

ИТЕРАЦИОННЫЙ АЛГОРИТМ ЛОКАЛИЗАЦИИ ЦЕНТРОВ ГЛАЗ НА ИЗОБРАЖЕНИИ ЛИЦА

Хрящев В.В., к.т.н., доцент, доцент кафедры динамики электронных систем Ярославский государственный университет им. П.Г. Демидова, e-mail: vhr@yandex.ru;

Приоров А.Л., д.т.н., доцент, доцент кафедры ДЭС ЯрГУ им. П.Г. Демидова, e-mail: andcat@yandex.ru;

Никитин А.Е., аспирант кафедры ДЭС ЯрГУ им. П.Г. Демидова, e-mail: dcslab@uniyar.ac.ru;

Степанова О.А., аспирант кафедры ДЭС ЯрГУ им. П.Г. Демидова, e-mail: dcslab@uniyar.ac.ru.

ITERATIVE ALGORITHM FOR THE EYE CENTERS LOCALIZATION ON FACIAL IMAGE

Khryashchev V., Priorov A., Nikitin A., Stepanova O.

A new iterative algorithm for the eye centers localization based on multi-block local binary patterns is proposed. This algorithm is based on machine learning using multi-block local binary patterns, previously used in the problem of face detection. The results of the comparative analysis of the given algorithm with other known methods of eyes localization (gradient and Bayesian) in the standard test image database – FERET and BioID are considered. Low computational complexity of considered algorithm can be used in different applications with real-time video processing.

Keywords: face recognition, eye center localization, multi-block local binary pattern, machine learning.

Ключевые слова: распознавание лиц, локализация центров глаз, мультиблочные локальные бинарные шаблоны, машинное обучение.

Введение

Важнейшую роль в современных системах охранного телевидения играют алгоритмы цифровой обработки изображений и технического зрения [1–3], позволяющие контролировать сотни и тысячи видеоканалов в режиме реального времени. Одним из важнейших направлений модернизации таких систем является решение задачи автоматического распознавания объектов [4–6]. Это необходимо для разработки и производства систем, способных интеллектуально оценивать внешнюю среду и выполнять в ней те или иные действия.

Задача точного определения положения глаз на изображении лица, или локализация глаз, важна для широкого спектра современных задач в области компьютерного зрения, таких как определение направления взгляда и угла поворота головы к камере, анализа выражения лица, мимики и т.д. [6–9]. Помимо этого, локализация глаз успешно используется как предварительный этап при распознавании лиц – координаты центров глаз помогают правильно провести нормировку изображения лица после его детектирования (рис. 1) [7, 8]. Исследования показывают, что точность локализации глаз оказывает существенное влияние на качество системы распознавания лиц [10, 11].

Довольно распространено ошибочное мнение, что локализация глаз является довольно простой задачей, так как глаза – всего лишь один из элементов лица. Однако глаза имеют свою уникальную геометрическую структуру, характеристики движения, фотометрические свойства, что делает их довольно трудным объектом

Предлагается итерационный алгоритм локализации центров глаз на основе мультиблочных локальных бинарных шаблонов. Основой для предлагаемого алгоритма является метод на основе машинного обучения с использованием мультиблочных локальных бинарных шаблонов, ранее применявшийся в задаче детектирования лиц. Приводятся результаты сравнительного анализа предложенного алгоритма с другими известными методами локализации (градиентным и байесовским) на стандартных тестовых базах изображений FERET и BioID. Низкая вычислительная сложность рассмотренного алгоритма позволяет использовать его в приложениях обработки видеоданных реального времени.

