

УДК 004.93'1

Комбинированный метод обнаружения и распознавания лиц в реальном режиме

*Вольнец М.Ю., магистр
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедра «Программное обеспечение ЭВМ и информационные технологии»*

*Научный руководитель: Майков К.А., д.т.н, профессор
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана
maikov@mx.bmstu.ru*

Введение

Задача распознавания лиц характеризуется рядом факторов, а именно: высокой вариативностью лиц, обусловленной анатомическими особенностями людей; различным уровнем освещенности объектов, зависящих от типа, количества и характеристик направленности источников света; необходимостью обнаружения лиц, имеющих различное пространственное положение.

Существующие потребности в создании систем, реализующих решение подобных задач, накладывают жесткие ограничения на скорость работы алгоритмов, которые должны работать в режиме, близком к реальному времени.

Для успешного решения задачи по распознаванию лиц обеспечение высокой скорости работы должно также сочетаться с малым (порядка 5%) количеством ложных распознаваний. В системах, реализующих существующие методы распознавания, при увеличении уровня распознаваний свыше 90% наблюдается существенный рост числа ложных решений, что затрудняет их практическое использование [1]. Прежде чем распознавать лицо, необходимо убедиться в его присутствии на изображении. Для чего применяются известные методы обнаружения и распознавания лиц на изображениях (метод главных компонент [12], нейронные сети [3], метод опорных векторов [4]). Результативность применения метода определяется спецификой решаемой задачи. Поэтому построение метода распознавания лиц, обеспечивающего высокий уровень достоверности решения при отсутствии ограничений на исходные изображения, является весьма актуальной задачей.

Целью данной работы является разработка методов распознавания и обнаружения лиц, обеспечивающих повышение достоверности распознавания объектов анализа,

снижение уровня ложных распознаваний, уменьшение времени обучения классификатора и времени предварительной обработки изображения.

Этапы решения задачи обнаружения и распознавания лиц

В задаче обнаружения и распознавания лиц следует выделить несколько этапов:

- извлечение данных из источников различного типа (видеофайл, последовательность изображений);
- последующая обработка извлеченной информации;
- RGB2GRAY- преобразование изображения из цветовой модели RGB в оттенки серого. Применяемого по причине отсутствия необходимости в использовании информации о цвете объектов анализа;
- Изменение размера (уменьшение размерности кадра в пикселях). и передача ее на следующий этап;
- обнаружение лиц на изображении и выявление характеристических черт лица;
- распознавание лиц на изображении путем сравнения с обучающим набором.

Методы обнаружения лиц

Большинство известных методов обнаружения лиц (метод адаптивного усиления [2], нейронные сети [3], машины опорных векторов [4], линейный дискриминантный анализ [5]) на изображении могут быть представлены двумя следующими категориями [3]:

1. методы, основанные на построении некоторого набора правил для обнаружения лица на изображении [6], одним из представителей которых является метод Виолы-Джонса [2]. Данные методы используют инвариантные свойства изображений лиц и включают в себя два основных этапа:

- обнаружение элементов и особенностей, характерных для изображения лица;
- анализ обнаруженных особенностей, формирование решения о количестве и пространственном положении лиц;

2. методы, в которых изображению (или его фрагменту) ставится в соответствие вычисленный вектор признаков, используемый в последствие для классификации изображений на два класса - лицо/не лицо. Каждый пиксель анализируемого изображения становится компонентом вектора, трансформируя черно-белое изображение в вектор пространства [7]. К таким методам относятся методы опорных векторов[4] и линейный дискриминантный анализ[5].

Выбор категории и метода для обнаружения зависит от ограничений и условий задачи обнаружения лиц. В качестве ограничений, влияющих на выбор метода решения задачи, следует выделить:

- наличие или отсутствие ограничений на возможные искусственные помехи на лице;
- пространственные характеристики положения лиц;
- цветность изображения;
- масштаб лиц и разрешение изображения;
- количество лиц на изображении;
- условия освещенности объектов;
- приоритет в минимизации ложных обнаружений или в количестве обнаруженных лиц.

Для решения задачи обнаружений лиц построим метод, представляющий собой комбинацию базовых компонентов метода адаптивного усиления (AdaptiveBoosting) [2] и метода нейронных сетей [3].

