ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

**ГОСТ Р 52633.2— 2010**

## Защита информации

# ТЕХНИКА ЗАЩИТЫ ИНФОРМАЦИИ

## Требования к формированию синтетических биометрических образов, предназначенных для тестирования средств высоконадежной биометрической аутентификации

Издание официальное

Москва
Стандартинформ
2011

## Предисловие

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

**Сведения о стандарте**

1 РАЗРАБОТАН Федеральным государственным унитарным предприятием «Пензенский научно-исследовательский электротехнический институт» (ФГУП «ПНИЭИ»), Федеральным государственным учреждением «Государственный научно-исследовательский испытательный институт проблем технической защиты информации Федеральной службы по техническому и экспортному контролю» (ФГУ «ГНИИИ ПТЗИ ФСТЭК России»)

2 ВНЕСЕН Управлением технического регулирования и стандартизации Федерального агентства по техническому регулированию и метрологии

3 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от 2 сентября 2010 г. № 215-ст

4 ВВЕДЕН ВПЕРВЫЕ

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

## Содержание

## Введение

Настоящий стандарт входит в комплекс стандартов ГОСТ Р 52633.X, устанавливающих требования к разработке и тестированию средств высоконадежной биометрической аутентификации, расширяет базовый стандарт ГОСТ Р 52633.0—2006, в большей части являющийся общим для двух стандартов (ГОСТ Р 52633.1, ГОСТ Р 52633.2).

Доверие к средствам высоконадежной биометрической аутентификации определяется результатами их тестирования, выраженными в форме гарантий производителя, подтвержденных при необходимости сертификационными документами.

Тестирование средств биометрической аутентификации проводится с использованием баз биометрических образов «Свой» и «Чужой», размеры которых являются достаточными для подтверждения характеристик тестируемых средств.

Используемые для достоверного тестирования размеры баз биометрических образов «Свой» малы, поэтому формирование таких баз легко осуществимо, а для образов «Чужой» велики ($10^{12}$ и больше образов). Соответственно, процесс создания баз естественных биометрических образов «Чужой» является крайне длительным и трудоемким. Создать базы такого размера в короткие сроки невозможно. В связи с этим, при тестировании приходится ограничиваться усеченными базами естественных образов «Чужой» размерами $10^3$ — $10^5$ образов, непосредственно полученными с тестируемого средства аутентификации. Дополнять такие базы можно естественными биометрическими образами «Чужой», ранее полученными при тестировании средств аутентификации с аналогичными биометрическими преобразователями. Кроме того, можно увеличивать размеры тестовой базы за счет применения искусственно синтезированных (синтетических) биометрических образов.

Настоящий стандарт устанавливает требования по формированию синтетических биометрических образов, которыми следует руководствоваться при пересчете биометрических образов, полученных при помощи первичных преобразователей того же типа, но с другими параметрами, и требования к размножению имеющихся естественных биометрических образов.

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

**Защита информации**

**ТЕХНИКА ЗАЩИТЫ ИНФОРМАЦИИ**

# Требования к формированию синтетических биометрических образов, предназначенных для тестирования средств высоконадежной биометрической аутентификации

Information protection.
Information protection technology.
Requirements for creation procedures for bases of synthetic biometric images, intended for high-reliability biometric authentication means testing

**Дата введения — 2010 — 10— 01**

## 1 Область применения

Настоящий стандарт распространяется на средства формирования баз синтетических биометрических образов или генераторы синтетических биометрических образов, а так же на процессы преобразования биометрических образов, полученных при помощи первичных преобразователей одного типа, но дающих другие по качеству данные. Настоящий стандарт применяется совместно с ГОСТ Р 52633 и ГОСТ Р 52633.1.

## 2 Нормативные ссылки

В настоящем стандарте использованы нормативные ссылки на следующие стандарты:

ГОСТ Р 52633.0 — 2006 Защита информации. Техника защиты информации. Требования к средствам высоконадежной биометрической аутентификации

ГОСТ Р 52633.1 — 2009 Защита информации. Техника защиты информации. Требования к формированию баз естественных биометрических образов, предназначенных для тестирования средств высоконадежной биометрической аутентификации

Примечание — При пользовании настоящим стандартом целесообразно проверить действие ссылочных стандартов в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет или по ежегодно издаваемому информационному указателю «Национальные стандарты», который опубликован по состоянию на 1 января текущего года, и по соответствующим ежемесячно издаваемым информационным указателям, опубликованным в текущем году. Если ссылочный документ заменен (изменен), то при пользовании настоящим стандартом следует руководствоваться заменяющим (измененным) стандартом. Если ссылочный стандарт отменен без замены, то положение, в котором дана ссылка на него, применяется в части, не затрагивающей эту ссылку.

## 3 Термины и определения

В настоящем стандарте применены следующие термины с соответствующими определениями.

3.1

**биометрические данные:** Данные с выходов первичных измерительных преобразователей физических величин, совокупность которых образует биометрический образ конкретного человека.

[ГОСТ Р 52633.0—2006, статья 3.5]

**Издание официальное**

3.2

**биометрический образ:** Образ человека, полученный с выходов первичных измерительных преобразователей физических величин, подвергающийся далее масштабированию и иной первичной обработке с целью извлечения из него контролируемых биометрических параметров человека.

Примечание — Биометрический образ — это континуум множества биометрических примеров, однако с конечной погрешностью континуум примеров может быть представлен всего несколькими различающимися примерами.

[ГОСТ Р 52633.0—2006, статья 3.6]

3.3

**биометрические параметры:** Параметры, полученные после предварительной обработки биометрических данных.

Примечание — Параметрами могут быть, например, коэффициенты Фурье кривых колебаний пера при воспроизведении человеком рукописного пароля.

[ГОСТ Р 52633.0—2006, статья 3.12]

3.4

**преобразователь «биометрия-код»:** Преобразователь, способный преобразовывать вектор нечетких, неоднозначных биометрических параметров «Свой» в четкий однозначный код ключа (пароля). Преобразователь, откликающийся случайным выходным кодом на воздействие случайного входного вектора, не принадлежащего множеству образов «Свой».

[ГОСТ Р 52633.0—2006, статья 3.19]

3.5

**естественный биометрический образ:** Биометрический образ донора, полученный в виде выходных биометрических данных первичного преобразователя и представленный одним или несколькими примерами.

[ГОСТ Р 52633.1—2009, статья 3.3]

3.6

**синтетический биометрический образ:** Биометрический образ, полученный путем имитационного моделирования естественных биометрических образов и представленный одним или несколькими примерами.

[ГОСТ Р 52633.1—2009, статья 3.4]

3.7

**база естественных биометрических образов «Чужой»:** Совокупность естественных биометрических образов, имитирующих предъявляемые средству аутентификации злоумышленником (злоумышленниками) случайные биометрические образы при реализации атаки подбора.

[ГОСТ Р 52633.1—2009, статья 3.5]

3.8

**база естественных биометрических образов «Свой»:** Совокупность естественных биометрических образов, состоящая из нескольких примеров одних и тех же биометрических образов, предназначенных для обучения или тестирования средств биометрической аутентификации.

