

01, январь 2016

УДК 612-087

Анализ методов распознавания человека по изображению лица

Нам Ди Хьян, бакалавр

*Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедра «Лазерные и оптико-электронные системы»*

Научный руководитель: Колючкин В.Я., профессор

*Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана
кафедра «Лазерные и оптико-электронные системы»
vkoluch@bmstu.ru*

В настоящее время, в задачах распознавания человека по изображению лица большое распространение получили метод главных компонент (МГК), а также редукция пространства признаков методом линейного дискриминантного анализа (ЛДА). Однако применение данных методов в задачах распознавания лиц вызывают сложности по следующим причинам. Во-первых, лицо человека – это динамический объект, имеющий высокую степень изменяемости связанной с вариациями мимики, ракурса регистраций. Во-вторых, вешают вариации условий освещения, зависящие от типа, направления и количества источников света.

Целью данной работы является исследование влияния условий освещенности и вариаций ракурсов регистраций лиц на эффективность методов распознавания, основанных на МГК, МГК+ЛДА с предварительной обработки входных изображений фильтром Габора.

Поставленная цель определила необходимость решения следующих задач:

- реализация алгоритмов в программной среде Matlab;
- выбор баз данных, которые включают в себя изображения лиц, снятые при различных условиях освещенности, при наличии неравномерного фона, а также при разных ракурсах регистраций.

Идея метода главных компонент (PCA – Principal Component analysis) заключается в разложении изображений по системе собственных векторов, соответствующих наибольшему собственному значению. Коэффициенты, полученные разложением изображения по главным компонентам, используются в качестве признаков для сравнения изображений. Преимуществом МГК является минимальная

размерность словаря признаков, необходимых для обеспечения распознавания с заданной вероятностью.[1]

Метод линейного дискриминантного анализа позволяет осуществить редукцию пространства признаков с целью минимизации внутриклассового и максимизации межклассового расстояния образов объектов в пространстве признаков. В качестве примера на рис.1 приведены 3 первых признака, полученных по МГК и ЛДА.[1]

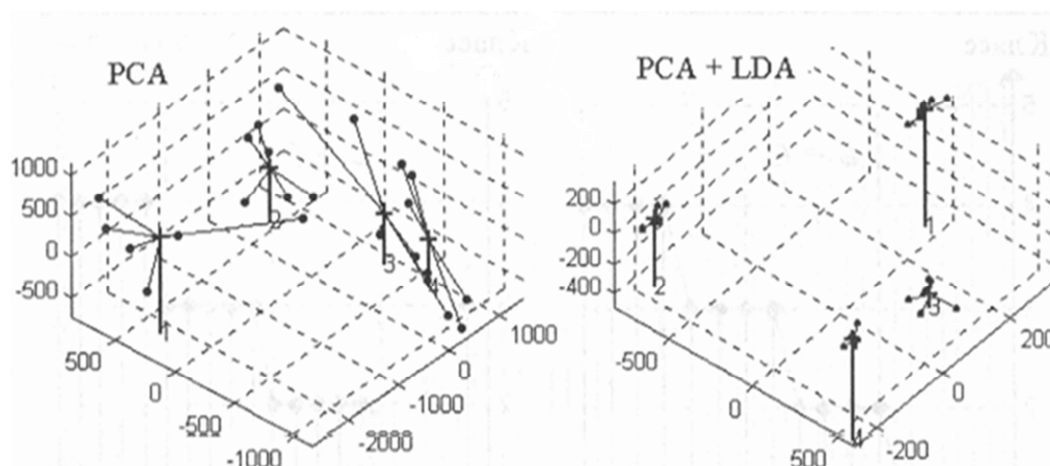


Рис. 1. Отображение 3 первых признаков, полученных по МГК и ЛДА

Фильтр Габора - линейный фильтр, импульсная характеристика которого определяется в виде гармонической функции, умноженной на гауссоиду. Чтобы наиболее полным образом подчеркнуть особенности изображений, используется набор фильтров, импульсные характеристики которых имеют различные параметры:

$$g(x, y; T, \theta, \psi, \sigma, \gamma) = e^{-\left(\frac{x'^2}{2\sigma^2} + \frac{y'^2}{2\sigma^2}\right)} \cos\left(2\pi \frac{x'}{T} + \psi\right),$$

где $x' = x\cos\theta + y\sin\theta$; $y' = -x\sin\theta + y\cos\theta$; T — период гармонической функций; θ — определяет ориентацию нормали параллельных полос функции Габора; ψ — сдвиг фаз; и γ — коэффициент сжатия, характеризующий эллиптичность функции Габора.

