

Методы выделения информативных признаков изображений лица в задаче распознавания лица

Петр Базанов, Оксана Джосан
кафедра автоматизация систем и вычислительных комплексов, Факультет ВМиК,
Московский Государственный Университет им. Ломоносова, Москва, Россия
{pbazanov, oxanad}@mail.ru

Аннотация

Данная работа относится к области компьютерного зрения и распознавания изображений. Представлены три различных способа извлечения информативных признаков изображений лица, производящие анализ целостных признаков, анализ локальных признаков и анализ структурных признаков изображений лица. Предложены эффективные модификации методов извлечения черт изображений на основе метода главных компонент, вейвлетов, скрытых марковских моделей. Поставлены эксперименты по распознаванию лица над нормализованной базой данной изображений, по результатам которых установлены достоинства и недостатки методов.

Ключевые слова : выделение признаков изображения, распознавание изображений лица, метод главных компонент, вейвлеты Габора, скрытые Марковские модели.

1. ВВЕДЕНИЕ

Технологии автоматического обнаружения и распознавания лица используются в ряде современных систем компьютерного зрения: биометрическая идентификация, человеко-машинный интерфейс [6], зрение роботов [7], компьютерная анимация, видеоконференции. Основное отличие данных приложений между собой – это целевые классы, которые являются объектами распознавания. Целевыми классами в задачи распознавания лица могут являться: лицо, лицо с элементами перекрытий, морда животного, лицо человека, живое лицо человека, мимика лица, пол, раса, личность человека. Выбор одного из таких целевых классов определяет специфику алгоритма распознавания, остальные классы являются второстепенными и скорее играют роль признаков при распознавании целевого класса.

2. ПОСТАНОВКА ЗАДАЧИ РАСПОЗНАВАНИЯ ЛИЦА

В данной работе целевым классом в процессе распознавания являлся класс личности пользователя, тем самым решалась задача биометрической аутентификации пользователя по изображению лица. Входным классом для алгоритма распознавания был класс нормализованных лиц, то есть подразумевалось то, что в изображении выделена область лица, выполнена нормализация освещенности, определен ракурс изображения лица и произведено согласование размеров изображений.

Задача биометрической системы идентификации состоит в том, чтобы принять правильное решение в пользу одной из гипотез:

H_0 - зарегистрированное лицо, которое система должна принять;

H_1 - незарегистрированное лицо, которое система должна отвергнуть.

Тогда ошибка первого рода, при которой мы отвергаем гипотезу H_0 и принимаем H_1 , будет характеризовать вероятность ошибочных отказов – FRR (false rejection rate), то есть процент зарегистрированных пользователей, отвергнутых системой. Ошибка второго рода, при которой мы принимаем H_0 и отвергаем H_1 , будет характеризовать вероятность ошибочных подтверждений – FAR (false acceptance rate), то есть процент незарегистрированных пользователей, пропущенных системой.

Метод распознавания изображения включает следующие процедуры:

- преобразование исходного изображения в начальное представление алгоритма (предобработка, сегментация и математические преобразования)
- выделение ключевых характеристик изображения
- механизм обобщения, сравнения, классификации признаков изображения, и принятия решений: кластерная модель, метрика, нейронная сеть.

