

Учреждение образования
«БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ»

Объект авторского права
УДК 004; 004.93'1; 004.932

АХУНДЖАНОВ
Умиджон Юнус угли

**ВЕРИФИКАЦИЯ РУКОПИСНОЙ ПОДПИСИ В
СТАТИЧЕСКОМ РЕЖИМЕ**

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

**по специальности 05.13.01 – Системный анализ, управление и обработка
информации**

Минск 2023

Научная работа выполнена в государственном научном учреждении «Объединенный институт проблем информатики Национальной академии наук Беларуси».

Научный руководитель

Старовойтов Валерий Васильевич, доктор технических наук, профессор, главный научный сотрудник лаборатории идентификации систем государственного научного учреждения «Объединенный институт проблем информатики НАН Беларуси»

Официальные оппоненты:

Богущ Рихард Петрович, доктор технических наук, доцент, заведующий кафедрой вычислительных систем и сетей учреждения образования «Полоцкий государственный университет имени Евфросинии Полоцкой»

Садов Василий Сергеевич, кандидат технических наук, доцент, профессор кафедры интеллектуальных систем Белорусского государственного университета

Оппонирующая организация **Белорусский национальный технический университет**

Защита состоится «24» октября 2023 г. в 11.00 на заседании совета по защите диссертаций Д 02.15.01 при учреждении образования «Белорусский государственный университет информатики и радиоэлектроники» по адресу: 220013, г. Минск, ул. П. Бровки, 6, корп. 1, ауд. 232, тел.: (017) 293-89-89, e-mail: dissovet@bsuir.by.

С диссертацией можно ознакомиться в библиотеке учреждения образования «Белорусский государственный университет информатики и радиоэлектроники».

Автореферат разослан « 22 » сентября 2023 г.

Ученый секретарь совета по защите диссертаций, кандидат технических наук, доцент



В. А. Рыбак

ВВЕДЕНИЕ

В настоящее время в мире наблюдается активное развитие научно-технического направления, связанного с исследованием и анализом различных биометрических технологий. Биометрические технологии по типу используемых параметров можно разделить на две группы: статические системы и динамические, или поведенческие. Статические системы базируются на индивидуальных признаках человека, определяемых его физиологией, например отпечатках пальцев, форме лица и уха, узоре радужной оболочки глаза, кровеносной системе на сетчатке глаза, ДНК. Динамические системы оценивают поведенческие характеристики человека, например рукописный почерк, подпись, голос, походку.

Верификация подписи – это проверка подлинности исполнения подписи определенным человеком. Для обеспечения процедуры верификации подписи необходимо иметь образцы нескольких подлинных подписей человека и процедуру ее анализа. Задача верификации подписи сложнее задачи распознавания, т. е. определения, какому из известного числа людей принадлежит подпись.

Для рукописной подписи человека характерна внутриклассовая изменчивость – различия между образцами подписей, сделанными в разное время и разном эмоциональном состоянии. Чем различия больше, тем ниже эффективность системы, так как расширяется область возможных значений параметров эталона, что предоставляет злоумышленнику больше возможностей подделки подписи.

Проверка подлинности подписей является одной из сложных задач распознавания образов. Проверка статических подписей выполняется в несколько этапов: предварительная обработка их изображений, вычисление локальных и глобальных признаков подписи, присущих конкретному человеку, их анализ и установление подлинности. Во всем мире большое внимание уделяется разработке программных средств предварительной обработки, вычислению признаков и верификации оцифрованных рукописных подписей. В развитие методов верификации рукописной подписи значительный вклад внесли Coetzer J., Luiz G., Yilmaz M.B., Kholmatov A., Gupta G., Plamondon R., Nalwa V.S., Fierrez-Aguilar J., Freitas C., Matsumoto T., Liu C. и др.

В последние годы был предложен ряд решений задачи верификации бумажной подписи, разработаны экспериментальные программы, но окончательное решение задачи еще не найдено. Объективная экспертиза подлинности подписи под документами требуется регулярно. Этот факт свидетельствует об актуальности настоящей диссертационной работы.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Связь работы с научными программами (проектами), темами

Тема диссертационной работы соответствует пункту 6 приоритетных направлений научных исследований Республики Беларусь на 2021-2025 гг. Указа Президента Республики Беларусь № 156 от 7 мая 2020 г. «О приоритетных направлениях научной, научно-технической и инновационной деятельности на 2021-2025 годы».