Метод адаптивного усиления (Adaptive Boosting)

В основе рассматриваемого метода лежит метод Виолы-Джонса, позволяющий обрабатывать изображения с довольно большой скоростью по сравнению с другими методами обнаружения лиц, при этом достигая высокого уровня распознавания [2]. Метод базируется на следующих трех основных принципах:

- использование интегрального представления изображения для быстрого вычисления признаков, используемых детектором;
- использование признаков Хаара, с помощью которых происходит поиск нужного объекта (лица и его черт) и последующий бустинг (от англ. boost – улучшение, усиление) для выбора наиболее подходящих признаков искомого объекта на анализируемой части изображения;
- комбинирование классификаторов, полученных на предыдущем этапе, в каскад для быстрого исключения окон, не содержащих изображение лица.

К преимуществам метода адаптивного усиления относится скорость решения задачи обнаружения, что дает возможность использовать его в режиме реального времени.

Однако, как показано в [2], алгоритмическая реализация метода адаптивного усиления предъявляет высокие требования к аппаратным и временным ресурсам, а также накладывает жесткие ограничения на пространственное положение лица и его освещенность.

Метод обнаружения с использованием нейронных сетей

На вход нейронной сети подается участок изображения, после чего алгоритм выносит решение о наличии или отсутствии лица на входном участке. Для обнаружения лица, размер которого больше входного участка, анализируемая область постепенно (в несколько итераций) уменьшается в размере, и обнаружение выполняется для каждой полученной структуры изображения.

Преимущественной особенностью данного метода является возможность обработки лицевых структур при различном угле наклона относительно вертикальной оси. В то же время алгоритм имеет ряд недостатков: ложное обнаружение объектов, отдаленно напоминающих лицо, и относительно большое время обнаружения лица не позволяет применять метод в режиме реального времени.

Данный метод требует предварительной обработки изображения, заключающийся в приведении анализируемых на изображении лиц к базовым размерам, а так же выравнивании яркости и контрастности на изображениях.

Представим лицо в виде эллипса: оси анализируемого объекта и оси лиц на тестовых изображениях должны отличаться не более чем на 12%, что позволяет уменьшить различия в положении, ориентации и масштабе лиц между тестовым изображением и изображениями из обучающего набора [7]. Объекты считаются расположенными на одной прямой, если квадратичное расстояние между парой соответствующих признаку точек минимально. Это может быть достигнуто путем масштабирования и поворота лица на изображении, что может быть описано в виде формулы

$$\begin{pmatrix} X' \\ Y' \end{pmatrix} = \begin{pmatrix} s \cos \theta & -s \sin \theta \\ s \sin \theta & s \cos \theta \end{pmatrix} * \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} = \begin{pmatrix} a & -b & t_x \\ b & a & t_y \end{pmatrix} * \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}. \quad (1)$$

Для нескольких наборов координат, формула (1) принимает следующий вид:

$$\begin{pmatrix} X_1 & -Y_1 & 1 & 0 \\ Y_1 & X_1 & 0 & 1 \\ X_2 & -Y_2 & 1 & 0 \\ Y_2 & X_2 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} * \begin{pmatrix} a \\ b \\ t_x \\ t_y \end{pmatrix} = \begin{pmatrix} X_1' \\ Y_1' \\ X_2' \\ Y_2' \\ \vdots \end{pmatrix}. \quad (2)$$

В результате преобразования по формуле (2) пространственное расположение лица на тестовом изображении трансформируется к виду, наиболее близкому к положению лиц на изображениях из обучающего набора.

В [3] показано, что различия в изображениях с различной яркостью могут быть убраны путем применения ряда методик, в частности, для начального выравнивания яркости на изображении строится функция, принимающая яркостные значения на всем изображении внутри овального окна. Пиксели вне окна отображают фон, поэтому их интенсивность не учитывается. Параметры a , b , c этой функции должны удовлетворять условию, показанному в формуле (3):

$$(X \ Y \ 1) * \begin{pmatrix} a \\ b \\ c \end{pmatrix} = I(X, Y), \quad (3)$$

где $I(x,y)$ - яркость пикселя (x,y) .

