

Использование метода преследования для повышения быстродействия алгоритма многоклассовой детекции объектов в видеопотоке каскадами Виолы-Джонса*

С.А. Усилин

Аннотация. В данной работе исследуется задача многоклассовой детекции объектов в видеопотоке с помощью каскадов Виолы-Джонса. Описывается алгоритм адаптивного выбора распознающего каскада, в основе которого лежит стратегия жадного выбора действия в задаче n -рукого бандита. Показана эффективность алгоритма на примере поиска и распознавания логотипов банковских карт в видеопотоке. Представленный алгоритм может быть эффективно применен в таких задачах, как локализация и идентификация документов, распознавание элементов дорожной сцены, поиск природных и техногенных объектов в прибрежной зоне Арктики, а также для решения ряда других задач идентификации ригидных объектов в потоках гетерогенных данных.

Ключевые слова: машинное обучение, детектирование образов, каскады Виолы и Джонса, задача об n -руком бандите, ε -жадный метод, *softmax*-метод экспоненциальное среднее.

Введение

В задачах поиска объектов важное значение имеет скорость обработки одного изображения. В простых ситуациях, когда ведется поиск объектов только одного типа метод Виолы-Джонса [1] обеспечивает стабильную скорость детектирования, достаточную для решения задачи в режиме реального времени. При этом проблему детектирования нескольких разнородных объектов решают с помощью обучения нескольких независимых классификаторов, каждый для своего типа объектов [2, 3]. В этом случае, закономерно, процедура поиска объектов замедляется линейно с ростом искомым объектов и часто выходит за допустимые временные рамки.

Обычной с практической точки зрения случаем является ситуация, когда несмотря на глобальную необходимость детектировать несколько разнородных объектов, в каждый момент времени на видеокадре может присутствовать не более одного объекта. В качестве примера рассмотрим задачу, когда банковский служащий с помощью веб-камеры регистрирует выдаваемые клиентам банковские карты, при этом требуется по нанесенному логотипу определить тип карты и распознать номер. Несмотря на то, что в настоящее время существует более десяти различных видов платежных систем (а, следовательно, более 10 различных логотипов), в каждый момент времени оператор показывает только одну карту. В этом

случае поиск всевозможных логотипов на видеокадре заведомо избыточен.

Другим важным примером является локализация различных элементов на серии изображений (кадров) Арктической зоны, получаемых со спутников или БПЛА. В этом случае, не смотря на большое количество возможных различных природных и техногенных объектов, которые могут встретиться на полученных изображениях, за счет низкой пространственной плотности скопления таких объектов можем считать, что в каждый момент времени в зоне интереса находится не более одного объекта.

В обоих случаях возникает задача выбора «правильного» распознающего классификатора для текущих видеокадров, которая может быть решена аналогично задаче об n -руком бандите [4-6], который в отличие от классического игрового автомата («однорукого бандита»), содержит n рычагов, при этом в каждой игре участвует только один рычаг. Можно провести аналогию между задачей выбора «правильного» классификатора и задачей выбора «лучшего» рычага n -рукого бандита, если заметить, что успешное детектирование объекта на очередном кадре в некотором смысле аналогично хорошему вознаграждению в игре, а видеофрагмент – последовательной серии игр.

В данной работе описан оригинальный метод адаптивного выбора распознающих каскадов Виолы-Джонса, который использует информацию, накопленную по уже обработанным видеокадрам с целью выбора наилучшего каскада для текущего видеокадра.

* Работа выполнена при поддержке РФФИ, проекты № 16-29-09508_офи_м, № 15-29-06079_офи_м.

1. Задача об n -руком бандите

Исходная форма задачи об n -руком бандите формулируется следующим образом. Пусть приходится многократно осуществлять выбор одной из n различных альтернатив (вариантов действий). Каждый выбор влечет за собой получение определенного вознаграждения, зависящего от выбранного действия. Среднее вознаграждение за выбор данного действия называется ценностью действия. Целью последовательности действий является максимизация ожидаемого полного вознаграждения за заданный период времени. Часто каждая такая попытка называется игрой.

