

## ИСПОЛЬЗОВАНИЕ КОМПЬЮТЕРНОГО ЗРЕНИЯ ДЛЯ РЕШЕНИЯ ПРОБЛЕМ РАСПОЗНАВАНИЯ ЖЕСТОВ

*Рябиков Мирослав Александрович*

*магистрант,*

*Гомельский государственный технический университет имени П.О. Сухого,  
Республика Беларусь, г. Гомель*

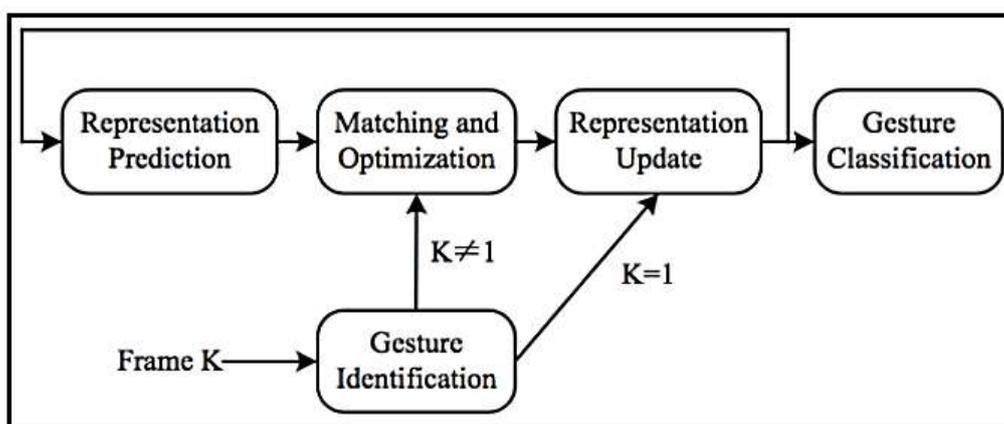
В настоящее время распознавание жестов играет очень важную роль в человеко-машинном взаимодействии благодаря их естественному и дружественному смысловому выражению. Для использования этой технологии машины должны распознавать их быстро и точно, чтобы пользователи чувствовали себя комфортно и были готовы к взаимодействию с машинами. Распознавание жестов остается наиболее сложной задачей из-за их разнообразия, схожести форм и сложности сценариев применения [1].

Проблема распознавания жестов имеет различные решения в научных работах. Существующие подходы обычно можно разделить на несколько категорий:

- распознавание на основе носимых устройств;
- распознавание на основе компьютерного зрения.

Одной из первых технологий, решивших проблему распознавания жестов, стали специальные перчатки. Они используются для сбора данных и последующей их передачи. Эти данные представляют собой движения рук, такие как скорость и угол поворота. Эти данные поступают на компьютер, и тот выполняет распознавание по специальному алгоритму. Например, Такахашаи и Кишино разработали перчатку для сбора данных, способную распознавать 46 типов жестов. Яншэн и др. использовали перчатку Saibo для точного распознавания 14 различных жестов и последующего управления роботом [1].

Таким образом, решение проблем, связанных с распознаванием сложных динамических жестов, находится на начальном уровне. Причем разнообразие жестов и способность человека понимать их настолько велики, что проблема их распознавания с помощью компьютера будет оставаться актуальной еще долгое время.



*Рисунок 1. Пример отслеживания жестов*

Отслеживание единичной гипотезы относится к оценке наилучшего соответствия при согласовании с минимальной ошибкой. Поэтому при использовании единичной гипотезы отслеживания, жест представлен только одной гипотезой. Большинство усовершенствованных алгоритмов отслеживания основаны на технологиях единичной гипотезы [1].

Отслеживание методом среднего сдвига – это базовая технология отслеживания. Метод выполняет сопоставление с гистограммами цвета *RGB*. Для каждого нового кадра средний сдвигающий трекер сравнивает расстояние Бхаттачарьи между гистограммами целевого окна нового кадра и кадрами старого кадра.