Также, настоящее диссертационное исследование в определенной степени служит для выполнения задач, предусмотренных указами Президента Республики Узбекистан №УП-4947 от 7 февраля 2017 года «О стратегии действий по дальнейшему развитию Республики Узбекистан», УП-5349 от 19 февраля 2018 года «О мерах по дальнейшему совершенствованию сферы информационных технологий и коммуникаций», УП-6079 от 5 октября 2020 года «Об утверждении стратегии «цифровой Узбекистан-2030» и мерах по ее эффективной реализации», постановлениями Президента №ПП-3245 от 29 августа 2017 года «О мерах по дальнейшему совершенствованию системы управления проектами в сфере информационно-коммуникационных технологий», ПП-4699 от 28 апреля 2020 года «О мерах по широкому внедрению цифровой экономики и электронного правительства», ПП-4996 от 17 февраля 2021 года «О мерах по созданию условий для ускоренного внедрения технологий искусственного интеллекта», а также другими нормативно-правовыми актами, принятыми в данной сфере.

Работа выполнялась в государственном научном учреждении «Объединенный институт проблем информатики Национальной академии наук Беларуси» в рамках НИР ГБ.

Работа выполнена в рамках:

1. Государственной программы научных исследований «Информатика, космос и безопасность», 2019-2020 гг., подпрограммы «Информатика и космические исследования», задания «Методы, алгоритмы и технологии интеллектуального анализа цифровых сигналов в компьютерных системах обработки данных и распознавания образов», № госрегистрации 20190953, период выполнения 2019-2020 гг.
2. Государственной программы научных исследований «Цифровые и космические технологии, безопасность человека, общества и государства», 2021-2023 гг., подпрограммы «Цифровые технологии и космическая информатика», задания «Методы и алгоритмы интеллектуального анализа

больших данных в компьютерных системах обработки информации», № госрегистрации 20210776, период выполнения 2021-2023 гг.

Цель, задачи, объект и предмет исследования

Целью исследования является автоматизация процесса верификации рукописной подписи, представленной на бумажном носителе. Для достижения поставленной цели требуется решить следующие задачи:

- выполнить обзор литературы по теме диссертации;
- собрать репрезентативное множество подлинных и поддельных рукописных подписей;
- разработать алгоритмы предварительной обработки изображений подписи и найти их оптимальные параметры;
- сформировать унифицированное представление изображения подписи;
- найти множество признаков, позволяющих различать подлинные и поддельные подписи произвольного человека;
- построить классификатор, решающий задачу верификации на малом множестве образцов подлинной подписи;
- разработать экспериментальное ПО системы верификации подписей и подтвердить корректность предложенных решений.

Объектом исследования являются изображения рукописных подписей одного человека.

Предметом исследования являются алгоритмы обработки изображений, вычисления признаков и верификации оцифрованных изображений рукописных подписей.

Научная новизна

1. Разработана методика предварительной обработки цифровых изображений произвольных рукописных подписей, не зависящая от их исполнения (размера, цвета, типа ручки, наклона и ориентации подписи).

2. Предложен новый признак описания локальных структурных особенностей подписи, присущих конкретному человеку, в виде массива, описывающего частотное распределение кодов локальных бинарных шаблонов (LBP) контуров подписи.

3. Предложен признак контуров подписи, описывающий, в отличие от известных, нормализованное частотное распределение значений локальной кривизны контуров подписи.

4. Разработана и экспериментально проверена методика верификации подписи на базе модели одноклассовой машины опорных векторов, построенной для N ($5 \leq N \leq 15$) подлинных подписей человека в двумерном признаковом пространстве. В этом пространстве представлены образы пар подписей в виде коэффициентов корреляции между описанными выше многомерными признаками для всех пар подлинных подписей. Модель SVM оценивает, имеются ли выбросы образов пар верифицируемой подписи с N подлинными относительно класса подлинных подписей человека.

Положения, выносимые на защиту

1. Методика предварительной обработки цифровых изображений рукописных подписей, отличающаяся от известных единообразием и независимостью от исполнения (размера, цвета, типа ручки, наклона и ориентации подписи) и позволяющая представлять подписи в виде бинарного контурного изображения фиксированного размера.

2. Новый признак особенностей формы подписи человека, который отличается от известных тем, что описывает подпись в виде нормализованного частотного распределения кодов локальных бинарных шаблонов (LBP), вычисляемых по контурам изображения подписи.

3. Признак контуров подписи, описывающий, в отличие от известных, нормализованное частотное распределение значений локальной кривизны контуров подписи.

4. Верификация подписи на базе модели одноклассовой машины опорных векторов (SVM), построенной для N ($5 \leq N \leq 15$) подлинных подписей человека в двумерном признаковом пространстве, которая отличается от известных тем, что обучение классификатора выполняется в пространстве образов пар подписей. Что приближает точность процедуры верификации к 100 %.

Личный вклад соискателя ученой степени

Основные результаты и положения, выносимые на защиту, получены лично автором. Научный руководитель и соавтор совместных работ доктор технических наук, профессор В. В. Старовойтов принимал участие в постановке целей и задач исследования, их предварительном анализе, планировании экспериментов, а также в обсуждении полученных результатов.