Для увеличения контрастности анализируемых областей применяется бинаризация методом Отсу [8], где значения яркостей пикселей изображения можно рассматривать как случайные величины, а их гистограмму – как оценку плотности распределения вероятностей принадлежности объекту либо фону, либо искомому объекту. Если плотности распределения вероятностей известны, то определяется оптимальный порог для сегментации изображения на два класса c_0 и c_1 (объекты и фон). Порог между классами задается таким образом, чтобы каждый из них был как можно более «плотным», что достигается минимизацией внутриклассовой дисперсии, определяющейся как взвешенная сумма дисперсий двух классов

$$\sigma_w^2 = w_1 * \sigma_1^2 + w_2 * \sigma_2^2, \quad (4)$$

где σ_w – внутриклассовая дисперсия, σ_1 и σ_2 – дисперсии, а w_1 и w_2 — вероятности принадлежности фрагмента анализируемого изображения либо к фону, либо к искомому объекту соответственно.

Комбинированный метод

Преимуществом метода AdaBoost по сравнению с другими методами обнаружения лиц [2] является высокая скорость обнаружения лиц на входном изображении. Это дает возможность использовать данный алгоритм в режиме реального времени. Основным недостатком метода является большое количество ложных обнаружений.

Нейронные сети показывают лучший результат для изображений, на которых лица расположены под значительным, более 20 градусов относительно вертикальной оси, углом и для изображений, сделанных при искусственном освещении.

Как следует из приведенного выше анализа для обнаружения лиц целесообразно использовать преимущества каждого из рассмотренных алгоритмов.

Предлагаемый метод обнаружения лиц представляет комбинацию двух методов: адаптивного усиления и нейронных сетей. Комбинация строится следующим образом: компонент метода адаптивного усиления обеспечивает принятие решения, есть ли на изображении лицо или нет, и отклоняет изображения без лиц. Метод нейронных сетей, получив на вход изображение, обработанное методом адаптивного усиления, формирует окончательное решение о наличии лица на входном изображении. Ниже на рисунке 1 представлена схема комбинированного метода. Применение на первом этапе комбинированного метода адаптивного усиления позволяет использовать предложенный метод в режиме реального времени. На втором этапе нейронные сети повторно выполняют проверку на присутствии лица на участке изображения, тем самым позволяя исправить ошибку ложного обнаружения метода адаптивного усиления.



Рис. 1. Функциональная схема комбинированного метода обнаружения лиц

Следующим этапом является решение задачи распознавания лиц.

Методы распознавания лиц

Выделим три базовых группы методов распознавания лиц [11]:

- группа методов, в основе которой лежит целостный подход. Методы этой группы обрабатывают всю поверхность лица как последовательность строк без учета индивидуальных анатомических признаков. Метод главных компонент (PCA)[9] и линейный дискриминантный анализ (LDA)[9] являются примерами целостного подхода по распознаванию лиц;

- группа методов, в основе которой лежит подход, основанный на анатомических признаках. Эти методы рассматривают анатомические признаки и выполняют сравнение этих признаков. Примерами таких методов являются метод распознавания, основанный на вейвлет преобразованиях Габора [10], метод эластичных связанных графов (EBGM) [11];

- группа методов, в основе которой лежит гибридный подход, представляющий объединение целостного подхода и подхода, основанного на признаках.

Для решения задачи распознавания лиц построим метод, представляющий собой комбинацию базовых компонентов метода подпространства линейного дискриминантного анализа[12] и метода, основанного на вейвлет преобразованиях Габора[10].

Метод «Подпространство LDA»

Метод «Подпространство LDA» является гибридным методом, объединяющим два независимых метода PCA[9] и LDA[9].

В методе PCA, предложенном Пентландом и Тарком, изображение рассматривается как точка в пространстве изображений путем преобразования изображения в вектор. Метод определяет пространство низшей размерности для представления лиц на изображениях с помощью устранения противоречий с изображениями, на которых отсутствуют лица. Признаки изображений из обучающей выборки получаются путем поиска максимального отклонения каждого изображения от значений, полученных от усредненных изображений.

Пространство собственных лиц получается с помощью применения метода главных компонент к обучающим изображениям. Затем обучающие изображения проектируются на пространство собственных лиц. Далее тестовое изображение проектируется на новое пространство и вычисляется расстояние между спроецированным тестовым изображением и изображениями из обучающего набора, которое используется для классификации тестового изображения. Принимается, что тестовое изображение, расстояние проекции которого является минимальным, содержит лицо из обучающей выборки.

