

Возможности алгоритма TLD по обеспечению трекинга целевого объекта в видеопотоке

В. А. Степанцов, email: mrstep@yandex.ru, М.А. Дынин

Воронежский государственный университет

***Аннотация.** Рассмотрены возможности алгоритма TLD (Tracking-Learning-Detection), обеспечивающие устойчивое слежение за целевым объектом в видеопотоке.*

***Ключевые слова:** трекинг, детектирование, объект в видеопотоке, TLD.*

Введение

Компьютерное зрение, обработка изображений и распознавание образов являются обширной областью научных и практических исследований, результаты которых широко используются в различных сферах человеческой деятельности, таких как дистанционное зондирование, медицинская диагностика, взаимодействие человек-компьютер и т.д. Системы безопасности, наблюдения, навигационные, ограниченного доступа, метеорологические являются системами анализа видеопотока. Важными задачами являются обнаружение, слежение, распознавание и прогнозирование поведения объекта на основе анализа видеопотока.

Слежение (трекинг) объектов используется в различных практических сферах:

– Автономная навигация – трекинг позволяет предсказать положение объекта, обеспечивая динамическое планирование маршрута.

– Системы внешнего наблюдения – трекинг позволяет отслеживать заданного человека с возможным исчезновением человека на видео.

– Человеко-компьютерный интерфейс – отслеживание определенной части тела позволяет использовать движения человека для ввода команд в компьютерных системах.

– Дополненная реальность – знание положения объекта на видео позволяет накладывать эффекты дополненной реальности.

В настоящее время сферы применения трекинга постоянно расширяются, что приводит к необходимости разработки или модификации ранее разработанных методов и алгоритмов (трекеров).

1. Представление объекта и способы слежения

В случае слежения за объектом выбирается нужный объект путем его выделения. Для выделения необходимо задать образ и характеристики объекта.

Существует несколько способов представления образа объекта:

– Точки. При данном подходе объект представляет из себя набор из одной, называемой центроидом, или нескольких точек. Трекер отслеживает перемещение каждой из точек.

– Прimitives геометрические фигуры. В данном случае перемещения объекта обычно моделируются как аффинное или проективное преобразование.

– Контур или силуэт объекта. Контур является границей объекта. Силуэт объекта – это площадь внутри контура.

– Сочлененные геометрические фигуры. В этой модели части объекта связаны вместе. Например, тело человека или животного можно представить, как набор соединенных геометрических фигур, представляющих различные части тела.

– Скелетонная модель. При этом подходе извлекается скелетон. Модель применима как для твердых, так и для подвижных объектов.

Наиболее часто применяемыми характеристиками объекта являются:

– Цвет. Следует отметить, что данная характеристика весьма чувствительна к шумам.

– Грани. Выделение (детектирование) наиболее изменяющихся границ объекта. Самым популярным методом выделения границ является детектор границ Canny [1].

– Оптический поток. Под оптическим потоком понимается изображение видимого движения объектов, поверхностей или краев сцены, получаемое в результате перемещения наблюдателя (глаз или камеры) относительно сцены. Алгоритмы расчета оптического потока реализуются на основе методов Horn–Schunck [2] и Lucas–Kanade [3].

Далее задача трекера, отслеживающего положение целевого объекта в видеопотоке, состоит в верном выделении объекта на последующих кадрах видеопоследовательности. Существующие алгоритмы располагают только информацией, полученной после выделения объекта, и имеют ряд недостатков, обусловленных отсутствием знаний об объекте слежения.

Отыскивать объект на видео возможно двумя способами: распознаванием и слежением.

В случае распознавания, программе, реализующей трекер, известно, как выглядит объект слежения, и в процессе работы

последовательно проверяются области изображения с целью нахождения похожих объектов. Очевидным недостатком данного способа является то, что практически невозможно отследить объект в случае его частичного или полного перекрытия другими объектами, а также в случаях сильного изменения его первоначального внешнего вида и характеристик видеосъемки из-за погодных или иных условий.

В случае слежения необходимо отслеживать первоначально выделенный объект оценивая оптический поток, который характеризует относительное изменение положение объектов на следующем кадре видео. На основе информации о нахождении объекта в предыдущих кадрах можно определить его скорость и направление движения, что дает возможность с определенной точностью прогнозировать следующее его положение. Недостатком является то, что даже кратковременное исчезновение объекта слежения из поля зрения объектива видеокамеры может привести к его полной потере.

Для устранения отмеченных недостатков необходимо использовать подход, в той или иной степени удовлетворяющий таким критериям как:

- устойчивая работа при полном или частичном пропадании объекта из поля зрения камеры;
- требуемая производительность;
- точность сопровождения.

В общем случае схема такого подхода может быть следующей:

- В начальный момент времени выделяется объект слежения.
- Классификатор обучается по выделенному и последующим изображениям объекта, получаемым трекером по относительному смещению объекта слежения.

