

## ПРИМЕНЕНИЕ ТЕХНОЛОГИЙ COMPUTER VISION ДЛЯ ОТСЛЕЖИВАНИЯ И МОДЕЛИРОВАНИЯ ДИНАМИЧЕСКИХ ОБЪЕКТОВ В РЕАЛЬНОМ РЕЖИМЕ ВРЕМЕНИ

Р.Р. Аббасов<sup>1</sup>, А.О. Юсифов<sup>1</sup>, О.Ф. Баширзаде<sup>1</sup>, П.Ф. Джаббарлы<sup>1</sup>, Л.И.  
Амирова<sup>2</sup>, Н.Ш. Гусейнова<sup>1,2</sup>

<sup>1</sup>Азербайджанский Технический Университет, Баку, Азербайджан

<sup>2</sup>Институт Прикладной Математики при Бакинском Государственном Университете,  
Баку, Азербайджан

e-mail: rft.abbas02@gmail.com, aliyusifov220@gmail.com, omarbashirzada@gmail.com,  
parvizjabbar@gmail.com, a\_leyla06@rambler.ru, nargiz\_huseynova@yahoo.com

**Резюме.** В данной статье рассматриваются современные алгоритмы компьютерного зрения, используемые для обнаружения и отслеживания движущихся объектов в режиме реального времени. Анализируются и сравниваются методы YOLO, фильтр Калмана и оптический поток. Приводятся эксперименты по точности модели и её производительности.

**Ключевые слова:** Компьютерное зрение, динамические объекты, отслеживание объектов, реальное время, YOLO, фильтр Калмана, Optical Flow.

**AMS Subject Classification:** 60G35, 93E11.

### 1. Введение

В наше время автоматическое обнаружение, отслеживание и моделирование динамических объектов стало одним из основных направлений исследований искусственного интеллекта и технологий обработки изображений. Технологии Computer Vision, в частности, открывают широкие возможности в этой области.

В данном исследовании были изучены современные алгоритмы, которые можно использовать для идентификации и отслеживания движущихся объектов (автомобилей, людей, животных и т. д.) в режиме реального времени. В качестве примеров были проанализированы такие алгоритмы, как YOLO (You Only Look Once), фильтр Калмана, Optical Flow, а также были сравнены результаты их применения.

Эксперименты показали, что подходы на основе глубокого обучения позволяют точно и быстро определять данные пространственных координат объектов. Однако интеграция таких методов, как временной анализ (temporal analysis) и сети LSTM для построения поведенческих моделей динамических объектов, обеспечивает более стабильные результаты.

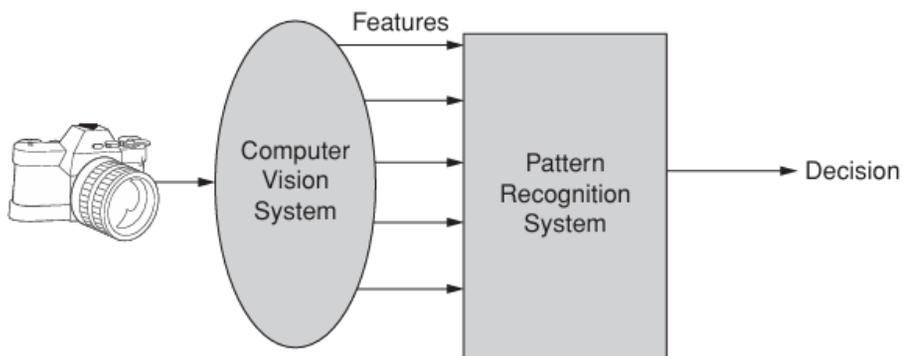
На основе экспериментов предложенная модель обеспечивает эффективное обнаружение объектов в реальном времени и анализ траектории движения. Результаты показывают, что алгоритм YOLO хорошо работает при быстром обнаружении объектов, но фильтр Калмана более эффективен при отслеживании объектов.

В качестве практических приложений эти технологии могут широко применяться в системах автоматического управления, проектах умных городов, робототехнике и автономных транспортных системах. Будущие направления исследований будут сосредоточены на применении этих технологий с более высокой точностью и оптимизации моделей, связанных с различными типами объектов.

Фильтр Калмана - это математическая модель, используемая для прогнозирования траектории динамических объектов. Он оценивает состояние объекта (например, координаты и скорость), минимизируя ошибки измерения. Преимущество фильтра Калмана - его низкие вычислительные требования и простая интеграция. Однако точность фильтра снижается при работе со сложными моделями движения (например, нелинейными траекториями). Чтобы преодолеть это ограничение, можно рассмотреть альтернативы, такие как расширенный фильтр Калмана (Extended Kalman Filter) или фильтр частиц (Particle Filter).

