УДК 004.932.2

Н. Л. Щеголева, В. А. Петрова
Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина)

Проблемы оценки биометрических систем

Обсуждается оценка биометрических систем обработки изображений, основные требования и параметры сравнения. Рассматриваются стандарты оценки биометрических систем FRVT, их развитие, оцениваемые характеристики. Исследуются особенности оценки систем, использующих несколько категорий изображений, систем поиска изображений по содержанию и мультибиометрических систем. Показана роль биометрических стандартов и их роль в совмещении технологий, разработанных в разных странах. Определяются основные проблемы оценивания: тестирование систем на нерепрезентативных базах, что приводит к несоответствию заявленных результатов при работе в реальных условиях; сложность определения критериев оценки для систем поиска изображений по содержанию, а также недостаточное развитие методик оценки биометрических систем. Использование программных средств для моделирования биометрических систем позволяет определить структуры возможных вариантов систем, состав блоков и их функции и более просто и экономически эффективно сравнивать варианты систем.

Биометрические системы распознавания лиц, оценка биометрических систем, биометрические стандарты

Стремительное развитие вычислительной техники обеспечило рост производительности современного аппаратного обеспечения мобильных устройств, а встраиваемые в них цифровые видеокамеры достаточно светочувствительны и обладают возможностью стабилизации изображений. Повсеместное использование систем видеонаблюдения и контроля доступа, а также простота видеосъемки с помощью смартфона привели к тому, что на сегодня получены значительные объемы цифровых визуальных данных, а самой большой базой, содержащей изображения лиц, являются социальные сети. Согласно Mary Meeker's «Internet Trends Report» (2015) [1], в день загружается более 3.25 млрд изображений (исследование проводилась в социальных сетях Facebook, Snapchat, Instagram). При этом большинство загружаемых файлов содержат изображения лиц, ссылку на автора и его геопозицию, что не только открывает широкие возможности для биометрических систем, но и позволяет решать новые задачи.

Открывшиеся технические возможности позволили крупным исследовательским центрам, международным корпорациям и целым государствам начать тестирование своих программных средств для идентификации личности по изображению лица. Однако в скором времени выяснилось, что существующие методы не пригодны для обработки такого количества изображений. Ярким примером может служить проект FindFace, который использует сверточные нейронные сети. Разработчикам удалось добиться 93 % точности на базе в 10 000 изображений, но на больших базах точность системы падала, так, например, на базе в 1 млн изображений точность составляла всего 73 % [2], [3].

Современные системы не дают стабильного высокого (более 90 %) качества распознавания в произвольных реальных условиях. Во многом это обусловлено тем, что тестирование проводится только в лабораторных условиях с использованием ограниченного количества баз изображений лиц, такой подход не может гарантировать сохранение характеристик работы при эксплуатации системы в реальных условиях. Примером этого может служить приложение FaceIt компании Visionics. Производитель заявил в спецификации устойчивость программного комплекса к таким факторам, как изменение освещения, наличие усов и бороды, очков и даже пластических операций. На конференции Биометрического консорциума в 2002 г. были представлены результаты трехмесячного тестирования FaceIt [4]: верная идентификация происходила только в 51 % случаев, что значительно отличалось от заявленной точности системы. И это не единственный при-

мер подобного рода. Количество разработанных биометрических систем обработки изображений лиц увеличивается с каждым годом, что приводит к необходимости решения проблемы оценки разных систем, решающих конкретную прикладную задачу.

Стандарты оценки биометрических систем. Ситуация изменилась с появлением ряда стандартов для оценки биометрических систем, разработанных National Institute of Standards and Technology Standards.

