

Аппаратная адаптация программного метода обнаружения цели по каскаду дескрипторов Хаара

И. А. Зайцев

Санкт-Петербургский государственный электротехнический университет
«ЛЭТИ» им. В.И. Ульянова (Ленина)

ivan-za-2000@mail.ru

Аннотация. Данный доклад рассматривает обнаружение крупных сложных объектов на радиолокационном изображении, занимающих несколько элементов разрешения с известным откликом в реальном времени. В литературе описан метод программного распознавания объектов на изображении, предложенный П. Виолой и М. Джонсом. Данная работа посвящена разработке цифрового вычислительного комплекса, реализующего аппаратное обнаружение сложных целей на основе метода П. Виолы и М. Джонса.

Ключевые слова: радиолокация; радиолокационное изображение; цифровая обработка изображения; обнаружение

I. ВВЕДЕНИЕ

В 2001 году П. Виолой и М. Джонсом был предложен метод распознавания объектов на изображении посредством каскадного последовательного порогового обнаружения простых дескрипторов [1], основанных на разности суммарных яркостей на различных участках рассматриваемой области (рис. 1). Метод относился к методам с машинным обучением [4], в котором порог обнаружения каждого дескриптора формировался в процессе предобучения. Вычисление последующего дескриптора происходило только в случае решения обнаружителя предыдущего дескриптора о наличии объекта. Поскольку метод был задуман как программный, данная особенность позволяла экономить процессорное время в случае априорного отсутствия объекта в рассматриваемой области изображения.

Сущность предложенного в данной статье метода заключается в распараллеливании вычисления каскада дескрипторов, т. к. аппаратно вычислением конкретного дескриптора занимается конкретный конечный автомат. Это позволяет как ускорить вычисления, так и разорвать зависимость времени вычисления от априорного наличия или отсутствия цели в рассматриваемой области изображения.

В данной работе рассматривается как задача обнаружения целей на фоне шума, так и перспективы распознавания целей предложенным методом.

II. ОСОБЕННОСТИ ПРОГРАММНОЙ РЕАЛИЗАЦИИ

Обобщенный алгоритм представлен на рис. 1. На первом этапе изображение переводится в интегральный вид, т. е. каждый пиксел представляет собой сумму значений всех пикселов выше и левее.

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y'),$$

где ii – интегральное представление изображения, i – оригинальное изображение.

Этап перевода в интегральный вид позволяет вычислять дескрипторы с меньшим временем доступа к оперативному запоминающему устройству, по сравнению с классическим подходом к вычислению сумм интенсивностей на прямоугольной области изображения.

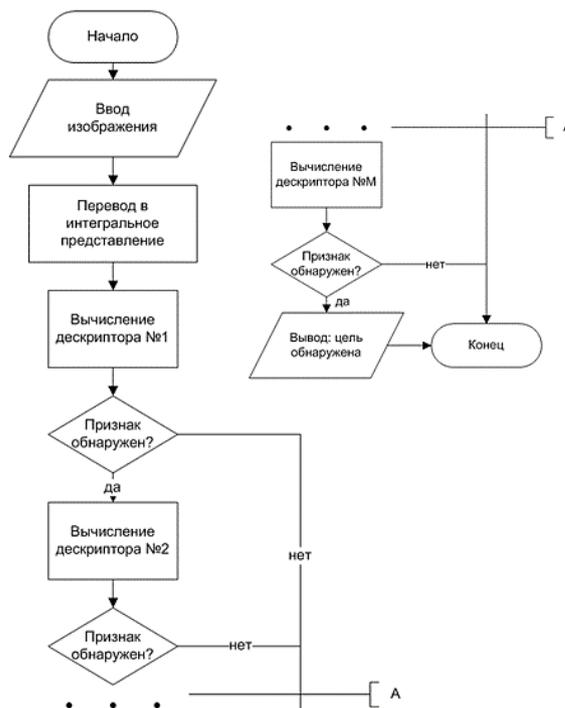


Рис. 1. Обобщенный алгоритм, предложенный П. Виолой и М. Джонсом

Затем вычисляется первый дескриптор. Результатом вычисления дескриптора служит число, которое сравнивается с порогом. Если порог обнаружения конкретного дескриптора преодолён (превышен, или наоборот не превышен в зависимости от относительного расположения оценок плотности вероятностей), то начинается вычисление последующего дескриптора, иначе – выдаётся решение об отсутствии цели. Порог выставляется как результат предобучения системы. Обучение проводится методом «AdaBoost» [3]. В случае,

если после вычисления каждого дескриптора, каждое пороговое устройство выдало решение в пользу наличия цели, вся система выдаёт решение о наличии цели в данной области.

