

Распознавание изображений лиц на основе сравнения с 2D и 3D эталонами

Кийко К.В., Кийко В.М., Недашковский А.В., Рябокони Д.И.

Рассматриваются новые алгоритмы идентификации человека по изображению лица на основе сравнения с эталонами, которые могут быть получены или путем обработки входных изображений, или на основе построения и использования 3D моделей распознаваемых лиц. Приведены результаты тестирования и сравнительного анализа разработанных алгоритмов.

ВВЕДЕНИЕ

Распознавание лиц имеет такую же давнюю историю, как и проблема компьютерного зрения. Однако, несмотря на значительный научный интерес и практическое значение, проблема распознавания человека по изображению его лица в реальных условиях остается все еще далекой от окончательного решения. Основная причина этого состоит в более высокой вариативности изображений лиц одного человека относительно поворотов и условий освещения по сравнению с вариативностью изображений лиц разных людей при приблизительно одинаковых положениях и условиях освещения.

Известные алгоритмы распознавания изображений лиц [1,2] состоят из двух частей:

- локализация лица на изображении;
- нормализация локализованного лица (по масштабу, повороту и яркости) и его распознавание.

Проведенные нами исследования показали, что эффективной является двухэтапная процедура локализации лиц на изображениях: сначала положения лиц определяются методом AdaBoost [3], а затем эти положения уточняются путем поиска совокупностей составных элементов лица в пределах прямоугольников, определенных на первом этапе.

Известные алгоритмы распознавания изображений локализованных лиц можно разделить на следующие основные группы:

1. Алгоритмы, которые выполняют поиск составных элементов лица на изображении и определяют их относительное положение и координаты особых точек лица [4,5].
2. Алгоритмы на основе сравнения изображения с эталонами [5].
3. Алгоритмы, использующие метод «собственных лиц» [6].
4. Нейросетевые алгоритмы распознавания [7].
5. Алгоритмы, которые а) формируют граф описания изображения, вершинам которого отвечают локальные признаки изображения, и б) вычисляют степень близости этого графа с эталонными графами путем поиска соответствия между ними [8].
6. Алгоритмы распознавания на основе использования 3D моделей лиц [9].

В статье рассматриваются алгоритмы распознавания лиц на основе сравнения входных изображений с эталонами. Известно, что надежность распознавания с помощью таких алгоритмов зависит от изменения условий освещения и поворотов распознаваемых лиц. Целью работы является разработка эффективных алгоритмов построения эталонов для уменьшения этой зависимости на основе построения и использования 3D моделей лиц, а также алгоритмов нормализации этих эталонов. Второй не менее важной целью является разработка эффективных алгоритмов поиска лиц на изображении.

Разработано два алгоритма распознавания, которые условно называются «Моно» и «Моно3D» и отличаются между собой способом формирования эталонных изображений. Для алгоритма «Моно» эталоны получают путем обработки изображений лиц. При реализации алгоритма «Моно3D» в качестве эталона используется 3D модель лица, с помощью которой генерируются эталонные изображения, отвечающие различным положениям головы человека в пространстве. Таким образом, этот алгоритм является более стойким по отношению к изменениям ориентации головы человека, чем алгоритм «Моно». При этом эталоны можно также вводить и хранить непосредственно, но при таком способе формирования эталонов

требуется больший объем памяти, и, кроме того, эталонам разных людей могут соответствовать существенно разные углы наклона лиц, что ухудшает надежность распознавания.

Ниже в разделе 1 рассматриваются алгоритмы поиска лиц на изображении, в разделах 2,3 - алгоритмы нормализации изображений, в разделе 4 - особенности вычисления расстояний между нормализованными изображениями, в разделе 5 - технология построения 3D модели лица, в разделах 6,7 - результаты тестирования алгоритмов и основные выводы.

1. Поиск лиц на изображении

Поиск лиц на изображении состоит из двух этапов. На первом этапе с помощью метода AdaBoost [3] определяются прямоугольники, которые покрывают квазифронтальные лица на изображении. С точки зрения видео наблюдения эти результаты являются вполне удовлетворительными, но с точки зрения последующего распознавания каждый прямоугольник недостаточно точно определяет координаты лица и центров зрачков глаз на изображении, необходимые для выполнения нормализации изображения. Вот почему на втором этапе положения лиц уточняются в пределах каждого из полученных прямоугольников на основе поиска признаков, описывающих пару радужных оболочек глаз лица. Рассмотрим основные этапы поиска лиц более подробно.

