
**ФЕДЕРАЛЬНОЕ АГЕНТСТВО
ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ**



**НАЦИОНАЛЬНЫЙ
СТАНДАРТ
РОССИЙСКОЙ
ФЕДЕРАЦИИ**

**ГОСТ Р ИСО/МЭК
29159-1-
201_**

**Информационные технологии
БИОМЕТРИЯ
Калибровка, аугментация и объединение
биометрических данных
Часть 1
Формат объединения информации**

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

(IDT)

**Москва
Стандартинформ
201_**

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) на основе собственного аутентичного перевода на русский язык 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 201_ г. № -ст

4 Настоящий стандарт идентичен международному стандарту ИСО/МЭК 29159-1:2010 «Информационные технологии. Данные биометрической калибровки, приращения и объединения. Часть 1. Формат объединения информации» (ISO/IEC 29159-1:2010 «Information technology – Biometric calibration, augmentation and fusion data – Part 1: Fusion information format»), за исключением приложения ДА.

Наименование настоящего стандарта изменено относительно наименования указанного международного стандарта для приведения в соответствие с ГОСТ Р 1.5 – 2012 (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 ГОСТ Р 1.0–2012 (раздел 8). 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.ru)

© Стандартиформ, 201_

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.....
2	Соответствие.....
3	Нормативные ссылки.....
4	Термины и определения.....
5	Обозначения и сокращения
6	Формат объединения информации.....
	6.1 Общие положения
	6.2 Порядок байтов.....
	6.3 Числовые значения
	6.4 Блок заголовка объединения
7	Общие элементы.....
	7.1 Общие положения.....
	7.2 Вид параметра.....
	7.3 Происхождение параметра.....
	7.4 Присутствующие распределения.....
	7.5 Число сравнений.....
	7.6 Признак предварительной нормализации данных
8	Запись типа 1.....
	8.1 Назначение
	8.2 Формат.....
	8.3 Применение (справочное).....
9	Запись типа 2.....
	9.1 Назначение
	9.2 Формат.....
	9.3 Применение (справочное).....
10	Запись типа 3.....
	10.1 Назначение
	10.2 Формат.....
Приложение А	(справочное) Обзор документа.....
Приложение В	(справочное) Примеры функций распределения.....
Приложение С	(справочное) Использование предварительной нормализации данных.....
Приложение D	(справочное) Исходный код для оценки сплайна.....
Приложение ДА	(справочное) Сведения о соответствии ссылочных международных стандартов национальным стандартам Российской Федерации.....
Библиография.....	

Введение

Разработка биометрических систем привела к появлению различных технологий и алгоритмов сравнения. Некоторые из них публикуются в открытом доступе, но большинство являются частной собственностью. В большинстве современных приложений для биометрической верификации или биометрической идентификации используется одна биометрическая модальность. Это означает, что используется одна какая-либо физическая или поведенческая характеристика человека с целью более или менее однозначной идентификации личности. Например, в системе контроля доступа может осуществляться регистрация изображения руки и использоваться ее геометрические характеристики. В рамках программы социальных льгот могут собираться отпечатки пальцев заявителей и использоваться 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 они обнаруживают 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 оценок схожести. Это объединение осуществляется после того как две или более системы обработали и сопостави-

ли биометрическую информацию индивида с одним или несколькими зарегистрированными образцами и предоставили скалярные результаты сравнения на выходе. Различают оценки схожести подлинного лица (тот же самый человек) и «самозванца» (другой человек). Схема объединения разрабатывается с целью комбинирования этих оценок схожести таким образом, чтобы уточнить границу классов оценок схожести 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 оценки схожести на выходе;

- в процессах, основанных на нормализации, объединению предшествует преобразование каждой оценки схожести 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 оценок схожести на входе. Формат объединения информации (ФОИ), 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 оценку схожести, которая подается в модуль объединения, инициализированного с помощью соответствующего экземпляра ФОИ.

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

В настоящем стандарте определены контейнеры для хранения информации о распределении оценок схожести, полученных от подсистемы сравнения. При этом не учитываются данные о совместном распределении, которые могут полностью получать статистические свойства многомерных оценок схожести (т.е. данные от двух и более подсистем или модальностей). Это означает, что мультимодальное объединение не поддерживается описанием совместных распределений биометрических оценок схожести. Это ограничение в большинстве случаев является незначительным, поскольку различные модальности часто считаются независимыми. Даже когда оценки схожести не являются независимыми, как и в случае мультиалгоритмических приложений, применение методов объединения на уровне степеней схожести 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, в большом 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 может включать в себя компоненты различной конструкции и производства, но по сути к ней не предъявляется требований обмена данными и функциональной совместимости с любой другой системой.

Применение стандартов форматов обмена биометрическими данными и стандартов биометрических интерфейсов является необходимым условием дос-

тижения полного обмена данными и функциональной совместимости в случае биометрического распознавания в среде открытых систем. Биометрические международные стандарты, разработанные в 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.

