

УДК 531/534: [57 + 61]; 616.713:616.12–089

Д. И. САМАЛЬ, И. И. ФРОЛОВ

## АЛГОРИТМ ПОДГОТОВКИ ОБУЧАЮЩЕЙ ВЫБОРКИ С ИСПОЛЬЗОВАНИЕМ 3D-МОДЕЛИРОВАНИЯ ЛИЦ

*Белорусский государственный университет информатики и радиоэлектроники*

*В работе представлен алгоритм подготовки выборки для обучения многоклассового классификатора машин опорных векторов (SVM). Описанный подход основан на моделировании возможных изменений внешности распознаваемого человека, путем введения дополнительных ракурсов съемки, условий освещения, углов наклона изображаемого фотопортрета человека. Полученные изменения искусственным путем непосредственно оказывают влияние на обучение классификатора, расширяя диапазон возможных вариаций исходного изображения. Классификатор, обученный на расширенной выборке, соответственно, готов в большей степени распознавать новые объекты. Основными рассматриваемыми параметрами для моделирования выбраны возраст, эмоциональные выражения лица, повороты головы, различные условия освещения, шумы, а также некоторые комбинации перечисленных параметров. Для моделирования дополнительных признаков использовано стороннее программное обеспечение FaceGen, позволяющее моделировать до 150 параметров и доступное в демо-версии для свободного скачивания.*

*В качестве классификатора для тестирования влияния вводимых модификаций обучающей выборки выбран SVM. Подготовка и предварительная обработка изображений выполнялась в несколько этапов: локализация области лица на изображении, оценка угла поворота и наклона, растяжение диапазона яркостей пикселей и эквализация гистограммы для выравнивания яркостно-контрастных характеристик обрабатываемых изображений, Масштабирование локализованной и обработанной области лица, Формирование вектора признаков отмасштабированного и обработанного изображения лица по методу главных компонент (алгоритм NIPALS), Обучение многоклассового SVM-классификатора.*

*Представленная методика расширения обучающей выборки носит практически ориентированный характер и позволяет с использованием 3D-моделей расширить диапазон обрабатываемых 2D-фотопортретов лиц, что положительно сказывается на результатах идентификации в системе распознавания лиц. Данный подход позволяет в определенной мере компенсировать недостаток исходных изображений для повышения качества идентификации.*

**Ключевые слова:** 3D-моделирование, предобработка, подготовка обучающей выборки

### Введение

Алгоритмы обнаружения области интереса (лица) на изображении ориентированы на устойчивость к качеству подаваемых на вход системы изображений – контрастности, яркости и т. д. Однако, даже при точной детекции области лица из-за разности условий освещения человека (рис. 1) при проведении фотосъемки процесс достоверной классификации остается весьма проблематичным.

Необходимо отметить, что на сегодняшний день не существует универсальных алгоритмов улучшения изображения и выравнивания яркостно-контрастных характеристик, применяемых для задач обработки изображений различного рода и, в частности, для распознава-

ния лиц. Остается открытым вопрос о выборе применяемых или разработке новых алгоритмов цифровой обработки изображений для задачи распознавания лиц.

В системе биометрической идентификации по фотопортрету компании Technest ([www.genextech.com](http://www.genextech.com)) использована технология 3D-моделирования для добавления в базу лиц фотопортретов недостающих ракурсов с отличающимися позами, освещением и выражениями лиц, однако данный ресурс ничего не сообщает о нивелировании возрастных отличий.

В работе [1] рассмотрена возможность использования фильтров, шумов с целью расширения обучающего множества для настройки биометрических систем распознавания. Одна-



Рис.1. Образцы изображений с различной освещенностью



Рис. 2. Сопоставление фотографии и 3D модели

ко данная работа направлена в большей степени на моделирование шумов и не рассматривает вопросы моделирования освещения, мимики, возрастных отличий.

