

ИСПОЛЬЗОВАНИЕ ГИСТОГРАММ ЯРКОСТИ ПИКСЕЛЕЙ В ПРОЦЕССЕ РАСПОЗНАВАНИЯ

USE OF HISTOGRAMS OF BRIGHTNESS OF PIXELS IN THE COURSE OF RECOGNITION

Аимбетова Д.Т., Муслимова А.З.

D.T.Aimbetova, A.Z.Muslimova

Костанайский государственный университет имени А.Байтурсынова, Костанай, Казахстан
Костанайский государственный университет имени А.Байтурсынова, Костанай, Казахстан

Kostanay state university of A.Baytursynov, Kostanay, Kazakhstan

Kostanay state university of A.Baytursynov, Kostanay, Kazakhstan

Аннотация

Предметом исследования данной статьи являлось применение формы гистограмм распределения яркости пикселей в качестве вектора признака для классификации изображений с лицами. В рамках задачи распознавания считается, что каждому образу ставится в соответствие единственное значение вектора признаков. Векторы образов содержат всю поддающуюся кодированию информацию об образе и рассматриваются в качестве точек n – мерного евклидова пространства. Классификация образов проводилась по критерию минимума расстояния между векторами. В статье также описывалось использование гистограмм яркости пикселей в стабилизации яркости изображения. Приведены результаты тестирования метода путем различных модификаций над вычислением гистограммы, результаты сравнения экспериментов распознавания для полной и составной гистограмм признаков. Полученные результаты подтверждают гипотезу использования яркостных гистограмм в качестве исходного вектора признаков. В свою очередь, сравнение двух схожих изображений, в которых имеются небольшие изменения в отдельных незначительных элементах сцены, может быть основано на применении гистограмм яркости, в которых эффективно подобран параметр уровней дискретизаций. В завершении статьи прилагается описание недостатков и преимуществ данного метода. Тестирование выполнялось над базой изображений лиц Olivetti Research Laboratory.

Abstract

Object of research of this article was application of the form of histograms of distribution of brightness of pixels as a sign vector for classification of images with persons. Within the task of recognition it is considered that to every image the unique value of a vector of signs is put in compliance. Vectors of images contain all information on an image striking to coding and are considered as n points – a measured Euclidean space. Classification of images was carried out by criterion of a minimum of distance between vectors. In article use of histograms of brightness of pixels in image brightness stabilizing was also described. Results of testing of a method by different modifications over computation of the histogram, results of comparing of experiments of recognition for complete and composite histograms of signs are given. Analyzing the received results confirm a hypothesis of use of brightness histograms as the initial vector of signs. In turn, comparing of two similar images in which there are little changes in separate insignificant elements of a scene can be based on application of the histogram of brightness in which the parameter of levels of samplings is picked effectively up. In completion of article the description of shortcomings and advantages of this method is attached. Testing was executed over a basis of images of the persons Olivetti Research Laboratory.

Ключевые слова: идентификация, признаки образа, процесс распознавания, гистограмма распределения яркости, вычислительные технологии.

Keywords: identification, signs of an image, recognition process, histogram of distribution of brightness, computing technologies.

Введение

На сегодня интеллектуализация методов обработки и анализа данных являются главными ориентирами вычислительных технологий «четвертого поколения». Всепроникающая компьютеризация обеспечивает разработку новых методов нулевого взаимодействия с компьютером, моделирующий интеллект человека. Главной задачей интерфейсов нового поколения является способность идентифицирования объектов. А одной из первых задач, послуживших развитию теории распознавания образов, была задача распознавания лица, где в качестве источника физического признака выступает лицо человека.

Концепция теории распознавания образов лежит в основе современных информационных систем, реализованных путём применения новейших компьютерных технологий [1]. Наиболее распространены системы распознавания, анализирующие зрительную информацию с физических объектов [2]. Нынешний уровень развития вычислительных технологий позволяет сочетать как подходы к описанию образов в системах распознавания, так и методы, участвующие в процессе распознавания. Главной целью разработчиков, работающих в этом направлении, является создание алгоритма системы распознавания для определения личности человека [3].

Как было доказано, еще в 1964 и 1965 Вудро Уилсон Бледсоу с Хелен Чан и Чарльзом Биссоном, что на процесс распознавания лиц огромное влияние оказывают видоизменения в освещении, ракурсе и мимике лица, а также биологическое старение. Поиски решения этого вопроса продлеваются путем измерения субъективных черт лица, таких как, расстояние между глазами или положение и ширина носа и т.п. Измерения во многих задачах машинного зрения проводятся над яркостью изображения, резкие перепады которой часто соответствуют чертам лица – границам рта, глаз, бровей и носа.