Метод LDA, в противоположность методу PCA, группирует как можно ближе изображения с одним и тем же лицом и одновременно отдаляет друг друга изображения с разными лицами.

«Подпространство LDA» – гибридный метод, одновременно использующий методы PCA и LDA: LDA используется как основной классификатор, принимающий решение о присутствии лица на изображении, а PCA – как этап для уменьшения размерности. Таким образом, изначально создаются собственные вектора в подпространстве PCA с помощью обучающих изображений. Полученные вектора подаются на вход алгоритма LDA, который создает новое подпространство. При классификации тестовое лицо сначала проецируется на подпространство PCA, затем на полученное при обучении подпространство LDA. После чего выбирается единообразная метрика, которая используется для сравнения расстояния между индивидуальными проекциями векторов и происходит классификация тестового лица.

Диаграмма метода «Подпространство LDA» представлено на рис. 2.

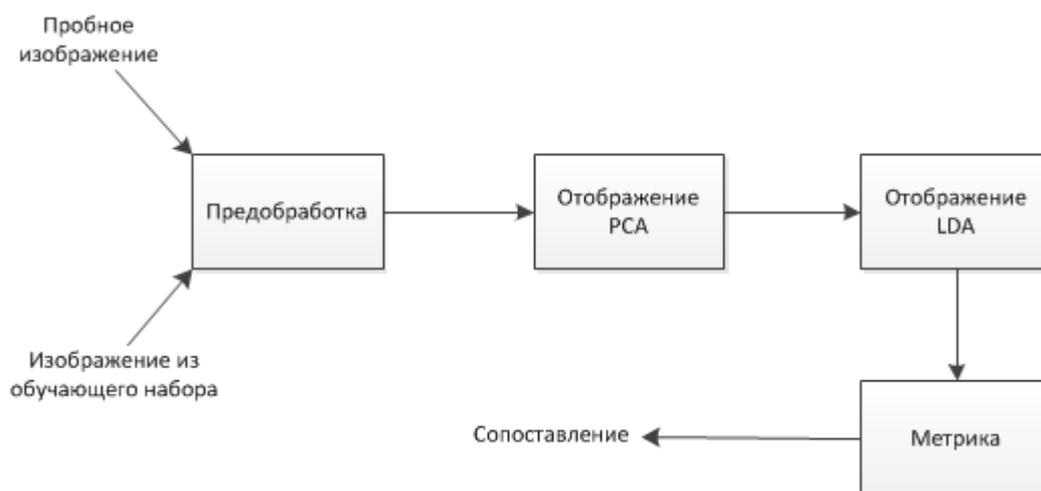


Рис. 2. Диаграмма метода подпространства LDA для распознавания лиц

Метод преобразования Габора

Рассмотрим второй компонент комбинации - метод, предложенный Брюсом, который использует вейвлеты Габора для обнаружения лицевых признаков на изображении [10]. Данный метод более устойчив к изменениям в освещении, поскольку не использует напрямую значения оттенков серого каждого пикселя, а извлекает признаки лица.

Вейвлет Габора (фильтр Габора) — линейный фильтр, импульсная переходная характеристика которого представляется в виде произведения функции Гаусса на гармоническую функцию [10]

$$g(x, y) = w(x, y)s(x, y). \quad (5)$$

Здесь составная синусоида $s(x, y) = e^{j(2\pi(ux+vy)+P)}$, где (6)

u, v – пространственные частоты;

В полярных координатах:

$$F_0 = \sqrt{u^2 + v^2}, \text{ где} \quad (7)$$

$$u = F_0 \cos w; \quad v = F_0 \sin w;$$

w - полярный угол; P – сдвиг фаз.