## **2. Основы и этапы развития технологии компьютерного зрения**

Компьютерное зрение — это автоматический анализ изображения машиной для понимания того, что находится на изображении. Например, она может распознавать деталь машины на изображении и проверять ее качество. В этой области часто используются и иногда путаются такие термины, как «Обработка изображений» и «Машинное зрение». Фактически, каждый из них описывает разные процессы. Обработка изображений делится на две основные части. В первой части в изображения вносятся основные изменения. Например, из изображения удаляется шум, оно становится более четким или его размер уменьшается. Этот тип обработки не распознает объекты на изображении, но подготавливает изображение для дальнейшего анализа. Во второй части объекты на изображении распознаются и интерпретируются. Это более сложный процесс, и обычно он выполняется на изображениях, подготовленных в первой части. Таким образом, системы компьютерного зрения могут анализировать изображения и понимать их содержимое [13].



**Рис. 1. Система компьютерного зрения**

Как показано на рисунке, процесс обработки обычно делится на две основные категории: низкоуровневая и высокоуровневая обработка изображений. Когда эти процессы рассматриваются с точки зрения теории сигналов и систем, их различия и сходства могут быть более четко определены на основе входных и выходных форматов. При низкоуровневой обработке изображений входное изображение в систему проходит различные этапы обработки, но выход по-прежнему имеет форму изображения — только с некоторыми изменениями. Примерами таких изменений являются очищенное от шума изображение, сжатая версия изображения, которая занимает меньше памяти, или более четкое изображение. Хотя низкоуровневая обработка изображений не дает прямой информации о содержании объектов на изображении, она является предварительным подготовительным этапом, необходимым для правильной работы высокоуровневых алгоритмов. С другой стороны, такие термины, как «Машинное зрение», «Понимание изображений» и «Высокоуровневая обработка изображений» часто используются для обозначения концепции компьютерного зрения. Хотя в этих процессах в качестве входных данных принимается изображение, на выходе получается уже не изображение, а информация, полученная в результате анализа и интерпретации изображения. Распознавание образов является одним из основных компонентов системы компьютерного зрения, и здесь на входе уже не изображение, а структурированная информация, описывающая объекты на изображении. Эта информация обычно представляет собой набор признаков, представленных в векторной форме. На выходе получается результат распознавания, полученный на основе этой информации. Таким образом, в процессе распознавания образов ни входные, ни выходные данные не имеют формы изображения, а скорее аналитического описания изображения и решения, принятого на его основе [13].

Рисунок 1 иллюстрирует принцип работы системы компьютерного зрения. В этой системе решение о том, что представляет собой объект и к какой категории он принадлежит, принимает система распознавания образов (PR). Решение системы распознавания образов основывается на измерениях, выполненных на входных данных. Эти измерения называются «признаками». Например, система распознавания образов определяет, является ли встречающийся ей объект топором или молотком, на основе таких признаков, как длина его ручки и масса его головки. Если длину ручки необходимо определить по изображению, то система, которая выполняет это измерение, уже является системой компьютерного зрения. Таким образом, система компьютерного зрения производит различные измерения на изображении и передает результаты этих измерений в систему распознавания образов. Если система компьютерного зрения достаточно развита и может напрямую выдавать результат, например, «количество слонов в комнате = 1», то, возможно, нет необходимости в работе системы распознавания образов. Из этих определений, которые мы представили, есть некоторые исключения. Лучшим примером в настоящее время являются исследования в области глубокого обучения. Эти системы используют вычислительные структуры, вдохновленные нейронными сетями, для идентификации конкретных объектов на изображениях. Хотя некоторые исследователи считают эти системы классификаторами образов, проблемы, которые они решают, явно являются проблемами компьютерного зрения. Поэтому было бы правильнее классифицировать эти системы как системы компьютерного зрения. Такой подход показывает, что с развитием технологий классические определения также меняются и расширяются. Методы глубокого обучения не только еще больше проясняют границы между компьютерным зрением и распознаванием образов, но и демонстрируют интеграцию обеих областей [11].

### **3. Роль фильтра Калмана в отслеживании динамических объектов в реальном времени**

Фильтр Калмана был предложен Рудольфом Э. Калманом в 1960 году [9]. Этот алгоритм используется для получения оптимальных оценок состояния из зашумленных наблюдений в линейных динамических системах. Фильтр Калмана основан на двух основных предположениях:

1. Система имеет линейную динамическую модель.
2. Шум (шум процесса и наблюдения) имеет гауссово распределение.

