

УДК 004.9

Анализ современных подходов к процедурам узнавания и распознавания лиц для систем биометрического контроля

А. Я. Аноприенко, С. В. Иваница, Р. В. Лычагин
Донецкий национальный технический университет
ivanitsa-serg@rambler.ru

Аноприенко А. Я., Иваница С. В., Лычагин Р. В. Анализ современных подходов к процедурам узнавания и распознавания лиц для систем биометрического контроля. Определены поэтапные задачи для качественного распознавания лица в технологии распознавании объектов. Проанализированы основные современные алгоритмы распознавания лиц. Предложена общая идея структуры базы данных и поиска в рамках разрабатываемой системы динамического распознавания лиц для идентификации и аутентификации личности.

Ключевые слова: распознавание лиц, фильтр Габора, метод Виолы-Джонса, графы, FERET.

Введение

Интерес к процессам идентификации и аутентификации личности в современном мире набирает все возрастающую практическую потребность: от систем охраны и верификации банковских карточек до криминалистической экспертизы и идентификации преступников при рецидиве. Поскольку, одним из однозначно различных критериев личности, является человеческое лицо, то в аспекте идентификации (а позже, и аутентификации) личности следует понимать процедуру динамического распознавания лиц.

Задача эффективного поиска и идентификации человеческого лица, не является для компьютерных систем тривиальной и вполне очевидной для разработчиков специализированного программного обеспечения. Работа со слабоконтрастными объектами (с точки зрения восприятия компьютерной системой человеческого лица, в отличие от естественной идентификации лица человеческим глазом) является наиболее значимой для развития искусственного интеллекта (ИИ), делающей акцент на системах компьютерного (кибернетического) зрения, которые кроме процедур распознавания обеспечивают кодирование/декодирование и хранение цифровых изображений лиц. Такие системы можно рассматривать как в свете классической проблемы восприятия, так и в свете новых подходов [1] к распознаванию объектов.

Общая структура распознавания лиц

Несмотря на большое разнообразие алгоритмов распознавания лиц, можно выделить общую структуру данного процесса [2], которая

представлена на рис. 1.

На первом этапе производится обнаружение и локализация лица на изображении (наиболее эффективным является использование метода Виолы-Джонса). Слежение подразумевает более упрощенные способы локализации лица (т.к. лицо к данному моменту однозначно определено) на последующих кадрах непрерывного видео.

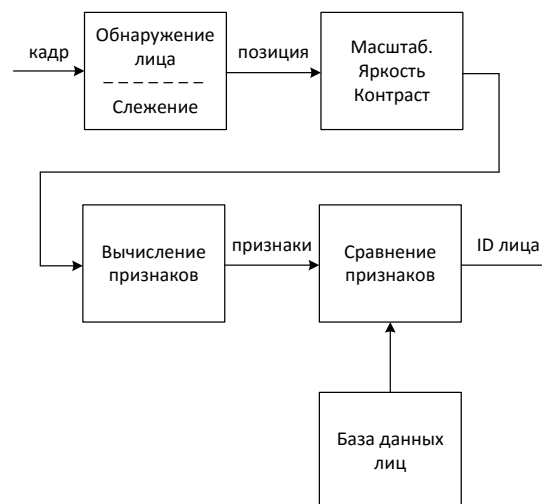


Рисунок 1 – Общая структура распознавания лиц

На втором этапе производится выравнивание изображения в найденной области (геометрическое и яркостное преобразование, применение фильтров). Вычисление и сравнение признаков варьируется между методами, и при этом все сводится к определенному сравнению вычисленных признаков с заложенными в базу данных эталонами. В данной работе будет также рассмотрено использование метода гибкого

сравнения на графах.

Метод Виолы-Джонса

В методе Виолы-Джонса используется интегральное представление изображения – матрица, которая совпадает по размерам с исходной и в каждом ее элементе хранится сумма всех элементов, находящихся левее и выше данного [3].

Элементы представлены в матричном виде и рассчитываются по следующей формуле (1):

$$I(x, y) = \sum_{\substack{x' < x \\ y' < y}} i(x', y') \quad (1)$$

где $i(x', y')$ — яркость пикселя исходного изображения.

Таким образом, каждый элемент $I(x, y)$ интегрального изображения содержит в себе сумму интенсивности пикселей в прямоугольнике от $(0, 0)$ до (x', y') .

