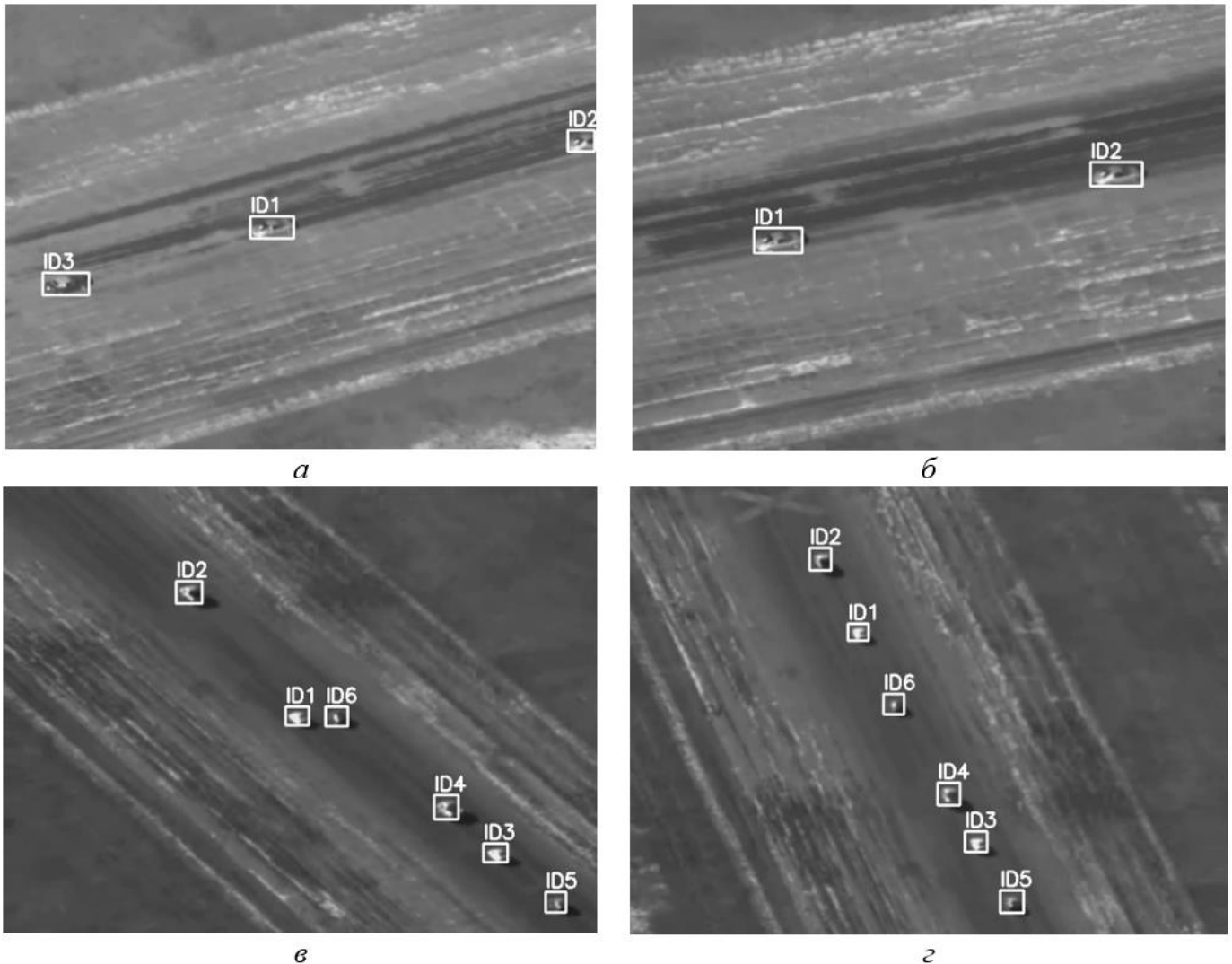


Рисунок 13 – Результат работы АСГО на тестовых последовательностях: *а, б* – видео 1, кадры 131 (*а*) и 326 (*б*); *в, г* – видео 6, кадры 23 (*в*) и 216 (*г*).