Первой базой данных для всесторонней проверки биометрических алгоритмов по распознаванию изображений лиц была база данных, разработанная в рамках проекта Face Recognition Technology (FERET) [5], [6]. Базу данных FERET собирали в период с августа 1993 г. по июль 1996 г. База данных содержит 14 126 изображений 1199 человек. Опыт создания этой базы данных показал, что необходимо минимизировать влияние людей на этот процесс. Необходимо гарантировать, чтобы база данных была не слишком сложная и не слишком простая. Если база данных будет слишком простой (т. е. будет содержать качественные изображения лиц анфас с маленьким разбросом внутриклассовых характеристик), то ошибочные показатели будут близки к нулю и будет очень трудно определить разницу между конкурирующими системами, и наоборот. В идеале база данных должна содержать типичную информацию и позволять оценивать ограничения разных систем.

Следующим шагом стала разработка в 2000 г. Face Recognition Vendor Tests (FRVT), которые состояли из двух частей: оценки качества распознавания и удобства использования. Проект FRVT стал началом серии тестов, которые постоянно усложнялись, отражая, таким образом, технический прогресс и достижения в обработке и распознавании изображений лиц: FRVT2002 – появилась оценка производительности на реальных базах данных большого размера; FRVT2006 – изображения высокого разрешения (5–6 мегапикселей), изображения 3D-сканеров лиц, MultiSample (представление изображения лица одного и того же человека через отдельные его фрагменты), алгоритмы предобработки для компенсации ракурса и освещения; FRVT2006 – идентификация в базе данных преступников, идентификация визовых документов, распознавание лиц на видео, распознавание близнецов, изображение лиц в профиль, определение пола, оценка возраста, оценка ракурса, нейтральное выражение лица и т. д. Каждый год в проекте добавляются новые оцениваемые характеристики и расширяется круг задач, решаемых биометрическими системами, например, возможность противодействия системы спуфингу (обману) или ее реализуемость в мобильных устройствах.

Заметим, что эффективность биометрической системы обработки изображений определяется использованными в ней методами и моделями представления изображений, алгоритмами реализации, структурой, используемой вычислительной техникой.

Поскольку проверка эффективности систем должна выполняться на больших выборках реальных изображений, то к используемым алгоритмам предъявляется ряд требований [7]: робастность – устойчивость к возможным искажениям изображений; точность – сохранение структуры информации, содержащейся в изображениях; вычислительная реализуемость алгоритмов для работы в реальном времени. Структура системы должна учитывать: объем и структуру базы изображений; сценарий решаемой задачи; временные ограничения; физическую реализуемость.

Удовлетворение данных требований позволит обеспечить решение прикладной задачи обработки изображений; адаптацию системы к изменяющимся условиям получения исходных изображений; требуемую пропускную способность системы; безотказную работу системы в заданном режиме; простоту эксплуатации и сопровождения системы.

Основные параметры сравнения и оценки биометрических систем. Оценка законченной биометрической системы – комплексная и сложная задача [8], которая требует знаний экспертов во многих областях, включая статистику, информатику, разработку, бизнес, психологию и юриспруденцию. В процессе оценки биометрической системы определяются: уровень ошибок; надежность, доступность и ремонтопригодность; слабые места; обеспечиваемый уровень безопасности; удобство использования; пропускная способность; влияние человеческого фактора: скорость привыкания, появление проблем частной жизни в процессе использования системы; стоимость; материальная выгода при ее внедрении.

Процесс сравнения и оценки биометрических систем рассматривается в [6], [9]–[11]. Идеальной является ситуация, когда оценка выполняется независимым третьим лицом, которое оценивает

качество проектирования, управления и результаты выполненных функциональных тестов. Общая оценка состоит:

– из технологической оценки, заключающейся в сравнении конкурирующих систем по одной технологии на стандартизированной базе данных;

– оценки сценариев использования;

– эксплуатационной оценки, которая должна показать степень соответствия выполняемых функций для решения требуемых задач.

В [11] дается ряд рекомендаций, включающих размеры тестовой базы изображений, способы тестирования, перечисляются факторы, которые могут повлиять на работу биометрической системы, методологии сбора данных, метрики для оценки работы системы.

