

УДК 004.93

Система автоматического контроля доступа на основе анализа изображений лица

Бобе А.С.¹, Жиденко М.В.², Прояев П.О.³

^{1,2,3} Студенты, кафедра «Биомедицинские технические системы»

МГТУ им. Н.Э. Баумана, г. Москва, Россия

*Научный руководитель: А.В. Самородов, к.т.н., доцент, доцент кафедры
«Биомедицинские технические системы» МГТУ им. Н.Э. Баумана, г. Москва, Россия*

МГТУ им. Н.Э. Баумана

Anatoly.Bobe@mail.ru

Введение

Системы автоматического контроля доступа играют важную роль в современном мире. В последние годы всё активнее проводится разработка и внедрение технологий, способных эффективно обеспечивать безопасность различного рода объектов. Одним из ведущих направлений исследований в этой сфере является разработка биометрических систем, осуществляющих анализ характерных биометрических особенностей человека с целью идентификации его личности и принятия решения о доступе.

В данной работе представлен аппаратно-программный комплекс, предназначенный для осуществления кооперативного распознавания человека и принятия решения о доступе путем анализа изображений лица по получаемому с камеры видеоряду. Система способна определять локальные и глобальные признаки лица для проведения верификации личности, а также различать движения головы путем определения изменений ракурса лица на изображениях. Программная часть комплекса реализует алгоритм верификации, а также алгоритм определения ракурса лица, основанный на объединении метода локализации биометрических (контрольных) точек лица и модифицированного метода представления активной модели формы (active shape model, ASM).

1. Алгоритм мультиклассификации для решения задачи верификации личности по изображению лица

В данной работе были проведены разработка и исследование алгоритмов верификации, построенных на основе объединения результатов работы базовых алгоритмов, применяемых к различным областям изображения лица. В качестве базовых были рассмотрены два алгоритма: алгоритм, основанный на методе локальных бинарных шаблонов, и алгоритм на основе фильтров Габора.

1.1 Базовый алгоритм верификации на основе локальных бинарных шаблонов

Метод локальных бинарных шаблонов (ЛБШ) применяется для выделения признаков изображений на основе описания локальной окрестности каждого пикселя в двоичной форме [1]. Модификация ЛБШ, применяемая для распознавания изображений лиц, была описана в работах [2, 3]. Оператор формирования ЛБШ применяется к каждому пикселю изображения и основан на сравнении значений интенсивности восьми пикселей его окрестности с его интенсивностью. По результатам данных сравнений формируется восьмиразрядный бинарный код, в котором единицы и нули соответствуют пикселям окрестности, интенсивность которых не меньше или, соответственно, меньше интенсивности рассматриваемого пикселя (рис.1).

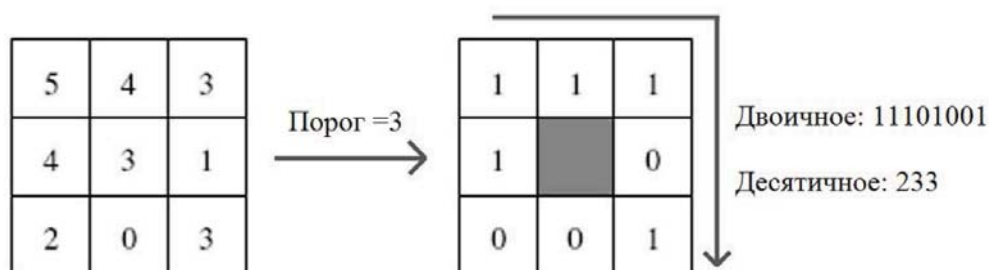


Рис.1. Пример работы ЛБШ-оператора [2]

В модификации алгоритма, используемой в данной работе, расстояние между центральным пикселем и пикселями окрестности было взято равным двум пикселям.

После формирования ЛБШ набираются те из них, которые имеют не более трех серий нулей и единиц, называемые равномерными ЛБШ (РЛБШ) [1].