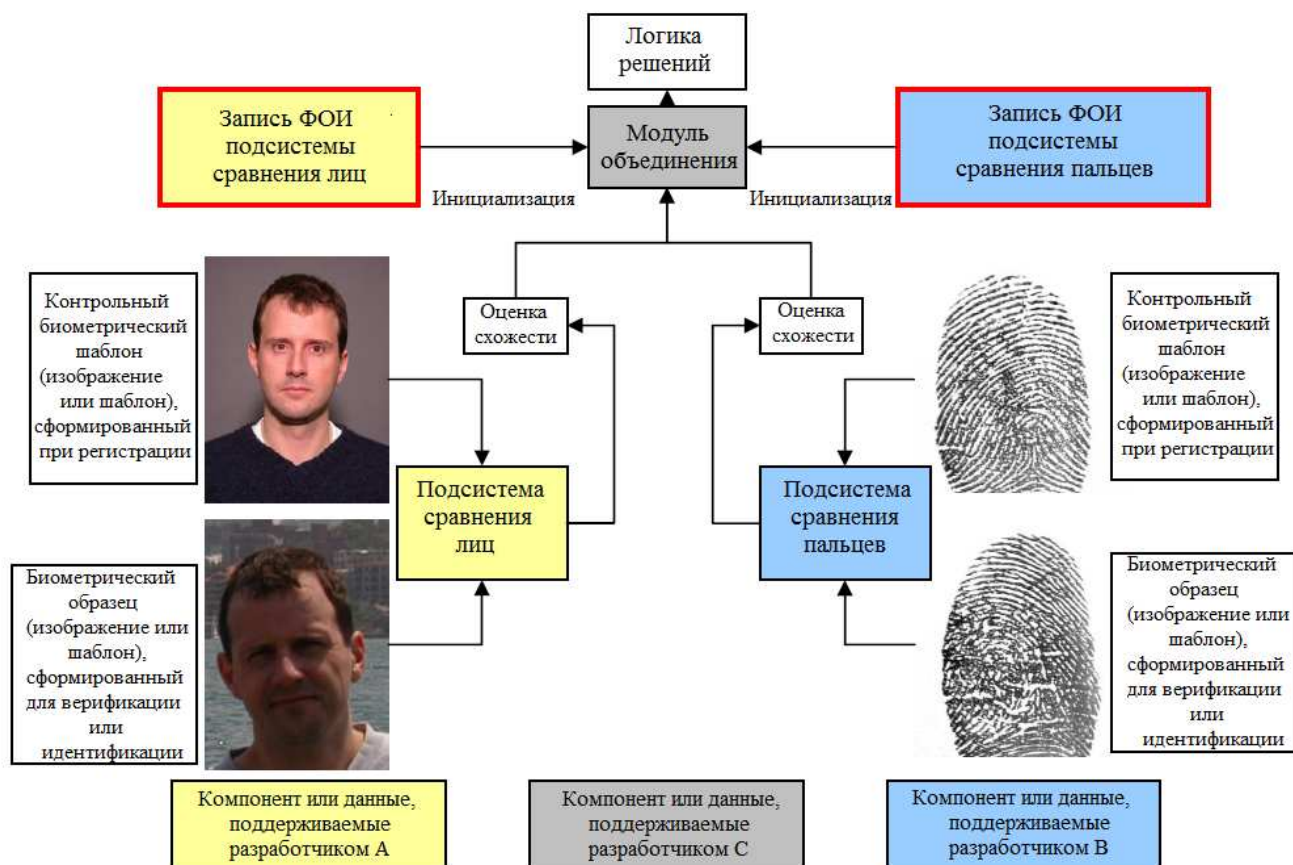


Рисунок 1 – Схема использования формата объединения информации

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

**Информационные технологии
БИОМЕТРИЯ****Калибровка, аугментация и объединение биометрических данных.****Часть 1. Формат объединения информации**

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

Дата введения – 201_–__–__

1 Область применения

Настоящий стандарт определяет формат объединения биометрической информации, который устанавливает машиночитываемые 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 определенным подразделам раздела 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.

ИИЭР 754 – 2008 Стандарт ИИЭР для арифметики с плавающей точкой (IEEE 754-2008, IEEE Standard for Floating-Point Arithmetic)

ИСО/МЭК 19785-1:2006 Информационные технологии. Единая структура формата обмена биометрическими данными. Часть 1. Спецификация элементов данных (ISO/IEC 19785-1:2006, Information technology – Common Biometric Exchange Formats Framework – Part 1: Data element specification)

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

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 выделению биометрических признаков.

Примечание – Устройство сбора биометрических данных это 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 оценка схожести подлинного лица (genuine score): Результат сравнения двух образцов от одного человека.

4.4 оценка схожести «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): Скалярная величина на выходе биометрической подсистемы сравнения.