Дескрипторы можно представить как две маски, накладываемые на изображение. На рис. 2 они обозначены как светлые и тёмные области. Вычисляется дескриптор как разность сумм яркостей всех пикселей в светлой области и сумм яркостей всех пикселей тёмной области. Из-за своей схожести с двумерными функциями Хаара, эти дескрипторы получили своё название – признаки Хаара.

Признаки Хаара вычисляются с шагом на ширину окна признака или на половину ширины окна признака [4].

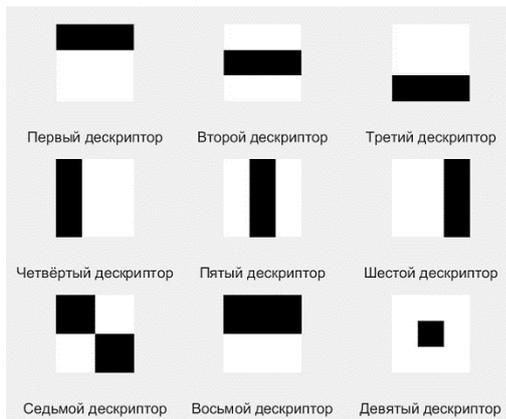


Рис. 2. Рассматриваемые в статье дескрипторы Хаара

III. АДАПТАЦИЯ ДЛЯ АППАРАТНОЙ РЕАЛИЗАЦИИ МЕТОДА

Аппаратный комплекс разрабатывался для обработки больших радиолокационных изображений с высоким разрешением в реальном времени. С АЦП приходит радиолокационный отклик, каждое зондирование служит строкой матрицы радиолокационного изображения. Смена кадра происходит, когда число строк сравнивается с числом дискретов дальности одного зондирования. Изображения потоково принимаются и записываются в оперативное запоминающее устройство. В это же время из оперативного запоминающего устройства соседнего банка происходит чтение и обработка предыдущего радиолокационного изображения таким образом, что изображение разбивается на квадратные окна, в которых и будет происходить поиск цели. Эти окна подаются на обработчик в формате телевизионной развёртки, как показано на рис. 3.

Соответственно, когда меняется разрешение радиолокационного изображения, то меняются пороги во всех пороговых устройствах. Если при этом меняется ещё и размер области поиска цели, то меняются не только пороги, но и эталоны на компараторах разрешения регистров-защёлок дескрипторов (рис. 6). Это связано со структурой работы дескрипторных вычислителей.

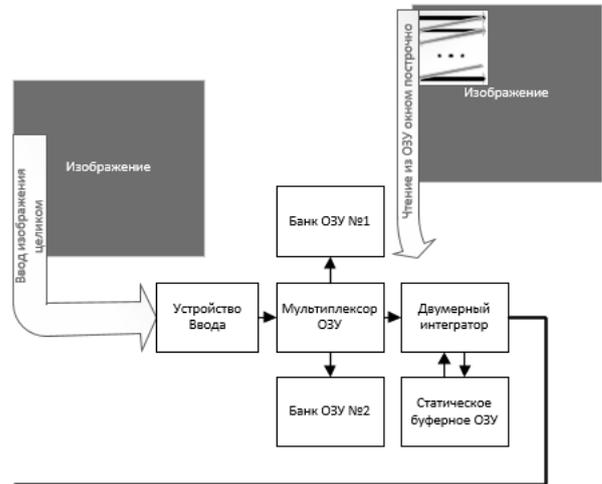


Рис. 3. Структурная схема аппаратного комплекса

Согласно рис. 3, обработка начинается с перевода изображения в интегральный вид. Двумерный интегратор представляет из себя цифровое интегрирующее звено с сумматором (рис. 4). На интегрирующее звено (вход 1) подаётся последовательно строка оригинального изображения. На нижний вход (вход 2) поступает последовательно предыдущая (уже прошедшая интегратор) строка. На выходе (выход 3) соответственно последовательно появляется проинтегрированный вид этой строки. С выхода данные идут в буферное статическое оперативное запоминающее устройство, откуда на следующем проходе строки подаются на нижний вход (вход 2) интегратора.