для исследований. Можно перечислить ряд факторов, которые оказывают значительное влияние на внешний вид глаз на изображении [12–16]:

- большое разнообразие цвета и видов глаз;
- различные эмоции и мимика: к примеру, у смеющегося человека глаза могут быть практически закрыты;
- наличие оптических препятствий на лице: глаза могут быть частично закрыты очками либо прядями волос;
- поза человека: различные варианты взаимного расположения человека и камеры (фронтальное, в профиль, снизу-вверх и т.д.) влияют на внешний вид глаз; например, при расположении головы в профиль один глаз может быть полностью закрыт;

- условия съемки и качество изображения: факторы внешней среды, например, освещение (его спектральный состав, расположение источников освещения и интенсивность) может значительно менять внешний вид глаз. Более того, очень распространенные на практике низкое разрешение изображений, размытость или низкая детализация текстуры могут сильно влиять на качество исходных данных, что становится большой проблемой для любых алгоритмов локализации глаз.

За последние три десятилетия предложено множество различных подходов к решению задачи локализа-

ции глаз. Тем не менее, несмотря на значительные успехи в данной области, стоит отметить, что данная проблема все еще далека от решения. Большинство методов локализации глаз, предложенных исследователями в последнее время, можно условно разделить на три категории [17]:

- методы, основанные на измерении параметров составных элементов глаза;
- методы, создающие статистическую модель глаз на основе обучения;
- методы, использующие информацию о пространственной структуре лица.

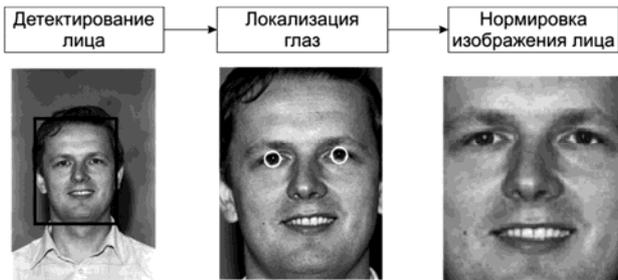


Рис. 1. Роль локализации положения глаз в процедуре распознавания лиц

Проведенный авторами анализ известных алгоритмов локализации глаз показал, что существующие алгоритмы при определении положения центров глаз на изображении лица могут допускать ошибки [11]. Неточности локализации известных алгоритмов связаны, прежде всего, с наличием на лице очков и недостаточным качеством изображений, имеющихся на практике. Погрешности в определении положения глаз серьезно увеличивают число ошибок всей системы распознавания лиц на телевизионном изображении. В работе предлагается новый алгоритм локализации положения центров глаз на основе локальных бинарных паттернов [18, 19], подстраивающийся к качеству и сложности изображения.

Для проведения тестирования использовались два известных алгоритма локализации глаз – градиентный [20] и байесовский [21] – в настоящее время часто используемых в практических приложениях и демонстрирующих приемлемую точность локализации. Градиентный алгоритм локализации использует априорную информацию о пространственной структуре лица. Байесовский алгоритм создан на основе статистического обучения на имеющейся выборке изображений глаз.

Алгоритм локализации центров глаз на основе мультиблочных локальных бинарных шаблонов

Мультиблочный локальный бинарный шаблон (МБ-ЛБШ) является оператором, применяемым к некоторой прямоугольной области изображения (рис. 2). Область разбивается на девять равных блоков. Средняя интенсивность g_c центрального блока рассматривается в качестве порогового значения, которое сравнивается с интенсивностями оставшихся восьми областей окрестности $\{g_1, \dots, g_8\}$ [19].

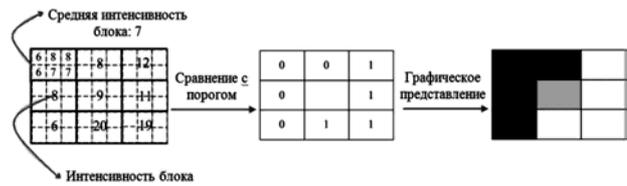


Рис. 2. Пример вычисления оператора МБ-ЛБШ

Выходное значение оператора МБ-ЛБШ центрального блока представляется в следующем виде:

$$МБ-ЛБШ_c = \sum_{i=1}^8 2^i s(g_i - g_c), \tag{1}$$

где s – сигма-функция:

$$s(z) = \begin{cases} 1, & z \geq 0 \\ 0, & z < 0. \end{cases} \tag{2}$$

Представим некоторое изображение в виде вектора признаков $\mathbf{x} = (x^1, \dots, x^S)$. Значения признаков x^1, \dots, x^S являются значениями всех возможных МБ-ЛБШ. Общее число возможных мультиблочных ЛБШ S определяется размером изображений. Так, для изображений разрешением 8x6 пикселей число возможных МБ-ЛБШ равно 45 (24 блока 3x3, 12 блоков 6x3, 6 блоков 3x6 и 3 блока 6x6, как показано на рис. 3). Число возможных признаков быстро возрастает с увеличением разрешения изображений (так, для изображений разрешением 21x15 это число $S = 2450$, а изображений 21x21 – $S = 4900$), поэтому при обучении лучше использовать изображения небольшого разрешения.

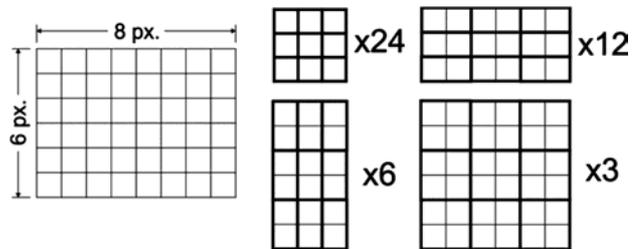


Рис. 3. Возможные МБ-ЛБШ для изображения разрешением 8x6 пикселей

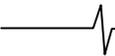
Построение и обучение классификатора

Построим классификатор $F(\mathbf{x})$, который по описанию изображения в виде вектора $\mathbf{x} = (x^1, \dots, x^S)$, состоящего из значений мультиблочных локальных шаблонов, относит это изображение либо к классу «глаз» e , либо к классу «не глаз» \bar{e} :

$$F(\mathbf{x}) = \begin{cases} +1, & \text{если } \mathbf{x} \in e \\ -1, & \text{если } \mathbf{x} \in \bar{e}. \end{cases} \tag{3}$$

Поиск итогового сильного классификатора $F(\mathbf{x})$ осуществляется с помощью алгоритма бустинга в виде суперпозиции T слабых классификаторов $f_i(\mathbf{x})$ по формуле

$$F(\mathbf{x}) = \text{sign}\left(\sum_{i=1}^T f_i(\mathbf{x})\right) \tag{4}$$



на основе обучающей выборки $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)$ из N изображений, представленных в виде вектора признаков \mathbf{x}_i и метки класса y_i . Значение $y_i = +1$ соответствует изображению «глаза», а $y_i = -1$ – изображению «не глаза».

На каждой стадии t необходимо найти все возможные слабые классификаторы $f^k(\mathbf{x})$, которые ищутся в виде деревьев решений [19]:

$$f^k(\mathbf{x}) = f^k(x^1, \dots, x^k, \dots, x^S) = \begin{cases} a_0, & \text{если } x^k = 0 \\ \dots \\ a_j, & \text{если } x^k = j \\ \dots \\ a_{255}, & \text{если } x^k = 255. \end{cases} \quad (5)$$

Коэффициенты a_j вычисляются по всей обучающей выборке согласно соотношению:

$$a_j = \frac{\sum_i^N w_i y_i \delta(x_i^k = j)}{\sum_i^N w_i \delta(x_i^k = j)}, \quad (6)$$

где w_i – веса элементов выборки на шаге t . Начальные веса элементов выборки, принадлежащих классу e , выбирались больше, так как количество экземпляров «глаз» в использованной обучающей выборке в несколько раз превышало число экземпляров «не глаз».