Формирование интегрального изображения занимает линейное время, пропорциональное числу пикселей исходного изображения, и осуществляется за один проход. Расчет интегрального изображения I можно производить по рекуррентной формуле (2):

$$I(x, y) = i(x, y) + I(x-1, y) + I(x, y-1) - I(x-1, y-1) \quad (2)$$

Важнейшим достоинством интегрального представления изображения является возможность быстрого вычисления суммы пикселей произвольного прямоугольника (3), а также любой другой фигуры, которую можно аппроксимировать несколькими прямоугольниками.

$$S = \sum_{\substack{x_0 \leq x \leq x_1 \\ y_0 \leq y \leq y_1}} i(x, y) = I(A) + I(C) - I(B) - I(D) \quad (3)$$

Для описания искоемых объектов (лицо, руки, или пр. предметы) используются каскады из признаков. Сам по себе каскад Хаара — это набор примитивов (рис. 2), для которых считается их свертка с изображением. Используются самые простые примитивы, состоящих из прямоугольников и имеющих всего два уровня, +1 и -1.

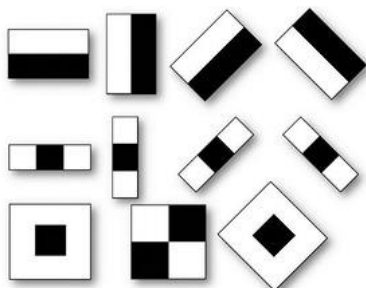


Рисунок 2 – Набор примитивов для описания объектов

При этом каждый прямоугольник используется несколько раз разного размера. Под сверткой тут подразумевается $s = X - Y$, где Y — сумма элементов изображения в темной области, а X — сумма элементов изображения в светлой области.

Такие свертки призваны структурировать информацию об объекте: например, в работе [6] показано что, для центра лица человека будет всегда отрицательная свертка, пример получения которой показан на рис. 3.

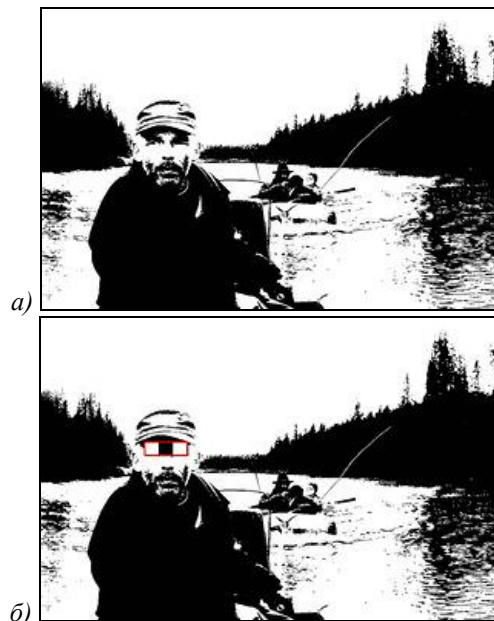


Рисунок 3 – Пример получения свертки для центра лица человека (а – исходное изображение, б – наложение свертки на центр лица) [6]

Формирование алгоритма слежения при повороте лица

Метод Виолы-Джонсона хоть и считается достаточно быстрым для использования в потоковых видео, но он имеет свои ограничения. При повороте лица на $\sim 15^\circ$ метод становится бесполезным. Поскольку полученные свертки достаточно точны, можно подгрузить дополнительные описатели либо фронтального лица, либо профильного, что приводит к потере результата, когда человек начинает поворачивать голову. Хотя подобное нельзя сказать при обнаружении более плоских объектов, например, глаза (он виден только во фронтальном расположении лица).

Так же если не настроить точность данного метода (размеры минимального и максимального квадрата поиска, количество пересечений, найденных областей и т. д.) уменьшается взаимодействие, но повышается вероятность правильного результата. Однако она также может варьироваться, поскольку качество изображения или потокового видео может быть получено с веб-

камер (плохое или среднее качество) или камер видео наблюдений до HD или HQ (высокое качество). При плохих разрешениях появляется много шумов или артефактов, которые могут исказить лицо (особенно находящиеся на расстоянии) настолько, что даже человек будет не в состоянии провести идентификацию по имеющемуся видеоряду. Также некоторые группы артефактов могут сформировать некоторое подобие лица, что приведет к ложному распознаванию, что потребует использовать более высокие настройки точности аппаратно-программного комплекса, чтобы в дальнейшем избежать подобных ошибок. Все эти особенности усложняют повторное обнаружение лица на следующих кадрах и делают слежение в реальном времени невозможным.