В компьютерном видении фильтр Калмана применяется для отслеживания траекторий движения динамических объектов (например, транспортных средств, людей, дронов) и прогнозирования их будущих состояний. Например, положение и скорость автомобиля можно оценить в реальном времени, даже если наблюдения (данные камеры или датчика) зашумлены.

Преимущества фильтра Калмана:

- Эффективность вычислений, подходит для приложений в реальном времени.
- Может генерировать точные оценки из зашумленных данных.
- Легко интегрируется в различные системы благодаря своей модульной структуре.

Однако фильтр Калмана ограничен линейными системами и гауссовым шумом. Для нелинейных систем используются альтернативы, такие как расширенный фильтр Калмана (ЕКФ) или Unscented фильтр Калмана (UKF).

### Математические основы фильтра Калмана

Фильтр Калмана [1-5] состоит из двух основных этапов: Прогнозирование (**Prediction**) и Обновление (**Update**). Эти этапы итеративно улучшают состояние системы и ее неопределенность (ковариацию).

Фильтр Калмана основан на следующей линейной динамической системе:

#### Уравнение состояния:

$$x_k = F_k x_{k-1} + B_k u_k + w_k,$$

здесь  $x_k$  вектор состояния системы в момент времени  $k$  (например, положение, скорость).  $F_k$ - матрица переходов состояний описывает динамику системы,  $B_k$ - Матрица управления представляет собой внешние сигналы управления ( $u_k$ ),  $w_k$ - процесс шума,  $w_k \sim N(0, Q_k)$ , с гауссовым распределением.

#### Уравнение наблюдения:

$$z_k = H_k x_k + v_k,$$

Здесь  $z_k$  - вектор наблюдения в момент времени  $k$  (например, координаты, полученные с камеры),  $H_k$ - Матрица наблюдений переводит ситуацию в наблюдения,  $v_k$ - Шум наблюдения,  $v_k \sim N(0, R_k)$ , с гауссовым распределением.  $Q_k$  и  $R_k$  - ковариационные матрицы шумов процесса и наблюдения соответственно.

### 4. Этап прогнозирования

На этапе прогнозирования состояние системы и ее неопределенность оцениваются с использованием предыдущего состояния:

1. Прогнозирование состояния:

$$\hat{x}_{k/k-1} = F_k \hat{x}_{k-1/k-1} + B_k u_k.$$

2. Ковариационный прогноз:

$$P_{k/k-1} = F_k P_{k-1/k-1} F_k^T + Q_k.$$

Фильтр Калмана итеративно оценивает состояние системы, используя эти уравнения. На этапе прогнозирования вычисляются состояние и его неопределенность (ковариационная матрица  $P_k$ ).

$$\hat{x}_{k/k-1} = A \hat{x}_{k-1/k-1} + B u_k$$

$$P_{k/k-1} = A P_{k-1/k-1} A^T + Q.$$

здесь  $Q$ - ковариационная матрица шума процесса.

На этапе обновления прогноз корректируется на основе измерений:

$$K_k = P_{k/k-1} H^T (H P_{k/k-1} H^T + R)^{-1},$$

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + K_k (z_k - H \hat{x}_{k/k-1}),$$

$$P_{k/k} = (I - K_k H) P_{k/k-1}.$$

Здесь  $K_k$ - выигрыш Калмана,  $R$ - это ковариационная матрица [6] измерения шума.

### 5.1 Отслеживание объекта

В системах компьютерного зрения отслеживание объектов обычно требует обработки последовательности видеок кадров. Фильтр Калмана используется в этом процессе для прогнозирования положения и скорости объекта. Например:

- **простое 2D отслеживание**

Для отслеживания движения транспортного средства на видео фильтр Калмана оценивает координаты  $(x, y)$  объекта и скорость их изменения  $(v_x, v_y)$ . Измерения (например, центр обнаруженного в кадре объекта) могут быть зашумлены, но фильтр уменьшает этот шум и вычисляет более точную траекторию.

- **Отслеживание нескольких объектов:**

Для отслеживания нескольких объектов одновременно фильтр Калмана применяет отдельную модель состояния к каждому объекту. Это особенно эффективно в сочетании с такими алгоритмами, как Deep SORT [14].

### 5.2 Фильтр шума

В системах компьютерного зрения изображения часто сопровождаются шумом (например, изменениями освещения, тенями или плохим качеством

камеры). Фильтр Калмана статистически моделирует этот шум и позволяет оценить истинное положение объекта. Например, в системе, отслеживающей движение дрона, измерения, сделанные камерой, могут быть неточными, но фильтр исправляет эти ошибки с помощью этапов прогнозирования и обновления.