Алгоритм [3] состоит из двух частей - обучения и распознавания. Во время обучения рассматривается большое количество (несколько тысяч) нормализованных изображений двух классов: а) лицо и б) не лицо, т.е. “все другое”. Для правильного распознавания как можно большего числа этих изображений последовательно выполняется автоматический поиск наиболее эффективных признаков в следующем классе простых признаков. Каждому из признаков соответствует определенный прямоугольник на изображении, состоящий из 2-х, 3-х или 4-х составных (по вертикали или горизонтали) прямоугольников. При этом для вычисления признака первого типа необходимо найти разность суммарных яркостей двух составных прямоугольников и сравнить ее с определенным порогом. Для вычисления признака второго типа необходимо от яркости центрального прямоугольника отнять яркости двух других прямоугольников, а значение признака 3-го типа зависит от разности суммарных яркостей диагональных пар 4-х составных прямоугольников.

На каждом шаге процесса обучения дополнительно привлекается тот признак, который является наиболее эффективным для правильной классификации тех изображений, которые были неправильно классифицированы совокупностью предварительно найденных признаков. Результат обучения есть одна или несколько совокупностей признаков, обеспечивающие правильное распознавание всех изображений учебной выборки.

Поиск лиц на изображении состоит из следующих операций. На изображении рассматриваются все прямоугольные участки, размеры которых принадлежат интервалу допустимых размеров лиц на этом изображении. В пределах каждого из этих прямоугольников выполняется проверка совокупностей простых признаков, определенных во время обучения. При этом решение о нахождении лица в пределах текущего прямоугольника принимается тогда и только тогда, когда выполняются все признаки из каждой из этих совокупностей. Для вычисления каждого из признаков необходимо выполнить лишь несколько простых операций, если предварительно сформировать так называемое “интегральное” представление изображения. Значение в каждой клетке интегрального изображения равно сумме яркостей всех клеток исходного изображения, расположенных слева и сверху от данной клетки. Как следствие, локализация лиц по этому алгоритму выполняется довольно быстро.

Алгоритм уточнения положения лица человека состоит из следующих частей или этапов:

1. Поиск сравнительно небольшого количества точек, каждая из которых может быть центром зрачка глаза на изображении. Определение для каждой точки наиболее соответствующего значения радиуса радужной оболочки с центром в этой точке и некоторых других параметров.

3. Рассмотрение всех совместимых пар найденных точек, т.е. таких, которые расположены приблизительно на горизонтальной линии и определенном расстоянии одна от

другой. Вычисление значения качества (правдоподобия) каждой из этих пар точек. Локализация лица человека на основе выбора той пары точек (центров зрачков), которая имеет наибольшее значение качества, превышающее заданное пороговое значение.

Результатом работы первого этапа алгоритма, который выполняется в процессе просмотра изображения вдоль его строк, есть координаты (x, y) центральной точки и половина длины r каждого из следующих горизонтальных отрезков $GL(x, y, r) = [(x - r, y), (x + r, y)]$ на изображении: а) значения градиента яркости на концах отрезка превышают заданные пороговые значения; б) средняя яркость $Bright(x, y, r)$ клеток изображения этого отрезка является сравнительно малой, а r относится к числу допустимых значений радиуса радужной оболочки глаза на снимке.

Во время второго этапа уточняются полученные на первом этапе значения координат центров (x, y) и радиусов r радужных оболочек. При этом для фиксированной точки (x, y) наиболее достоверным считается то значение радиуса радужной оболочки r_0 , которому соответствует наибольшее значение $Grad(x, y, r_0)$ изменения в радиальном направлении суммарной яркости клеток изображения на границе левого $[-45^\circ, 60^\circ]$ и правого $[120^\circ, 225^\circ]$ секторов круга с центром (x, y) и радиусом r_0 . В процессе уточнения рассматриваются 4 значения радиуса r_i из интервала $[r - 1, r + 2]$ и 9 значений координат центра (x_i, y_i) , $x_i = x \pm 1$, $y_i = y \pm 1$. Для каждой из троек (x_i, y_i, r_i) сначала вычисляются усредненные яркости клеток изображения $Br_L(x_i, y_i, r_i)$ и $Br_R(x_i, y_i, r_i)$ вдоль границ соответственно левого и правого секторов круга с центром (x_i, y_i) и радиусом r_i . После этого находятся средние яркости вдоль соответствующих секторов круга с этим же центром и радиусом $(r_i + 2)$, а также значение градиента яркости $Grad(x_i, y_i, r_i) = Grad_L(x_i, y_i, r_i) + Grad_R(x_i, y_i, r_i)$, где

$$Grad_L(x_i, y_i, r_i) = Br_L(x_i, y_i, r_i + 2) - Br_L(x_i, y_i, r_i),$$

$$Grad_R(x_i, y_i, r_i) = Br_R(x_i, y_i, r_i + 2) - Br_R(x_i, y_i, r_i),$$

если выполняются условия:

$$\max(Grad_L(x_i, y_i, r_i), Grad_R(x_i, y_i, r_i)) > 10$$

$$\min(Grad_L(x_i, y_i, r_i), Grad_R(x_i, y_i, r_i)) > 0.$$

Окончательно $Grad(x, y, r_0) = Grad(x_{ic}, y_{ic}, r_{ic})$ и $r_0 = r_{ic}$, где $Grad(x_{ic}, y_{ic}, r_{ic})$ – наибольшее вычисленное значение градиента. После этого, если выполняются определенные условия симметрии и однородности яркости клеток в пределах радужной оболочки, показатель того, что точка $t(x, y)$ является центром радужной оболочки глаза с радиусом r вычисляется по формуле:

$$Sim(t) = Grad(x, y, r) - Bright(x, y, r) * 0.25 + r \quad (1)$$

В конце второго этапа выполняется сокращение списка точек – кандидатов на центр радужной оболочки глаза. При этом некоторая точка исключается из списка, если рядом с ней есть другая точка с большим значением показателя (1). На третьем этапе алгоритма рассматриваются все пары (t_1, t_2) оставшихся точек, удовлетворяющие следующим условиям:

а) расстояние между двумя точками больше, чем $4.5 * D$, и меньше, чем $9 * D$, где $D = (r_1 + r_2)$ - сумма радиусов радужных оболочек, соответствующих этим точкам;

б) $|x_2 - x_1| > |y_2 - y_1| * 2.8$;

в) $|r_2 - r_1| \leq 2$.

Для каждой пары таких точек определяется показатель качества $SIM(t_1, t_2) = Sim(t_1) + Sim(t_2) + Sim_a(t_1, t_2)$ того, что t_1 и t_2 - центры радужных оболочек глаз

на снимке. Значение дополнительного показателя $Sim_a(t_1, t_2)$ уменьшается, если угол наклона соединяющей эти точки линии больше 15^0 или параметры двух радужных оболочек существенно отличаются между собой. Значение этого показателя зависит также от выполнения условий относительной яркости определенных составных участков лица на изображении, соответствующих положению текущей пары точек t_1 и t_2 .

В конце работы алгоритма выполняется локализация лица человека на основе выбора той пары точек – центров зрачков глаз, которая имеет наибольшее значение их общего показателя $SIM(t_1, t_2)$, превышающее заданное пороговое значение. Примеры результатов работы алгоритма показаны на Рис. 1.

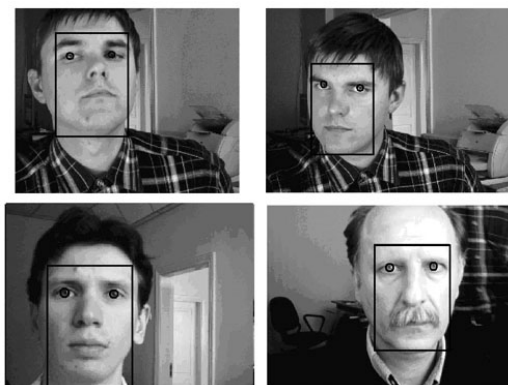


Рис. 1. Примеры результатов поиска лиц и зрачков глаз на полутоновых изображениях, полученных с помощью веб камеры.

2. Нормализация изображения лица по масштабу и повороту

Перед сравнением двух изображений лица человека необходимо привести эти изображения к стандартному виду, при котором координаты центров зрачков глаз на изображении имеют определенные фиксированные значения. Линия, которая соединяет эти центры, является горизонтальной. Такую нормализацию можно выполнить с помощью операций поворота и масштабирования изображения.

Угол поворота вычисляется по формуле:

$$\alpha = -\arctan(b_y / b_x),$$

где b – вектор, который соединяет центры зрачков глаз.

Коэффициент масштабирования вычисляется по формуле:

$$t = \frac{eye_dist}{|b|},$$

где $eyes_dist$ – фиксированное расстояние между центрами зрачков глаз.

Поворот - это преобразование клеток изображения, которое задается матрицей

$$T = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix},$$

где α – угол поворота.

Новые координаты клетки (x, y) после поворота вычисляются по формуле:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = T * \begin{pmatrix} x \\ y \end{pmatrix}.$$

С целью минимизации внесенных искажений используется алгоритм поворота с помощью сдвигов. Матрица поворота при этом представляется в виде произведения трех матриц:

$$\begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix} \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix},$$

где

$$a = \operatorname{tg} \alpha / 2; \quad b = -\sin \alpha; \quad c = \operatorname{tg} \alpha / 2.$$

Первая матрица соответствует сдвигу строк клеток изображения по горизонтали, вторая - сдвигу столбцов по вертикали и третья - снова сдвигу строк по горизонтали. Вследствие этого результирующее изображение состоит из тех же клеток, что и начальное, но перенесенных на другое место. Размывание изображения при этом не происходит.