Из всех возможных слабых классификаторов $f^k(\mathbf{x})$ на шаге t выбирается классификатор $f_t(\mathbf{x})$ с минимумом среднеквадратической ошибки:

$$f_t(\mathbf{x}) = \min_{f^k \in F} \sum_{i=1}^N w_i (y_i - f^k(\mathbf{x}_i))^2. \quad (7)$$

После этого элементам выборки присваиваются новые веса согласно соотношению:

$$w_i' = w_i e^{-y_i f_t(\mathbf{x}_i)}, \quad (8)$$

что позволяет увеличить вес сложных примеров, неверно классифицированных на данном шаге. Стартовые веса и веса на каждом шаге нормируются согласно условию:

$$\sum_{i=1}^N w_i = 1. \quad (9)$$

Так как каждый слабый классификатор, отобранный на каждой из t стадий, требует для вычисления только одного МБ-ЛБШ, значение функции $F(\mathbf{x})$ может быть найдено с помощью T отобранных МБ-ЛБШ. Таким образом, количество необходимых признаков для оценки схожести некоторого изображения с изображением глаза сокращается с величины S (порядка нескольких тысяч) до T (около 10–20).

Итерационный алгоритм локализации центров глаз

Обучим три разных классификатора $F'(\mathbf{x})$, $F''(\mathbf{x})$, $F'''(\mathbf{x})$ согласно описанной выше схеме для трех масштабов изображения глаз (рис. 4). Каждый классификатор является суперпозицией слабых классификаторов и представляется в виде выражения (4). Опуская в нем функцию $sign$, получим функцию схожести некоторого изображения с изображением глаза. Чем выше значение, тем более похоже изображение, представленное вектором признаков \mathbf{X} , на изображение глаза.

Предлагаемый новый алгоритм локализации определяет положение глаза согласно итерационному подходу, позволяющему на каждой стадии улучшать оценку местоположения центра глаза и подстраиваться к качеству и сложности изображения лица (рис. 5).

В начале работы алгоритма изображение лица представляется в грубом масштабе. Внутри области поиска центра глаза для каждого пикселя вычисляется значение функции $F'(\mathbf{x})$. Отбирается m_1 пикселей с наибольшими значениями этой функции. Выбранные точки являются грубой оценкой положения глаза.

Если разрешение изображения лица является очень низким (линейный размер лица h меньше заданного порога H'_{\min}) или функция $F'(\mathbf{x})$ имеет низкое значение для последнего из отобранных пикселей $F'_{m_1}(\mathbf{x}) < \Pi'$, то это говорит о крайне невысоком качестве изображения или о наличии на лице очков со значительными бликами или перекрытием глаз. В этом случае работа алгоритма прекращается, и в качестве положения глаза выбирается пиксель с максимальным значением $F'(\mathbf{x})$. Если же качество и размер изображения лица являются удовлетворительными, происходит переход ко второй ступени алгоритма. Здесь изображение масштабируется, для отобранных пикселей и соседних с ними в новом масштабе считается значение функции $F''(\mathbf{x})$. Отбирается m_2 пикселей с наибольшими значениями этой функции. Выбранные пиксели являются более точной оценкой положения глаза.

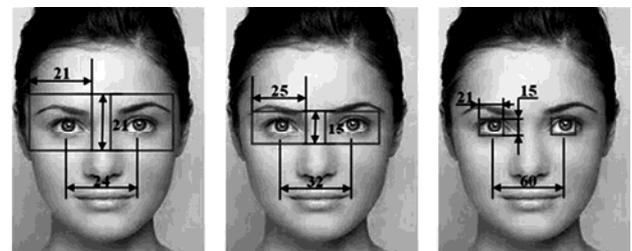


Рис. 4. Три масштаба изображения глаз для построения классификаторов $F'(\mathbf{x})$, $F''(\mathbf{x})$, $F'''(\mathbf{x})$

Если изображение лица довольно качественное и имеет относительно высокое разрешение, в качестве итогового решения о положении центра глаза выбирается пиксель с наивысшим значением функции $F'''(\mathbf{x})$, в противном случае – функции $F''(\mathbf{x})$.

При работе с изображениями лиц очень высокого разрешения число стадий и соответствующих классифи-

каторов $F(x)$ можно увеличить, что позволит определить положение центра глаза с еще большей точностью. Для этого надо предварительно обучить классификаторы для новых масштабов изображений глаз.

