

РАСПОЗНАВАНИЕ ЛИЦ НА ОСНОВЕ МНОГОМАСШТАБНОГО ЗНАКОВОГО ПРЕДСТАВЛЕНИЯ ИЗОБРАЖЕНИЙ

Гончаров А.В., ведущий математик Технологического института Южного федерального университета,
г. Таганрог, ag.tsure@gmail.com

Ключевые слова: распознавание образов, анализ изображений, информативные признаки, алгоритм идентификации, знаковое представление.

Введение

Распознавание лиц является актуальной задачей искусственного интеллекта и машинного зрения в таких областях применения, как системы видео наблюдения, автоматизированный поиск в больших фотоархивах и др. Под распознаванием лиц принято понимать широкий спектр задач: детекцию лиц, локализацию антропометрических признаков (контуры бровей, глаз, носа, рта и овала лица), идентификацию и верификацию лиц. В данной статье основное внимание уделяется задаче идентификации лиц: из коллекции изображений необходимо осуществлять поиск поступающих на вход лиц, и строить на выходе упорядоченный набор изображений из базы, наиболее похожих на лицо-запрос.

Многие методы идентификации лиц основаны на принципе, который впервые был применен для задачи распознавания, изложенном в [1]. Поскольку данный метод оперирует с исходными значениями яркости изображения, он чувствителен к изменению освещения, и при использовании в реальных условиях более близкими оказываются изображения, полученные при схожих условиях освещения, а не изображения лица одного и того же человека. Таким образом, основной проблемой при решении задач распознавания лиц является изменение условий освещения [2 – 4]. Подходы к решению проблемы освещенности можно разделить на две группы — подходы, основанные на компенсации эффектов освещенности при помощи предварительной обработки изображения, и подходы, основанные на разработке специальных представлений изображений, устойчивых к изменению условий освещения. По результатам сравнительного анализа алгоритмов идентификации лиц, приведенным в [2], наилучшее качество идентификации достигается при предварительной обработке изображения, в частности, методом саморазделенного изображения (Self Quotient Image, SQI) [5]. Среди второй группы методов следует отметить представление изображения в виде локальных бинарных паттернов (Local Binary Patterns, LBP) [6].

В данной работе исследуется новый подход к формированию информативных признаков изображения, устойчивых к изменению освещенности, — *многомасштабное знаковое представление* изображения, в осно-

Предлагается новый подход к построению системы информативных признаков для задач распознавания образов и анализа изображений. Рассматривается вопрос о выборе весовых коэффициентов для взвешенной метрики Хемминга на знаковых представлениях. Предлагается алгоритм идентификации лиц на основе введенного представления изображений. Проводится оценка качества идентификации лиц на изображениях базы университета Essex в терминах полноты и точности. Приводятся результаты сравнения предложенного алгоритма с современными аналогами.

ве которого лежит анализ отношений порядка на смежных пикселях изображения. Знаковое представление изображений рассматривалось в работах [7,8] и применялось для решения таких задач, как детекция лиц на изображении, идентификация лиц, локализация антропометрических признаков лица, поиск нечетких дубликатов в больших коллекциях изображений. Многомасштабное знаковое представление — это естественное обобщение знакового представления изображений для одновременного анализа нескольких масштабов изображения (пирамиды изображений). Например, при детекции лиц, производится анализ нескольких масштабов изображения для поиска лиц произвольного размера. В задаче локализации антропометрических признаков лица анализ изображения на нескольких масштабах используется с целью последовательного уточнения координат искомых признаков лица.

Таким образом, многомасштабное знаковое представление можно рассматривать как специальное представление изображения, обеспечивающее устойчивость к изменениям условий освещения за счет перехода от исходных значений яркости изображения к рассмотрению отношений порядка на смежных пикселях. Анализ изображения на нескольких масштабах при построении знакового представления позволяет одновременно учитывать как крупные, так и мелкие детали лица, что способствует повышению качества идентификации лиц. К крупным деталям лица можно отнести его форму («овальное», «треугольное», «ромбовидное» и т.п.) и относительные размеры лицевых признаков (рот, нос, глаза и брови). К мелким деталям лица относится форма лицевых признаков, например, форма бровей, разрез глаз, контуры носа и рта.