Хотя точные значения ценности действий не известны, их оценки могут быть получены с каждой новой игрой. Тогда в каждый момент времени найдется хотя бы одно действие, для которого такая оценка будет наибольшей (такие действия называются жадными). Под применением понимается выбор жадного действия в очередной игре. Под изучением понимается выбор одного из нежадных действий в данной игре с целью более точной оценкой ценности нежадного действия.

Обозначим через $V_i(a)$ предполагаемое значение ценности действия a в i -ой игре. Если к моменту времени t начала t -ой игры действие a было выбрано k_a раз, что привело к последовательному получению вознаграждения r_1, r_2, \dots, r_{k_a} , то ценность действия будет оцениваться по следующей формуле:

$$V_i(a) = \begin{cases} \frac{r_1 + r_2 + \dots + r_{k_a}}{k_a}, & k_a > 0 \\ 0, & k_a = 0 \end{cases}. \quad (1)$$

Несмотря на то, что проще всего в каждой игре выбирать действие, обладающее максимальной предполагаемой ценностью, такой подход не приводит к успеху из-за отсутствия процедуры исследования. Наиболее простой и самый популярный способ избавиться от этой проблемы – это использовать ϵ -жадный метод [7-9], когда равномерным образом с вероятностью ϵ выбирается произвольное действие, вне зависимости от его ценности.

Хотя ϵ -жадный выбор действия является достаточно эффективным и распространенным способом достижения баланса между изучением и применением в задаче о n -руком бандите, его недостаток состоит в том, что при изучении выбор среди всех действий осуществляется равновероятно. Если для некоторых задач выбор наилучшего действия даже на этапе изучения является нежелательным, то ϵ -жадный выбор действия может не подойти. В этом случае можно варьировать вы-

бором действия посредством некоторой функции от предполагаемых значений действий. Для жадного действия вероятность выбора по-прежнему будет наибольшей, но все остальные действия будут ранжированы согласно их предполагаемой ценности. Такой принцип выбора действия называется правилом на основе операции softmax или softmax-методом [9, 10]. В самом распространенном варианте softmax-метода используется распределение Гиббса, тогда выбор действия a в игре t осуществляется с вероятностью, посчитанной по формуле:

$$\pi(a) = \frac{e^{\frac{V_i(a)}{\tau}}}{\sum_{b=1}^n e^{\frac{V_i(b)}{\tau}}}, \quad (2)$$

где τ – некоторый положительный параметр, называемый температурой. Высокие температуры делают все действия примерно равновероятными. Низкие температуры вызывают более значительное различие в вероятности выбора действий, имеющих разную ценность.

2. Задача выбора распознающего каскада

Пусть есть видеосекция, содержащий T кадров. На каждом кадре видеопоследовательности может находиться не более одного объекта, который требуется детектировать. Всего N различных типов объектов, причем для каждого типа обучен собственный каскад Виолы-Джонса. Объекты на видеопоследовательности появляются и исчезают естественным путем. Не существует мгновенного появления или исчезновения объектов (например, в задаче регистрации банковских карт с помощью веб-камеры, описанной во введении. Естественность заключается в плавном последовательном появлении и исчезновении карты в кадре). Необходимо построить алгоритм выбора распознающего классификатора для каждого кадра видеосекции, обеспечивающий максимально точное детектирование объектов.

Пусть $V = \{F_i\}_{i=1}^T$ – видеосекция, содержащий T кадров. На каждом кадре может присутствовать не более одного объекта, то есть каждому кадру F_i можно поставить в соответствие число $j_i^* \in \{0, 1, \dots, N\}$, причем $j_i^* = 0$ соответствует отсутствию какого-либо объекта на рассматриваемом кадре. Пусть $C = \{C^j(F) : F \rightarrow \{0, j\}\}_{j=1}^N$ – множество обученных каскадов Виолы-Джонса. Требуется построить правило адаптивного выбора классификатора для очередного кадра из видео-

фрагмента $a: V \rightarrow C$ так, чтобы максимизировать количество правильно найденных объектов.