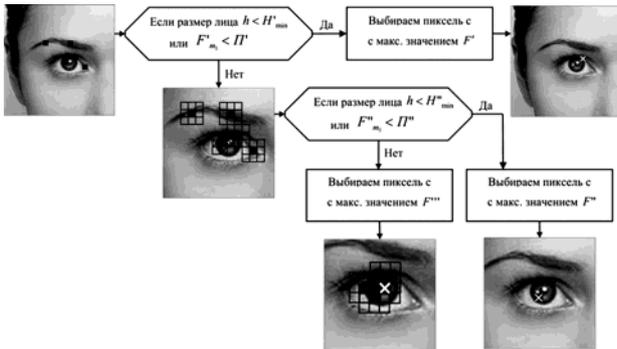


Рис. 5. Блок-схема работы

итерационного алгоритма локализации центров глаз

Результаты моделирования

Предлагаемый в работе алгоритм на основе мультиблочных локальных бинарных шаблонов, а также байесовский и градиентный алгоритмы локализации центров глаз [20, 21] программно реализованы и протестированы на базах изображений BioID [22] и FERET [23]. Для обучения байесовского локализатора и классификаторов, используемых в алгоритме на основе МБ-ЛБШ, из 3363 фронтальных изображений базы FERET с подписанными координатами левого и правого глаз выбиралась 1000 изображений. Для градиентного алгоритма локализации обучение не требуется. Для оставшихся 2363 изображений базы FERET производилось детектирование лиц стандартным алгоритмом Виолы-Джонса. Он сработал корректно на 2350 изображениях, которые и использовались для тестирования алгоритмов.

База BioID состоит из 1521 изображения с подписанными координатами глаз. Для тестирования на ней использовались алгоритмы, ранее обученные на первой тысяче изображений базы FERET, а также градиентный алгоритм, не требующий обучения. На 1469 изображениях из базы детектор лиц сработал корректно. На этих изображениях и проводилось тестирование алгоритмов локализации глаз.

В качестве критерия оценки точности локализации положений центров глаз используется нормализованная ошибка [17]:

$$err = \frac{\max(\|l - l_g\|, \|r - r_g\|)}{\|l_g - r_g\|}, \quad (10)$$

где l_g и r_g – координаты центров левого и правого глаза соответственно, определенные человеком вручную, а l и r – координаты, найденные с помощью алгоритма.

Как уже отмечалось, неточная локализация глаз может привести к существенному росту ошибок распознавания на последующем этапе обработки изображения. Тесты на стандартной базе FERET показали, что ошибки локализации $err > 0,15$ приводят к снижению уровня верного распознавания с 98% до 50% и ниже, поэтому такие ошибки будем считать грубыми. От алгоритма локализа-

ции требуется избегать таких ошибок.

На рис. 6 и рис. 7 представлены графики сравнения алгоритмов при тестировании на базах FERET и BioID соответственно. По оси абсцисс отсчитывается нормализованная ошибка, по оси ординат – доля изображений в базе, на которых алгоритм локализации глаз дал ошибку, меньшую соответствующего значения на оси абсцисс. Например, доля изображений, равная 0,96, при $err = 0,15$ показывает, что на 96% изображений из базы ошибка локализации глаз не превышает показателя 0,15.

Анализ зависимостей показывает, что разработанный итерационный алгоритм на основе МБ-ЛБШ практически не дает грубых ошибок локализации ($err > 0,15$, что соответствует смещению в среднем на половину ширины глаза от истинного положения). Только на 1% изображений из базы FERET и на 4% из базы BioID ошибка при локализации глаз превышает 0,15. Байесовский локализатор с аналогичной ошибкой работает на 2% изображений из базы FERET и 12% изображений из базы BioID. Для градиентного локализатора результаты еще хуже: 10% и 17% соответственно.