2. Алгоритм TLD (Tracking-Learning-Detection)

В наибольшей степени этому подходу соответствует алгоритм TLD (Tracking-Learning-Detection) [4], состоящий из трех компонентов: краткосрочного Tracking (Сопровождения), Learning (Обучения) и Detection (Обнаружения). Алгоритм, в дополнение к основной задаче трекера, выполняет детектирование и моделирование объекта, тем самым повышая качество отслеживания. Это также позволяет алгоритму редетектировать объект в случае его исчезновения в видеопотоке. Схема взаимодействия модулей показана на рис. 1.

Модуль Tracking осуществляет сопровождение объекта от кадра к кадру. Модуль Detection локализует все появления объекта, которые наблюдались в последнее время и в случае необходимости корректирует модуль Tracking.

Модуль Detection на основе заданной модели объекта реализует алгоритм, позволяющий обнаружить объект на изображении.

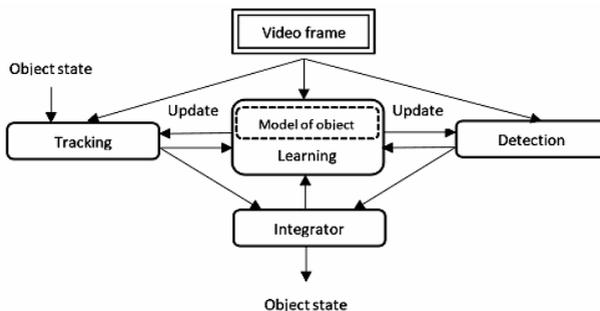


Рис. 1. Схема взаимодействия модулей алгоритма TLD

В алгоритме TLD используется стратегия сканирующего окна и модель, представленная набором патчей, на основе двухбитных бинарных шаблонов. Текущий видеокادر последовательно сканируется данным окном и для каждого фрагмента изображения бинарный классификатор определяет наличие или отсутствие там объекта. Классификатор содержит некоторый набор параметров, по каждому из которых на основе сравнения определяется соответствует ли данный фрагмент этому параметру. Если для фрагмента изображения положительно классифицировано более половины параметров, то классификатор делает вывод о принадлежности данного фрагмента объекту.

Модуль Learning оценивает ошибки модуля Detection и обновляет его, чтобы избежать их в будущем. Алгоритм является самообучающимся, при этом для обучения используются помеченные данные. В качестве данных о положении объекта выступают результаты работы модуля Tracking. В случае, если расстояние между положением, указанным модулем Detection и положением, указанным модулем Tracking, минимально, то патч определяется как положительный. В противном случае как отрицательный. Модель в алгоритме модуля Learning представляется в виде рандомизированного леса.

Обучение происходит путем применения двух событий: расширяющего, в случае если патч определен как положительный и урезающего, в случае если патч определен как отрицательный. Модель объекта инициализируется первым кадром, увеличивается расширяющими событиями, уменьшается урезающими событиями постепенно приближаясь к реальному объекту. Расширение модели необходимо для добавления новых правильных возможных состояний модели, и происходит добавлением и расширением веток. Урезание модели необходимо для удаления неверных сэмплов, и происходит

путем удаления веток из леса. Схема изменения модели приведена на рис. 2.

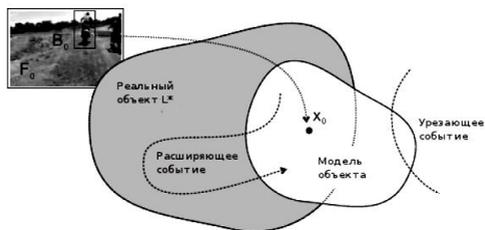


Рис. 2. Схема изменения модели

Модуль Integrator объединяет ограничивающую рамку модуля Tracking и ограничивающую рамку модуля Detection в единую ограничивающую прямоугольную рамку. Если ни модуль Tracking, ни модуль Detection не выводят ограничивающую рамку, объект объявляется невидимым.

Заключение

Комбинированные методы, к которым относится алгоритм TLD, позволяют объединять достоинства разных методов по наивысшим показателям различных критериев. К достоинствам данного алгоритма следует отнести высокий уровень обнаружения объектов, высокую производительность, устойчивость к шуму и продолжительным перекрытиям отслеживаемого объекта, а также к различным видам искажений внешнего вида.

Список литературы

1. Canny J., A Computational Approach to Edge Detection – [Электронный ресурс]. – Режим доступа: https://perso.limsi.fr/vezien/PAPIERS_ACS/canny1986.pdf
2. Horn B. K. P., Schunck B. G., Determining Optical Flow – [Электронный ресурс]. – Режим доступа: <https://www.caam.rice.edu/~zhang/caam699/opt-flow/horn81.pdf>
3. Lucas Bruce D., Kanade Takeo, An Iterative Image Registration Technique with an Application to Stereo Vision – [Электронный ресурс]. – Режим доступа: <http://cseweb.ucsd.edu/classes/sp02/cse252/lucaskanade81.pdf>
4. Kalal Z., Mikolajczyk K., Matas J., Tracking-Learning-Detection – [Электронный ресурс]. – Режим доступа: http://vision.stanford.edu/teaching/cs231b_spring1415/papers/KalalPAMI.pdf