

ИНФОРМАТИКА

УДК 004.931

ОБЗОР МЕТОДОВ РАСПОЗНАВАНИЯ ЛИЦА НА ИЗОБРАЖЕНИИ

А.В. АНТОНЧИК, А.А. ДЕРЮШЕВ

*Белорусский государственный университет информатики и радиоэлектроники
П. Бровка, 6, Минск, 220013, Беларусь*

Поступила в редакцию 11 ноября 2008

Приведена краткая характеристика существующих методов распознавания лица на статических изображениях. Рассмотрены методы с явным и неявным выделением черт лица. Предложена классификация методов распознавания лица. Выделены кластерные, модельные и признаковые методы. Определены основные принципы работы каждой категории методов с позиций экстрагирования и классификации.

Ключевые слова: метод распознавания, способ распознавания, экстрактор признаков, классификация изображений лиц, сравнение лиц.

Введение

С развитием информационных технологий все большей популярностью пользуются системы ограничения доступа и идентификации личности. Традиционные способы идентификации, такие как ключи или пароли, не обеспечивают высокую степень надежности идентификации, поэтому наиболее динамично развиваются системы биометрической идентификации. Область применения этих технологий с каждым годом существенно расширяется. По данным International Biometric Group, распознавание личности по изображению лица является третьим по популярности после систем идентификации по папиллярному узору и форме ладони.

Задача распознавания

В работе любого алгоритма распознавания [1, 2] или обнаружения [3] лица следует выделить два логических блока: экстрактор характерных признаков и механизм классификации [4, 5]. Действие экстрактора основано на выделении из огромного потока входных данных полезной для классификатора информации. При идентификации личности этой информацией могут являться характеристики однозначно определяемых признаков (например, применяющееся в криминалистике относительное расположение глаз, бровей, губ и носа). Классификатор при принятии решения о назначении метки класса распознаваемому объекту должен руководствоваться именно этими признаками. Выбор признаков является наиболее важной задачей. Очевидно, что при их выборе учитываются наиболее уникальные свойства, так как по ним возможно наиболее достоверно судить о принадлежности объекта к тому или иному классу. Существует множество разных подходов к получению признаков класса. Применительно к задаче распознавания лица можно выделить три основных подхода к экстрагированию признаков.

К категории признаковых можно отнести методы, основанные на эмпирических знаниях и попытках формализовать процессы распознавания, происходящие в головном мозге человека. К таким методам можно отнести методы с явным выделением признаков, методы шаблонного сравнения. В этих методах классификация осуществляется с использованием параметров структурных элементов лица, их относительных характеристик и косвенных проявлений.

Исторически первыми в признаковой категории были геометрические методы. В ходе распознавания выделялись ключевые точки с использованием детекторов глаза, носа, рта, и анализировалось их взаимное расположение. Эти методы применялись тогда, когда вычислительные способности компьютерных систем не позволяли использовать более сложные алгоритмы. Гораздо более эффективны признаковые методы распознавания на основе карты краевых линий [6]. При этом в исходном изображении выделяют края с разными пороговыми значениями, и полученный набор сравнивают с картой краев модели посредством вычисления расстояния по Хаусдорфу. Признаками классов этого метода являются яркостные перепады и неоднородности. Метод показывает высокие результаты [2].

Другим ранним методом распознавания был метод сравнения эталонов. Метод обладал существенными недостатками, обусловленными простотой алгоритма, которые успешно решались с появлением сравнения эластичных графов лица [7]. Преимущества использования эластичных графов состоят в низкой чувствительности к вариативности положения, масштаба и выражения лица. Для этого подготавливается обобщенный граф лица по имеющимся в базе лицам. С помощью вейвлетов Габора выделяются локальные признаки лица. Затем, сравнивая заданное изображение с обобщенным графом с помощью гибкого сравнения, создается граф изображения лица, состоящий из малого набора подграфов шаблонных изображений, находящихся в его вершинах. Процесс назначения метки класса заданному изображению состоит в сравнении созданного графа с графами базы лиц.