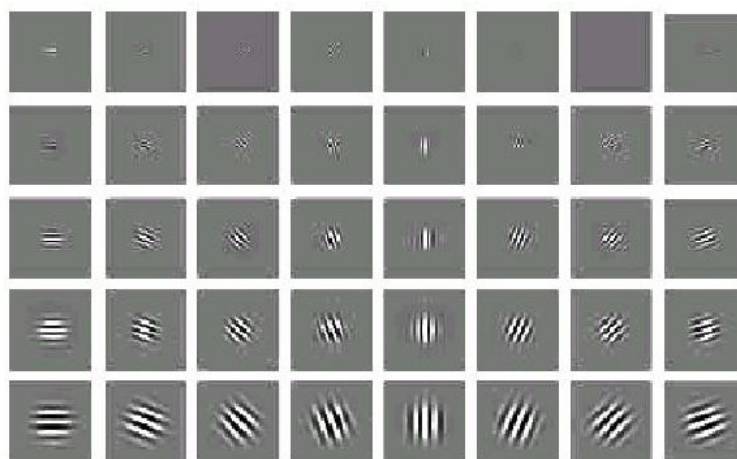


Рис. 2. Примеры импульсных характеристик фильтров Габора, имеющих пять различных размеров и восемь различных ориентаций

В качестве примера на рис. 3 представлен набор импульсных характеристик Габора различными значениями σ (от 0,062 до 0,25) и пространственной ориентации $\theta = \frac{\pi k}{4}$, где $k = 0, 1, 2, \dots, 7$, а также результат применения соответствующих фильтров к изображениям представлен на рис. 3.

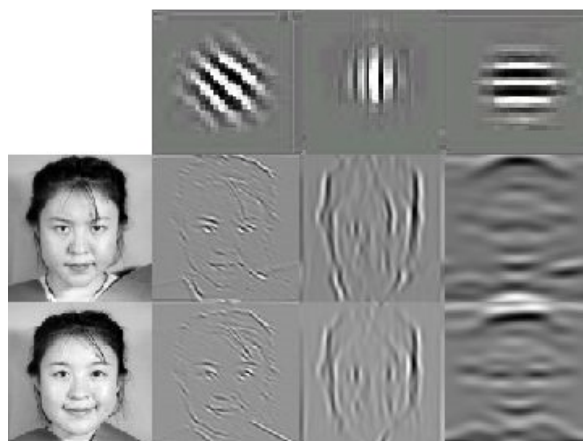


Рис. 3. Пример применения фильтра на изображениях лица

При проведении численных экспериментов использованы базы данных изображений лица ORL (Olivetti Research Laboratory) [2] и FERET(The Facial Recognition Technology) [3].

База данных ORL включала изображения 40 человек по 10 фотографий каждого человека, т.е. в сумме 400 фотографий. Особенностью данной базы данных является, то, что все изображения получены при одинаковых условиях освещения, с небольшими поворотами лица и с небольшими вариациями мимики (см. рис. 4).

База данных FERET включала от 5 до 40 фотографий каждого человека изображений 564 человек. Объем этой базы данных - около 10 тыс. изображений. В отличие от ORL, данная база содержала изображения, полученные при различных условиях освещенности и при различных ракурсах регистрации в пределах от 20 до 60 градусов. На практике чаще используют базу данных FERET для анализа и тестирования различных алгоритмов распознавания.



Рис. 4 Пример изображений из баз данных ORL

Алгоритмы распознавания были реализованы в программе Matlab. Для удобства использования был разработан графический интерфейс пользователя.

Для оценки эффективности алгоритмов распознавания человека по изображениям снятым в одинаковых условиях, база данных ORL была разделена на 2 части. Одна часть использовалась для тренировок – выбора главных компонент. Она содержала 3 изображения лица каждого человека – всего, $3 \cdot 40 = 120$ изображений. Вторая часть базы данных, предназначенная для тестирования, содержала 7 изображений каждого человека - всего $7 \cdot 40 = 280$ изображений.

Отнесение каждого тестового изображения к какому либо классу из базы данных производилась на основе вычисления евклидова расстояния в пространстве признаков. Для построения кривой ошибок устанавливался порог, в соответствии с которым принималось решения об отнесении тестового изображения к одному из классов из базы данных, при котором обеспечивался заданный уровень вероятности ошибочного решения – FAR [1].

