ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГОСТ Р ИСО/МЭК 29159-1— 2017

Информационные технологии

БИОМЕТРИЯ

Калибровка, аугментация и объединение данных в биометрии

Часть 1

Формат объединения данных

**(ISO/IEC 29159-1:2010,
Information technology — Biometric calibration, augmentation and fusion data — Part 1: Fusion information format, IDT)**

Издание официальное

Москва
Стандартинформ
2017

gost.e-ecolog.ru

Предисловие

1 ПОДГОТОВЛЕН Научно-исследовательским и испытательным центром биометрической техники Московского государственного технического университета имени Н.Э. Баумана (НИИЦ БТ МГТУ им. Н.Э. Баумана) и Федеральным государственным унитарным предприятием «Всероссийский научно-исследовательский институт стандартизации и сертификации в машиностроении» (ВНИИНМАШ) на основе собственного перевода на русский язык англоязычной версии стандарта, указанного в пункте 4

2 ВНЕСЕН Техническим комитетом по стандартизации ТК 098 «Биометрия и биомониторинг»

3 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от 22 июня 2017 г. № 577-ст

4 Настоящий стандарт идентичен международному стандарту ИСО/МЭК 29159-1:2010 «Информационные технологии. Калибровка, аугментация и объединение данных в биометрии. Часть 1. Формат объединения данных» (ISO/IEC 29159-1:2010 «Information technology — Biometric calibration, augmentation and fusion data — Part 1: Fusion information format», IDT).

Наименование настоящего стандарта изменено относительно наименования указанного международного стандарта для приведения в соответствие с ГОСТ Р 1.5—2012 (пункт 3.5).

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

5 ВВЕДЕН ВПЕРВЫЕ

6 Некоторые элементы настоящего стандарта могут быть объектами патентных прав. Международная организация по стандартизации (ИСО) и Международная электротехническая комиссия (МЭК) не несут ответственности за установление подлинности каких-либо или всех таких патентных прав

*Правила применения настоящего стандарта установлены в статье 26 Федерального закона от 29 июня 2015 г. № 162-ФЗ «О стандартизации в Российской Федерации». Информация об изменениях к настоящему стандарту публикуется в ежегодном (по состоянию на 1 января текущего года) информационном указателе «Национальные стандарты», а официальный текст изменений и поправок — в ежемесячном информационном указателе «Национальные стандарты». В случае пересмотра (замены) или отмены настоящего стандарта соответствующее уведомление будет опубликовано в ближайшем выпуске ежемесячного информационного указателя «Национальные стандарты». Соответствующая информация, уведомление и тексты размещаются также в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет (*gost.e-ecolog.ru

© Стандартинформ, 2017

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 ... 1
5 Обозначения и сокращения ... 2
6 Формат объединения данных ... 2
6.1 Общие положения ... 2
6.2 Порядок следования байтов ... 4
6.3 Числовые значения ... 4
6.4 Блок «Заголовок объединения» ... 4
7 Общие элементы ... 8
7.1 Общие положения ... 8
7.2 Вид параметра ... 8
7.3 Происхождение параметра ... 9
7.4 Присутствующие распределения ... 10
7.5 Число сравнений ... 10
7.6 Флаг предварительной нормализации данных ... 10
8 Запись типа 1 ... 11
8.1 Назначение ... 11
8.2 Формат ... 11
8.3 Применение (справочно) ... 12
9 Запись типа 2 ... 12
9.1 Назначение ... 12
9.2 Формат ... 13
9.3 Применение (справочно) ... 14
10 Запись типа 3 ... 14
10.1 Назначение ... 14
10.2 Формат ... 15
Приложение А (справочное) Общие положения документа ... 17
Приложение В (справочное) Примеры функций распределения ... 19
Приложение С (справочное) Использование предварительной нормализации данных ... 20
Приложение D (справочное) Исходный код для оценки сплайна ... 22
Приложение ДА (справочное) Сведения о соответствии ссылочных международных стандартов национальным стандартам ... 23
Библиография ... 24

Введение

Разработка биометрических систем привела к появлению различных технологий и алгоритмов сравнения. Некоторые из них публикуются в открытом доступе, но большинство являются частной собственностью. Во многих современных приложениях для биометрической верификации или биометрической идентификации применена одна биометрическая модальность. Это означает, что применена одна определенная биологическая или поведенческая характеристика человека с целью более или менее однозначной идентификации личности. Например, в системе контроля доступа может быть осуществлена регистрация изображения руки и использованы ее геометрические характеристики. В рамках программы социальных льгот могут проводить сбор отпечатков пальцев заявителей и использовать их в качестве входных данных для поиска дубликата «один-ко-многим». Применение разных биометрических методов предполагает использование различного количества отличительной информации, а также влечет за собой разные проблемы, связанные с получением данных. Все биометрические системы в какой-то степени подвержены отказам; при этом существуют различные режимы отказа. Таким образом, представляется целесообразным комбинировать технологии или алгоритмы с целью повышения производительности и/или удобства использования биометрических систем. Такое комбинирование называется объединением. Объединение может быть мультимодальным (например, при использовании в качестве биометрических характеристик лица и пальца), мультиалгоритмическим (например, при использовании алгоритмов А и В), мультиэкземплярным (например, при использовании указательного и большого пальца), мультидатчиковым (например, при использовании оптического и ультразвукового датчика отпечатков пальцев) или являться мультипредставлением (например, при использовании трех изображений лица пользователя).

В настоящем стандарте рассмотрен наиболее общий и легко реализуемый метод объединения: объединение на уровне результатов сравнения. Это объединение осуществляется после того, как две или более системы обработали и сравнили биометрические данные индивида[1)] с одним или несколькими зарегистрированными образцами и предоставили скалярные результаты сравнения на выходе. Различают результаты сравнения подлинного лица (тот же самый человек) и «самозванца» (другой человек). Схему объединения разрабатывают с целью комбинирования этих результатов сравнения таким образом, чтобы уточнить границу классов результатов сравнения подлинных лиц и «самозванцев».

Распределения результатов сравнения являются уникальными для каждой биометрической подсистемы сравнения. Диапазоны результатов сравнения и формы распределений могут сильно отличаться. Объединение часто осуществляется в двух направлениях:

- в процессах, основанных на классификации, доступные результаты сравнения комбинируются напрямую для получения решения или результата сравнения на выходе;

- в процессах, основанных на нормализации, объединению предшествует преобразование каждого результата сравнения к общей области определения. Иногда эффективны простые методы нормализации, основанные на статистических параметрах, таких как среднее значение и стандартное отклонение, но в более сложных методах используется детальное знание о распределении результатов сравнения на входе. Формат объединения данных (ФОД), определенный в настоящем стандарте, предназначен для гибкой поддержки любого из доступных преобразований. Несмотря на установление стандартизированных средств обмена данными, настоящий стандарт поддерживает модульный подход к интеграции биометрических систем, при котором и алгоритмы сравнения, и алгоритмы объединения остаются защищенными «черными ящиками», что позволяет избежать нарушения прав интеллектуальной собственности. В настоящем стандарте рассмотрено приложение, в котором каждая из двух или более лежащих в основе технологий сбора данных и сравнения (например, геометрии руки и отпечатка пальца) генерирует результат сравнения, который подается в модуль объединения, инициализированного с помощью соответствующего экземпляра ФОД.

На рисунке 1 представлена логическая роль записей в процессе мультимодального объединения.

В настоящем стандарте определены контейнеры для хранения информации о распределении результатов сравнения, полученных от подсистемы сравнения. При этом не учитываются данные о совместном распределении, которые могут полностью получать статистические свойства многомерных результатов сравнения (т. е. данные от двух и более подсистем или модальностей). Это означает, что мультимодальное объединение не поддерживается описанием совместных распределений биометрических результатов сравнения. Это ограничение в большинстве случаев является незначительным, поскольку различные модальности часто считаются независимыми. Даже когда результаты сравнения

[1)] В биометрии термин «индивид» относится только к человеку.

не являются независимыми, как и в случае мультиалгоритмических приложений, применение методов объединения на уровне результатов сравнения остается эффективным, пусть и не оптимальным.

Настоящий стандарт предназначен для поддержки функциональной совместимости и обмена данными между биометрическими приложениями и системами. Он определяет требования, которые позволяют решать проблемы, возникающие при применении биометрических технологий, во множестве приложений распознавания личности вне зависимости от того, работают ли такие приложения в среде открытых систем или состоят из одной закрытой системы. Открытые системы строятся на основе стандартов, устанавливающих общепризнанные форматы данных, интерфейсы и протоколы, чтобы обеспечивать обмен данными и совместимость с другими системами, которые могут включать компоненты различной конструкции и производства. Закрытая система также может быть построена на основе общепризнанных стандартов и включать в себя компоненты различной конструкции и производства, но по сути к ней не предъявляются требования обмена данными и функциональной совместимости с любой другой системой.

Применение стандартов форматов обмена биометрическими данными и стандартов биометрических интерфейсов является необходимым условием достижения полного обмена данными и функциональной совместимости в случае биометрического распознавания в среде открытых систем. Биометрические международные стандарты, разработанные в рамках ИСО/МЭК СТК 1/ПК 37, образуют многоуровневый набор международных стандартов, состоящий из форматов обмена биометрическими данными и биометрических интерфейсов, а также профилей приложений, которые описывают порядок использования данных международных стандартов в конкретных областях применения.