Примечание – Термин «результат сравнения» применяется в настоящем стандарте как для результатов мер различия (меньшая указывает на большую вероятность того, что образцы получены от одного и того же человека), так и для результатов оценки схожести (большая указывает на большую вероятность того, что образцы получены от одного и того же человека). Эта разница подробно описана в пункте 6.4.9 о типах смысловых значений оценок схожести.

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» - что поле является 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);
ФОИ	– формат объединения информации (fusion information format, FIF);
ФПРВ	– функция плотности распределения вероятностей (probability density function, PDF);
ID	– идентификатор (identifier).

6 Формат объединения информации

6.1 Общие положения

6.1.1 Структура записи

Запись ФОИ используется для поддержки модульности в мультимодальных биометрических системах и системах поддержки принятия решений. Ее формат приведен в таблице 1.

П р и м е ч а н и е – Приложение должно установить профиль на основе настоящего стандарта. Профиль по умолчанию будет вызывать один из типов записей.

Т а б л и ц а 1 – Структура записи формата объединения информации

Блок «Заголовок объединения» (Fusion header block)	Запись типа 1 (Type 1 Record) (см. 6.1.3)	и/или ли	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 заголовка

Структура блока «Заголовок объединения» определяет формат записи и указывает на содержание. Ее формат приведен 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 биометрии (Biometric Type)				
подраздел 6.4		6.4.2	6.4.3	6.4.4	6.4.5				
25 байтов		4	4	4	3				
необработанное дополнение									
				Идентификатор продукта 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 оценки схожести (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 содержание. Ее формат приведен 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	Положение распределения	Масштаб
подраздел 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 содержание. Ее формат приведен в таблице 4.

Т а б л и ц а 4 – Структура записи типа 2

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

6.1.5 Структура записи типа 3

Структура «Заголовок объединения» определяет формат записи и указывает на содержание. Ее формат приведен в таблице 5.

Т а б л и ц а 5 – Структура записи «Тип 3»

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

6.2 Порядок байтов

В записи ФОИ и всех однозначно определенных блоках данных в ней все многобайтовые величины должны храниться в обратном порядке следо-

вания 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 настоящего стандарта, 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).

Т а б л и ц а 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.

П р и м е ч а н и е – 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 ФОИ.

Таблица 7 – Блок «Заголовок объединения»

Поле	Статус	Размер, байт	Допустимые значения	Примечание
Идентификатор формата	М	4	0x46494600	Строка ASCII «FIF» с нулевым терминатором
Номер версии стандарта	М	4	0x30313000	Строка ASCII «010» с нулевым терминатором
Длина записи	М	4	$1 \leq L \leq 2^{32}-1$	Длина входной записи в байтах
Тип биометрии	М	3	$0 \leq t \leq 0x080000$	Модальность, от которой получена запись
Идентификатор продукта подсистемы сравнения	М	4	≥ 0	Текущий идентификатор, определенный разработчиком
Идентификатор базы данных	М	2	≥ 0	Текущий идентификатор, определенный разработчиком
Качество базы данных регистрации	М	1	[0-100], 254, 255	Совокупные качества образцов, используемые для вычисления статистических данных результатов сравнения
Качество базы данных верификации	М	1	[0-100], 254, 255	
Смысловое значение оценки схожести	М	1	0 или 1	Различие или схожесть? См. 6.4.9
Число экземпляров типа	М	1	$1 \leq N \leq 4$	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 биометрии

Это значение (3 байта) должно быть взято из перечня ЕСФОБД биометрических модальностей, приведенного в ИСО/МЭК 19785-1:2006, 6.5.6. Это значение позволяет приложению определить, какая биометрическая модальность представлена, с помощью экземпляра формата объединения информации.

Пример: Для оценок схожести от реализации, использующей изображения венозного русла, это значение будет 0x040000.

6.4.6 Идентификатор продукта подсистемы сравнения

Идентификатор алгоритма подсистемы сравнения (т.е. сравнения биометрических данных с контрольными биометрическими шаблонами) (4 байта = 2 байта+2 байта), который позволяет получать информацию об оценках схожести, содержащуюся в данной записи ФОИ, должен быть записан в соответствии с требованиями, приведенными в таблице 8. Эти два значения яв-

ляются идентификаторами продукта ЕСФОБД, описанными в ИСО/МЭК 19785-1:2006.

Таблица 8 – Идентификаторы продуктов ЕСФОБД

	Кем определяется	Длина, байт	Описание
Идентификатор продукта (сохраненный в первых двух байтах)	Международная ассоциация биометрии и идентификации (International Biometrics & Identification Association, IBIA)	2	См. ИСО/МЭК 19785-1:2006
Номер версии стандарта (сохраненный во вторых двух байтах)	Разработчик	2	

Примечание 1 – Можно утверждать, что поле идентификатора продукта 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.

