

В.С. ЧИЖОВ,
Ю.А. КОВАЛЕВ, А.Д. ВАРЛАМОВ

**Разработка метода повышения
качества поиска лиц на
изображениях анализом их
биометрических признаков**

УДК 004.931

Муромский институт
(филиал) ФГБОУ ВПО
«Владимирский
государственный
университет имени
А.Г. и Н.Г. Столетовых»,
г. Муром

В статье показана актуальность повышения качества алгоритмов поиска лиц. Разработан метод, который предполагает возможную корректировку результата поиска лица после анализа его биометрических признаков на основе вероятностной модели.

Введение

В процессе эволюции вычислительной техники стали активно развиваться биометрические технологии как методы и технические средства получения и использования биометрических данных человека в целях его идентификации [4]. Это привело к разработке автоматизированных систем идентификации человека на основе биометрических данных [7].

Одной из причин повышенного внимания к биометрическим технологиям является существование обширного круга коммерческих и социальных приложений, где автоматическая идентификация человека будет воспринята весьма успешно. Так, например, идентификация человека по его лицу может применяться в системах контроля удостоверений личности (паспортов, водительских прав), информационной безопасности (доступ к ЭВМ, базам данных), наблюдения и расследования криминальных событий, в банковской сфере (банкоматах, системах удалённого управления счётом) и так далее.

Анализ проблемы и постановка задачи

Известно, что существует множество различных алгоритмов поиска и идентификации лиц. Но, как показала практика, все они не

идеальны, так как возникают ошибки при идентификации лиц, не-точности их распознавании [1,3,8]. Возникают случаи, при которых лицо не находится или иные объекты ошибочно принимаются за лицо. Это ведет к ухудшению качеств распознавательных систем, построенных на основе этих алгоритмов. Следовательно, задача повышения их качества является актуальной в настоящее время.

Общим для большинства алгоритмов распознавания является выделение ключевых точек и вычисление по ним набора признаков, по которым принимается решение о схожести с другими лицами [2,6]. Ключевыми точками могут быть уголки глаз, губ, кончик носа, центр глаза и тому подобное (рисунок 1). Признаками могут быть, например, расстояние между ключевыми точками (толщина губ, расстояние между углами глаз), отношение таких расстояний.



Рис. 1. Пример выделения ключевых точек для расчета биометрических признаков

В процессе распознавания сравниваются признаки анализируемого лица с признаками, хранящимися в базе. Корректное нахождение ключевых точек на изображении во многом определяет успех распознавания. Поэтому изображение лица человека должно быть без помех, мешающих процессу поиска ключевых точек. К таким помехам относят очки, бороды, украшения, элементы прически и макияжа. Кроме того, изображение лица должно иметь фронтальный ракурс, возможно, с небольшими отклонениями. Выражение лица должно быть нейтральным. Это связано с тем, что в большинстве методов нет модели учета таких изменений.

В работе требуется разработать и исследовать метод, который на основе статистического анализа значений биометрических при-

знаков позволит повысить качество алгоритмов поиска лиц. Если система ошибается и принимает за лицо иной объект, то значения биометрических признаков, статистически сильно отличающиеся от значений характеристик для правильной идентификации лица, могут свидетельствовать о данной ошибке алгоритма. Таким образом, задача сводится к установлению для каждого признака доверительного интервала.

Разработка метода на основе анализа биометрических признаков

Чтобы установить границы допустимых значений биометрических признаков, необходимо знать их статистические данные, то есть законы распределения значений признаков и их параметры. Для этого требуется проанализировать значения признаков, полученных из множества изображений с лицами.

Для примера рассмотрим биометрический признак расстояния между углами глаз. Выдвинем гипотезу, что его значения подчинены нормальному закону распределения. Подтвердим или опровергнем гипотезу с помощью критерия согласия Пирсона. Для этого выполним расчеты значений признака на выборке, состоящей из 100 изображений с лицами ($n=100$), примеры которых приведены на рисунке 2.



Рис. 2. Примеры изображений обучающей выборки

Полученные данные сведены в таблицу 1.

