ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГОСТ Р ИСО/МЭК 29794-1— 2012

Информационные технологии

БИОМЕТРИЯ

Качество биометрических образцов

Часть 1

Структура

ISO/IEC 29794-1:2009
Information technology — Biometric sample quality — Part 1: Framework
(IDT)

Издание официальное

Москва
Стандартинформ
2013

gost.e-ecolog.ru

Предисловие

Цели и принципы стандартизации в Российской Федерации установлены Федеральным законом от 27 декабря 2002 г. № 184-ФЗ «О техническом регулировании», а правила применения национальных стандартов Российской Федерации — ГОСТ Р 1.0—2004 «Стандартизация в Российской Федерации. Основные положения»

Сведения о стандарте

1 ПОДГОТОВЛЕН Научно-исследовательским и испытательным центром биометрической техники Московского государственного технического университета имени Н.Э. Баумана (НИИЦ БТ МГТУ им. Н.Э. Баумана) на основе собственного аутентичного перевода на русский язык международного стандарта, указанного в пункте 4, при консультативной поддержке Ассоциации автоматической идентификации «ЮНИСКАН/ГС1 РУС»

2 ВНЕСЕН Техническим комитетом по стандартизации ТК 355 «Технологии автоматической идентификации и сбора данных и биометрия»

3 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от 18 сентября 2012 г. № 351-ст

4 Настоящий стандарт идентичен международному стандарту ИСО/МЭК 29794-1:2009 «Информационные технологии. Качество биометрического образца. Часть 1. Структура» (ISO/IEC 29794-1:2009 «Information technology — Biometric sample quality — Part 1: Framework»).

Наименование настоящего стандарта изменено относительно наименования указанного международного стандарта для приведения в соответствие с ГОСТ Р 1.5—2004 (подраздел 3.5).

При применении настоящего стандарта рекомендуется использовать вместо ссылочных международных стандартов соответствующие им национальные стандарты Российской Федерации, сведения о которых приведены в дополнительном приложении ДА

5 ВВЕДЕН ВПЕРВЫЕ

6 Некоторые элементы настоящего стандарта могут быть объектами патентных прав. Организации ИСО и МЭК не несут ответственности за установление подлинности каких-либо или всех таких патентных прав

Информация об изменениях к настоящему стандарту публикуется в ежегодно издаваемом информационном указателе «Национальные стандарты», а текст изменений и поправок — в ежемесячно издаваемом информационном указателе «Национальные стандарты». В случае пересмотра (замены) или отмены настоящего стандарта соответствующее уведомление будет опубликовано в ежемесячно издаваемом информационном указателе «Национальные стандарты». Соответствующая информация, уведомление и тексты размещаются также в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии

© Стандартинформ, 2013

Настоящий стандарт не может быть полностью или частично воспроизведен, тиражирован и распространен в качестве официального издания без разрешения Федерального агентства по техническому регулированию и метрологии

Содержание

1 Область применения . . . 1
2 Соответствие . . . 1
3 Нормативные ссылки . . . 1
4 Термины и определения . . . 2
5 Обозначения и сокращения . . . 4
6 Общие сведения о биометрической системе . . . 4
7 Критерии качества биометрического образца . . . 4
7.1 Базовая модель . . . 4
7.2 Элементы качества: качество источника, точность, полезность . . . 5
7.3 Практическая значимость данных о качестве . . . 6
8 Определение поля формата обмена данными . . . 8
8.1 Поля данных . . . 8
8.2 Показатель качества . . . 8
8.3 Идентификация алгоритма оценки качества . . . 9
8.4 Стандартизированный обмен результатами работы алгоритмов оценки качества . . . 10
9 Нормализация . . . 10
Приложение А (справочное) Процедуры построения набора данных для нормализации показателя качества . . . 12
Приложение В (справочное) Пример стандартизированного обмена результатами работы алгоритмов оценки качества . . . 16
Приложение С (справочное) Процедуры для объединения показателей качества, основанных на полезности . . . 18
Приложение ДА (справочное) Сведения о соответствии ссылочных международных стандартов ссылочным национальным стандартам Российской Федерации . . . 20
Библиография . . . 20

Введение

Комплекс стандартов ИСО/МЭК 29794 (далее ИСО/МЭК 29794) имеет заголовок «Информационные технологии. Биометрия. Качество биометрических образцов» и включает в себя следующие части:

- часть 1 — Структура;
- часть 4 — Данные изображения отпечатка пальца (технический отчет);
- часть 5 — Данные изображения лица (технический отчет).

Последующие части ИСО/МЭК 29794 будут посвящены другим биометрическим модальностям, определенным в комплексе стандартов ИСО/МЭК 19794 (далее ИСО/МЭК 19794), причем номера и названия частей ИСО/МЭК 29794 будут соответствовать номерам и названиям частей ИСО/МЭК 19794. Поскольку настоящий стандарт распространяется на все биометрические модальности, наличие специальной части ИСО/МЭК 29794, описывающей использование показателей качества для конкретной биометрической модальности, не требуется.

В последующих версиях каждой из частей ИСО/МЭК 19794 будут содержаться нормативные ссылки на настоящий стандарт и соответствующие обновленные поля данных.

Измерение качества рекомендуется для нескольких приложений в области биометрии. В ИСО/МЭК 19794-1 определена структура и представлены рекомендации по распределению показателей качества по группам, а ИСО/МЭК 29794 устанавливает и описывает методы объективной количественной оценки качества, интерпретацию результатов и их обмен. Настоящий стандарт предназначен для повышения эффективности приложений, а именно:

а) для поддержания конкуренции, новаторства, способности к взаимодействию и совершенствования эксплуатационных характеристик;

б) для предотвращения систематических погрешностей отдельных приложений, модальностей или технологий.

В настоящем стандарте предоставлено несколько способов оценки показателей качества биометрических образцов, использование которых обычно необязательно, но может быть определено как обязательное отдельными профилями приложения или индивидуальными программными реализациями.

Использование данных о качестве биометрических образцов рекомендуется для ряда приложений, например оценка качества в режиме реального времени для организации обратной связи в процессе регистрации биометрических образцов с целью повышения эффективности работы и улучшения эксплуатационных характеристик биометрической системы. Комплекс данных о качестве биометрических образцов является важным компонентом стандартизации измерения качества. Поля данных о качестве, в соответствии с подразделом 8.1 настоящего стандарта, будут включены в форматы обмена данными. При наличии заголовка ЕСФОБД для характеристик данных о качестве можно дополнительно использовать элемент данных CBEFF_BDB_quality. Анализ путей улучшения эксплуатационных характеристик биометрической системы может быть выполнен на основе данных о качестве совместно с другими данными. Например, сопоставление данных о качестве с другими показателями системы может использоваться для выявления проблем и определения возможных путей улучшения эксплуатационных характеристик.

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

Информационные технологии

БИОМЕТРИЯ

Качество биометрических образцов

Часть 1

Структура

Information technology. Biometrics. Biometric samples quality. Part 1. Framework

Дата введения — 2013—07—01

1 Область применения

В настоящем стандарте для всех типов биометрических образцов:

- установлены термины и определения, необходимые для определения и измерения качества;
- представлены назначение и методы интерпретации показателей качества биометрических образцов;
- определены формат и расположение полей данных о качестве в форматах обмена биометрическими данными;
- представлены методы разработки наборов данных о биометрических образцах для нормализации показателей качества;
- представлен формат для обмена результатами работы алгоритмов оценки качества.

За рамками области применения настоящего стандарта находятся:

- определение минимальных требований к показателям качества образца, модуля или системы;
- определение эксплуатационных характеристик алгоритмов оценки качества;
- стандартизация алгоритмов оценки качества.

2 Соответствие

Блок данных о качестве соответствует настоящему стандарту в том случае, если он отвечает требованиям, установленным в разделе 8.

3 Нормативные ссылки

В настоящем стандарте использованы нормативные ссылки на следующие стандарты, которые необходимо учитывать при использовании настоящего стандарта. В случае ссылок на документы, у которых указана дата утверждения, необходимо пользоваться только указанной редакцией. В случае, когда дата утверждения не приведена, следует пользоваться последней редакцией ссылочных документов, включая любые поправки и изменения к ним:

ИСО/МЭК 19794-1:2006 Информационные технологии. Форматы обмена биометрическими данными. Часть 1. Структура (ISO/IEC 19794-1:2006, Information technology — Biometric data interchange formats — Part 1: Framework)

ИСО/МЭК 19785-2:2006 Информационные технологии. Единая структура форматов обмена биометрическими данными. Часть 2. Процедуры действий регистрационного органа в области биометрии (ISO/IEC 19785-2:2006, Information technology — Common Biometric Exchange Formats Framework — Part 2. Procedures for the operation of the Biometric Registration Authority)

Издание официальное

4 Термины и определения

В настоящем стандарте применены следующие термины с соответствующими определениями:

4.1 **точность сбора данных** (acquisition fidelity): Точность параметров образца, обусловленная процессом сбора данных.