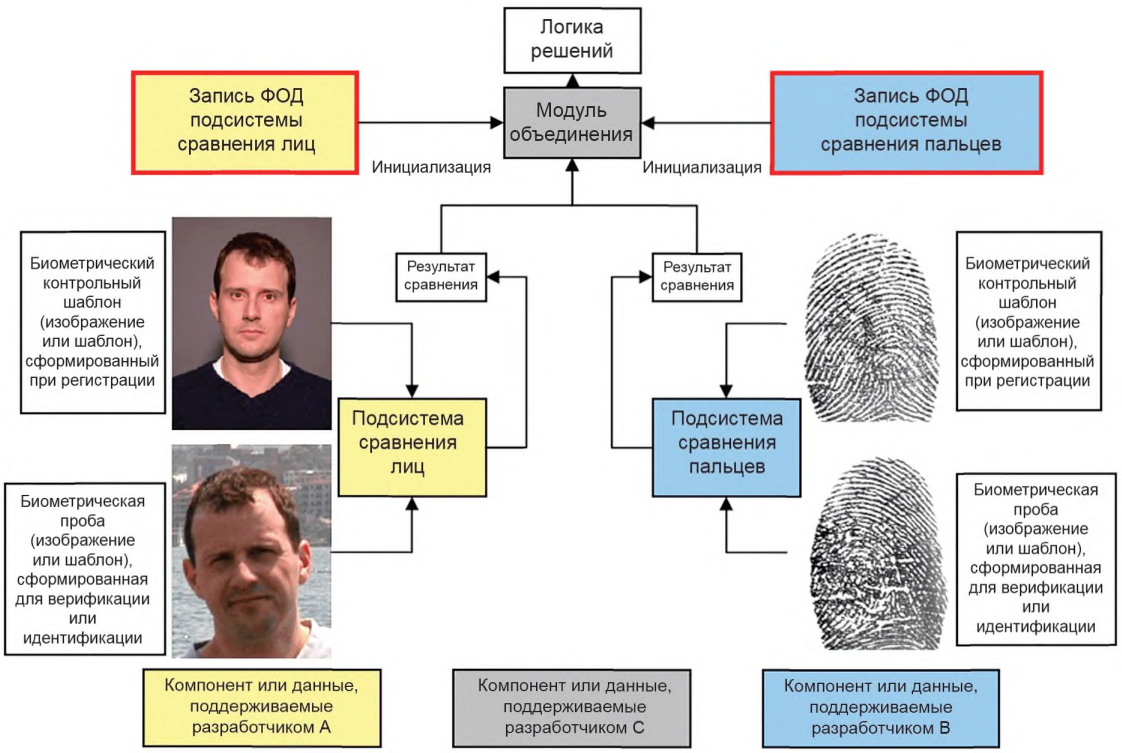


Рисунок 1 – Схема использования ФОД

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

Информационные технологии

БИОМЕТРИЯ

Калибровка, аугментация и объединение данных в биометрии

Часть 1
Формат объединения данных

Information technology. Biometrics.
Biometric calibration, augmentation and fusion data. Part 1. Fusion information format

Дата введения — 2017—10—01

1 Область применения

Настоящий стандарт определяет формат объединения данных в биометрии, который устанавливает машиносчитываемые форматы данных для описания статистических данных результатов сравнения, подаваемых на вход процесса объединения.

В настоящем стандарте не рассматриваются:

- процессы нормализации результатов сравнения;
- процессы объединения.

2 Соответствие

Записи соответствуют требованиям настоящего стандарта, если они удовлетворяют всем нормативным требованиям раздела 6. В разделе 6 предъявлены требования соответствия одному из разделов 8, 9 или 10, в которых, в свою очередь, предъявлены требования соответствия разделу 7.

3 Нормативные ссылки

В настоящем стандарте использованы нормативные ссылки на следующие стандарты, которые необходимо учитывать при его использовании. В случае датированных ссылок необходимо пользоваться только указанной редакцией. В случае недатированных ссылок следует пользоваться последней редакцией ссылочных документов, включая любые поправки и изменения к ним:

IEEE 754-2008 IEEE Standard for floating-point arithmetic (Стандарт ИИЭР для арифметики с плавающей точкой)

ISO/IEC 19785-1:2006 Information technology — Common biometric exchange formats framework — Part 1: Data element specification (Информационные технологии. Единая структура форматов обмена биометрическими данными. Часть 1. Спецификация элементов данных)

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

4 Термины и определения

В настоящем стандарте применены термины и определения по ИСО/МЭК 19794-1, а также следующие термины с соответствующими определениями:

4.1 **биометрический образец** (biometric sample)**:** Аналоговое или цифровое представление биометрических характеристик, предшествующее извлечению биометрических признаков.

Издание официальное

Примечание — Устройство сбора биометрических данных (биометрический сканер) — подсистема сбора биометрических данных, состоящая из одного компонента.

4.2 **функция распределения** (cumulative distribution function): Вероятность того, что случайная величина примет значение, меньше или равное произвольному числу.

4.3 **результат сравнения подлинного лица** (genuine score): Результат сравнения двух образцов от одного человека.

4.4 **результат сравнения «самозванца»** (impostor score): Результат сравнения двух образцов от разных людей.

4.5 **параметр положения** (location parameter): Общий количественный показатель позиции распределения.

Примечание — Параметр расположения не обязательно является средним значением распределения.

4.6 **функция плотности распределения вероятностей** (probability density function): Производная функции распределения.

4.7 **параметр масштаба** (scale parameter): Общий количественный показатель ширины распределения.

Примечание — Параметр масштаба, как правило, не является ни дисперсией, ни стандартным отклонением распределения.

4.8 **результат сравнения** (comparison score): Численное значение (множество значений), являющееся результатом процесса сравнения, полученное на выходе биометрической подсистемы сравнения.

Примечание — Термин «результат сравнения» применен в настоящем стандарте как для оценки степени различия (меньшая указывает на большую вероятность того, что образцы получены от одного и того же человека), так и для оценки степени схожести (большая указывает на большую вероятность того, что образцы получены от одного и того же человека). Эта разница подробно описана в 6.4.9.

5 Обозначения и сокращения

В настоящем стандарте применены обозначения и сокращения, приведенные ниже. В таблицах, определяющих структуры двоичных записей, символ «М» в столбце «статус» означает, что поле является обязательным, символ «О» — необязательным. Это значит, что байты для этих полей могут отсутствовать. В любом случае записи могут быть проанализированы, поскольку информация о наличии или отсутствии необязательного поля записывается в предыдущем поле.

ЕСФОБД — единая структура форматов обмена биометрическими данными (common biometric exchange formats framework, CBEFF);

ФР — функция распределения (cumulative distribution function, CDF);

ЭФР — эмпирическая функция распределения (empirical cumulative distribution function, ECDF);

ВЛД — вероятность ложного допуска (false acceptance rate, FAR);

ВЛС — вероятность ложного совпадения (false match rate, FMR);

ФОД — формат объединения данных (fusion information format, FIF);

ФПРВ — функция плотности распределения вероятностей (probability density function, PDF);

ID — идентификатор (identifier).

6 Формат объединения данных

6.1 Общие положения

6.1.1 Структура записи

Запись ФОД использована для поддержки модульности в мультимодальных биометрических системах и системах поддержки принятия решений. Структура записи приведена в таблице 1.

Примечание — Приложение должно установить профиль на основе настоящего стандарта. Профиль по умолчанию будет вызывать один из типов записей.

Таблица 1 — Структура записи ФОД

Блок «Заголовок объединения» (Fusion header block)	Запись типа 1 (Type 1 Record) (см. 6.1.3)	и/или	Запись типа 2 (Type 2 Record) (см. 6.1.4)	и/или	Запись типа 3 (Type 3 Record) (см. 6.1.5)

6.1.2 Структура блока «Заголовок объединения»

Структура блока «Заголовок объединения» определяет формат записи и указывает на ее содержание. Структура блока «Заголовок объединения» приведена в таблице 2.

Таблица 2 — Структура блока «Заголовок объединения»

Блок «Заголовок объединения»	=	Идентификатор формата (Format identifier)	Номер версии стандарта (Version number)	Длина записи (Record length)	Тип биометрической модальности (Biometric type)				
6.4		6.4.2	6.4.3	6.4.4	6.4.5				
25 байтов		4	4	4	3				
Необработанное дополнение				Идентификатор алгоритма подсистемы сравнения (Comparison subsystem product ID)	Идентификатор базы данных (Database ID)	Качество базы данных регистрации (Enrolment database quality)	Качество базы данных верификации (Verification database quality)	Смысловое значение результата сравнения (Score sense)	Число экземпляров типа (Number of type instances)
				6.4.6	6.4.7	6.4.8	6.4.8	6.4.9	6.4.10
				4	2	1	1	1	1

6.1.3 Структура записи типа 1

Структура блока «Заголовок объединения» определяет формат записи и указывает на ее содержание. Структура записи типа 1 приведена в таблице 3.

Таблица 3 — Структура записи типа 1

Структура записи типа 1 (Type 1 Record Structure)	=	Тип (Type)	Наличие распределений (Distributions Present)	Распределение «самозванцев» (Impostor Distribution)			Распределение подлинных лиц (Genuine Distribution)		
				Число сравнений (Num Comp)	Положение распределения (Loc)	Масштаб (Scale)	Число сравнений (Num Comp)	Положение распределения (Loc)	Масштаб (Scale)
8.2		8.2.3	8.2.4	8.2.5	8.2.6	8.2.7	8.2.8	8.2.9	8.2.10
26 или 50 байтов		1	1	4	10	10	4	10	10

6.1.4 Структура записи типа 2

Структура блока «Заголовок объединения» определяет формат записи и указывает на ее содержание. Структура записи типа 2 приведена в таблице 4.

