

РАСПОЗНАВАНИЕ ЛИЦ ПО ФОТОГРАФИЯМ НА ОСНОВЕ ИНВАРИАНТНЫХ МОМЕНТОВ

¹Хачумов М.В., ²Нгуен Зуй Тхань

¹Федеральное государственное учреждение «Федеральный исследовательский центр «Информатика и управление» Российской академии наук», Москва, Россия (117312 Россия, г. Москва, пр-т 60-летия Октября, 9), e-mail: khmike@inbox.ru

²Российский университет дружбы народов, Москва, Россия, (117198 Россия, г. Москва, ул. Миклухо-Маклая, д.6)

Настоящая статья посвящена анализу и разработке методов распознавания образов, которые могут быть применены в системе биометрической идентификации человека по изображению его лица. Дана математическая постановка задач кластеризации и распознавания лиц. Разработан комбинированный метод решения задачи распознавания лиц по фотографиям на основе инвариантных моментов полутоновых изображений в комплексе с метрикой Евклида-Махаланобиса и вероятностной искусственной нейронной сетью. Особенностью условий решения задачи является работа с фотографиями, подверженными различным искажениям (гримасы, мимика, эмоции) и шумам. Выполнена программная реализация и проведены экспериментальные исследования, подтверждающие перспективность предложенного подхода. Указаны достоинства и недостатки различных типов классификаторов. Даны предложения по дальнейшему развитию исследований в направлении использования 3D моделей лиц.

Ключевые слова: кластеризация, распознавание лиц, метрика, инвариантные моменты, искусственные нейронные сети.

FACE RECOGNITION IN PHOTOES BASED ON THE METHOD OF INVARIANT MOMENTS

¹ Khachumov M.V., ² Nguyen Duy Thanh

¹ Federal State Institution 'Federal Research Centre of "Informatics and Control" of the Russian Academy of Sciences', Moscow, Russia, (117312 Russia, Moscow, 60-letiya Oktyabrya, 9), e-mail: khmike@inbox.ru

² People's Friendship University of Russia, Moscow, Russia (117198 Russia, Moscow, Miklukho-Maklaya str. 6)

This article is dedicated to the analysis and development of pattern recognition methods that can be applied in the system of biometric identification of a person according to the image of his face. A mathematical formulation of the tasks of clustering and facial recognition is given. The combined method for solving the problem of face recognition in photos on the basis of the invariant moments of grayscale images in complex with the Euclidean-Mahalanobis metric and probabilistic neural network is developed. Special conditions for solving the problem is work with photos subject to various distortions (grimaces, facial expressions, emotions) and to noise. Software implementation is realized and experimental studies confirming the prospects of the proposed approach are executed. Advantages and disadvantages of different types of classifiers are specified. Proposals for the further development of research towards the use of 3D models of faces are given.

Keywords: clustering, face recognition, metrics, invariant moments, artificial neural network.

Среди современных работ, посвященных распознаванию лиц, отметим статью [6], в которой построена теоретико-вероятностная модель полутонового изображения и применен метод идентификации личности по фотографии лица на основе оптимального байесовского правила. В другой работе [5] решалась задача автоматического распознавания на основе принципа минимума информационного рассогласования. В работе [4] предложен оригинальный алгоритм распознавания лиц в режиме реального времени. В задачах идентификации лиц часто прибегают к использованию инвариантных моментов в качестве признаков. Так, в работе [8] применяется вектор особенностей, состоящий из 11 различных

моментов, а в работе [1] исследуются свойства инвариантных моментов Ну. Показано, что инварианты имеют разную чувствительность к изменениям входных данных. Отметим капитальный в этой области труд [10], в котором систематизируются знания о применении инвариантных моментов. Укажем также на интересную работу [3], посвященную обработке трехмерных изображений, однако задача 3D-распознавания лиц еще недостаточно исследована. Разумеется, здесь нельзя не отметить подходы, связанные с выделением таких особенностей лица, как губы, нос, овал или профиль лица, но ввиду множества факторов, усложняющих анализ изображений (зашумления, повороты лица, выражения различных эмоций и др.) пока нет подхода, гарантирующего точное решение проблемы.