4.2 **отказ регистрации** (biometric failure to enrol): Отказ биометрической системы во время регистрации сохранить приемлемый биометрический эталон вследствие нехватки биометрических данных, полученных в процессе регистрации.

Примечание 1 — Нехватка биометрических данных может возникнуть вследствие отказа в получении приемлемого биометрического образца, отказа извлечь приемлемый биометрический признак из образца, или отказа сформировать на основании биометрических признаков приемлемый биометрический эталон.

Примечание 2 — Наиболее актуальное определение представлено в SC 37 NSD2.

4.3 **вероятность отказа регистрации** (biometric failure to enrol rate): Доля сессий биометрической регистрации, результатом которых становится отказ регистрации по различным причинам, за исключением небиометрических.

Примечание 1 — Если знаменатель доли основан на количестве сессий биометрической регистрации, то значение вероятности отказа регистрации окажется выше, чем в том случае, когда знаменатель доли основан на количестве субъектов сбора биометрических данных.

Примечание 2 — Знаменатель доли представляет собой число сессий биометрической регистрации, исключая сессии, которые не завершились по небиометрическим причинам.

Примечание 3 — Наиболее актуальное определение представлено в SC 37 NSD2.

4.4 **качество источника** (character): Фактор, влияющий на качество образца, связанный с особенностями источника.

4.5 **окружающая среда** (environment): Окружающая обстановка и условия, в которых производится сбор биометрических данных, включая такие факторы, как квалификация оператора и степень его содействия процессу регистрации.

4.6 **точность извлечения данных** (extraction fidelity): Элемент точности образца, обусловленный процессом извлечения биометрических признаков.

4.7 **внешний показатель качества** (extrinsic): Показатель качества, при котором для полноценной интерпретации и нормализации требуется ссылка на внешний источник, например стандарт, журнал записей или технические условия.

4.8 **точность** (fidelity): Степень соответствия биометрического образца исходной биометрической характеристике.

Примечание — Точность образца включает в себя элементы, присущие одному или нескольким этапам обработки: получение данных, извлечение, обработка сигнала.

4.9 **внутренний показатель качества** (intrinsic): Показатель качества, передающий полностью интерпретированные нормализованные данные, при котором не требуется дополнительная внешняя информация для нормализации показателя качества.

4.10 **интерпретация** (interpretation): Процесс анализа показателя качества совместно с иными данными с целью придания этому показателю контекстно-зависимого смыслового содержания.

4.11 **вероятность отказа сбора данных** (failure to acquire rate): Доля попыток, результатом которых становится отказ получения надлежащего или приемлемого биометрического образца по различным причинам, за исключением небиометрических.

Примечание 1 — Знаменатель доли представляет собой число попыток биометрической регистрации, исключая сессии, которые не завершились по небиометрическим причинам.

Примечание 2 — Наиболее актуальное определение представлено в SC 37 NSD2.

4.12 **вероятность ложного совпадения;** ВЛС (false match rate; FMR): Доля завершенных попыток сопоставления биометрических образцов и контрольных образцов, результатом которых становится ложное совпадение.

Примечание 1 — Значение, вычисленное для вероятности ложного совпадения, зависит от пороговых значений и иных параметров процесса сравнения и протокола, в котором описаны попытки сопоставления биометрических образцов и контрольных образцов. В частности, подход к сравнению:

- однояйцовых близнецов;

- совершенно разных биометрических характеристик разных личностей, как, например, топография лица и гребни Гальтона (отпечаток пальца);

- разных, но связанных биометрических характеристик одной и той же личности, как, например, данные геометрии контура кистей правой и левой рук требуют тщательного рассмотрения (см. ИСО/МЭК 19795-1).

Примечание 2 — Термин «завершенная попытка» относится к процессам вычисления, необходимым для принятия решения о сравнении, то есть отказ принятия решения исключен.

Примечание 3 — Наиболее актуальное определение представлено в SC 37 NSD2.

4.13 **вероятность ложного несовпадения;** ВЛНС (false non-match rate; FNMR): Доля завершенных попыток сопоставления биометрических образцов и контрольных образцов, результатом которых становится ложное несовпадение.

Примечание 1 — Значение, вычисленное для вероятности ложного несовпадения, зависит от пороговых значений и иных параметров процесса сравнения и протокола, в котором описаны попытки сопоставления биометрических образцов и контрольных образцов.

Примечание 2 — Термин «завершенная попытка» относится к процессам вычисления, необходимым для принятия решения о сравнении, то есть отказ принятия решения исключен.

Примечание 3 — Наиболее актуальное определение представлено в SC 37 NSD2.

4.14 **оператор** (operator): Человек, осуществляющий или контролирующий процесс получения или повторного получения биометрического образца пользователя биометрической системы.

4.15 **эксплуатационные характеристики** (performance): Оценка ВЛС и ВЛНС, вероятности отказа регистрации и вероятности отказа сбора данных биометрической системы.

4.16 **качество** (quality): Степень, в которой биометрический образец отвечает определенным требованиям целевого приложения.

Примечание — Определенные требования могут предъявляться к таким аспектам качества, как фокус, разрешение и т. д. Неявно выраженные требования к качеству относятся к вероятности получения правильного результата сопоставления.

4.17 **показатель качества** (quality score): Количественное выражение качества.

4.18 **нормализация показателя качества** (quality score normalization): Преобразование показателей качества для согласования масштаба и облегчения интерпретации.

4.19 **блок данных для нормализации показателя качества;** БДНПК (quality score normalization dataset; QSND): Блок данных о биометрических образцах, снабженный комментариями о показателях качества и используемый для нормализации показателей качества.

Примечание — Целевые показатели качества могут быть установлены на основе результатов работы биометрической системы при помощи рассматриваемого образца или могут быть основаны на факторах качества, записанных при составлении блока данных.

4.20 **процентильный ранг показателя качества;** ПРПК (quality score percentile rank; QSPR): Процентильный ранг показателя качества биометрического образца, полученный из собственного показателя и показателей иных образцов в определенном контрольном блоке данных для нормализации показателя качества (для сравнения – блок данных для нормализации показателя качества, БДНПК).

4.21 **необработанный показатель качества** (raw quality score): Показатель качества, который не был интерпретирован ни создателем показателя, ни получателем показателя и в действительности не способен самостоятельно предоставить контекстную информацию.

4.22 **образец** (sample): Интерпретация физической характеристики человека, основанная на изображении, сигнале или шаблоне и используемая для идентификации или верификации на основе биометрических технологий.

4.23 **источник** (source): Физическая часть или функция тела, представленная биометрическим образцом.

4.24 **полезность** (utility): Наблюдаемые эксплуатационные характеристики биометрического образца или набора образцов в одной или более биометрических системах.

Примечание 1 — Качество источника образца и точность обработанного образца вносят вклад в полезность образца.

Примечание 2 — Полезность может совмещать измерения эксплуатационных характеристик, как, например ВЛС, ВЛНС, вероятность отказа регистрации и вероятность отказа сбора данных.

5 Обозначения и сокращения

В настоящем стандарте применены следующие обозначения и сокращения:

ББД (BDB) — блок биометрических данных (biometric data block);
ЗБИ (BIR) — запись биометрической информации (biometric information record);
ЕСФОБД (CBEFF) — единая структура форматов обмена биометрическими данными (common biometric exchange formats framework) (ИСО/МЭК 19785);
FERET — база данных технологии распознавания лица (facial recognition technology database);
ВЛС (FMR) — вероятность ложного совпадения (false match rate);
ВЛНС (FNMR) — вероятность ложного несовпадения (false non-match rate);
ИАОК (QAID) — идентификатор алгоритма оценки качества (quality algorithm identification);
НДНПК (QSND) — набор данных для нормализации показателя качества (quality score normalization dataset);
ПРПК (QSPR) — процентильный ранг показателя качества (quality score percentile rank);
XML — расширяемый язык разметки (extensible markup language).