Системы, использующие несколько биометрических характеристик. Следует отметить, что по статистике в реальной базе данных, содержащей 100 000 и более портретов, примерно 0.5 % будут иметь «двойников», т. е. будут очень похожи. Для решения этой проблемы используются гетерогенные изображения лиц (Heterogeneous Face Recognition). Этот термин был введен (2013) B. F. Klare и A. K. Jain и означает представление изображений лиц разными категориями: в видимом и/или инфракрасном свете (разных длин волн), композиционными фотороботами и рисованными скетчами, в форме «rang image» и объектов 2D, 2.5D и 3D. Это позволяет расширить область использования лицевой биометрии, однако существенное различие характеристик категорий усложняет решение прикладных задач.

Другим подходом, позволяющим решить проблему сходства разных людей, является совместное использование различных биометрических характеристик человека (физических и поведенческих) в рамках мультибиометрических систем.

Основными преимуществами мультибиометрических систем являются:

– разнообразие используемых характеристик;

– возможность охвата большей части населения;

– альтернативное использование характеристик;

– произвольное комбинирование характеристик;

– интеграция всех биометрических характеристик как уникального идентификатора человека;

– ускорение процесса идентификации;

– повышение надежности и качества идентификации;

– расширение круга задач, решаемых биометрической системой;

– интеграция в единый комплекс нескольких систем: видеонаблюдения, дактилоскопческих, голосовых и т. д.

Проблемы оценки мультимодальных систем. Одно из основных отличий между оценкой одно- и мультимодальной биометрической системы – природа используемой базы данных. Значения метрик для оценки работы биометрической системы [8], [12]–[14], такие как точность, пропускная способность, масштабируемость, могут быть определены с высокой степенью точности, только если система проверена на большой представительной базе данных.

Большинство существующих мультимодальных систем проверены только на небольших базах данных. Связано это в первую очередь со значительными ресурсами, требующимися для создания репрезентативных выборок различных модальностей, используемых в таких базах данных. Следует отметить, что результаты исследований на реальных и виртуальных базах данных значтельно отличаются.

Другой проблемой, уникальной для мультимодальных систем [8], является тестирование сложной системы, имеющей и каскадные, и параллельные подсистемы. Это обусловлено сложной взаимосвязью между удобством пользователя и безопасностью в данных системах.

Системы поиска изображений по содержанию. Активно развивающиеся в последнее время системы поиска изображений по содержанию (content-based image retrieval – CBIR) используют описание изображения с помощью некоторых его характеристик, например цветовой гаммы, текстуры, формы и взаимного расположения объектов и трансформации полученной количественной информации в текстовую форму, описывающую изображение. Оценка подобия исходного изображения и других изображений базы данных биометрической системы выполняется на основании сравнения значений перечисленных характеристик. Для эффективного поиска необходима сложная «трансформация базы данных в лингвистическую модель с взвешенным набором весов для терминологических переменных» [15].

Выделяют две основные проблемы CBIR:

1) сенсорный разрыв (Sensory gap) – несоответствие описаний между объектом в окружающем мире и информационным описанием, полученным из его цифрового изображения;

2) семантический разрыв (Semantic gap) – отсутствие совпадения между информацией, которую можно извлечь из визуальных данных, и их интерпретацией для человека в данной ситуации.

Таким образом, CBIR опирается на визуальное сходство при определении семантического сходства. Все перечисленное делает достаточно сложным процесс оценки таких систем, поскольку основной проблемой является определение критериев, по которым можно оценить качество полученного результата. Это требует дополнительных исследований и стандартизации.

Биометрические стандарты. На сегодняшний день практически каждая страна мира реализует большое количество проектов развития и совершенствования биометрических систем. Этот процесс сопровождается активной разработкой биометрических стандартов. Более того, в последнее время бизнес рассматривает стандартизацию «как инструмент поддержки инновационной деятельности» [16]. Это позволит создать грамотную технологическую инфраструктуру, повысить конкурентоспособность производства, оптимизировать производственные затраты.