Вектор признаков всего изображения формируется на основе алгоритма, описанного в работе [4]. Сначала изображение лица делится на $k*k$ блоков (в данной работе $k=6$, т.е. всего 36 блоков) и в каждом блоке для каждого пикселя изображения вычисляются ЛБШ. Затем для каждого блока изображения отбираются РЛБШ и строится гистограмма значений ЛБШ, причем все ЛБШ, не являющиеся равномерными,

учитываются как имеющие одинаковые значения. Гистограмма каждого блока изображения содержит $P \cdot (P-1) + 3$ отсчетов, где $P = 8$ – количество рассматриваемых точек окрестности пикселя. Набор гистограмм всех блоков рассматриваемого изображения является его вектором признаков.

При сопоставлении двух изображений лица для каждой пары гистограмм соответствующих областей вычисляется расстояние Кульбака-Лейблера [3]:

$$d_n = \sum_{i=1}^L \lg \left(\frac{x_{ni}}{y_{ni}} \right) \cdot (x_{ni} - y_{ni}),$$

где x_{ni}, y_{ni} – i -е отсчеты гистограмм ЛБШ, соответствующих n -му блоку первого и второго изображений лица;

L – количество отсчетов гистограммы ЛБШ.

В результате формируется $k \cdot k$ – мерный вектор $D = \{d_n\}$, характеризующий степень отличия двух изображений.

Окончательное решение принимается по значению степени схожести сопоставляемых изображений:

$$s = \frac{y_1}{y_1 + y_2},$$

где y_1 – расстояние в $k \cdot k$ -мерном пространстве между текущим вектором D рассматриваемой пары изображений и средним вектором D , полученным ранее по группе обучения для пар изображений, соответствующих классу «чужих»;

y_2 – аналогичное расстояние до среднего вектора пар изображений, соответствующих классу «своих».

Указанные расстояния в $k \cdot k$ – мерном пространстве вычисляются как расстояния Махаланобиса:

$$y_c = \sqrt{(D - \mu_c)^T \cdot C^{-1} \cdot (D - \mu_c)},$$

где μ_c – средний вектор D для класса $c = \{\text{«свои»}, \text{«чужие»}\}$;

C – корреляционная матрица вектора D .

Сопоставляемые изображения считаются принадлежащими одному человеку, если мера схожести $s > s_{\text{thr}}$, где пороговое значение s_{thr} определяется по группе обучения исходя из требований к уровню ошибки второго рода (вероятность ложного совпадения, ВЛС).

1.2 Базовый алгоритм верификации на основе фильтров Габора

Для реализации второго базового алгоритма распознавания проводится предварительная фильтрация изображения в оттенках серого по следующему алгоритму. Окно фильтра достаточно большого размера (в данной работе 13×13 пикселей) проходит

по всем пикселям изображения. Для каждого положения окна производится бинаризация соответствующей окрестности изображения по порогу, равному значению интенсивности ее центрального пикселя (Рис 2). Количество пикселей со значениями интенсивности, превышающими пороговое, записывается в качестве значения центрального пикселя выходного изображения. Результат такой фильтрации показан на рис. 2.



Рис. 2. Предобработка изображения лица

Полученные изображения затем обрабатываются фильтром Габора:

$$g(x, y) = K \cdot \exp(-\pi(a^2(x - x_0)^2_r + b^2(y - y_0)^2_r)) \cdot \exp(j(2\pi(u_0x + v_0y) + P)),$$

где K – масштабирующий коэффициент;

a, b – масштабы по осям x, y ;

x_0, y_0 – координаты центра гауссовского окна;

u_0, v_0 – частоты фильтра по осям x, y ;

P – начальная фаза.

При формировании векторов признаков изображения лица фильтрация проводится на 5 различных частотах и при 8 различных ориентациях фильтра. Таким образом формируются 40 каналов преобразования, каждый из которых дает свой вектор признаков.

На следующем этапе к каждому из 40 векторов признаков применяется метод главных компонент (МГК), с оставлением 60 отсчетов для каждого вектора. Затем по каждому из 40 каналов рассчитываются степени схожести сравниваемых изображений как косинусное расстояние между соответствующими векторами признаков:

$$d = \frac{(x, y)}{\|x\| \cdot \|y\|}.$$

Итоговая степень схожести представляет собой сумму 40 степеней схожести по каждому из каналов. Как и для первого базового алгоритма, ее значение сравнивается с порогом, полученным для данного алгоритма по группе обучения.