Предположим, на очередном кадре F_i в качестве действия выбран классификатор $C_i = C^j$. Будем считать, что данный кадр распознан правильно, если $C_i(F_i) = j_i^*$. Тогда функционал качества для правила адаптивного выбора классификатора a определяется следующим образом:

$$Q(a, V) = \sum_{i=1}^T L(a, F_i), \quad (3)$$

где $L(a, F_i)$ – индикатор ошибки, определяемый следующим образом:

$$L(a, F_i) = \begin{cases} 1, & a(F_i) = j_i^* \\ 0, & a(F_i) \neq j_i^* \end{cases}. \quad (4)$$

3. Жадный алгоритм выбора каскада

Алгоритм жадного выбора распознающего каскада реализует идею жадного выбора действия, представленного выше, с учетом особенностей задачи поиска объектов.

Как говорилось выше, в жадном методе учитываются оценки ценности действий, которые вычисляются по вознаграждениям в отдельных играх. И если в задаче об n -руком бандите вознаграждение возникает естественным образом, то в задаче поиска объектов понятие «вознаграждение» требует отдельного определения. Интуитивно понятно, что вознаграждение за выбор действия (расознающего каскада) должно быть положительным, если удалось правильно найти объект, или нулевым в противном случае. Однако в момент игры достоверной информации об имеющихся на кадрах объектов нет, так как в противном случае задача поиска объектов на изображении не имеет смысла. Тогда, предполагая, что имеющиеся каскады обладают хорошей точностью и полнотой, будем поощрять каскады только за найденные на изображении объекты (то есть вознаграждение в очередной игре равно 1 при нахождении объекта на кадре и 0 в противном случае).

Вторым отличием предлагаемого алгоритма от оригинального жадного метода является модификация в части адаптации к нестационарной задаче. Способ вычисления ценности действия по формуле (1) хорошо подходит только для стационарных задач. В нестационарных задачах более поздние вознаграждения обладают большим приоритетом, чем более ранние. Наиболее распростра-

ненный способ этого добиться состоит в использовании экспоненциального среднего [11, 12]. Тогда ценность действия a при получении очередного вознаграждения r_{k+1} определяется по следующей рекурсивной формуле:

$$V_{k+1}(a) = \alpha r_{k+1} + (1 - \alpha)V_k(a), \quad (5)$$

где $\alpha \in [0; 1]$ – размер шага (чем больше значение α , тем больший вес в ценности действия имеет новое вознаграждение). С физической точки зрения параметр α регулирует насколько быстро текущее действие становится жадным.

В рассматриваемой задаче поиска объектов величина параметра α может быть оценена при наличии некоторых знаний о входной видеопоследовательности. Значение параметра α должно быть таким, чтобы за время смены объектов в зоне интереса, текущий жадный каскад перешел из состояния с большим значением ценности (которое не может превышать 1) в начальное (когда значение ценности близко к нулю). Тогда, если предположить, что смена объекта в зоне видимости происходит за M кадров (то есть M кадров в зоне видимости нет никаких объектов), вознаграждение на таких кадрах нулевое, справедливо следующее выражение:

$$V_M(a) = \alpha r_M + (1 - \alpha)V_{M-1}(a) = (1 - \alpha)V_{M-1}(a) = (1 - \alpha)^M V_0. \quad (6)$$

Зная, что начальное значение ценности $V_0 \leq 1$, а конечно – $V_M = \varepsilon \approx 0$, то для параметра α справедливо следующее выражение:

$$\alpha \leq 1 - \sqrt[M]{\varepsilon}. \quad (7)$$

С учетом вышесказанного алгоритм адаптивного выбора классификатора представим в виде следующей последовательности шагов. Блок-схема описанного алгоритма представлена на рис. 1.

Шаг 1 (Инициализация). Пусть есть N классификаторов. Зададим начальное значение ценности для классификатора $V^1 = V^2 = \dots = V^N = 0$. Оценим, исходя из имеющихся данных, значение параметра α .