В [17] пределяются следующие цели стандартизации:

1. Техническая совместимость и корректное взаимодействие между устройствами и программным обеспечением, разработанным российскими и иностранными производителями.

2. Возможность обмена биометрической информацией между организациями разных стран.

3. Взаимозаменяемость технических частей биометрических систем, выполняющих одинаковую функцию.

4. Унификация условий регистрации биометрических характеристик пользователей.

5. Идентификация производителей биометрических продуктов по единому перечню характеристик в международном реестре.

6. Возможность модернизации и усложнения биометрических систем.

7. Сопоставимость результатов испытаний.

8. Поддержание фонда национальных и межгосударственных стандартов в области биометрической идентификации на современном научно-техническом уровне.

Каждая страна разрабатывает национальные стандарты, обязательные для своих производителей биометрических систем. Однако проводимые международные конференции, посвященные внедрению биометрических технологий, выявили проблему, связанную с тем, что совмещение технологий, разработанных в разных странах, подчас является чрезвычайно сложной технической задачей.

Данная проблема может быть решена за счет разработки международных стандартов. Наиболее активными участниками этого процесса являются ANSI/NIST – American National Standards Institute / National Institute of Standards and Technology Standards, ISO/IEC – International Organization for Standardization / International Electro-Technical Commission Standards, INCITS – Inter-National Committee for Information Technology Standards, British Standards Institute, BSI. Разработка мировых стандартов осуществляется в рамках Международной организации по стандартизации (International Standards Organization – ISO).

В нашей стране также идет активная работа над проектами национальных и межгосударственных стандартов по биометрии в целом и лицевой в частности [17] (НИИЦ БТ МГТУ им. Н. Э. Баумана). Наиболее обеспеченными стандартами в российской биометрии являются процессы обмена биометрическими данными, биометрический программный интерфейс, процесс эксплуатационных испытаний.

Оценка качества биометрических систем обработки изображений. Несмотря на широкое распространение биометрических систем, практически отсутствуют работы, посвященные оценке качества биометрической системы обработки изображений лиц как *программной системы*. Этот вопрос требует тщательного изучения, поскольку системы подобного рода имеют ряд особенностей, которые необходимо учитывать. Необходима дальнейшая разработка метрик для оценки характеристик и их взаимосвязи. При этом необходимо учитывать качество: данных (изображений на входе системы), программного обеспечения (корректное использование алгоритмов, используемых для экстракции, селекции/редукции пространства признаков, классификации изображений, а также выбор оптимальной структуры системы распознавания, обеспечивающей максимальную точность распознавания и минимальное время обработки), информации, продуцируемой системой (точность решения задачи обработки изображения), инфраструктуры (качество программного обеспечения, использованного для разработки системы распознавания, а также используемого аппаратного обеспечения), сервиса (насколько

просто и быстро возможно обучение работе с системой), административного управления (качество менеджмента, а также наличие разработчиков, понимающих особенности биометрической обработки изображений лиц).

Использование моделирования для оценки биометрических систем. Современные промышленные системы идентификации и распознавания лиц имеют множество настраиваемых параметров и гибкую структуру модулей, что обеспечивает их адаптацию к реальным условиям эксплуатации, однако, чтобы оценить выбранные архитектурные решения и выполнить первоначальную настройку системы, на ранних этапах разработки применяют структурные прототипы, которые представляют собой «вертикальный» срез системы, т. е. затрагивают все уровни реализации, а не только интерфейс пользователя.

Разработка биометрической системы – это сложный и трудоемкий процесс. При этом необходимо учитывать условия эксплуатации (освещенность, используемую аппаратуру, возрастные и мимические изменения зарегистрированных в базе лиц, наличие модных аксессуаров и т. п.), а также требования конечных пользователей (производительность, итоговая стоимость, точность распознавания, объем обрабатываемых данных).

