DOI: 10.17117/na.2015.12.02.040

<http://ucom.ru/doc/na.2015.12.02.040.pdf>

Поступила (Received): 24.12.2015

Буй Ван Шон, Архипенко И.Е., Литвинов Ю.В. Алгоритм для отслеживания объектов по видеоизображению

Bui Van Sean, Arkhipenko I.E., Litvinov Yu.V.
An algorithm for tracking video image

В статье рассмотрена проблема отслеживания неизвестного объекта в видеопотоке, где объект меняет направление и часто пропадает из поля зрения камеры. Разработана новая структура, которая раскладывает задачу на три компоненты: отслеживание, обучение и обнаружение. Компонент «Обучение» детально проанализирован. Показано, что детектор объекта может быть обучен на одном примере выбранного видеопотока, используя следующую стратегию: оценивают детектор, оценивают свои ошибки с помощью пары экспертов и обновляют классификатор

Ключевые слова: отслеживание, видеопоток, детектор объекта, обучение и обнаружение

Буй Ван Шон

Магистрант

Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики
г. Санкт-Петербург, Кронверкский проспект, 49

Архипенко Игорь Евгеньевич

Магистрант

Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики
г. Санкт-Петербург, Кронверкский проспект, 49

Литвинов Юрий Володарович

Кандидат технических наук, доцент, старший преподаватель

Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики
г. Санкт-Петербург, Кронверкский проспект, 49

The article considers the problem of tracking an unknown object in a video stream where the object changes direction and often disappears from the camera view. Developed a new structure that decomposes the task into three components: tracking, learning and detection. A Training component is analyzed in detail. It is shown that the object detector can be trained on one example, the selected video stream using the following strategy: assess the detector, estimate its errors by a pair of experts and update the classifier

Key words: tracking, the video stream object detector training and detection

Bui Van Shawn

Master

Saint-Petersburg national research university of information technologies, mechanics and optics
Saint-Petersburg, Kronverkskiy ave., 49

Archipenko Igor Evgenievich

Master

Saint-Petersburg national research university of information technologies, mechanics and optics
Saint-Petersburg, Kronverkskiy ave., 49

Litvinov Yuri Volodarovich

Candidate of Technical Sciences, Associate Professor, Senior Lecturer

Saint-Petersburg national research university of information technologies, mechanics and optics
Saint-Petersburg, Kronverkskiy ave., 49

В статье описывается исследование алгоритма надёжного длительного сопровождения заранее неизвестных объектов в естественной среде. Алгоритм убирает разрывы между кадрами, обеспечивает быстрое движение камеры, полное исчезновение, а затем появление объекта в кадре. Подход, который использован в данном алгоритме называется *опровождение-Моделирование-Обнаружение (Tracking-Modeling-Detection (TMD))*, он сочетает адаптивное сопровождение объекта с обучением детектора объекта в процессе распознавания. После того как объект был захвачен при помощи какого-либо алгоритма захвата, траектория объекта начинает наблюдаться двумя процессами (*расширяющее и урезающее события*). Они строят детектор объекта на лету. Оба процесса делают ошибки, стабильность системы достигается отменой событий. Обучение на лету и классификация производятся при помощи рандомизированного леса.

Математическое описание алгоритма

Пусть A^t и B^t – кадр видео ряда и описанный прямоугольник сопровождаемого объекта в момент времени t . Пиксели внутри прямоугольника B^t описываются вектором признаков x^t , который содержит информацию о наличии объекта сопровождения. Множество последовательных описанных прямоугольников определяет трек $T_t = \{B_0, B_1, \dots, B_t\}$ длины $t + 1$, который определяет траекторию объекта в пространстве изображений. T_t^f описывает соответствующую траекторию в пространстве признаков U , которое является подпространством L^* . L^* представляет все возможные состояния наблюдаемого объекта. L^* неизвестно в момент начала сопровождения, когда выбрано первое измерение $x_0 \in L^*$ [1]. Это первое измерение описывает начальное состояние модели в момент времени $t = 0$: $L_0 = \{x_0\}$.

Алгоритм работает следующим образом: выделенный объект сопровождается при помощи краткосрочного трекера. Траектория в пространстве признаков анализируется двумя событиями, которые непрерывно пытаются расширить или уменьшить пространство, описываемое моделью. L_t расширяется измерениями, которые скорее всего содержат сопровождаемый объект.