Шаг 2. Выберем классификатор C с максимальным текущим значением ценности V_{curr} . Если таких классификаторов несколько, выберем один из них произвольным образом.

Шаг 3. Применим на очередном кадре F выбранный классификатор C . Определим выигрыш классификатора следующим образом:

$$r = \begin{cases} 1, & \text{если на } F \text{ есть объект} \\ 0, & \text{если на } F \text{ нет объекта} \end{cases}. \quad (8)$$

Шаг 4. Обновим значение ценности V_{new} для классификатора C по следующей формуле:

$$V_{new} = \alpha r + (1 - \alpha)V_{curr}. \quad (9)$$

Шаг 5. Перейдем на шаг 2, если есть еще кадры для поиска объектов. Иначе закачиваем обработку.

4. Экспериментальные результаты

Эффективность описанного алгоритма оценивалась путем применения для задачи определения типа банковской карточки в видеопотоке. Оператор фотографирует с помощью web-камеры банковскую карту, а система распознавания должна определить тип карты по имеющемуся логотипу. Всего тестовый видеофрагмент содержал 29 представлений банковских карт следующих пяти типов: VISA, MasterCard, American Express, Discover и UnionPay. Общее количество кадров видеофрагмента составило 1116.

Распределение кадров по содержанию объектов, а также общее количество кадров тестового видеофрагмента представлено в табл. 1.

Таблица 1
Состав видеофрагмента

Тип объекта	Кол-во карт	Общее кол-во кадров
VISA	16	302
MasterCard	8	139
American Express	2	37
UnionPay	1	17
Discover	2	34
Кадры без карты	-	587
Итого:	29	1116

Детектирование карт выполнялось путем использования каскадов логотипов соответствующих платежных систем, построенных с помощью инструментария objed [13]. В качестве положительной обучающей выборки использовались изображения логотипов, вырезанные из фотографий соответствующих банковских карт (данные о количестве фотографий представлены во втором столбце табл. 2. Из-за малого количества обучающих образцов использовалась аугментация (синтез обучающих данных на основе существующих) [14-16] данных с помощью геометрических и яркостных преобразований: поворот изображения логотипа вокруг центра на углы $\pm 5^\circ$ и $\pm 3^\circ$, и приме-

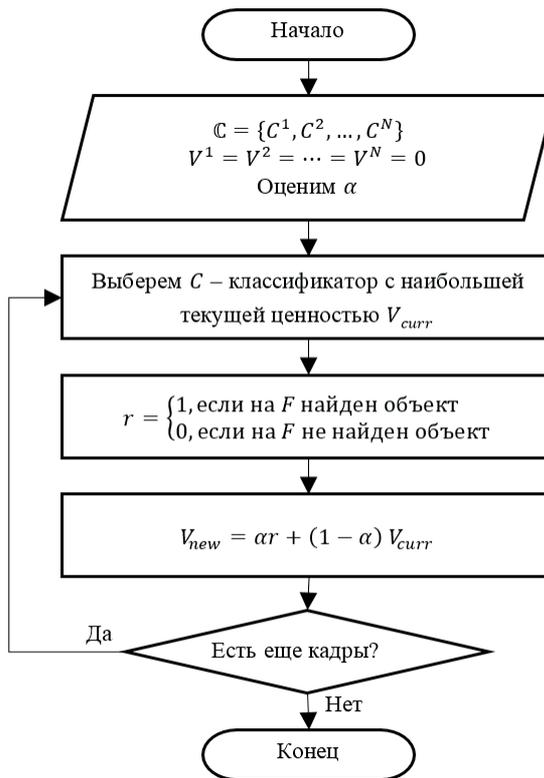


Рис. 1. Блок-схема алгоритма выбора распознающего каскада Виолы-Джонса

нение гамма-коррекции с коэффициентами 0,5 и 2,0, что позволило увеличить обучающий набор изображений в 15 раз (см. третий столбец табл. 2). В качестве отрицательной обучающей выборки использовались фрагменты фотографий банковских карточек, на которых отсутствовал логотип платежной системы.