Использование дорогостоящей сложной биометрической системы является неоправданным в ситуациях, когда возможно применение простых решений. Однако определить, какие именно методы и решения позволят эффективно решить конкретную задачу, на первом этапе достаточно сложно. С экономической точки зрения для поиска таких решений хорошо подходит компьютерное моделирование, которое активно используется в научных целях, а также может быть применено для промышленной разработки программных средств. Моделирование, благодаря своей способности глубокого исследования биометрической системы, может помочь в выборе наилучшего прототипа и выборе значений параметров. Основным преимуществом средств моделирования является возможность специалистов работать на максимально близком к предметной области уровне абстракции, т. е. описывать и решать задачи в терминах лицевой биометрии. Это позволяет экономить время на низкоуровневой разработке, быстро программировать новые методы, а также исследовать системы в целом для подбора наилучшей комбинации модулей и настройки параметров системы.

Вопросам разработки программных средств для моделирования биометрических систем обработки изображений лиц посвящен ряд работ [18]–[20], в которых подробно рассмотрены требования, предъявляемые к подобным системам, концепция их построения, архитектура, состав и функции компонентов, а также методика использования при решении практических задач обработки изображений. Использование программных средств для моделирования биометрических систем позволяет наглядно описывать их структуру, а также выполнить оценку разных вариантов: сравнить структуры разработанных систем, определить вычислительную сложность их реализации, а также реализуемость на имеющихся аппаратных средствах.

Большое количество биометрических систем, разработанное к настоящему времени, определяет необходимость потребителей в разработке рекомендаций, программных средств, стандартов и т. д., которые позволят оценить существующие системы с точки зрения их соответствия требованиям сценария конкретной задачи. Рассмотренные подходы: определение основных требований к системам и параметров их сравнения, проект FRVT, биометрические стандарты позволят выполнить только самые общие оценки. Основными проблемами являются тестирование систем на нерепрезентативных базах, сложность определения критериев оценки для систем поиска изображений по содержанию, тестирование систем, имеющих сложную структуру (содержащих и параллельные, и каскадные подсистемы), а также недостаточное развитие методик оценки биометрических систем в соответствии с типом решаемой задачи.

Более эффективно можно решать задачу оценки, если используются программные средства для моделирования биометрических систем. Они позволяют не только определить структуры возможных вариантов систем, состав блоков и их функции, но и более просто и экономически эффективно сравнивать варианты систем, оценивать сложность их реализации, что делает возможным разработку биометрических систем с учетом всех требований и современных технологий.

СПИСОК ЛИТЕРАТУРЫ

1. Collection of Facial Images (faces94, faces95, faces96, grimace). URL: http://cswww.essex.ac.uk/mv/allfaces/index.html.

2. Meeker M., Mary Meeker's 2016 internet trends report: All the slides, plus analysis. 2016. URL: http://www.recode.net/2016/6/1/11826256/mary-meeker-2016-internet-trends-report.

3. N-tech Lab. Your face is big data. 2016. URL: http://ecrforum.com/files/presentations/ECR_Forum_2016_15.1_NTechLab_Технология%20распознавания%20лиц%20для%20ритейла.pdf.

4. Guarino B. Russia's new FindFace app identifies strangers in a crowd with 70 percent accuracy // Washington Post. 2016-07-05. URL: https://gadgets.ndtv.com/apps/news/russias-new-findface-app-identifies-strangers-in-a-crowd-with-70-percent-accuracy-839474.

5. Face Recognition Technology (FERET). URL: http://www.nist.gov/itl/iad/ig/feret.cfm.

6. An Introduction to Evaluating Biometric Systems / P. J. Phillips, A. Martin, C. L. Wilson, M. Przybocki // IEEE Computer. 2000. № 33(2). P. 56–63.