3. МЕТОД СОБСТВЕННЫХ ФИЛЬТРОВ ДЛЯ РАСПОЗНАВАНИЯ ЛИЦА

3.1 Построение фильтров методом главных компонент

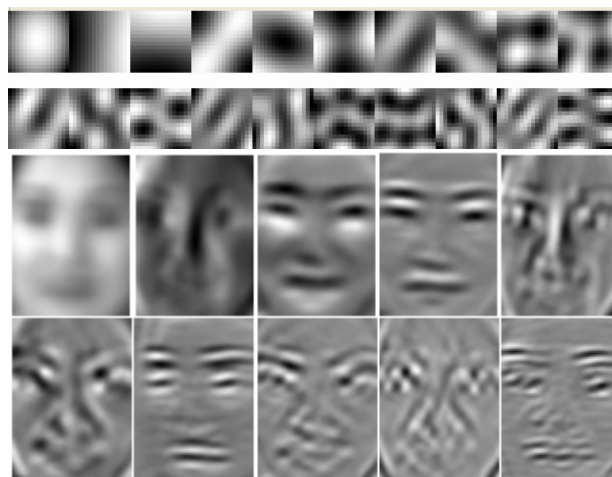


Рис. 1 Собственные фильтры, полученные в результате самоорганизации на БД изображений лиц и результат выделения целостных признаков.

Для определения целостных признаков на изображении использовался метод главных компонент в следующей модификации. На изображении лица случайным образом выбирались блоки размером 16x16. Собирались статистика блоков изображения - 10000 блоков. По данной статистике – строились собственные фильтры. В дальнейшем полученные фильтры использовались при распознавании для извлечения черт на изображении.

Более конкретно, пусть x – вектор размерности $N \times 1$, соответствующий, например, изображению $N=16 \times 16$ и задана статистика векторов x (например 200 изображений размерности $16 \times 16 = 256 \times 1$).

u – вектор той же размерности (256×1), что и x на него мы будем проектировать входные вектора x . [1][2]

Пусть $\bar{x} = E\{x\}$

$a = (x - \bar{x})^T u = u^T (x - \bar{x})$ - проекция X на U .

$\sigma^2 = E[a^2] = E[(u^T (x - \bar{x}))((x - \bar{x})^T u)] = u^T E[(x - \bar{x})(x - \bar{x})^T] u$ - дисперсия для данной статистики образов и есть функция от u :

$\psi(u) = \sigma^2 = u^T R u$, где $R = E[(x - \bar{x})(x - \bar{x})^T]$ - матрица ковариации размерности $N \times N$.

Задача анализа главных компонент выбрать базис пространства $u = (u_1, u_2, \dots, u_k)$ так, чтобы вариация (дисперсия) проекции была минимальной. Минимум этой функции достигается решением задачи на собственные значения $Ru = \lambda u$ (доказательство проведено в книге [3]). Для этого используется метод Якоби.[3]. На Рис. 1 Собственные фильтры, полученные в результате самоорганизации на БД изображений лиц и результат выделения целостных признаков. показан результат применения метода главных компонент к блокам изображений, результат их самоорганизации в определенные типы фильтров, а также результат фильтрации – первые 10 главных компонент.

3.2 Сравнение образов на основе нейросетевого классификатора

Для данного алгоритма [1] число собственных фильтров варьировалось в диапазоне от 4-20 – лучшие результаты были получены для 8 собственных фильтров. Соответственно для классификации образов использовались 8 нейронных сетей архитектуры типа двухслойный перцептрон, с количеством нейронов 16×20 – входной слой, 30 – скрытый слой, 50 – выходной слой”. Входной слой выбирался под размерность главных компонент. Число нейронов скрытого слоя зависело от регистрируемой статистики. Выходной слой соответствовал количеству целевых классов – регистрируемых персон.

4. МЕТОД СРАВНЕНИЯ ЭЛАСТИЧНЫХ ГРАФОВ НА ОСНОВЕ ВЕЙВЛЕТОВ ГАБОРА

4.1 Фильтры Габора

Для определения локальных признаков на изображении использовались фильтры на основе вейвлетов

Габора. Разложение Габора было предложено в качестве инструмента анализа текстур. Этот подход был мотивирован современными представлениями о психофизической природе человеческого зрения. Декомпозиция Габора представляет собой так называемый “направленный микроскоп” с набором ориентаций и масштабом.

В результате преобразование Габора спектральная плоскость разбивается на m частотных и n ориентационных диапазонов [4]:

$$\hat{g}_{ij}(\omega) = \exp\left(-\frac{1}{2}(\omega - \omega_{ij})^T R_j \Sigma_{ij}^{-1} R_j^T (\omega - \omega_{ij})\right) \quad (1)$$

- вейвлет, соответствующий (i,j) ориентации и масштабу

где $1 \leq i \leq m$, $1 \leq j \leq n$, R_j – матрица поворота, а Σ_{ij} – диагональная матрица.