Чтобы оценить эффективность, качество работы представленного алгоритма выбора распознающих каскадов (далее – АВРК) со значением параметра $\alpha=0.9$ было сравнено с

Таблица 2
Обучающий набор

Тип объекта	К-во изоб. (исходно)	К-во изоб. (аугмент.)
VISA	190	2850
MasterCard	160	2400
Am. Express	81	1215
UnionPay	8	120
Discover	9	135
Изо-я без карты	1792	-
Итого:	2240	6720

Таблица 3

Метод	Q (Q _r)			
	300	600	900	1116
Поочередный	191 (0.64)	379 (0.63)	547 (0.61)	676 (0.61)
ϵ -жадный ($\epsilon=0.1$)	210 (0.70)	432 (0.72)	612 (0.68)	773 (0.69)
softmax ($\tau=0.1$)	201 (0.67)	411 (0.68)	581 (0.65)	724 (0.65)
АВРК ($\alpha=0.9$)	216 (0.72)	446 (0.74)	682 (0.76)	852 (0.76)

качеством работы следующих трех методов: поочередным методом (обученные каскады применяются к кадрам по очереди), ϵ -жадным методом при $\epsilon = 0.1$ и softmax-методом при $\tau = 0.1$. Значение параметра α для представленного алгоритма было рассчитано по формуле (7), опираясь на следующую информацию: минимальный интервал между демонстрациями соседних карт составляет 16 кадров, а значение $\epsilon = 2^{-52} \approx 2.20 \cdot 10^{-16}$ (что соответствует значению машинного эпсилон [17] для типа double 64 бита в языке Си). Параметры для ϵ -жадного метода и softmax-метода выбраны исходя из экспериментальных исследований.

Графические характеристики и поведение различных методов показано на рис. 2. В табл. 3 представлено значение функционала качества Q , а также отношение функционала качества к общему числу обработанных кадров Q_r для отдельных интервальных точек. Полученные результаты свидетельствуют о том, что методы решения задачи об n -руком бандите применимы для повышения быстродействия алгоритмов поиска объектов (простой поочередный алгоритм демонстрирует результаты существенно хуже всех остальных методов). Кроме того, важно отметить, что предложенный алгоритм АВРК опережает ϵ -жадный и softmax-методы на всех уровнях эксперимента.