Так для ускорения повторного поиска можно уменьшить область повторного поиска. При частоте 30 кадров в секунду, предполагая нахождение объекта на предыдущем кадре, можно утверждать, что за 33 миллисекунды он сдвинется незначительно и тем самым при повторном поиске искать уже не на всем кадре, а лишь в увеличенном (x1.5) квадрате предыдущего результата (ROI). Так и при приближении или сдвиге объекта он будет найден в этой области.

Для решения проблем слежения при развороте объекта неизвестной стороной и потери лица методом Виолы-Джонсона, можно использовать алгоритм шаблонного соответствия (Template Matching) [9]. Сохраняя предыдущее найденное изображение и поиск подобного на ROI нового кадра. Особенность метода в том, что он всегда возвращает результат. Но и данный метод имеет свои проблемы так как со временем окно может сползти в область где нет лица и продолжать сравнение с последовательно предыдущим кадром тем самым оставаться на месте потеряв лицо. Поэтому вводятся ограничения по времени использования для этого метода. Как только отказывает основной метод распознавания, то начинает работать шаблонное соответствие, и при этом запускается таймер. Если таймер превышает некоторое ограничение (например, 2 секунды), лицо считается потерянным и сбрасываются все предыдущие значения. Однако подобное ограничение будет неприемлемым если лицо долго остается повернутым на некоторый неприемлемый для метода угол, так же при этом программа может утверждать, что затылок человека тоже лицо.

Таким образом, описанный выше алгоритм работы программы представлен в виде блок-схемы на рис. 4.

Метод гибкого сравнения на графах (Elastic graph matching)

Суть метода сводится к эластичному

сопоставлению графов, описывающих изображения лиц [2, 5]. Лица представлены в виде графов со взвешенными вершинами и ребрами. На этапе распознавания один из графов — эталонный — остается неизменным, в то время как другой деформируется с целью наилучшей подгонки к первому. В подобных системах распознавания графы могут представлять собой как прямоугольную решетку, так и структуру, образованную характерными (антропометрическими) точками лица (рис. 5). В вершинах графа вычисляются значения признаков, чаще всего используют комплексные значения фильтров Габора или их упорядоченных наборов (рис. 6) — Габоровских вейвлет (строки Габора), которые вычисляются в некоторой локальной области вершины графа локально путем свертки значений яркости пикселей с фильтрами Габора (рис. 7).

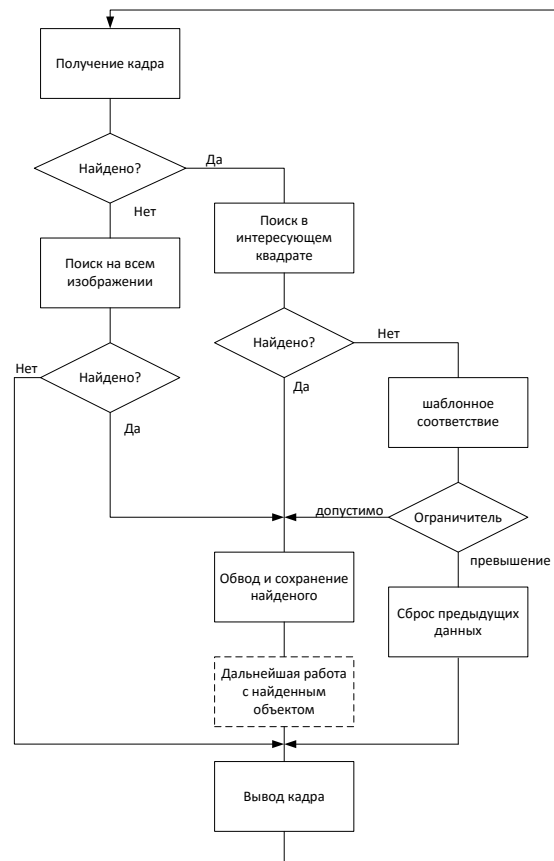
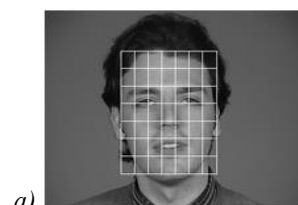