7. Обработка и анализ изображений в задачах машинного зрения / Ю. В. Визильтер, С. Ю. Желтов, А. В. Бондаренко и др. М.: Физматкнига, 2010. 672 с.

8. Ross A., Nandakumar Karthik, Jain Anil K. Handbook of Multibiometrics. Springer, 2006. 202 p.

9. Face Recognition Vendor Test 2002 Performance Metrics / P. J. Phillips, P. Grother, R. J. Micheals, D. M. Blackburn, E. Tabassi, J. Bone // 2003. URL: http://ws680.nist.gov/publication/get_pdf.cfm?pub_id=50759.

10. The FERET Evaluation Methodology for Face Recognition Algorithms / P. J. Phillips, H. Moon, P. J. Rauss, S. Rizvi // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2000. № 22(10). P. 1090–1104.

11. Mansfield A. J., Wayman, J. L. Best Practices in Testing and Reporting Performance of Biometric Devices // Technical Report NPL Report CMSC 14/02. London: National Physical Laboratory, 2002.

12. Jain A., Ross A. Learning user-specific parameters in multibiometric system // Proc. Intern. Conf. on Image Proc. (ICIP), Sept. 22–25, 2002. New York: Rochester, 2002. P. 57–70. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.19.5890&rep=rep1&type=pdf.

13. Ross A., Jain A. K. Iformation Fusion in biometrics // Pattern Recognition Letters. Special issue: Audio- and video-based biometric person authentication (AVBPA 2001). 2003. Vol. 24. P. 2115–2125.

14. Ross A., Jain A. K. Multimodal Biometrics // 12th European Signal Proc. Conf. (EUSIPCO). Vienna, Austria, 2004. P. 1221–1224.

15. Дорогов А. Ю., Курбанов Р. Г., Разин В. В Быстродействующий алгоритм семантической классификации JPEG–изображений // Электронный журн. «Нейроинформатика». 2006. Т. 1, № 2. С. 124–144. URL: http://www.ni.iont.ru/Journal/N2/.

16. Горский М. Инновационное развитие и политика в области отраслевых стандартов: биометрия в Великобритании. 2013. URL: http://polit.ru/article/2013/12/22/ps_innov_3/.

17. Николаев Д. Е. Нормативно-техническая база в области биометрических технологий. Состояние и перспективы. URL: http://smart-city.org/img/Nikolaev29.10.2013.pdf.

18. Petrova V. A., Shchegoleva N. L. Rapid prototyping of face retrieval systems in Simulink // Proc. XX Intern. Conf. on Soft Computing and Measurements (SCM). Moscow, 2015. P. 312–314.

19. Щеголева Н. Л. Концепция построения комплекса программных средств для моделирования систем поиска изображений лиц // Изв. СПбГЭТУ «ЛЭТИ». 2016. № 5. С. 40–47.

20. Методы обработки и распознавания изображений лиц в задачах биометрии / под ред. М. В. Хитрова. СПб.: Политехника, 2013. 388 с.

N. L. Shchegoleva, V. A. Petrova
Saint Petersburg Electrotechnical University «LETI»

PROBLEMS OF BIOMETRIC SYSTEM ESTIMATION

The evaluation of biometric image processing systems, basic requirements and comparison parameters are discussed. Standards for the evaluation of FRVT biometric systems, their development, and estimated characteristics are considered. The features of the evaluation of systems using several categories of images, image search systems for content and multibiometric systems are studied. The role of biometric standards and their role in combining technologies developed in different countries is shown. The main evaluation problems are determined: testing of system on non-representative bases, which lead to inconsistency of the declared results when operating in real conditions; the complexity of determining the evaluation criteria for CBIR systems, as well as the inadequate development of biometric systems estimation techniques. The use of software tools for modeling biometric systems makes it possible to determine the structures of possible system options, the composition of the units and their functions, and to compare the variants of systems more simply and cost-effectively.

Biometric systems of face recognition, biometric systems estimation, biometric standards