В настоящей работе предлагается комбинированный подход, в котором сочетаются следующие инструментальные средства: методы выделения инвариантных моментов Ну; методы формирования эталонных классов лиц; метрика Евклида-Махаланобиса [2] и аппарат искусственных нейронных сетей (ИНС).

Постановка задач кластеризации и распознавания лиц

Целью кластеризации является объединение изображений лиц в непересекающиеся классы. Каждое лицо, представляется множеством выделенных из фотографий особенностей и является отдельной точкой в пространстве признаков. Задача кластеризации формулируется следующим образом [7]: разбить множество точек $X \subset R^p$, $|X| = n$ на k

классов C_1, C_2, \dots, C_k , $C_i \cap C_j = \emptyset \quad \forall i, j, i \neq j$, $X = \bigcup_{i=1}^k C_i$ так, чтобы $H = \sum_{i=1}^k \sum_{x \in C_i} d(a_i, x) \rightarrow \min$.

Здесь $a_i \in R^p$ – ядро класса C_i , а $d(a_i, x)$ – расстояние между a_i и любой точкой $x \in X$.

Задача распознавания лиц формулируется следующим образом. Требуется построить распознающую функцию $F(\omega)$, $F(\omega) = (F_1(\omega), F_2(\omega), \dots, F_k(\omega))$, выход которой определяет класс произвольного изображения ω , представленного вектором признаков $(x_1(\omega), \dots, x_n(\omega))$.

$$F_k(\omega) = \begin{cases} 1, & \text{если } \omega \in \Omega_k \\ 0, & \text{если } \omega \notin \Omega_k \\ \Delta, & \text{если неизвестно } \omega \in \Omega_k \text{ или } \omega \notin \Omega_k \end{cases}$$

Для решения задачи используются методы дискриминантного анализа, ИНС, деревья решений, байесовский классификатор и другие методы. Узловым моментом является выбор метрики, от которой существенным образом зависит результат классификации.

Получение инвариантов как признаков для распознавания лиц

Инвариантные моменты используются как характерные признаки, извлекаемые из снимков. Для идеальных изображений их значения не зависят от позиции, масштаба и ориентации на плоскости графического объекта. На практике лица на фотографиях

зашумлены и подвергаются различным деформациям, свойственным мимике человека. В таких условиях имеем «псевдо-инварианты». Для создания классов необходимо иметь представительный набор фотографий каждого человека. Центральные моменты цифрового изображения лица определяются следующим образом $m_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y)$, где \bar{x}, \bar{y} – центр тяжести, моменты m_{pq} – центральные моменты порядка не выше чем $(p + q) = 3$, $f(x, y)$ – функция яркости, определяющая исходное растровое изображение. Целесообразно выполнить перевод цветных изображений в полутоновый вид. После предварительной обработки и нормализации образец представляет собой матрицу пикселей, каждый из которых имеет значение яркости в диапазоне $[0, 1]$. Моменты H_u , инвариантные к повороту и сдвигу представлены в таблице 1.