1.3 Объединение базовых алгоритмов верификации

Для повышения качества работы базовых алгоритмов верификации каждый из них вначале применяется независимо к пяти разным областям лица: области глаз, лба, носа, рта и лица в целом. Полученные по каждой из областей степени схожести нормируются следующим образом:

$$sn_i = \frac{S - S_{thr,i}}{S_{thr,i}}$$

где $S_{thr,i}$ – порог принятия решения для i -й области.

Знак нормированных таким образом степеней схожести может интерпретироваться как решение классификатора, а абсолютная величина – как степень уверенности в этом решении. Нормированные степени схожести складываются с весовыми коэффициентами w_i для получения итоговой степени схожести. Схема объединения результатов работы базового алгоритма, реализованная в данной работе, показана на рис. 3.

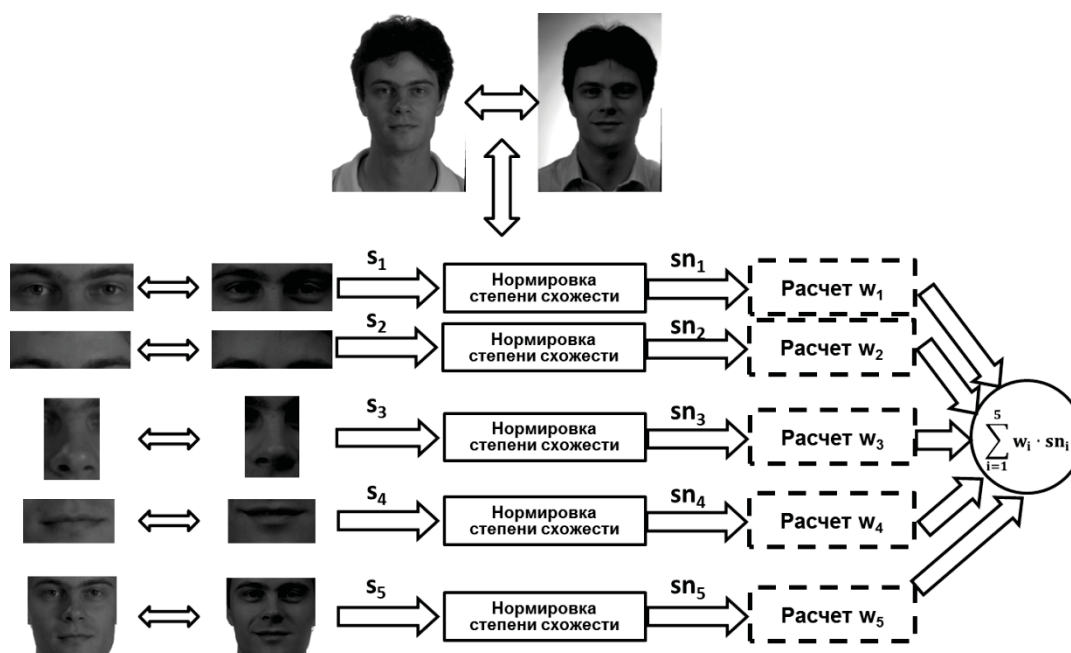


Рис.3. Схема объединения результатов работы алгоритма по 5 областям изображения лица

Существует несколько способов нахождения весовых коэффициентов. В данной работе на основании ранее проведенных исследований весовой коэффициент i -й области вычислялся следующим образом:

$$w_i = \begin{cases} \frac{sn_i}{\max_2_i}, & sn_i \geq 0 \\ \frac{-sn_i}{\min_1_i}, & sn_i < 0 \end{cases}$$

где \max_2i , \min_1i – границы диапазонов возможных значений нормированной степени схожести sn_i .

Построенные два алгоритма верификации представляют собой мультиклассификаторы на основе объединения степеней схожести.

После расчета объединенных по областям лица степеней схожести для каждого из базовых классификаторов выполняется их итоговое объединение. Для этого значения объединенных по областям лица степеней схожести нормировались следующим образом:

$$sn_{ik} = \frac{s_{ik} - \mu(s_i)}{\sigma(s_i)},$$

где $\mu(s_i)$, $\sigma(s_i)$ – соответственно среднее значение и среднеквадратичное отклонение значений объединенных степеней схожести i -го мультиклассификатора, полученные для тестовой выборки.