Анализ локальных признаков (АЛП) [2] является методом, использующим топографическое индексирование для описания лица. По сути АЛП представляет собой банк локально коррелированных детекторов. Детекторы осуществляют разложение по методу главных компонент (МГК). Ошибочность такого представления равна ошибке разложения по методу главных компонент, однако такое представление позволяет резко уменьшить критичность к вариациям освещения лица, понизить чувствительность к вращению и уменьшить вычислительную сложность по сравнению с методом главных компонент.

Принципиально иной подход к распознаванию лежит в основе категории кластерных методов. Их первоначальный принцип заключается в корреляционном сравнении изображений. Эти методы не ставят своей задачей выделить структурные элементы; признаками тут являются все пиксели изображения. Однако, так как корреляционная схема избыточна, обладает большой вычислительной сложностью и пригодна для распознавания недеформируемых объектов с неизменным положением и освещением, то эти методы практически не используют для распознавания лиц. Тем не менее, сужая пространство исходных признаков, можно сильно улучшить показатели быстродействия алгоритма. Стоит отметить, что все кластерные методы осуществляют попытку максимального разделения классов изображений лиц и преимущественно являются именно инструментом классификации изображений. Признаками экстрактора во всех кластерных методах являются наиболее отличающиеся от среднего лица области распознаваемого лица.

Метод главных компонент является классическим кластерным методом распознавания [8, 9]. Сам метод и некоторые принципы, используемые в нем, часто используются на стадиях предварительной цифровой обработки как в кластерных, так и в других способах распознавания. Поэтому метод будет рассмотрен более подробно.

Для проведения классификации в МГК любое изображение (рис. 1,а) представляется как вектор в n -мерном пространстве, где n — количество пикселей изображения.

Теоретическое основание метода заключается в переводе исходного изображения в пространство признаков

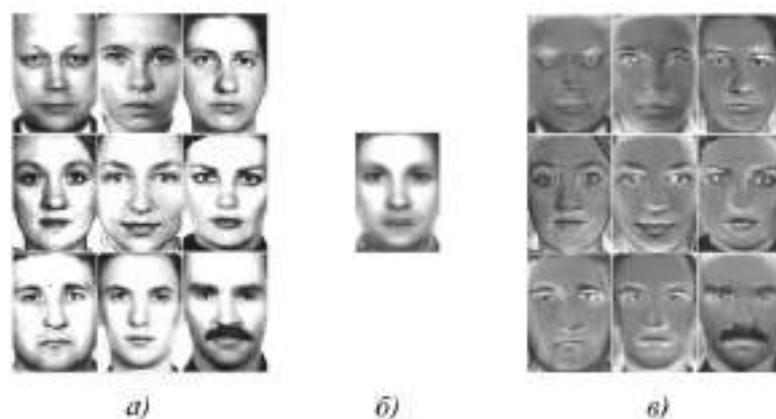


Рис. 1. Подготовка изображений лиц к классификации в МГК

с гораздо меньшей размерностью, в котором его можно максимально полно охарактеризовать. Применение метода к набору векторов линейного пространства позволяет перейти к такому базису, что основная дисперсия будет направлена вдоль нескольких первых осей базиса, называемых главными осями. Это позволяет представить уникальность векторов тренировочного набора с помощью нескольких компонент. Полученное пространство максимально информативно описывает объект. На практике для осуществления метода используется преобразование Карунена-Лоэва [10]. На первом шаге находится среднее лицо Ψ из изображений тренировочного набора S (рис. 1,б):

$$S = \{ \Gamma_1, \Gamma_2, \Gamma_3, \dots, \Gamma_M \}, \quad (1)$$

$$\Psi = \frac{1}{M} \sum_{n=1}^M \Gamma_n. \quad (2)$$

На втором шаге вычисляются изображения лиц Φ_i , центрированные относительно среднего лица посредством вычитания среднего изображения из исходных:

$$\Phi_i = \Gamma_i - \Psi. \quad (3)$$

Именно эта разность (рис. 1,в) трансформируется в форму вектора многомерного пространства лиц. Полученные векторы образуют сгустки (кластеры), так как их компоненты коррелируются между собой. Пространство лиц может быть описано собственными векторами, и, исходя из этого, требуется меньшее число измерений для представления всего пространства лиц. Собственные векторы являются собственными лицами, которые характеризуют отличия лица относительно среднего лица. Собственные векторы вычисляются из ковариационной матрицы C и сортируются в порядке убывания собственных значений.

$$C = \frac{1}{M} \sum_{n=1}^M \Phi_n \Phi_n^T = A A^T, \quad (4)$$

где матрица $A = [\Phi_1 \Phi_2 \dots \Phi_M]$.