[ГОСТ Р 52633.1—2009, статья 3.6]

3.9

**критерий Хемминга:** Мера сравнения двух кодов одинаковой длины, вычисляемая путем подсчета различающихся разрядов сравниваемых кодов.

[ГОСТ Р 52633.1—2009, статья 3.16]

3.10

**полная база естественных биометрических образов «Чужой»:** Совокупность биометрических образов «Чужой», содержащая достаточное число случайных естественных образов ($N_{полн}$) для достоверной оценки ожидаемой вероятности ошибки второго рода средства высоконадежной биометрической аутентификации, тестируемого прямым подбором.

[ГОСТ Р 52633.1—2009, статья 3.17]

3.11

**полная база естественных биометрических образов «Свой»:** Совокупность биометрических образов «Свой», содержащая их достаточное число для тестирования, и принадлежащих ко всем классам показателей стабильности, уникальности и качества биометрических параметров

[ГОСТ Р 52633.1—2009, статья 3.20]

3.12 **биометрический пример:** Совокупность биометрических данных, полученная с выхода первичного преобразователя при однократном предъявлении человеком своего биометрического образа.

3.13 **морфинг биометрических образов:** Создание промежуточного синтетического биометрического образа(ов), основанное на нахождении некоторого промежуточного значения каждого из биометрических параметров пары биометрических образов-родителей.

Примечание — Обычно морфингом размножается пара биометрических образов, но эту операцию можно повторять многократно для каждой пары биометрических образов, может быть использована любая пара образов «Свой» и «Чужой» либо пара образов «Чужой».

3.14 **морфинг биометрических примеров:** Создание промежуточного синтетического биометрического примера(ов), основанное на нахождении некоторого промежуточного значения каждого из биометрических параметров пары биометрических примеров-родителей.

Примечание — Биометрические примеры-родители могут принадлежать как одному биометрическому образу, так и различным биометрическим образам.

3.15 **биометрический образ-родитель:** Один из биометрических образов, используемый для создания биометрических образов-потомков в процессе морфинга.

3.16 **биометрический пример-родитель:** Биометрический пример, используемый для создания биометрических примеров-потомков в процессе морфинга примеров биометрических образов.

3.17 **биометрический образ-потомок:** Синтетический биометрический образ, создаваемый в процессе морфинга пары биометрических образов-родителей.

3.18 **биометрический пример-потомок:** Биометрический пример, создаваемый в процессе морфинга пары биометрических примеров-родителей.

3.19 **мутация биометрического образа:** Создание синтетического биометрического образа(ов), основанное на случайном изменении биометрических параметров биометрического образа-родителя.

3.20 **мутация биометрического примера:** Создание синтетического биометрического примера, основанное на случайном изменении биометрических параметров биометрического примера-родителя.

3.21 **кодовый центр биометрического образа:** Код, найденный путем выявления наиболее вероятного состояния каждого из разрядов выходных откликов преобразователя биометрия-код на предъявленные биометрические примеры биометрического образа.

## 4 Обозначения

v — значение биометрического параметра;

$h$ — значение расстояния Хэмминга;

$s$ — расстояние между двумя биометрическими параметрами;

$E(.)$ — оператор вычисления математического ожидания (ГОСТ Р 50779.10);

$\sigma(.)$ — оператор вычисления стандартного отклонения (ГОСТ Р 50779.10);

$P_1$ — вероятность ошибки первого рода (ошибочное непризнание «Своего»);

$P_2$ — вероятность ошибки второго рода (ошибочный допуск «Чужого»).

## 5 Общие положения

5.1 Для повышения качества разрабатываемых средств высоконадежной биометрической аутентификации необходимо проводить их тестирование. Для тестирования средств высоконадежной биометрической аутентификации необходимо сформировать базы биометрических образов, размеры которых должны гарантировать подтверждение заданных характеристик тестируемых средств. ГОСТ Р 52633.1 устанавливает требования к формированию баз естественных биометрических образов. Также этим стандартом определяется необходимость дополнения баз естественных биометрических образов синтетическими биометрическими образами.

5.2 Синтетические биометрические образы должны дополнять неполную базу естественных биометрических образов «Свой» или «Чужой» до полной базы, используя которую при тестировании можно будет с заданными достоверностями оценить вероятности ошибок первого и второго рода исследуемого преобразователя биометрия-код.

5.3 Дополнение баз естественных биометрических образов может производиться двумя вариантами.

1) С использованием заранее созданной базы синтетических биометрических образов. Недостатком этого варианта является необходимость хранения больших объемов синтезированной информации. Подобный способ необходимо использовать, если требуется высокая скорость работы с базой и возможность ее предварительного полного статистического исследования.

2) Через генерирование синтетических биометрических образов непосредственно во время тестирования средства биометрической аутентификации. Недостатком подобного варианта является необходимость дополнительных затрат вычислительных ресурсов в процессе тестирования и невозможность предварительного проведения статистических исследований. Подобный способ необходимо использовать, если требуется хранить меньший объем данных.

5.4 Расчет необходимого количества синтетических биометрических образов производится в соответствии с требованиями ГОСТ Р 52633.1. Пример расчета представлен в приложении А.

## 6 Использование естественных биометрических образов, ранее полученных с первичных преобразователей того же типа, но дающих другие по качеству данные

### 6.1 Общие положения

Для дополнения имеющихся баз естественных биометрических образов могут быть использованы естественные биометрические образы, ранее полученные при тестировании других средств биометрической аутентификации. При этом первичные преобразователи других средств биометрической аутентификации могут иметь другие технические характеристики (масштаб, разрешающая способность, шаг дискретизации, уровень шумов, формат представления данных).

### 6.2 Использование естественных биометрических образов, ранее полученных от других первичных преобразователей с теми же или более качественными данными

6.2.1 В случае если другие первичные преобразователи биометрического образа имеют те же самые или более качественные показатели разрешения считывания (погрешностей считывания), ранее собранные образы можно использовать при тестировании без их модификации. В данном случае тестирование производят в соответствии с блок-схемой, изображенной на рисунке 1.

6.2.2 Если естественные биометрические образы, собранные для другого первичного преобразователя, представлены в ином формате хранения, необходимо использовать конвертор форматов (блок *5* рисунка 1).

6.2.3 Критерием возможности использования данных других первичных преобразователей является совпадение статистик распределения каждого из контролируемых биометрических параметров дополняемых и дополняющих биометрических образов. Проверка совпадения осуществляется через предъявление двум сравниваемым первичным преобразователям одинаковых или близких биометрических образов. Близость биометрических образов в данном случае должна оцениваться экспертом.