Алгоритм масштабирования построен по принципу ближайшего соседа. Яркость нового изображения в клетке (x', y') подсчитывается за формулой:

$$f(x', y') = g\left(\left[\frac{x'}{k}\right], \left[\frac{y'}{k}\right]\right)$$

где $[x]$ – ближайшее к x целое число, $g(x, y)$ – функция яркости начального изображения, k – коэффициент масштабирования, $k > 0$.

3. Нормализация изображения по яркости

Цель нормализации изображений по яркости состоит в выполнении таких преобразований яркости этих изображений, которые уменьшают зависимость результатов распознавания от условий освещения и тем самым обеспечивают повышение надежности распознавания. При распознавании изображений лиц используются приведенные ниже алгоритмы их нормализации по яркости, которые в большинстве случаев улучшают результаты распознавания, но не всегда, поскольку поиск информативного и вместе с тем независимого от условий освещения представления входного изображения в общем случае является некорректной задачей [10]. Каждый из этих алгоритмов обеспечивает инвариантность результатов распознавания к определенному классу изменений яркости изображения, и эти классы не совпадают между собой.

Нормализация по средней яркости и дисперсии. Согласно этому преобразованию функция яркости клеток изображения $f(x, y)$ меняется на функцию $f_{norm}(x, y) = f(x, y) * a + b$, которая имеет заданные значения средней яркости и дисперсии (соответственно 0 и 1). Такая нормализация является инвариантом к аффинным преобразованиям яркости клеток изображения и не зависит от изменений яркости и контраста изображения. Аффинные преобразования яркости, с одной стороны, составляют довольно широкий класс преобразований, но с другой – это лишь малая часть возможных изменений яркости даже в таком сравнительно простом случае: $f(x, y)$ есть изображение плоского объекта, который освещается только одним источником света, положение и интенсивность которого могут меняться.

Нормализация методом выравнивания гистограммы. Под выравниванием (эквализацией) гистограммы понимается такое преобразование входного изображения, которое приводит к более равномерному распределению яркостей клеток изображения в интервале $[0, 255]$ допустимых значений этих яркостей. Алгоритм нормализации изображения, который использует нелинейное выравнивание гистограммы, состоит из следующих этапов.

Сначала вычисляют гистограмму $h(j)$ яркостей клеток изображения $f(x, y)$:

$$h(j) = \sum_{x,y} [f(x, y), j], \quad [f(x, y), j] = \begin{cases} 1, & \text{если } f(x, y) = j \\ 0, & \text{если } f(x, y) \neq j \end{cases}$$

На втором шаге получают так называемую “кумулятивную” гистограмму $C(i)$:

$$C(i) = \sum_{j=1}^i (h(j))^r,$$

где r – параметр, который определяет степень нелинейности. В большинстве случаев этот параметр равен 1, но при нормализации изображений лиц лучше использовать $r = 0.5$.

На заключительном, третьем этапе, вычисляют новые значения яркостей клеток изображения:

$$f_{norm}(x, y) = \frac{C(f(x, y))}{C(255)} \cdot 255.$$



Рис. 2. Изображение (слева) и результат его нормализации после выполнения выравнивания гистограммы.

Нормализация изображения методом выравнивания гистограммы является инвариантной к любой монотонной функции от яркостей клеток изображения, в том числе и к любому аффинному преобразованию яркости. При этом сохраняется упорядоченность клеток по значениям их яркости и изменяются только самые значения: выравнивание не нарушает общую форму гистограммы, а только определенным чином растягивает ее вдоль всего интервала $[0, 255]$. Нормализация большинства изображений по приведенному выше алгоритму улучшает результаты их распознавания, но не является желательной для очень контрастных изображений.

Нормализация на основе локальных изменений яркости изображения. Такая нормализация является эффективной, например, в таких случаях: освещение не является фронтальным, т.е. источник освещения расположен не перед лицом, а например, сбоку; изображение имеет малый динамический диапазон яркости, например, при недостаточном или вечернем освещении;

Алгоритм нормализации основывается на вычислении локального контраста $Lc(x, y)$ входного изображения в клетке (x, y) как отношения яркости этой клетки к средней яркости изображения в сравнительно малой ее окрестности [5]:

$$Lc(x, y) = \begin{cases} f_c(x, y), & \text{если } f_c(x, y) \leq 1 \\ 2 - \frac{1}{f_c(x, y)}, & \text{если } f_c(x, y) > 1, \end{cases}$$

где $f_c(x, y) = \frac{f(x, y)}{f(x, y) \otimes K}$.

Символом \otimes обозначено свертку изображения с двумерным Гауссовым ядром K , дисперсия δ которого определяет размер окрестности, в пределах которого выполняется усреднение яркости изображения. Нормализованное таким образом изображения меньше зависит от условий освещения и в тот же время является достаточно информативным. На Рис. 3 показан пример нормализации изображения лица при размере фильтра, равном значению радиуса радужной оболочки глаза на изображении. Яркости клеток изображения на этом рисунке равны $127 * Lc(x, y)$.