6 Общие сведения о биометрической системе

Общие сведения о биометрической системе представлены в Постоянно действующем документе 11, ИСО/МЭК СТК 1/ПК 37 Часть 1. Гармонизация стандартов. Общее руководство (Standing Document 11, ISO/IEC JTC 1/SC 37 Part 1 Overview Standards Harmonization Document (SC 37 N-SD11)).

7 Критерии качества биометрического образца

7.1 Базовая модель

Термин «качество» в биометрии применяется для описания нескольких различных свойств биометрического образца, которые вносят вклад в обобщенные эксплуатационные характеристики биометрической системы. В целях стандартизации в настоящем документе установлены термины, определения и базовая модель для определения различий между аспектами качества, представленными на рисунке 1. На рисунке 2 показана взаимосвязь между качеством источника, точностью, качеством, полезностью и эксплуатационными характеристиками системы.



Качество образца зависит от качества источника и компонентов точности.

Полезность отражает воздействие качества единичного образца на эксплуатационные характеристики системы.

Рисунок 1 — Изображение базовой модели качества

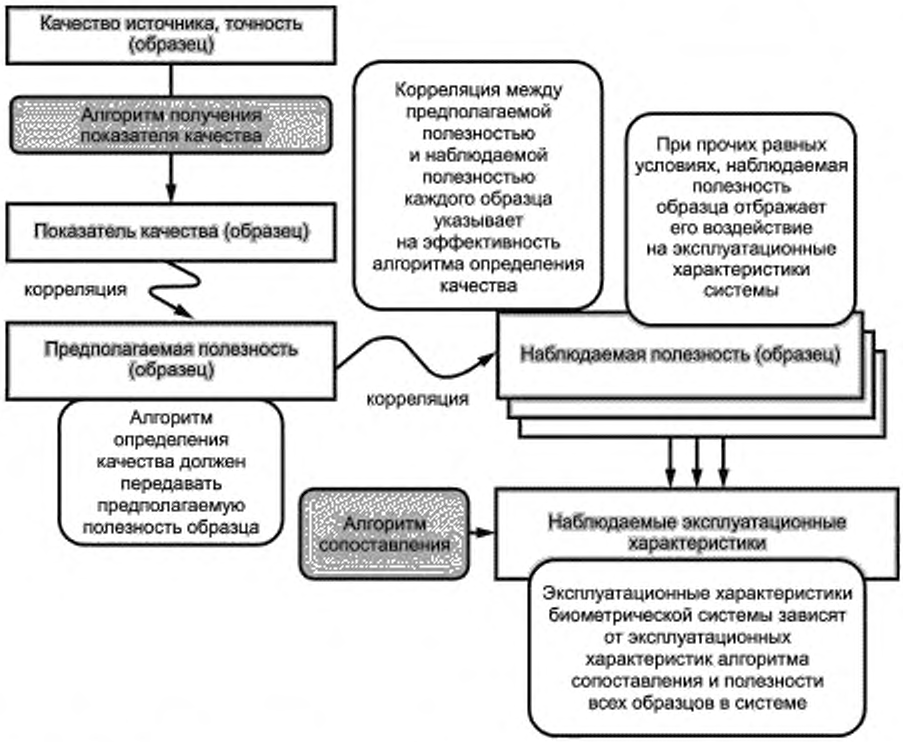


Рисунок 2 — Взаимосвязь между качеством и эксплуатационными характеристиками системы

7.2 Элементы качества: качество источника, точность, полезность

Термин «качество» в том виде, в котором он применяется в области биометрии, имеет несколько дополнительных значений, зависящих от контекста. Ниже представлено субъективное определение трех наиболее распространенных значений:

1) **качество источника** образца. Выражение качества, основанное на признаках источника, от которого получен биометрический образец. Например, низкое качество источника изображения отпечатка пальца в случае наличия пореза на пальце; низкое качество источника изображения радужной оболочки глаза в случае блефароптоза (опущение верхнего века);

2) **точность** воспроизведения образцом источника, от которого он был получен. Выражение качества, основанное на точности, отражает степень схожести образца и его источника. Точность образца состоит из нескольких элементов точности, представленных разными процессами;

3) **полезность** образца в рамках биометрической системы. Выражение качества, основанное на полезности, отражает предполагаемый положительный или отрицательный вклад отдельного образца в эксплуатационные характеристики биометрической системы. Качество, основанное на полезности, зависит как от качества источника образца, так и от его точности. Оценка качества, основанного на полезности, используется для того, чтобы сделать эксплуатационные характеристики системы более предсказуемыми, например, в отношении ВЛС, ВЛНС, вероятности отказа регистрации и вероятности отказа сбора данных, чем при оценке качества, основанного только на качестве источника или точности (см. таблицу 1).

Термин «качество» не должен применяться исключительно к параметрам получения образца, как, например, разрешение, размер в пикселях, глубина цвета и глубина цвета в градации серого или количество признаков. Однако данные факторы могут оказать влияние на полезность образца и обобщенный показатель качества.

Качество источника и полезность полученного образца зависят от тех признаков, которые используются компаратором. Например, у одного и того же изображения пальца может быть низкое качество источника и полезность, если используется распознавание минуций (вследствие слишком малого количества минуций), но высокое качество источника и полезность, если используется распознавание по спектральным данным изображения отпечатка пальца.

Таблица 1 — Иллюстрация взаимосвязи между точностью, полезностью и качеством источника

Качество источника	Точность источника	
	Низкая	Высокая
Низкое	Низкая точность и низкое качество источника при низком уровне полезности. Увеличение полезности достигается путем повторного сбора данных. По возможности рекомендуется использование других биометрических характеристик	Высокая точность и низкое качество источника при низком уровне полезности. Повторный сбор данных не приведет к увеличению полезности. Рекомендуется использование других биометрических характеристик
Высокое	Образцы с высоким качеством источника и низкой точностью, как правило, свидетельствуют о низкой полезности. Увеличение полезности достигается путем повторного сбора данных или использования техники повышения качества изображений	Образцы с высоким качеством источника и высокой точностью свидетельствуют о том, что полученный образец является пригодным, т. е. он обладает высокой полезностью

7.3 Практическая значимость данных о качестве

7.3.1 Оценка качества в режиме реального времени

Данные о качестве, поступающие в режиме реального времени, могут быть использованы оператором, автоматической системой или пользователем для улучшения среднего качества образцов, представленных при регистрации. Подобная обратная связь может выявлять качество образца, точность, полезность и возможность совершенствования образца. Таким образом, эксплуатационная эффективность и эксплуатационные характеристики системы могут быть улучшены при содействии оператору или автоматической системе контроля качества в принятии решения:

a) принять образец,

b) отвергнуть образец,

c) повторить сбор данных или

d) заявить об отказе сбора данных или отказе регистрации.

Данные о качестве могут быть сохранены для дальнейшего использования, например для принятия решения о том, нужно ли заменить зарегистрированный образец, когда будет получен следующий образец.

7.3.2 Применение в различных приложениях

Вследствие того, что возможно использование определенного биометрического образца в различных приложениях, данные о его качестве должны быть применимы ко всем этим приложениям. Предполагается выполнение сравнений как один-к-одному, так и один-ко-многим при помощи алгоритмов сравнения, предоставленных различными поставщиками, которые по-разному интерпретируют признаки образца и выдают различные результаты сравнения. Проблема определения универсального стандарта качества состоит в выборе метрики, которую можно применять во всех приложениях, в которых используется конкретный образец, учитывая то, что метрики полезности в разных приложениях различаются. Таким образом, проблема определения единой метрики, которая бы достоверно передавала данные о полезности биометрического образца всем приложениям, в которых он может использоваться, является технической, что следует учитывать при разработке стандартов качества. Метрика качества, которая определяет эксплуатационные характеристики компаратора или группы компараторов, скорее всего, может быть разработана только для определения режима отказа сбора данных и чувствительности ограниченного количества биометрических систем. Возможно, стоит применять несколько метрик для определения различных режимов отказа.

Для получателей данных о показателях качества крайне важно иметь возможность различать показатели, сгенерированные при использовании разных алгоритмов оценки качества и оборудования для сбора данных. Эти данные о показателях качества могут быть применены для настройки программного обеспечения, производящего получение данных, так, чтобы к показателям, сгенерированным разными алгоритмами, можно было бы применить разные пороговые значения или классификации качества. Более того, имея возможность разграничивать показатели, сгенерированные разными алгоритмами, получатель данных может выделить результаты, сгенерированные разными алгоритмами и, исходя из этого, применить эти данные для оптимизации пороговых значений.