Таблица 4 — Структура записи типа 2

Структура записи типа 2 (Type 2 Record Structure)	=	Тип	Присутствующие распределения	ФР «самозванцев»	ФР подлинных лиц
9.2		9.2.3	9.2.4	9.2.5	9.2.6
16*N*+13 байтов или 32*N*+22 байтов		1	1	16*N*+11	16*N*+11

6.1.5 Структура записи типа 3

Структура блока «Заголовок объединения» определяет формат записи и указывает на содержание. Структура записи типа 2 приведена в таблице 5.

Таблица 5 — Структура записи «Тип 3»

Структура записи типа 3 (Type 3 Record Structure)	=	Тип	Присутствующие распределения	ФР «самозванцев»	ФР подлинных лиц
10.2		10.2.3	10.2.4	10.2.5	10.2.6
16*N*-18 или 32*N*-38 байтов		1	1	16*N*-20	16*N*-20

6.2 Порядок следования байтов

В записи ФОД и всех однозначно определенных блоках данных в ней все многобайтовые величины должны храниться в обратном порядке следования байтов, то есть старшие байты должны иметь более низкие уровни адресов памяти, чем младшие байты.

Пример — Значение 1025 (два в степени 10 плюс один) будет храниться в двух байтах: первый байт = 00000100b и второй байт = 00000001b.

6.3 Числовые значения

Все числовые значения, присутствующие в определенных типах записей в настоящем стандарте, являются целыми величинами фиксированной длины без знака, если не указано иное.

Все числовые значения, приведенные в тексте настоящего стандарта, являются десятичными, кроме тех, которым предшествует 0x (они являются шестнадцатеричными) и тех, которые оканчиваются на «b» (они являются двоичными).

Таблица 6 — Текстовое представление числового значения

Пример значения	Основание системы счисления	Десятичное значение
1010b	2	10
39	10	39
0xF5	16	245

Числа двойной точности должны соответствовать ИИЭР 754.

Примечание — Чтобы избежать численной неточности, спецификации ИИЭР 754 может быть недостаточно.

6.4 Блок «Заголовок объединения»

6.4.1 Общие положения

Блок «Заголовок объединения», приведенный в таблице 7, должен присутствовать в качестве первого блока всех записей ФОД.

Таблица 7 — Блок «Заголовок объединения»

Поле	Статус	Длина, байт	Допустимые значения	Примечание
Идентификатор формата	M	4	0x46494600	Нуль-терминированная строка «FIF» с символами ASCII
Номер версии стандарта	M	4	0x30313000	Нуль-терминированная строка «010» с символами ASCII
Длина записи	M	4	$1 \le L \le 2^{32}-1$	Длина входной записи в байтах
Тип биометрической модальности	M	3	$0 \le t \le$ 0x080000	Модальность, от которой получена запись

Окончание таблицы 7

Поле	Статус	Длина, байт	Допустимые значения	Примечание
Идентификатор алгоритма подсистемы сравнения	M	4	≥ 0	Текущий идентификатор, определенный разработчиком
Идентификатор базы данных	M	2	≥ 0	Текущий идентификатор, определенный разработчиком
Качество биометрического образца при регистрации	M	1	[0-100], 254, 255	Совокупные качества биометрических образцов, используемые для вычисления статистических данных результатов сравнения
Качество биометрического образца при верификации	M	1	[0-100], 254, 255	
Смысловое значение результата сравнения	M	1	0 или 1	Степень различия или степень схожести? (См. 6.4.9)
Число экземпляров типа	M	1	$1 \le N \le 4$	Значение 0 недопустимо

6.4.2 Идентификатор формата

Поле «Идентификатор формата» (4 байта) является нуль-терминированной строкой с тремя символами ASCII «FIF» в начале для соответствия настоящему стандарту.

6.4.3 Номер версии стандарта

Поле «Номер версии стандарта» (4 байта) является нуль-терминированной строкой с тремя символами ASCII.

Первый и второй символы обозначают номер версии стандарта, третий символ — номер поправки или изменения данной редакции.

Номер версии настоящего стандарта будет 0x30313000, т. е. 010 — номер версии 1, номер редакции 0.

6.4.4 Длина записи

Поле «Длина записи» (4 байта) должно содержать полную длину в байтах всей записи. Она определяется как сумма длины блока «Заголовок объединения» (25 байтов) и длины следующей за ним записи типа 1, 2 или 3.

6.4.5 Тип биометрической модальности

Значение поля «Тип биометрической модальности» (3 байта) должно быть взято из перечня ЕСФОБД биометрических модальностей, приведенного в ИСО/МЭК 19785-1, 6.5.6. Это значение позволяет приложению определить, какая биометрическая модальность представлена, с помощью экземпляра ФОД.

Пример — Для результатов сравнения в реализации, использующей изображения венозного русла, это значение будет 0x040000.

6.4.6 Идентификатор алгоритма подсистемы сравнения

Идентификатор алгоритма подсистемы сравнения (т. е. сравнения биометрических данных с биометрическими контрольными шаблонами) (4 байта = 2 байта + 2 байта), который позволяет получать информацию о результатах сравнения, содержащуюся в данной записи ФОД, должен быть записан в соответствии с требованиями, приведенными в таблице 8. Эти два значения являются идентификаторами биометрических продуктов ЕСФОБД, описанных в ИСО/МЭК 19785-1[1)].

Таблица 8 — Идентификаторы биометрических продуктов ЕСФОБД

Поле	Кем определяется	Длина, байт	Описание
Идентификатор биометрического продукта (сохраненный в первых 2 байтах)	МАБП[2)]	2	См. ИСО/МЭК 19785-1
Номер версии стандарта (сохраненный во вторых 2 байтах)	Разработчик	2	

[1)] Заменен на ISO/IEC 19785-1. Однако для однозначного соблюдения требования настоящего стандарта, выраженного в датированной ссылке, рекомендуется использовать только указанное в этой ссылке издание.

[2] МАБП — Международная ассоциация биометрической промышленности [International biometric industry association (IBIA)]. В настоящий момент данная организация называется «Международная ассоциация биометрии и идентификации» [International Biometrics & Identification Association (IBIA)].

Примечания

1 Можно утверждать, что поле идентификатора алгоритма не требуется, потому что статистическая информация включает в себя все, что требуется модулю объединения. Данное поле однако является полезным для осуществления поддержки, например контроля версий и кэширования.

2 Одного идентификатора биометрического продукта, скорее всего, недостаточно для выполнения нормализации результатов сравнения. Тем не менее он может быть полезен для приложений, использующих записи, определенные в настоящем стандарте.

6.4.7 Идентификатор базы данных

Определенный разработчиком идентификатор базы данных (2 байта), используемый подсистемой сравнения разработчика для получения результатов сравнения, которые применяются при вычислении данных, содержащихся в полях 4 и 5, должен быть записан в соответствии с требованиями, приведенными в таблице 9. Поле «Идентификатор базы данных» может быть использовано:

- разработчиком подсистемы сравнения, чтобы произвольно указать, какие данные использовались при вычислении содержимого записи ФОД;
- в профиле приложения или документах с требованиями, чтобы обязать разработчиков использовать определенные входные данные [например, (55) — база данных калибровки отпечатков пальцев MINEX POEBVA (MINEX POEBVA fingerprint calibration database)].

Примечание — ИСО/МЭК СТК 1/ПК 37 поддерживает процесс регистрации других баз данных для их использования в создании записей, определенных в настоящем стандарте. Реестр идентификаторов баз данных включает в себя базы данных, перечисленные в таблице 9.

Таблица 9 — Идентификаторы баз данных

Идентификатор базы данных	Описание	Модальность
0	Не определена	
1	Неизвестена	
2	Составная	
3—15	Зарезервировано	
16	FERET Face Database gost.e-ecolog.ru	Лицо
17	Yale Face Database http://cvc.yale.edu/projects/yalefaces/yalefaces.html	Лицо
18	PIE Database CMU gost.e-ecolog.ru _418.html	Лицо
19	AR Database gost.e-ecolog.ru _face_DB.html	Лицо
20	FRGC I Database http://face.nist.gov/frvt/	Лицо
21	FRGC II Database http://face.nist.gov/frvt/	Лицо
22	FRVT 2002 HCInt (i.e. FRVT 2006 Low Resolution) Database http://face.nist.gov/frvt/	Лицо
23	FRVT 2006 High Resolution Database http://face.nist.gov/frvt/	Лицо
24	FRVT 2006 Very High Resolution Database http://face.nist.gov/frvt/	Лицо
25	FVC 2000	Палец
26	FVC 2002 DB1 http://bias.csr.unibo.it/fvc2002/	Палец
27	FVC 2002 DB2 http://bias.csr.unibo.it/fvc2002/	Палец
28	FVC 2002 DB3 http://bias.csr.unibo.it/fvc2002/	Палец
29	FVC 2002 DB4 (synthetic) http://bias.csr.unibo.it/fvc2002/	Палец
30	FVC 2004 DB1 http://bias.csr.unibo.it/fvc2004/	Палец
31	FVC 2004 DB2 http://bias.csr.unibo.it/fvc2004/	Палец
32	FVC 2004 DB3 http://bias.csr.unibo.it/fvc2004/	Палец
33	FVC 2004 DB4 (synthetic) http://bias.csr.unibo.it/fvc2004/	Палец
34	FVC 2006 DB1 http://bias.csr.unibo.it/fvc2006/	Палец