Рис. 1. Модель объекта инициализируется первым кадром. Она увеличивается расширяющими событиями или уменьшается урезающими событиями. Постепенно она приближается к реальному объекту

Эти измерения определяются при помощи расширяющего события. Из L_t удаляются измерения, которые определены как не содержащие объект. Эти измерения определяются при помощи обрезающих событий. События работают параллельно, стремясь достичь $L_t \rightarrow L^*$ [2] (рис. 1).

Главная цель построения L_t – это предоставить алгоритму память, чтобы создать детектор объекта, который непрерывно обновляется и улучшается. Он сканирует входное изображение F_t и выдаёт на выход множество описанных прямоугольников, которые содержат измерения, входящие в L_t . Эти прямоугольники описывают альтернативные гипотезы к позиции, возвращённой трекером. Слияние гипотез производится взятием позиции, которая минимизирует расстояние до L_t . Отсюда следует, что если положение от трекера очень близко к L_t , то ложные отклики детектора не влияют на трек (пока они не станут ещё ближе к L_t , чем позиция от трекера). Минимальное расстояние до L_t становится очками недоверия к результату, выданному алгоритмом. Опираясь на эти очки, алгоритм решает, виден объект или нет [3, 7, 8].

Алгоритм 1 TMD

Require Выбрать x_0 , $L_0 = \{x_0\}$

for $t=1:\infty$ **do**

Отслеживать последний патч $x_t - 1$.

Определить патчи, содержащиеся в модели $L_t - 1$.

$L_t \leftarrow L_t - 1 \cup$ Положительные измерения из расширения.

$L_t \leftarrow L_t - 1 \setminus$ Отрицательные измерения из урезания.

$x_t \leftarrow$ Наиболее подходящий патч (определённый или от трекера).

end for

Ранее было сказано о событиях, которые наблюдают краткосрочный трекер. Рассмотрим как они работают. Мы различаем две части L_t : правильную часть $L_t^c \subset L^*$ и ошибочную часть $L_t^e \not\subset L^*$, $L_t^c \cup L_t^e = L_t$, $L_t^c \cap L_t^e = \emptyset$. *Покрывание (coverage)* является долей измерений с объектом уже открытых неконтролируемым процессом обучения, т.е. $coverage(L_t) = |L_t^c|/|L^*|$. *Примесью (impurity)* называется доля L_t , которая неверна, т.е. $impurity(L_t) = |L_t^e|/|L^*|$. Оператор $||$ обозначает размер множества [4].

Расширяющие события. В момент времени t краткосрочный трекер создал траекторию $T_t^f = \{x_0, x_1, \dots, x_t\}$. Расширяющее событие сначала выбирает определённую часть траектории, которая рассматривается положительной, $P \subset T_t^f$. Модель объекта L_t затем обновляется, т.е. $L_t = L_t - 1 + P$. После этого обновления покрытие модели увеличивается, если P содержит хотя бы одно измерение из L^* . Стратегию выбора части траектории обсудим несколько ниже [5].

Обрезающие события. Невозможно придумать стратегию выбора, которая бы выбирала только верные измерения. По этой причине примесь L_t постоянно увеличивается. Так как нашей целью является условие $L_t \rightarrow L^*$, то события, уменьшающие примесь модели, являются необходимыми. Если модель характеризуется определённым уровнем примеси, обрезающее событие необходимо, чтобы идентифицировать подмножество N , которое рассматривается как некорректное и удалить его из модели, т.е. $L_t = L_t - 1 - N$ [5].

Расширяющие события в одиночку приводят к высокому уровню примеси и, следовательно, к детектору низкой точности. Обрезающие события служат отрицательной обратной связью: чем выше уровень примесей, тем больше измерений в модели определяются как неверные и удаляются. Динамическое взаимодействие расширяющих и обрезающих событий является решающим в придании системы стабильности, это будет показано эмпирически в экспериментальном разделе [6].

Реализация

Краткосрочный трекер основывается на методе Лукаса-Канаде (ЛК). Вначале множество ключевых точек извлекается из прямоугольной решётки внутри описанного вокруг сопровождаемого объекта прямоугольника. Затем ЛК сопровождает точки от одного кадра до другого, строя разреженное поле движения. Основываясь на поле движения, смещение и изменение масштаба ограничивающего прямоугольника могут быть надёжно оценены как средние значения по распределению. В каждом новом кадре сопровождается новый набор точек, это делает трекер адаптивным.