В данной работе предложен комплексный подход для решения задачи по формированию и расширению обучающей выборки из изображений с разными характеристиками яркости, контрастности, условий освещенности, эмоциональных выражений лица, возрастных отличий. Демонстрируемый результат достигается путем синтеза 3D-моделирования лица с изменяемыми параметрами и методов цифровой обработки изображений для имитации шумов, а также приведения к единым условиям обработки.

SVM-классификатор, так же как и нейронные сети, относится к классификаторам статистического типа, поэтому чем больше изображений лица в различных его вариациях (освещение, мимика, возраст, поворот, наклон) будет представлено системе на этапе обучения, тем большей обобщающей способностью будет обладать данный классификатор и тем выше процент распознавания будет на этапе распознавания лиц (данное утверждение будет проверено на этапе тестирования алгоритмов).

#### Параметры 3D-моделирования лица

Для построения 3D-модели лица необходимо наличие минимум одного фотопортрета полного фронтального типа с разрешением, не

ниже требуемого (не менее 90 пикселей между центрами зрачков глаз) и без отклонений, поворотов и наклонов головы. Построение 3D-модели выполняется с использованием стороннего программного обеспечения FaceGen <http://www.facegen.com> и [2].

При использовании технологии 3D-моделирования наглядно был установлен факт высокой степени визуального соответствия фотографий реальных людей и их 3D-моделей (рис. 2). Построение виртуального набора лиц различных ракурсов, эмоций, освещения, шумов, возраста позволяет расширить обучающую выборку, а также нивелировать влияние искажающих факторов.

Самым распространенным фактором, влияющим на точность распознавания является изменение источников освещения, которые приводят к появлению теней на изображениях. Если обучающая выборка не содержит изображений, снятых различными вариантами освещения, то уровень распознавания системой таких лиц резко падает. На рис. 3 представлены варианты освещения модели лица (слева направо): сверху, рассеянное прямое освещение, источник света справа, источник света слева.

Важную роль в идентификации личности по цифровому фотопортрету играет также возраст при получении фотоснимков и момент проведения процедуры идентификации. В данном случае лицо человека подвержено процес-



Рис. 3. Моделирование различных условий освещения



Рис. 4. Модель изменения возраста, слева-направо: 30 лет, 40 лет, 50 лет, 60 лет

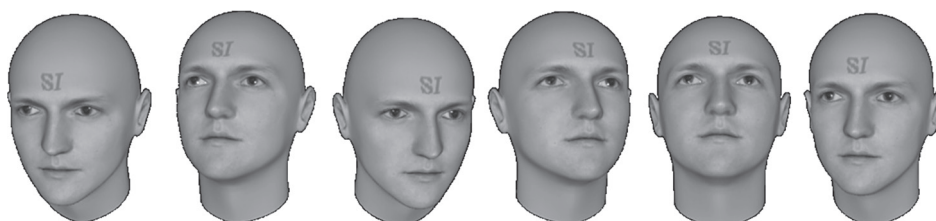


Рис. 5. Наклон и поворот головы в трех плоскостях при равномерном рассеянном освещении



Рис. 6. Эмоции: улыбка закрытая, удивление, страх, гнев

су старения и, как следствие, изменению черт лица (рис. 4).

Кроме того, важным моментом при распознавании является ракурс фотосъемки: отклонение головы от вертикального положения вправо-влево, поворот головы вокруг вертикальной оси вправо-влево, наклон головы вокруг горизонтальной оси вверх-вниз. Перечисленные изменения положения головы могут иметь место как в одной плоскости (рис. 5), так и одновременно во всех трех перечисленных (рис. 5).

Следующим фактором, также меняющим черты лица, является проявление человеческих эмоций (рис. 6).