Апробация диссертации и информация об использовании ее результатов

Основные положения и результаты диссертационной работы докладывались на следующих конференциях: Международной научно-практической конференции «Актуальные проблемы и инновационные технологии в области естественных наук» (20–21 ноября 2020 г., Ташкент, Узбекистан); Международной научной конференции молодых учёных «Наука и инновации» (26 ноября 2020 г., 25 ноября 2021 г., Ташкент, Узбекистан); 5-й Международной научно-практической конференции «САПР и моделирование в современной электронике» (21–22 октября 2021 г., Брянск, Россия); 3-й Международной научно-практической конференции «Компьютерные технологии и анализ данных (CTDA'2022)» (21–22 апреля 2022 г., Минск, Беларусь); 8-й Международной научно-практической конференции «BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня» (11–12 мая 2022 г., Минск, Беларусь); 20-й Белорусско-Российской научно-технической конференции «Технические средства защиты информации» (7 июня 2022 г., Минск, Беларусь); 10-й научно-технической интернет-конференции «Информационные технологии в образовании, науке и производстве» (21–22 ноября 2022 г., Минск, Беларусь); 58-й научной конференции аспирантов, магистрантов и студентов БГУИР (18–22 апреля 2022 г., Минск, Беларусь); 3-й Международной научно-практической конференции «Судебная экспертиза: теория и практика в современных условиях» (26–27 апреля 2023 г., Минск, Беларусь); 5-й Международной научно-практической конференции «Наука молодых – наука будущего» (5 июня 2023 г., Петрозаводск, Российская Федерация).

Опубликованность результатов диссертации

По результатам выполненных исследований, представленных в диссертации, опубликовано 17 научных работ, в том числе: 5 статей в рецензируемых научных журналах, рекомендованных ВАК Республики Беларусь для опубликования результатов научных исследований, 12 статей в сборниках материалов докладов научных конференций. Общий объем опубликованных работ – 3,5 авторского листа.

Структура и объем диссертации

Диссертационная работа состоит из перечня сокращений и обозначений, введения, общей характеристики работы, четырех глав, заключения,

библиографического списка и приложений. Полный объем диссертационной работы составляет 156 страниц, включая 56 рисунков и 29 таблиц на 65 странице; 5 приложений на 36 страницах; список использованных источников, включающий 169 наименований на 14 страницах и список публикаций соискателя по теме диссертации – 17 наименований на 3 страницах.

ОСНОВНАЯ ЧАСТЬ

Во введении обоснована актуальность темы диссертационного исследования, которая обуславливается необходимостью ее комплексного научного изучения и практического исследования, определена и рассмотрена степень изученности проблемы, рассматриваемой в диссертационном исследовании.

В **первой главе** описаны основные понятия систем верификации подписей, проанализированы достоинства и недостатки известных подходов к созданию систем верификации рукописных подписей [1–А, 6–А, 10–А]. Сложность задачи верификации подписи определяется следующими факторами. Основные трудности заключаются в несоответствиях между подписями одного и того же человека, сделанными в разное время: различия могут возникать из-за расположения и ориентации подписи, ширины пера, качества ручки, стресса, настроения человека. Почерки разных людей бывают похожи, а подпись человека всегда вариативна.

При анализе литературы установлено, что эксперты, занимающиеся почерковедческими исследованиями подписей в Беларуси и России, в основном анализируют рукописные подписи визуально и дают субъективные оценки их подлинности, опираясь на собственный опыт и служебные методики.

Представлена общая структура системы верификации рукописных подписей. Определено, что наиболее важным этапом работы системы распознавания, определяющим качество ее работы и обучения, является выделение информативных признаков. Разные исследователи используют разные признаки.

Приведены основные оценки качества бинарной классификации подписей на два класса подлинные и поддельные. Определено, что чаще всего результаты классификации подлинности подписей оцениваются по матрице ошибок (confusion matrix) функциями «точность» (Accuracy) и «равный уровень ошибки» (EER).

Описаны общедоступные базы отсканированных изображений подлинных и поддельных подписей, которые используются исследователями