Функция Габора $w(x, y) = K e^{-\pi(a^2x_r + b^2y_r)}$, где (8)

K – константа масштабирования;

(a, b) – константы масштабирования вдоль осей координат;

θ - определяет ориентацию нормали параллельных полос функции Габора;

$$x_r = x \cos \theta + y \sin \theta; \quad (9)$$

$$y_r = -x \sin \theta + y \cos \theta. \quad (10)$$

Вейвлеты Габора используются для обнаружения анатомических признаков лица на изображениях оттенка серого [10]. Для недопущения потери признаков лица из-за различий в освещении между тестовым изображением и изображениями из обучающего набора формула приобретает вид [7]

$$g(x, y) = K e^{-\pi\sigma^2(x^2+y^2)} \left(e^{j(2\pi F_0(x \cos w + y \sin w))} - e^{-\left(\frac{\sigma^2}{2}\right)} \right). \quad (11)$$

Метод включает в себя два этапа обработки: извлечение признаков и их сопоставление с обучающим набором. Извлечение признаков отвечает за автоматическую

идентификацию признаков лица (нос, рот, уши и т.п.) и генерирует вектор признаков для каждой точки признака. Полученные пики вейвлета Габора являются точками признака. Поэтому метод позволяет идентифицировать различные структурные особенности лиц. Сопоставление признаков включает в себя сравнение векторов признаков тестового изображения с векторами признаков изображения из обучающей выборки, рис. 3.

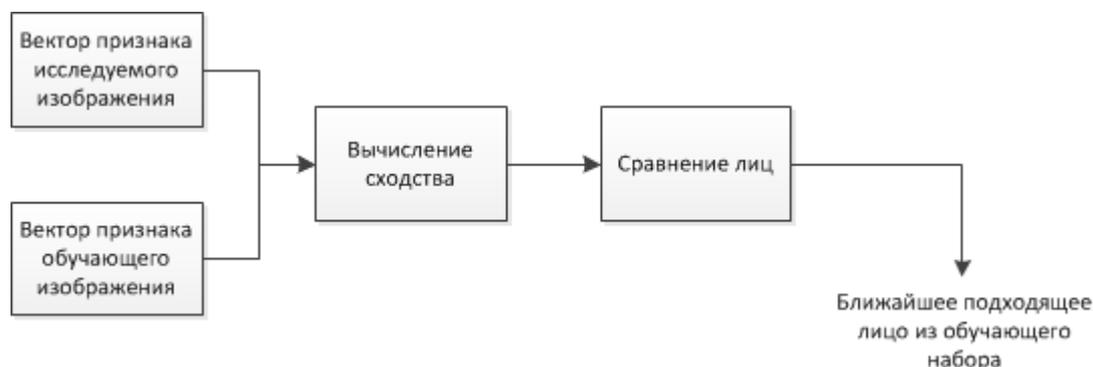


Рис. 3. Диаграмма, показывающая процедуру сопоставления признака

Сравнительный анализ выбранных методов-прототипов

Как показано в [12] достоверность распознавания метода «Подпространство LDA», который является методом целостного подхода, не зависит от разрешения тестового изображения. Метод способен распознать лица до тех пор, пока на изображении содержится общая структура лица, что позволяет использовать его для распознавания лиц в видеопотоках и в режиме реального времени с приемлемым уровнем точности. Однако производительность метода уменьшается при значительном различии в освещении на обучающем и пробном изображениях и при повороте распознаваемого лица.

Процесс переобучения для метода «Подпространство LDA» довольно сложен и требует временных затрат – нуждается в повторном вычислении подпространства LDA при добавлении новых изображений.

Производительность распознавания метода Кепенекси, в основе которого лежит подход, основанный на признаках, зависит от разрешения обучающих (пробных) изображений с лицами [10]: высокое разрешение изображений позволяет обнаруживать различные признаки лица (борода, шрам, родинка) и использовать их во время процесса сравнения. Большое количество выявленных признаков гарантирует высокую точность в процессе распознавания. Однако зависимость от качества изображения и сравнение большого количества признаков пробного изображения с признаками всех обучающих изображений делает его не применимым для использования в режиме реального времени.

Для метода, предложенного Кепенекси, производительность зависит от выбора параметров комбинации размера окна и от частоты синусоиды (w), что осложняет задачу подбора оптимальных значений параметров для наилучших характеристик производительности.

Для метода Кепенекси время переобучения невелико [10], т.к. учитывается время вычисления признаков новых изображений и дополнение новыми данными обучающей базы.

