*УДК 681.3*

# ПРОТОКОЛЫ СИСТЕМ БИОМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ

И.И. ФРОЛОВ

*Белорусский государственный университет информатики и радиоэлектроники*
*П. Бровки, 6, Минск, 220013, Беларусь*

*Поступила в редакцию 8 октября 2015*

Приведены результаты анализа протоколов оценки и эксплуатационных испытаний биометрических систем, а также основных метрик оценки, используемых при исследовании характеристик в биометрии. Исследованы главные факторы влияния на эксплуатационные показатели биометрических систем.

*Ключевые слова:* распознавание лиц, статистические оценки, протоколы испытаний в биометрии.

## Введение

На сегодняшний день многих заказчиков биометрических систем контроля доступа интересуют определенные параметры и характеристики системы, которая может быть построена на базе уже известных методов и алгоритмов машинного обучения и обработки изображений.

Для оценки алгоритмов машинного обучения (алгоритмов классификации) созданы и положительно зарекомендовали себя такие онлайн-платформы как «Полигон», TunedIt, MLcomp [1]. Для выполнения тестирования системы распознавания лиц в целом признанным стандартом являются протоколы «Face Recognition Vendor Test» [2]. Кроме того, создан целый ряд как зарубежных стандартов обработки биометрической информации и тестирования систем распознавания [3, 4], так и русскоязычных аналогов (например, в России [5]).

Тем не менее, перечисленные подходы могут оказаться недостаточными для целей построения практико-ориентированной системы, направленной на решение вполне определенной задачи распознавания лица, т.к. не предоставляют инструментария для построения системы по функциональному или модульному принципу. Близким решением, с данной точки зрения, является инструментарий National Instruments Lab View [6], предназначенный для построения системы технического зрения с использованием готовых функциональных блоков.

Однако все вышеперечисленные инструменты все-таки оставляют свободную нишу для создания онлайн-платформы, позволяющей выполнять задачу построения системы (или прототипа системы) технического зрения по модульному принципу на уровне функционального программирования, а также проводить тестирование заданных параметров выбранных модулей (алгоритмов) как по отдельности, так и системы в целом.

## Основные типы испытаний

Согласно [7] при тестировании биометрических систем распознавания лиц можно выделить следующие три основных типа оценки: технологическая, постановочная (сценарная) и операционная оценки.

Основная цель технологической оценки – выполнить сравнение оцениваемых алгоритмов или системы в целом в одинаковых условиях: база изображений для тестирования всех алгоритмов должна быть получена заранее посредством одного и того же считывающего устройства. Хотя даже в данном случае результаты тестирования алгоритмов могут отличаться в

зависимости от используемой базы. Тестирование выполняется в оффлайн режиме на одинаковых вычислительных мощностях. Учитывая, что база является фиксированной, результаты тестирования являются воспроизводимыми.

Назначение постановочного или сценарного тестирования сводится к определению производительности всей системы в рамках прототипа или приложения, симулирующего систему. Выполняется тестирование всей системы в условиях, максимально приближенных к реальным эксплуатационным условиям. В ходе тестирования каждая система обладает собственным устройством захвата изображения, что в результате вносит определенные отличия для одних и тех же распознаваемых образов (лиц). Ограничение накладывается на выборку людей – это должны быть одни и те же личности, проходящие через тестируемые системы с минимально возможным интервалом временем (для сохранения идентичности входного изображения, подаваемого на вход каждой системы). В зависимости от выделенных вычислительных ресурсов и устройств хранения каждой тестируемой системы оценка может выполняться как в оффлайн, так и в онлайн режиме. Результаты тестирования можно воспроизвести только в случае обеспечения идентичных сценариев теста.

Оценка в условиях эксплуатации предназначена для получения характеристик цельной биометрической системы в реальных условиях на целевой аудитории. В зависимости от вычислительных ресурсов, а также от загруженности системы, распознавание в оффлайн-режиме не всегда может быть доступно в данном случае. Результаты данного вида тестирования в общем случае являются невоспроизводимыми в связи с, как правило, не документируемыми и неизвестными изменениями между сессиями проведения тестов. В достоверности результатов тестирования в таком случае удостовериться достаточно проблематично.