Рис. 3. Изображение (слева) и результат его нормализации по локальному контрасту.

Нормализация изображения по приведенному выше алгоритму является инвариантом к определенному классу изменений функции яркости изображения. Так, например, если изображение имеет области с постоянным градиентом яркости, то после выполнения нормализации эти области будут иметь значения $Lc(x, y) = 1$. Нормализованное изображение является также инвариантом к аффинным преобразованиям яркости, если предварительно прологарифмировать входное изображение. Перед вычислением меры близости двух изображений, нормализованных по локальному контрасту, необходимо их дополнительно нормализовать по средней яркости и дисперсии.

4. Распознавание нормализованных изображений на основе вычисления сходства между эталоном и входным изображением

При распознавании используются два типа эталонов лиц (Рис. 4,б-в), которые состоят соответственно из 3-х и 4-х частей эталонных изображений (Рис. 4,а) из базы данных.

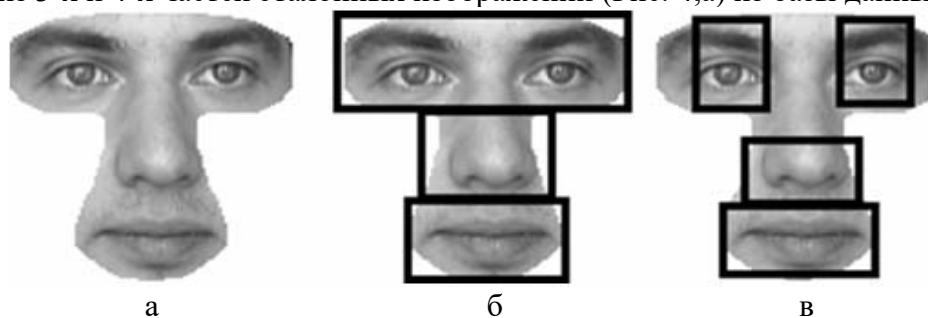


Рис. 4. Эталоны двух типов (б, в), состоящие из частей эталонного изображения (а) лица.

В процессе распознавания находится оптимальное положение каждой составной части эталона на входном изображении, при котором расстояние между соответствующими участками изображения и эталона является наименьшим. Это расстояние имеет вид суммы расстояний между соответствующими клетками изображения и эталона, а расстояние между эталоном и изображением определяется как взвешенная сумма расстояний его составных частей, разделенная на суммарную яркость клеток эталона. Значения этих расстояний зависят от выбора меры близости между яркостями клеток. При этом расстояние между изображением x и эталоном e по общепринятой L_p -метрике вычисляется по формуле:

$$L_p(x, e) = \left(\sum_{t \in T} |x(t) - e(t)|^p \right)^{\frac{1}{p}}.$$

В случае, когда $p = 2$, $L_2(x, e)$ является евклидовым расстоянием между изображением и эталоном. Экспериментальные исследования показывают, что использование этого расстояния менее целесообразно в сравнении с расстоянием $L_1(x, e)$, равном сумме модулей разностей яркостей соответствующих клеток изображения и эталона, и расстоянием $L_1^*(x, e)$, равном сумме модулей разностей яркостей клеток, превышающих заданное пороговое значение. Именно это расстояние $L_1^*(x, e)$ было выбрано за основу для оценки близости между изображением и эталоном.

Распознавание изображений организовано в два этапа. Сначала входное полутоновое изображение распознается путем сравнения с эталонами первого типа. При этом изображение и эталоны первого типа нормализуются по яркости методом выравнивания гистограммы, которая вычисляется по всему изображению. Если в результате первого этапа принимается решение об отказе от распознавания, то выполняется второй этап, на котором изображение сравнивается с эталонами второго типа. При этом последовательно используются два способа локальной (независимо по каждой из 4-х составных частей эталона) нормализации по яркости

распознаваемого изображения и эталонов - методом выравнивания гистограммы и методом вычисления локального контраста, а расстояние между изображением и эталоном считается равным меньшему из двух расстояний, соответствующих использованию различных способов нормализации. Экспериментальные результаты показывают, что совместное использование эталонов упомянутых выше двух типов уменьшает зависимость результатов распознавания от условий освещения и повышает надежность распознавания.