Рис. 2. Двухбитные бинарные шаблоны – признаки, используемые в детекторе объекта

Модель представлена множеством 15×15 нормированных по интенсивности патчей. Расстояние между патчами определено при помощи нормированной кросс-корреляцией, т.е. $distance(x_i, x_j) = 1 - NCC(x_i, x_j)$. Расстояние от измерения x_i до модели L_t определено как $distance(x_i, L_t) = \min_{x \in L_t} (distance(x_i, x))$. Модель основывается на патчах, чтобы построить эффективный и способный к обобщению детектор объекта.

Крайне важно построить быстрый и надёжный детектор, способный локализовать патчи, содержащиеся в модели, и способный эффективно изменять свою границу решений при помощи расширяющих и обрезающих событий. Хорошие детекторы объектов в реальном времени обычно базируются на алгоритме AdaBoost. Они требуют обучающие множества большого размера, а обучение обладает высокой вычислительной сложностью. Данный подход не применим к обучению на лету. Детектор объекта основан на двухбитных бинарных шаблонах. Эти признаки измеряют ориентацию градиента внутри определённой зоны, квантуют её и выдают 4 возможных кода [7] (рис. 2).

Заключение

В статье рассмотрена проблема отслеживания неизвестного объекта в видеопотоке, где объект меняет направление и часто пропадает из поля зрения камеры. Разработана новая структура, которая раскладывает задачу на три компоненты: отслеживание, обучение и обнаружение. Компонент «Обучение» детально проанализирован. Показано, что детектор объекта может быть обучен на одном примере выбранного видеопотока, используя следующую стратегию: оценивают детектор, оценивают свои ошибки с помощью пары экспертов и обновляют классификатор. Каждый эксперт ориентирован по выявлению конкретного типа ошибки классификатора, при этом допускаются ошибки самого эксперта. Стабильность обучения достигается путем разработки экспертов, которые взаимно компенсируют их ошибки. Теоретический вклад: формализация этого процесса в дискретной динамической системы, что позволило указать условия, при которых процесс обучения гарантирует улучшение классификатора. Показано, что эксперты могут использовать пространственно-временные отношения в видео. Был выполнен обширный набор экспериментов и описана реализация в режиме реального времени.

Список используемых источников:

1. Omar Javed, Mubarak Shah (2006) *Object Tracking: A Survey*. ACM Computing Surveys, Vol. 38. No. 4, Article 13.
2. Random Forests Leo Breiman and Adele Cutler.
URL: http://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm
3. Zdenek Kalal, Jiri Matas, Krystian Mikolajczyk (2009) *Online learning of robust object detectors during unstable tracking*. Computer Vision Workshop 2009, Kuomo, Япония, IEEE CS.
4. Zdenek Kalal, Jiri Matas, Krystian Mikolajczyk (2010) *P-N Learning: Bootstrapping Binary Classifiers by Structural Constraints*. Conference on Computer Vision and Pattern Recognition, CVPR, Июнь 13-18, Сан-Франциско, США, 2010
5. Zdenek Kalal, Krystian Mikolajczyk, Jiri Matas (2010) *Forward-Backward Error: Automatic Detection of Tracking Failures*. International Conference on Pattern Recognition, 23-26 Август 2010, Стамбул, Турция.
6. Donald Knuth (1969) *The Art of Computer Programming. Volume 2*. AddisonWesley.

7. Буй В.Ш., Бушуев А.Б., Шмигельский Г.М., Литвинов Ю.В., Щаев Е.Г. Алгоритмы управления летающим роботом при слежении за подвижным объектом // Известия высших учебных заведений. Приборостроение. 2015. Т. 58. № 8. С. 593-599.
8. Буй В.Ш., Литвинов Ю.В., Щаев Е.Г. Использование технического зрения для управления летающим роботом при слежении за подвижным объектом // Перспективы развития науки и образования. Ч. 12. 2015. С. 33-36.

© 2015, Буй Ван Шон, Архипенко И.Е.,
Литвинов Ю.В.
Алгоритм для отслеживания объектов по
видеоизображению

© 2015, Bui Van Sean, Arkhipenko I.E.,
Litvinov Yu.V.
An algorithm for tracking video image