Рисунок 4 – Блок-схема работы алгоритма слежения



а)

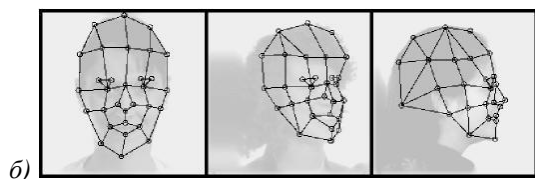


Рисунок 5 – Пример структуры графа для распознавания лиц (а – регулярная решетка, б – граф на основе антропометрических точек лица)

Ребра графа взвешиваются расстояниями между смежными вершинами. Различие (расстояние, дискриминационная характеристика) между двумя графами вычисляется при помощи некоторой ценовой функции деформации, учитывающей как различие между значениями признаков, вычисленными в вершинах, так и степень деформации ребер графа.

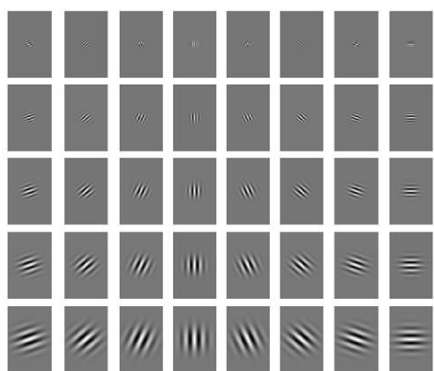


Рисунок 6 – Набор (банк, jet) фильтров Габора

Входное изображение	Пример двух фильтров Габора	Результат свертки входного изображения лица и фильтров Габора

Рисунок 7 – Пример свертки изображения лица с двумя фильтрами Габора

Деформация графа происходит путем смещения каждой из его вершин на некоторое расстояние в определенных направлениях относительно ее исходного местоположения и выбора такой ее позиции, при которой разница между значениями признаков (откликов фильтров Габора) в вершине деформируемого графа и соответствующей ей вершине эталонного графа будет минимальной. Данная операция

выполняется поочередно для всех вершин графа до тех пор, пока не будет достигнуто наименьшее суммарное различие между признаками деформируемого и эталонного графов. Значение ценовой функции деформации при таком положении деформируемого графа и будет являться мерой различия между входным изображением лица и эталонным графом. Данная «релаксационная» процедура деформации должна выполняться для всех эталонных лиц, заложенных в базу данных системы. Результат распознавания системы — эталон с наилучшим значением ценовой функции деформации (рис 8).

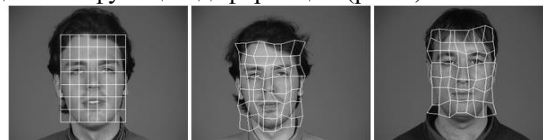


Рисунок 8 – Пример деформации графа в виде регулярной решетки

В отдельных публикациях указывается 95-97%-ая эффективность распознавания даже при наличии различных эмоциональных выражений и изменении ракурса лица до 15°. Однако разработчики систем эластичного сравнения на графах ссылаются на высокую вычислительную стоимость данного подхода. Например, для сравнения входного изображения лица с 87 эталонными тратилось приблизительно 25 секунд при работе на параллельной ЭВМ с 23 транспьютерами [4]. В других публикациях по данной тематике время либо не указывается, либо говорится, что оно велико.

К недостаткам данного метода относится высокая вычислительная сложность процедуры распознавания, а также низкая технологичность при запоминании новых эталонов и линейная зависимость времени работы от размера базы компьютерного изображения лица.

Основные проблемы [4, 7], связанные с разработкой систем распознавания лиц представлены на рис. 9, 10.



Рисунок 9 – Проблема неравномерной освещенности лица



Рисунок 10 – Проблема положения головы
(в подобном случае лицо – трехмерный объект)

С целью оценки эффективности предложенных методов распознавания лиц, агентство DARPA и исследовательская лаборатория армии США разработали программу FERET (FacE REcognition Technology) [8].

В масштабных тестах программы FERET принимали участие алгоритмы, основанные на гибком сравнении на графах и всевозможные модификации метода главных компонент (PCA). Эффективность всех алгоритмов была примерно одинаковой. В этой связи трудно или даже невозможно провести четкие различия между ними (особенно если согласовать даты тестирования).