Расстояние d между прямоугольным участком эталона $(x_1 \leq x \leq x_2, y_1 \leq y \leq y_2)$ и соответствующим участком изображения вычисляется последовательно по столбикам клеток изображения по формуле:

$$d = \sum_{x=x_1}^{x_2} \sum_{y=y_1}^{y_2} dif(x, y), \quad (2)$$

где (x_1, y_1) и (x_2, y_2) – координаты соответственно верхнего левого и нижнего правого углов участка эталона, а

$$dif(x, y) = \begin{cases} |f(x, y) - e(x, y) + \beta(x)|, & \text{если } dif(x, y) > \delta \\ 0, & \text{в противном случае,} \end{cases} \quad (3)$$

где δ – пороговое значение;

$f(x, y)$ – яркость изображения в клетке (x, y) ;

$e(x, y)$ – яркость эталона в клетке (x, y) ;

$\beta(x)$ – разность средних яркостей в столбце x изображения и эталона.

При таком способе вычисления расстояния уменьшается зависимость результатов распознавания от условий освещения распознаваемого лица. После нахождения оптимальных положений составных частей эталона на изображении окончательное расстояние между изображением и эталоном вычисляется следующим образом:

$$dif''(x, y) = \min_{(x', y') \in \Delta(x, y)} dif(x', y'),$$

где $\Delta(x, y)$ – окрестность 3×3 клетки (x, y) .

Критерий принятия решения. Пусть $(d_1, p_1), (d_2, p_2), \dots (d_N, p_N)$ – упорядоченный по расстоянию список пар, где

$d_i \in R$ – расстояние между распознаваемым изображением и эталоном лица человека p_i ;

$$d_{(1)} = \min_{1 \leq i \leq N} d_i, \quad d_{(2)} = \min_{i \neq i_1} d_i, \quad i_1 = \arg \min_i d_i,$$

$$a = \frac{d_{(2)} - d_{(1)}}{d_{(1)}} \cdot 100.$$

Решение об отказе от распознавания принимается, когда $d_{(1)}$ больше некоторого порога θ_1 (большие расстояния между изображением и эталонами), или когда a меньше порога θ_2 (близкие расстояния между изображением и эталонами разных людей). Во всех других случаях принимается решение о распознавании человека p_j .

5. Получение эталонных изображений на основе пространственной реконструкции поверхности лица по стереопаре изображений

Для распознавания повернутых влево-право или вверх-вниз изображений лиц необходимо иметь соответствующие эталонные изображения, так как нормализация изображений по таким поворотам не может быть выполнена на основе значений координат центров зрачков глаз лица человека на изображении. В этом разделе рассматривается способ получения эталонных изображений путем восстановления 3D модели лица по стереопаре изображений и использования этой модели для генерирования двумерных проекций, отвечающих наиболее распространенным вариантам положения лица человека перед камерой.

Разработана автоматическая технология стереорекострукции поверхности лица. Входными данными этой технологии являются два изображения лица (неидеальная стереопара), введенные двумя некалиброванными камерами, которые сдвинуты одна относительно второй в пространстве. На этих изображениях сначала выполняется поиск лиц. После этого выделенные изображения лиц поступают на вход алгоритма стереорекострукции [11,12], который восстанавливает карту относительных высот поверхности лица, на основе которых затем строится триангулированная поверхность лица. Для раскрашивания этой поверхности берется левое изображение исходной стереопары.

При повороте и масштабировании восстановленной 3D модели лица каждая ее вершина (вектор (x,y,z)) умножается на матрицу R размерностью 3×3 , которая является результатом произведения матриц поворота и масштабирования:

$$R = Sz(t) \cdot Tz(\alpha_z) \cdot Ty(\alpha_y) \cdot Tx(\alpha_x),$$

где $Tx(\alpha_x)$, $Ty(\alpha_y)$, $Tz(\alpha_z)$ – матрицы поворота соответственно вокруг осей X, Y, Z на углы α_x , α_y , α_z ; $Sz(t)$ – матрица масштабирования в t раз.

Новые координаты точки $\vec{p} = (x, y, z)$ в пространстве вычисляются по формуле: $\vec{p}^* = R \cdot \vec{p}$; точки на спроектированном изображении имеют координаты (x,y) , что соответствует значениям $(0,0, -\infty)$ координат камеры.

Для получения двумерных проекций 3D модели используется библиотека OpenGL.

6. Результаты тестирования разработанных алгоритмов

Для проверки алгоритмов была сформирована база данных, которая состоит из 1600 приблизительно фронтальных изображений 80 людей: 1) 1200 изображений лиц, введенных с помощью видеокамеры при разных условиях освещения, в разных помещениях и в разное время на протяжении двух лет и 2) 400 изображений из базы данных ORL (<http://www.cam-orl.co.uk/facedatabase.html>). На этой базе данных выполнялось тестирование алгоритмов Моно, Моно3D, а также трех известных алгоритмов распознавания основанных на наиболее эффективных модификациях метода «собственных» лиц [6]: анализ главных компонент (PCA) [13], байесовский классификатор максимального правдоподобия (BML) [14] и байесовский классификатор максимальной апостериорной вероятности (BMAP) [14]. В методе Моно3D эталонами являются восстановленные по стереопаре изображений пространственные конфигурации лиц. Из-за отсутствия внешних баз данных стереоизображений лиц для тестирования алгоритма Моно3D было создано семнадцать 3D эталонов 14-и сотрудников МНУЦИТиС. Каждый из этих 3D эталонов использовался для получения 11 эталонных изображений, соответствующих сравнительно небольшим отклонениям (до 25°) от фронтальной ориентации лица человека. При тестировании алгоритма Моно в большинстве случаев использовалось только одно эталонное изображение лица каждого человека.