В областях обработки изображений и машинного зрения немаловажную роль играют гистограммы распределения уровней яркостей цифрового изображения. Изображение представляет собой двумерную функцию $f(x, y)$, где x, y – координаты на плоскости, f – амплитуда в любой точке с парой координат (x, y) называется *интенсивностью* или *яркостью* цвета изображения в этой точке. Таким образом, если координаты x, y и величина амплитуды f принимают значения из дискретного множества, то говорят о *цифровом изображении* [4 – 18 с].

В математической статистике гистограмму рассматривают как функцию, приближающую плотность вероятности некоторого распределения (*например, распределения пикселей определенной яркости*), построенной на основе выборки из него. Гистограмма цифрового изображения представляет собой график, на котором указано число пикселей на каждом уровне интенсивности цвета, где интенсивность – это уровень концентрации цвета, то есть преобладание того или иного тона. Тональный диапазон изображений содержит определенное число пикселей во всех областях с глубиной цвета в 256 градаций. В цифровом мире большинство устройств оперируют 8-битовыми изображениями, следовательно, кодируется 256 различных состояний.

Использование гистограмм яркости пикселей в стабилизации яркости изображения

На процесс распознавания огромное влияние оказывает освещенность. При недостаточной освещенности мы сталкиваемся с узким смещенным диапазоном яркости пикселей. Это часто встречается в темных изображениях и в изображениях, имеющих пересвет яркости пикселей. Большинство пикселей сконцентрированы в какой-то одной области, вместо того, чтобы занимать весь диапазон значений яркости от 0 до 255. К такому изображению применяется преобразование яркостей, компенсирующее нежелательный эффект. Другая причина освещенности изображения является более проблемной, когда изображение состоит из пикселей самых темных тонов, и самых светлых тонов, но тем не менее большинство пикселей концентрируется вокруг определенной яркости.

Исходя из определения гистограммы яркости пикселей, возможно численно оценить неравномерность освещения. И если в изображении имеется неправильный контраст, то можно применить к нему глобальное преобразование яркости – тональную коррекцию.

Преобразование яркости изображения описывается следующей функцией: $f^{-1}(y) = x$, где y – яркость пикселя на исходном изображении, а x – яркость пикселя после корреляции. Первую причину яркости изображения можно осилить с помощью линейной коррекции, где компенсацией узкого диапазона яркостей будет *линейное растяжение гистограммы*. Линейная коррекция – это преобразование, которое переводит самые темные пиксели в черные, а самые светлые в белые. Изображение становится более контрастным

при «раздвиге» гистограммы. Но к изображению, в котором большая доля пикселей очень темная/светлая, метод линейной коррекции не подходит, так как минимальное значение пикселя в изображении будет равно нулю, а максимальное равно 255. После применения линейной коррекции мы получим тоже, самое изображение. В таком случае необходимо использовать нелинейную коррекцию значений яркости: гамма – коррекция, логарифмическая коррекция.

Однако существуют классы изображений, для которых линейное или нелинейное преобразование яркостных диапазонов не всегда эффективно. В первую очередь это касается изображений, градации яркостей которого занимают максимально возможный диапазон, а яркостной диапазон потенциально информативных участков изображения - небольшую часть диапазона. Для таких изображений рекомендуется применять метод кусочно-линейных или кусочно-нелинейных преобразований яркостных диапазонов. [5]

Вычисление гистограммы распределения значений яркости пикселей

В нахождении подобия между изображениями важным является *форма гистограммы*. Особенность формы заключается в том, что если исходное изображение будет повернуто на плоскости на любой угол или будет масштабировано по любой из осей, то форма остается одинаковой. Количество уровней дискретизации определяется параметром BIN . Параметр BIN разбивает гистограмму на x – интервалов, после в каждом j – ом элементе которой ведется подсчет количества пикселей, принадлежащих заданному интервалу с определенной яркостью. Элемент j гистограммы $H(j)$ строится из суммы количества пикселей, имеющих соответствующую яркость со значениями $j = 0, 1, \dots, 255$.