Примечание 2 – Одного идентификатора продукта, скорее всего, недостаточно для выполнения нормализации оценок схожести. 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 создания оценок схожести, которые применяются при вычислении данных, содержащихся в полях 4 и 5, должен быть записан в соответствии с требованиями, приведенными в таблице 9. Это поле может быть использовано:

– разработчиком подсистемы сравнения, чтобы произвольно указать, какие данные использовались при вычислении содержимого записи ФОИ;

– 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 использования их 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 набора данных	Описание	Модальность
0	Не определено	
1	Неизвестно	
2	Смешанно	
3-15	Зарезервировано	
16	FERET Face Database http://www.itl.nist.gov/iad/humanid/colorferet/	Лицо
17	Yale Face Database http://cvc.yale.edu/projects/yalefaces/yalefaces.html	Лицо
18	PIE Database CMU http://www.ri.cmu.edu/projects/project_418.html	Лицо
19	AR Database http://cobweb.ecn.purdue.edu/~aleix/aleix_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

Продолжение таблицы 9

Идентификатор набора данных	Описание	Модальность
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
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 http://fingerprint.nist.gov	Латентный палец
39	NIST Special Database 29 http://fingerprint.nist.gov	Палец
40	MCYT Fingerprint subcorpus http://atvs.ii.uam.es/bbdd_EN.html	Палец
41	MCYT Signature subcorpus http://atvs.ii.uam.es/bbdd_EN.html	Подпись
42	BANCA still http://www.ee.surrey.ac.uk/banca/	Лицо
43	BANCA video http://www.ee.surrey.ac.uk/banca/	Голос
44	BANCA high quality http://www.ee.surrey.ac.uk/banca/	Голос
45	BANCA low quality http://www.ee.surrey.ac.uk/banca/	Голос
46	NIST Speaker Verification http://www.nist.gov/speech/tests/spk/2005/	Голос
47	NIST Speaker Verification http://www.nist.gov/speech/tests/spk/2006/	Голос
48	CASIA Iris http://www.nlpr.ia.ac.cn/english/irids/irisdatabase.htm	Радужная оболочка глаза
49	Bath Iris http://www.irisbase.com	Радужная оболочка глаза

Окончание таблицы 9

Идентификатор набора данных	Описание	Модальность
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 http://fingerprint.nist.gov	Палец
54	NIST MINEX DOS calibration set http://fingerprint.nist.gov	Палец
55	NIST MINEX POEBVA calibration set http://fingerprint.nist.gov	Палец
56	TURBINE GUC100 http://www.nislabs.no/guc100	Палец
57 - 2047	Зарезервировано	
2048 - 65535	Определено разработчиком, возможно, не публичным и не уникальным среди разработчиков	

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

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

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

6.4.8 Качество базы данных

Качества 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

ние 0 представляет минимально возможное качество, а значение 100 – максимально возможное качество. В таблице 10 приведены допустимые значения для каждого поля.

Т а б л и ц а 1 0 – 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 оценки схожести

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

Т а б л и ц а 1 1 – Коды смыслового значения оценки схожести

Смысловое значение	Значение (1 байт)
Различие	0
Схожесть	1

П р и м е ч а н и е – 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 «ФОИ». 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.

Обратите внимание, 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	Производитель пренебрегает указанием величины (как правило, неприемлемо)
1	Неизвестно	Неизвестная величина
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	

Окончание таблицы 12

Вид	Описание	Примечание
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 $i = 1 \dots N$, которые используются для записей типа 2
97	Функция распределения	Как B-сплайн распределение, которое используется для записей типа 3
98 - 255	Не определено	

Примечание 1 – Приложение также может повлиять на то же самое преобразование независимо от вида значения, которое содержит запись, например, линейное

преобразуется в $(x-\text{mean})/\text{stdev}$ и $(x-\text{median})/\text{mad}$, где mean – среднее значение, stdev – стандартное отклонение, median – медиана, mad - медианное абсолютное отклонение.

Примечание 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

Это значение (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 (значения), чье происхождение не разглашается
1	Неизвестно - значение (значения) с неустановленным происхождением
2	Эмпирическое - значение (значения), оцененные на основе экспериментальных и выборочных данных
3	Известно априори - значение (значения), известные на этапе проектирования или исходя из теоретических соображений. Для распределения это означает, что оно известно в замкнутой форме
4-254	Зарезервировано

7.4 Присутствующие распределения

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

Таблица 14 – Информационное существование распределения

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

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.

7.6 Признак предварительной нормализации данных

Это значение (1 байт), приведенное в таблице 15, следует использовать, чтобы указать, существуют ли оценки схожести, 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 оценок схожести без необходимости интерполяции. В приложении С приведена информация о значении этого поля.