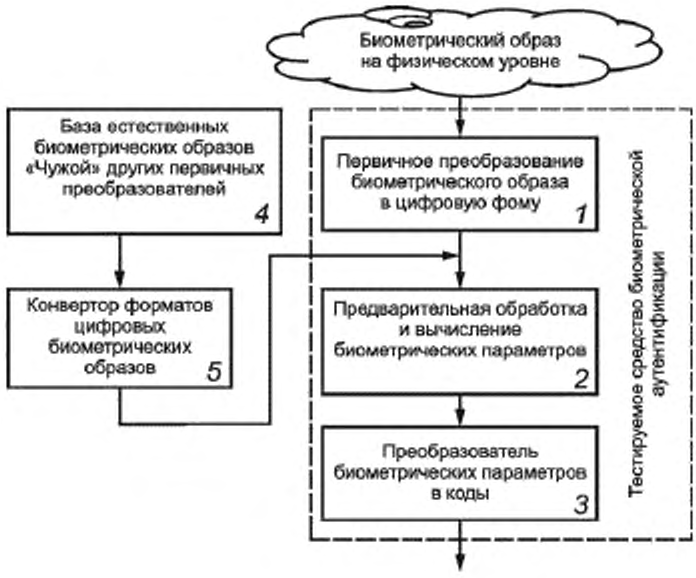


Рисунок 1 — Блок-схема тестирования с использованием естественных биометрических образов датчиков близкого типа с близкими характеристиками

**6.3 Использование частично-синтетических биометрических образов «Чужой», частично полученных от других первичных преобразователей с близкими по качеству данными**

6.3.1 Если разрешающая способность первичного преобразователя, для которого собраны дополнительные образы, мала (например недостаточны дискретизация пространства возможных состояний или частота опроса датчика), их данные нельзя использовать описанным в 6.2 образом.

6.3.2 Необходимо классифицировать получаемые биометрические параметры по их качеству. При этом часть контролируемых биометрических параметров, обладающих достаточным качеством, возможно применять для формирования примеров в дополняемой базе естественных биометрических образов. Другая часть параметров оказывается неприменима из-за их плохого качества. Данные параметры должны быть улучшены и добавлены к вектору естественных параметров. Улучшение качества может быть осуществлено через:

- сглаживание данных и их повторную дискретизацию;
- удаление грубых ошибок и наибольших отклонений;
- синтез данных, воспроизводящих статистику более высокого качества (синтез производится в соответствии с алгоритмами 7.2);
- замещение данных данными из других примеров достаточного качества.

6.3.3 Использование полученных частично-синтетических образов для дополнения базы естественных биометрических образов производится в соответствии с 6.2.

***Примеры***

***1 При анализе динамики рукописного слова недостаточная разрешающая способность первичного измерительного преобразователя по чувствительному полю экрана карманного компьютера приводит к тому, что старшие коэффициенты Фурье вычисляются со значительными погрешностями. Искажается статистика распределений значений образов «Чужой» по параметрам, которые вычисляются через расчет старших коэффициентов Фурье — старшие коэффициенты занижаются. В связи с этим, необходимо заменить естественные параметры с искаженными статистиками на искусственные параметры, повторяющие статистику естественных параметров более высокого качества, например, увеличив значения старших коэффициентов. В итоге формируется база векторов***

***биометрических образов, каждый из которых имеет часть естественных параметров и дополняющую их часть искусственно улучшенных параметров.***

***2 Собранные ранее образы других первичных преобразователей имеют требуемое разрешение по динамике проекций кривых колебаний пера на ортогональные оси X(t), Y(t), но у этих преобразователей отсутствует канал контроля давления пера на поверхность — P(t). В данном случае необходимо синтезировать отсутствующие данные и тем самым получать образы «Чужой», частично являющиеся естественными и частично являющиеся синтетическими.***

## 7 Размножение биометрических примеров одного биометрического образа

### 7.1 Общие положения

7.1.1 Размножение биометрических примеров одного биометрического образа может производиться с помощью синтеза случайных биометрических примеров с сохранением статистических характеристик биометрического образа, частичной мутации имеющихся биометрических примеров, морфинга между биометрическими примерами одного биометрического образа, перестановки фрагментов биометрических примеров.

7.1.2 Возможно сочетание перечисленных методов, например морфинг между биометрическими примерами одного биометрического образа с последующими мутациями полученного примера или синтез двух случайных биометрических примеров с последующим морфингом между ними и мутациями полученного биометрического примера.

### 7.2 Синтез случайных биометрических примеров одного биометрического образа

7.2.1 Синтез биометрических примеров производят по следующему алгоритму.

7.2.1.1 По всем имеющимся примерам образа вычисляются математические ожидания каждого *i*-го контролируемого биометрического параметра $E_{образ}(\nu_i)$.

7.2.1.2 Вычисляется стандартное отклонение каждого из контролируемых биометрических параметров $\sigma_{образ}(\nu_i)$.

7.2.1.3 Генератором нормальных случайных чисел со статистическими характеристиками, определенными в 7.2.1.1 — 7.2.1.2, формируются векторы параметров заранее заданного числа биометрических примеров.

Примечание — Синтез случайных биометрических примеров одного образа не учитывает корреляционные связи между параметрами биометрических примеров. Его использование приводит к нарушению естественных корреляционных связей параметров размножаемого биометрического образа. Подобный способ применим в тех случаях, когда корреляционные связи параметров синтетических биометрических образов не учитываются средством биометрической аутентификации.

### 7.3 Мутации биометрических примеров

7.3.1 Мутацию биометрического примера производят в соответствии со следующим алгоритмом.

7.3.1.1 По всем имеющимся примерам вычисляют математические ожидания каждого *i*-го контролируемого биометрического параметра $E_{образ}(\nu_i)$.

7.3.1.2 Вычисляют стандартное отклонение каждого из контролируемых биометрических параметров $\sigma_{образ}(\nu_i)$.

7.3.1.3 Каждый параметр $\nu_i$ исходного (мутируемого) биометрического примера изменяется на величину $\Delta\nu_i$, полученную от случайного генератора с нормальным законом распределения значений с математическим ожиданием $\nu_i$ и стандартным отклонением, равным трети $\Delta\nu_{i\,max}$, где $\Delta\nu_{i\,max}$ — расстояние до ближайшей границы $E_{образ}(\nu_i) \pm 3\sigma_{образ}(\nu_i)$ от значения $\nu_i$.

Графическая интерпретация мутации одного параметра биометрического примера приведена на рисунке 2.

Примечание — Мутации биометрических примеров ослабляют естественные корреляционные связи между параметрами биометрических примеров. Подобный способ применим в тех случаях, когда корреляционные связи параметров синтетических биометрических образов учитываются слабо или не учитываются средством биометрической аутентификации.