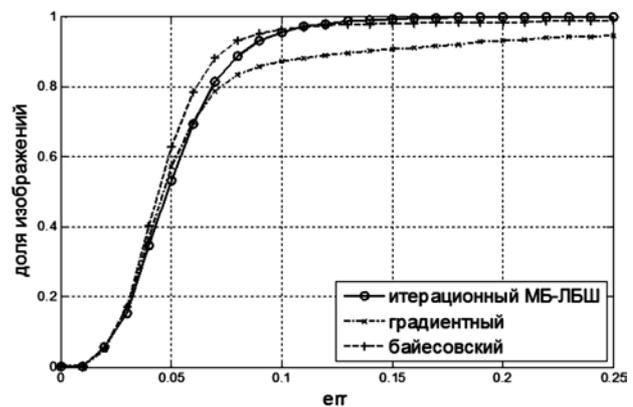


Рис. 6. Результаты тестирования алгоритмов локализации центров глаз на базе FERET

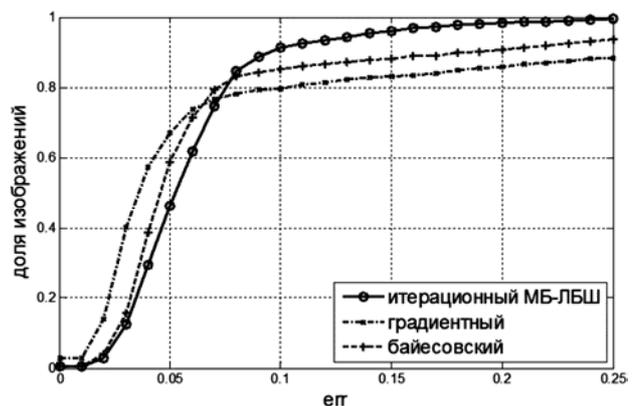


Рис. 7. Результаты тестирования алгоритмов локализации центров глаз на базе BioID

Снижение числа грубых ошибок разработанного алгоритма локализации на основе МБ-ЛБШ происходит за счет первичной оценки местоположения глаза классификатором $F'(x)$ в самом крупном масштабе. За счет адаптивной подстройки к изображениям низкого качества преимущество нового метода особенно заметно на изображениях из более сложной базы BioID. Использование



классификаторов $F''(x)$ и $F'''(x)$ позволяет новому алгоритму иметь сопоставимые результаты с другими методами локализации на простых изображениях, где нормализованная ошибка всех алгоритмов не превышает 0,1.

Результаты тестирования производительности алгоритмов представлены в табл. 1. В качестве параметра сравнения вычислялось среднее время поиска глаз на лице из базы FERET (среднее разрешение изображения лица 170x170 пикселей).

Таблица 1. Сравнение алгоритмов локализации глаз по производительности

Алгоритм	Время поиска глаз на изображении лица
Градиентный	587 мс
Байесовский	367 мс
Итерационный МБ-ЛБШ	44 мс

Видно, что новый алгоритм практически на порядок превосходит по скорости работы другие алгоритмы локализации, позволяя производить поиск положения глаз в видеопотоке в режиме реального времени.

Заключение

Предложен итерационный алгоритм локализации центров глаз на основе мультиблочных локальных бинарных шаблонов, адаптирующийся к качеству и сложности изображения лица. Разработанный алгоритм практически не дает грубых ошибок локализации ($err > 0,15$). Только на 1% изображений из базы FERET и на 4% из базы BioID ошибка при локализации глаз превышает 0,15. Рассматриваемый итерационный алгоритм практически на порядок превосходит по скорости работы другие используемые алгоритмы локализации глаз, позволяя производить поиск положения глаз в видеопотоке в режиме реального времени.

Работа выполнена при поддержке проекта № 1060 в рамках базовой части государственного задания на НИР ЯРГУ им. П.Г. Демидова и грантов РФФИ (№ 15-07-08674 и № 15-08-99639).