7.3.3 Использование данных о качестве как статистических данных об исследовании

Показатели качества могут быть использованы для отслеживания качества функционирования. Например, объединенные показатели качества могут быть сопоставлены с предварительно установленными граничными показателями или с функциональными требованиями. В том случае, если показатели качества сгенерированы при обработке биометрических образцов, собранных из многих источников, или в течение различных периодов времени, то они могут быть применены для распознавания отклоняющегося от нормы действия. Например, если качество изображения лица определяют в отделе по выдаче водительских прав службы регистрации транспортных средств, то упорядоченный перечень объединенных показателей качества следует применять для выявления отделов, которые демонстрируют уровень качества ниже среднего, или для отслеживания тенденции за неделю или месяц.

7.3.4 Суммирование соответствующих статистик

Надежные показатели качества могут быть использованы для исследования пользователей и транзакций для накопления статистик, предоставляющих условные вероятности вида «если есть образец с качеством X, полученный от пальца A, какова вероятность получения образца с качеством Y от пальца A (или пальца B)». С помощью данной информации система и/или оператор определяет, возможно ли получить образец с более высоким качеством при повторном сборе данных.

7.3.5 Улучшение контрольного набора данных

Соотнесение данных о качестве с образцом, который должен быть внесен в контрольный набор данных, необходимо для поддержания и улучшения качества контрольного набора данных. Отслеживание данных о качестве образца может привести к обнаружению потенциального ухудшения подготовки оператора или выявлению ухудшения в эксплуатационных характеристиках оборудования по сбору данных. Отслеживание данных о качестве образца должно занимать существенную часть процесса функционирования биометрической системы. Данные о качестве также могут быть использованы для улучшения качества контрольного файла и, следовательно, эксплуатационных характеристик биометрической системы в целом. Улучшения могут быть произведены путем замены или дополнения, если возможно, сохраненной информации для того, чтобы иметь в распоряжении наиболее качественные данные из доступных. Как правило, решения о замене в системе связаны с характеристиками компаратора, осуществляющего обработку данных.

7.3.6 Выбор обработки на основе качества

Биометрические образцы могут быть обработаны по-разному в зависимости от значений показателей качества. В частности, данные низкого качества могут быть обработаны при помощи нестандартных алгоритмов или порогов.

7.3.7 Обмен данными о качестве между несопоставимыми системами

Стандартизированный обмен данными о качестве между несопоставимыми системами применяется для сохранения способности взаимозаменяемости локальных или удаленных компонентов программного или аппаратного обеспечения систем и для сохранения целостности данных о качестве в случае подобной замены.

Например, при использовании стандартизированного обмена данными о качестве системы получатели данных о качестве от компонента нуждаются в минимальной модификации, если данный компонент заменяется.

8 Определение поля формата обмена данными

8.1 Поля данных

В таблице 2 представлена общая структура блока данных о качестве биометрических данных.

Таблица 2 — Поля данных

Описание		Размер	Допустимые значения	Примечание
Количество блоков данных о качестве		1 байт	[0,255]	После данного поля следует количество 5-байтовых блоков качества, равное значению данного поля (см. рисунок 3). Значение «ноль» (0) означает, что попыток присвоить показатель качества произведено не было. В этом случае блоки качества отсутствуют
Блок качества	Показатель качества	1 байт	[0,100] 255	0: самое низкое 100: самое высокое 255: ошибка присвоения показателя качества
	ID поставщика алгоритма оценки качества	2 байта	[1,65535]	ID поставщика алгоритма оценки качества должен быть зарегистрирован МАБИ (Международной ассоциацией биометрии и идентификации) в качестве организации, занимающейся биометрией и ЕСФОБД. См. процедуры регистрации ID поставщика ЕСФОБД, представленные в ИСО/МЭК 19785-2
	ID алгоритма оценки качества	2 байта	[1,65535]	ID алгоритма оценки качества может быть опционально зарегистрирован МАБИ в качестве кода продукта ЕФСОБД. См. процедуры регистрации ID продукта ЕСФОБД, представленные в ИСО/МЭК 19785-2

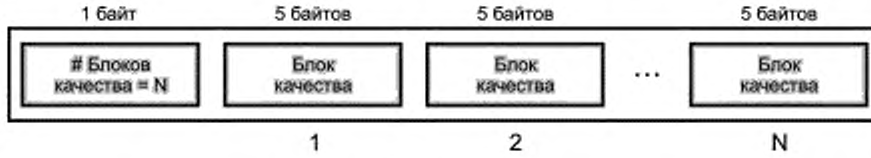


Рисунок 3 — Структура поля качества

Показатели качества должны всегда вноситься в поле показателя качества блока биометрических данных (ББД), связанного с образцом, как установлено в ИСО/МЭК 19794. Поля качества ЕСФОБД не должны использоваться вместо полей качества, определенных в ИСО/МЭК 19794; они должны использоваться в качестве дополнительной информации. Применение полей качества ЕСФОБД может быть представлено в любом стандарте ЕСФОБД и не рассматривается в настоящем стандарте. Необходимо отметить, что несколько показателей качества, вычисленных при помощи одного и того же алгоритма, не должны присутствовать в одном и том же ББД.

8.2 Показатель качества

8.2.1 Назначение

Результатом работы алгоритма должны быть показатели качества, определяющие такие эксплуатационные характеристики, как ложное совпадение или ложное несовпадение. В тех случаях, когда в системе используются элементы, предоставленные несколькими поставщиками, при применении метода вычисления показателя качества должны учитываться все аспекты функционирования каждого используемого алгоритма. Как указано в 7.3.2, сложно разработать единую метрику качества, которая была бы универсальной, не зависела от поставщика и при этом надлежащим образом описывала эксплуатационные характеристики, поэтому может быть полезным применение более одного алгоритма оценки качества.

8.2.2 Факторы, которые необходимо учитывать при преобразовании данных

Преобразование данных, производимое приложением, способно оказать воздействие на качество данных (например, субдискретизация или дальнейшее сжатие). Воздействие подобных преобразований на качество данных может быть вычислено заново прикладной системой в соответствии с руководством, представленным в настоящем стандарте. Каждый раз, когда биометрический образец подвергается преобразованию, качество преобразованного образца должно быть заново оценено и сопоставлено с преобразованным образцом. Так, во всей системе управления идентификацией биометрический образец может храниться в нескольких форматах (например, изображение отпечатка пальца с высоким разрешением может храниться в основном накопителе, а представление отпечатка пальца на основе минуций может храниться на смарт-карте).

8.2.3 Режимы отказа

Для контролирования эксплуатационных характеристик разработчику алгоритмов следует создать показатели качества, с помощью которых можно моделировать известные режимы отказа или изменять чувствительность биометрического компаратора и алгоритмов обработки изображений/сигналов. В дальнейшем для достижения некоторой степени общности показатель качества должен быть основан на наборе чувствительностей, которые были бы общими для систем определенного класса (например, систем сопоставления минуций).

8.2.4 Разрешающая способность

В устройстве по оценке качества должны использоваться по крайней мере четыре значения, которые соответствуют следующим уровнями эксплуатационных характеристик: «отличный», «достаточный», «минимально допустимый» и «неприемлемый».

8.2.5 Суммирование

В приложении В представлены процедуры объединения показателей качества, основанных на полезности множества образцов, например суммирование по всей компании. Результатом становится суммарное значение показателя качества, с помощью которого можно оценивать качество. Суммирование показателей качества должно осуществляться в рамках области использования, например суммирование показателей качества всех зарегистрированных образцов организации или суммирование показателей качества всех верификационных образцов организации. В операциях, когда пользователи часто контактируют с биометрической системой (приложения учета рабочего времени), показатели качества могут объединяться для каждого пользователя в отдельности. Такой подход позволит выявить пользователей, которые предоставляют образцы низкого качества.

8.3 Идентификация алгоритма оценки качества

8.3.1 Общие сведения

Идентификатор алгоритма оценки качества (ИАОК) используется для вычисления показателя качества образца. Поскольку единого критерия оценки качества не существует, крайне важно предоставить получателю ЗБИ возможность различать показатели качества, сгенерированные разными алгоритмами оценки качества, и вносить поправки по необходимости в соответствии с отличиями в обработке и анализе. В реестре ИАОК должна быть ссылка с указанием поставщика и номера версии идентифицированного алгоритма оценки качества. Использование ИАОК является удобным, но не позволяет решить все задачи стандартизации показателей качества.