Таблица 1

Инвариантные моменты и их чувствительность к флуктуациям

| Инвариантный момент | Чувствительность |
|---|------------------|
| $M_1 = m_{20} + m_{02}$ | δ^2 |
| $M_2 = (m_{20} - m_{02})^2 + 4m_{11}^2$ | δ^4 |
| $M_3 = (m_{30} - 3m_{12})^2 + (3m_{21} + m_{03})^2$ | δ^6 |
| $M_4 = (m_{30} + m_{12})^2 + (m_{21} + m_{03})^2$ | δ^6 |
| $M_5 = (m_{30} - 3m_{12})(m_{30} + m_{12})[(m_{30} + m_{12})^2 - 3(m_{21} + m_{03})] + (3m_{21} - m_{03})(m_{21} + m_{03})[3(m_{30} + m_{12})^2 - (m_{21} + m_{03})^2]$ | δ^{12} |
| $M_6 = (m_{20} + m_{02})[(m_{30} + m_{12})^2 - (m_{21} + m_{03})^2] + 4m_{11}(m_{30} + m_{12})(m_{21} + m_{03})$ | δ^6 |
| $M_7 = (3m_{21} - m_{03})(m_{30} + m_{12})[(m_{30} + m_{12})^2 - 3(m_{21} + m_{03})] - (m_{30} - 3m_{12})(m_{21} + m_{03})[3(m_{30} + m_{12})^2 - (m_{21} + m_{03})^2]$ | δ^{12} |

Здесь δ , $0 < \delta \leq 1$ – коэффициент, определяющий точность измерения координат точек объекта. Чувствительность инвариантов к флуктуациям является неодинаковой, что может быть учтено при отборе признаков для решения задачи распознавания [1]. Для формирования и корректного сравнения векторов признаков необходимо их нормализовать:

$$M'_i = \frac{M_i - M_i^{(\min)}}{M_i^{(\max)} - M_i^{(\min)}} \cdot 10, \quad i = \overline{1, 7}, \quad \text{где } M_i^{(\min)}, M_i^{(\max)} - \text{минимальное и максимальное значения.}$$

Рассмотрим задачу выбора адекватной метрики.

Выбор метрики

Для решения задачи о принадлежности ω к классу Ω_p можно применить свертку в

виде евклидовой меры $d(\omega, \Omega_p) = \sqrt{\sum_{i=1}^7 (M_i^\omega - M_i^{\Omega_p})^2}$, где $M_i^{\Omega_p} = (\sum_{l=1}^{m_i} M_{i,l}^{\Omega_p}) / m_i$ – среднее

инвариантных моментов всех изображений, входящих в класс Ω_p . Расстояние Евклида не учитывает корреляционные связи изображений, что сказывается на качестве распознавания, поэтому предлагается воспользоваться расстоянием Евклида-Махаланобиса [2]. Метрику можно использовать для измерения расстояний между изображением ω (представленным вектором признаков x) и классом Ω_p : $d_{E-M}(\omega, \Omega_p) = \sqrt{(x - \bar{y})^T (S_p + E)^{-1} (x - \bar{y})}$, где S_p – матрица ковариации класса p , \bar{y} – центр класса. В качестве альтернативного классификатора предлагается использовать ИНС прямого распространения и вероятностную нейронную сеть.

Постановка задач классификации нейронными сетями

Решение задачи классификации предполагает обучение типовой ИНС. Пусть дано множество M изображений лиц $\{\omega_1, \dots, \omega_m\}$, каждое из которых формализуется вектором значений признаков $X_i = (x_{i1}, \dots, x_{in})$, $i = 1, \dots, m$, $x_j, j = \overline{1, n}$, где n – количество признаков.

Вектора признаков отнесены экспертами к некоторым классам $\Omega_l, l = 1, \dots, k, M = \bigcup_{l=1}^k \Omega_l$,

следующим образом $\Omega_1 = \{\omega_1, \dots, \omega_{m_1}\}, \Omega_2 = \{\omega_{m_1+1}, \dots, \omega_{m_1+m_2}\}, \dots, \Omega_k = \{\omega_{m_{k-1}}, \dots, \omega_m\}$, где $m_1 + m_2 + \dots + m_k = m, m = |M|$. Вся выборка подразделена на два непересекающихся

подмножества: обучающее и тестовое. После обучения ИНС ее качество проверяется на тестовой части. Предлагается ИНС прямого распространения с сигмоидальной функцией

активации $\sigma(x) = \frac{1}{1 + e^{-x}}$ нейронов скрытого слоя и линейной функцией активации