Для тестирования алгоритмов PCA, BML, BMAP использовалось программное обеспечение Colorado State University Face Identification Evaluation System (версия 5.0) [15], которое реализует эти алгоритмы. При тестировании алгоритмов Моно и Моно3D использовалась разработанная автоматизированная технология тестирования программного обеспечения. Согласно этой технологии все изображения из базы данных последовательно распознаются в автоматическом режиме, а результаты распознавания записываются в HTML файл.

В Таблице 1 представлены результаты работы пяти вышеупомянутых алгоритмов на одних и тех же входных данных. При этом параметр $\theta_2 = 0$ (см. раздел 4), поскольку в программном обеспечении [15] решение о распознавании принимается по методу ближайшего соседа, и отказ от распознавания не предусмотрен. В Таблице 2 приведены результаты работы алгоритмов Моно и Моно3D в двух режимах при различных значениях параметра θ_2 . В первом из этих режимов эталон распознаваемого лица имеется в базе данных. При этом правильным

решением является определение лица перед камерой; ошибочным решением является принятие одного человека за другого, а отказ от распознавания означает, что система не может узнать человека перед камерой. Во втором режиме (тест на отказ в доступе) база данных эталонов не содержит эталон распознаваемого лица. При этом условия единственно правильным ответом является отказ от распознавания, т.е. недопущение чужого.

Таблица 1. Результаты тестирования алгоритмов распознавания лиц.

Алгоритм	θ_2	Тест на идентификацию		
		Верно, %	Отказ, %	Ошибка, %
Моно3D	0	97.3	-	2.7
Моно-Моно	0	91.6	-	8.4
Bayesian ML	0	89.7	-	11.3
Bayesian MAP	0	76.5	-	23.5
PCA	0	65.0	-	35.0

Таблица 2. Результаты тестирования алгоритмов Моно и Моно3D при разных значениях θ_2 .

Алгоритм	θ_2	Тест на идентификацию			Тест на отказ в доступе	
		Верно, %	Отказ, %	Ошибка, %	Верно, %	Ошибка, %
Моно-Моно	40	84.3	14.5	1.2	92.0	8.0
	50	81.3	17.7	1.0	94.9	5.1
	55	80.3	19.0	0.7	96.0	4.0
Моно3D	40	91.6	8.4	0	94.3	5.7
	50	91.3	8.7	0	97.0	3.0
	55	90.0	10.0	0	97.6	2.4

Как видно из результатов тестирования, алгоритм Моно3D обеспечивает большую надежность распознавания по обоим тестам. При этом основная часть ошибок связана с существенно разными условиями освещения эталонного и входного изображений, а также с изменениями выражения лица. Существенно разными условиями освещения считаются, например, такие: эталон введен при фронтальном естественном освещении, а входное изображение снято при верхнем искусственном свете. Если при таких условиях наблюдается ухудшение качества распознавания, необходимо ввести не один, а два эталона соответственно при дневном и искусственном освещении. Источником другой, значительно меньшей части ошибок (1.1%), является неправильное нахождение лица или глаз лица на изображении. Третий источник ошибок – наличие изображений с углами поворота лица, существенно большими, чем 25°. Примеры изображений, нераспознанных алгоритмом Моно (отказ от распознавания), показаны на Рис. 5.

Алгоритм Моно по результатам распознавания уступает алгоритму Моно3D, но превышает известные алгоритмы, реализованные в CSU Face Identification Evaluation System. Результаты распознавания алгоритмом Моно можно улучшить, если использовать несколько эталонных изображений (вместо одного) каждого человека. Эти эталоны могут отличаться между собой, например, поворотом или выражением лица, условиями освещения и т.п.. Разработанное программное обеспечение позволяет вводить новые эталоны в базу данных „на лету” в процессе распознавания.

Общее время распознавания всех 1600 изображений алгоритмами Моно и Моно3D на компьютере с процессором Athlon XP 2600 составляет соответственно 5.7 и 62 минут, т.е. в

среднем 0.13 и 2.3 сек на одно входное изображение. Это время может быть уменьшено путем разбиения процесса распознавания изображения на два этапа с использованием сначала небольшой, а затем большей разрешающей способности этого изображения. Необходимо также отметить, что в режиме верификации решение о доступе человека принимается за значительно меньшее время в результате сравнения входного изображения лишь с несколькими эталонами.