### **5.3 Применение в моделировании**

Фильтр Калмана не ограничивается только отслеживанием, но также используется для моделирования динамических объектов. Например, в алгоритмах SLAM (Simultaneous Localization and Mapping) фильтр Кальмана применяется для создания карты собственного положения робота и окружающих его объектов. Расширенный фильтр Калмана (EKF) является адаптированной версией для нелинейных систем и широко применяется в робототехнике [12].

## **6. Преимущества и ограничения фильтра Калмана**

### **6.1. Преимущества**

- **Скорость:** фильтр Калмана работает в режиме реального времени с низкими вычислительными требованиями, что делает его идеальным для устройств с ограниченными ресурсами (например, устройств IoT).
- **Точность:** эффективно компенсирует шумные измерения.
- **Простота:** обеспечивает простую и интуитивно понятную модель для линейных систем.
- **Гибкость:** может быть адаптирован к различным приложениям (транспорт, робототехника, видеонаблюдение).

### **6.2. Ограничения**

- **Ограничение линейными системами:** фильтр Калмана оптимизирован для линейных динамических систем. В нелинейных системах используются такие вариации, как EKF или Unscented Kalman Filter (UKF), но это увеличивает вычислительную сложность [8].
- **Предположение о распределении шума:** фильтр основан на предположении о нормально распределенном шуме. Если шум не соответствует этой модели, результаты могут быть неточными.
- **Сложные среды:** в средах со множеством объектов или там, где объекты могут перекрывать друг друга (окклюзия), одного фильтра Калмана может быть недостаточно, и его необходимо сочетать с алгоритмами глубокого обучения.

## **7. Примеры практического применения**

- **Автономный транспорт:** фильтр Калмана используется в автономных транспортных средствах для прогнозирования траектории движения других

транспортных средств и пешеходов. Например, измерения LiDAR и радаров могут быть зашумленными, но фильтр объединяет эту информацию для обеспечения точного отслеживания.

- Дроны: в системах навигации дронов фильтр Калмана оценивает положение, объединяя данные GPS и инерционных датчиков.
- Видеонаблюдение: фильтр Калмана вычисляет траектории объектов в реальном времени для отслеживания подозрительных движений в камерах безопасности.
- Робототехника: фильтр Калмана играет ключевую роль в алгоритмах SLAM, позволяющих роботам распознавать и отображать объекты в своем окружении.

### **8. Результат и перспективы на будущее**

Применение фильтра Калмана в компьютерном зрении расширяется с появлением новых технологий. Интеграция с глубоким обучением (например, объединение фильтра Калмана с нейронными сетями) позволяет более точно отслеживать и моделировать в сложных средах.

### **Литература**

1. Aliev F.A., Larin V.B. Optimization of Linear Control Systems: Analytical Methods and Computational Algorithms, Amsterdam: Gordon and Breach Sci , (1998).
2. Aliev F.A., Larin V.B., Velieva N.I. Algorithms of the Synthesis of Optimal Regulators, USA, Outskirts Press , ISBN 197724985X, 9781977249852, (2022).
3. Aliev F.A., Mutallimov M.M., Velieva N.I., Safarova N.A., Jafarov A.G. Algorithms for solving the stabilization problems of the quadcopter motion by GPS data, Applied and Computational Mathematics, V.22, N.4, (2023), pp.520-527.
4. Aliev F.A., Ozbek L. Comments on "Adaptive fading Kalman filter with an application" , AUTOMATICA, V.34, N.12, (1998), pp.1663-1664.
5. Aliev F.A., Ozbek L. Evaluation of convergence rate in the central limit theorem for the Kalman filter, IEEE Transactions on Automatic Control, V. 44, N.10, (1999), pp.1905-1909.
6. Grewal M.S. and Andrews A.P. Kalman Filtering Theory and Practice Using MATLAB. 4th Edition, John Wiley & Sons Inc., New York, (2014).
7. Intille S.S., & Bobick A.F. Closed-World Tracking, Proceedings of the IEEE International Conference on Computer Vision (ICCV), (1995), pp.672-678.
8. Julier S.J., & Uhlmann J.K. A New Extension of the Kalman Filter to Nonlinear Systems, Proceedings of SPIE, 3068, (1997), pp.182-193.
9. Kálmán R.E. A New Approach to Linear Filtering and Prediction Problems, Journal of Basic Engineering, V.82, N.1, (1960), pp.35-45.
10. Krishnan R.G., Shalit U., & Sontag D. Structured Inference Networks for Nonlinear State Space Models, arXiv preprint arXiv:1510.05787, (2015).
11. McCulloch W. & Pitts W. A logical calculus of the ideas immanent in nervous activity, Bulletin of Mathematical Biophysics, (1943).