При распознавании лица с помощью МГК вычисляются две характеристики: расстояние от проекции тестового вектора до среднего вектора тренировочного набора и расстояние от тестового вектора до его проекции в подпространство главных компонент. Исходя из этих параметров, выносится решение о принадлежности тестового объекта классу, образованному тренировочным набором. К недостаткам метода следует отнести высокую критичность к вариациям масштаба лица (40% ошибок), пространственной ориентации (15% ошибок) и освещения (5% ошибок) [4]. Метод главных компонент можно рассматривать как частный случай факторного анализа, когда все специфические факторы приняты равными нулю, а общие факторы ортогональны. Коренное отличие состоит в том, что результатом факторного анализа является модель, в явном виде описывающая зависимость наблюдаемых переменных от скрытых факторов, а МГК проводит только описательный анализ данных, без получения модели.

В отличие от метода главных компонент, метод опорных векторов [11] увеличивает размерность пространства с целью поиска базиса, в котором кластеры изображений будут являться линейно разделимыми. Для линейного разделения необходимо спроецировать кластер в пространство большей размерности. Осуществляется это с помощью ядерных функций.

Еще одним мощным решением для разделения векторов исходного изображения лица является анализ независимых компонент [10]. Метод позволяет неявно найти источник различий изображений лица. С помощью этого метода находится линейное преобразование исходного изображения, в котором полученные векторы будут максимально статистически независимыми. Результатом такого представления является облегчение задачи классификации и, следовательно, увеличение показателей правильной и уменьшение показателей неправильной классификации.

Линейный дискриминантный анализ (ЛДА) [12] ставит своей целью найти пространство, в котором кластеры изображений лиц окажутся наиболее компактно представленными и расстояние между ними будет максимальным. С помощью ЛДА удается получить маломерное

пространство, в котором кластеры лиц пересекаются минимально. Производить классификацию в таком пространстве значительно проще. Для повышения результатов распознавания и устранения недостатков одномерного ЛДА используют двумерный ЛДА [13]. Для повышения быстродействия используют совместные методы двумерных МГК и ЛДА [14], позволяющие несущественно улучшить результаты распознавания [15].

Нейронные сети успешно применяются для решения многих задач распознавания. Для распознавания лица обычно используется структура многослойного персептрона. Математически работа персептрона сводится к классификации входных сигналов, принадлежащих n -мерному гиперпространству по некоторому числу классов [16]. Это происходит путем разбиения гиперпространства гиперплоскостями. Нейронная сеть может демонстрировать очень высокие результаты, но известным недостатком любой нейронной сети является необходимость тщательной подстройки весовых коэффициентов w_{mn} (рис. 2).

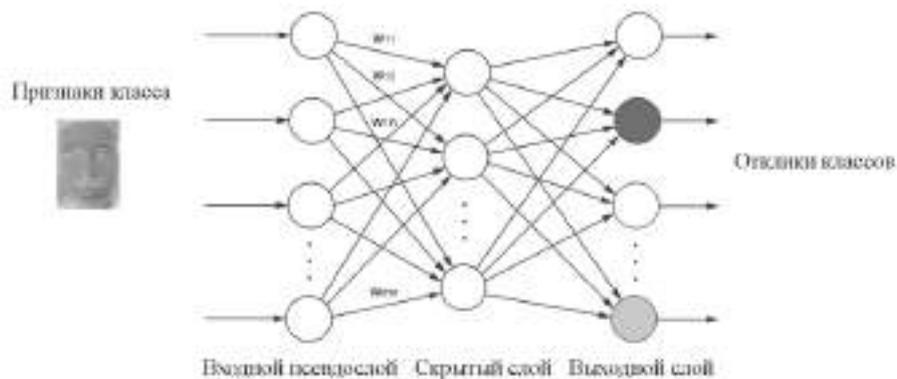


Рис. 2. Нейронная сеть прямого распространения. Максимально активный нейрон (на рисунке более темный) указывает на принадлежность к распознанному классу

Пожалуй, наиболее успешным и гармоничным решением задачи распознавания лица является использование категории модельных методов. Во всех методах этого класса создается модель, прямо или косвенно отражающая классы изображений лиц.