В работе предложен алгоритм обнаружения и сопровождения группы наземных неточечных объектов на последовательности аэрофотоснимков в условиях изменяющегося фона и изменения положения объектов и камеры на основе свёрточной нейронной сети *YOLOv2*. *YOLOv2* обучена на разработанном наборе обучающих данных. В результате работы выявлена зависимость сложности структур свёрточных нейронных сетей семейства *YOLO* от сложности модели сцены.

Для обработки ситуаций выхода одного из объектов из кадра применяется сравнение цветовых гистограмм.

По результатам проведённых исследований показано, что предложенный алгоритм превосходит алгоритм «*TLD*» и «*IOU Tracker*», поскольку позволяет сопровождать более одного объекта и корректно решить проблему с исчезновением объекта из поля зрения камеры.

Заключение

Основные результаты работы. В ходе исследований получены следующие результаты:

1. Разработан нейросетевой метод обнаружения ИКО на звёздном фоне и оценены его помехоустойчивость и сложность реализации для применения в бортовом вычислителе.
2. Разработан метод адаптации кадровой частоты фотоприёмника.
3. Разработан нейросетевой метод адаптивного кодирования изображений в задачах автоматического обнаружения ИКО.
4. Разработан нейросетевой метод обнаружения, распознавания и сопровождения группы наземных объектов с борта летательного аппарата в условиях сложного фона.

Общий вывод. Результаты работы внедрены в ряде НИОКР АО «НИИ телевидения» и АО «ЛОМО», а также в учебный процесс СПбПУ Петра Великого.

Таким образом, совокупность выводов и научных положений показывает, что поставленная задача разработки и исследования методов автоматической селекции, классификации и сопровождения искусственных космических и наземных объектов, адаптивное сжатие видеоинформации на основе принципа доминантной информации решена.

**Список работ, опубликованных по теме научно-квалификационной работы
(диссертации)**

Публикации в изданиях, рецензируемых ВАК

1. Павлов, В. А. Применение свёрточных нейронных сетей для обнаружения и распознавания искусственных космических и наземных объектов [Текст] / Павлов В. А. // Радиотехника. – 2019. – Т. 83. – № 12(20). – С. 58–67.

2. Павлов, В. А. Классификация малоразмерных изображений космических объектов по признакам движения с помощью обучаемого алгоритма [Текст] / Цыцулин А. К., Морозов А. В., Бобровский А. И., Баскакова Ю. В., Павлов В. А. // Вопросы радиоэлектроники. Серия Техника телевидения. – 2018. – № 3. – С. 72–80.

3. Павлов, В. А. Адаптация кадровой частоты к этапам наблюдения в системе контроля сближения космических аппаратов [Текст] / Цыцулин А. К., Девяткин А. В., Бобровский А. И., Морозов А. В., Горшанов Д. Л., Павлов В. А. // Вопросы радиоэлектроники. Серия Техника телевидения. – 2019. – № 1. – С. 31–38.

4. Павлов, В. А. Адаптивное кодирование изображений, делимых на доминантный объект и фон [Текст] / Цыцулин А. К., Павлов В. А., Бобровский А. И., Морозов А. В., Зубакин И. А. // Вопросы радиоэлектроники. Серия Техника телевидения. – 2019. – № 3. – С. 75–85.

5. Павлов, В. А. Информационные оценки в задачах обнаружения – оценивания – передачи сигнала в космическом телевидении [Текст] / Цыцулин А. К., Павлов В. А., Бобровский А. И., Морозов А. В. // Вопросы радиоэлектроники. Серия Техника телевидения. – 2019. – № 3. – С. 61–74.

6. Павлов, В. А. Применение свёрточных нейронных сетей для автоматической селекции малоразмерных искусственных космических объектов на оптических изображениях звёздного неба [Текст] / Цыцулин А. К., Бобровский А. И., Морозов А. В., Павлов В. А., Галеева М. А. // Оптический журнал. – 2019. – Т. 86. № 10. – С. 30–38.