С точки зрения исследовательской оценки разработанных алгоритмов наиболее приемлемой представляется технологическая оценка в связи с возможностью унификации проводимых испытаний, а также наименьшим влиянием внешних факторов и воздействий.

**Основные факторы влияния на эксплуатационные показатели биометрических систем**

В стандартах и наиболее известных протоколах испытаний [3–5, 8] описываются 2 основных режима работы биометрических систем: идентификация и верификация [5]. С учетом этих режимов работы биометрических систем выработан перечень оценок производительности системы. Необходимо отметить, что более полная оценка производительности системы включает в себя как метрики качественной оценки системы/алгоритма, так и оценки загруженности вычислительных ресурсов, удобства использования, архитектурное решение. Коммерческие разработки также ориентированы и на использование специальных аппаратных решений, позволяющих улучшить определенные характеристики отдельных вычислительных алгоритмов или системы в целом.

Основные вычислительные метрики, используемые в биометрических системах, как и обрабатываемые данные, также стандартизированы [3–5]. Однако оценка вероятностей ошибок и пропускной способности биометрических систем при попытке умышленного обмана (т.е. активной попытке «самозванца») находится вне области российского стандарта [5], однако в ряде случаев может служить полезной оценочной информацией при эксплуатации системы.

Главным аспектом оценки биометрической системы является ее точность в выполнении процедур идентификации/верификации. С точки зрения пользователя происходит ошибка распознавания, когда, например, система не подтверждает подлинность личности зарегистрированного лица или когда система ошибочно подтверждает подлинность личности злоумышленника. Есть несколько факторов, которые могут повлиять на возникновение ошибок и их уровень при тестировании/эксплуатации.

*Человеческий фактор.* Первым этапом каждой итерации использования биометрической системы является захват биометрических данных, в частности, если речь идет об обработке данных в ходе реальной эксплуатации. Для захвата изображения лица качественный захват изображения лица предполагает наложения некоторых требований расположения лица, в частности, стандарт [9] регламентирует требования, предъявляемые к данным, обрабатываемым в биометрических системах. Как следствие, человеческое поведение в момент захвата биометрических данных является критической проблемой для последующих шагов и точности

распознавания. Несоответствующий захват вызывает дефект в извлеченных биометрических чертах, по сути, в исходных данных для обработки, или даже приводит к неспособности такого извлечения.

*Аппаратный фактор.* Биометрические датчики также могут быть источниками ошибок, даже когда люди выполняют требования и предоставляют в нужном виде биометрические данные, нивелируя отрицательное воздействие человеческого фактора и проходя соответствующие процедуры захвата. Условия окружающей среды (температура, влажность, скорость ветра, подвижность платформы установки и т.д.) и технологические свойства устройств (например, разрешающая способность) также влияют на качество захваченных биометрических записей.

*Алгоритмический фактор.* У типичного биометрического алгоритма есть три компонента.

1. Компонент «Регистрация»: данный компонент анализирует биометрическую запись, извлекает ее индивидуальные характеристики и упаковывает их в компактной структуре, называемой биометрическим шаблоном.

2. Компонент «Вычисления»: данный компонент устанавливает степень подобия путем вычисления определенных алгоритмом метрик при сравнении между двумя шаблонами. При выполнении идентификации выполняется сравнение «один-ко-многим», при верификации «один-против-одного».

3. Компонент «Принятие решений»: этот компонент использует рассчитанную метрику, а также заданные доверительные интервалы для принятия решения о результате сравнения распознавания.

Вышеперечисленные компоненты удобно отображать на схеме обобщенной биометрической системы, описанной в стандарте [5], представленной на рис. 1.