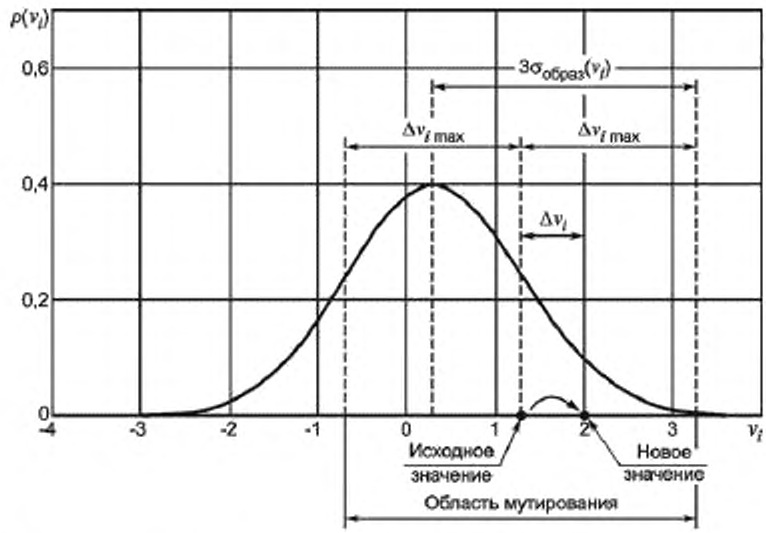


Рисунок 2 — Мутация параметра $v_i$ при синтезе биометрического примера

**7.4 Морфинг биометрических примеров одного биометрического образа**

**7.4.1 Общие положения**

7.4.1.1 Морфинг биометрических примеров одного биометрического образа заключается в нахождении промежуточных значений (биометрических примеров-потомков) для каждого из параметров пары биометрических примеров-родителей.

7.4.1.2 Количество потомков для каждой пары родителей зависит от заранее заданного количества синтетических биометрических образов и расстояния между биометрическими примерами-родителями.

7.4.1.3 В случае если исходные биометрические примеры имеют различный набор параметров либо различный смысл параметров, такие биометрические примеры нужно преобразовать таким образом, чтобы они имели одинаковый набор параметров. Если такое преобразование невозможно, такие биометрические примеры нельзя использовать при морфинге.

7.4.1.4 В случае если исходные биометрические примеры получены на первичных преобразователях, дающих другие по качеству данные, необходимо предварительно провести процедуры, описанные в разделе 6.

**7.4.2 Оценка количества биометрических примеров образов-потомков для каждой пары биометрических примеров-родителей**

7.4.2.1 Оценку количества биометрических примеров-потомков для каждой пары биометрических примеров-родителей производят в соответствии со следующим алгоритмом.

7.4.2.2 Выбирают $N$ пар биометрических примеров-родителей ($N > 2$), участвующих в морфинге.

7.4.2.3 Необходимое число синтетических биометрических примеров $N_{синт} > N$ рассчитывают в соответствии с разделом 8 ГОСТ Р 52633.1.

7.4.2.4 Для каждой пары $A$, $B$ биометрических примеров-родителей рассчитывают расстояние между их биометрическими параметрами

$$s_{AB} = \frac{1}{n}\sum_{i=1}^{n}|v_{i,A} - v_{i,B}|, \qquad (1)$$

где $v_{i,A}$ — $i$-й параметр биометрического примера $A$ и $v_{i,B}$ — $i$-й параметр биометрического примера B;

$n$ — общее число параметров биометрического примера.

7.4.2.5 По множеству пар A, B биометрических примеров-родителей рассчитывают среднее значение расстояния между их биометрическими параметрами $E(s)$.

7.4.2.6 Выбираемое количество потомков $k$ для пары биометрических примеров-родителей A, B рассчитывают по формуле

$$k_{AB} = \frac{2N_{сумм}}{N} P(s_{AB}), \quad (2)$$

где $P(s_{AB})$ — вероятность появления расстояния между параметрами $s_{AB}$ из множества всех возможных расстояний между параметрами.

При этом значение $k_{AB}$ округляют до ближайшего целого числа.

Примечание — Вероятность $P(s_{AB})$ может быть аппроксимирована нормальным законом распределения либо иным законом распределения. При этом количество потомков для минимального значения $s_{AB}$ нулевое, а для максимального значения $s_{AB}$ — удваивается по сравнению со средним числом потомков.

**7.4.3 Морфинг конкретной пары биометрических примеров-родителей**

7.4.3.1 Морфинг пары биометрических примеров-родителей A, B производят с помощью линейной интерполяции параметров биометрических примеров-родителей.

7.4.3.2 Морфинг пары биометрических примеров-родителей A, B производят по следующему правилу. Значения каждого $i$-го биометрического параметра каждого из биометрических примеров биометрических образов-потомков вычисляют по формуле:

$$v_{i\,j} = \frac{(k_{AB} + 1) - j}{k_{AB} + 1} v_{i,A} + \frac{j}{k_{AB} + 1} v_{i,B}, \quad (3)$$

где $j$ — порядковый номер потомка ($j = 1,2, \ldots, k_{AB}$);

$k_{AB}$ — рассчитанное количество потомков примеров-родителей A и B.

Примечание — Морфинг биометрических примеров сохраняет естественные корреляционные связи, присутствующие у биометрических примеров-родителей. Подобный способ применим в тех случаях, когда корреляционные связи параметров синтетических биометрических образов учитываются средством биометрической аутентификации.

**7.5 Размножение биометрических примеров перестановкой фрагментов**

7.5.1 Для размножения биометрических примеров перестановкой фрагментов необходимо разделить исходные биометрические примеры на фрагменты. Возможны два варианта такого разделения.

1) Использование естественной фрагментации исходных биометрических примеров (таких, как примеры голосовых образов, легко разделяемые на фонемы, или примеры рукописных образов, легко разделяемые на отдельные символы, буквы).

2) Использование фрагментации биометрических примеров, заложенной производителем средства при вычислении биометрических параметров, и связанной с тем, что каждый из множества полученных параметров сам по себе является некоторым фрагментом исходного примера.

7.5.2 Для примеров, легко разделимых на естественные фрагменты с номерами 1, 2, 3..., синтетический биометрический пример составляется из той же последовательности фрагментов с номерами 1, 2, 3..., принадлежащих случайно выбранным исходным биометрическим примерам.

7.5.3 При невозможности разделения исходных биометрических примеров на естественные фрагменты, биометрические параметры исходных примеров случайным образом разделяются на фрагменты с номерами 1, 2, 3... . Синтетический биометрический пример составляется из той же последовательности фрагментов 1, 2, 3... биометрических параметров, принадлежащих случайно выбранным исходным биометрическим примерам.

7.5.4 При размножении $N$ исходных биометрических примеров, как при использовании естественной фрагментации, так и при фрагментации биометрических параметров, необходимо для формирования каждого синтетического биометрического примера использовать приблизительно $1/N$ фрагментов каждого из исходных биометрических примеров. Пример размножения биометрических примеров представлен в приложении В.

## 8 Размножение биометрических образов

**8.1 Общие положения**

8.1.1 Размножение биометрических образов производят с помощью переноса биометрических образов, морфинга пары биометрических образов-родителей, перестановки фрагментов биометрических образов.

8.1.2 Возможно сочетание перечисленных методов, например, морфинг пары биометрических образов-родителей с последующим многократным случайным переносом полученного биометрического образа-потомка.

**8.2 Перенос биометрического образа**

8.2.1 Случайный перенос биометрического образа заключается в смещении всех параметров каждого биометрического примера исходного биометрического образа на заданную для каждого параметра константу.