Расчет критерия согласия Пирсона

Пределы эмпирического распределения		Пределы теоретического распределения		Значение функции Лапласа		Оценка вероятности попадания в интервал	Слагаемые для критерия согласия Пирсона
Левый предел	Правый предел	Левый предел	Правый предел	Левый предел	Правый предел		
1	2	3	4	5	6	7	8
0,2995	0,3062	$-\infty$	-2,53	0	0,0056	0,0108	0,9263
0,3063	0,3126	-2,53	-2,29	0,0056	0,0108		
...							
0,4117	0,4167	1,39	1,58	0,9177	0,9429	0,0252	0,951
0,4168	0,4258	1,58	1,92	0,9429	0,9725	0,0296	0,033
0,4259	0,4348	1,92	2,26	0,9725	0,988	0,0275	0,36
0,4349	0,4669	2,26	$+\infty$	0,988	1		

В 3-ем и 4-ом столбцах таблицы показаны результаты вычисления границ интервалов, выраженные через нормированную случайную величину. Значения функции Лапласа для границ интервалов отображены в 5-м и 6-м столбцах. Оценка вероятности попадания значений в интервалы как разность значений функции Лапласа на правой и левой границе интервала приведены в столбце 7. В последнем столбце приведены оценки математических ожиданий числа наблюдений по интервалам, которые определены умножением оценки вероятности на общее число изображений в выборке $n=100$.

Основные статистические параметры, необходимые для вычисления значения критерия согласия Пирсона, указаны в таблице 2.

Таблица 2

Основные статистические параметры эмпирических данных

Параметр	Значение
Среднее значение	0,313943
Дисперсия	0,001498
СКО	0,038709

Проверяя гипотезу с помощью критерия согласия χ^2 , примем уровень значимости $\alpha=0,05$. После расчетов значение величины

$$\chi^2 = 24,932$$

При количестве степеней свободы 19 и уровне значимости 0,05 для доказательства выдвинутой гипотезы о подчинении исходных данных нормальному закону распределения необходимо, чтобы по-

лученное значение было меньше 30,14353. Так как это условие выполняется, верно и утверждение:

$$\chi^2 \leq \chi_{\alpha}^2$$

что позволяет сделать вывод о доказательстве гипотезы.

На рисунке 3 можем визуальнo наблюдать подобие эмпирического распределения значений признака нормальному закону распределения случайных величин.

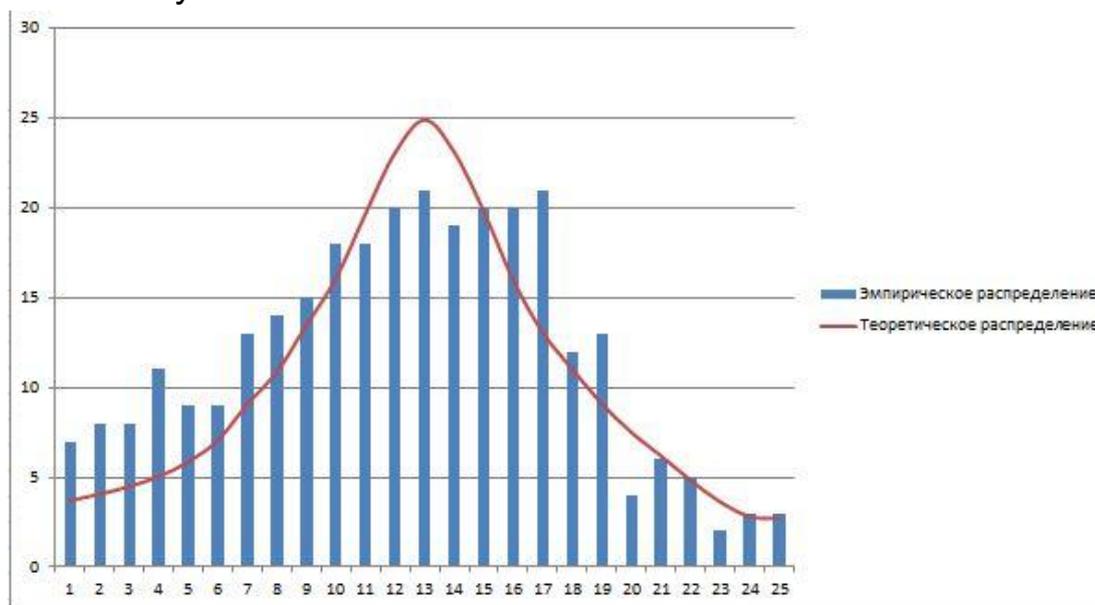


Рис. 3. Визуальное подобие значений биометрического признака нормальному закону распределения

Аналогичным образом были исследованы распределения для других биометрических признаков лица (ширина носа, расстояние между зрачками глаз, толщина губ, расстояние между углами губ) и подтверждена их принадлежность к нормальному закону распределения.