Рис. 5. Примеры нераспознанных изображений.

Изображения лиц из сформированной нами базы данных по своим свойствам близки к изображениям категорий *fc* и *dupl* известной базы данных FERET [16] (*fc* – фронтальные изображения, снятые в один день, но при разных условиях освещения; *dupl* – изображения, снятые в разные дни, отличающиеся между собой масштабом, углом поворота, выражением лица и освещением). По результатам тестирования лучших на 1997 год алгоритмов на базе данных FERET было установлено [16], что надежности распознавания изображений категорий *fc* и *dupl* этими алгоритмами составляют соответственно 80% и 50%. Тестирование разработанных алгоритмов на базе данных FERET планируется выполнить на следующем этапе исследований с целью дальнейшего сравнительного анализа этих алгоритмов.

7. Выводы

Разработаны алгоритмы поиска лиц, обеспечивающие высокую надежность и точность локализации лиц на изображениях со сложным фоном и различными условиями освещения.

Разработаны алгоритмы распознавания, которые используют операции только с целыми числами и обеспечивают меньшую зависимость результатов распознавания от условий освещения и поворотов лица человека в сравнении с известными алгоритмами. Это достигается путем построения эталонов на основе восстановленных 3D моделей лиц, а также использования более эффективных алгоритмов нормализации этих эталонов.

Разработанные алгоритмы и программное обеспечение могут быть использованы в системах идентификации человека и контроля его доступа.

Литература

1. *W.Zhao, R.Chellappa, A.Rosenfeld, and P.J.Phillips*, "Face Recognition: A Literature Survey," pp.1-66, 2000.
2. *M.-H. Yang, D.Kriegman, and N.Ahuja*, "Detecting Faces in Images: A Survey," IEEE Trans. on PAMI, vol. 24, no. 1, pp. 34-58, Jan. 2002.
3. *P.Viola and M.Jones*, "Robust real-time object detection," In Proc. of IEEE Workshop on Statistical and Computational Theories of Vision, pp. 1-25, 2001.

4. *I.Craw, H.Ellis, J.R.Lishman*, "Automatic extraction of face features", Pattern Recognition Letters, Vol. 5, pp. 183-187, 1987.
5. *R.Brunelli, T.Poggio*, "Face Recognition: Features versus Templates", IEEE Trans. on PAMI, vol. 15, No. 10, pp.1042-1052, 1993.
6. *M.A.Turk and A.P.Pentland*, "Eigenfaces for recognition," Journal of Cognitive Neuroscience, vol. 3, pp. 72-86, 1991.
7. *S.Y.Kung, M.Fang, S.P.Liou, M.Y.Chiu, and J.S.Taur*, "Decision-based neural network for face recognition system," in Proc. ICI'96, vol. 1. Washington, pp. 430-437, 1997.
8. *L.Wiskott, J.Fellows, N.Kruger and C.Marlborg*, "Face recognition by Elastic Bunch Graph Matching," IEEE Trans. on PAMI, vol. 19, no.7, pp. 775-779, 1997.
9. *C.Beumier, M.Acheroy*, "Automatic 3D Face Authentication", In Image and Vision Computing, vol. 18, no. 4, pp. 315-321, 2000.
10. *Hansen F. Chen, Peter N. Belhumeur and David W. Jacobs*. "In search of illumination invariants", In Proc. of CVPR-2000, USA, pp. 1254-1261, June 2000.
11. *Д. Рябоконт*, "Пространственная реконструкция поверхностей по стереопаре изображений с помощью алгоритма поиска минимального сечения на графе", УСиМ, 2004, №3, с. 47-51.
12. *Д. Рябоконт*, "Восстановление пространственной конфигурации объектов и сцен по их стереоизображениям", УСиМ, 2005, №1, с. 22-31.
13. *Wendy S. Yambor*, "Analysis of PCA-Based and Fisher Discriminant-Based Image Recognition Algorithms", M.S. Thesis, Technical Report CS-00-103, Computer Science, p. 70, July 2000.
14. *David Bolme*, "The Bayesian Intrapersonal/Extrapersonal Classifier", Masters Thesis, CSU Computer Science Department, p. 42, July 2003.
15. *Ross Beveridge*. Evaluation of face recognition algorithms web site.
<http://cs.colostate.edu/evalfacerec>.
16. *P.J.Phillips, H.Moon, S.A.Rizvi and P.J.Rauss*, "The FERET Evaluation Methodology for Face-Recognition Algorithms," IEEE Trans. On PAMI, vol. 22, no. 10, pp. 1090-1104, Oct. 2000.