8.3.2 Методология

Следующие положения обязательны для данного метода:

Если показатель качества не оценивается, то значение поля «Количество блоков данных о качестве» равно 0 и блоки качества отсутствуют. В том случае, если значение числа блоков данных о качестве находится в пределах между 1 и 255, ИАОК и идентификатор поставщика алгоритма оценки качества, используемого для генерации показателя, должны быть внесены в соответствующие поля (таблица 2). Применение данного метода не исключает, а дополняет дальнейшие работы по стандартизации универсального метода оценки показателя качества (то есть создание показателя, который по своей природе заключает в себе некоторую степень нормализации), такого как набор данных для нормализации показателя качества (НДНПК), см. раздел 9.

Особенность метода идентификации алгоритма оценки качества состоит в том, что получателю данных о необработанных показателях следует провести локальный анализ и/или обработку данных для получения полноценной интерпретации показателей. Другими словами, посылающий данные о показателе не интерпретирует показатель качества для потенциально неизвестного приложения или получателя. Поставщик алгоритма оценки качества может предоставить информацию о том, каким обра-

зом устанавливается показатель качества, а получатель — разработать надлежащие средства по автоматическому проведению различий между показателями, сгенерированными разными алгоритмами оценки качества, и интерпретировать их должным образом.

8.4 Стандартизированный обмен результатами работы алгоритмов оценки качества

Поставщики алгоритмов оценки качества должны иметь возможность предоставить биометрическому обществу стандартизированные результаты работы своих алгоритмов оценки качества. С другой стороны, потребители, использующие записи обмена данными, соответствующие ИСО/МЭК 19794, имеют возможность получить и обработать информацию так, чтобы оценить значение выходных данных этого алгоритма оценки качества при использовании. У данного подхода существуют следующие преимущества:

- поставщики и потребители алгоритмов оценки качества имеют возможность извлекать пользу из технических усовершенствований, необходимых на начальной стадии распространения системы показателей качества;
- в некоторых приложениях произведенные обновления могут извлекаться автоматически при наличии необходимой инфраструктуры;
- попытка оценки, связанная с идентификацией алгоритма оценки качества, перейдет обратно от потребителя или интегратора к поставщикам алгоритма оценки качества (которые в любом случае производят оценку);
- происходит развитие и совершенствование стандартизированных наборов тестов, так как:

1) в интересах поставщика алгоритмов оценки качества использовать набор (наборы) тестов, которым пользуются потребители,

2) со временем потребность в новых наборах тестов отпадет, а применение подобных наборов тестов будет тщательно анализироваться биометрическим сообществом;

- совершенствование наборов тестов будет способствовать развитию НДНПК.

Для осуществления обмена необходимы:

- идентификатор поставщика алгоритма оценки качества;
- ИАОК;
- минимальное и максимальное теоретические значения выходных данных алгоритма,
- уникальное имя применяемого набора тестов, например «FERET-grayscale» в случае распознавания лица;
- перечень образцов, которые были обработаны, например для базы данных FERET «Duplicate 1» в случае распознавания лица.

Каждый получит возможность публиковать новые наборы тестов (биометрические образцы совместно со схемой наименований).

Самоописываемый язык, например XML, будет использоваться для описания наборов данных и результатов оценки. Результаты оценки могли бы содержаться в основном реестре или на сайте поставщика, доступ к которому осуществляется посредством ссылки, размещенной в основном реестре.

Пример реализации языка с использованием языка XML приведен в приложении В настоящего стандарта.

9 Нормализация

Нормализация данных о показателях качества — это процесс, при помощи которого данные о показателе качества обрабатываются их получателем, для того чтобы показатели качества, сгенерированные разными алгоритмами, имели одинаковое смысловое значение.

В результате работы алгоритма оценки качества биометрическому образцу присваивается необработанный показатель качества. В целях интерпретации необработанного показателя получателю показателя необходимо иметь в своем распоряжении некоторую контекстную информацию о нем, которая предоставляется из следующих источников.

1 Внешние источники в форме метаданных или автономных данных (например, стандарт), представляющие собой руководство по интерпретации показателя. Если показатель качества сопровождается идентификатором алгоритма, примененного для генерации показателя качества соответствующего образца (например, ИОАК), то программное обеспечение получателя следует настроить для использования данных, предоставленных поставщиком (например, с помощью предложенных порогов) для наилучшей обработки образца. Также алгоритм может быть применен для осуществления анализа, в результате которого происходит полноценная оптимизация интерпретации

показателей при наличии локальных приложений и данных. С помощью идентификации алгоритма показатели, созданные разными алгоритмами, могут быть дифференцированы таким образом, что к образцу применяются разные пороги в зависимости от источника показателя качества.

2 Внутренние источники в форме нормализированного показателя качества. Посредством нормализации данных о показателе качества предоставляется контекстная информация о показателе. Примером является показатель качества, представляющий собой вероятность того, что результатом сопоставления образца станет ложное несовпадение.

Использование ИАОК позволяет осуществлять заданное поставщиком масштабирование. Например, получатель файла прибегнет к анализу степени взаимосвязи показателей качества с ВЛС и ВЛНС образцов, обработанных его компаратором. Результат может быть использован, например, для определения пороговых значений допуска. Данный метод обеспечивает получателя информацией, которая необходима для интерпретации показателей в соответствии с условиями эксплуатации и приложением получателя, и предоставляет возможность применять в одной и той же системе множество разных алгоритмов или версий алгоритмов.

С помощью совокупности данных по нормализации показателя качества осуществляется представление соответствующей интерпретации показателей качества посредством нормализации показателей качества или процентильного ранга показателя качества (ПРПК). ПРПК обеспечивает универсальное выражение и интерпретацию количественного показателя качества образца, которая состоит в том, что на выходе алгоритма оценки качества «X», обрабатывающего образец «Y», получается процентильный ранг качества «Z». Перевод необработанных показателей качества в показатели процентильного ранга достигается путем обработки совокупности стандартизированных образцов алгоритмом оценки качества и попарного соединения всех возможных выходных необработанных показателей качества с показателями процентильного ранга.

Более подробная информация представлена в приложении А настоящего стандарта.

Приложение А
(справочное)

Процедуры построения набора данных для нормализации показателя качества

А.1 Общие сведения о наборе данных для нормализации показателя качества

Цель метода стандартизации при помощи НДНПК — обеспечение согласующейся и допускающей взаимодействие интерпретации показателей качества. Для достижения поставленной цели необходима стандартизация набора(ов) данных и обеспечение каждой биометрической модальности собственным набором(ами) данных. Для каждого набора данных необходимо определить несколько классов показателей, с каждым из которых сопоставляется диапазон показателей качества. Данные о нескольких образцах, попадающие в каждый класс, собираются и вносятся в стандарты, предназначенные для разработчиков алгоритмов и получателей (пользователей) показателей качества. Поставщикам алгоритмов рекомендуется модифицировать или предоставить такую функцию преобразования, чтобы выходные данные алгоритмов качества согласовывались с рекомендуемым диапазоном показателя для данных о каждом образце, предоставленном в базе данных стандарта. Однако при наличии более одной версии стандартизированного набора данных важно предоставить сведения о том наборе данных, который используется для нормализации показателя качества. В дальнейшем набор данных вносится в стандарт и становится доступным для разработчиков и получателей (пользователей) показателей качества.

В настоящем приложении установлена процедура составления контрольной базы данных, ориентированной на эффективность. Результатом выполнения процедуры является набор образцов, снабженный целевым показателем качества. Результат является классификацией согласованных показателей схожести, полученных от множества компараторов. Подобные контрольные наборы используются при такой нормализации показателя качества, сгенерированного разными алгоритмами оценки качества, что показатели качества от разных алгоритмов оценки качества могут быть единообразно интерпретированы в рамках целесообразного диапазона. Метод применяется при настройке алгоритма оценки качества в соответствии с условиями эксплуатации, когда известны тип компаратора и тип данных. Разработчики алгоритма оценки качества, выполняющие общую задачу, также могут использовать данный метод. Входными данными для процедуры является репрезентативная база данных образцов. Выходными данными является сопоставленный каждому образцу целевой скалярный показатель качества. Метод предполагает доступность репрезентативного алгоритма сопоставления, который используется для сопоставления образов для создания как истинного, так и ложного показателя схожести. Таким образом, подразумевается, что на каждого человека может приходиться два и более образца.

А.2 Данные

Данные, собранные в целевом рабочем приложении, являются наиболее показательными. Рекомендуется создать контрольный набор с большей долей образцов, представляющих сложность для компаратора, чем присутствует в генеральной совокупности. Контрольный набор образцов должен быть максимально сбалансирован, например, в случае с отпечатками пальцев — в отношении положения пальца (правый/левый указательный/большой/средний), типа отпечатка (прокатный/оттисковый), пола, возраста и сканера отпечатков пальцев. Недостаточное количество данных делает создание сбалансированного набора данных проблематичным.