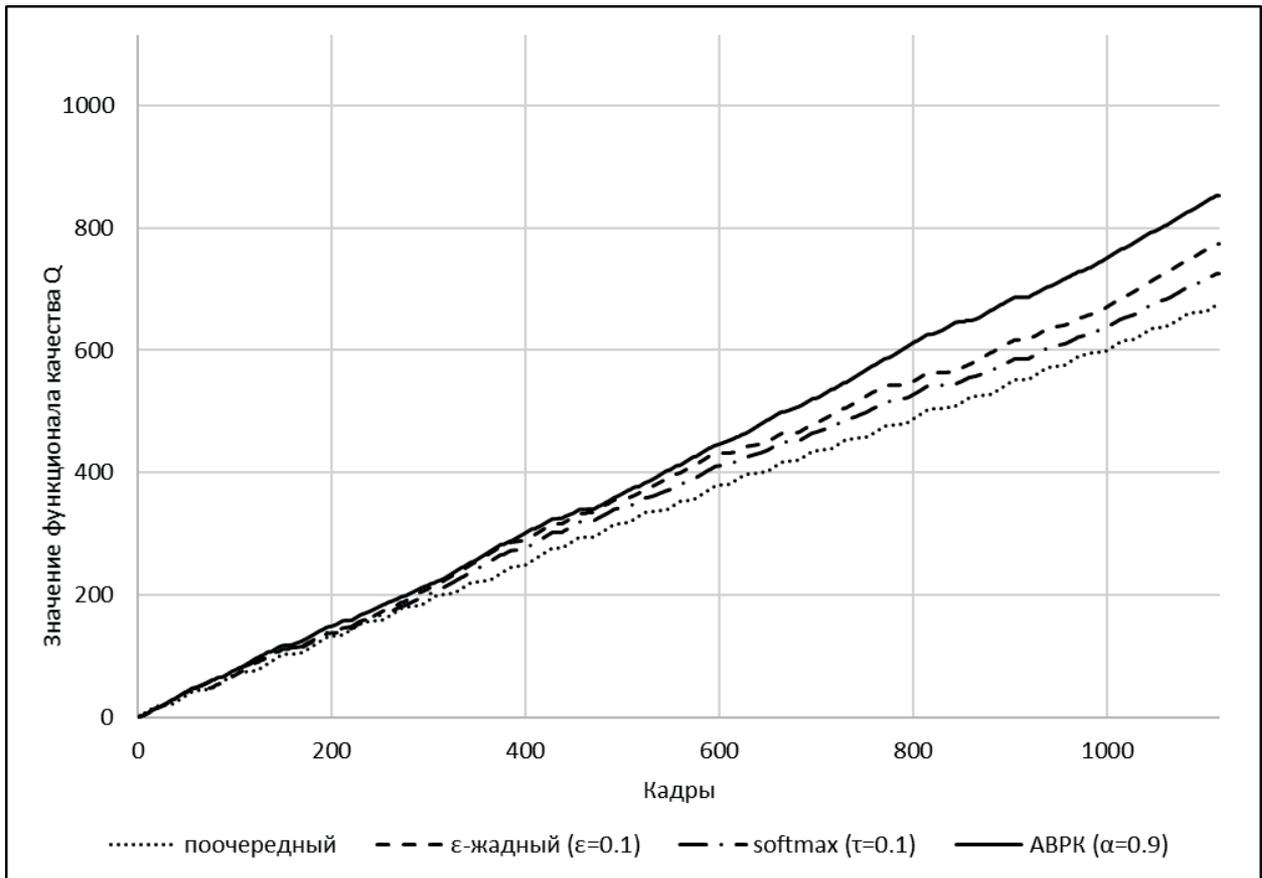


Рис. 2. График зависимости значения функционала качества Q от номера кадра

Несмотря на то, что представленный алгоритм по сравнению с рассматриваемыми методами обеспечивает наилучшее качество детектирования, тем не менее идеальное качество не достигается. Отчасти это связано с собственными параметрами качества используемых каскадов Виолы-Джонса, определяемые параметрами точности и полноты [18, 19]. Так как каскады для всех видов логотипов обучены, был проведен эксперимент по применению к каждому изображению видеопотока, в результате которого было получено максимально возможное значение функционала качества $Q^* = 1030$ ($Q_r^* = 0.92$). Следовательно, улучшение качества детектирования каскадов приведет к повышению качества работы представленного алгоритма.

Заключение

В задачах поиска объектов в видеопотоке важное значение имеет скорость обработки отдельного кадра. Существуют задачи, для которых даже применение быстрых алгоритмов, таких как метод Виолы-Джонса, не обеспечивает необходимой скорости. В качестве важного с практической точки зрения примера можно привести проблему многоклассовой детекции объектов в видеопотоке с помощью каскадов Виолы-Джонса.

В данной работе представлен алгоритм адаптивного выбора распознающего каскада, в основе которого лежит стратегия жадного выбора действия в задаче n -рукого бандита. Эффективность описанного алгоритма была продемонстрирована на примере поиска и распознавания логотипов банковских карт в видеопотоке путем сравнения с классическими для задачи n -рукого бандита стратегиями: ϵ -жадным методом и softmax-методом.

Представленный в работе алгоритм может быть эффективно применен в таких задачах, как локализация и идентификация документов, распознавание элементов дорожной сцены, поиск природных и техногенных объектов в прибрежной зоне Арктики, а также для решения ряда других задач идентификации ригидных объектов в потоках гетерогенных данных.