В МГК было выбрано 120 компонент. Как следует из рис. 6 вероятность правильного распознавания (Verification Rate) при вероятности ложного распознавания FAR=1% составила 86.79%.

После редукции пространства признаков методом ЛДА размерность пространства признаков сократилась до 40 компонент. При этом, как следует из рис. 6 вероятность правильного распознавания при вероятности ложного распознавания 1% составила 90%.

Следующий эксперимент заключался в предварительной обработке изображений набором из 40 фильтров Габора. Размерность 40 полученных изображений была уменьшена в 8 раз, после чего эти изображения были представлены в виде векторов. Путём их объединения получен один вектор признаков для каждого входного изображения, причем размерность вектора для каждого исходного изображения составила 10240. По МГК выбрано 100 компонент. При дальнейшей редукции по методу ЛДА была получена размерность пространства признаков 35 компонент. В методе с использованием набора фильтров Габора и последующего ЛДА вероятность правильного распознавания при вероятности ложного распознавания FAR=1% составила 96% (см рис. 5).

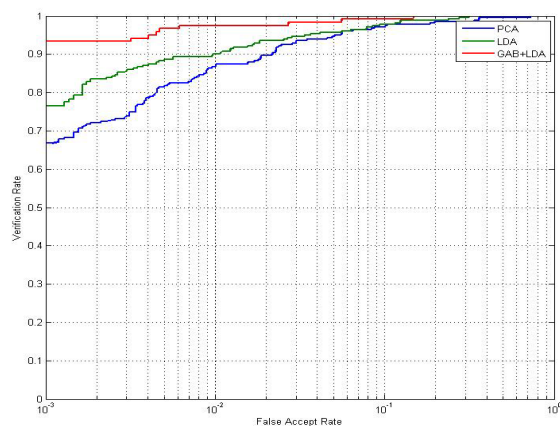


Рис. 5. Кривая ошибок для МГК, ЛДА и Габор+МГК+ЛДА

В следующей серии экспериментов к базе данных ORL были добавлены изображения из базы данных FERET, полученные при одинаковых условиях регистрации. По этой объединенной базе данных произведена оценка эффективности распознавания с использованием методов МГК, ЛДА и Габор+МГК+ЛДА (рис. 7).

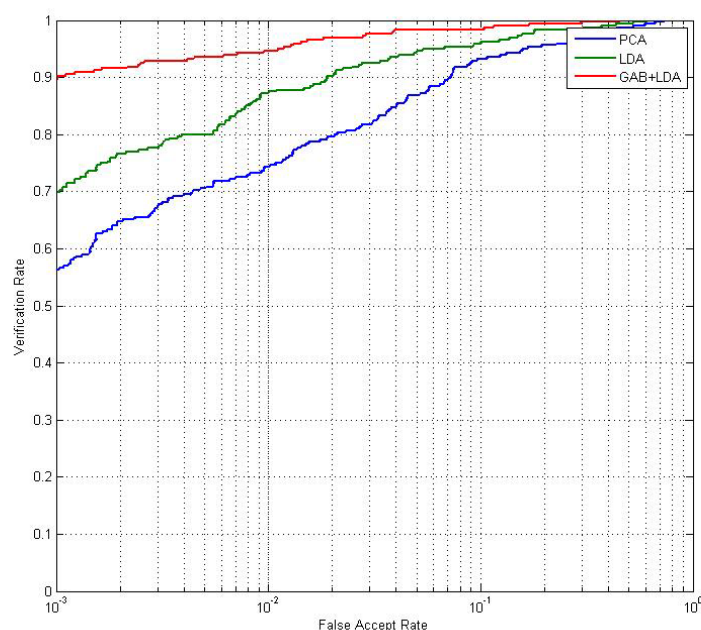


Рис. 6. Кривая ошибок для МГК, ЛДА и Габор+МГК+ЛДА (для баз данных ORL+FERET)

Из результатов, представленных на рис. 6, следует, что эффективность методов распознавания уменьшилась на 6%.

Чтобы проанализировать влияние изменения ракурсов база, данных FERET разделена на 2 части. Первая часть, содержащие все фронтальные изображения, использовалась в качестве тренировочного набора. Вторая часть в которую были включены все изображения снятые под углом от 20 до 60 градусов, использовалась для тестирования. Как следует из рис. 8 вероятность правильного распознавания снизилась более чем на 40% для всех методов.