Примечание 1 – Это поле присутствует в записях типа 2 или 3 для распределений подлинных лиц и «самозванцев». Если подсистема сравнения представляет предварительно нормализованные оценки схожести «самозванцев», оценки схожести подлинных лиц будут неравномерны, и наоборот. Таким образом, это поле в записях настоящего стандарта не сможет указать на то, что оценки схожести и «самозванцев», и подлинных лиц предварительно нормализованы.

Примечание 2 – Реализации, совместимые с БиоАПИ [1], возвращают «значения ВЛД». Это означает, что полученные внутри системы оценки схожести «самозванцев» либо изначально распределены на отрезке [0,1], либо были нормализованы с помощью ожидаемой ФР «самозванцев» этих оценок схожестей. В результате можно рас-

считывать, что оценки схожести «самозванцев» на выходе будут равномерно распределены на отрезке [0,1].

Таблица 15 – Коды предварительной нормализации

Статус предварительной нормализации	Диапазон оценок схожести	Значение (1 байт)
Не нормализованы	$-\infty < x < \infty$	0
Предварительно нормализованы	$0 \leq x \leq 1$	1

8 Запись типа 1

8.1 Назначение

Запись типа 1 содержит минимальную статистическую информацию об оценках схожести «самозванцев» и/или подлинных лиц, полученных от биометрической системы. Эта информация может быть использована для масштабирования оценок схожести перед операцией объединения.

8.2 Формат

8.2.1 Поддерживаемые типы данных – подтип А

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

Таблица 16 – Формат подтипа А

Поле	Статус	Описание	Длина, байт	Тип данных	Допустимые значения	Пример
1	М	Вид параметра	1	uint (целое число без знака)	См. таблицу 12	3 (медиа-на)
2	М	Происхождение параметра	1	uint (целое число без знака)	0-3	2 (эмпирическое значение)
3	М	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	Допустимые значения	Пример
1	М	Тип	1	uint (целое число без знака)	1	1
2	М	Присутст- вующие распреде- ления	1	Упакован- ное битовое поле	0x01, 0x02, 0x03	0x01
3	О	Число сравнений «самозван- цев»	4	uint (целое число без знака)	$[0, 2^{32} - 1]$	40000
4		Положение распреде- ления «са- мозванцев»	10	Подтип А		(3; 1; 2,998)
5		Масштаб распреде- ления «са- мозванцев»	10	Подтип А		(34; 1; 0,308)
6	О	Число сравнений подлинных лиц	4	uint (целое число без знака)	$[0, 2^{32} - 1]$	240
7		Располо- жение рас- пределения подлинных лиц	10	Подтип А		(3; 1; 8,310)
8		Масштаб распреде- ления под- линных лиц	10	Подтип А		(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 оценок схожести «самозванцев», должно быть записано в поле «Число сравнений».

8.2.6 Положение распределения «самозванцев»

Поле «Подтип А», при наличии, должно содержать вид, происхождение и значение параметра положения распределения оценок схожести «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 А», при наличии, должно содержать вид, происхождение и значение параметра масштаба распределения оценок схожести «самозванцев».

8.2.8 Число сравнений подлинных лиц

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

8.2.9 Положение распределения подлинных лиц

Поле «Подтип А», при наличии, должно содержать вид, происхождение и значение параметра положения распределения оценок схожести 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 А», при наличии, должно содержать вид, происхождение и значение параметра масштаба распределения оценок схожести подлинных лиц.

8.3 Применение (справочно)

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

$$s = (a - a_{\text{mean}}) / a_{\text{sigma}} + (b - b_{\text{mean}}) / b_{\text{sigma}}$$

где a – необработанная оценка схожести, полученная от подсистемы сравнения отпечатков пальцев;

b - необработанная оценка схожести, полученная от подсистемы сравнения радужных оболочек глаз;

a_{mean} – оценка среднего значения оценок схожести «самозванцев», полученных от подсистемы сравнения А;

a_{sigma} - оценка стандартного отклонения оценок схожести «самозванцев», полученных подсистемы сравнения А;

b_{mean} - оценка среднего значения оценок схожести «самозванцев», полученных от подсистемы сравнения В;

b_{sigma} - оценка стандартного отклонения оценок схожести «самозванцев», полученных от подсистемы сравнения В.

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.

Примечание 1 – Примеры ФР приведены в приложении А.

Примечание 2 – Модули объединения, требующие наличия функций плотности распределения вероятностей (ФПРВ), могут оценить их путем численного дифференцирования ФР. Тип 2 изначально не включает в себя ФПРВ из-за необходимости указания интервала.

Примечание 3 – Если $F(x)$ обозначает ФР, то ее значение для оценки схожести «а», включает в себя поиск оценки схожести «а» в структуре данных двумерного массива подтипа В. Это может быть реализовано с помощью двоичного поиска первого вектора для обнаружения индекса i для интервала $x_{i-1} \leq a < x_i$ 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 В

Подтип В 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) $x_1 \leq x_i \leq x_N$ для $i = 1 \dots N$.

Формат должен соответствовать требованиям, приведенным таблице 18.