8.2.2 Перенос одного биометрического примера исходного биометрического образа производят по следующему алгоритму.

8.2.2.1 Вычисляют математическое ожидание $E_{все}(\nu_i)$ и стандартное отклонение $\sigma_{все}(\nu_i)$ дополняемой базы биометрических образов «Все образы».

8.2.2.2 Вычисляют математическое ожидание $E_{образ}(\nu_i)$ каждого параметра для всех примеров переносимого биометрического образа.

8.2.2.3 Для всех примеров переносимого биометрического образа вычисляют отклонение каждого параметра $\Delta_i$ от математического ожидания $E_{образ}(\nu)$.

8.2.2.4 Генератором случайных данных с нормальным распределением генерируют константы переноса $\Delta\nu_i$ для каждого $i$-го параметра каждого примера переносимого образа. При этом $|\Delta\nu_i| < \Delta\nu_{i\,max}$, где $\Delta\nu_{i\,max}$ — расстояние до ближайшей границы $E_{образ}(\nu_i) \pm 3\sigma_{все}(\nu_i)$ от значения $\nu_i$. Математическое ожидание генератора данных принимается равным $E_{образ}(\nu_i)$, стандартное отклонение принимается равным одной трети $\Delta\nu_{i\,max}$.

8.2.2.5 Синтезируется множество биометрических примеров со значением $i$-го параметра $E_{образ}(\nu_i) + \Delta\nu_i + \Delta_i$. Графическая интерпретация переноса биометрического образа приведена на рисунке 3.

Примечание — Перенос биометрического образа сохраняет естественные корреляционные связи между параметрами биометрических примеров образа.

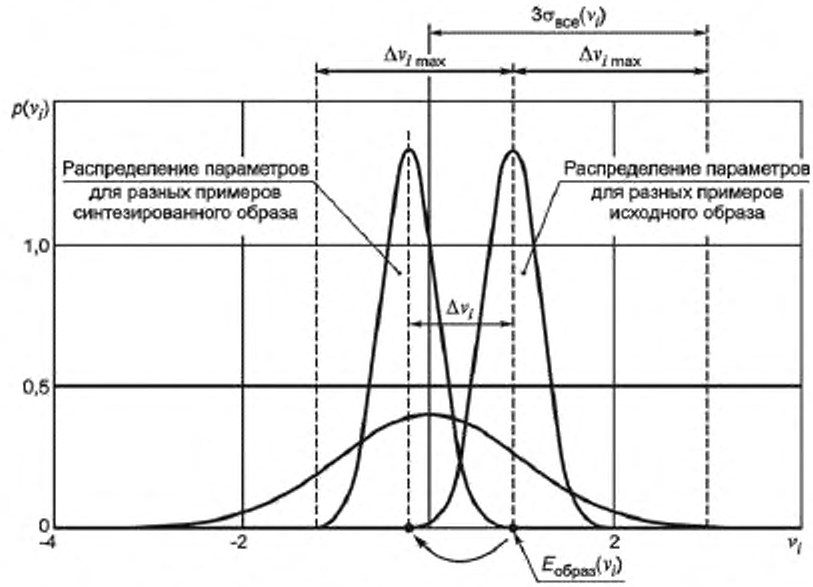


Рисунок 3 — Перенос примеров биометрического образа на константу $\Delta\nu_i$

**8.3 Морфинг двух биометрических образов-родителей**

8.3.1 Морфинг биометрических образов заключается в нахождении промежуточных значений (биометрических образов-потомков) для каждого из параметров двух биометрических образов-родителей.

Количество потомков для каждой пары родителей зависит от необходимого количества синтетических биометрических образов и расстояния между биометрическими образами-родителями, определяемого через расстояние Хэмминга между проявленными центрами биометрических образов-родителей.

**8.3.2 Проявление кодовых центров биометрических образов-родителей «Чужой»**

8.3.2.1 Проявление кодовых центров биометрических образов-родителей «Чужой» производят в соответствии со следующим алгоритмом.

8.3.2.2 При наличии менее чем 100 примеров каждого из биометрических образов-родителей «Чужой» необходимо увеличить количество биометрических примеров до 100 или более за счет добавления дополнительных синтетических биометрических примеров, полученных любым из описанных в разделе 7 алгоритмов.

8.3.2.3 Для каждого примера каждого из образов-родителей «Чужой» находят кодовый отклик преобразователя биометрия-код.

8.3.2.4 Если количество разрядов со значением «0» превышает количество разрядов со значением «1», соответствующему разряду проявленного кодового центра присваивают значение «0».

8.3.2.5 Если количество разрядов со значением «1» превышает количество разрядов со значением «0», соответствующему разряду проявленного кодового центра присваивают значение «1».

8.3.2.6 Стабильность проявленного разряда тем больше, чем больше превышение одних значений разрядов над другими. В случае отсутствия превышения одних значений над другими проявленный разряд считается нестабильным. Значение нестабильного разряда устанавливается по значению соответствующего разряда у последнего рассмотренного при проявке биометрического примера.

Пример проявления кодового центра биометрического образа «Чужой» представлен в приложении Б.

**8.3.3 Оценка количества биометрических образов-потомков для каждой пары биометрических образов-родителей**

8.3.3.1 Количество биометрических образов-потомков для каждой пары биометрических образов-родителей зависит от необходимого числа синтетических образов и расстояния Хэмминга между проявленными в соответствии с пунктом 8.3.2 кодовыми центрами биометрических образов-родителей.

8.3.3.2 Количество биометрических образов-потомков для каждой пары биометрических образов-родителей должно выбираться так, чтобы равномерно заполнить базу биометрических образов.

8.3.3.3 Оценку количества биометрических образов-потомков для каждой пары биометрических образов-родителей производят в соответствии со следующим алгоритмом.

8.3.3.4 Выбирают $N$ пар ($N > 2$) биометрических примеров-родителей, участвующих в морфинге.

8.3.3.5 Необходимое число синтетических биометрических примеров $N_{синт} > N$ рассчитывают в соответствии с разделом 8 ГОСТ Р 52633.1.

8.3.3.6 Для каждого биометрического образа-родителя в соответствии с пунктом 8.3.2. находят кодовый центр.

8.3.3.7 Для каждой пары биометрических образов-родителей (например, A, B) находят расстояние Хэмминга между кодовыми центрами $h_{AB}$ (см. рисунок 4).