Для фронтальных изображений, сделанных в один и тот же день, приемлемая точность распознавания, как правило, составляет 95%. Для изображений, сделанных различными средствами при разном освещении, точность, как правило, падает до 80%. Для изображений, сделанных с разницей в год, точность распознавания составила примерно 50%. При этом стоит заметить, что 50% — это более чем приемлемая точность работы системы подобного рода.

Ежегодно FERET публикует отчет о сравнительном испытании современных систем распознавания лиц [8] на базе лиц более одного миллиона. При этом в последних отчетах не раскрываются принципы построения систем распознавания, а публикуются только результаты работы коммерческих систем. На сегодняшний день лидирующей является система NeoFace [10] разработанная компанией NEC.

Общая идея структуры базы данных и поиска

В ходе работы над данным проектом планируется разработать систему динамического распознавания лиц и идентификации личности.

Для этого с помощью метода Виолы-

Джонсона локализовать лицо на изображении. Используя фильтр Габора с набором динамических параметров определить антропологические точки лица. По ним построить граф. Используя различные методы сравнения графов определить или разработать наиболее эффективный.

Создать динамическую базу данных с набором описателей лиц, привязанных к личностям. Предполагается хранение минимум двух записей для каждого лица: одна — с коэффициентами, рассчитанными на меньшем изображении (хуже качество), вторая — в оригинальном или большем формате (выше качество). Причем второй записи у отдельных пользователей может и не быть (рис. 11), так как оригинал не был достаточно велик для его создания (увеличение изображения приводит к размытию и потере качества).

Подобная структура записей (рис. 12) позволит теоретически ускорить поиск, сделав упрощенное сравнение сначала на всех записях, а после уточнить уже на полученном множестве.

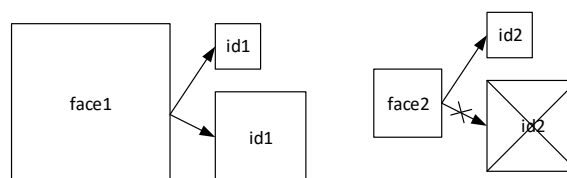


Рисунок 11 – Предложенное хранение минимум
двух записей для каждого лица

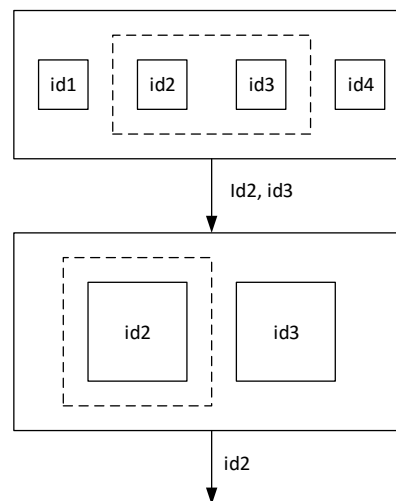


Рисунок 12 – Предложенная структура
записей лица

Планируется, сделать возможность динамически без участия пользователя создавать новые записи, нераспознанных лиц, и привязывать их к датам обнаружения.

Предполагается, возможность применения или разработки фильтров для повышения качества

изображения (в рамках данной темы) и решения проблем с плохим освещением.

Выводы

Компьютерное зрение — развивающаяся отрасль программирования, но при этом востребована и имеет большой спектр применения. Функцию идентификации людей на фотографиях активно используют в программном обеспечении для управления фотоальбомами (Picasa, iPhoto и др.). Скомбинировав ее с реальными параметрами, можно составлять альбомы по отдельному человеку. Идентификация также находит свое применение в системах безопасности, например при распознавании сотрудников объекта (учреждения).

При разработке программного обеспечения данного проекта идентификация личности до конца не реализована, однако выполнены нахождение лица и распознавание части признаков с помощью всего двух параллельных вейвлет Габора. С увеличением ряда фильтров и калибровки их параметров, возможны нахождение большего числа признаков и отсеки шумов. На основе полученных точек также появится возможность построения графа для выполнения сравнительного анализа и идентификации личности.

Основными трудностями на данном этапе выполнения проекта является настройка параметров фильтра Габора и разработка алгоритмов построения и сравнения графов.

В дальнейшем планируется использование средств и методов постбинарного компьютерного зрения [11, 12] для повышения точности вычислений. Так, например, некоторые помехи можно обозначать как неопределенность, вследствие чего помехи не будут просто отбрасываться, а сохраняться как возможные особенности лица.