Итоговое объединение нормированных степеней схожести двух мультиклассификаторов заключается в их суммировании с весовыми коэффициентами, вычисляемыми на основе значений вероятности совпадения (TMR) для каждого из мультиклассификаторов. Итоговый алгоритм также является алгоритмом мультиклассификации.

Характеристики работы полученных мультиклассификаторов исследовались на базе ColorFERET, причем для обучения и тестирования были выбраны только изображения, полученные в разные даты (341 изображение в обучающей выборке и 100 в тестовой). Задача верификации на данной выборке существенно усложняется по сравнению со стандартной выборкой из базы ColorFeret из-за существенных различий во внешности людей на изображениях.

Полученные вероятностные характеристики (вероятность совпадения TMR при заданной ошибке ложного совпадения FMR) представлены в таблице 1.

Таблица 1

Сравнительные характеристики построенных алгоритмов верификации

Алгоритм	TMR, %		
	FMR=0,1 %	FMR=0,5 %	FMR=1,0 %
Мультиклассификатор по областям лица на основе LBP	55,0	68,4	77,1

Мультиклассификатор по областям лица на основе фильтров Габора	88,3	96,5	97,4
Итоговый мультиклассификатор	84,0	92,6	96,1

Как следует из таблицы, для базы ColorFeret вероятностные характеристики общего мультиклассификатора оказались ниже, чем характеристики мультиклассификатора на основе признаков Габора, который и был реализован в разработанной системе автоматического контроля доступа.

2. Алгоритм определения ракурса головы на изображении

В данной работе были проведены разработка и исследование алгоритма слежения за ракурсом головы человека по двум степеням свободы – повороты вправо и влево (YAW) и наклоны вниз и вверх (PITCH). Алгоритм построен на основе объединения метода локализации контрольных точек (КТ) лица, представленного в работе [5], и модифицированного метода представления активной модели формы – ASM, представленного в работе [6].

В ходе обучения алгоритма решаются задачи расчета эталонных признаков КТ лица, а также построение модели определения ракурса. При практическом применении алгоритма осуществляется поиск КТ лица на распознаваемом кадре, определение ракурса лица на этом кадре, прослеживание изменения ракурса лица в видеопоследовательности и принятие решения о наличии нужного паттерна движения головы.

3.1 Обучение алгоритма определения ракурса лица

3.1.1 Расчет эталонных векторов признаков КТ лица

Для обучения использовался набор изображений с размеченными КТ лица, представленный в базе данных Pointing04. В основе расчета признаков КТ лежит алгоритм фильтрации, представленный в [5]. Фильтр представляет собой набор из 16-ти так называемых ядер, каждое из которых чувствительно только к одному из 16-ти направлений и находится как разность между двумя функциями Гаусса:

$$G_{pqf_k}^- = \frac{1}{2\pi\sigma^2} * \exp\left(-\frac{(p - q * \cos(f_k))^2 + (p - \sin(f_k))^2}{2\sigma^2}\right),$$

$$G_{pqf_k}^+ = \frac{1}{2\pi\sigma^2} * \exp\left(-\frac{(p + q * \cos(f_k))^2 + (p + \sin(f_k))^2}{2\sigma^2}\right),$$

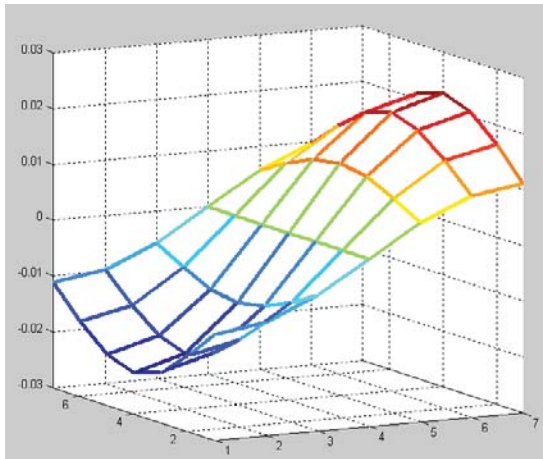
где σ - СКО, f_k - угол поворота, $f_k = k * 22.5^\circ$, $k=0..15$, $p \in [-3; 3]$, $q \in [-3; 3]$.