В некоторых областях использование предлагаемого метода полностью зависит от стандартизации наборов данных биометрических образцов и их доступности. Наборы данных, как минимум, должны формироваться таким образом, чтобы: 1) оставалось достаточно места для хранения файлов, 2) осуществлялась защита целостности сохраненных в наборе данных файлов и 3) обеспечивалась безопасность файлов посредством контроля доступа к ним квалифицированных работников.

А.3 Присвоение целевого показателя качества

Каждому изображению, находящемуся в контрольном наборе данных, необходимо присвоить основанный на эффективности показатель качества. Соответствующие показатели схожести гарантируют эффективность показателей качества. Рассмотрим биометрический набор данных, содержащий $N_i \geq 2$ образцов, $d_i^{(1)}, d_i^{(2)}, \ldots, d_i^{(N_i)}$, для каждого из M субъектов $i = 1,\ldots, M$, где каждое изображение содержит лишь одну характеристику, например изображение лишь одного пальца или одной радужной оболочки. При помощи данной процедуры присваиваются значения качества $q_i^{(1)}, q_i^{(2)}, \ldots, q_i^{(N_i)}$, $i = 1,\ldots, M$ всем изображениям в контрольном наборе данных.

I Для каждого компаратора V_k, $k = 1,\ldots, K$, из K доступных компараторов и каждой записи экземпляра $d_i^{(u)}$ (то есть u-го образца субъекта i) необходимо:

1) сгенерировать набор всех возможных парных показателей сравнения при помощи k-го компаратора:

$$S_k = \{S_{i,i}^{u,v} \mid S_{i,i}^{u,v} = V_k(d_i^{(u)}, d_i^{(v)})\}$$

$$u = 1, \ldots, N_i,\ v = u + 1, \ldots, N_i; \qquad (1)$$

$$i = 1, \ldots, M,$$

где V_k — k-й из всех доступных $k = 1, \ldots, K$ компараторов.

В результате получаем $P(N_i, 2) = N_i (N_i - 1)$элементов в наборе S_{ii}, где для каждого образца субъекта i присутствует $N_i - 1$ парных сопоставлений.

При этом полагаем, что генерируются несимметричные показатели сравнения, то есть

$$V_k(d_i^{(u)}, d_j^{(v)}) \neq V_k(d_j^{(v)}, d_i^{(u)}).$$

Если появляются симметричные показатели сравнения, то следует осуществлять процедуру с использованием лишь половины показателей сравнения;

2) сгенерировать набор непарных показателей сравнения при помощи k-го компаратора посредством сравнения образцов, взятых у индивида i, и образцов, взятых у всех $j = 1, \ldots, M$ и $i \neq j$ остальных индивидов:

$$S_{ij} = \{S_{i,j}^{u,v} \mid S_{i,j}^{u,v} = V_k(d_i^{(u)}, d_j^{(v)})\}$$

$$u = 1, \ldots, N_i; v = 1, \ldots, N_j, \tag{2}$$

$$i = 1, \ldots, M; j = 1, \ldots, M; i \neq j.$$

В результате получаем $\sum_{j=1, j \neq i}^{M} N_j$ непарных показателей сравнения для образца $d_i^{(u)}$;

3) внести (i, u) в набор T, если его парные показатели сравнения больше, чем все его непарные показатели сравнения, то есть

$$S_{i,j}^{u,v} > S_{i,j}^{u,w} \quad \forall j \neq i, \; v \neq u, \; w,$$

это состояние ранга 1;

4) вычислить целевую степень полезности для образца $d_i^{(u)}$ таким образом.

$$utility_i^u = \frac{m_{i,u}^{\text{парный}} - m_{i,u}^{\text{непарный}}}{\sigma_{i,u}^{\text{парный}} + \sigma_{i,u}^{\text{непарный}}}, \tag{3}$$

где $m_{i,u}^{\text{парный}}$ является средним значением парных показателей сравнения образца $d_i^{(u)}$, вычисленным следующим образом:

$$m_{i,u}^{\text{парный}} = \frac{\sum_{v=1, v \neq u}^{N_i} S_{i,i}^{u,v}}{N_i - 1}, \tag{4}$$

а $m_{i,u}^{\text{непарный}}$ является средним значением непарных показателей сравнения образца $d_j^{(u)}$, вычисленным следующим образом:

$$m_{i,u}^{\text{непарный}} = \frac{\sum_{j=1, j \neq i}^{M} \sum_{v=1}^{N_j} S_{i,j}^{u,v}}{\sum_{j=1, j \neq i}^{M} N_j}, \tag{5}$$

схожесть $\sigma_{i,u}^{\text{парный}}$ является стандартным отклонением парных показателей сравнения образца $d_i^{(u)}$, вычисленным следующим образом:

$$\sigma_{i,u}^{\text{парный}} = \sqrt{\frac{\sum_{v=1, v \neq u}^{N_i} (S_{i,i}^{u,v} - m_{i,u}^{\text{парный}})^2}{N_i - 1}}, \tag{6}$$

а $\sigma_{i,u}^{\text{непарный}}$ является стандартным отклонением парных показателей сравнения образца $d_i^{(u)}$, вычисленным следующим образом:

$$\sigma_{i,u}^{\text{непарный}} = \sqrt{\frac{\sum_{j=1, j \neq i}^{M} \sum_{v=1}^{N_j} (S_{i,j}^{u,v} - m_{i,u}^{\text{непарный}})^2}{\sum_{j=1, j \neq i}^{M} N_j}}. \tag{7}$$

Как только все целевые показатели степени полезности ($utility_i^u$ $\forall u = 1, \ldots, N_i$ и $\forall i = 1, \ldots, M$) вычислены, необходимо осуществить следующие процедуры.

a) Вычислить две эмпирические суммирующие функции распределения — одну для парных показателей сравнения высшего ранга набора T:

$$C(z) = \frac{|\{utility_i^u : (i,u) \in T, utility_i^u \le z\}|}{|\{utility_i^u : (i,u) \in T, utility_i^u \le \infty\}|}, \quad (8)$$

а другую для тех показателей, которых нет в наборе:

$$W(z) = \frac{|\{utility_i^u : (i,u) \notin T, utility_i^u \le z\}|}{|\{utility_i^u : (i,u) \notin T, utility_i^u \le \infty\}|}. \quad (9)$$

Данные суммирующие функции распределения изображены на рисунке А.1 на примере изображений указательных пальцев правой руки, полученных методом прямого сканирования 6000 субъектов, и показателях коммерческого компаратора отпечатков пальцев, где $N_i = 2$ для всех $i = 1, \ldots, 6000$.

b) Выбрать разрешающую способность целевого показателя качества (L). Этой разрешающей способностью является количество уровней качества для целевого показателя качества, и поэтому целевые показатели качества представляют собой значения $q = 1, \ldots, L$, где 1 является наименьшим показателем качества, а L — наибольшим показателем качества. Значение L может равняться любому целому числу от 2 до 100, например 5. Следует отметить, что чем больше значение L, тем больше образцов (большее M) требуется для верного присвоения показателей качества.

c) Сгруппировать целевые показатели полезности образцов в L ячейках, основанных на квантилях распределений $C(.)$ и $W(.)$ целевых показателей в уравнениях (8) и (9). Границы ячеек могут быть выбраны таким образом, чтобы стать точками перегиба функций суммирующего распределения, то есть $C(.)$ и $W(.)$. Один из вариантов, для $L = 5$, представлен в таблице А.1, где W^{-1} и C^{-1} являются квантильными функциями, а $C^{-1}(0)$ и $C^{-1}(1)$ обозначают эмпирически полученные наименьшее и наибольшее значения, соответственно (аналогично $W^{-1}(0)$ и $W^{-1}(1)$); x и y являются надлежащими процентильными точками, выбранными на основании формы $C(.)$. Границы ячеек при $x = 0{,}25$ и $y = 0{,}75$ представлены на рисунке А.1.

d) Образцу $d_j^{(u)}$ необходимо присвоить целевое качество $g_i^{(u)}$, соответствующее ячейке его целевого показателя полезности из уравнения (3).