Кроме того, для формирования более полной обучающей выборки с многочисленными вариантами освещения, наклонов/поворотов

головы, эмоциональных выражений лица, реализованы комбинации перечисленных вариантов для каждого фотопортрета. Учитывая фактор различных условий получения, хранения и передачи изображений, для нивелирования различных возможных шумов на изображении выполнено совмещение 3D-моделей лица и нескольких распространенных видов шумов (рис. 7).

Применение данных подходов позволяет расширить тренировочный набор фотопортретов в процессе формирования и обработки обучающей выборки, получая большее количество возможных вариаций условий съемки. Таким образом, повышается количество векторов признаков, описывающих изображения, и, соответственно, учитывая их отличия между собой, увеличивается и количество опорных



Рис. 7. Моделирование вариаций освещения, эмоций, возраста, шумов

векторов при обучении классификаторов, что повышает обобщающую способность классификаторов и, следовательно, приводит к увеличению коэффициента распознавания.

Если рассматривать ограниченный набор изменяемых параметров, то можно вычислить общее количество дополнительных изображений, которые строятся для расширения обучающей выборки по формуле  $images = n \times (n - 1) / 2$ , где  $images$  – общее число дополнительных изображений для каждого исходного фотопортрета;  $n$  – количество варьируемых параметров.

В проведенном исследовании использовано 17 параметров для расширения обучающей выборки – минимальный набор вариаций, позволяющий охватить сразу несколько влияющих факторов: угол поворота и наклона (7 позиций), возраст (2 позиции), эмоциональные выражения лица (4 позиции), освещение (4 позиции), что позволило увеличить обучающую выборку в 136 раз (т. к. для каждого изображения сгенерировано именно по 136 дополнительных изображений). Для того, чтобы не делать значительной долю зашумленных изображений, эффекты шума и фильтры применялись выборочно только для изображений лиц с нормальным рассеянным освещением без дополнительных измененных параметров, что позволило сформировать для 40 изображений каждого класса (случайно отобранных) эффекты шумов (рис. 7). В работе были использованы следующие преобразования: гауссов шум, импульсный шум (типа «соль и перец»). Таким образом, общее количество изображений, сгенерированных для каждого исходного изображения, составило 176 экземпляров.

Представленную последовательность операций по формированию обучающей выборки можно описать в виде следующего алгоритма.

Шаг 1. Формирование исходной выборки изображений лиц для обучения SVM-классификатора.

Шаг 2. Построение 3D-модели обрабатываемого лица с использованием стороннего программного обеспечения FaceGen ([www.facegen.com](http://www.facegen.com)) и генерирование дополнительных 176 изображений (количество и виды параметров – освещение, возраст, эмоциональные выражения лица, наличие усов, очков, наклонов/поворотов – устанавливается в зависимости от условий использования системы и требований пользователя) для каждого исходного фотопортрета в соответствии с выбранными изменяемыми параметрами фотосъемки.

Шаг 3. Оценка ракурса для каждого изображения (исходного и сгенерированных): предварительный поиск лица на изображении, поиск зрачков, вычисление угла поворота изображения и поворот до расположения глаз на одной горизонтальной линии, вычисление углов поворота  $\alpha_1^{new}$  и наклона  $\alpha_2^{new}$  головы.

Шаг 4. Локализация области лица по границам надбровных дуг с помощью алгоритма локализации лица Виола-Джонса.

Шаг 5. Предварительная обработка и нормализация изображений: растяжение диапазона яркостей пикселей и эквализация гистограммы для выравнивания яркостно-контрастных характеристик обрабатываемых изображений.

Шаг 6. Масштабирование локализованной и обработанной области лица до размеров  $200 \times 200$  пикселей (данное разрешение соответствует требованию о минимальном расстоянии в 90 пикс. между центрами зрачков глаз при локализации лица по границам надбровных дуг).

Шаг 7. Добавление локализованной и обработанной области лица в обучающий набор фотопортретов и соответствующую базу лиц системы распознавания.