Далее вычисляется набор значений свертки области вокруг текущей точки по формуле:

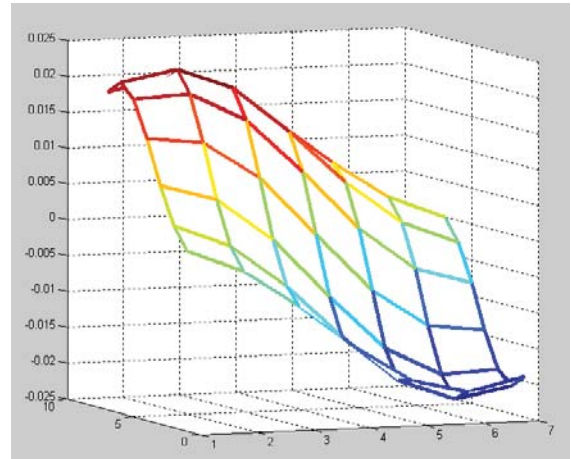
$$CE_{ijfk} = \sum_{p,q} C_{i-p,j-q} * (G_{pqfk}^- - G_{pqfk}^+),$$

где C_{ij} – значение интенсивности в данной точке.

Максимальный отклик среди 16-ти значений CE_{ijfk} определяет максимальное изменение локальной границы вокруг данной точки. На рис. 4 представлена визуализация функции $(G_{pqfk}^- - G_{pqfk}^+)$ для углов 0 и 157,5 градусов.



а) Угол $f=0^\circ$



б) Угол $f=157,5^\circ$

Рис.4. Вид ядра фильтра для различных направлений

Для каждой КТ выполнялся расчет среднего по обучающей выборке вектора признаков, состоящего из 49 элементов, каждый из которых представляет собой результат фильтрации окрестности соответствующего узла сетки (рис.5) по описанному выше алгоритму.

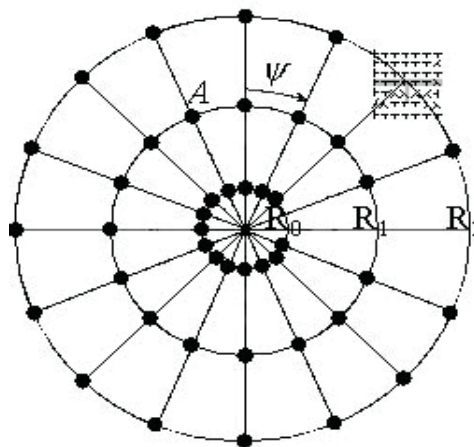


Рис.5. К вычислению вектора признаков КТ (в центре). $\Psi=22.5^\circ$, R_0, R_1, R_2 выбираются исходя из размеров кадра

Средние вектора признаков КТ, полученные по обучающей выборке, считаются эталонными.

3.1.2 Построение модели формы лица при изменениях ракурса

Для построения модели формы лица при изменениях ракурса была использована база данных Pointing04, представляющая собой набор фотоснимков пятнадцати людей, сфотографированных с известными ракурсами. Изменения ракурса проводились только по двум степеням свободы – наклон головы (PITCH) и поворот в стороны (YAW). В дальнейшем для удобства ракурс будет обозначаться двумя числовыми значениями: первое характеризует наклоны вперед и назад, второе – повороты вправо и влево. Таким образом, запись $-15+30$ означает наклон головы вниз на 15 градусов и поворот вправо на 30 градусов (см. рис. 6).

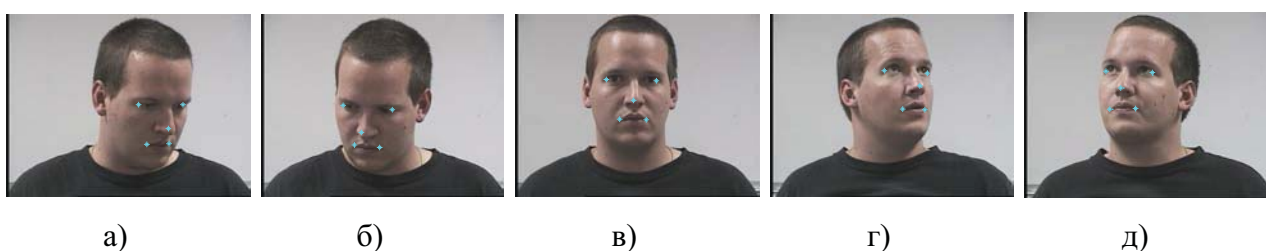


Рис.6. Пример серии изображений базы данных Pointing04

В качестве вектора данных каждого изображения лица были использованы координаты пяти КТ лица: внешних углов глаз, вершины носа и углов рта. Указанный вектор данных характеризует относительное расположение пяти КТ, то есть форму лица. С использованием МГК размерность вектора данных была снижена до двух. При этом первый признак нового двумерного вектора признаков формы лица пропорционален углу поворота головы, второй – углу наклона (рис. 7).