$R_j = \begin{pmatrix} \cos(\phi_j) & \sin(\phi_j) \\ -\sin(\phi_j) & \cos(\phi_j) \end{pmatrix}$ положительно определенная матрица, главные оси которой определяют эллипс, ориентированный в соответствии с направлением, задаваемым j . Конкретнее,

$$R_j = \begin{pmatrix} \cos(\phi_j) & \sin(\phi_j) \\ -\sin(\phi_j) & \cos(\phi_j) \end{pmatrix} \quad (2)$$

$$\Sigma_{ij} = \begin{pmatrix} \sigma_{\phi_j}^2 & 0 \\ 0 & \sigma_{\omega_j}^2 \end{pmatrix}$$

$$\omega_{ij} = R_j \begin{pmatrix} \omega_{\phi_j} \\ 0 \end{pmatrix} = \begin{pmatrix} \cos(\phi_j) \omega_{\phi_j} \\ \sin(\phi_j) \omega_{\phi_j} \end{pmatrix}$$

где ω_{ij} – центральная частота фильтра.

Благодаря разложению Габора, вычисление соответствующей функции точечного распределения можно провести аналитически. Кроме описанного разложения можно также использовать другие разложения с использованием логарифмическо-полярных отображений.

Точечный отклик, соответствующий \hat{g}_{ij} , вычисляется по формуле:

$$g_{ij}(x) = \exp\left(-\frac{1}{2} x^T R_j \Sigma_{ij}^{-1} R_j^T x\right) \exp(i 2\pi \omega_{ij}^T x). \quad (3)$$

Займемся определением остальных параметров. Разложение по n направлениям и m масштабам выбирается так, чтобы соседние фильтры пересекались под одинаковыми значениями $1/\sqrt{e}$ вдоль своих главных осей. Направления выбираются эквидистантными, вследствие чего значения σ_{ϕ_j} постоянны для всех j :

$$\sigma_{\phi_j} = \sigma_{\phi} = \frac{\pi}{2n}. \quad (4)$$

Угол между фильтрами равен π/n . Направления задаются формулой:

$$\omega_{\phi_j} = 2\sigma_{\phi} (j - 1). \quad (5)$$

Радиальные диапазоны частот распределены так, что частотный диапазон удваивается на каждом шаге, покрывая интервал от $\omega_{r_{\min}} > 0$ до $\omega_{r_{\max}} < 1/2$ в нормализованных частотах, где единица соответствует частоте Найквиста.

Определим промежуточную переменную:

$$\sigma_0 = \frac{\omega_{r_{\max}} - \omega_{r_{\min}}}{2(2^m - 1)} \quad (6)$$

Тогда радиальные центры задаются формулами:

$$\begin{aligned} \omega_{r_i} &= \omega_{r_{\min}} + \sigma_0(1 + 3(2^{i-1} - 1)) \\ \sigma_{r_i} &= \sigma_0 2^{i-1} \end{aligned} \quad (7)$$

В экспериментах параметры максимальной и минимальной центральной частоты вейвлетов Габора выбирались: $\omega_{r_{\min}} = 1/16$, $\omega_{r_{\max}} = 1/4$

Ориентация вейвлетов выбиралась эквидистантно с круговой частотой - для 6 ориентаций, то есть центры эллипсов, характеризующие частотную область действия вейвлета, задавались со следующими ориентациями

$$\omega_{\phi_j} = 2 \times \sigma_{\phi}(j-1) = \left\{0, \frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}, \frac{5\pi}{6}\right\}$$

Визуализация показана на Рис. 2 Фильтры Габора и результат их применения локальные признаки изображения.