ГОСТ Р ИСО/МЭК 19795-1—2007

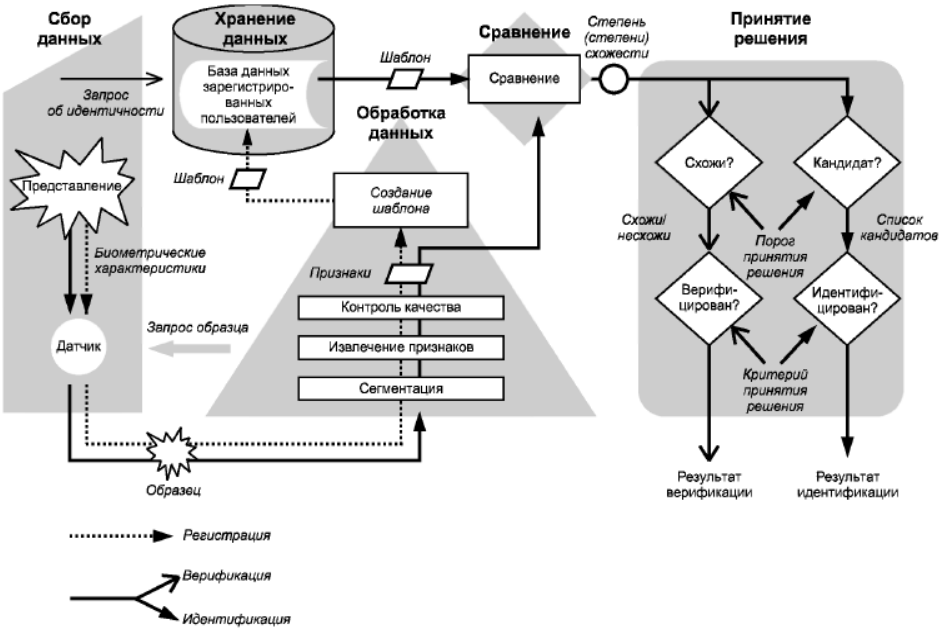



Рис.1. Компоненты обобщенной биометрической системы

В случае оценки системы, работающей в режиме идентификации, решение принимается не просто по значению меры расположения к распознаваемому шаблону, например, чем меньше значение метрики сравнения, т.е. распознаваемый образ считается ближе всего к определенному классу, а рассматривается некоторый диапазон, перечень или возвращаемый список ближайших кандидатов. Число образцов, входящих в данный список, как правило, определяется экспериментальным путем и зависит от величины базы распознаваемых лиц, качества обрабатываемых изображений, требований безопасности и иных факторов.

Таким образом, одним из главных источников ошибки точности распознавания в биометрической системе является неспособность алгоритма принять верное решение о

соответствии шаблонов при решении задачи верификации или при ранжировании списка кандидатов при идентификации.

Рассматриваются несколько типов биометрических системных ошибок как результаты воздействия человеческого, аппаратного и алгоритмических факторов:

– Failure to capture error (FCE): неспособность устройства ввода захватить биометрическую информацию. Данная ошибка может быть результатом двух факторов – устройство и человек;

– Failure to enroll error (FEE): неспособность системы извлечь информацию шаблона после успешного захвата устройства. Данная ошибка связана с алгоритмом, выполняющим обработку и извлечение шаблонов, поступающих от устройства ввода. Уровень данного типа ошибок напрямую зависит от устройства и человеческих факторов в момент биометрического захвата.

**Основные метрики оценки биометрических систем в режиме верификации**

Оценка работы системы в **режиме верификации** выполняется по расчету ошибок I и II рода. К ошибкам I рода (False Rejection rate, далее – FRR) относятся ошибки типа «отказ в правомерном доступе клиенту», а к ошибкам II рода (False Acceptance rate, далее – FAR) относятся, соответственно, ошибки типа «предоставление доступа мошеннику». Классификация ошибок приведена в соответствии с терминологией протокола [8]. Расчет ошибок FAR и FRR выполняется по формулам (1) и (2):

$$FAR = \frac{EM}{M} * 100\,\%, \quad (1)$$

$$FRR = \frac{EK}{K} * 100\,\%, \quad (2)$$

где $EM$ – число предоставлений доступа мошенникам, $M$ – общее число попыток получения доступа мошенниками, $EK$ – число отказов в доступе клиентам, $K$ – общее число попыток получения доступа клиентами.