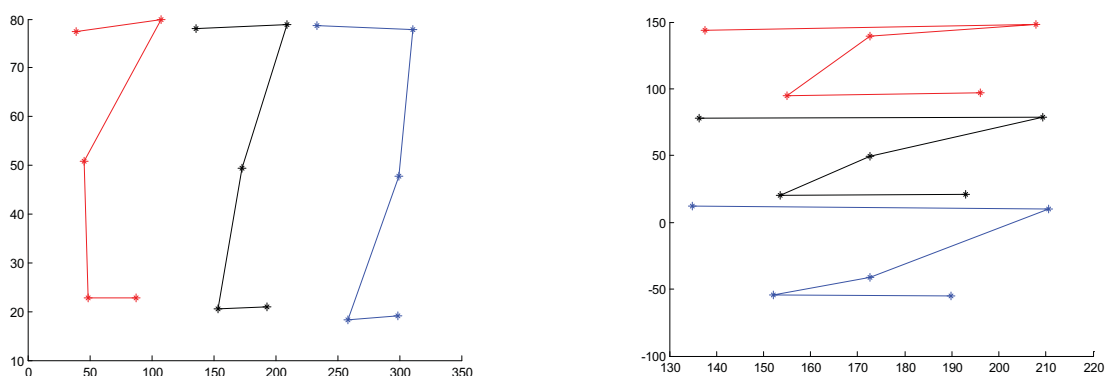


Рис.7. Изменение формы для одного человека в зависимости от изменения значений признаков формы

3.2 Определение ракурса лица

Для локализации указанных выше пяти КТ лица на новом кадре видеопоследовательности используются признаки КТ, вычисляемые аналогично п.3.1.1. В некоторой окрестности от положения каждой КТ на предыдущем кадре выполняется поиск максимума степени схожести вектора признаков пикселя – кандидата и эталонного вектора признаков соответствующей КТ. Степень схожести текущего и эталонного вектора признаков рассчитывается как [5]:

$$K = \sum_{i=1}^{49} th\left(\frac{1}{16} \sum_{j=1}^{16} Sg(A_{ij}^b; A_{ij}^{rw})\right),$$

где A_{ij} является значением свертки для j -ой ориентации ядра фильтра в i -м узле сетки (рис.5);

b – эталон,

rw – текущий вектор признаков;

$$Sg(a, b) = \begin{cases} 1, & \text{если } (a > 0) \wedge (b > 0), \\ 1, & \text{если } (a = 0) \wedge (b = 0), \\ 0, & \text{в других случаях;} \end{cases}$$

$$th(x) = \begin{cases} x, & \text{если } x > x_{thr}, \\ 0, & \text{в других случаях;} \end{cases}$$

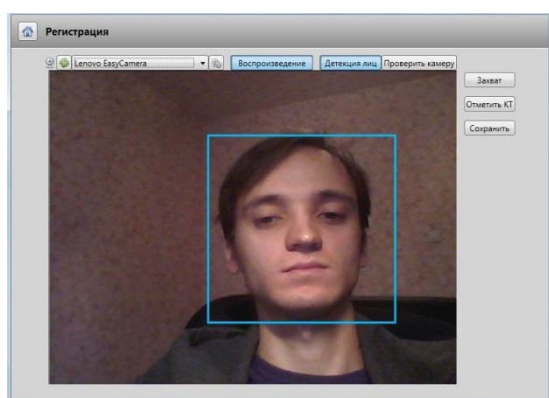
$x_{thr} = 0,2$ – пороговое значение, подобранное эмпирически.