Литература

1. Дворкович В.П., Дворкович А.В. Цифровые видеоинформационные системы (теория и практика). – М.: Техносфера, 2012. – 1009 с.
2. Форсайт Д.А., Понс Д. Компьютерное зрение. Современный подход // М.: «Вильямс», 2004.
3. Szeliski R. Computer Vision: Algorithms and Applications – Springer, 2010.
4. Алпатов Б.А., Муравьев В.С., Стротов В.В., Фельдман А.Б. Исследование эффективности применения алгоритмов анализа изображений в задаче навигации беспилотных летательных аппаратов // Цифровая обработка сигналов. 2012. № 3. С. 29–34.
5. Алпатов Б.А., Бабаян П.В., Смирнов С.А., Масленников Е.А. Алгоритм предварительного оценивания пространственной ориентации объекта с помощью дескриптора внешнего контура // Цифровая обработка сигналов. 2014. № 3. С. 43–46.
6. Никитин А.Е., Хрящев В.В., Приоров А.Л., Матвеев Д.В. Разработка и анализ алгоритма распознавания лиц на основе локальных квантованных шаблонов // Нели-

нейный мир. 2014. № 8. С. 35–42.

7. Kriegman D., Yang M.H., Ahuja N. Detecting faces in images: A survey // IEEE Trans. on Pattern Analysis and Machine Intelligence. 2002. V. 24. № 1. P. 34–58.

8. Hjelmas E. Face detection: A Survey // Computer vision and image understanding. 2001. V. 83. № 3. P. 236–274.

9. Zhao W., Chellappa R., Phillips P., Rosenfeld A. Face recognition: A literature survey // ACM Computing Surveys (CSUR). 2003. V. 35, № 4. P. 399–458.

10. Marques J., Orlans N.M., Piszcz A.T. Effects of eye position on eigenface- based face recognition scoring // Technical Paper of the MITRE Corporation. October 2000. 7 p.

11. Никитин А.Е., Степанова О.А., Студенова А.А., Хрящев В.В. Локализация положений центров глаз на изображении лица // Докл. 17-й межд. конф. «Цифровая обработка сигналов и ее применение» (DSPA-2015). Москва, 2015. Т. 2. С. 719–723.

12. Riopka T., Boulton T. The eyes have it // Proc. of the ACM SIGMM Multimedia Biometrics Methods and Applications Workshop. 2003. P. 9–16.

13. Zhu Z., Fujimura K., Ji Q. Real-time eye detection and tracking under various light conditions // Proc. of the Symposium on Eye Tracking Research and Applications. 2002. V. 25. P. 139–144.

14. Zhu Z., Ji Q., Robust real-time eye detection and tracking under variable lighting conditions and various face orientations // Computer Vision and Image Understanding. 2005. 98 (1). P. 124–154.

15. Wang P., Green M., Ji Q., Wayman J. Automatic eye detection and its validation // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition. 2005. V. 3. P. 164–172.

16. Li G. An Efficient Face Normalization Algorithm Based on Eyes Detection // Proc. of the IEEE International Conference on Intelligent Robots and Systems. 2006. P. 3843–3848.

17. Song F., Tan X., Chen S., Zhou Z.H. A literature survey on robust and efficient eye localization in real-life scenarios // Pattern Recognition. 2013. V. 46(12). P. 3157–3173.

18. Ефимов И.Н. Локальные бинарные шаблоны медианного пикселя – эффективные информативные признаки технологии распознавания образов // Цифровая обработка сигналов. 2015. № 1. С. 61–65.

19. Zhang L., Chu R., Xiang S., Liao S., Li S.Z. Face Detection Based on Multi-Block LBP Representation // Advances in Biometrics, Lecture Notes in Computer Science. 2007. P. 11–18.

20. Timm F., Barth E. Accurate Eye Centre Localisation by Means of Gradients // Proc. of the International Conference on Computer Theory and Applications (VISAPP), 2011. V. 1. P. 125–130.

21. Everingham M.R., Zisserman A. Regression and classification approaches to eye localization in face images // IEEE International Conference on Automatic Face & Gesture Recognition. 2006. P. 441–446.

22. BioID face database // URL: <https://www.bioid.com/About/BioID-Face-Database>.

23. Phillips P.J., Moon H., Rauss P.J., Rizvi S. The FERET evaluation methodology for face recognition algorithms // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2000. V. 22(10). P. 1090–1104.