Окончание таблицы 9

Идентификатор базы данных	Описание	Модальность
35	FVC 2006 DB2 http://bias.csr.unibo.it/fvc2006/	Палец
36	FVC 2006 DB3 http://bias.csr.unibo.it/fvc2006/	Палец
37	FVC 2006 DB4 (synthetic) http://bias.csr.unibo.it/fvc2006/	Палец
38	NIST Special Database 27 gost.e-ecolog.ru	След пальца
39	NIST Special Database 29 gost.e-ecolog.ru	Палец
40	MCYT Fingerprint subcorpus gost.e-ecolog.ru _EN.html	Палец
41	MCYT Signature subcorpus http://atvs.ii.uam.es/bbdd_EN.html	Подпись
42	BANCA still http://gost.e-ecolog.ru	Лицо
43	BANCA video http://www.ee.surrey.ac.uk/banca/	Голос
44	BANCA high quality http://gost.e-ecolog.ru	Голос
45	BANCA low quality gost.e-ecolog.ru	Голос
46	NIST Speaker Verification gost.e-ecolog.ru	Голос
47	NIST Speaker Verification gost.e-ecolog.ru	Голос
48	CASIA Iris http://gost.e-ecolog.ru	Радужная оболочка глаза
49	Bath Iris gost.e-ecolog.ru	Радужная оболочка глаза
50	ICE 2005 http://iris.nist.gov/ice/	Радужная оболочка глаза
51	ICE 2006 http://iris.nist.gov/ice/	Радужная оболочка глаза
52	NIST MINEX DHS2 calibration set http://fingerprint.nist.gov	Палец
53	NIST MINEX POE calibration set gost.e-ecolog.ru	Палец
54	NIST MINEX DOS calibration set gost.e-ecolog.ru	Палец
55	NIST MINEX POEBVA calibration set http://fingerprint.nist.gov	Палец
56	TURBINE GUC100 http://gost.e-ecolog.ru	Палец
57—2047	Зарезервировано	
2048—65535	Определено разработчиком, возможно, не публичным и не уникальным среди разработчиков	

Примечания

1 Хотя некоторые из этих баз данных не находятся в открытом доступе, они перечислены здесь для поддержки калибровки в будущем.

2 Хотя некоторые из этих баз данных не находятся в открытом доступе, они перечислены здесь в качестве справочной информации для пользователей настоящего стандарта.

3 Производительность систем объединения будет зависеть от используемых баз данных, так как распределения результатов сравнения, полученных путем сравнения образцов из одной базы данных, как правило, будут отличаться от распределений результатов сравнения, полученных путем сравнения образцов из другой базы данных. Стабильность распределений рассмотрена в приложении С.

6.4.8 Качество базы данных

Качества биометрических образцов в базе данных могут быть объединены для формирования скалярного значения качества базы данных. Два значения суммарного качества (1 байт) должны быть сохранены в последовательных полях: первое — для контрольных шаблонов; второе — для полученных образцов. Значение 0 представляет минимально возможное качество, а значение 100 — максимально возможное качество. В таблице 10 приведены допустимые значения для каждого поля.

Таблица 10 — Значения качества базы данных

Значение	Описание
0—100	Определенное значение
254	Неопределенное значение, т. к. не было попыток его определить
255	Неопределенное значение, т. к. попытка определить качество не удалась

6.4.9 Смысловое значение результатов сравнения

Необработанные данные, полученные от биометрических подсистем сравнения, представляют собой оценку степени различия либо оценку степени схожести. Маленькие значения оценки степени различия указывают на большую вероятность того, что они являются результатами сравнения подлинного лица; в случае оценки степени схожести на это указывают большие значения. Для записи смыслового значения результатов сравнения должно использоваться однобайтовое поле со значениями, приведенными в таблице 11.

Таблица 11 — Коды смыслового значения результатов сравнения

Смысловое значение	Значение (1 байт)
Оценка степени различия	0
Оценка степени схожести	1

Примечание — Использование слова «различие» не обязательно подразумевает, что значения обладают метрическим свойством.

6.4.10 Число экземпляров типа

Блок «Число экземпляров типа» (1 байт) должен представлять количество записей типа 1, 2 или 3, включенных в запись ФОД. Запись должна содержать ноль или один экземпляр каждого типа, но всегда должен присутствовать хотя бы один экземпляр одного из них. Таким образом, число экземпляров типа должно быть 1, 2 или 3.

Пример — Если присутствует и экземпляр типа 1, и экземпляр типа 3, значение будет 2.

Примечание — Профиль приложения (или эквивалентная спецификация) может надлежащим образом вызывать один определенный тип.

7 Общие элементы

7.1 Общие положения

В настоящем разделе описаны общие структуры данных для поддержки записей типа 1, 2 или 3 настоящего стандарта. В настоящий раздел также включены таблицы, в которых перечислены значения и приведены их соответствующие описания.

Необходимо обратить внимание на то, что в некоторых таблицах первый столбец озаглавлен «Поле» и содержит числовые данные. Они приведены только для облегчения ссылок на строки в таблицах; эти данные не должны включаться в бинарные записи стандарта.

7.2 Вид параметра

В таблице 12 приведены целые значения в качестве идентификаторов для величин, необходимых для описания распределения.

Таблица 12 — Идентификаторы для статистических величин

Вид	Описание	Примечание
0	Не определено	Изготовитель пренебрегает указанием величины (как правило, неприемлемо)
1	Неизвестно	Неизвестная величина

Окончание таблицы 12

Вид	Описание	Примечание
2	Среднее значение	
3	Медиана	
4	Мода	
5	Минимальное	Предельные значения не являются настоящими параметрами положения
6	Максимальное	
7	(Минимум + максимум)/2	
8	Критерий Тьюки	
9	Общий параметр положения	
10—31	Не определено	
32	Дисперсия	
33	Стандартное отклонение	
34	Медианное абсолютное отклонение	На самом деле, 1,4826 (med|x-med|), где med — медиана, а постоянная 1,4826 используется для нормализации, чтобы ожидаемое значение медианы было равно стандартному отклонению нормального распределения
35	(Максимум - минимум)	Предельные значения не являются настоящими параметрами положения
36	Общий параметр масштаба	
37	Асимметрия	
38	Эксцесс	
39—65	Не определено	
66	Общий параметр I	
67	Общий параметр II	
68—95	Не определено	
96	Функция распределения	Как дискретные пары (x_i, $F(x_i)$), где $i = 1...N$, которые используются для записей типа 2
97	Функция распределения	Как B-сплайн распределение, которое используют для записей типа 3
98—255	Не определено	

Примечания

1 Приложение также может повлиять на то же самое преобразование независимо от вида значения, которое содержит запись, например линейное преобразуется в (x-mean)/stdev и (x-median)/mad, где mean — среднее значение, stdev — стандартное отклонение, median — медиана, mad — медианное абсолютное отклонение.

2 Диапазон параметров очень широк. Для отдельного приложения их необходимо определять более точно. Они, по сути, должны быть определены в профиле приложения или другом документе с требованиями.

7.3 Происхождение параметра

Значение поля «Происхождение» (1 байт), приведенное в таблице 13, следует использовать для указания происхождения параметра.

Таблица 13 — Происхождение статистических данных

Происхождение	Распределение каких значений было получено
0	Не определено — значение(я), чье(ьи) происхождение(я) не разглашается
1	Неизвестно — значение(я) с неустановленным происхождением
2	Эмпирическое — значение(я), оцененное(ые) на основе экспериментальных и выборочных данных

Окончание таблицы 13

Происхождение	Распределение каких значений было получено
3	Известно априори – значение(я), известное(ые) на этапе проектирования или исходя из теоретических соображений. Для распределения это означает, что оно(и) известно(ы) в замкнутой форме
4—254	Зарезервировано

7.4 Присутствующие распределения

Значение поля «Присутствующие распределения» (1 байт) должно быть использовано для обозначения того, какие именно распределения «самозванцев» и подлинных лиц включены в запись. Допустимые значения приведены в таблице 14. Поле «Присутствующие распределения» должно содержать значение 0x01, 0x02 или 0x03. Когда присутствуют оба распределения, распределение «самозванцев» должно предшествовать распределению подлинных лиц.

7.5 Число сравнений

Число сравнений подлинных лиц или «самозванцев» (4 байта), используемое для оценки информации типа 1, 2 или 3, должно быть закодировано в этом поле. Нулевое значение должно использоваться, когда это число неизвестно.

Таблица 14 — Допустимые значения поля «Присутствующие распределения»

Значение	Присутствующие распределения
0x01	«Самозванцы»
0x02	Подлинные лица
0x03	«Самозванцы» и подлинные лица
0x04 — 0xFF	Зарезервировано

7.6 Флаг предварительной нормализации данных

Значение поля «Флаг предварительной нормализации данных» (1 байт), приведенное в таблице 15, следует использовать, чтобы указать, находятся ли результаты сравнения, полученные от подсистемы сравнения, в произвольном диапазоне или они были предварительно нормализованы. В контексте настоящего стандарта это означает, что модуль объединения может рассматривать результаты, полученные от подсистемы сравнения, как равномерно распределенные на отрезке [0, 1]. Это делает возможным прямое использование результатов сравнения без необходимости интерполяции. В приложении С приведена информация о значении поля «Флаг предварительной нормализации данных».