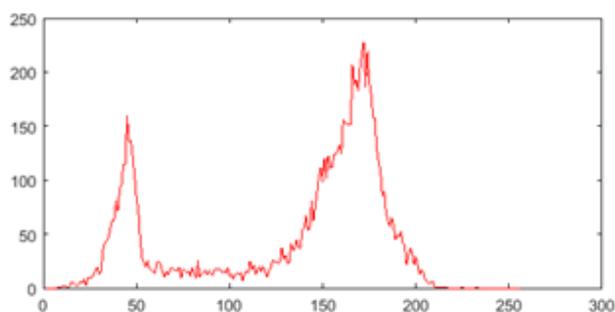
Вычисление яркостных признаков производится следующим образом:

$$H(x) = \sum_{j=(x-1)\frac{256}{BIN}}^{(x\frac{256}{BIN})-1} H(j), \quad x = 1, 2, \dots, BIN. \quad (1)$$

На рисунке 1 изображен диапазон яркости пикселей, которые составляют изображение. Параметр BIN , как упоминалось ранее, предназначен для деления распределенной информации на j столбцов. Высота столбца характеризуется количеством пикселей, попавших в соответствующий интервал. В первом пике функции сконцентрированы основные пиксели, формирующие фон изображения. Ширина тонального диапазона также зависит от однородности фона. В максимальном пике, сформированном в конце гистограммы, концентрируются пиксели относящиеся к объекту исследования (лицо). Основная яркость которого фиксируется в точке ограниченного максимума. Высота функции распределения яркости зависит от однородности интенсивности цвета объекта. Малая доля пикселей, рассеянная в середине гистограммы, образуются наличием шума.



а)



б)

Рисунок 1 – а) исходное изображение лица, б) вычисленная гистограмма распределения яркости

На рисунка 2 представлена идентификация тестового образа №10 первого класса тестовой базы Olivetti Research Laboratory (ORL). Процедура идентификации образа №10 опирается на нахождении минимума расстояния между гистограммами эталонов всех классов, которые представлены в виде векторов - столбцов. В результате каждому классу изображений лиц сопоставляется некоторое множество векторов - столбцов образов. Признаковое пространство при этом разбивается на области, соответствующие классам, которые называют *кластерами* [6]. В результате всех этапов кластерного анализа создаются кластеры «похожих» образов. В качестве функции расстояния векторов образов, интерпретируемых как точки в евклидовом пространстве, можно использовать следующую евклидову метрику [7]:

$$d(x,y) = \|x-y\| = \|y-x\| = [(x_1-y_1)^2 + \dots + (x_n-y_n)^2]^{1/2} \quad (2)$$

В системе Matlab данная метрику можно описать следующим образом:

$$DIST(j) = \text{sum}(\text{abs}(HQF - \text{BASE}(:,j))),$$

где *HQF* является гистограммой распределения яркости тестового образа, *BASE* – массив, состоящий из гистограмм эталонов каждого класса.