Таблица 18 – Формат подтипа В

Поле	Статус	Описание	Длина, байт	Тип данных	Допустимые значения	Пример
1	М	Вид параметра	1	uint (целое число без знака)	96	96 (ФР)
2	М	Происхождение параметра	1	uint (целое число без знака)	0-3	1
3	М	Предварительная нормализация	1	uint (целое число без знака)	0, 1	0

Окончание таблицы 18

Поле	Статус	Описание	Длина, байт	Тип данных	Допустимые значения	Пример
4	М	Число сравнений	4	uint (целое число без знака)	$[0, 2^{32} - 1]$	$1,5 \cdot 10^9$
5	М	Число элементов, N	4	uint (целое число без знака)	$0 \leq N \leq 2^{32} - 1$	800
6	М	N значений x	8N	double (действительное число двойной точности)	Диапазон действительных чисел двойной точности	(0,2; 0,4; 0,8)
7	М	N значений f(x)	8N	double (действительное число двойной точности)	Диапазон действительных чисел двойной точности	(7; -4; -1)

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 $(13 + 16N)$ байтов или $(24 + 32N)$ байтов).

Таблица 19 – Формат записи типа 2

По- ле	Ста- тус	Описание	Дли- на, байт	Тип дан- ных	Допусти- мые зна- чения	Пример
1	М	Тип	1	uint (целое число без знака)	2	2
2	М	Присутствую- щие распреде- ления	1	Упакован- ное битовое поле	0x01 - 0x03	0x01
3	О	Распределение «самозванцев»	11 + 16N	Подтип В		Требует- ся одно или оба
4	О	Распределение подлинных лиц	11 + 16N	Подтип В		

Примечание 1 – Значение N в полях «Длина» это число пар (x, F(x)).

Примечание 2 – Значения, содержащиеся в записи подтипа В 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 В», при наличии, должно содержать ФР оценок схожести «самозванцев». Поле должно содержать монотонно возрастающие данные.

9.2.6 Распределение подлинных лиц

Поле «Подтип В», при наличии, должно содержать ФР оценок схожести подлинных лиц. Поле должно содержать монотонно возрастающие данные.

9.3 Применение (справочное)

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

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

где a – необработанная оценка схожести, полученная от подсистемы сравнения отпечатков пальцев;

b – необработанная оценка схожести, полученная от подсистемы сравнения радужных оболочек глаз;

M'_A – ФПРВ оценок схожести подлинных лиц, вычисленная как численная производная ФР оценок схожести подлинных лиц, полученных от подсистемы сравнения А;

N'_a – ФПРВ оценок схожести «самозванцев», вычисленная как численная производная ФР оценок схожести «самозванцев», полученных от подсистемы сравнения А;

M'_B – ФПРВ оценок схожести подлинных лиц, вычисленная как численная производная ФР оценок схожести подлинных лиц, полученных от подсистемы сравнения В;

N'_B – ФПРВ оценок схожести «самозванцев», вычисленная как численная производная ФР оценок схожести «самозванцев», полученных от подсистемы сравнения В.

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

Примечание 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. подраздел 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 В-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. В-сплайн должен быть рассчитан в соответствии с алгоритмом, приведенным в [5]. Степень В-сплайна должна равняться 3 (т.е. быть кубической).

10.2 Формат

10.2.1 Поддерживаемые типы данных – подтип С

Подтип С является структурой для коэффициентов и вершин представления функции на основе В-сплайна. Сплайн вычисляется на основе n точек данных при условии, что $N \ll n$, где N - число вершин (значений оценок схожести). 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 С

Поле	Статус	Описание	Длина, байт	Тип данных	Допустимые значения	Пример
1	М	Вид параметра	1	uint (целое число без знака)	97	
2	М	Происхождение параметра	1	uint (целое число без знака)	0, 1, 2, 3	

Окончание таблицы 20

Поле	Статус	Описание	Длина, байт	Тип данных	Допустимые значения	Пример
3	М	Предварительная нормализация	1	uint (целое число без знака)	0, 1	0
4	М	Число сравнений, n	4	uint (целое число без знака)	$[0, 2^{32} - 1]$	$1,5 \cdot 10^9$
5	М	Степень сплайна, K	1	uint (целое число без знака)	$1 \leq K \leq 255$	3
6	М	Число вершин, N	4	uint (целое число без знака)	$1 \leq N \leq 2^{32} - 1$	100
7	М	Вершинные значения x	8N	double (действительное число двойной точности)	Диапазон действительных чисел двойной точности	
8	М	Коэффициенты	8C (см. примечание 2)	double (действительное число двойной точности)	Диапазон действительных чисел двойной точности	

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

Примечание 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 В 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.