Примечания

1 Поле «Флаг предварительной нормализации данных» присутствует в записях типа 2 или 3 для распределений подлинных лиц и «самозванцев». Если подсистема сравнения представляет предварительно нормализованные результаты сравнения «самозванцев», результаты сравнения подлинных лиц будут неравномерны, и наоборот. Таким образом, поле «Флаг предварительной нормализации данных» в записях, соответствующих настоящему стандарту, не сможет указать на то, что результаты сравнения и «самозванцев», и подлинных лиц предварительно нормализованы.

2 Реализации, совместимые с BioAPI [1)] [1], возвращают «значения ВЛД». Это означает, что полученные внутри системы результаты сравнения «самозванцев» либо изначально распределены на отрезке [0, 1], либо были нормализованы с помощью ожидаемой ФР «самозванцев» этих результатов сравнения. В результате можно рассчитывать, что результаты сравнения «самозванцев» на выходе будут равномерно распределены на отрезке [0, 1].

Таблица 15 — Коды предварительной нормализации

Статус предварительной нормализации	Диапазон результатов сравнения	Значение (1 байт)
Не нормализованы	$-\infty < x < \infty$	0
Предварительно нормализованы	$0 \le x \le 1$	1

[1)] BioAPI — биометрический прикладной программный интерфейс (biometric application programming interface).

8 Запись типа 1

8.1 Назначение

Запись типа 1 содержит минимальную статистическую информацию о результатах сравнения «самозванцев» и/или подлинных лиц, полученных от биометрической системы. Эта информация может быть использована для масштабирования результатов сравнения перед операцией объединения.

8.2 Формат

8.2.1 Поддерживаемые типы данных — подтип *A*

Подтип *A* содержит вид (см. 7.2), происхождение (см. 7.3) и значение скалярной статистической величины. Формат должен соответствовать требованиям, приведенным в таблице 16.

Таблица 16 — Формат подтипа *A*

Поле	Статус	Описание	Длина, байт	Тип данных	Допустимое значение	Пример
1	M	Вид параметра	1	uint (целое число без знака)	См. таблицу 12	3 (медиана)
2	M	Происхождение параметра	1	uint	0—3	2 (эмпирическое значение)
3	M	Значение параметра	8	double (действительное число двойной точности)	Диапазон действительных чисел двойной точности	$2{,}998 \cdot 10^8$

8.2.2 Определение

Формат записи типа 1 должен соответствовать требованиям таблицы 17. Должны быть представлены данные либо «самозванцев», либо подлинных лиц, либо и те и другие данные (т. е. допустимые значения 26 или 50 байтов).

Таблица 17 — Формат записи типа 1

Поле	Статус	Описание	Длина, байт	Тип данных	Допустимое значение	Пример
1	M	Тип	1	uint	1	1
2	M	Присутствующие распределения	1	Упакованное битовое поле	0x01, 0x02, 0x03	0x01
3	O	Число сравнений «самозванцев»	4	uint	$[0,2^{32} - 1]$	40000
4		Положение распределения «самозванцев»	10	Подтип *A*		(3; 1; 2,998)
5		Масштаб распределения «самозванцев»	10	Подтип *A*		(34; 1; 0,308)
6	O	Число сравнений подлинных лиц	4	uint	$[0,2^{32} - 1]$	240
7		Расположение распределения подлинных лиц	10	Подтип *A*		(3; 1; 8,310)
8		Масштаб распределения подлинных лиц	10	Подтип *A*		(34; 1; 1,406)

8.2.3 Тип

Тип записи типа 1 должен быть 1 и храниться в 1 байте.

8.2.4 Присутствующие распределения

Распределения, присутствующие в записи типа 1, должны быть записаны в битах поля «Присутствующие распределения».

8.2.5 Число сравнений «самозванцев»

Число сравнений, используемых при вычислении статистических результатов сравнения «самозванцев», должно быть записано в поле «Число сравнений».

8.2.6 Положение распределения «самозванцев»

Поле «Подтип *A*», при наличии, должно содержать вид, происхождение и значение параметра положения распределения результатов сравнения «самозванцев». За этим значением, при наличии, должна следовать информация о масштабе, описанном в 8.2.7.

Примечание — Если это целесообразно, то схема нормализации может включать в себя только перенос (без масштабирования); в этом случае параметр масштаба, описанный в 8.2.7, будет 1.

8.2.7 Масштаб распределения «самозванцев»

Поле «Подтип *A*», при наличии, должно содержать вид, происхождение и значение параметра масштаба распределения результатов сравнения «самозванцев».

8.2.8 Число сравнений подлинных лиц

Число сравнений, используемых при вычислении статистических результатов сравнения подлинных лиц, должно быть записано в поле «Число сравнений».

8.2.9 Положение распределения подлинных лиц

Поле «Подтип *A*», при наличии, должно содержать вид, происхождение и значение параметра положения распределения результатов сравнения подлинных лиц. За этим значением, при наличии, должна следовать информация о масштабе, описанном в 8.2.10.

8.2.10 Масштаб распределения подлинных лиц

Поле «Подтип *A*», при наличии, должно содержать вид, происхождение и значение параметра масштаба распределения результатов сравнения подлинных лиц.

8.3 Применение (справочно)

Пусть дано две записи типа 1, одна из которых получена от подсистемы сравнения отпечатков пальцев (*A*), а другая из подсистемы сравнения радужных оболочек глаз (*B*). В этом случае объединение на уровне результатов сравнения может быть представлено в виде суммы *z*-нормализованных результатов сравнения как:

$$s = (a - \text{amean})/\text{asigma} + (b - \text{bmean})/\text{bsigma},$$

где *a* — необработанный результат сравнения, полученный от подсистемы сравнения отпечатков пальцев;

b — необработанный результат сравнения, полученный от подсистемы сравнения радужных оболочек глаз;

amean — оценка среднего значения результатов сравнения «самозванцев», полученных от подсистемы сравнения *A*;

asigma — оценка стандартного отклонения результатов сравнения «самозванцев», полученных от подсистемы сравнения *A*;

bmean — оценка среднего значения результатов сравнения «самозванцев», полученных от подсистемы сравнения *B*;

bsigma — оценка стандартного отклонения результатов сравнения «самозванцев», полученных от подсистемы сравнения *B*.

9 Запись типа 2

9.1 Назначение

Запись типа 2 должна содержать функцию распределения одного или обоих распределений результатов сравнения «самозванцев» и подлинных лиц. Если присутствуют оба распределения, распределение «самозванцев» должно предшествовать распределению подлинных лиц.

Примечания

1 Примеры ФР приведены в приложении А.

2 Модули объединения, требующие наличия ФПРВ, могут оценить их путем численного дифференцирования ФР. Тип 2 изначально не включает в себя ФПРВ из-за необходимости указания интервала.

3 Если $F(x)$ обозначает ФР, то ее значение для результата сравнения «а», включает в себя поиск результата сравнения «а» в структуре данных двумерного массива подтипа *B*. Это может быть реализовано с помощью двоичного поиска первого вектора для обнаружения индекса i для интервала $x_{i-1} \le a < x_i$ таким образом, что для получения значения $F(a)$ может быть использована интерполяция между $F(x_{i-1})$ и $F(x_i)$.

9.2 Формат

9.2.1 Поддерживаемые типы данных — подтип *B*

Подтип *B* является структурой для дискретных образцов вещественной одномерной функции. Он устанавливает таблицу поиска путем включения двух массивов одинаковой длины *N*: значения *x* и значения $f(x)$. Значения сортируются по *x* в порядке возрастания. Таким образом:

1) если i-й элемент первого вектора x_i, тогда i-й элемент второго вектора $f(x_i)$;

2) $x_1 \le x_i \le x_N$ для $i = 1 \ldots N$.

Формат должен соответствовать требованиям, приведенным в таблице 18.

9.2.2 Определение

Формат записи типа 2 должен соответствовать требованиям таблицы 19. Должны быть представлены данные либо «самозванцев», либо подлинных лиц, либо и те и другие данные [т. е. допустимые значения (13 + 16*N*) байтов или (24 + 32*N*) байтов].

Таблица 18 — Формат подтипа *B*

Поле	Статус	Описание	Длина, байт	Тип данных	Допустимое значение	Пример
1	M	Вид параметра	1	uint	96	96 (ФР)
2	M	Происхождение параметра	1	uint	0—3	1
3	M	Предварительная нормализация	1	uint	0, 1	0
4	M	Число сравнений	4	uint	$[0,2^{32} - 1]$	$1{,}5 \cdot 10^9$
5	M	Число элементов *N*	4	uint	$0 \le N \le 2^{32} - 1$	800
6	M	*N* значений *x*	8*N*	double	Диапазон действительных чисел двойной точности	(0,2; 0,4; 0,8)
7	M	*N* значений $f(x)$	8*N*	double	Диапазон действительных чисел двойной точности	(7; -4; -1)

Таблица 19 — Формат записи типа 2

Поле	Статус	Описание	Длина, байт	Тип данных	Допустимое значение	Пример
1	M	Тип	1	uint	2	2
2	M	Присутствующие распределения	1	Упакованное битовое поле	0x01 – 0x03	0x01
3	O	Распределение «самозванцев»	11 + 16*N*	Подтип *B*		Требуется одно или оба
4	O	Распределение подлинных лиц	11 + 16*N*	Подтип *B*		

Примечания

1 Значение N в полях «Длина» это число пар [x, $F(x)$].

2 Значения, содержащиеся в записи подтипа B для одного из распределений (например, «самозванцев»), как правило, получены эмпирически. Если известно лежащее в основе распределение, значения могут быть образцами известных ФР. Для нормального распределения ФР = 0,5(1 + erf($x/\sqrt{2}$)), где erf — интеграл вероятности.