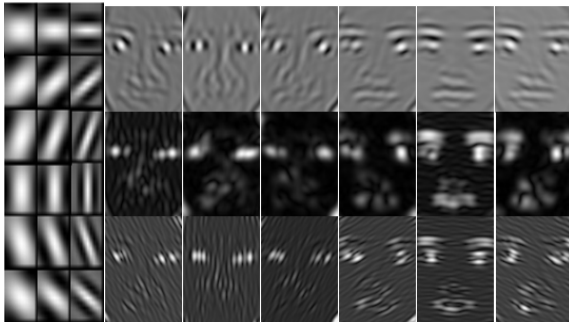


Рис. 2 Фильтры Габора и результат их применения локальные признаки изображения

4.2 Сравнение образов на основе нейросетевого классификатора

Для данного алгоритма использовались все 18 фильтров. Соответственно для классификации образов использовались 18 нейронных сетей архитектуры типа двухслойный перцептрон, с количеством нейронов 20x28 – входной слой, 30 – скрытый слой, 50 – выходной слой. Выбор количества нейронов в слое обуславливался аналогичными соображениями как для метода собственных фильтров.

5. МЕТОД РАСПОЗНАВАНИЯ НА ОСНОВЕ СКРЫТЫХ МАРКОВСКИХ МОДЕЛЕЙ

Скрытые Марковские Модели (СММ) – это статистическая модель, используемая для описания

статических характеристик сигнала [5]. Каждая модель $\lambda = (A, B, \pi)$ представляет собой набор N состояний $S = \{S_1, S_2, \dots, S_N\}$, между которыми возможны переходы.

5.1 Встроенные Скрытые Марковские Модели

В работе используются встроенные СММ, в которых каждое состояние представлено суперсостоянием, имеющим внутри себя набор состояний. В модели пять суперсостояний, соединенных в линейную СММ, количество встроенных состояний 3-6-6-6-3, соответствующих последовательности лоб – глаза – нос – рот – подбородок, внутренние состояния в каждом суперсостоянии объединены также в линейную СММ:

Внешняя модель задается следующими параметрами:

$$\begin{aligned} S_0 &= \{S_{0i}, i = 1, 2, \dots, N_0\} \\ \pi_0 &= \{\pi_{0i}, i = 1, 2, \dots, N_0\} \\ A_0 &= \{a_{0ij}, i, j = 1, 2, \dots, N_0\} \end{aligned}$$

где S_0 – набор суперсостояний, π_0 – определяет начальное состояние, A_0 – вероятность переходов по суперсостояниям

Каждая внутренняя модель: Рис. 3 Схема сегментирования изображения на основе скрытой Марковской модели и Структурные признаки изображения, выделенные скрытой Марковской моделью

$$\begin{aligned} S_1^k &= \{S_{1i}^k, i = 1, 2, \dots, N_1^k\} \\ \Pi_1^k &= \{\pi_{1i}^k, i = 1, 2, \dots, N_1^k\} \\ A_1^k &= \{a_{1ij}^k, i, j = 1, 2, \dots, N_1^k\} \\ B^k &= \{b_j^k(O_{t_0 t_1})\} \\ b_j^k(O_{t_0 t_1}) &= \sum_{m=1}^M c_{im}^k N(O_{t_0 t_1}, \mu_m^k, U_m^k) \\ \Lambda^k &= \{\Pi_1^k, A_1^k, B^k\} \end{aligned}$$

где N – гауссиан, μ – вектор средних значений, U – матрица ковариации, $b_j^k(O_{t_0 t_1})$ – мера близости наблюдения к эталонному, c – коэффициенты разложения по гауссианам (в работе использовалось разложение по 3-м гауссианам)

$$\begin{aligned} \Lambda &= \{\Lambda^1, \Lambda^2, \dots, \Lambda^{N_0}\} \\ \text{Модель СММ:} \quad \lambda &= (\Pi_0, A_0, \Lambda) \end{aligned}$$

5.2 Подготовка входных данных.