Список использованной литературы

1. Технология распознавания лиц / Data Систем. Товары и технологии XXI. [Электронный ресурс]. — Режим доступа: <http://hardbroker.ru/pages/recognition>. — Заглавие с экрана.
2. Коломиец В. Анализ существующих подходов к распознаванию лиц. [Электронный ресурс] / Блог компании Синезис. — Режим доступа: <http://habrahabr.ru/company/synesis/blog/238129>.
3. OpenCV шаг за шагом. Интегральное изображение. [Электронный ресурс]. — Режим доступа: <http://robocraft.ru/blog/computervision/53-6.html>
4. Lades M. Distortion Invariant Object Recognition in the Dynamic Link Architecture. [Электронный ресурс] / Martin Lades, Jan C. Vorbruggen, Joachim Buhmann, Jorg Lange, Christoph v.d. Malsburg, Rolf P. Wurtz, Wolfgang Konen. — IEEE Transactions on Computers, vol. 42, No. 3, March 1993. — Режим доступа: http://www.cse.psu.edu/~rtc12/CSE597E/papers/o_bjrecLadesMarlsberg93.pdf.
5. Grother P. Face Recognition Vendor Test (FRVT). Performance of Face Identification Algorithms. / Patrick Grother, Mei Ngan. — Information Access Division National Institute of Standards and Technology. — May 26, 2014 — p. 138.
6. Мальцев А. Использование каскада Хаара для сравнения изображений [Электронный ресурс] / Мальцев Антон. — Режим доступа: <https://habrahabr.ru/post/198338/>. — Заглавие с экрана.
7. Face Recognition [Электронный ресурс] / NISTC Subcommittee on Biometrics and Identity Management Room. — 7 August, 2006. — Режим доступа: <http://www.biometrics.gov/documents/>.
8. Face Recognition Technology (FERET). Instructions on getting FERET database [Электронный ресурс]. / National Institute of Standards and Technology's web site. — Режим доступа: <http://www.nist.gov/itl/iad/ig/feret.cfm>.
9. Fast and robust face detection and tracking: mc-jesus/face_detect_n_track. [Электронный ресурс]. — Режим доступа: https://github.com/mc-jesus/face_detect_n_track.
10. NeoFace® Watch Generate Real-time Alerts Using Highly Accurate Face Recognition Technology. [Электронный ресурс]. / NEC Corporation of America, 2013. — Режим доступа: <https://www.necam.com/docs/?id=c8a08fd5-e79c-4f00-9f37-9919318cc772>.
11. Аноприенко А. Я. Постбинарный компьютерный и интервальные вычисления в контексте кодологической эволюции. / А. Я. Аноприенко, С. В. Иваница — Донецк, ДонНТУ, УНИТЕХ, 2011. — 248 с.
12. Аноприенко А. Я. Тетралогия, травывчисления и ноокомпьютинг. Исследования 2010–2012. / А. Я. Аноприенко, С. В. Иваница — Донецк: ДонНТУ, Технопарк ДонНТУ УНИТЕХ, 2012. — 308 с.

Anopriyenko A., Ivanitsa S., Lychagin R. A analysis of modern approaches to the procedures for recognition and face recognition for biometric systems. Defined milestone objectives for quality facial recognition technology in the recognition of objects. Analyzed the main current face recognition algorithms. The general idea of the database structure and retrieval developed in the framework of the dynamic face recognition systems for identification and authentication of identity.

Keywords: face recognition, Gabor filter, Viola-Jones method, graphs, FERET.

Анопрієнко О. Я., Іваниця С. В., Личагін Р. В. Аналіз сучасних підходів до процедур впізнавання та розпізнавання обличчя для систем біометричного контролю. Визначено поетапні завдання для якісного розпізнавання обличчя в технології розпізнавання об'єктів. Проаналізовано основні сучасні алгоритми розпізнавання обличчя. Запропоновано загальну ідею структури бази даних і пошуку в рамках розроблюваної системи динамічного розпізнавання обличчя для ідентифікації і аутентифікації особистості.

Ключові слова: розпізнавання осіб, фільтр Габора, метод Віолі-Джонса, графи, FERET.

Стаття поступила в редакцію 21.05.2016
Рекомендована к публікації д-ром техн. наук В.Н. Павльшином