9.2.3 Тип

Тип записи типа 2 должен быть 2 и храниться в 1 байте.

9.2.4 Присутствующие распределения

Распределения, присутствующие в записи типа 2, должны быть записаны в битах поля «Присутствующие распределения».

9.2.5 Распределение «самозванцев»

Поле «Подтип B», при наличии, должно содержать ФР результатов сравнения «самозванцев». Поле должно содержать монотонно возрастающие данные.

9.2.6 Распределение подлинных лиц

Поле «Подтип B», при наличии, должно содержать ФР результатов сравнения подлинных лиц. Поле должно содержать монотонно возрастающие данные.

9.3 Применение (справочно)

Пусть дано две записи типа 2, одна из которых получена от подсистемы сравнения отпечатков пальцев (A), а другая из подсистемы сравнения радужных оболочек глаз (B). В этом случае объединение на уровне результатов сравнения может быть представлено как:

$$s = -\log\left(M'_A(a)/N'_A(a)\right) - \log\left(M'_B(b)/N'_B(b)\right),$$

где a — необработанный результат сравнения, полученный от подсистемы сравнения отпечатков пальцев;

b — необработанный результат сравнения, полученный от подсистемы сравнения радужных оболочек глаз;

M'_A — ФПРВ результатов сравнения подлинных лиц, вычисленная как численная производная ФР результатов сравнения подлинных лиц, полученных от подсистемы сравнения A;

N'_A — ФПРВ результатов сравнения «самозванцев», вычисленная как численная производная ФР результатов сравнения «самозванцев», полученных от подсистемы сравнения A;

M'_B — ФПРВ результатов сравнения подлинных лиц, вычисленная как численная производная ФР результатов сравнения подлинных лиц, полученных от подсистемы сравнения B;

N'_B — ФПРВ результатов сравнения «самозванцев», вычисленная как численная производная ФР результатов сравнения «самозванцев», полученных от подсистемы сравнения B.

Примечания

1 Эта формула подробно рассмотрена в [7] и ИСО/МЭК ТО 24722: 2007, приложение А [3].

2 ФПРВ «самозванцев» для системы радужной оболочки глаза описана в замкнутой форме в [4]. ФР представляла бы собой неполную бета-функцию (от суммы двучленов) и была бы записана как имеющая происхождение 3 (см. 7.3). Она была бы включена в запись типа 2 с помощью соответствующей (т. е. достаточно точной) выборки.

10 Запись типа 3

10.1 Назначение

Запись типа 3 должна содержать функцию распределения одного или обоих распределений результатов сравнения «самозванцев» и подлинных лиц. Если присутствуют оба распределения, распределение «самозванцев» должно предшествовать распределению «подлинных лиц». Запись использует представление на основе B-сплайна парных элементов (x, $F(x)$) записи типа 2 и, как правило, вычисляется на основе этих данных. Эта запись является более компактной альтернативой записи типа 2 и используется, когда число пар (x, $F(x)$) слишком большое для записи типа 2. B-сплайн должен быть рассчитан в соответствии с алгоритмом, приведенным в [5]. Степень B-сплайна должна равняться 3 (т. е. быть кубической).

10.2 Формат

10.2.1 Поддерживаемые типы данных — подтип *C*

Подтип *C* является структурой для коэффициентов и вершин представления функции на основе *B*-сплайна. Сплайн вычисляется на основе *n* точек данных при условии, что $N << n$, где *N* — число вершин (значений результатов сравнения). Результатом является набор коэффициентов для вершин. Сплайн вычисляется с использованием метода минимизации наименьших квадратов с линейными ограничениями. Результатом является функция *O*(*N*), где *N* выступает в качестве параметра «точности». Формат должен соответствовать требованиям, приведенным в таблице 20.

Таблица 20 — Формат подтипа *C*

Поле	Статус	Описание	Длина, байт	Тип данных	Допустимое значение	Пример
1	M	Вид параметра	1	uint	97	
2	M	Происхождение параметра	1	uint	0, 1, 2, 3	
3	M	Предварительная нормализация	1	uint	0, 1	0
4	M	Число сравнений *n*	4	uint	$[0,2^{32} - 1]$	$1{,}5 \cdot 10^9$
5	M	Степень сплайна *K*	1	uint	$1 \le K \le 255$	3
6	M	Число вершин *N*	4	uint	$1 \le N \le 2^{32} - 1$	100
7	M	Вершинные значения *x*	8N	double	Диапазон действительных чисел двойной точности	
8	M	Коэффициенты	8*C* (см. примечание 2)	double	Диапазон действительных чисел двойной точности	

Примечания

1 Эта структура подходит для сплайна, описанного в [5].

2 Число коэффициентов составляет $C = N - K - 1$. Для сплайна кубической степени $K = 3$. Размер записи вершины и части коэффициентов будет меньше, чем *n* пар $(x, F(x))$ записи подтипа *B* при $N + C < 2n$. Таким образом, для $K = 3$ пространство будет сохранено при $N < n - 2$. Применение сплайна является эффективным, так как могут быть использованы многие вершины, меньшие, чем точки данных, при сохранении точности интерполяции. Продукт, ссылающийся на запись, должен определить соответствующее значение для *N*.

3 «Допустимые значения» являются здесь общими. Когда используется экземпляр подтипа *C*, например в записи типа 3 (см. 10), степень $K = 3$.

10.2.2 Определение

Формат записи типа 3 должен соответствовать требованиям таблицы 21. Должны быть представлены данные либо «самозванцев», либо подлинных лиц, либо и те и другие данные [т. е. допустимые значения $(16N - 18)$ байтов или $(32N - 38)$ байтов].

Таблица 21 — Формат записи типа 3

Поле	Статус	Описание	Длина, байт	Тип данных	Допустимое значение	Пример
1	M	Тип	1	uint	3	3
2	M	Присутствующие распределения	1	Упакованное битовое поле	0x01 — 0x03	0x01
3	O	Распределение «самозванцев»	$12 + 8N + 8(N - 4)$	Подтип *C* K=3		Требуется одно или оба
4	O	Распределение подлинных лиц	$12 + 8N + 8(N - 4)$	Подтип *C* K=3		

Примечания

1 Первая производная представления ФР на основе сплайна сама является сплайном. Она представляет собой ФПРВ. Она непрерывна для всех x. Уравнение (4) в [5] представляет собой формулу для производной.

2 Одним из требований настоящего стандарта является использование линейных ограничений, указанных в [5], для обеспечения монотонности сплайна.

3 *B*-сплайн, как правило, не равен исходным данным в вершинах, т. е. сплайн $(x_i) \neq F(x_i)$.

4 Кубическая функция имеет достаточно высокую степень для обеспечения точной интерполяции.

5 *B*-сплайн является более точным представлением функции, чем сглаживание N вершин, т. к. он вычисляется по n точкам при $n >> N$.

6 В приложении D приведен код C ++ для оценки функции сплайна.

10.2.3 Тип

Тип записи типа 3 должен быть 3 и храниться в 1 байте.

10.2.4 Присутствующие распределения

Распределения, присутствующие в записи типа 2, должны быть записаны в битах поля «Присутствующие распределения».

10.2.5 Распределение «самозванцев»

Поле «Подтип *C*», при наличии, должно содержать ФР результатов сравнения «самозванцев». Поле должно содержать монотонно возрастающие данные.

10.2.6 Распределение подлинных лиц

Поле «Подтип *B*», при наличии, должно содержать ФР результатов сравнения подлинных лиц. Поле должно содержать монотонно возрастающие данные.

Приложение А
(справочное)

Общие положения документа

А.1 Общие положения

Настоящий стандарт определяет запись ФОД для хранения статистических данных результатов сравнения, полученных от биометрических подсистем сравнения. В разделе 6 рассмотрена структура записи и определен общий заголовок ФОД. Как показано в таблице А.1, запись может содержать три типа данных, которые определены в разделах 8—10. Эти типы имеют некоторые общие элементы, которые представлены в разделе 7. Приложение может профилировать настоящий стандарт путем вызова одного из трех типов и ограничения его дополнительного содержания.

Таблица А.1 — Таксономия типов ФОД

Тип	Уровень	Описание	Предусмотренное применение	Пример нормализации результатов сравнения
1	Результат сравнения	Параметры положения и масштаба одного или более распределений подлинных лиц и «самозванцев»	Поддержка любой функции масштабирования, которая использует только параметры положения и масштаба распределений	Масштабирование, достигаемое путем использования одного или более распределений, т. е. $y = (x - m_I)/s_I$, где m_I и s_I – среднее значение и стандартное отклонение распределения «самозванцев» соответственно
2	Результат сравнения	ЭФР одного или более распределений подлинных лиц и «самозванцев»	Выполнение любой операции масштабирования на основе полной информации о распределении	Масштабирование путем использования распределения «самозванцев» $y = N_I(x)$, где $N_I(x)$ — ФР «самозванцев» без учета оценки ВЛС подсистемы сравнения
3	Результат сравнения	*B*-сплайн аппроксимация функции распределения типа 2	Выполнение любой операции масштабирования на основе полной информации о распределении	Как для типа 2

А.2 Выбор типов

Пользователям настоящего стандарта следует отметить, что тип 1, будучи очень компактным, поддерживает только рудиментарные методы объединения на уровне результатов сравнения. Альтернативой является использование типов 2 и 3, которые полностью кодируют распределение результатов сравнения и способны поддерживать достаточно сложные схемы объединения. Этим типам должно отдаваться предпочтение при использовании в силу возможности достижения более высокой точности.