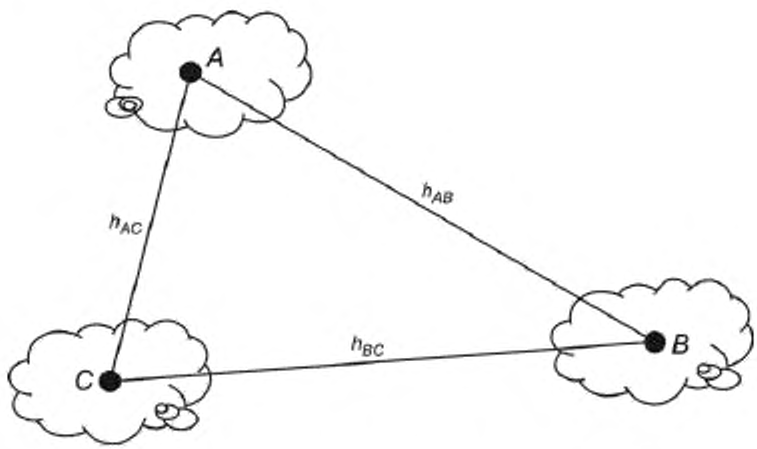


Рисунок 4 — Расстояния Хэмминга между кодовыми центрами биометрических образов A, B, C

8.3.3.8 По всем $N$ парам биометрических примеров-родителей рассчитывают среднее значение расстояния Хэмминга между кодовыми центрами биометрических образов-родителей $E(h)$.

8.3.3.9 Оптимальное количество потомков k для конкретной пары биометрических образов-родителей $A$, $B$ рассчитывают по формуле

$$k_{AB} = \frac{2N_{синт}}{N} P(h_{AB}), \quad (4)$$

где $P(h_{AB})$ — вероятность появления расстояния $h_{AB}$ между кодовыми центрами образов A, B.

При этом значение $k_{AB}$ округляют до ближайшего целого числа.

Примечание — Вероятность $P(h_{AB})$ может быть аппроксимирована нормальным законом распределения, либо иным законом распределения. При этом количество потомков для минимального значения $h_{AB}$ нулевое, а для максимального значения $h_{AB}$ — удваивается по сравнению со средним числом потомков.

**8.3.4 Морфинг двух биометрических образов-родителей**

8.3.4.1 Морфинг двух биометрических образов-родителей производят через морфинг любых пар биометрических примеров биометрических образов-родителей (см. рисунок 5).

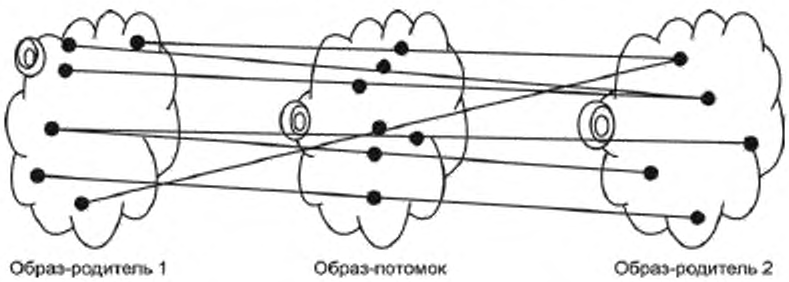


Рисунок 5 — Морфинг пары биометрических образов и получение 1 образа-потомка

Пример морфинга пары примеров различных рукописных образов и получения 10 образов-потомков приведен на рисунке 6.

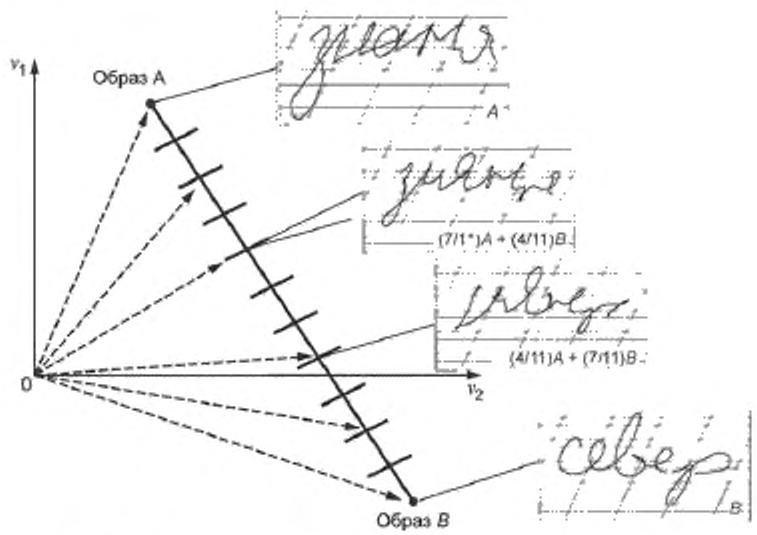


Рисунок 6 — Пример морфинга пары примеров различных биометрических рукописных образов и получение 10 образов-потомков (изображены два параметра $v_1$, $v_2$ из 416)

8.3.4.2 Возможно дальнейшее размножение полученных биометрических примеров биометрического образа-потомка любыми способами раздела 7 (например, мутациями или морфингом между примерами биометрического образа-потомка).

Примечание — Морфинг биометрических образов сохраняет естественные корреляционные связи между параметрами биометрических примеров.

**8.4 Размножение биометрических образов перестановкой фрагментов**

8.4.1 При размножении биометрических образов перестановкой фрагментов необходимо разделить примеры исходных биометрических образов на фрагменты. Варианты подобного разделения описаны в 7.5.1. Фрагменты последовательно нумеруются.

8.4.2 Выбирается случайная комбинация номеров фрагментов образов, в которой допускается повторение одинаковых номеров. Все примеры синтетического образа формируются в соответствии с выбранной последовательностью фрагментов. Каждый из примеров синтетического образа получают, извлекая соответствующие фрагменты из различных примеров исходных образов.

8.4.3 Примеры синтетического биометрического образа могут состоять из большего или меньшего по сравнению с исходными биометрическими образами числа фрагментов. Исключение, добавление, изменение последовательности фрагментов при формировании примеров синтетических биометрических образов приводят к появлению других синтетических биометрических образов. Размножение биометрических образов перестановкой фрагментов представлено в приложении Г.

## 9 Выявление и заполнение слабо заполненных фрагментов выходного поля кодовых откликов базы тестовых биометрических образов

**9.1 Выявление слабо заполненных фрагментов выходного поля кодовых откликов базы тестовых биометрических образов**

9.1.1 После формирования базы синтетических биометрических образов или в процессе ее формирования необходимо контролировать равномерность заполнения поля выходных кодов кодовыми откликами синтетических биометрических образов.

9.1.2 Контроль заполнения базы биометрических образов состоит в следующем.

9.1.3 Выбирают не менее $N = 1000$ случайных биометрических образов из базы биометрических образов.

9.1.4 Для каждого разряда каждого образа вычисляется вероятность появления значений «1» и «0». В случае отклонения вероятности появления значений «1» или «0» от значения 0,5 более чем на $\frac{3}{2\sqrt{N}}$ (где $N$ — количество случайных биометрических образов в выборке) должно производится дополнение выявленного слабо заполненного фрагмента выходного кодового поля синтетическими биометрическими образами.

**9.2 Дополнение выявленных слабо заполненных фрагментов выходного поля кодовых откликов базы тестовых биометрических образов**

9.2.1 В случае обнаружения повышенной вероятности появления значений «1» в контролируемом разряде, необходимо произвести размножение биометрических образов базы, кодовые центры которых в заданном разряде имеют значение «0». При этом размножаются биометрические образы, имеющие в контролируемом разряде наибольшую стабильность. Размножение ведут при помощи морфинга, описанного в 8.3.