Многомасштабное знаковое представление изображений

Пусть изображение — это функция $f: [0, a] \times [0, b] \rightarrow [0, \infty)$, действующая из некоторого прямоугольника во множество неотрицательных вещественных чисел. Данную функцию можно построить из

реально получаемого дискретизированного изображения с помощью алгоритма интерполяции. На прямоугольнике $[0, a] \times [0, b]$ будем рассматривать равномерные сетки различной скважности, определяемые длиной ячейки Δx вдоль оси X и длиной ячейки Δy вдоль оси Y . Целесообразно выбрать длины ячеек так, чтобы значения $N = a/\Delta x$ и $M = b/\Delta y$ были натуральными. Тогда сеточное представление $f_{\Delta x, \Delta y}$ изображения f — это функция $f_{\Delta x, \Delta y} : \{0, \dots, N\} \times \{0, \dots, M\} \rightarrow [0, \infty)$, определяемая равенством:

$$f_{\Delta x, \Delta y}(i, j) = \frac{1}{\Delta x \Delta y} \int_{\Delta x i}^{\Delta x(i+1)} \int_{\Delta y j}^{\Delta y(j+1)} f(x, y) dx dy,$$

где $i \in \{0, \dots, N\}$, $j \in \{0, \dots, M\}$. Можно рассматривать не одно, а несколько сеточных представлений изображения, а именно, $f_{\Delta x_k, \Delta y_k}$, $k = 1, \dots, n$, которые назовем *масштабами изображения* f .

Знаковым представлением сеточной части изображения $f_{\Delta x, \Delta y}$ будем считать «двухслойную» матрицу

$$\Pi_{\Delta x, \Delta y}(f) = \left(\left(\pi_{i,j}^x, \pi_{i,j}^y \right) \right)_{N \times M},$$

состоящую из знаков разностей –

$$\pi_{i,j}^x = \text{sgn} \left(f(\Delta x i, \Delta y j) - f(\Delta x(i-1), \Delta y j) \right) \text{ и}$$

$$\pi_{i,j}^y = \text{sgn} \left(f(\Delta x i, \Delta y j) - f(\Delta x i, \Delta y(j-1)) \right), \quad i = 1, \dots, N,$$

$$j = 1, \dots, M.$$

Упорядоченный набор знаковых представлений

$$\left\{ \Pi_{\Delta x_k, \Delta y_k} \right\}_{k=1}^n,$$

построенных для масштабов $f_{\Delta x_1, \Delta y_1}, f_{\Delta x_2, \Delta y_2}, \dots, f_{\Delta x_n, \Delta y_n}$ изображения f является *многомасштабным знаковым представлением* изображения f .

Для решения задач распознавания образов введем понятие меры близости на изображениях с помощью метрики на многомасштабных знаковых представлениях. Пусть изображения $f : [0, a] \times [0, b] \rightarrow [0, \infty)$, $g : [0, a] \times [0, b] \rightarrow [0, \infty)$ и соответствующие им многомасштабные знаковые представления

$$\left\{ \Pi_{\Delta x_k, \Delta y_k}^{(f)} \right\}_{k=1}^n \text{ и } \left\{ \Pi_{\Delta x_k, \Delta y_k}^{(g)} \right\}_{k=1}^n, \text{ тогда}$$

$$d(f, g) = \sum_{k=1}^n \alpha_k \rho \left(\Pi_{\Delta x_k, \Delta y_k}^{(f)}, \Pi_{\Delta x_k, \Delta y_k}^{(g)} \right),$$

где ρ — метрика на матрицах изменения яркостей, а $\alpha_k > 0$ — весовые коэффициенты, определяющие приоритет каждого из масштабов. В качестве метрики на знаковых представлениях изображений используем взвешенную метрику Хэмминга, расстояние по которой соответствует количеству различных соответствующих элементов матриц

$$\Pi_{\Delta x_k, \Delta y_k}^{(f)} \text{ и } \Pi_{\Delta x_k, \Delta y_k}^{(g)} :$$

$$\rho \left(\Pi_{\Delta x_k, \Delta y_k}^{(f)}, \Pi_{\Delta x_k, \Delta y_k}^{(g)} \right) =$$

$$= \sum_{i=1}^N \sum_{j=1}^M \omega_{kij} \theta \left(\Pi_{\Delta x_k, \Delta y_k}^{(f)}(i, j), \Pi_{\Delta x_k, \Delta y_k}^{(g)}(i, j) \right), \quad (1)$$

где $\omega_{kij} \geq 0$ — весовые коэффициенты,

$$\theta(a, b) = \begin{cases} 0, & a = b, \\ 1, & a \neq b. \end{cases}$$

Отметим, что поскольку на весовые коэффициенты накладывается нестрогое неравенство, выражение (1), вообще говоря, по своим свойствам является псевдометрикой.