Найденные таким образом координаты пяти КТ лица затем уточняются в процессе «выравнивания» модели на основе так называемых жестких и нежестких преобразований модели активной формы лица [6]. Отличием реализованного метода уточнения координат КТ от известного является использование только двух параметров нежесткого преобразования, соответствующих двум признакам формы лица, пропорциональным углам поворота и наклона головы и определенным в ходе обучения алгоритма при построении модели формы (п.3.1.2). Найденные в ходе «выравнивания» модели признаки формы задают ракурс лица на текущем кадре.

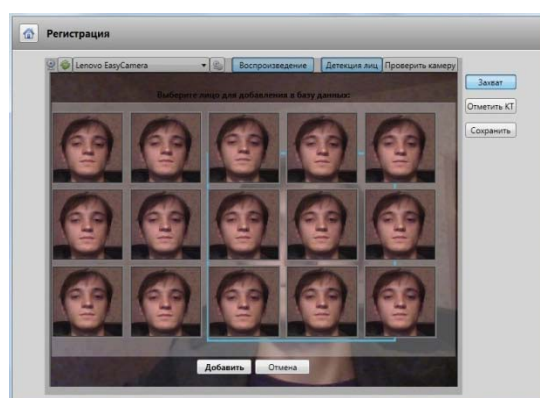
По результатам оценки ракурса лица на каждом кадре видеоряда определяется паттерн движения (например, наклон вниз, поворот влево), который используется для подтверждения человеком своего права доступа после его верификации.

4. Порядок работы системы контроля доступа

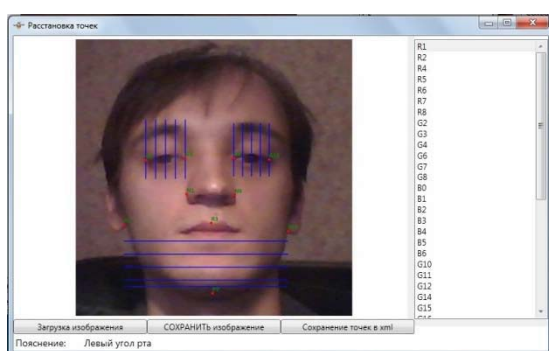
Факт появления человека перед камерой фиксируется детектором лиц (рис. 8а), реализованным на основе детектора Viola-Jones. Затем производится захват, нормализация и сохранение установленного числа кадров для создания выборки изображений, пригодных для верификации (рис. 8б). В случае работы в режиме ввода нового человека в базу данных людей, которым разрешен доступ, производится выбор изображения-шаблона, ручная расстановка на нем КТ и запись шаблона в базу данных (рис 8в). В случае работы в режиме контроля доступа производится ввод персонального пароля, по которому из базы данных извлекается биометрический шаблон данного человека. Затем производится процедура сравнения зарегистрированных изображений с этим шаблоном (верификация). По результатам верификации, в случае подтверждения личности человека запускается процедура подтверждения права доступа, состоящая из совершения человеком контрольного движения головой и его распознавания системой автоматического контроля доступа (рис 8г).



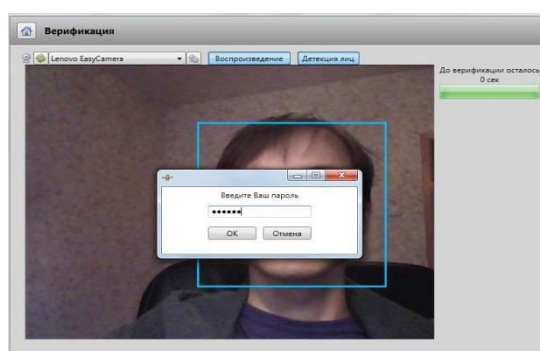
а)



б)



в)



г)

Рис.8. Этапы работы автоматической системы контроля доступа

Основные технические элементы автоматической системы контроля доступа представлены на рис. 9.