Процедура применяется ко всем образцам $d_j^{(u)}$, $u = 1, \ldots, N_j$, $i = 1, \ldots, M$ и всем K компараторам. Поскольку распределение парных показателей сравнения и непарного показателя сравнения одного образца отличается от распределения парных показателей сравнения и непарного показателя сравнения другого образца, у разных образцов от одного и того же субъекта могут быть разные целевые показатели полезности и, следовательно, разные целевые показатели качества.

II Необходимо объединить результаты группировки для всех K компараторов. Выбор функции объединения зависит от того, желательно ли и, если желательно, то до какой степени, обобщение целевых показателей качества. Ниже представлен перечень возможных вариантов:

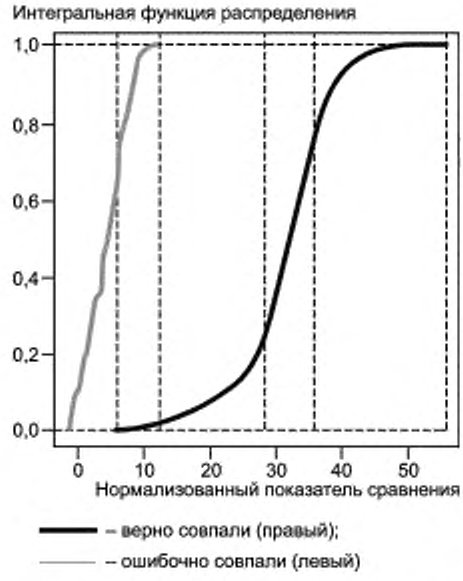
a) единогласие. Образцы, которым всеми K компараторами присвоены одинаковые показатели качества, становятся элементами контрольного набора данных о качестве. Образцы, не получившие единогласного решения, могут быть отвергнуты;

b) медианная или иная определенная процентильная точка. Образцы, которым более X процентов K компараторов присвоили одинаковые показатели качества, становятся элементами контрольного набора данных о качестве. Остальные образцы могут быть отвергнуты. Стоит отметить, что $X = 100$ означает единогласие, а если $X = 50$, то применяется правило решения большинством голосов;

c) среднее арифметическое. Конечный целевой показатель качества каждого из образцов представляет собой среднее арифметическое присвоенных K компараторами показателей качества.

Таблица А.1 — Группировка целевых показателей полезности

Группа	Диапазон целевых показателей полезности
1	$\{z_j : -\infty < z_j < C^{-1}(0.01)\}$
2	$\{z_j : C^{-1}(0.01) \le z_j < W^{-1}(1)\}$
3	$\{z_j : W^{-1}(1) \le z_j < C^{-1}(x)\}$
4	$\{z_j : C^{-1}(x) \le z_j < C^{-1}(y)\}$
5	$\{z_j : C^{-1}(y) \le z\}$



Примечание — Вертикальные линии показывают один из способов группировки нормализированных показателей сравнения

Рисунок А.1 — Эмпирические и интегральные функции распределения истинных и ложных показателей

А.4 Размер набора данных для нормализации показателя качества

Набор данных для нормализации показателя качества должен быть достаточно большим, чтобы его можно было использовать как для обучения, так и для испытания различных собственных алгоритмов оценки качества с целью нормализировать выходные данные. Они должны соответствовать категориям НДНПК, представленным в настоящем стандарте. Размер образца N в случае контролируемого испытания может быть оценен следующим образом:

$$N > 32 \frac{s^2}{d^2}, \tag{10}$$

где s — типичная ошибка (шум); d — наименьший полезный эффект (сигнал).

Поскольку количество категорий равно четырем, можно принять $d = 1$, так как при переходе от одной категории к соседней величина изменяется на единицу. Наиболее распространенная ошибка заключается в неверном размещении показателя качества в ближайшей к нему категории, вместо правильной категории, что дает $s = 2$. Таким образом, на основании уравнения (10) минимальный требуемый размер набора данных на категорию составляет 128 или 512 для НДНПК. Половина набора данных может быть использована для обучения, а вторая половина — для испытания. Подход, заключающийся в двойной перекрестной проверке, может быть использован для вычисления общей вероятности ошибки для случая, когда наборы тестов для обучения и испытания меняются местами и средняя вероятность ошибки вычисляется на основе выходных данных каждого набора тестов, как предлагается в литературе, посвященной статистике и машинному обучению.

Приложение В
(справочное)

Пример стандартизированного обмена результатами работы алгоритмов оценки качества

Ниже представлен пример кодирования в XML для поставщика «SampleVendor» с id = 123, публикующего результаты алгоритма «SampleAlgo_v10» с id = 456 на наборе тестов «FERET-grayscale» и «FERET-color».

```
<?xml version="1.0" encoding="UTF-8"?>
<iso-vendor-quality-report
    xmlns:iso="http://gost.e-ecolog.ru
    xmlns:xsi="xmlns:xsi=http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="gost.e-ecolog.ru
                http://www.iso.org/WD29794-1.xsd"
    quality-vendor-id="123"
    quality-algorithm-id="456"
    quality-algorithm-min-value="0.0"
    quality-algorithm-max-value="100.0">
    <iso:testset
      name="FERET-grayscale"
      location="gost.e-ecolog.ru                                    >
      <iso:sample
        name="00001fa010_930831"
        quality-value="73.64"/>
      <iso:sample
        name="00002fa010_930831"
        quality-value="48.91"/>
    </iso:testset>
    <iso:testset
      name="FERET-color"
      location="gost.e-ecolog.ru                    rferet/home.html">
      <iso:sample
        name="00002_931230_fa"
        quality-value="51.26"/>
      <iso:sample
        name="00002_931230_fb"
        quality-value="82.17"/>
    </iso:testset>
</iso-vendor-quality-report>

<?xml version="1.0» encoding=»UTF-8"?>
<xs:schema
    xmlns:xs="http://gost.e-ecolog.ru
    targetNamespace="gost.e-ecolog.ru
    elementFormDefault="qualified">
    <xs:annotation>
      <xs:documentation xml:lang="en">
        ISO WD 29794-1 Vendor Quality Report
      </xs.documentation>
    </xs:annotation>
    <xs:element
      name="iso-vendor-quality-report"
      type="iso-vendor-quality-report-type">
    </xs:element>
    <xs:complexType
      name="iso-vendor-quality-report-type">
      <xs:sequence minOccurs="0" maxOccurs="unbounded">
        <xs:element
          name="iso:testset"
```

```
            type="iso:testset-type"/>
    </xs:sequence>
    <xs:attribute
        name="quality-vendor-id"
        type="xs:ID"
        use="required"/>
    <xs:attribute
        name="quality-algorithm-id"
        type="xs:ID"
        use="required"/>
    <xs:attribute
        name="quality-algorithm-min-value"
        type="xs:float"
        use="required"/>
    <xs:attribute
        name="quality-algorithm-min-value"
        type="xs:float"
        use="required"/>
    </xs:complexType>
    <xs:complexType
        name="testset-type">
        <xs:sequence minOccurs="0" maxOccurs="unbounded">
            <xs:element
                name="iso:sample"
                type="iso:sample-type"/>
        </xs:sequence>
        <xs:attribute
            name="name"
            type="xs:ID"
            use="required"/>
        <xs:attribute
            name="location"
            type="xs:anyURI"
            use="required"/>
    </xs:complexType>
    <xs:complexType
        name="sample-type">
        <xs:attribute
            name="name"
            type="xs:ID"
            use="required"/>
        <xs:attribute
            name="quality-value"
            type="xs:float"
            use="required"/>
    </xs:complexType>
</xs:schema>
```

Приложение С
(справочное)

Процедуры для объединения показателей качества, основанных на полезности

С.1 Назначение

В приложении С представлены процедуры объединения показателей качества, основанных на полезности, по множеству образцов, например суммирование по всей организации. Результатом становится суммарное значение, позволяющее контролировать качество. Суммирование качества должно осуществляться в рамках соответствующего применения, например суммирование показателей качества всех зарегистрированных образцов организации, или суммирование показателей качества всех верификационных образцов организации. В операциях, когда пользователи часто контактируют с биометрической системой (приложение учета рабочего времени), показатели качества могут быть объединены для каждого пользователя в отдельности. Такой подход позволит выявить исполнителей, систематически предоставляющих образцы низкого качества.