Предполагается, что мошенник пытается получить доступ, используя поочередно все идентификаторы клиентов, т.е. выдавая себя за нового клиента каждую новую попытку. Таким образом, общее число попыток получения доступа мошенниками в рамках тестирования вычисляется по формуле (3):

$$M = M_S * N_M * K_S, \quad (3)$$

где $M_S$ – общее количество мошенников, $N_M$ – общее число попыток каждого мошенника выдать себя за одного из клиентов (по одной попытке), $K_S$ – общее количество клиентов в базе.

False non-match rate (FNMR) – вероятность ложного несовпадения: доля образцов, полученных в результате транзакций верификации зарегистрированного в системе подлинного пользователя, которые ошибочно признаны не совпадающими с шаблоном тех же биометрических данных пользователя, который проходит верификацию и представил биометрический образец.

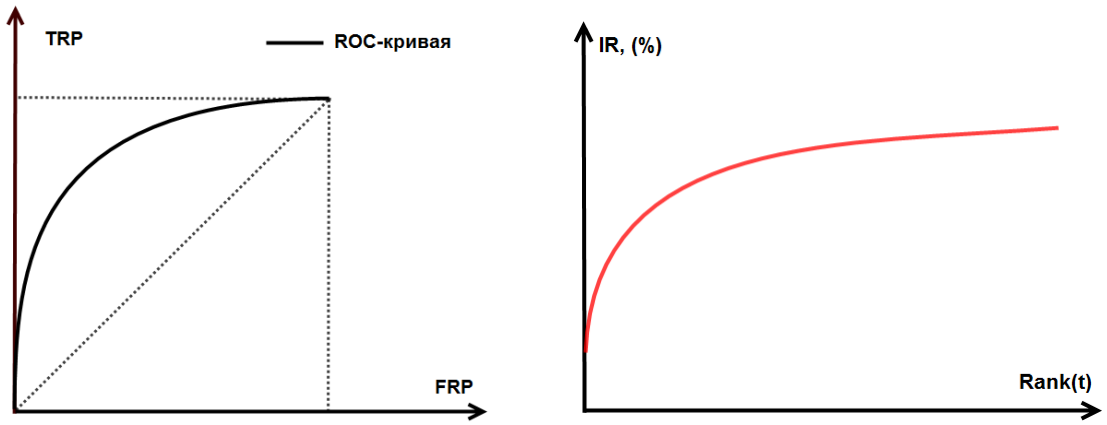
False match rate (FMR) – вероятность ложного совпадения: доля образцов, полученных в результате пассивных попыток «мошенника», которые ошибочно признаны совпадающими с шаблоном другого пользователя [5].

В общем случае метрики FMR и FNMR не являются синонимами FAR и FRR, однако зачастую используются в эквивалентном смысле.

Наиболее распространенной графической характеристикой бинарной классификации является кривая рабочей характеристики (receiver operating characteristic) или ROC-кривая (рис. 2, *а*). ROC-кривая показывает зависимость количества верно классифицированных положительных примеров от количества неверно классифицированных отрицательных примеров. В терминологии ROC-анализа первые называются истинно положительным, вторые – ложно отрицательным множеством. При этом предполагается, что у классификатора имеется некоторый параметр, варьируя который будем получать то или иное разбиение на два класса.

Этот параметр часто называют порогом, или точкой отсечения (cut-off value). В зависимости от него будут получаться различные величины ошибок I и II рода [10].

Важной графической характеристикой при идентификации является кривая характеристики совокупной схожести (Cumulative Match Characteristic) (рис. 2, *б*) или CMC-кривая [10], показывающая зависимость вероятности идентификации от ранга идентификации на замкнутом множестве.