Скрытые Марковские модели (СММ) [17] используются для описания изображений лиц на основе только статистических данных. СММ позволяют учитывать пространственно-временные характеристики сигнала, поэтому широко используются для распознавания речи. В распознавании лиц СММ применяются следующим образом (рис. 3). Каждому классу изображений лиц соответствует своя модель. Распознаваемое изображение лица представляется в виде последовательности событий с вероятностями перехода a_{mn} (рис. 3,а). Для каждой модели вычисляется вероятность генерации представленной последовательности. Модель, получившая максимальную вероятность генерации, представляет класс распознаваемого лица. Для сокращения вычислительной сложности обычно используют одномерные [17] и псевдодвумерные [18] структуры скрытых Марковских моделей.

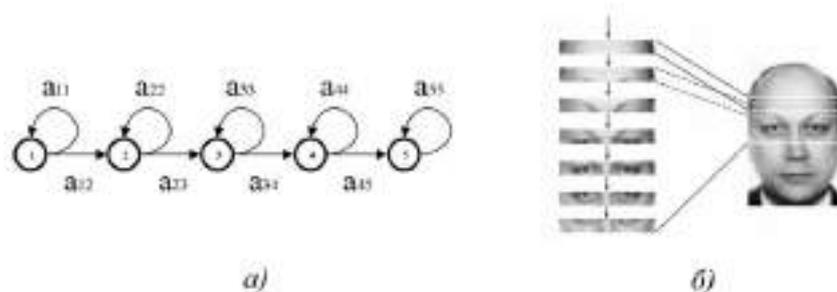


Рис. 3. Представление лица в виде одномерной СММ:
 а — конфигурация одномерной СММ; б — разбиение лица на последовательность наблюдений

В качестве наблюдений берутся блоки изображения с некоторым перекрытием (рис. 3,б). Далее вычисляются наборы коэффициентов дискретно-косинусного преобразования с низкими частотами для каждого блока. Полученные наборы наиболее полно и наименее за-

тратно представляют изображение лица. При использовании одномерной модели блоки выбираются путем разбиения изображения на несколько горизонтальных областей: подбородок, губы, нос, глаза, лоб. Метод СММ хорошо работает при варьировании условий освещения, изменении выражения лица или его ориентации. Недостатком является низкая различающая способность классов, что может ограничить использование при больших количествах пользователей.

Еще одним модельным методом является использование активных моделей внешнего вида. Активные модели внешнего вида используются для моделирования лиц, подверженных как жесткой (поворот, масштабирование), так и нежесткой (изменение выражения лица) деформации. При распознавании лица модели базы подвергаются деформации для нахождения активной модели, степень близости которой с распознаваемым лицом была бы максимальной. Одновременно с этим изменяется текстура активной модели. Наиболее передовым примером метода активных моделей внешнего вида является распознавание на основе подгонки деформируемых 3D моделей [19, 20]. В ходе работы метода на исходном изображении выделяются стандартные элементы лица, изображению лица ставится в соответствие 3D модель среднего лица из всех лиц базы и происходит ряд трансформаций, чтобы подогнать модель к оригиналу. Сначала меняется ориентация и размер средней модели, затем ее освещение. Автоматизированный алгоритм, путем изменения нескольких параметров для подгонки формы модели, находит наиболее похожую форму. Вычисляя расстояние между лицами из базы и подогнанной 3D моделью, можно найти модель наиболее соответствующего лица базы. При сравнении моделей используется метод главных компонент, при этом сама форма лица и его текстура сравниваются отдельно и корреляции между ними не учитываются. Подгонка деформируемых 3D моделей показывает многообещающие результаты и позволяет получить информацию не только о принадлежности лица, но и о поле его владельца, эмоциональном выражении, а также других параметрах. Из недостатков алгоритма сильно выделяется время подгонки лиц. Оно составляет около 4,5 минут на процессоре Pentium 4 с частотой 2 ГГц [19].