Шаг 8. Формирование вектора признаков отмасштабированного и обработанного изображения лица по методу главных компонент (алгоритм NIPALS), используя 100 главных компонент для представления каждого век-



Рис. 8. Пример процедуры предобработки изображения

тора. Сохранение вектора признаков в обучающем наборе и соответствующей базе системы.

Шаг 9. Обучение многоклассового SVM-классификатора (4). Завершение алгоритма.

При разработке системы были проведены исследования и экспериментальным путем установлена целесообразность выполнения на первом этапе работы операции выделения области интереса – лица – на исходном изображении, а только на последующем шаге – реализация процедуры предварительной обработки цифрового изображения. Для демонстрации влияния именно фактора смоделированных изображений на конечный результат в данной работе в качестве нормализации изображения применялась только эквализация гистограммы изображения, без поворотов и иных алгоритмов увеличения робастности к шумам. Эффективность выполнения данных процедур в указанном порядке отображена на рис. 8. Изображения справа имеют более контрастный и насыщенный вид, а значит, содержат более информативные признаки по сравнению с более однотонным и низкоконтрастным изображением слева. Данные результаты объясняются влиянием фона и «лишних» частей фотопортрета (фон, одежда) при применении алгоритмов эквализации и нормализации гистограммы изображения [3], т. к. при их применении расчёты проводятся с учетом пикселей всего изображения, на котором фон зачастую оказывается либо темнее, либо светлее области интереса.

Выполнение процедуры улучшения изображения только на заданной области интереса (лицо) позволяет получить более контрастные изображения, в большей степени пригодные для дальнейшей процедуры редуцирования исходного пространства данных и распознавания образов.

С целью оценки эффективности применения предложенного подхода по формированию

и расширению обучающей выборки для обучения SVM-классификатора проведен соответствующий эксперимент. Для распознавания был обучен SVM-классификатор с параметрами обучения  $C = 8$ ,  $\gamma = 0,003125$ . Результаты эксперимента приведены в табл. 1.

Таблица 1. Влияние расширения объема обучающей выборки на коэффициент распознавания

Обучающая выборка	Число классов	Общее число изображений для обучения	Число изображений каждого класса	Коэффициент распознавания при идентификации, %
По 2 изображения каждого класса	700	1400	2	89,71
По 2 изображения каждого класса + дополнительный виртуальный набор лиц	700	124600	178	96,86

Необходимо отметить, что с увеличением числа обучающих изображений наблюдается и рост времени обучения классификатора. Однако данный недостаток не является определяющим, т. к. обучения происходит на стадии настройки системы и не влияет на скорость работы в режиме идентификации.

### Заключение

Полученные результаты подтверждают целесообразность применения разработанного алгоритма формирования обучающей выборки для обучения классификатора в задаче распознавания лиц. Применение 3D-моделирования лица позволяет в большой степени компенсировать отсутствие множества реальных фотопортретов с различными условиями фотосъемками, возрастными отличиями и присутствием эмоциональных выражений лица, на-

личием шумов. Разработанный алгоритм отличается от известных подходов в системах биометрической идентификации устойчивостью не только к изменениям освещения, ракурса и эмоций, но и к возрастным отличиям и шумам.

### Литература

1. Старовойтов В. В. Расширение обучающего множества для настройки биометрических систем распознавания / В. В. Старовойтов, Никодимов Д. Ю. // Обработка информации и управление в чрезвычайных и экстремальных ситуациях: материалы 4-й международной конференции. – Минск, 2004. – С. 204–209.
2. Yanushkevich S. N. Biometric-Based Decision Support Assistance in Physical Access Control Systems / S. Chague, B. Droit, O. Boulanov, S. N. Yanushkevich, V. P. Shmerko, A. Stoica // Bio-inspired Learning and Intelligent Systems for Security: Proceedings of ECSIS Symposium, Aug. 4–6, 2008, Edinburgh, UK. – Edinburg, 2008. – P. 11–16.
3. Гонсалес Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. – Москва: Техносфера. – 2005. – 1072 с.
4. Хайкин С. Нейронные сети: полный курс / С. Хайкин. – Москва: Вильямс. – 2006. – 1104 с.

