

Метод быстрой корреляции с использованием множества шаблонов в задачах анализа изображений

Чичева М. А., Глумов Н. И., Копенков В. Н., Мясников Е. В.

mchi@smr.ru, nglu@smr.ru, vkop@smr.ru, mevg@smr.ru

Самара, Институт систем обработки изображений РАН

В настоящей работе рассматривается задача поиска и распознавания на изображениях фрагментов, соответствующих одному из множества шаблонов. Такая задача возникает, например, при распознавании текста. Другим очевидным примером является поиск характерных точек на изображениях лиц, таких как координаты зрачков, уголков губ, и т. п. Эти задачи похожи тем, что искомый образец может принимать существенно разный вид, при этом, возможно, следует определить не только его местоположение, но и меру сходства с каждым из образцов, которая впоследствии может служить критерием при принятии решения или использоваться в качестве признака.

Общепринятым подходом для решения таких задач является хорошо известный из литературы метод [1, 2], когда ищется корреляция входного изображения с каждым из шаблонов, после чего анализируются полученные корреляционные поля. Независимо от того, какой способ вычисления корреляции будет выбран (прямое вычисление, через дискретное преобразование Фурье и т. п.), такой подход требует решения ряда проблем. Первая из них — это высокая вычислительная сложность, которая естественно возрастает с ростом числа шаблонов. Вторая — поиск критерия, который позволит из множества точек корреляционных полей выбрать наиболее соответствующую истинному значению. Наконец, в связи с тем, что на реальных изображениях объекты, как правило, находятся на сложном фоне, искажены, зашумлены, необходимо уменьшить влияние мешающих факторов.

К настоящему времени существует ряд методов [3], в той или иной мере решающих изложенные проблемы. Так, использование параллельно-рекурсивной реализации при помощи аппроксимации шаблонов специальным базисом позволяет существенно снизить время обработки. Упрощение шаблонов (например, бинаризация), изменение формы учитываемой области шаблона (уход от прямоугольного «окна») позволяет как частично решить проблему быстродействия, так и уменьшить влияние мешающих факторов. Специальная обработка корреляционных полей может облегчить поиск и расчет критериев. Однако, ни один из названных методов в чистом виде не позволяет эффективно решить поставленную задачу. В докладе предлагается метод, сочетающий в себе перечисленные подходы, а также демонстрируется его применение к двум задачам анализа изображений.

Предлагаемый метод быстрой корреляции со множеством шаблонов состоит из следующих шагов.

1. *Предварительные действия* включают в себя подготовку набора шаблонов (выполняется один раз для всего набора изображений) и определение (поиск или задание) области, в которой будет производиться сравнение с шаблонами (выполняется для каждого изображения).
2. *Формирование корреляционных полей* заключается в расчете корреляции с каждым из шаблонов для всех его положений в области поиска.
3. *Обработка корреляционных полей* выполняется с целью отбора перспективных точек.
4. *Принятие решения* делается на основе анализа корреляционных полей в области выбранных точек.

Характерной особенностью метода является введение тринарных шаблонов, которые содержат только три значения: 1 — для области объекта, -1 — для области фона и 0 — для части, неучитываемой при анализе. Такой способ формирования шаблонов позволяет реализовать алгоритм быстрой корреляции, при котором вычисление свертки изображения символа с шаблоном реализуется без операций умножения, что позволяет существенно ускорить процесс распознавания символов. При этом количество отсчетов изображения, реально участвующих в формировании ответа, невелико. Наличие же неучитываемой части позволяет снизить влияние шумов, искажения толщины линий, исключить иные мешающие факторы.

Задача распознавания машиночитаемых строк на сканированных изображениях документов. Для документов, удостоверяющих личность (паспортов международного образца), Международной Ассоциацией Гражданской Авиации (ИКАО) разработан стандарт MRTD [4], который распространяется на документы, предъявляемые при путешествиях (паспорта и визы). В соответствии с ним, изображение документа личности, в частности, содержит семантическую информацию в виде машиночитаемых строк.

Подробно алгоритм распознавания машиночитаемых строк на изображении будет изложен в докладе. В частности, будут показаны сформированные шаблоны, изложен алгоритм выбора решения из списка возможных вариантов, приведены результаты экспериментального исследования.

Задача поиска положения глаз на документальных фотографиях лиц. Эта задача возникает в рамках практически всех алгоритмов распознавания лиц, первым шагом в которых выполняется геометриче-

ская нормализация с привязкой центров роговиц к заранее заданным точкам.

Предварительно формируются тринарные шаблоны для набора различных радиусов роговиц. В расчете свертки участвует небольшое количество отсчетов, причем с ростом радиуса роговицы число таких отсчетов растет несущественно. Это обеспечивает высокую скорость вычисления.

В докладе будет подробно изложен метод определения зоны поиска глаз, метод обработки корреляционных полей, а также критерий, по которому осуществляется выбор верного варианта. Особенностью решения этой задачи является учет ограничений, которые обусловлены расположением глаз, и требований к фотографиям на документы.

Экспериментальное исследование алгоритма показало, что с вероятностью 0.967 максимальное отклонение найденных координат центров роговиц от их истинных значений не превышает $0.25r$, где r — радиус роговицы.

Работа выполнена при поддержке РФФИ, проекты №06-01-00722, №07-01-96612, №07-07-97610.

Литература

- [1] Прэтт У. К. Цифровая обработка изображений. В 2 томах. Пер. с англ. — М.: Мир, 1982. — 310 с.
- [2] Гонсалес Р., Вудс Р. Цифровая обработка изображений. Пер. с англ. — М.: техносфера, 2005. — 1072 с.
- [3] Методы компьютерной обработки изображений // Под ред. Сойфера В. А. — М.: Физматлит, 2003. — 784 с.
- [4] Документ ICAO 9303 «Machine Readable Travel Documents (MRTD)» <http://www.icao.int/mrtd/publications/doc.cfm>.