Для определения весовых коэффициентов ω_{kij} используем подход, основанный на измерении количества информации, соответствующей элементам матрицы знакового представления. При анализе реальных коллекций изображений каждое изображение можно моделировать как реализацию многомерной случайной величины f . Тогда элементы матрицы знакового представления изображения f также являются случайными величинами, и для оценки их информативности можно, например, использовать энтропию Шеннона, которая широко применяется при анализе и кодировании изображений [9, 10]. Пусть $\pi = \{(1, 1), (1, -1), (-1, 1), (-1, -1)\}$ — множество значений, принимаемых элементами матрицы знакового представления, и $F = \{f_1, \dots, f_N\}$ — случайная выборка изображений заданного класса. Тогда $P \left\{ \Pi_{\Delta x_k, \Delta y_k}^{(f)}(i, j) = \pi_r \right\}$ — вероятность того, что элемент $\Pi_{\Delta x_k, \Delta y_k}^{(f)}(i, j)$ матрицы знакового представления изображения f примет фиксированное значение π_r , можно оценить по выборке F изображений заданного класса как частоту появления данного значения:

$$p_r^{(k)}(i, j) = \frac{\left| \left\{ f \in F \mid \Pi_{\Delta x_k, \Delta y_k}^{(f)}(i, j) = \pi_r \right\} \right|}{|F|},$$

где $|F|$ — мощность множества F .

Считая, что элементы матрицы знакового представления являются независимыми случайными величинами, количество информации можно определить с помощью энтропии Шеннона следующим образом:

$$S^{(k)}(i, j) = - \sum_{r=1}^{|\pi|} p_r^{(k)}(i, j) \log p_r^{(k)}(i, j). \quad (2)$$

Таким образом, в качестве весовых коэффициентов в выражении (1) для k -го масштаба возьмем значение энтропии признаков на k -м масштабе $\omega_{kij} = S^{(k)}(i, j)$, тогда выражение для меры близости на изображениях, задаваемое метрикой на многомасштабных знаковых представлениях, примет вид:

$$\rho(f, g) = \sum_{k=1}^n \alpha_k \sum_{i=1}^N \sum_{j=1}^M \omega_{kij} \theta \left(\Pi_{\Delta x_k, \Delta y_k}^{(f)}(i, j), \Pi_{\Delta x_k, \Delta y_k}^{(g)}(i, j) \right). \quad (3)$$

Мера близости (3) по своим свойствам является псевдометрикой на изображениях, поскольку из $\rho(f, g) = 0$, вообще говоря, не следует $f = g$, и задает разбиение множества изображений Ψ на непересекающиеся классы эквивалентности $\Psi_f = \{g \in \Psi \mid \rho(f, g) = 0\}$, причем, если $g(x, y) = \varphi(f(x, y))$ для всех $(x, y) \in [0, a] \times [0, b]$, где φ — монотонно возрастающая функция, то $g \in \Psi_f$. Таким образом, знаковое представление является инвариантным относительно строго монотонно возрастающего преобразования яркости изображения. Отметим, что такие преобразования яркости изображения являются характерными при формировании и обработке изображения. Например, изменение яркости и контрастности изображения, гамма-коррекция, искажения, возникающие при формировании изображения цифровыми фото сенсорами, естественным образом учитываются многомасштабным знаковым представлением.

Идентификация лиц на основе многомасштабного знакового представления

Задача состоит в необходимости выполнить поиск по изображению-запросу и выдавать пользователю упорядоченный набор лиц, наиболее похожих на запрос.

Прежде всего, необходимо рассчитать весовые коэффициенты метрики (3), используя выражения для оценки энтропии (2) по выборке изображений лиц. На данном этапе все изображения лиц из обучающей выборки приводятся к единому размеру, и для них рассчитывается многомасштабное знаковое представление. Для того чтобы фон не влиял на значение меры близости, используется маска, которая зануляет область изображения, содержащую элементы фона. Размер изображения на каждом масштабе выбирается, исходя из характерного размера деталей лица в пикселях по аналогии с идеей кратномасштабного анализа. «Мелкие» масштабы (размер изображения лица от 10x10 до 20x20 пикселей) характеризуют интегральные особенности лица и отвечают за существенные различия в форме лица (овальное лицо, ромбовидное, треугольное и т.д.). «Средние» масштабы (размер изображения лица от 20x20 до 50x50 пикселей) отвечают за локальные особенности лица — форму глаз, контур носа, относительные размеры глаз, бровей, носа и рта. «Крупные» масштабы (размер изображения лица 50x50 пикселей и выше) отвечают за «текстурные» особенности лица.