Литература

1. *Viola P., Jones M.* Robust Real-time Object Detection // International Journal of Computer Vision. 2002.
2. *Jones M., Viola P.* Fast Multi-view Face Detection // Mitsubishi Electric Research Lab TR2000396. 2003. No. July.
3. *Thompson W.R.* On the likelihood that one unknown probability exceeds another in view of the evidence of two samples // Biometrika. 1933. Vol. 25, No. 3–4. P. 285–294.
4. *Robbins H.* Some Aspects of the Sequential Design of Experiments // Bulletin of the American Mathematical Society. 1952. Vol. 58, No. 5. P. 527–535.
5. *Sutton P.C., Барто Э.Г.* Обучение с подкреплением / Пер. с англ. М.: БИНОМ. Лаборатория знаний, 2012. 399 p. (Sutton R.S., Barto A.G. Reinforcement Learning. London: The MIT Press. 2012. 399 p.)
6. *Gittins J.C.* Bandit processes and dynamic allocation indices // Journal of the Royal Statistical Society, Series B. 1979. Vol. 41, No 2. P. 148–177.
7. *Thathachar M.A.L., Sastry P.S.* A New Approach to the Design of Reinforcement Schemes for Learning Automata // IEEE Transactions on Systems, Man and Cybernetics. 1985. Vol. SMC-15, No 1. P. 168–175.
8. *C. J. C. H. Watkins.* Learning from Delayed Rewards. Ph.D. thesis. Cambridge University, 1989.
9. *Tokic M., Palm G.* Value-difference based exploration: Adaptive control between epsilon-greedy and softmax // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2011. Vol. 7006 LNAI. P. 335–346.
10. *D. Luce.* Individual Choice Behavior. Wiley, 1959.
11. *Гришилов А.А., Стакун В.А., Стакун А.А.* Математические методы построения прогнозов. М.: Радио и связь, 1997. 112 с.
12. *Nikolaev D.P., Gladkov A., Chernov T., Bulatov K.* Diamond recognition algorithm using two-channel x-ray radiographic separator // Proc. SPIE 9445, Seventh International Conference on Machine Vision (ICMV 2014).
13. *Objed* [Электронный ресурс]. URL: <https://github.com/usilinsergey/objed> (дата обращения: 01.01.2017).
14. *Жуковский А.Е., Тарасова Н.А., Усилин С.А., Николаев Д.П.* Синтез обучающей выборки на основе реальных данных в задачах распознавания изображений // Информационные технологии и системы (ИТиС'12): сборник трудов конференции. М., 2012. С. 377–382.
15. *Ilin D., Krivtsov V.* Creating training datasets for OCR in mobile device video stream // Proceedings 29th European Conference on Modelling and Simulation. 2015.
16. *Арлазаров В.В., Решетняк Н., Славин О.* Формирование множества графических образов символов в задачах обучения классификатора символов // Труды ИСА РАН. 2014. Т. 64. № 4. С. 73-79.

17. *Машинный ноль* – Wikipedia [Электронный ресурс]. URL: https://ru.wikipedia.org/wiki/Машинный_ноль (дата обращения: 01.09.2015).
18. *Precision and recall* – Wikipedia [Электронный ресурс]. URL: https://en.wikipedia.org/wiki/Precision_and_recall (дата обращения: 01.09.2015).
19. *Powers D.M.W.* Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation // *Journal of Machine Learning Technologies*. 2011. Vol 2. No 1. P. 37–63.

Усилин Сергей Александрович. Научный сотрудник ИСА ФИЦ ИУ РАН. Исполнительный директор ООО «Смарт Энджинс Сервис». В 2009г. окончил МФТИ. Количество печатных работ: 15. Область научных интересов: обработка изображений, распознавание образов. E-mail: usilin.sergey@gmail.com

Using greedy strategy of Viola-Jones cascade choosing for improving performance of multi-class object detection in video stream

S.A. Usilin

Abstract. This paper aims to study the problem of multi-class object detection in video stream with Viola and Jones cascades. An adaptive algorithm of choosing Viola-Jones cascade based on greedy choice strategy in N-armed bandit problem is proposed. The efficiency of the algorithm is shown on the problem of detection and recognition the logo of bank card in the video stream. The proposed algorithm can be effectively used in documents localization and identification, recognition of road scene elements, search of natural and man-made objects in the coastal zone of the Arctic, and for solving other problems of rigid object detection in a heterogeneous data flows.