Заключение

Предложенная классификация методов распознавания лица позволяет достаточно четко разграничить существующие подходы. В кластерной категории к задаче подходят с позиций общей классификации, в признаковой — делают попытки морфологической классификации, в модельной — проводят верификацию с моделями. Каждый из методов сочетает в себе компромисс скорости и качества. При необходимости создания системы, не критичной к изменению положения лица, следует отказаться от использования методов кластерной категории. Но использование их в системах реального времени вполне оправдано. Также необходимо помнить об ограничениях, которые могут накладывать некоторые методы на общее число пользователей. Среди всех рассмотренных способов распознавания наиболее качественными и перспективными видятся модельные методы. Моделирование позволяет максимально полно отражать реальность и совершенно обоснованно требует высоких затрат ресурсов и вычислительной производительности.

A SURVEY OF METHODS FOR FACE RECOGNITION ON STATIC IMAGE

A.V. ANTONCHIK, A.A. DERYUSHEV

Abstract

Short-form description of existing methods of face recognition on static images is given. The methods with explicit and implicit feature extraction are contemplated. Classification of face recognition methods is proposed. Cluster, model and feature methods are contemplated. The basic work principles of each method category are defined from a position of extraction and classification.

Литература

1. *Delac K., Grgic M.* Face Recognition. I-Tech, 2007.
2. *Kong S., Heo J., Abidi B. et al.* // Computer Vision and Image Understanding. 2005. Vol. 97, № 1. P. 103–135.
3. *Yang M.H., Kriegman D.J., Ahuja N.* // IEEE Trans. on Pattern Analysis and Machine Intelligence. 2002. Vol. 24, № 1. P. 34–58.
4. *Форсайт Д.А., Поинс Ж.* Компьютерное зрение. Современный подход. М., 2004.
5. *Шапиро Л., Стокман Дж.* Компьютерное зрение. М., 2006.
6. Method and system for recognition people with model-based face detection: patent US 699606 B1, Int. Cl. G06K 9/00 / R. Frischholz: assignee HumanScan GmbH.
7. *Wiskott L., Fellous J.-M., Krueger N., Malsburg C.* // Intelligent Biometric Techniques in Fingerprint and Face Recognition. 1999. Ch. 11. P. 355–396.
8. *Turk M., Pentland A.* // J. Cognitive Neuroscience. 1991. Vol. 3, № 1. P. 71–86.
9. *Kirby M., Sirovich L.* // IEEE Trans. on Pattern Analysis and Machine Intelligence. 1990. Vol. 12, № 1. P. 103–108.
10. *Liu C., Wechsler H.* // Proc. of the 2nd Int. Conf. on Audio- and Video-based Biometric Person Authentication. 1999. P. 211–216.
11. *Guo G., Li S.Z., Chan K.* // IEEE Int. Conf. on Automatic Face and Gesture Recognition. 2000. P. 196–201.
12. *Zhao W.Y.* // Pattern Recognition 2000. Proc. 15th Int. Conf. 2000. Vol. 2. P. 818–821.
13. *Jieping Y., Ravi J., Qi L.* // Neural Information Processing Systems. 2004. P. 1569–1576.
14. *Sanguansat P., Asdornwised W., Jitapunkul S., Marukatat S.* // IEICE Trans. on Information and Systems. 2006. Vol. E89-D. P. 2164–2170.
15. *Delac K., Grgic M., Grgic S.* // Int. J. of Imaging Systems and Technology. 2005. Vol. 15, № 5. P. 252–260.
16. *Кружлов В.В., Борисов В.В.* Искусственные нейронные сети. Теория и практика. М., 2002.
17. *Nefian A.V.* // Int. Conf. on Acoustics, Speech and Signal Proceeding. 1998. P. 2721–2724.
18. *Othman H., Aboulnasr T.* // IEEE Trans. on Pattern Analysis and Machine Intelligence. 2003. Vol. 25, №10. P. 1229–1238.
19. *Blanz V., Vetter T.* // IEEE Trans. on Pattern Analysis and Machine Intelligence. 2003. Vol. 25, № 9. P. 1–12.
20. *Xiaoguang Lu, Anil K.J.* // Proc. of the 2006 IEEE Computer Soc. Conf. on Computer Vision and Pattern Recognition. 2006. Vol. 2. P. 1377–1383.