Что касается типов 2 и 3, следует отметить, что запись типа 3 получается по сути на основе данных записи типа 2 и является ее функциональным эквивалентном. Вычисление записи типа 3, как правило, должно быть параметризовано для получения более компактных результатов по сравнению с записью типа 2. Тип 3, с другой стороны, требует использования численного метода для вычисления представления сплайна. Хотя в настоящем стандарте не приведен исходный код для типа 3, тип 3 использовать предпочтительнее, чем тип 2, в силу его компактности.

А.3 Совместимость типов

Взаимодействие в рамках ФОД относится к использованию с помощью модуля объединения, например записи типа 3 от разработчика *A* совместно с записью типа 2 от разработчика *B*. Пользователям настоящего стандарта следует обратить внимание, что тип 1 ограниченно совместим с самим собой и с другими типами. Например, запись типа 1, содержащая информацию о среднем значении и стандартном отклонении, будет представлять худшие характеристики при использовании совместно с записью типа 1, содержащей информацию о медиане и медианном абсолютном отклонении. Кроме того, взаимодействие типа 1 с типами 2 и 3 затруднено. В тех случаях, когда в модуль объединения поступают записи типа 1 и типа 2, совместимость может быть достигнута двумя способами:

- способ с понижением точности заключается в вычислении статистических данных записей типа 1 на основе данных типа 2. Например, медиана — это 50-й перцентиль, и среднее значение может быть оценено путем выборки;

- способ с повышением точности заключается в построении записи типа 2 или 3 на основе записи типа 1 при наличии предположения о форме распределения (например, нормальное), а затем параметризации с помощью параметров положения и масштаба типа 1.

Использовать приведенные способы не рекомендуется. Они могут применяться только в том случае, когда данные типа 2 или 3 недоступны. По этим причинам пользователям настоятельно рекомендуется профилировать настоящий стандарт до его использования. Это означает, что должен быть сформирован официальный документ, который включает в себя требование, что все разработчики должны обеспечивать, например, чтобы запись типа 2 содержала распределение как «самозванцев», так и подлинных лиц. Документ может быть формальным профайлом, или документом с требованиями (техническое задание), или каким-либо другим обязательным руководством.

Пользователи также должны предоставить автору профайлы для поддержки перехода с одного продукта на другой. В таком профайле будет содержаться, например, требование, чтобы разработчики обеспечивали как библиотеку алгоритмов сравнения, так и соответствующую запись типа 3 формата объединения данных подлинных лиц и «самозванцев».

А.4 Расширяемость

Настоящий стандарт является расширяемым. В частности, при пересмотре в стандарт могут быть включены другие типизированные записи. Эти записи должны устанавливать форматы для поддержки новых, альтернативных и более сложных процессов объединения (возможно, например, на основе совместных плотностей вероятностей) или различных приложений (например, для случайного объединения [8]).

А.5 Качество для обеспечения объединения

Стандартный ФОД не включает запись для статистических данных результатов сравнения, зависящих от измерений качества биометрических образцов [2]. Однако следует отметить, что модуль объединения, инициализируемый записями ФОД, тем не менее может использовать значения качества биометрических образцов на входе как часть каждой последующей операции объединения. Кроме того, значения качества могут быть использованы разработчиком в процессе начальной подготовки собственных записей ФОД.

Приложение В
(справочное)

Примеры функций распределения

В.1 Общие положения

На рисунке В.1 показаны две функции распределения типовых коммерческих биометрических устройств сравнения. В каждом случае шкала оси абсцисс отражает внутренний характер лежащей в основе подсистемы сравнения. Значения вертикальной оси располагаются на отрезке [0, 1]. Ступенчатые участки являются ЭФР, гладкие линии — кубическими сплайнами ЭФР. ЭФР представляет собой, по определению, долю результатов сравнения, меньших или равных абсциссе. Поэтому сплайн проходит через вершины «ступеней».

График на рисунке В.1 а) был построен путем вычисления 44 пар (x, $F(x)$) на участке значений x [-1, 54]. График на рисунке В.1 б) был построен путем вычислений 571 пары (x, $F(x)$) на участке значений x [-1, 871]. Длинные правые участки графика не представлены, так как они были эмпирически усечены 95 %-ной квантилью для удобства использования. Гладкие кривые представляют собой распределения сплайна типа 3.

Использование сплайнов на основе записей ФОД поддерживает оценку ФР при произвольных значениях, в том числе:

- нецелочисленных значениях;
- значениях за пределами рассматриваемого диапазона;
- интегральных значениях, которые не были получены в исходном образце результатов сравнения.

Но если подсистема сравнения изначально предоставляет только интегральные результаты сравнения, то почему сплайн является актуальным? В качестве альтернативы может быть использована справочная таблица для типа 2, но она имеет недостаток, который заключается в ее большом размере и отсутствии всех возможных целых значений. Это обстоятельство потребует выполнения интерполяции (или экстраполяции) в модуле объединения. Функция сплайна оценивает значения за пределами первоначального интервала как 0 или 1. Если ФР была вычислена с большим набором результатов сравнения, то будут наблюдаться более экстремальные значения (больше максимума и меньше минимума), и тогда ФР будет оцениваться по-другому.

Таким образом, рекомендуется использовать сплайн, т. к.:

- он является интерполяцией;
- первые два дифференциала являются непрерывными (поддерживают методы объединения, которые от этого зависят).

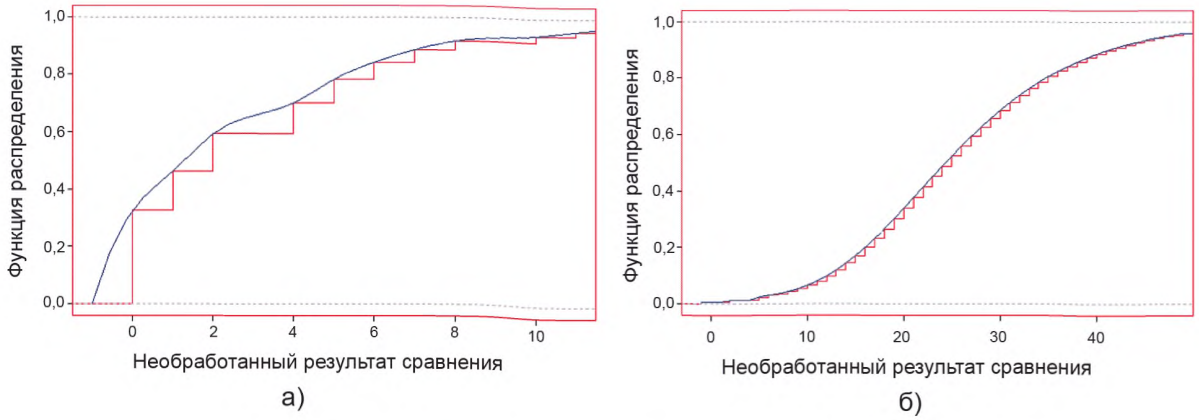


Рисунок В.1 — Пример ФР и их представлений на основе сплайнов

Приложение С
(справочное)

Использование предварительной нормализации данных

С.1 Общие положения

Предположим, что биометрическая подсистема сравнения внутри себя преобразует необработанные результаты сравнения так, что результаты сравнения «самозванцев» на выходе равномерно распределены на отрезке [0, 1]. Подсистема будет выполнять внутри себя преобразование, применяя некоторые оценки ФР «самозванцев» к своим внутренним результатам сравнения. Это практика по умолчанию использована в BioAPI [1], где результаты сравнения на самом деле расцениваются как вероятность ложного совпадения т. е.

$$s_{output} = 1 - N_{calibration}(s_{internal}),$$

где s_{output} — результат сравнения на выходе;
$s_{interna}$ — внутренний результат сравнения;
$N_{calibration}$ — оценка ФР «самозванцев», выступающая в качестве функции калибровки.

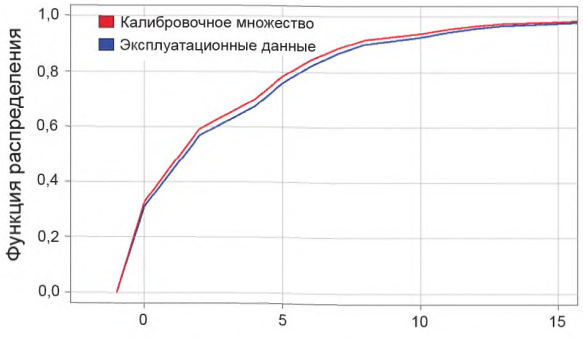
Распределение «самозванцев» используется таким образом, потому что считается, что оно не чувствительно к изменениям свойств выборки, таких как окружающая среда и популяция. Эта стабильность распределения «самозванцев» позволяет осуществлять нормализацию результатов сравнения портативных кросс-приложений.

С.2 Пример

На практике в распределении «самозванцев» будет присутствовать некоторое изменение в зависимости от приложений. На рисунке С.1 а) показаны:

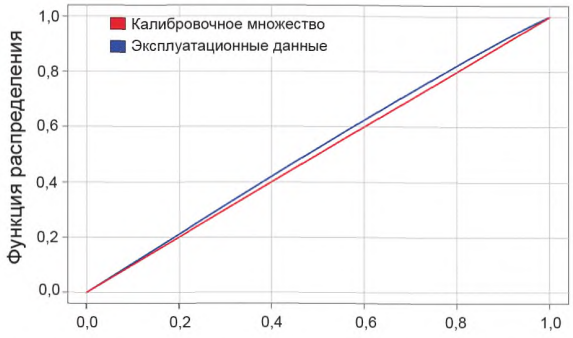
- ФР результатов сравнения «самозванцев», полученных путем сравнения образцов из калибровочного множества изображений;
- ФР результатов сравнения «самозванцев», полученных путем сравнения образцов из непересекающегося множества изображений, собранных при различных условиях.