Следующий этап — это формирование базы, по которой осуществляется поиск лиц. На данном этапе для каждого лица вычисляется многомасштабное знаковое представление и сохраняется в базе. Отметим, что построение многомасштабного знакового представления является вычислительно простой операцией и требует всего $O(nNM)$ арифметических операций, где n — количество масштабов, N и M — размер сетки самого крупного масштаба. Кроме того, добавление очередного лица в базу никак не затрагивает уже имеющиеся изображения в базе, необходимо лишь посчитать многомасштабное знаковое представление для данного изображения.

Рассмотренные выше этапы (расчет весовых коэффициентов и формирование базы лиц) проводятся один раз перед использованием алгоритма идентификации.

Непосредственно процедура поиска лица-запроса в базе данных заключается в вычислении псевдометрики (3) между изображением-запросом и всеми элементами базы лиц. Изображения из базы, расстояние до которых от изображения-запроса не превышает заданного порога, упорядоченные по мере возрастания псевдометрики, предоставляются пользователю в качестве результата поиска.

Оценка качества идентификации лиц и результаты экспериментов

Определение качества работы алгоритмов поиска лиц будем производить методом статистической оценки показателей полноты (Recall) и точности (Precision) поиска, широко используемых в информационном поиске [11]. Полнота и точность оцениваются по тестовой коллекции $\Omega = (f_k, l_k)_{k=1}^L$, в которой каждому изображению f_k соответствует идентификатор персоны l_k . Пусть $\Omega_l = \{(f, n) \in \Omega | n = l\}$ — множество изображений, соответствующих персоне с идентификатором l , а

$$\Psi_{l_k} = \{(f, n) \in \Omega | \rho(f, q_{l_k}) < t\} -$$

множество изображений, найденных по изображению-запросу q_{l_k} , соответствующему персоне с идентификатором l , где t — пороговая величина, которая задает максимальное расстояние между похожими лицами. Тогда полнота r_{l_k} и точность p_{l_k} поиска изображений лиц по запросу l_k , целью которого является поиск изображения с идентификатором l , определяются следующим образом:

$$r_{l_k} = \frac{|\Omega_l \cap \Psi_{l_k}|}{|\Omega_l|}, \quad p_{l_k} = \frac{|\Omega_l \cap \Psi_{l_k}|}{|\Psi_{l_k}|}.$$

Для интегральной оценки метода поиска лиц производится усреднение показателей r_{l_k} и p_{l_k} по всем запросам l_k :

$$P = \frac{1}{L} \sum_{k=1}^L p_{l_k}, \quad R = \frac{1}{L} \sum_{k=1}^L r_{l_k}.$$

Таким образом, полнота поиска характеризует долю найденных изображений искомого человека среди всех изображений этого человека в базе, а точность характеризует долю изображений искомого человека среди всех найденных изображений.

Оценка качества предложенного алгоритма поиска лиц производилась на изображениях базы Essex [12], которая содержит 7900 изображений — по 20 фотографий каждого из 395 персон. Множество изображений было разделено на два непересекающихся подмножества: обучающее Ω_1 и тестовое Ω_2 , $|\Omega_1|/|\Omega_2| = 1/2$. На Ω_1 осуществлялось обучение алгоритма, состоящее в настройке весовых коэффициентов (2), а на Ω_2 оценивалось качество работы в терминах полноты и точности.

Для оценки качества были использованы алгоритмы идентификации, основанные на знаковом представлении как для одного масштаба (на рис. 1 им соответствуют графики SR1, размер лица 24x24 пикселя, и SR2, размер лица 64x64 пикселя), так и для совокупности двух масштабов (на рис. 1 — графики ML1, ML2 и ML3). Кроме того, для сравнения приведены результаты оценки качества для метода идентификации лиц, основанного на методе главных компонент [1] (на рис. 1 ему соответствует график PCA, размер лица 64x64 пикселя).