12. Thrun S., Burgard W., & Fox D. Probabilistic Robotics, MIT Press, (2006).
13. Wesley E. Snyder & Hairong Qi. Fundamentals of Computer Vision, Cambridge University Press, (2017).
14. Wojke N., Bewley A., & Paulus D. Simple Online and Realtime Tracking with a Deep Association Metric, IEEE International Conference on Image Processing (ICIP), (2017), pp.3645-3649.

## APPLICATION OF COMPUTER VISION TECHNOLOGIES FOR TRACKING AND MODELING DYNAMIC OBJECTS IN REAL TIME

**R.R. Abbasov<sup>1</sup>, A.O. Yusifov<sup>1</sup>, O.A. Bəşirzadə<sup>1</sup>, P.A. Cabbarlı<sup>1</sup>, L.İ. Amirova<sup>2</sup>, N.Sh. Huseynova<sup>1,2</sup>**

<sup>1</sup>Azerbaijan Technical University, Baku, Azerbaijan

<sup>2</sup>Institute of Applied Mathematics of Baku State University, Baku, Azerbaijan

e-mail: rft.abbas02@gmail.com, aliyusifov220@gmail.com, omarbashirzada@gmail.com, parvizjabbar@gmail.com, a\_leyla06@rambler.ru, nargiz\_huseynova@yahoo.com

**Abstract.** This paper investigates modern Computer Vision algorithms used for real-time detection and tracking of dynamic objects. Methods such as YOLO, Kalman filter, and Optical Flow are analyzed and compared for performance and efficiency. Experiments on model accuracy and performance are presented.

**Keywords:** Computer Vision, dynamic objects, object tracking, real-time, YOLO, Kalman filter, Optical Flow.

## References

15. Aliev F.A., Larin V.B. Optimization of Linear Control Systems: Analytical Methods and Computational Algorithms, Amsterdam: Gordon and Breach Sci , (1998).
16. Aliev F.A., Larin V.B., Velieva N.I. Algorithms of the Synthesis of Optimal Regulators, USA, Outskirts Press , ISBN 197724985X, 9781977249852, (2022).
17. Aliev F.A., Mutallimov M.M, Velieva N.I., Safarova N.A., Jafarov A.G. Algorithms for solving the stabilization problems of the quadcopter motion by GPS data, Applied and Computational Mathematics, V.22, N.4, (2023), pp.520-527.
18. Aliev F.A., Ozbek L. Comments on "Adaptive fading Kalman filter with an application" , AUTOMATICA, V.34, N.12, (1998), pp.1663-1664.
19. Aliev F.A., Ozbek L. Evaluation of convergence rate in the central limit theorem for the Kalman filter, IEEE Transactions on Automatic Control, V. 44, N.10, (1999), pp.1905-1909.

20. Grewal M.S. and Andrews A.P. Kalman Filtering Theory and Practice Using MATLAB. 4th Edition, John Wiley & Sons Inc., New York, (2014).
21. Intille S.S., & Bobick A.F. Closed-World Tracking, Proceedings of the IEEE International Conference on Computer Vision (ICCV), (1995), pp.672-678.
22. Julier S.J., & Uhlmann J.K. A New Extension of the Kalman Filter to Nonlinear Systems, Proceedings of SPIE, 3068, (1997), pp.182-193.
23. Kálmán R.E. A New Approach to Linear Filtering and Prediction Problems, Journal of Basic Engineering, V.82, N.1, (1960), pp.35-45.
24. Krishnan R.G., Shalit U., & Sontag D. Structured Inference Networks for Nonlinear State Space Models, arXiv preprint arXiv:1510.05787, (2015).
25. McCulloch W. & Pitts W. A logical calculus of the ideas immanent in nervous activity, Bulletin of Mathematical Biophysics, (1943).
26. Thrun S., Burgard W., & Fox D. Probabilistic Robotics, MIT Press, (2006).
27. Wesley E. Snyder & Hairong Qi. Fundamentals of Computer Vision, Cambridge University Press, (2017).
28. Wojke N., Bewley A., & Paulus D. Simple Online and Realtime Tracking with a Deep Association Metric, IEEE International Conference on Image Processing (ICIP), (2017), pp.3645-3649.