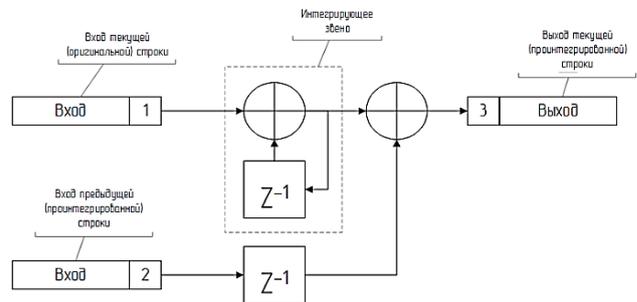


Рис. 4. Двумерный интегратор

Перевод области изображения в интегральный вид позволяет вычислить сумму значений яркости пикселей в любом прямоугольном фрагменте области.



Рис. 5. Пример областей интегрального изображения

Пример расчёта суммы значений всех пикселей в области «Г» рис. 5:

$$S_i(\Gamma) = S_i(АВВГ) - S_i(АБ) - S_i(АВ) + S_i(А) = ii(4) - ii(2) - ii(3) + ii(1)$$

где $S_i(\Gamma)$ – сумма значений элементов в области Г оригинальной матрицы, $ii(4)$ – значение элемента интегральной матрицы по координатам точки №4.

Дескрипторы (рис. 3) состоят из следующих крупных структурных элементов: устройство потокового выбора данных и арифметико-логическое устройство.

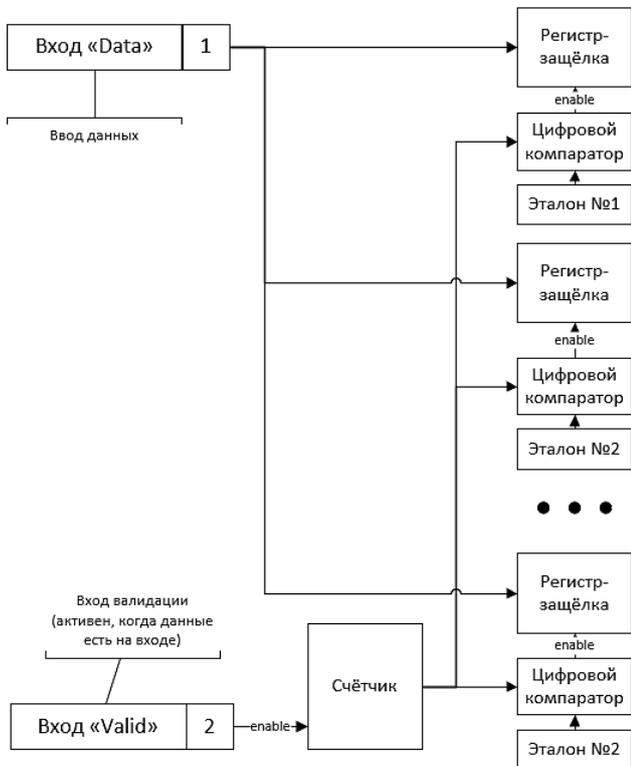


Рис. 6. Обобщенная структурная схема устройства потокового выбора данных в дескрипторных вычислителях

Устройство потокового выбора данных представлено на рис. 6. Во время поэлементного ввода пикселей изображения счётчик отсчитывает номера пикселей. Как только номер пикселя (координата в телевизионной развертке) совпадёт с эталоном, значение этого пикселя записывается в синхронный статический регистр-защёлку (рис. 6).