9.2.2 В случае обнаружения повышенной вероятности появления значений «0» в данном разряде размножают биометрические образы, имеющие в контролируемом разряде наиболее стабильное значение «1».

9.2.3 Размножение производится до тех пор, пока описанным в 9.1 алгоритмом будут определяться слабо заполненные фрагменты выходного кодового поля.

**Приложение А**
**(справочное)**

## Расчет необходимого количества синтетических биометрических образов

**А.1 Оценка необходимого числа примеров синтетических биометрических образов «Свой»**

Предположим, что требуется тестировать вероятность ошибок первого рода $P_1$ = 0,01, заявленную некоторым производителем средства биометрической аутентификации. При этом мы имеем естественный биометрический образ «Свой», представленный $N_{ест}$ = 30 примерами. В соответствии с требованиями ГОСТ Р 52633.1 имеющиеся естественные биометрические примеры необходимо дополнять искусственно синтезированными примерами количеством:

$$N_{синт} = \frac{10}{P_1} - N_{ест} = 970 \text{ образов.} \quad (А.1)$$

Для размножения примеров биометрического образа «Свой» можно применять любые изложенные в стандарте способы. В случае использования морфинга количество потомков каждой пары рассчитывают следующим образом. Существующие 30 примеров могут образовать максимум (30×29)/2 = 435 пар. Из данного количества нужно отобрать 245 наиболее различающихся пар, каждая из которых в среднем должна давать по 970/245 = 4 потомка.

**А.2 Оценка необходимого числа синтетических биометрических образов «Чужой»**

Предположим, что требуется тестировать вероятность ошибок второго рода $P_2 = 10^{-6}$, заявленную некоторым производителем средства биометрической аутентификации. При этом мы имеем базу естественных биометрических образов «Чужой», состоящую из $N_{ест}$ = 10 000 образов. В соответствии с требованиями ГОСТ Р 52633.1 имеющиеся естественные биометрические образы необходимо дополнять искусственно синтезированными образами количеством:

$$N_{синт} = \frac{10}{P_2} - N_{ест} \approx 9{,}99 \cdot 10^6 \text{ образов.} \quad (А.2)$$

Для размножения биометрических образов «Чужой» можно применять любые изложенные в стандарте способы. В случае использования морфинга биометрических образов количество пар биометрических образов-родителей выбирается в зависимости от степени заполненности исходной базы естественных биометрических образов, а количество потомков для каждой пары определяется требованиями 8.3. Например, можно выбрать 6000 образов-родителей, составляющих максимально различные пары, и для каждой пары биометрических образов-родителей генерировать одного потомка:

$$N = (6000 \cdot 5999)/2 \approx 18 \cdot 10^6. \quad (А.3)$$

**Приложение Б**
**(справочное)**

## Проявление кодового центра биометрического образа

Графическая интерпретация процесса проявления кодового центра биометрического образа представлена на рисунке Б.1.

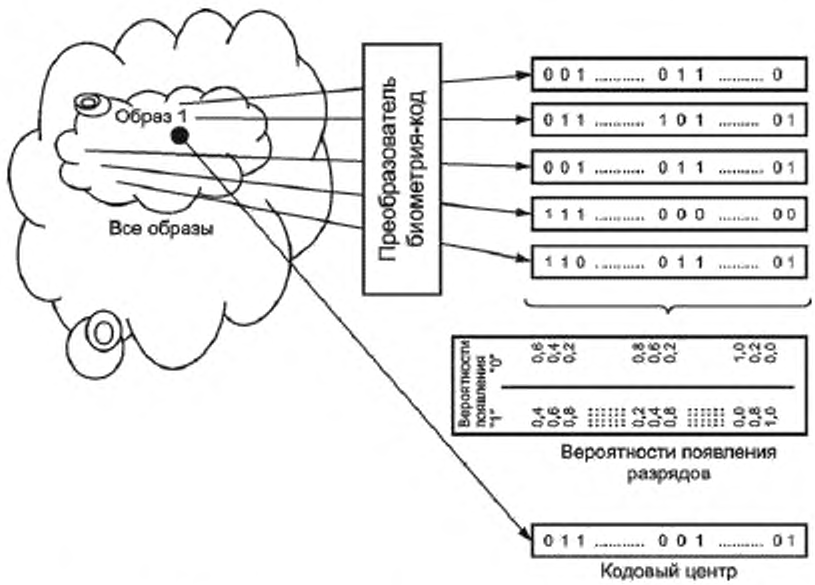


Рисунок Б.1 — Проявление кодового центра биометрического образа

Проявление осуществляется следующим образом. Для каждого разряда определяются вероятности появления в этом разряде единичных и нулевых битов. В случае превалирования значений в определенном разряде его значение считается стабильным. При появлении нестабильных разрядов необходимо увеличивать количество используемых при проявлении биометрических примеров до тех пор, пока разряд не станет стабильным. Наличие 1—2 % нестабильных разрядов в проявленном центре допускается.

**Приложение В**
**(справочное)**

## Пример размножения примеров изображения отпечатка пальца перестановкой фрагментов векторов биометрических параметров

При формировании синтетического примера изображения отпечатка пальца совмещают исходные примеры изображения отпечатка пальца и затем производят выделение особенностей отпечатка.

Например, для этого на изображение может быть наложена прямоугольная сетка, состоящая из 12·19 = 228 ячеек. Затем в каждой ячейке сетки производится проверка на наличие особенности.

После этого данные, полученные при анализе ячеек сетки, преобразовывают в вектор биометрических параметров, значения которого характеризуют наличие или отсутствие особенности в одном из фрагментов исходного изображения, ограниченного ячейкой сетки (см. рисунок В.1).

Полученные для исходных биометрических примеров векторы параметров имеют некоторую стабильную часть, общую для всех примеров, а также нестабильную часть, различающуюся для разных примеров.

Для формирования синтетического биометрического примера выбирают случайные фрагменты биометрических параметров исходных биометрических примеров (см. правую часть рисунка В.1). Синтетический биометрический пример формируется путем замещения выделенных фрагментов одного примера на соответствующие фрагменты других примеров. При размножении двух примеров выделенные фрагменты включают примерно половину особенностей изображения отпечатка пальца. При использовании трех примеров из каждого исходного примера выделяются фрагменты, включающие примерно 1/3 особенностей изображения отпечатка пальца.

На рисунке В.1 изображена ситуация синтеза нового примера из двух исходных примеров рисунков отпечатков пальца. Первый пример отображен на рисунке и соответствует случайному выделению трех фрагментов параметров, включающих примерно половину замещаемых особенностей анализируемого рисунка.