Примечание 3 – «Допустимые значения» являются здесь общими. Когда используется экземпляр подтипа С, например, в записи типа 3 (см. раздел 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	Допустимые значения	Пример
1	М	Тип	1	uint (целое число без знака)	3	3
2	М	Присутствующие распределения	1	Упакованное битовое поле	0x01 - 0x03	0x01
3	О	Распределение «самозванцев»	12 + 8N + 8(N - 4)	Подтип С K=3		Требуется одно или оба
4	О	Распределение подлинных лиц	12 + 8N + 8(N - 4)	Подтип С K=3		

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

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

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

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

Примечание 5 – В-сплайн является более точным представлением функции, чем сглаживание N вершин, т.к. он вычисляется по n точкам при $n \gg N$.

Примечание 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 С», при наличии, должно содержать ФР оценок схожести «самозванцев». Поле должно содержать монотонно возрастающие данные

10.2.6 Распределение подлинных лиц

Поле «Подтип В», при наличии, должно содержать ФР оценок схожести подлинных лиц. Поле должно содержать монотонно возрастающие данные

Приложение А
(справочное)
Обзор документа

А.1 Обзор

Настоящий стандарт определяет запись формата объединения информации (ФОИ) для хранения статистических данных результатов сравнения, полученных от биометрических подсистем сравнения. В разделе 6 рассмотрена структура записи и определен общий заголовок ФОИ. Как показано в таблице А.1, запись может содержать три типа данных, которые определены в разделах 8, 9 и 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 форматов объединения информации

Тип	Уровень	Описание	Предусмотренное применение	Пример нормализации оценок схожести
1	Оценка схожести	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. $y = (x - m_1) / s_1,$ где m_1 и s_1 - среднее значение и стандартное отклонение распределения «самозванцев» соответственно
2	Оценка схожести	Эмпирические функции распределения (ЭФР) одного или более распределений подлинных лиц и «самозванцев»	Выполнение любой операции масштабирования на основе полной информации о распределении	Масштабирование путем использования распределения «самозванцев» $y = N_1(x),$ где $N_1(x)$ – ФР «самозванцев» без учета оценки вероятности ложных совпадений (ВЛС) подсистемы сравнения
3	Оценка схожести	В-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 оценок схожести. Альтернативой является использование типов 2 и 3, которые полностью кодируют распределения оценок схожести 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 эквивалентом. 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 ФОИ относится к использованию с помощью модуля объединения, например, записи типа 3 от разработчика А совместно с записью типа 2 от разработчика В. 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 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 информации 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 ФОИ не включает запись для статистических данных результатов сравнения, зависящих от измерений качества биометрических образцов [2]. Однако следует отметить, что модуль объединения, инициализируемый записями ФОИ, тем не менее, может использовать значения качества образцов на входе как часть каждой последующей операции объединения. Кроме того, значения качества могут быть использованы разработчиком в процессе начальной подготовки собственных записей ФОИ.

Приложение В

(справочное)

Примеры функций распределения

В.1 Обзор

На рисунке В.1 показаны две функции распределения типовых коммерческих биометрических устройств сопоставления. 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 эмпирическими функциями распределения (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 %-ой квантилью для удобства использования. Гладкие кривые представляют собой распределения сплайна типа 3.

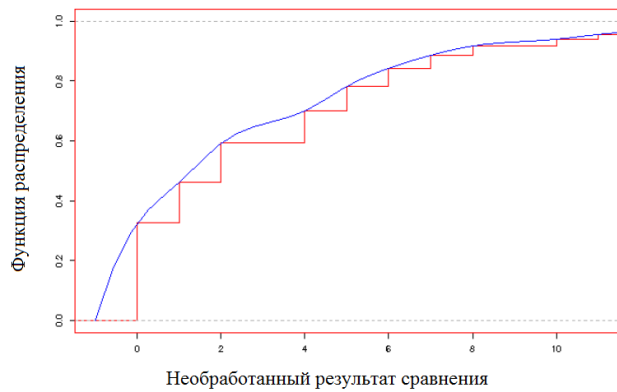
Использование сплайнов на основе записей ФОИ поддерживает оценку ФР при произвольных значениях, в том числе:

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

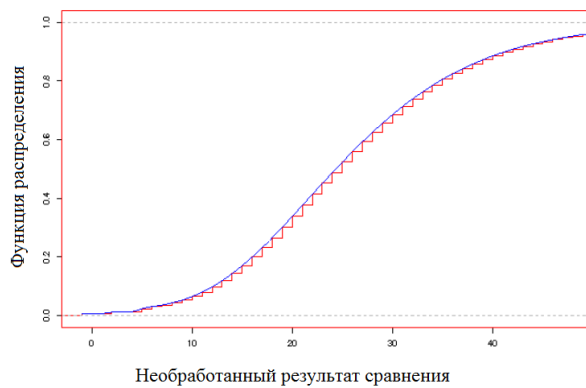
Но если подсистема сравнения изначально предоставляет только интегральные оценки схожести, то почему сплайн является актуальным? В качестве альтернативы может быть использована справочная таблица для типа 2, но она имеет недостаток, который заключается в ее большом размере и отсутствии всех возможных 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 оценок схожести, 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).