выходного слоя. ИНС должна быть настроена таким образом, чтобы при подаче вектора признаков X_i , отнесенного экспертами к классу Ω_i , на выходе с номером i сеть выдавала значение «1», а на всех остальных выходах – «0». Это достигается настройкой сети методом обратного распространения ошибки. С учетом особенностей формирования классов для распознавания лиц представляет интерес вероятностная нейронная сеть. В ней образцы классифицируются на основе оценок их близости к классам с учетом особенностей вероятностного распределения значений признаков. Для каждого класса на основе учебных данных определяется функция плотности распределения признака, которая характеризуется математическим ожиданием и дисперсией. Функция активация нейрона, соответствующего

j -ому классу, имеет вид $f_j = \exp\left(-\sum_{i=1}^n (w_{ij} - x_i)^2 / \sigma_{ij}\right)$, где: σ_{ij} – среднее квадратическое отклонение i -го признака j -го класса, w_{ij} – математическое ожидание i -го признака j -го класса. При подаче на вход вектора признаков неизвестного образца будет выбран класс, которому соответствует максимальное выходное значение функции.

Программная реализация и эксперименты

Для экспериментов были использованы фотографии лиц размером 380x420 пикселей из открытой базы данных [9]. В качестве среды разработки использовались программные средства C# и MATLAB. Из базы данных экспертами были отобраны изображения 30 человек на белом фоне с разрешением 380x420 пикселей. При создании классов на каждого человека приходилось по три фотографии с различными эмоциями и поворотами головы. В таблице 2 приведены примеры классов и образцов, которые подаются для распознавания.

Таблица 2

Пример фотографий после предварительной обработки.

| Примеры фотографий из обучающих выборок трех классов | | | Образцы для распознавания | |
|--|---|---|---|---|
| Класс 1 |  |  |  |  |
| Класс 2 |  |  |  |  |
| Класс 3 |  |  |  |  |

В таблице 3 представлены выборочно математические ожидания моментов различных классов и соответствующие дисперсии.

Таблица 3

Математические ожидания и дисперсии инвариантных моментов

| Класс | M_1 | σ_1^2 | M_2 | σ_2^2 | M_3 | σ_3^2 | M_4 | σ_4^2 | M_5 | σ_5^2 | M_6 | σ_6^2 | M_7 | σ_7^2 |
|-------|-------|--------------|-------|--------------|-------|--------------|-------|--------------|-------|--------------|-------|--------------|-------|--------------|
| 1 | 9.7 | 0.1 | 0.4 | 0.1 | 0.2 | 0.1 | 0.2 | 0.1 | 7.9 | 0.0 | 3.5 | 0.0 | 1.1 | 0.0 |
| 2 | 5.8 | 0.5 | 4.6 | 16.8 | 0.5 | 0.1 | 1.0 | 0.4 | 7.8 | 0.0 | 4.1 | 0.3 | 1.1 | 0.0 |
| 3 | 6.7 | 0.0 | 4.3 | 0.8 | 0.7 | 0.0 | 0.4 | 0.0 | 7.8 | 0.0 | 3.7 | 0.0 | 1.1 | 0.0 |

| | | | | | | | | | | | | | | |
|---|-----|-----|-----|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 4 | 5.3 | 0.3 | 5.6 | 13.8 | 1.0 | 0.0 | 2.9 | 0.2 | 7.7 | 0.0 | 4.6 | 0.4 | 1.2 | 0.1 |
| 5 | 7.4 | 0.2 | 6.1 | 3.3 | 0.8 | 0.0 | 0.9 | 0.2 | 7.8 | 0.0 | 4.0 | 0.1 | 1.1 | 0.0 |

В таблице 4 показаны результаты экспериментального исследования качества распознавания лиц различными методами.