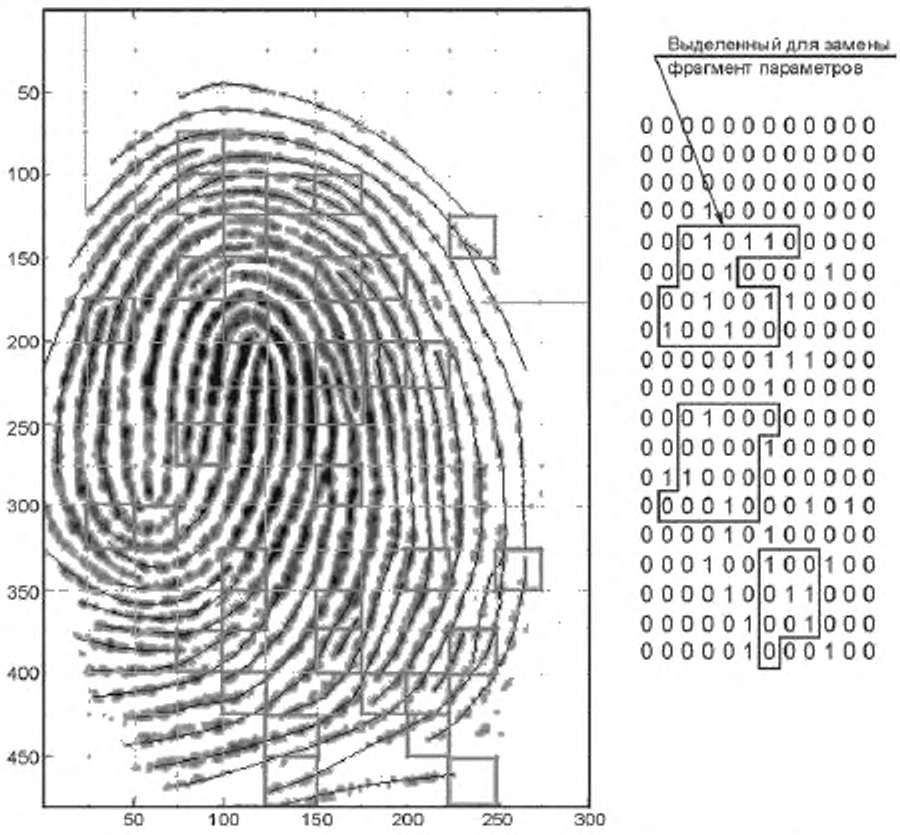


Рисунок В.1 — Выделение фрагментов вектора биометрических параметров изображения отпечатка пальца при размножении двух примеров

**Приложение Г**
**(справочное)**

## Пример размножения трех голосовых образов перестановкой фонем

Для синтеза голосового образа из трех голосовых образов (фраза-1, фраза-2, фраза-3 на рисунке Г.1) осуществляют разбиение каждой исходной фразы на фонемы (см. рисунок Г.1).

Синтетический голосовой образ формируется путем добавления 1/3 фрагментов каждого из исходных образов, например, двух фонем фразы-1 (фонемы 1, 2), двух фонем фразы-2 (фонемы 7, 9), двух фонем фразы-3 (фонемы 20, 21).

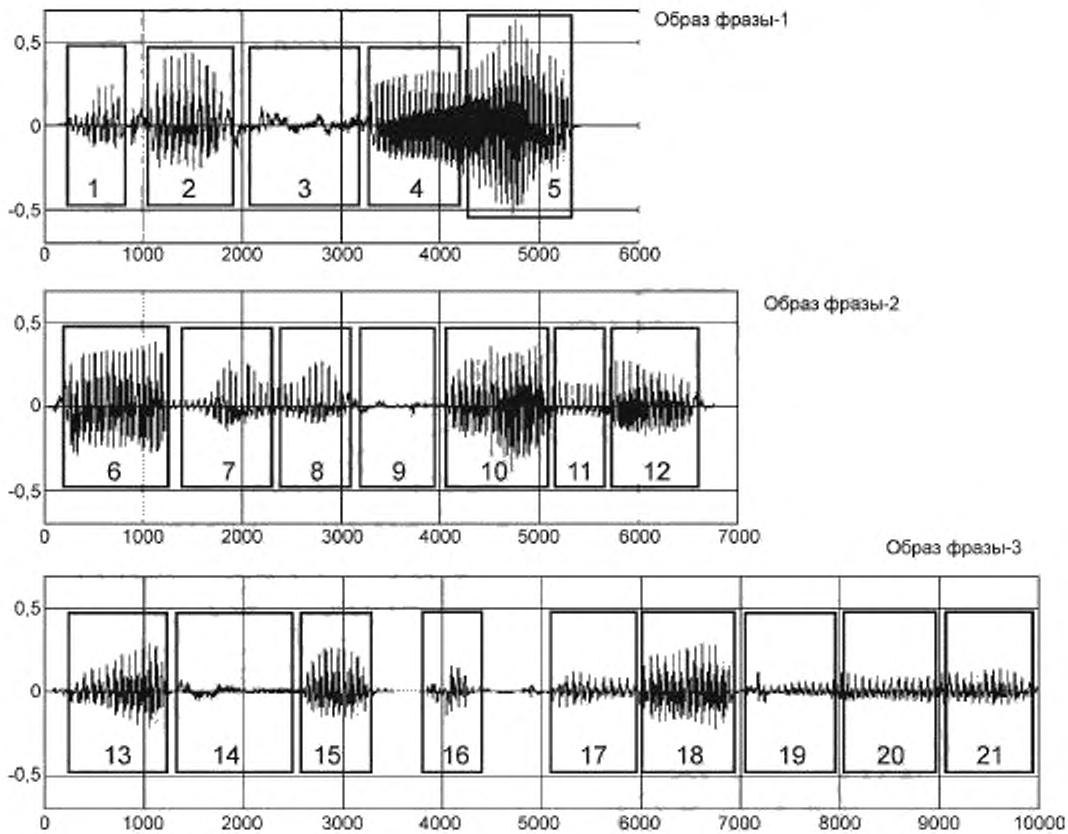


Рисунок Г.1 — Разбивание голосовых образов на фонемы

Число примеров синтетического голосового образа определяется числом примеров исходных биометрических образов (фраза-1, фраза-2, фраза-3).

Каждый пример синтетического голосового образа, собранный из фонем 1, 2, 7, 9, 20, 21, формируется из разных комбинаций примеров исходных голосовых образов.

Исключение, добавление, замещение и изменение последовательности фонем приводит к появлению других синтетических голосовых образов.

УДК 681.18:006.354 ОКС 35.040 Т00

Ключевые слова: техническая защита информации, биометрия, тестирование, синтетические биометрические образы

Редактор *П. М. Смирнов*
Технический редактор *В. Н. Прусакова*
Корректор *Н. И. Гаврищук*
Компьютерная верстка *Т. Ф. Кузнецовой*

Сдано в набор 15.02.2011. Подписано в печать 01.03.2011. Формат 60×84 1/8. Бумага офсетная. Гарнитура Ариал. Печать офсетная. Усп. печ. л. 2,79. Уч.-изд. л. 1,90. Тираж 114 экз. Зак 123

ФГУП «СТАНДАРТИНФОРМ», 123995 Москва, Гранатный пер., 4.
www.gostinfo.ru info@gostinfo.ru
Набрано и отпечатано в Калужской типографии стандартов, 248021 Калуга, ул. Московская, 256.