Таким образом, предложен метод распознавания лиц, представляющий комбинацию двух методов – «Подпространство LDA» и метода, основанного на вейвлет преобразованиях Габора (алгоритм Кепенекси). Комбинация строится следующим образом: алгоритм Кепенекси обрабатывает только те пробные изображения, расстояние которых, подсчитанное методом «Подпространства LDA», для каждого класса лиц приблизительно равно. В этом случае разрешение входных изображений должно быть среднего качества – не менее 64*64 пикселей. Структура алгоритма представлена на рис. 4. Применение на первом этапе комбинационного метода подпространства LDA позволяет использовать предложенный метод в режиме реального времени. На втором этапе метод Кепенекси выполняет сравнение обнаруженного лица с лицами из обучающего набора на наличие сходства.



Рис. 4. Функциональная схема комбинированного метода распознавания лиц

Заключение

С использованием преимущественных особенностей методов адаптивного усиления и нейронных сетей предложен комбинированный метод обнаружения лиц, позволяющий решить задачу обнаружения с повышенной скоростью обработки изображений и достоверностью обнаружения.

Для ускорения обучения каскадной модели в методе адаптивного усиления предложено:

- задавать количество классификаторов, с которых необходимо начинать отсчет при обучении, и количество классификаторов, которые необходимо добавлять для выполнения условия останова.

- в случае добавления нового классификатора, начинать обучение с добавленного классификаторов посредством сохранения весов последнего уровня каскадной модели.

- в случае возникновения ошибки при обучении, дальнейшее обучение начинать с этапа, на котором возникла ошибка.

Предложенный комбинированный метод распознавания лиц, включающий в себя базовые процедуры известных методов – метод, основанный на вейвлет преобразованиях Габора, и подпространство LDA, позволяет повысить качество распознавания лиц, сохранив скорость распознавания, которой будет достаточно для работы метода в режиме реального времени.

«Подпространство LDA» показывает лучшую производительность для изображений низкого качества (разрешения), однако производительность заметно ухудшается для изображений, содержащих отклонение в освещении и при повороте распознаваемых лиц. Метод Кепенекси, напротив, устойчив к изменению освещения, но результаты распознаванию приемлемы только при высоком разрешении входного изображения.

Список литературы

1. Кудряшов П.П. Алгоритм обнаружения лица человека для решения прикладных задач анализа и обработки изображений // Издательство физико-математической литературы. — М., 2007. — Т.2. — С.132-134;
2. Viola P., Jones M.J. Rapid object detection using a boosted cascade of simple features // In conference on computer vision and pattern. — 2001;
3. Henry R. Neural Network-Based Face Detection // School of Computer Science. — 2009;
4. Maydt J., Lienhart R. Face Detection with Support Vector Machines and a Very Large Set of Linear Features // IEEE ICME. — 2002;
5. Rahat M., Nazari M., Bafandehkar A., Ghidary S.S. Improving 2D boosted classifier using depth LDA classifier for robust face detection // International Journal of Computer Science Issues. — May 2012. — Vol. 9. No 2;
6. Kotropoulos C., Pitas I. Rule-Based Face Detection in Frontal Views // Proc. Int'l Conf. Acoustics, Speech and Signal Processing (ICASSP). — 1997. — Vol. 4. — P. 2537-2540;
7. Yang M.H., Kriegman D.J., Ahuja N. Detecting faces in images // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2002. — Vol. 24. — P. 34-58;
8. Otsu N. A Threshold selection method from gray-level histograms // IEEE Transactions on Systems and Cybernetics, — 1979. — Vol. 9, — P. 62- 66;

9. Zhao W., Chellapa R., Krishnaswamy A., Swets D., Weng J. Discriminant analysis of principle components for face recognition // 2nd International Conference on Automatic Face and Gesture Recognition. — April 1998. — P. 336-341;
10. Etemand K., Chellapa R. Discriminant analysis for recognition of human faces image // Journal of optical society of America A. — 2005. — P. 1724-1733;
11. Kepenekci B. Face recognition using gabour wavelet transform // Thesis, the Middle East Technical University. — September 2001;
12. Wiskott L., Fellous J.M., Kruger N., Malsburg C. Face Recognition by Elastic Bunch Graph Matching // I9 Transactions on Pattern Analysis and Machine Intelligence. — 1997. — Vol. 19. — P. 775-779.