Таблица 4

Сводная таблица результатов экспериментов

| Метод | Метрика Евклида | Метрика Евклида-Махаланобиса | ИНС прямого распространения | Вероятностная ИНС |
|---------------|-----------------|------------------------------|-----------------------------|-------------------|
| Кол-во ошибок | 6 | 2 | 4 | 2 |

Из результатов эксперимента следует, что приоритет по качеству классификации принадлежит метрике Евклида-Махаланобиса и вероятностной нейронной сети.

Заключение

Проведение вычислительных экспериментов показало перспективность комплексного подхода распознавания лиц по полутоновым фотографиям на основе метода инвариантных моментов и классификаторов на основе метрики Евклида-Махаланобиса и вероятностной нейронной сети. При использовании метрики Евклида-Махаланобиса система справляется с поворотами и/или наклонами головы, а также с яркостными изменениями изображения. Вероятностная нейронная сеть лучше справляется с такими «сложными» для системы факторами как закрытые глаза, измененная мимика лица (улыбка, гримаса и т.п.). Недостатком нейронных сетей являются существенные временные затраты на обучение при больших объемах данных. Коренного улучшения результата, очевидно, следует ожидать после перехода на 3-D технологию распознавания лиц и применение высокопроизводительных вычислений для достижения реального времени при полном цикле обработки изображений.

Результаты исследований планируется применить в системе биометрической идентификации человека по изображению его лица, создаваемой в рамках проекта Программы фундаментальных исследований ОНИТ 1 РАН «Интеллектуальные информационные технологии, системный анализ и автоматизация».

Список литературы

1. Абрамов Н. С., Хачумов В. М. Распознавание на основе инвариантных моментов // Вестник РУДН. Серия: Математика. Информатика. Физика. — 2014. — № 2. — С.142-149.

2. Амелькин С. А., Захаров А. В., Хачумов В. М. Обобщенное расстояние Евклида-Махаланобиса и его свойства // Информационные технологии и вычислительные системы. — 2006. — № 4. — С. 40-44.
3. Баев А. А., Морозовский К. В., Роженцов А. А. Модифицированное обобщённое преобразование Хафа для обработки трёхмерных изображений с неизвестными параметрами вращения и масштабирования // Автометрия. — 2013. — № 2. — С. 30-41.
4. Буй Т. Ч., Спицын В. Г., Фан Н. Х. Распознавание лиц на основе применения метода Виолы-Джонса, вейвлет-преобразования и метода главных компонент // Известия Томского политехнического университета. — 2012. — Т. 320, №5. — С. 54-59.
5. Савченко А. В. Метод направленного перебора альтернатив в задаче распознавания полутоновых изображений // Автометрия. — 2009. — Т. 45, № 3. — С. 90-98.
6. Савченко А. В. Теоретико-вероятностная модель полутонового изображения для задачи распознавания образов без учителя на основе метода направленного перебора // Компьютерная оптика. — 2011. — Т. 35, № 3. — С. 385-394.
7. Хачумов М. В. Задача кластеризации текстовых документов // Информационные технологии и вычислительные системы. — 2010. — № 2. — С. 42-49.
8. Baker F., Dib R., Zaeri N. Thermal Face Recognition Using Moments Invariants // International Journal of Signal Processing Systems. — 2015. — Vol.3, № 2. — P. 94-99.
9. Casati J., Moraes D., Rodrigues E. SFA: A Human Skin Image Database based on FERET and AR Facial Images // In: IX Workshop de Visão Computacional, 2013. URL: <http://www.sel.eesc.usp.br/sfa/> (дата обращения: 29.10.2015).
10. Flusser J., Suk T., Zitová B. Moments and Moment Invariants in Pattern Recognition / John Wiley & Sons Ltd., 2009. — 296 p.

Рецензенты:

Славин О.А., д.т.н., рук. проектов ООО «Смарт Энджинс Рус», г. Москва;

Морозова Т.Ю., д.т.н., профессор, зав. кафедрой «Автоматизированные системы управления» МИРЭА, г. Москва.