Учитывая, что в практических задачах обнаружение лиц часто сопровождается дальнейшим их распознаванием, а подавляющее большинство идентифицирующих алгоритмов в своей основе используют биометрические признаки лица, полученные результаты предлагается использовать в алгоритмах распознавания лиц, окончательно идентифицировав человека только тогда, когда значения всех его биометрических признаков войдут в доверительные интервалы.

Для получения доверительных интервалов использовалось правило трех сигм. Согласно ему границы интервала составляют

$$[s-3\sigma; s+3\sigma],$$

где s - среднее значение признака; σ - среднеквадратическое отклонение. Если один из биометрических признаков не входит в доверительный интервал, то система делает вывод, что лицо было найдено ошибочно и отменяет положительный результат поиска лица.

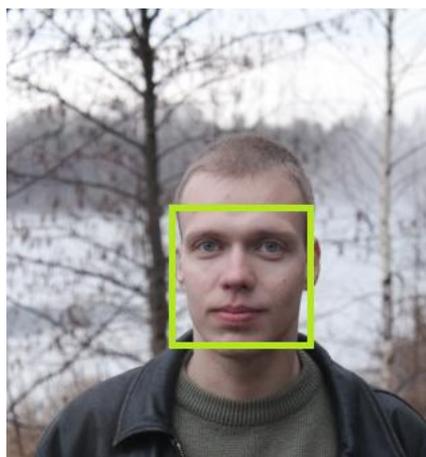
В соответствии с методом поиск лица реализуется автоматическим выполнением системой следующих этапов:

1. Выполнение алгоритма поиска лица на изображении.
2. Если лицо найдено, то рассчитываются значения его биометрических признаков.
3. Если значение хотя бы одного биометрического признака выходит за границы доверительного интервала, аннулируются результаты поиска лица - лицо считается не найденным. Границы доверительных интервалов устанавливаются предварительным анализом коллекции изображений с лицами.

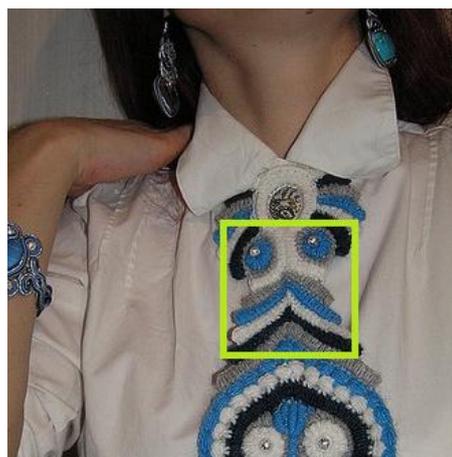
На первом этапе выполняется алгоритм поиска лица, который часто предшествует алгоритму его идентификации. Второй этап (вычисление признаков) в классических схемах используется как составная часть алгоритма распознавания, поэтому для распознающих систем метод не приведет к значительному увеличению вычислительной сложности. На третьем этапе по конечному набору полученных значений биометрических данных подтверждается или опровергается наличие лица на изображении. Фактически третий этап представляет собой классическую задачу распознавания образов с разделением их на 2 класса - верно найденное лицо и ложное нахождение лица.

Исследование разработанного метода

Для исследования мы использовали алгоритмы библиотеки компьютерного зрения с открытым исходным кодом OpenCV. Первоначально было взято 2 изображения, на одном из которых лицо присутствует (рисунок 4а), а на другом отсутствует (рисунок 4б), но система указала наличие лица на обоих изображениях.