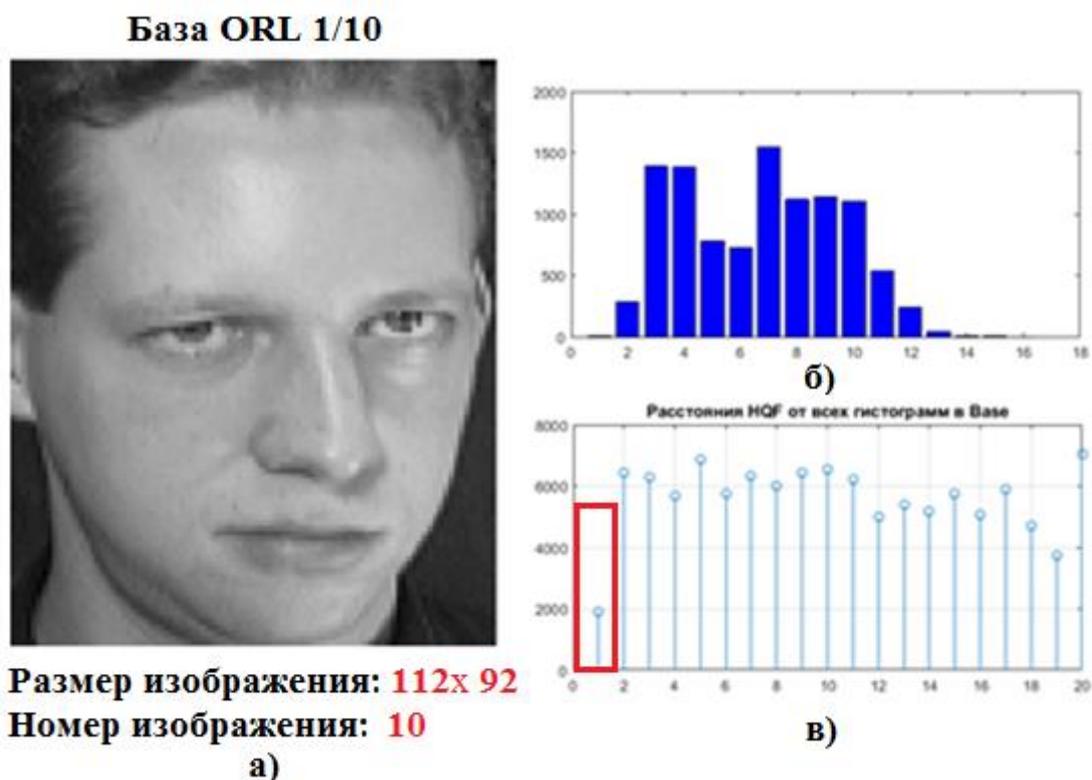


Рисунок 2 – а) изображение №10, 1 класса, б) гистограмма эталонного изображения №1 1 класса; в) расстояние тестового образа (изображение №10, 1 класс) от образов, хранящихся в базе эталонов всех классов

Тестирование метода распознавания изображений путем сопоставления форм гистограммы распределения значений яркости пикселей

Тестирование проводилось на базе изображений лиц ORL, параметры которой были приведены в статье [8].

В итоге тестирования наивысший результат распознавания (RR) равен 64%, при заданном количестве эталона $L = 1$ в каждом классе, с уровнем дискретизации $BIN = 32$;

наивысший RR равен 72%, при заданном $L = 3$ с параметром $BIN = 128$; при заданном $L = 5$ наивысший результат 92,5%. Распознавания дал тест с параметром $BIN = 256$.

Можно сказать, что полученные вектора признаков отличаются от первичных (смотрите табл.1) в следствии различной дискретизации распределения пикселей, составляющих изображение. Соответственно с этим меняется и значение минимального расстояния между векторами, а результат распознавания подвергается изменению. В заданиях распознавания изображений с лицами значения BIN выбирается в пределах от 8 до 64 [9 – 206 с.].

Таблица 1. Зависимость результата распознавания для 40 классов от количества эталонов и уровней дискретизации гистограммы

Количество эталонов в классе	Количество эталонных/тестовых изображений в базе	BIN	Recognition result, %
1	40/360	16	58,3333 %
		32	64,1667 %
		64	60,8333 %
		128	60,8333 %
		256	61,1111 %
3	120/280	16	69,6429 %
		32	71,0714 %
		64	71,0714 %
		128	72,1429 %
		256	71,4286 %
5	200/200	16	91 %
		32	90,5 %
		64	91,5 %
		128	92 %
		256	92,5 %

Еще одной особенностью распознавания форм гистограмм изображений является то, что гистограммы двух семантически разных изображений могут быть подобными. Метод основан на представлении исходных изображений как совокупности строк и столбцов. Метод не требует предварительного уменьшения размеров исходных изображений, не является итерационным и прямо реализуется по двум направлениям координат – а именно по строкам и столбцам исходного изображения. [10]

Для убеждения в том, что действительно на изображении представлено лицо, мы можем вычислить вектор признаков для полной гистограммы всего изображения ($M \times N$) и сравнить полученный вектор с вектором признаков для составной гистограммы ($M/2 \times N$, $M/2+1 \times N$). Значение минимального расстояния двух гистограмм (составной и полной) совпадет, то означает, что на изображениях представлено одно и то же лицо. Таким образом, можно предположить, что метод сравнения формы полученных гистограмм яркостей обладает семантической сегментацией.

Гистограммы верхней и нижней половин лица по горизонтали вычисляется методом разбиения строк пополам (смотрите рис.3). Далее вычисляются гистограммы значений яркости пикселей для каждой половины, после чего полученные гистограммы соединяют друг с другом в порядке следования половин и получают полную гистограмму исходного изображения.



Рисунок 3 – Формы двух гистограмм верхней и нижней части изображения лица

Тестирования для полной и составной гистограмм показал лучшее качество распознавания. В таблице 2 представлены следующие данные: значение BIN, результаты распознавания для полного и составного методов, процентное улучшение результата распознавания с применением метода вычисления гистограммы верхней и нижней части изображения лица по отношению к результату распознавания форм полной гистограммы. Тест проводился с базой ORL для 40 классов, состоящих из 10 изображений, которые разделялись случайным выбором на 5 тестовых и 5 эталонных образов.