Следует отметить, что для метода Габор+МГК+ЛДА вероятность правильного распознавания при вероятности ложного распознавания FAR=1% составила 50%.

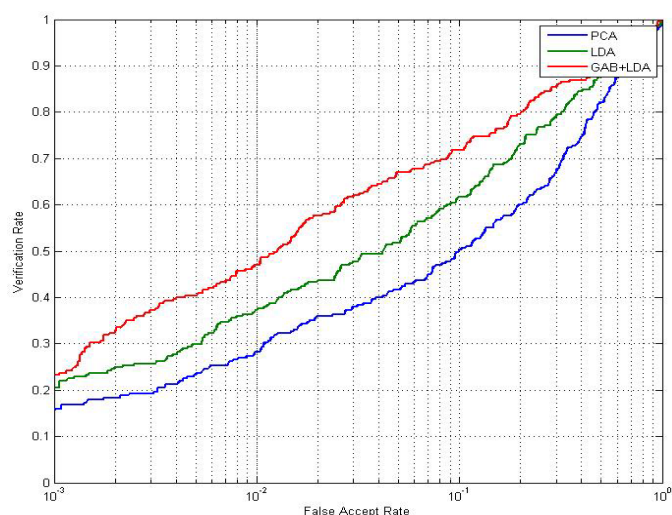


Рис. 7. Кривая ошибок для МГК, ЛДА и Габор+МГК+ЛДА (влияние изменения ракурса)

Для оценки эффективности методов распознавания изображений, снятых при различных условиях освещения при наличии неравномерного фона, использовалась база данных FERET. В качестве тренировочного набора использованы все фронтальные изображения. А в качестве тестирующего набора использовали фронтальные изображения с разными освещенностями и при наличии неравномерного фона.

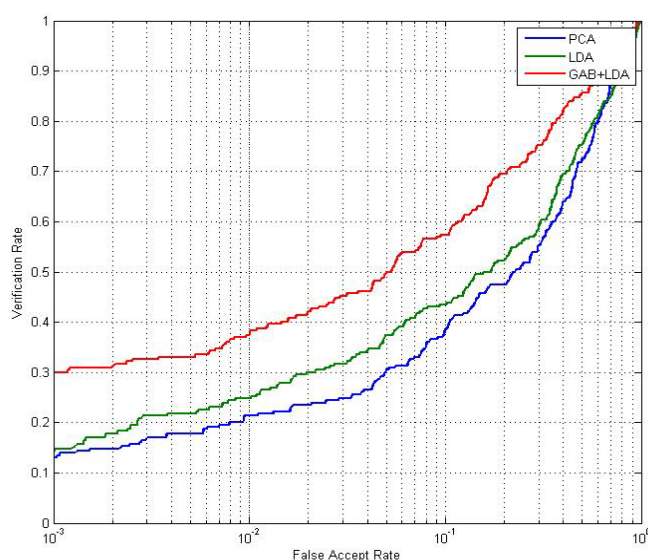


Рис. 8. Кривая ошибок для МГК, ЛДА и Габор+МГК+ЛДА (влияние внешних условий)

Как видно из рис. 8, вероятности правильного распознавания для всех методов снизилась более чем на 50%. Наилучший результат показал метод Габор+МГК+ЛДА.

Для данного метода вероятность правильного распознавания при вероятности ложного распознавания FAR=1% составила 40%.

На основе проведенных исследований можно сделать следующие выводы:

- 1) эффективность распознавания МГК возрастает при использовании ЛДА, т.к. этот метод минимизирует внутриклассовое и максимизирует межклассовое расстояние в пространстве признаков;
- 2) вариации освещенности и наличие неравномерного фона оказывает более сильное влияние на эффективность распознавания, чем изменение ракурсов изображений лиц;
- 3) при предварительной обработке фильтром Габора эффективность распознавания возрастает, т.к. происходит подчеркивание локальных особенностей текстуры изображений. Недостатком этого метода является высокая трудоёмкость. Например, при обработке изображений на ЭВМ Intel i5 3470, 3.2ГГц для реализации МГК потребовалось 5сек., а для метода Габор+МГК+ЛДА - около 30сек.

Список литературы

- [1] Кухарев Г.А. Биометрические системы; Санкт-Петербург: Политехника, 2001. 240 с.
- [2] База данных ORL. Режим доступа: <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html> (дата обращения 25.3.2015).
- [3] База данных FERET: Режим доступа http://www.itl.nist.gov/iad/humanid/feret/feret_master.html (дата обращения 26.3.2015).