Арифметико-логическое устройство по окончании выбора из потока данных выполняет дескрипторный расчёт. Каждый дескрипторный вычислитель обладает арифметико-логическим устройством со своей структурой. Структурные схемы некоторых арифметико-логических устройств приведены на рис. 7.

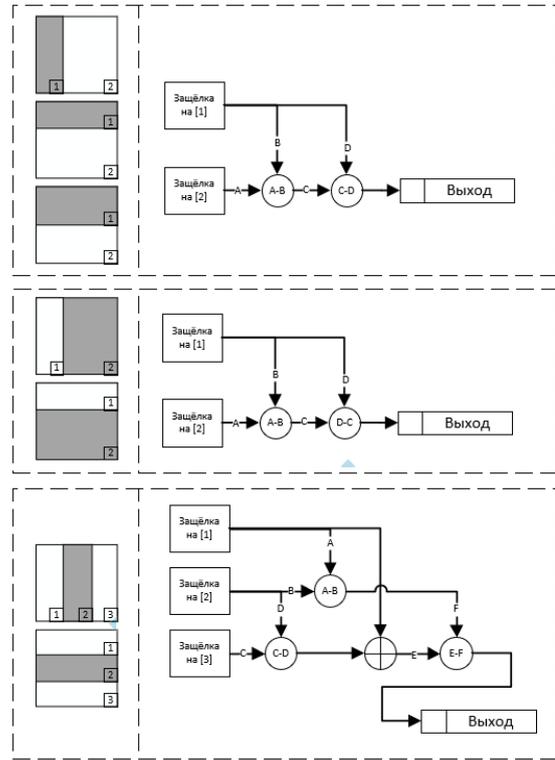


Рис. 7. Структурные схемы некоторых арифметико-логических устройств в дескрипторных вычислителях

Адаптация метода в данной статье предлагает не только структурные изменения к вычислению результата, но и несколько иной подход к построению решающего устройства. В оригинальной статье П. Виолы и М. Джонса [1] предлагался простой порог обнаружения, выбираемый в процессе обучения по модели AdaBoost.

В рамках имитационного моделирования был применён следующий подход. По обучающей выборке составлялись оценки плотностей вероятности откликов дескрипторов на область с целью. Далее выставлялись два порога, ограничивающие область, где находится 99% площади фигуры оценочной плотности вероятности. Таким образом вероятность правильного обнаружения комплексом, состоящим из девяти дескрипторов, составила:

$$D = 99\% ^ 9 = 0.99^9 \approx 91.4\%$$

Далее проводилась такая же оценка плотностей вероятностей откликов дескрипторов на белый шум и ложные (визуально похожие) цели. По данным оценкам определялись ожидаемые вероятности ложной тревоги по каждому дескриптору и для всего комплекса в целом. Данные оценки приведены в табл. 1. Для каждого дескриптора приведен процент ложных срабатываний, при обработке фрагментов, содержащих ложные цели и процент ложных срабатываний, при обработке шумового изображения. В конце таблицы рассчитан процент ложных срабатываний для всего комплекса из девяти дескрипторов.

ТАБЛИЦА I.

Номер дескриптора	Вероятность ложной тревоги по шуму	Вероятность тревоги по ложной цели
1	82%	76%
2	2%	99%
3	44%	86%
4	98%	96%
5	3%	95%
6	6%	99%
7	95%	82%
8	6%	75%
9	2%	41%
Весь комплекс	0,00015%	14,7%

На рис. 8 представлены гистограммы показаний дескрипторов по соответствующим выборкам. В качестве эталонных целей выступали размытые прямоугольные фигуры, имитирующие здания на радиолокационном изображении. В качестве ложных целей выступали линии под разным наклоном, имитирующие дороги, линии передач, трассы на радиолокационном изображении.