*а* *б*

Рис. 2. Кривые оценок качества модели: *а* – ROC-кривая; *б* – CMC-кривая

Визуальное сравнение кривых ROC не всегда позволяет выявить наиболее эффективную модель. Своеобразным методом сравнения ROC-кривых является оценка площади под кривыми. Теоретически она изменяется от 0 до 1,0, но поскольку модель всегда характеризуются кривой, расположенной выше положительной диагонали, то обычно говорят об изменениях от 0,5 («бесполезный» классификатор) до 1,0 («идеальная» модель). Эта оценка может быть получена непосредственно вычислением площади под многогранником, ограниченным справа и снизу осями координат и слева вверху – экспериментально полученными точками (рис. 5). Численный показатель площади под кривой называется AUC (Area Under Curve) [10]. Вычислить его можно, например, с помощью численного метода трапеций по формуле

$$AUC = \int f(x)\,dx = \sum_i \left[\frac{X_{i+1} + X_i}{2}\right] * \left(Y_{i+1} + Y_i\right). \qquad (4)$$

С большими допущениями можно считать, что чем больше показатель *AUC*, тем лучшей прогностической силой обладает модель. Однако следует знать, что:

– показатель AUC предназначен скорее для сравнительного анализа нескольких моделей;

– AUC не содержит никакой информации о чувствительности и специфичности модели.

В литературе иногда приводится следующая экспертная шкала для значений AUC, по которой можно судить о качестве модели по значениям, приведенным в таблице.

**Значения показателя AUC для характеристики качества модели**

| Интервал AUC | Качество модели |
|---|---|
| 0,9–1,0 | Отличное |
| 0,8–0,9 | Очень хорошее |
| 0,7–0,8 | Хорошее |
| 0,6–0,7 | Среднее |
| 0,5–0,6 | Неудовлетворительное |

### Основные метрики оценки биометрических систем в режиме идентификации

Кроме того, стандартами предусмотрены формальные определения для следующих метрик режима идентификации.

True positive identification rate (TPIR) – вероятность истинно положительной идентификации/вероятность идентификации: ожидаемая доля идентификационных транзакции пользователей, зарегистрированных в системе, в результате которых корректный идентификатор

пользователя будет присутствовать среди возвращаемых *t* идентификаторов. Если выходными данными биометрической системы являются *t* ближайших кандидатов-совпадений, соответствующая оценка TPIR также известна как *t*-ранг идентификации.

False positive identification rate (FPIR) – вероятность ложноположительной идентификации: ожидаемая доля идентификационных транзакции пользователей, НЕ зарегистрированных в системе, в результате которых возвращается идентификатор. Это означает, что мошенник, даже не зарегистрированный в системе, отправив свои биометрические данные на обработку, получает от системы положительный ответ-допуск к объекту. Тогда как в данной ситуации ожидаемой реакцией системы не предполагается возвращение никакого ближайшего кандидата. Это своего рода аналог ошибки второго рода для процедуры верификации. Ошибка FPIR зависит как от числа зарегистрированных пользователей (*N*), так и от величины порога (*n*), по которому определяется допустимая величина меры близости между кандидатом и шаблонами базы. Кроме того, на замкнутом множестве невозможно определить FPIR для процедуры идентификации, т.к. все пользователи системы являются зарегистрированными в системе.

False negative identification rate (FNIR) – вероятность ложноотрицательной идентификации: ожидаемая доля идентификационных транзакции пользователей, зарегистрированных в системе, в результате которых корректный идентификатор пользователя НЕ будет присутствовать среди возвращаемых *t* идентификаторов. FNIR зависит от числа зарегистрированных пользователей (*N*), от величины порога (n), по которому определяется допустимая величина меры близости между кандидатом и шаблонами базы, а также от количества возвращаемых кандидатов – от ранга идентификации.