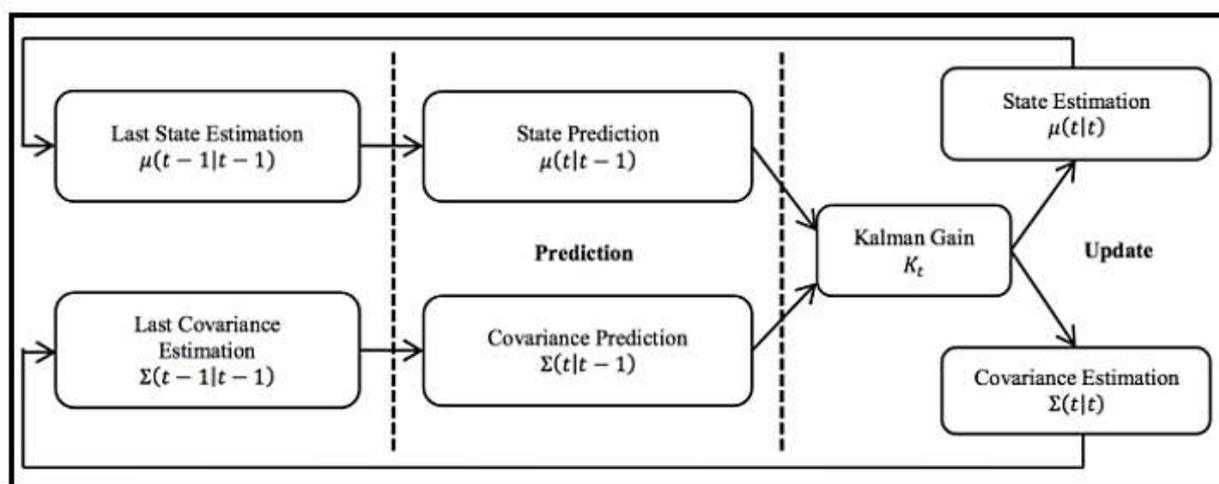


Рисунок 2. Процесс работы фильтра Калмана

Фильтр Калмана (*KF*) представляет собой рекурсивный алгоритм реального времени, используемый для оптимальной оценки лежащих в основе состояний ряда шумовых и неточных результатов измерений, наблюдаемых с течением времени. Технологический поток *KF* показан на рисунке 2. В настоящее время *KF* активно развивается и применяется в различных областях, таких как аэрокосмическая промышленность, робототехника и экономика.

*KF* предполагает, что вектор состояния является линейной моделью. Расширенный фильтр Калмана (*Extend Kalman Filter – EKF*) является функциональным алгоритмом отслеживания, даже если модель нелинейна. Другим алгоритмом, который решает ту же проблему под другим углом, является сигма-точечный фильтр Калмана (*Unscented Kalman Filter – UKF*). *UKF* решает проблему, применяя детерминированный подход с взвешенной выборкой. Распределение состояний представлено с использованием минимального набора выбранных точек выборки [2].

Многие исследователи пытались объединить *PF* с другими алгоритмами. Исследователи объединили *PF* со средним сдвиговым трекером, генетическим алгоритмом, *PSO*, *Ant Colony Optimization* и другими алгоритмами машинного обучения для решения проблемы вырождения и истощения выборки. Некоторые другие исследователи также улучшили стратегию передискретизации *PF*.

За последнее время появилось много продвинутых методов отслеживания. Некоторые из этих передовых методов использовали часть алгоритмов отслеживания, упомянутых выше. Другие методы улучшили отслеживаемость с помощью алгоритмов обнаружения или обучения.

Для долгосрочных проблем отслеживания многие алгоритмы отслеживания терпят неудачу, поскольку цель поддерживает фиксированные модели. Расширенное отслеживание модели сохраняет целевое поведение или внешний вид из последних кадров изображения. Поэтому для целевой оценки зарезервировано больше целевой информации. Инкрементный визуальный трекер использует расширенную модель для сохранения большего объема информации для процесса отслеживания. Квон представил отслеживание с помощью выборочного отслеживания. Расширенная модель сохраняется в процессе выборки. Трекер выбирает из многих трекеров и, соответственно, выбирается наиболее подходящий трекер [2].

Еще один вид алгоритмов отслеживания строится на базе алгоритмов обучения идентификации жестов. Для этих алгоритмов отслеживания в кадрах изображений применяется классификатор или детектор, чтобы идентифицировать жест из справочной информации. Одним из характерных подходов является трекер отслеживания, обучения и обнаружения. Этот подход объединяет результаты детектора объекта с устройством оптического отслеживания потока. Еще одна типичная технология отслеживания по обнаружению – применять множественное обучение экземпляров. Алгоритм обучения может повысить надежность трекера и уменьшить количество параметров [3].

**Список литературы:**

1. Anderson F., Grossman T., Matejka J., Fitzmaurice G. YouMove: enhancing movement training with an augmented reality mirror, in: Proceedings of the 26th annual ACM symposium on User interface software and technology [Text] / F. Anderson, T. Grossman, J. Matejka, G. Fitzmaurice. – ACM, 2013 – 320 с.
2. Adithya V., Rajesh R.A. Deep Convolutional Neural Network Approach for Static Hand Gesture Recognition. Procedia Computer Science. [Text] – 2020. – vol. 171. – P. 2353–2361.
3. Smeulders A.W., Chu D.M., Cucchiara R., Calderara S., Dehghan A., Shah M. Visual tracking: an experimental survey, Pattern Analysis and Machine Intelligence [Text]: A.W. Smeulders, D.M. Chu, R. Cucchiara, S. Calderara, A. Dehghan, M. Shah, IEEE Transactions on, 36 (2014) 1442–1468.