### References

1. Starovojtov V. V. Rasshirenie obuchajushhego mnozhestva dlja nastrojki biometricheskikh sistem raspoznavanija / В. В. Старовойтов, Никодимов Д. Ю. // Obrabotka informacii i upravlenie v chrezvychajnyh i jekstremal'nyh situacijah: materialy 4-j mezhdunarodnoj konferencii. – Minsk, 2004. – S. 204–209.
2. Yanushkevich S. N. Biometric-Based Decision Support Assistance in Physical Access Control Systems / S. Chague, B. Droit, O. Boulanov, S. N. Yanushkevich, V. P. Shmerko, A. Stoica // Bio-inspired Learning and Intelligent Systems for Security: Proceedings of ECSIS Symposium, Aug. 4–6, 2008, Edinburgh, UK. – Edinburg, 2008. – P. 11–16.
3. Gonsales R. Cifrovaja obrabotka izobrazhenij / R. Gonsales, R. Vuds. – Moskva: Tehnosfera. – 2005. – 1072 s.
4. Hajkin S. Nejrornyie seti: polnyj kurs / S. Hajkin. – Moskva: Vil'jams. – 2006. – 1104 s.

Поступила  
10.10.2016

После доработки  
23.11.2016

Принята к печати  
28.11.2016

Samal D. I., Frolov I. I.

## ALGORITHM OF PREPARATION OF THE TRAINING SAMPLE USING 3D-FACE MODELING

*The algorithm of preparation and sampling for training of the multiclass qualifier of support vector machines (SVM) is provided. The described approach based on the modeling of possible changes of the face features of recognized person. Additional features like perspectives of shooting, conditions of lighting, tilt angles were introduced to get improved identification results. These synthetic generated changes have some impact on the classifier learning expanding the range of possible variations of the initial image. The classifier learned with such extended example is ready to recognize unknown objects better. The age, emotional looks, turns of the head, various conditions of lighting, noise, and also some combinations of the listed parameters are chosen as the key considered parameters for modeling. The third-party software 'FaceGen' allowing to model up to 150 parameters and available in a demoverison for free downloading is used for 3D-modeling.*

*The SVM classifier was chosen to test the impact of the introduced modifications of training sample. The preparation and preliminary processing of images contains the following constituents like detection and localization of area of the person on the image, assessment of an angle of rotation and an inclination, extension of the range of brightness of pixels and an equalization of the histogram to smooth the brightness and contrast characteristics of the processed images, scaling of the localized and processed area of the person, creation of a vector of features of the scaled and processed image of the person by a Principal component analysis (algorithm NIPALS), training of the multiclass SVM-classifier.*

*The provided algorithm of expansion of the training selection is oriented to be used in practice and allows to expand using 3D-models the processed range of 2D – photographs of persons that positively affects results of identification in system of face recognition. This approach allows to compensate in a certain measure a lack of initial images for identification improvement of quality.*



**Самаль Дмитрий Иванович**, кандидат технических наук, доцент. Заведующий кафедрой ЭВМ учреждения образования «Белорусский государственный университет информатики и радиоэлектроники».

*Сфера научных интересов: машинное обучение, обработка и анализ изображений, системы технического, распознавание образов, зрения, робототехника, параллельные вычисления, анализ данных.*

E-mail: samal@bsuir.by, dmitry\_samal@mail.ru,



**Фролов Игорь Иванович**, кандидат технических наук. Доцент кафедры ЭВМ учреждения образования «Белорусский государственный университет информатики и радиоэлектроники», г. Минск, Беларусь. Сфера научных интересов: машинное обучение, обработка изображений, распознавание образов, системы технического зрения.

E-mail: [frolov@bsuir.by](mailto:frolov@bsuir.by)