а)



б)

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

Приложение С

(справочное)

Использование предварительной нормализации данных

С.1 Введение

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

$$S_{\text{output}} = 1 - N_{\text{calibration}}(S_{\text{internal}}),$$

где S_{output} – оценка схожести на выходе;

S_{internal} – внутренняя оценка схожести;

$N_{\text{calibration}}$ - оценка ФР «самозванцев», выступающая в качестве функции калибровки.

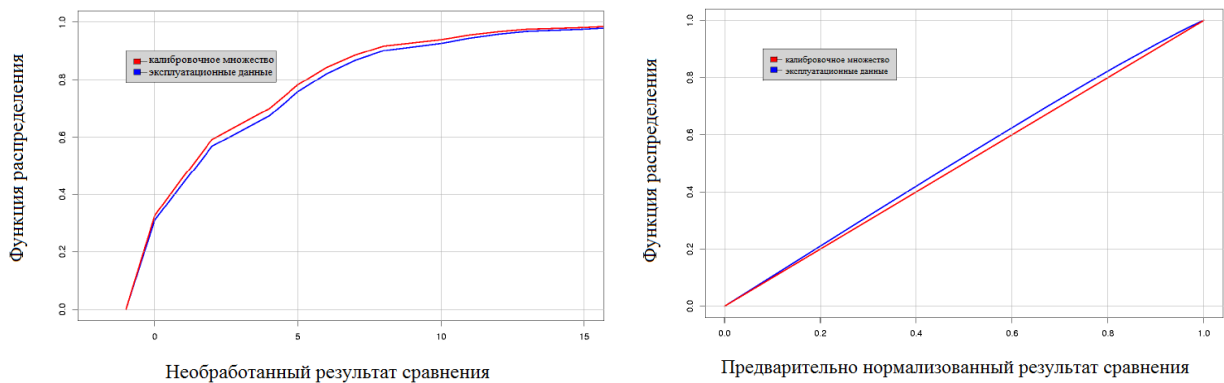
Распределение «самозванцев» используется таким образом, потому что считается, что оно не чувствительно к изменениям свойств выборки, таких как окружающая среда и популяция. Эта стабильность распределения «самозванцев» позволяет осуществлять нормализацию оценок схожести портативных кросс-приложений.

С.2 Пример

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

- ФР оценок схожести «самозванцев», полученных путем сравнения образцов из калибровочного множества изображений;
- ФР оценок схожести «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 оценок схожести является равномерным (по определению), но ФР эксплуатационных данных находится выше.



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

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 рассматриваются оценки схожести 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 оценок схожести «самозванцев» зависит от стабильности оценок схожести. ФОИ подходит к этому путем определения альтернативных форматов:

- для произвольно распределенных оценок схожести;
- для предварительно нормализованных оценок схожести, для которых оценки схожести «самозванцев» калибруются таким образом, чтобы быть равномерными на отрезке $[0, 1]$.

Если разработчик обеспечивает запись ФОИ, предоставляющую предварительно нормализованные оценки схожести, то под этим подразумевается, что в процессе функционирования равномерно распределенные оценки схожести «самозванцев» будут реализовываться в течение длительного периода эксплуатации. Если разработчик обеспечивает записи типа 2 или 3 с ФР «самозванцев», тогда заявленные оценки схожести «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 оценок схожести «самозванцев» и формирование записи ФОИ, 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, содер-
жащего 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 Российской Федерации

Таблица ДА.1

Обозначение ссылочного международного стандарта	Степень соответствия	Обозначение и наименование соответствующего национального стандарта
ИИЭР 754-2008		
ИСО/МЭК 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»
ИСО/МЭК 19794-1	IDT	ГОСТ Р ИСО/МЭК 19794-1—2015 «Информационные технологии. Биометрия. Форматы обмена биометрическими данными. Часть 1. Структура»
<p>Примечание – В настоящей таблице использовано следующее условное обозначение степени соответствия стандарта:</p> <p>- IDT – идентичный стандарт.</p>		

Библиография

- [1] ISO/IEC 19784-1 Information technology – Biometric application programming interface –Part 1: BioAPI Specification (ИСО/МЭК 19784-1 «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. <http://www.itl.nist.gov/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 (ИСО/МЭК ТО 24722:2007 «Информационные технологии. Биометрия. Мультимодальные и другие мульти биометрические 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)

,

Ключевые слова: информационные технологии, биометрия, мультибиометрические технологии, мультибиометрическое объединение, оценка схожести, результат сравнения, функция распределения

Организация–разработчик:

Научно–исследовательский и испытательный центр биометрической техники Московского государственного технического университета имени Н.Э. Баумана (НИИЦ БТ МГТУ им. Н.Э. Баумана).