Между этими двумя ФР есть небольшие, но значащие различия. Эффект предварительной нормализации данных смоделирован на рисунке С.1, где показано, что после нормализации $1 - N(x)$, распределение калибровочного множества результатов сравнения является равномерным (по определению), но ФР эксплуатационных данных находится выше.



а)

а) верхний график — калибровочное множество;
нижний график — эксплуатационные данные



б)

б) верхний график — эксплуатационные данные;
нижний график — калибровочное множество

Рисунок С.1 — Примеры ФР с внутренними результатами сравнения и предварительно нормализованными результатами сравнения

Примечание — В данном приложении не рассмотрены результаты сравнения подлинных лиц, поскольку они будут иметь разные распределения и не будут равномерными ни при наличии, ни при отсутствии предварительной нормализации.

С.3 Передовая практика

Таким образом, использование предварительно нормализованных данных сопровождается предупреждением, что фактическая равномерность распределения предварительно нормализованных результатов сравнения «самозванцев» зависит от стабильности результатов сравнения. ФОД подходит к этому путем определения альтернативных форматов для:

- произвольно распределенных результатов сравнения;
- предварительно нормализованных результатов сравнения, для которых результаты сравнения «самозванцев» калибруются таким образом, чтобы быть равномерными на отрезке [0, 1].

Если разработчик обеспечивает запись ФОД, предоставляющую предварительно нормализованные результаты сравнения, то под этим подразумевается, что в процессе функционирования равномерно распределенные результаты сравнения «самозванцев» будут реализовываться в течение длительного периода эксплуатации. Если разработчик обеспечивает записи типа 2 или 3 с ФР «самозванцев», тогда заявленные результаты сравнения «самозванцев», преобразованные с помощью этой ФР, также должны быть равномерными. Разница в том, кто выполняет преобразование: разработчик ли обеспечивает выполнение преобразования внутри системы или изначально, или приложение (т. е. интегратор). В любом случае равномерность результатов является темой для испытаний.

Пример передовой практики — сбор приложением эксплуатационных результатов сравнения «самозванцев» и формирование записи ФОД, относящейся к конкретной ситуации. Она может периодически обновляться для учета трендов данных (например, сезонные изменения или долговременный тренд).

Приложение D
(справочное)

Исходный код для оценки сплайна

Для оценки представления сплайна ФР [5], содержащей запись типа 3, может быть использован следующий фрагмент кода:

```
// Оценить B-сплайн функцию k-ой степени для значения t.
// Значение t находится между вершинами j и j + 1.
// Функция является рекурсивной.
double B(const unsigned int j, const unsigned int k, const double t)
{
(k == 0)
return (knots[j] <= t && t < knots[j+1]) ? 1.0 : 0.0;
const double c1 = (knots[j+k] == knots[j] ) ? 0 : (t - knots[j]) / (knots[j+k] - knots[j] );
const double c2 = (knots[j+k+1] == knots[j+1]) ? 0 : (knots[j+k+1] - t) / (knots[j+k+1]
- knots[j+1]);
return c1*B(j, k-1, t) + c2*B(j+1, k-1, t);
}

// линейный поиск по возрастающим вершинам для интервала, содержащего значение t
unsigned int i = 0;
for ( ; knots[i+1] <= t ; i++ );

//дан интервал, вычислить значение сплайна при данном значении t
double value = 0.0;
for ( unsigned int j = i - k ; j <= i ; j++ )
value += x[j+1] * B(j, k, t);
```

Приложение ДА
(справочное)

Сведения о соответствии ссылочных международных стандартов национальным стандартам

Таблица ДА.1

Обозначение ссылочного международного стандарта	Степень соответствия	Обозначение и наименование соответствующего национального стандарта
IEEE 754:2008	—	*
ISO/IEC 19785-1:2006	IDT	ГОСТ Р ИСО/МЭК 19785-1—2008 «Автоматическая идентификация. Идентификация биометрическая. Единая структура форматов обмена биометрическими данными. Часть 1. Спецификация элементов данных»
ISO/IEC 19794-1	IDT	ГОСТ ISO/IEC 19794-1—2015 «Информационные технологии. Биометрия. Форматы обмена биометрическими данными. Часть 1. Структура»

* Соответствующий национальный стандарт отсутствует. До его принятия рекомендуется использовать перевод на русский язык данного международного стандарта.

Примечание — В настоящей таблице использовано следующее условное обозначение степени соответствия стандартов:

- IDT — идентичные стандарты.

Библиография

[1] ISO/IEC 19784-1 Information technology — Biometric application programming interface — Part 1: BioAPI Specification (Информационные технологии. Биометрический программный интерфейс. Часть 1. Спецификация биометрического программного интерфейса)

[2] Discriminative multimodal biometric authentication based on quality measures. J. Fierrez-Aguilar, J. Ortega-Garcia, J. Gonzalez-Rodriguez, and Josef Bigun. Pattern Recognition, 38(5):777—779, May 2005. gost.e-ecolog.ru iad/894.03/quality/biblio.html (Различная мультимодальная биометрическая аутентификации на основе показателей качества. Дж. Фиеррец—Агилар, Х. Ортега-и-Гарсиа, Дж. Гонсалес-Родригес и Йозеф Бигун. Распознавание образов, 38(5):777—779, май 2005 г.)

[3] ISO/IEC TR 24722:2007 Information technology — Biometrics — Multimodal and other multibiometric fusion (Информационные технологии. Биометрия. Мультимодальные и другие мультибиометрические объединения)

[4] The importance of Being Random, J. Daugman, Pattern Recognition, Vol. 36, No. 2, pp. 279—291, February 2003 (Важность быть случайным. Дж Даугман. Распознавание образов, Том 36, № 2, стр. 279-291, февраль 2003 г.)

[5] Constrained Monotone Regression of ROC Curves and Histograms Using Splines and Polynomials, T. Kanungo, D. M. Gay, and R. M. Haralick. In Proceedings of the International Conference on Image Processing, Washington D.C., October 23-26, 1995. Volume 2, IEEE Computer Society, 1995 (Ограниченная монотонная регрессия ROC-кривых и гистограмм, использующих сплайны и полиномы, Т. Кануго, Д.М. Гей, и Р.М. Хэрэлик. Труды международной конференции по обработке изображений, Вашингтон, округ Колумбия, 23-26 октября 1995 года, том 2, Компьютерное общество ИИЭР, 1995 г.)

[6] Monotone Piecewise Cubic Interpolation, Fred Fritsch and R Carlson, SIAM Journal on Numerical Analysis, Volume 17, Number 2, April 1980, pages 238-246 (Кусочно-монотонная кубическая интерполяция, Фред Фрич и Р. Карлсон, Журнал численного анализа SIAM, том 17, № 2, апрель 1980 г., стр. 238-246)

[7] Handbook of Multibiometrics, A. Ross, K. Nandakumar, and A. K. Jain, Springer International Series on Biometrics, Vol. 6, 2006, XXII, 202 p., 65 illus., Hardcover. ISBN: 0-387-22296-0 (Справочник мультбиометрии, А. Росс, К. Нандакумар и А.К Международная серия по биометрии издательства Springer, том 6, декабрь 2006 г., 202 стр., 65 илл., в твердом переплете. ISBN: 0-387-22296-0)

[8] When to Fuse Two Biometrics, E. Tabassi, G.W. Quinn, and P. Grother, IEEE CS Conference on Computer Vision and Pattern Recognition (CVPR '06), June 2006 (Когда объединять две биометрические характеристики, Е. Табассии, Г.У. Куинн, П. Грозер, Конференция ИИЭР по компьютерному зрению и распознаванию образов, июнь 2006 г.)

[9] ISO/IEC JTC1/SC 37 Standing Document 2 — Harmonized Biometric Vocabulary. This document has been developed within Working Group 1 of SC 37 (ИСО/МЭК СТК 1/ПК 37 Постоянный документ 2. Гармонизированный биометрический словарь. Этот документ был разработан рабочей группой 1 ПК 37)[1)]

[1)] Заменен на ISO/IEC 2382-37:2012. Однако для однозначного соблюдения требования настоящего стандарта, выраженного в датированной ссылке, рекомендуется использовать только указанное в этой ссылке издание.

УДК 004.93'1:006.354 ОКС 35.040 П85

Ключевые слова: информационные технологии, биометрия, мультибиометрические технологии, формат объединения данных, мультибиометрическое объединение, результат сравнения, функция распределения

БЗ 11—2016/79

Редактор *Е.С. Неверова*
Технический редактор *И.Е. Черепкова*
Корректор *Р.А. Ментова*
Компьютерная верстка *Е.О. Асташина*

Сдано в набор 26.06.2017. Подписано в печать 12.07.2017. Формат 60×84 1/8. Гарнитура Ариал.
Усл. печ. л. 3,72. Уч.-изд. л. 3,36. Тираж 20 экз. Зак. 1153.
Подготовлено на основе электронной версии, предоставленной разработчиком стандарта

Издано и отпечатано во ФГУП «СТАНДАРТИНФОРМ», 123001 Москва, Гранатный пер., 4.
gost.e-ecolog.ru gost.e-ecolog.ru

gost.e-ecolog.ru