С.2 Метод

Предположим, некоторая организация собирает отпечатки пальцев и вычисляет степень качества каждого из них при помощи алгоритма оценки качества. При этом показатели качества разбиты на L уровней (без ограничения общности) так, что $q = 0, \ldots, L$, где $q = 0$ и $q = L$ указывают наименьшее и наибольшее значения показателей качества соответственно. Если количество отпечатков, собранных за некоторый период в оперативной обстановке, равно n и включает n_q отпечатков качества q, то можно вычислить средний показатель качества всех n образцов. Однако среднее арифметическое не является предпочтительным способом суммирования показателей качества, так как при этом все образцы, независимо от их показателей качества, имеют одинаковое весовое значение. Если, напротив, прогнозируемой полезностью отпечатка пальца с качеством q является $u_q = U(q)$, то итоговый показатель качества предпочтительнее вычислять по формуле

$$\overline{q} = \frac{\sum_{q=0}^{L} u_q n_q}{\sum_{q=0}^{L} n_q}. \qquad (11)$$

Если значение полезности u_q по сути является оценкой вероятности ложного недопуска для образцов с показателем качества q контрольной системы по верификации с использованием отпечатков пальцев, работающей с использованием некоторого порога, то $\overline{q}$ является оценкой ожидаемой вероятности отказа. Далее представлена процедура для вычисления значения полезности u_q для различных уровней алгоритма оценки качества, при использовании которой суммарный показатель качества представляет собой оценку ожидаемой вероятности ошибки.

Предположим, что биометрический набор содержит $2N$ пар изображений, полученных от N субъектов. Первый образец представляет собой регистрационный образец, а второй — аутентификационный образец. Показатели качества образцов являются целочисленными $q_j^{(1)}$ и $q_j^{(2)}$ для $j = 1, \ldots, N$. Применяя V алгоритмов сравнения, получаем

- N показателей истинной схожести $S_{jj}^{(v)}$ и
- вплоть до $N(N-1)$ показателей ложной схожести $S_{jk}^{(v)}$ при $j \neq k$, где $v = 1, \ldots, V$ и $V > 1$.

1 Для всех алгоритмов сравнения v и показателей качества q вычисляется ВЛНС (τ, i) аутентификационных образцов с качеством i с регистрационными образцами с качеством, равным или большим i при пороге срабатывания τ при помощи показателей истинной схожести алгоритма сравнения v. При этом считается, что чем выше показатель качества, тем выше качество

для $(v = 1, \ldots, V)$

для $(i = 1, \ldots, L)$

$$\text{ВЛНС}^{v}(\tau, i) = \frac{|\{s_{jj}^{(v)} : s_{jj} \leq \tau, q_j^{(1)} \geq i, q_j^{(2)} = i\}|}{|\{s_{jj}^{(v)} : s_{jj} < \infty, q_j^{(1)} \geq i, q_j^{(2)} = i\}|} \qquad (12)$$

конец цикла,

результатом чего становится следующий массив данных

$$\begin{pmatrix} \text{ВЛНС}^1(\tau,1)\text{ВЛНС}^2(\tau,1)\ldots\text{ВЛНС}^V(\tau,1) \\ \text{ВЛНС}^1(\tau,2)\text{ВЛНС}^2(\tau,2)\ldots\text{ВЛНС}^V(\tau,2) \\ \text{ВЛНС}^1(\tau,L)\text{ВЛНС}^2(\tau,L)\ldots\text{ВЛНС}^V(\tau,L) \end{pmatrix}$$

2 Вычисляется весовое значение u_i:

$$u_i = \frac{\sum_{v=1}^{V} \text{ВЛНС}^v(\tau, i)}{\sum_{q=0}^{L}\sum_{v=1}^{V} \text{ВЛНС}^v(\tau, q)} \tag{13}$$

Таким образом, объединенный показатель качества по всей организации:

$$Q = \sum_{i=0}^{L} u_i p_i, \tag{14}$$

где u_i — апостериорные вероятности, указанные выше.

В качестве вероятностей данные значения не входят в известный пользователям диапазон. Стоит отметить, что если бы качество всех образцов было бы наивысшей степени (то есть $q = L$), то результатом было бы $Q = u_L$. Аналогично наихудшим событием было бы то, если бы качество всех образцов на предприятии было $q = 0$, результатом чего было бы $Q = u_0$. Таким образом, результатом данного формульного выражения было бы суммирование показателей качества в диапазоне $[u_L, u_0]$. Пользователи должны воспринимать уравнение (14) как меру ожидаемой общей ВЛНС. Однако рекомендуется трансформировать $[u_L, u_0]$ в предпочтительный диапазон [0, 100], где 0 представляет собой наименьший показатель качества, а 100 — наибольший показатель качества. Этого можно добиться посредством:

a) либо соотнесения суммарного показателя качества Q (то есть ожидаемой вероятности ошибки) с собственным диапазоном качества при помощи обратной функции полезности:

$$\overline{Q} = U^{-1}(Q) = U^{-1}\left(\sum_{i=0}^{L} u_i p_i\right), \tag{15}$$

где U^{-1} — аппроксимирующая функция (например, кусочно-линейной аппроксимацией) пар (i, u_i);

b) либо преобразования (например, линейного отображения) диапазона $[u_L, u_0]$ в диапазон [0, 100]. Таким образом, суммарные показатели качества, преобразованные к диапазону [0, 100], имеют вид:

$$\overline{Q} = \frac{100u_0}{u_0 - u_L} - \sum_{i=0}^{L} \frac{100u_i}{u_0 - u_L} p_i. \tag{16}$$

Примечание 1 — Весовые значения в уравнении (13) представляют собой оценки наблюдаемых вероятностей ложного несовпадения, вычисленных с использованием некоторого фиксированного порогового значения. В результате данные весовые значения в наивысшей степени точны в отношении данного конкретного порогового значения, но не так точны в отношении биометрических систем, функционирующих при иных пороговых значениях. В приложениях по верификации, где порог срабатывания установлен на уровне τ, пользователи алгоритма оценки качества должны следовать установленным процедурам для установления специальных весов.

Примечание 2 — Весовые значения в уравнении (13) представляют собой согласованные оценки. Это значит, что оценки были произведены при помощи наблюдаемых в наборе алгоритмов сравнения вероятностей ложного несовпадения. Указанные весовые значения не являются весовыми значениями, которые применялись бы для какого-либо одного алгоритма или установленного набора алгоритмов. Весовые значения в уравнении (13) рассматриваются в качестве наилучшего метода оценки и предназначены для использования до тех пор, пока неизвестны другие подробности о приложении. В приложениях по верификации, в которых известен и доступен определенный набор из одного или более алгоритмов сравнения, для установки назначенного весового значения пользователи алгоритма должны следовать установленным процедурам.

Приложение ДА
(справочное)

Сведения о соответствии ссылочных международных стандартов ссылочным национальным стандартам Российской Федерации

Таблица ДА.1

Обозначение ссылочного международного стандарта	Степень соответствия	Обозначение и наименование соответствующего национального стандарта
ИСО/МЭК 19794-1:2006	IDT	ГОСТ Р ИСО/МЭК 19794-1—2008 «Автоматическая идентификация. Идентификация биометрическая. Форматы обмена биометрическими данными. Часть 1. Структура»
ИСО/МЭК 19785-2:2006	IDT	ГОСТ Р ИСО/МЭК 19785-2—2008 «Автоматическая идентификация. Идентификация биометрическая. Единая структура форматов обмена биометрическими данными. Часть 2. Процедуры действий регистрационного органа в области биометрии»
Примечание — В настоящей таблице использовано следующее условное обозначение степени соответствия стандартов: - IDT — идентичные стандарты.		

Библиография

[1] ISO/IEC 19795-1, Information technology — Biometric performance testing and reporting — Part 1: Principles and framework
[2] SC 37 SD2

УДК 004.93'1:006.89 ОКС 35.040 П85

Ключевые слова: информационные технологии, биометрия, качество биометрических образцов, критерии качества биометрического образца, показатель качества алгоритма оценки качества

Редактор *К.С. Савинова*
Технический редактор *В.Н. Прусакова*
Корректор *М.И. Першина*
Компьютерная верстка *И.А. Налейкиной*

Сдано в набор 16.05.2013. Подписано в печать 05.07.2013. Формат 60 × 84 1/8. Гарнитура Ариал.
Усл. печ. л. 2,79. Уч.-изд. л. 2,14. Тираж 79 экз. Зак. 762.

ФГУП «СТАНДАРТИНФОРМ», 123995 Москва, Гранатный пер., 4.
Набрано во ФГУП «СТАНДАРТИНФОРМ» на ПЭВМ.
Отпечатано в филиале ФГУП «СТАНДАРТИНФОРМ» — тип. «Московский печатник», 105062 Москва, Лялин пер., 6.

gost.e-ecolog.ru gost.e-ecolog.ru gost.e-ecolog.ru