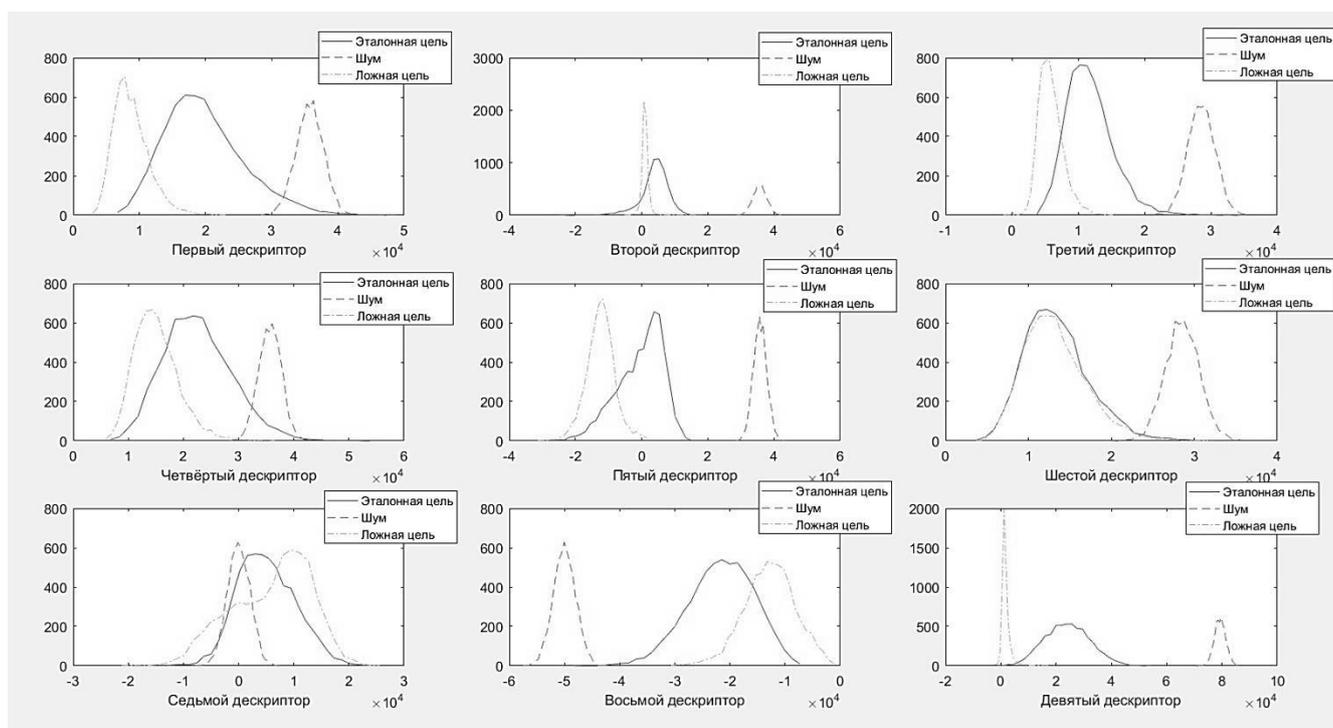


Рис. 8. Гистограммы показаний дескрипторов по соответствующим выборкам

IV. ЗАКЛЮЧЕНИЕ

Подобную оценку работы метода трудно назвать объективной, т. к. обрезанные под размер окон данные подготавливались заранее, и выборка была менее 5000 изображений, однако можно утверждать, что аппаратный комплекс теоретически способен эффективно определять высококонтрастные цели на фоне шума и так же может с определённой точностью различать ложные цели.

БЛАГОДАРНОСТЬ

Выражается благодарность Смирнову Б. И., доценту кафедры РС СПбГЭТУ «ЛЭТИ» за консультацию и помощь в подготовке материала.

СПИСОК ЛИТЕРАТУРЫ

- [1] Viola P. and Jones M.J. «Rapid Object Detection using a Boosted Cascade of Simple Features» // Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2001), 2001.
- [2] Гонсалес Р., Вудс Р. «Цифровая обработка изображений», ISBN 5-94836-028-8, изд-во: Техносфера, Москва, 2005. 1072 с.
- [3] Viola P. and Jones M.J. «Robust real-time face detection» // International Journal of Computer Vision, vol. 57, no. 2, 2004, pp.137–154.
- [4] Прокопеня А.С., Азаров И.С. «Современные методы распознавания изображений», BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня // сб. материалов V международной науч.-практ. конф. (Республика Беларусь, Минск, 13-14 марта 2019 года), с. 351-359. ISBN 978-985-543-483-3