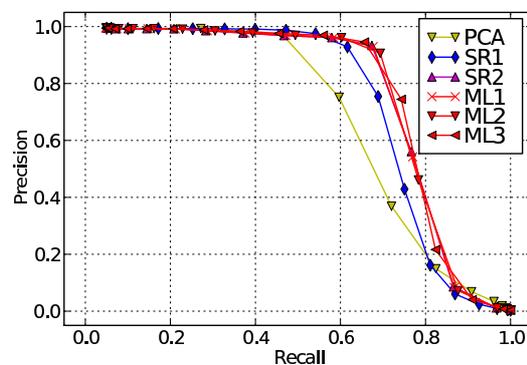


Рис. 1. Результаты оценки качества идентификации лиц на изображениях базы университета Essex. PCA — метод главных компонент, SR1 и SR2 — методы на основе знакового представления изображений, ML1 - ML3 — методы на основе многомасштабного знакового представления изображений.

Заключение

В работе предложено многомасштабное знаковое представление изображений. Рассматривается введение метрики на многомасштабных знаковых представлениях и настройке весовых коэффициентов на основе информативности признаков. Результат проведенного вычислительного эксперимента свидетельствует об эффективности и перспективности рассмотренного подхода. Разработанные методы идентификации лиц превосходят метод главных компонент по показателям полноты и точности. Кроме того, результаты идентификации лиц на основе многомасштабного знакового представления превосходят результаты идентификации на основе знакового представления, рассчитанного для одного масштаба, что особенно заметно при полноте поиска более 70%, когда точность резко падает при небольшом приросте полноты.

Литература

1. Turk M., Pentland A. Eigenfaces for recognition. // J. Cognitive Neuroscience. 3(1). 1991. P. 71 – 86.
2. [Ruiz] Ruiz-del-Solar J., Quintero J. Illumination compensation and normalization in eigenspace-based face recognition: A comparative study of different pre-processing approaches // Pattern Recognition Letters. Vol 29(14). 2008. P. 1966 – 1979
3. Zhou S., Chellappa R., Zhao W. Unconstrained Face Recognition. Series: International Series on Biometrics. Vol. 5. Springer Science & Business. 2006.
4. Delac K., Grgic M. Face Recognition. I-TECH Education and Publishing. 2007.
5. Wang H., Li S., Wang Y. Face Recognition under Varying Lighting Conditions Using Self Quotient Image. FGR. P. 819 – 824. 2004.
6. Fröba B., Ernst A. Face Detection with the Modified Census Transform // Proc. 6th Int. Conf. on Face and Gesture Recognition – FG 2004, Seoul, Korea. 2004. P. 91 – 96.
7. Гончаров А.В., Губарев В.В. Выделение характерных признаков лиц на цифровых изображениях с использованием

- знакового представления // Математические методы распознавания образов: 14-я Всероссийская конференция. Суздаль, 21-26 сентября 2009г.: Сборник докладов. М.: МАКС Пресс. 2009. С. 325 – 328
8. Гончаров А.В., Каркищенко А.Н. Влияние освещенности на качество распознавания фронтальных лиц // Известия ЮФУ. Технические науки. Тематический выпуск «Интеллектуальные САПР». Тагангр: Изд-во ТТИ ЮФУ. №4(81). 2008. С. 88 – 92
 9. [Pratt] Pratt W. Digital Image Processing. Wiley. 1978
 10. Escolano F., Suau P., Bonev B. Information Theory in Computer Vision and Pattern Recognition. Springer Verlag. 2009.
 11. Российский семинар по Оценке Методов Информационного Поиска. Труды РОМИП 2009 (Петрозаводск, 16 сентября 2009г.) / Под ред. И. Некрестьянова, М. Некрестьяновой. — Санкт-Петербург: НУ ЦСИ, 2009, 198 с.
 12. Spacek L. Collection of Facial Images. <http://cswww.essex.ac.uk/mv/allfaces/index.html>

FACE RECOGNITION ON THE BASIS OF SIGN-BASED IMAGE REPRESENTATION

Goncharov A.V., ag.tsure@gmail.com

A new approach to the feature extraction in pattern recognition and image analysis is proposed. The problem of choosing of weights coefficients in weighted Hamming's similarity measure on sign-based image representation is discussed. The face identification algorithm based on introduced image representation is proposed. Quality estimation of face identification algorithms in terms of recall and precision is carried out on face database of the Essex University. Results of the quality assessment of introduced face identification algorithm are compared to the state of the art.