а) изображение с лицом



б) изображение без лица

Рис. 4. результаты поиска лица

Каждый признак выделенной области лица вычислялся следующим образом: брались координаты двух определенных точек среди базовых (например, верхний и нижний края губ на центральной линии, центры зрачков и так далее); вычислялось эвклидово расстояние между ними; найденное расстояние нормировалось к размеру области лица.

Таблица 3

Проверка вхождений значений признаков в доверительные интервалы

Признак	Среднее значение, s	СКО, σ	Доверительный интервал, $[s-3\sigma; s+3\sigma]$	Значение признака изображения	
				с лицом	без лица
Расстояние между углами глаз	0,1981	0,0175	[0,1456; 0,2506]	0,1862	0,2273
Расстояние между зрачками	0,3741	0,0268	[0,2937; 0,4545]	0,3458	0,4663
Расстояние между углами губ	0,3139	0,0387	[0,1978; 0,4300]	0,2949	0,2988
Толщина губ	0,0966	0,0271	[0,0153; 0,1719]	0,0775	0,0671
Ширина носа	0,2278	0,0217	[0,1627; 0,2929]	0,2091	0,1578

В таблице 3 показано, что разработанный метод позволяет исправить сделанную алгоритмом ошибку. Можно наблюдать, что все значения биометрических признаков изображения с лицом попали в доверительные интервалы, а на изображении без лица в довери-

тельные интервалы попали только три биометрических признака из пяти.

Дальнейшее исследование проведем на коллекции изображений, содержащей 500 экземпляров. Все фотографии после алгоритмического сканирования лиц разделены на 4 группы по двум критериям: наличие лица (по факту) и обнаружение лица (по "мнению" алгоритма). Результаты распределения по группам сведены в таблицу 4.

Таблица 4

Результаты исследования метода на множестве изображений

Без использования метода			С использованием метода		
	Есть лицо	Нет лица		Есть лицо	Нет лица
Лицо найдено	47,6%	2,8%	Лицо найдено	47,6%	2,2%
Лицо не найдено	5,2%	44,4%	Лицо не найдено	5,2%	45%

Полученные данные позволили рассчитать значения основных критериев качества [5], применяемых к поисковым алгоритмам.

Таблица 5

Исследование метода по критериям качества

Критерий	Без использования метода	С использованием метода
Полнота поиска	90,2	90,2
Точность поиска	94,4	95,6
Ошибка поиска	8,0%	7,4%

Полученные значения неопровержимо свидетельствуют, что разработанный метод повышает качество распознавания лиц.

Заключение

Таким образом, в статье разработан метод повышения качества поиска лиц на изображениях анализом значений их биометрических признаков. Был подробно описан данный метод, приведены результаты его применения к алгоритмам, подтверждено улучшение качества распознавания лиц, а также была доказана гипотеза, что значения основных биометрических признаков лица подчинены нормальному закону распределения.

Литература

1. *Alexey Varlamov, Ruslan Sharapov*. Machine Learning of Visually Similar Images Search // Proceedings of the 14th All-Russian Scientific Conference "Digital libraries: Advanced Methods and Technologies, Digital Collections", vol. 934, Pereslavl-Zalessky, Russia, October 15-18, 2012, pp. 113-120.
2. *Li H., Toh K.-A., Li L.* Advanced Topics in Biometrics, World Scientific, 2012. — 500 pages.
3. *Yang J., Xie S.J.* New Trends and Developments in Biometrics, InTech – 2012, 342 pages.
4. *Болл Р.М.* Руководство по биометрии. Изд-во Техносфера, 2007. – 367 с.
5. *Варламов А.Д.* Основные метрики, оценивающие качество работы систем поиска изображений // Алгоритмы, методы и системы обработки данных. Выпуск 2(24), 2013, с. 3-12.
6. *Ивантер Э.В., Коросов А.В.* Элементарная биометрия: учеб. пособие — Петрозаводск: Изд-во ПетрГУ, 2010. — 104 с.
7. *Кухарев Г.А.* Биометрические системы: Методы и средства идентификации личности человека. – СПб.: Политехника, 2001. – 240 с.
8. *Шарапов Р.В., Варламов А.Д.* Сравнительный анализ систем поиска графических данных // Современные наукоемкие технологии. 2013. № 1. С. 27-31.