Keywords: machine learning, object detection, Viola-Jones cascades, N-armed bandit problem, epsilon-greedy method, softmax method, exponential moving average.

References

1. *Viola P., Jones M.* Robust Real-time Object Detection // International Journal of Computer Vision. 2002.
2. *Jones M., Viola P.* Fast Multi-view Face Detection // Mitsubishi Electric Research Lab TR2000396. 2003. No. July.
3. *Thompson W.R.* On the likelihood that one unknown probability exceeds another in view of the evidence of two samples // Biometrika. 1933. Vol. 25, No. 3–4. P. 285–294.
4. *Robbins H.* Some Aspects of the Sequential Design of Experiments // Bulletin of the American Mathematical Society. 1952. Vol. 58, No. 5. P. 527–535.
5. *Sutton R.S., Barto A.G.* Reinforcement Learning. London: The MIT Press. 2012. 399 p.
6. *Gittins J.C.* Bandit processes and dynamic allocation indices // Journal of the Royal Statistical Society, Series B. 1979. Vol. 41, No 2. P. 148–177.
7. *Thathachar M.A.L., Sastry P.S.* A New Approach to the Design of Reinforcement Schemes for Learning Automata // IEEE Transactions on Systems, Man and Cybernetics. 1985. Vol. SMC-15, No 1. P. 168–175.
8. *C. J. C. H. Watkins.* Learning from Delayed Rewards. Ph.D. thesis. Cambridge University, 1989.
9. *Tokic M., Palm G.* Value-difference based exploration: Adaptive control between epsilon-greedy and softmax // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2011. Vol. 7006 LNAI. P. 335–346.
10. *D. Luce.* Individual Choice Behavior. Wiley, 1959.
11. *Grishilov A.A., Stakun V.A., Stakun A.A.* Matematicheskie metody postroeniya prognozov [Mathematical methods of forecasting]. Moscow: Radio and Communications, 1997. 112 p.
12. *Nikolaev D.P., Gladkov A., Chernov T., Bulatov K.* Diamond recognition algorithm using two-channel x-ray radiographic separator // Proc. SPIE 9445, Seventh International Conference on Machine Vision (ICMV 2014).
13. *Objed.* Available at: <https://github.com/usilinsergey/objed> (accessed: January 1, 2017).
14. *Hukovskij A. E., Tarasova N. A., Usilin S. A., Nikolaev D. P.* Sintez obuchayushchey vyborki na osnove realnyh dannyh v zadachah raspoznavaniya izobrazhenij [Synthesis of the training sample based on real data in image recognition problems] // Informacionnye tekhnologii i sistemy ITiS 12 sbornik trudov konferencii [Proceedings of Information Technology and Systems (ITaS'12)]. Moscow, 2012. P. 377–382
15. *Ilin D., Krivtsov V.* Creating training datasets for OCR in mobile device video stream // Proceedings 29th European Conference on Modelling and Simulation. 2015.
16. *Arlazarov V. V., Reshetnyak N., Slavin O.* Formirovanie mnozhestva graficheskikh obrazov simvolov v zadachah obucheniya klassifikatora simvolov [Forming a plurality of graphic images of characters in the character classifier learning problems] // Trudy ISA RAN [ISA RAN Proceedings]. 2014. Vol. 64. No. 4. P. 73–79.
17. *Machine epsilon* – Wikipedia. Available at: https://en.wikipedia.org/wiki/Machine_epsilon (accessed: September 1, 2015).
18. *Precision and recall* – Wikipedia. Available at: https://en.wikipedia.org/wiki/Precision_and_recall (accessed: September 1, 2015).
19. *Powers D.M.W.* Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation // Journal of Machine Learning Technologies. 2011. Vol 2. No 1. P. 37–63.

S. A. Usilin Researcher at ISA FRC CSC RAS. Executive Director at Smart Engines Service LLC. Graduated Moscow Institute of Physics and Technology (MIPT) in 2009. Number of publications: 15. Research interests: image processing, pattern recognition. E-mail: usilin.sergey@gmail.com