Из изображения выделяется последовательность наблюдений: изображение сканируется окошком (16x16) слева на право сверху вниз с перекрытием 14 пикселей. Получается набор блоков.

Для каждого блока получаем коэффициенты дискретного косинусного разложения.

$$C(u, v) = \frac{1}{4} w(u)w(v) \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x, y) \cos\left[\frac{\pi(2x+1)u}{MN}\right] \cos\left[\frac{\pi(2y+1)v}{MN}\right]$$

где x, y - пространственные координаты пикселя, $f(x, y)$ - значения пикселей исходного макроблока, u, v - координаты пикселя в частотном представлении, $w(u) = 1/\sqrt{2}$ при $u=0$, в остальных случаях $w(u)=1$, $w(v) = 1/\sqrt{2}$ при $v=0$, в остальных случаях $w(v)=1$, $N=16$, $M=16$ – ширина и высота окошка сканирования. Первые девять коэффициентов составляют входной вектор.

5.3 Построение Встроенной Скрытой Марковской Модели

- Равномерное разбиение входных векторов по состояниям
- Инициализация модели
- Применение разбиения Витерби для встроенной СММ[5]. На этом этапе происходит перераспределение входных векторов по состояниям.
- Проводим оценку полученной модели, и либо возвращаемся к предыдущему шагу, либо говорим, что искомая модель построена.

Фактически происходит сегментация изображения: входные вектора разделяются на группы, т.е. каждый вектор относится к некоторому внутреннему состоянию.

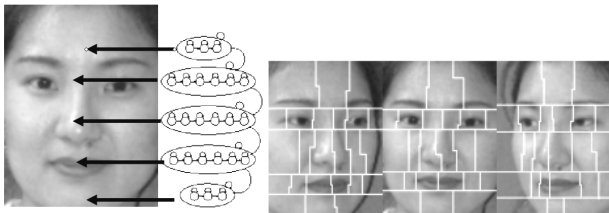


Рис. 3 Схема сегментирования изображения на основе скрытой Марковской модели и Структурные признаки изображения, выделенные скрытой Марковской моделью

5.4 Процесс распознавания в Скрытой Марковской Модели

- Системе подается на вход некоторое изображение
- Система строит из него входную последовательность
- Система строит вероятности построения такой входной последовательности всеми моделями
- Выдается наиболее вероятный ответ и вероятность соответствия

6. ЭКСПЕРИМЕНТЫ ПО РАСПОЗНАВАНИЮ ЛИЦА НА НОРМАЛИЗОВАННОЙ БД

Эксперимент по идентификации лица производился на базе HongIk [1]. Данная база содержит изображения 250 персон по 10 изображений для каждой персоны. Данная БД отвечает качеству нормализации образов: лица выделены, согласованы размеры, нормализовано освещение.

В тестах использовались следующие выборки. Рисунок 4 Обучающая выборка и тестирующая выборка 150=50*3 – обучающая выборка для 50 регистрируемых персон на трех разных ракурсах. 350=50*7 – тестирующая

выборка для 50 регистрируемых персон на остальных изображениях. 1000=10*10 – выборка отрицательных незарегистрированных лиц.



Рисунок 4 Обучающая выборка и тестирующая выборка

Метод	FAR	FRR
Метод собственных фильтров + двуслойный перцептрон	2,9 %	36,5 %
Метод вейвлетов Габора + двуслойный перцептрон	2,9%	30%
Скрытая марковская модель	2,9%	31 %

7. ЗАКЛЮЧЕНИЕ

Предварительные результаты экспериментов над алгоритмами распознавания свидетельствуют:

- 1) алгоритм на основе метода главных компонент и нейронных сетей позволяет распознавать лицо даже при наличии значительных шумов и внешних объектов (очки). Тем не менее, по точности он уступает методу на основе локальных признаков и Марковских моделей. Это объясняется тем, что главные компоненты несут в себе главную информацию о признаках некоторого обобщенного лица, нежели чем позволяют определять уникальные признаки лица конкретного человека.
- 2) алгоритм на основе вейвлетов и нейронных сетей имеет наилучшую точность, и позволяет наиболее точно определять уникальные локальные черты лица человека. Тем не менее, алгоритм более чувствителен к шумам на изображении, нежели метод главных компонент.
- 3) алгоритм на основе марковских моделей показывает точность сравнимую с анализом локальных признаков. Алгоритм позволяет определять структурные особенности лица и учитывать характер

искажений лица. Данный метод более устойчив к шумам по сравнению с методом локальных признаков

В настоящий момент ведется работа по созданию обобщенного алгоритма распознавания, использующего основные достоинства предложенных методов распознавания.

8. БИБЛИОГРАФИЯ

- [1] Peter Bazanov, Tae-Kyun Kim, Seok Cheol Kee, Sang Uk Lee "Hybrid and Parallel Face Classifier based on artificial neural networks and principal component analysis" Image Processing. 2002. Proceedings. 2002 International Conference on , Volume: 1 , 2002 Page(s): 916 -919 (International Conference of Image Processing 2002 Rochester , New York September 22-25 2002
- [2] P. McGuire, G.M.T. D'Eleuterio "Eigenpaxels and a Neural-Network Approach to Image Classification" IEEE Transactions on Neural Networks, Vol.12, NO. 3, May 2001.
- [3] Simon Haykin "Neural Networks A Comprehensive Foundation", стр.363-370 , 1994, Macmillan College Publishing Company, USA 1994
- [4] Benoit Duc, Stefan Fischer, Joseph Bigun "Face Authentication with Gabor Information on Deformable Graphs" IEEE Transaction on Image Processing Vol 8, NO 4, April 1999
- [5] A.V.Nefian and M.H.Hayes, "Face recognition using an embedded HMM", in Proceedings of the IEEE Conference on Audio and Video-based Biometric Person Authentication, pp. 19-24, 1999
- [6] Vezhnevets V. Face and facial feature tracking for natural human-computer interface // Труды конференции Графикон 2002.
- [7] И.А.Матвеев, А.Б.Мурынин "Принципы построения системы стереоскопического зрения для управления движением робота", Известия РАН. Теория и системы управления. 2002

Об авторах

Петр Валерьевич Базанов - аспирант факультета вычислительной математики и кибернетики Московского Государственного Университета им. М.В. Ломоносова

Адрес: Москва, 119899, Воробьевы горы, МГУ, 2-й учебный корпус, факультет ВМиК, кафедра автоматизации систем и вычислительных комплексов.

Его электронный адрес: pbazanov@mail.ru

Оксана Джосан, студентка 4-го курса факультета вычислительной математики и кибернетики Московского Государственного Университета им. М.В. Ломоносова

Адрес: Москва, 119899, Воробьевы горы, МГУ, 2-й учебный корпус, факультет ВМиК, кафедра автоматизации систем и вычислительных комплексов.

Ее электронный адрес: oxanad@mail.ru

Face Feature Extraction Methods For the Task of Face Recognition

Abstract

The paper is devoted to the field of computer vision & face image recognition. In this paper it is presented three different methods of face feature extraction methods for the task of face recognition. These methods perform extraction of holistic, local & structured features of face object. It is presented efficient modifications of well-known methods of image feature extraction based on principal component analysis, Gabor wavelets, hidden Markov models. Face recognition methods has been tested on normalized face database as a result it is outlined advantages & disadvantages of these methods.

Keywords: Image Feature Extraction Methods, Face Recognition, Principal Component Analysis, Gabor Wavelets, Hidden Markov Models

About the authors

Peter Valeryevich Bazanov is a Ph.D. student at Moscow State University, Department of Computational Mathematics & Cybernetics. His contact email is pbazanov@mail.ru.

Oxana Vasilevna Josan is a 4-year student at Moscow State University, Department of Computational Mathematics & Cybernetics. Her contact email is oxanad@mail.ru.