Таблица 2. Сравнение результатов распознавания для полной и составной гистограмм признаков

BIN	Результаты распознавания, %		Улучшение распознавания, %
	Полная	Составная	
16	90 %	96 %	6 %
32	91 %	98 %	7 %
64	92 %	97 %	5 %
128	92 %	97 %	5 %
256	93 %	97 %	4 %

По данным улучшения процента распознавания, представленных в третьей колонке таблицы 2, метод выделения признаков на базе построения составных гистограмм имеет лучшие показатели качества распознавания.

Вывод

Гистограммы распределения уровней яркости являются преобразованием изображения в многомерный вектор признаков, с помощью которого возможно сопоставить изображения на основе функций расстояний. Можно сказать, что принадлежность вектора

тестового образа к конкретному классу определяется тем, что этот вектор находится ближе к векторам образов этого класса.

Таким образом, формы гистограмм яркости пикселей являются хорошими признаками для сопоставления изображений, так как обладают устойчивостью к небольшим деформациям, таким как поворот и масштабирование изображения. Преимущество метода распознавания с помощью гистограмм заключается в простоте вычислений. Недостаток метода: если два структурно или текстурно – идентичных изображения будут иметь различную яркость, то формы данных изображений будут разными, так как гистограмма вычисляет распределение яркостей пикселей. Поэтому возникают искажения циклического сдвига и дополнительные искажения на границах гистограмм. Этот факт исключает возможность использования данного метода в процессах распознавания в условиях неконтролируемого освещения.

Список использованных источников:

1. Жумагалиева А.Ж. Построение математической модели распознавания образов. Алматинская Академия Экономики и Статистики. Вып. 1 (56), 2015 г.
2. Денисенко О.А. Математическая теория распознавания образов: курс.работа, 2013г.
3. Компьютерное зрение для развития биометрических технологий. Агентство инноваций и развития экономических и социальных проектов". Ссылка: <https://www.innoros.ru/publications/foreign-innovations/16/kompyuternoe-zrenie-dlya-razvitiya-biometricheskikh-tekhnologii>
4. Гонсалес Р., Вудс Р., Эддинс С.. Цифровая обработка изображений в среде Matlab. Изд: Техносфера, Москва, 2006 г., 621 с.
5. Обработка сигналов и изображений\Image processing toolbox. Ссылка: <http://matlab.exponenta.ru/imageprocess/book3/10/imadjust.php>
6. Людмила Николаевна Чабан. Теория и алгоритмы распознавания образов. Учебное пособие. Москва, 2004 г., 70 с.
7. Гонсалес Р., Дж.Ту. Принципы распознавания образов. Изд. «Мир», Москва, 1978 г., 414 с.
8. Аимбетова Д.Т, Муслимова А.З., Жарлыкасов Б.Ж. Распознавание изображений лиц для идентификации личности// Актуальные научные исследования в современном мире. iScience. Вып. 12, Декабрь 2017, с.164 -168
9. Кухарев Г.А., Е.И. Каменская, Ю.Н. Матвеев., Н.Л. Щеголева. Методы обработки и распознавание изображений лиц в задачах биометрии. Изд: Политехника, 2013 г., 388 с.
10. Кухарев Г.А., Н.Л. Щеголева. Методы представления и сравнения семантически разных классов изображений//Анализ данных и интеллектуальные системы. Бизнес – информатика №4(26) -2013 г.

Сведения об авторах

Аимбетова Диана Талгатовна – магистрант, Костанайский государственный университет имени А. Байтурсынова, г. Костанай, Казахстан. E-mail: aimbetova-di94@mail.ru .

Муслимова Агима Зайнагатдиновна – кандидат педагогических наук, доцент, заведующая кафедрой информатики Костанайского государственного университета имени А. Байтурсынова, г.Костанай. E-mail: muslimova_agima@mail.ru .

Diana Talgatovna Aimbetova – Master’s student, Kostanay state university of A.Baytursynov, Kostanay, Kazakhstan. E-mail: aimbetova-di94@mail.ru .

Agima Zainagatdinovna Muslimova – The candidate of pedagogical sciences, the associate professor, head of the department of informatics of the Kostanay state university of A. Baytursynov, Kostanay. E-mail: muslimova_agima@mail.ru .