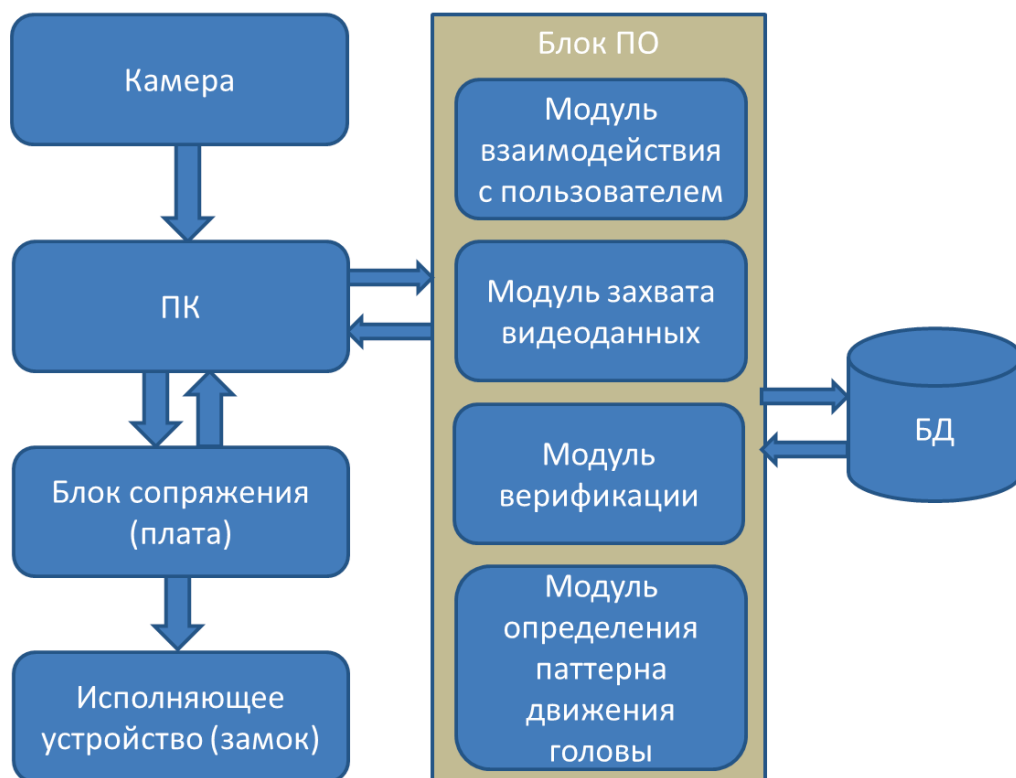


Рис.9. Структура системы контроля доступа

Видеосигнал с камеры поступает на ПК, программное обеспечение которого проводит анализ поступающей информации. Блок захвата видеоданных с помощью детектора лиц осуществляет первичный отбор данных, передавая на дальнейшую обработку модулю верификации нормализованные изображения лиц (в случае необходимости, осуществляется также запись изображений в базу данных). По результатам обработки данных через блок сопряжения генерируется управляющий сигнал, осуществляющий контроль исполняющего устройства.

Заключение

В работе была рассмотрена система автоматического контроля доступа на основе анализа характеристик лица по видеоряду. Система, состоящая из камеры для регистрации изображений, ПК с программным обеспечением для обработки получаемых данных в реальном времени и исполняющего устройства, регулирующего доступ в помещение, способна успешно решать задачу обеспечения безопасности объекта. Были достигнуты относительно высокие (до 97,4%) вероятностные характеристики алгоритма верификации, которые должны еще более повыситься при практическом применении системы благодаря наличию дополнительной степени защиты в виде распознавания контрольного движения головы. Разработанная система работает в режиме реального времени.

Список литературы

1. Ojala, T., Pietikainen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition* 29(1), 51–59 (1996).
2. Маслий Р. В.: Использование локальных бинарных шаблонов для распознавания лиц на полутоновых изображениях. *Наукові праці ВНТУ*, № 4 (2008)
3. Петрук В.И., Самородов А.В., Спиридонов И.Н. Применение локальных бинарных шаблонов к решению задачи распознавания лиц //Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение. Спец. выпуск «Биометрические технологии». – 2011. – С.58-63.
4. Ahonen T., Hadid A., Pietikainen M.: Face description with local binary patterns: application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28(12): 2037-2041. (2006).
5. Gao X.W., Anishenko S., Shaposhnikov D.: High-precision Detection of Facial Landmarks to Estimate Head Motions Based on Vision Models //Journal of Computer Science, Vol.3, No.7. – P.528-532, 2007
6. Cootes, T.F., Taylor, C.J.: Active Shape Models – Their Training and Application// *Computer Vision and Understanding* Vol. 61, No 1, January, pp. 38-59, 1995.