Вышеперечисленные характеристики связаны следующими соотношениями (5)–(7):

$$FNIR = 1 - TPIR, \quad (5)$$

$$FNIR = FMNR. \quad (6)$$

Вероятность, что входные данные ложно признаны как несоответствующие шаблону пользователя, метрика FNIR в данном случае аналогична метрике FNMR в режиме верификации.

$$FNIR = 1 - \left(1 - FMR\right)^{N}. \quad (7)$$

FPIR определяется, когда входной образец ошибочно совпадает с одним или несколькими шаблонами из базы. Тогда данная ошибка вычисляется как единица минус вероятность того, что не произошло совпадений ни с одним из шаблонов базы (*N* – количество зарегистрированных шаблонов в базе). В случае, если FMR является малой величиной порядка $\left( << \left( \frac{1}{N} \right) \right)$, ошибка FPIR может быть аппроксимирована и представлена в виде формулы $FPIR \approx N^{*} FRM$.

## Результаты и их обсуждение

Представлены наиболее актуальные и общепринятые метрики и подходы к оценке биометрических систем в области распознавания лиц. Также представлен ряд факторов, эксплуатационных условий, оказывающих на выходные характеристики непосредственное влияние. В работе приведены ссылки на актуальные стандарты, протоколы испытаний и онлайн-системы оценки алгоритмов машинного зрения. Однако ниша функционального онлайн-программирования биометрической системы с возможностью анализа и вычисления статистических метрик по оценке системы остается свободной, что является областью интересов и дальнейшего направления работы.

## Заключение

Развитие биометрических технологий в области распознавания лиц, связанное как с разработкой новых, более эффективных алгоритмов технического зрения и машинного обучения, так и с появлением доступных высокопроизводительных вычислительных ресурсов, привело к росту разрабатываемых систем, требующих корректной оценки их работы. Одним из наиболее важных аспектов разработки данных биометрических систем является их соответствие требованиям условий эксплуатации и удовлетворение заданным требованиям заказчиков. Объективные оценки могут быть получены при выполнении требований соответствующих стандартов и протоколов проведения испытаний.

## PROTOCOLS OF THE BIOMETRIC IDENTIFICATION SYSTEMS

I.I. FROLOV

### Abstract

The survey of protocols of estimation and performance tests of the biometrical systems, and general metrics of the investigation used in case of research of characteristics in biometry are provided. The primary factors of impact on performance criteria of biometrical systems are investigated.

*Keywords*: face recognition, statistical estimation, test protocols in biometrical systems.

### Список литературы

1. *Лисица А.В., Воронцов К.В., Ивахненко А.А. и др.* // Матер. Междунар. конф. «Интеллектуализация обработки информации ИОИ-8». Кипр, 2010. С. 157–160.
2. Face Recognition Vendor Test (FRVT) 2013 [Electronic resource] – Mode of access: http://www.nist.gov/itl/iad/ig/frvt-2013.cfm.– Date of access: 09.12.2015.
3. ISO/IEC 19794-5:2005. Information technology – Biometric data interchange formats – Part 5: Face image data.
4. ANSI/INCITS 385-2004. Information technology – Face Recognition Format for Data Interchange.
5. ГОСТ Р ИСО/МЭК 19795-1-2007. Автоматическая идентификация. Идентификация биометрическая. Эксплуатационные испытания и протоколы испытаний в биометрии. Часть 1. Принципы и структура.
6. NI Vision 2013 Concepts Help [Electronic resource] – Mode of access: http://zone.ni.com/reference/en-XX/help/372916P-01/.– Date of access: 09.12.2015.
7. *Mansfield A.J., Wayman J.L.* // NPL Report CMSC 14/02. 2002. P. 1–36.
8. *Messer K.* // Proceedings of Second International Conference on Audio and Video-based Biometric Person Authentication. 1999. P. 965–966.
9. ГОСТ Р ИСО/МЭК 19794-5-2006. Автоматическая идентификация. Идентификация биометрическая. Форматы обмена биометрическими данными. Часть 5. Данные изображения лица.
10. *Паклин Н.Б.* Логистическая регрессия и ROC-анализ – математический аппарат [Электронный ресурс]. – Режим доступа: https://basegroup.ru/community/articles/logistic. Дата доступа: 09.12.2015.