7. Павлов, В. А. Выбор архитектуры нейронной сети для обнаружения и классификации изображений космических объектов [Текст] / Бобровский А. И., Морозов А. В., Чепелев А. Г., Павлов В. А., Галанов В. В. // Вопросы радиоэлектроники. Серия Техника телевидения. – 2020. – № 1. – С. 64–73

8. Павлов, В. А. Применение свёрточной нейронной сети для сопровождения объектов на последовательности аэрофотоснимков [Текст] / Макаров С. Б., Бобровский А. И., Павлов В. А., Безбородов А. К. // Вопросы радиоэлектроники. Серия Техника телевидения. – 2020. – № 2. – С. 3–18.

9. Павлов, В. А. Помехоустойчивость классификации объектов по признаку скоростного смаза [Текст] / Цыцулин А. К., Павлов В. А., Бобровский А. И., Морозов А. В., Рогачёв В. А. // Вопросы радиоэлектроники. Серия Техника телевидения. – 2020. – № 3. – С.3–17.

Публикации в изданиях, индексируемых в Scopus

1. Pavlov, V. Automatic detection of objects on star sky images by using the convolutional neural network [Текст] / Bobrovsky A., Galeeva M., Morozov A., Pavlov V., Tsytsulin A. // Journal of Physics: Conference Series. – 2019. – № 1236. – С. 1–6.

2. Pavlov, V. Adaptation of frame frequency to observation stages at control of spacecraft convergence [Текст] / Devyatkin A., Tsytsulin A., Bobrovsky A., Morozov A., Gorshanov D., Pavlov V. // Journal of Physics: Conference Series. – 2019. – № 1236. – С. 1–4.

3. Pavlov, V. A. Detection and recognition of objects on aerial photographs using convolutional neural networks [Текст] / Pavlov V. A., Galeeva M. A. // Journal of Physics: Conference Series. – 2019. – № 1326. – С. 1–6.

4. Pavlov, V. A. Taking into Account Noise and Background Information During Transmitting of Decisions on Signal Detection-Estimation [Текст] / Tsytsulin A. K., Morozov A. V., Bobrovsky A. I., Pavlov V. A., Galeeva M. A. // IEEE International Conference on Electrical Engineering and Photonics (EExPolytech), St. Petersburg, Russia. – 2019. – С 124–127.

Свидетельства о государственной регистрации программы для ЭВМ

1. Свид. 2019616555 Российская Федерация. Свидетельство об официальной регистрации программы для ЭВМ. Программа для сжатия фрагментов изображения с разным коэффициентом сжатия для доминантной и фоновой информации с автоматической селекцией с сегментацией доминантной информации / Цыцулин А. К.,

Бобровский А. И., Павлов В. А., Зубакин И. А.; заявитель и правообладатель Акционерное общество «Научно-исследовательский институт телевидения». – № 2019616555; заявл. 13.05.2019; опубл. 24.05.2019, Реестр программ для ЭВМ. – 8 с.

Другие статьи и материалы конференций

1. Павлов, В. А. Реализация модуля сопровождения объектов в видеопотоке, получаемом с летательного аппарата [Текст] / Павлов В. А., Завьялов С. В., Волвенко С. В. // Неделя науки СПбПУ 2015. Материалы научного форума с международным участием. Институт физики, нанотехнологий и телекоммуникаций. – 2015. – С. 66–69

2. Павлов, В. А. Реализация программного модуля сопровождения объектов в видеопотоке с беспилотного летательного аппарата [Текст] / Павлов В. А., Завьялов С. В., Нужный С. П., Волвенко С. В. // Доклады 18-й международной конференции Цифровая обработка сигналов и ее применение – DSPA–2016, Москва. – 2016 – № 18, – Т. 2. – С. 629–633.

3. Павлов, В. А. Применение свёрточной нейронной сети для обнаружения и сопровождения нескольких объектов в видеопотоке с борта летательного аппарата [Текст] / Павлов В. А., Крылов С. С., Бобровский А. И. // Материалы 15-й международной конференции «Телевидение: передача и обработка изображений» 2018. – С. 138–140.

4. Павлов, В. А. Алгоритм классификации малоразмерных изображений космических объектов по признакам движения [Текст] / Морозов А. В., Бобровский А. И., Баскакова Ю. В., Павлов В. А. // Пулково-2018. Тезисы докладов. – 2018. – С. 30.

Аспирант _____ Павлов Виталий Александрович