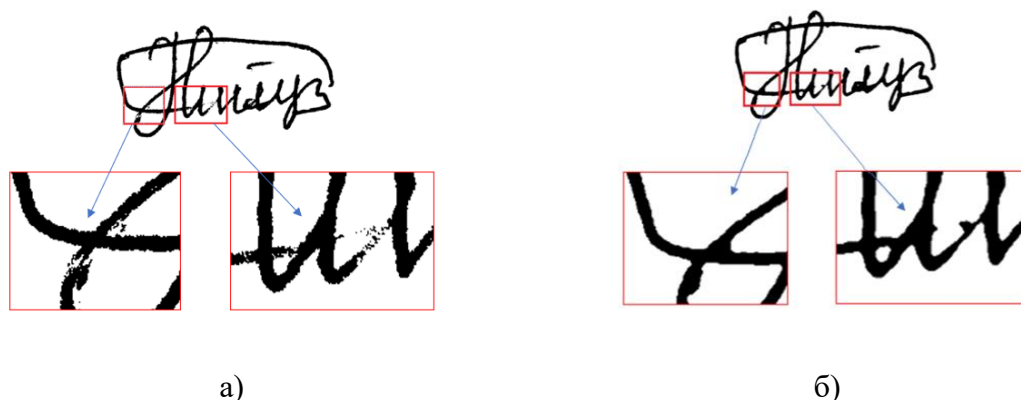
для сравнения систем верификации. Приведены известные результаты точности верификации подписей из трех общедоступных баз.

По результатам анализа литературы сделан вывод о необходимости разработки и экспериментального исследования единой методики предварительной обработки изображений рукописных подписей, выбора универсальных индивидуальных признаков подписи человека и метода построения модели классификатора, обученного на ограниченном числе подлинных подписей этого человека.

Во **второй главе** описаны исследования, посвященные выбору конкретных методов предварительной обработки изображений отсканированных подписей для их унифицированного цифрового представления с целью последующего вычисления признаков и выполнения верификации.

В результате экспериментальных исследований были сформулированы рекомендации по преобразованию цветных изображений рукописных подписей, выполненных разными типами ручек и разного цвета в полутоновое представление [2–А, 4–А, 11–А]. Для преобразования полутонового изображения подписи, выполненной на однородном фоне, в бинарное выбран метод Отсу.

Экспериментально обоснованы алгоритмы фильтрации бинарного изображения подписи и их параметры для стабилизации толщины линий (рисунок 1).



**Рисунок 1 – Фильтрация бинарного изображения рукописной подписи:
а) исходное изображение; б) результат морфологической и медианной
фильтрации**

Нормализация ориентации подписи выполняется с помощью метода главных компонент (principal component analysis, PCA). Он находит направление максимальной дисперсии в многомерных данных, которое описывается главной компонентой. Для поворота цифрового изображения подписи методом PCA используется координатное представление множества

черных пикселей подписи на бинарном изображении. Благодаря этому размерность их представления равна двум, что значительно сокращает вычислительные затраты.

Определено, что для более точного сравнения растровых изображений подписей их лучше приводить к единому размеру. Экспериментально определен размер такого шаблона, равный 300×150 пикселей.

Предложен морфологический алгоритм выделения границ подписи. Бинарное изображение границ подписи в виде контуров толщиной в один 8-связный пиксел предложено использовать в качестве ее унифицированного цифрового представления фиксированного размера (рисунок 2).

Экспериментально показано, что такое представление сохраняет индивидуальные особенности подписи человека.

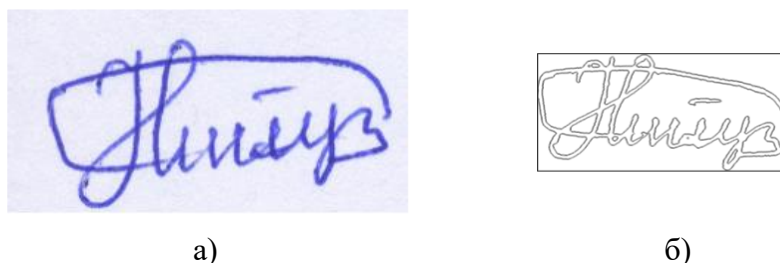


Рисунок 2 – Унифицированное цифровое представление рукописной подписи:
а) исходное изображение подписи; б) ее нормализованное контурное представление

В результате выполнения экспериментальных исследований предложена методика предварительной обработки изображений рукописных подписей, состоящая из следующих шагов:

- 1) Если подписи представлены на бумажном носителе, оцифровать их с разрешением 600 dpi в виде цветных изображений.
- 2) Если подписи представлены в виде цифровых изображений, проанализировать тип исходных изображений:
 - если изображение полутоновое, преобразовать его в бинарное, перейти на шаг 3;
 - если изображение бинарное, перейти на шаг 3;
 - если изображение цветное, преобразовать его в бинарное изображение подписи (БИП) с учетом доминирующего цвета.
- 3) При необходимости удалить фрагменты изображения, не относящиеся к собственно подписи.
- 4) Оценить качество цифрового представления подписи. Если оно неудовлетворительное, данную подпись следует признать непригодной для верификации.

- 5) Отфильтровать БИП методами математической морфологии с использованием структурирующего элемента в форме креста размером 3×3 и медианного фильтра с маской 5×5 .
- 6) Повернуть БИП с использованием метода главных компонент.
- 7) Применить масштабирование описывающего прямоугольника БИП в шаблон фиксированного размера 300×150 .
- 8) Выделить контуры подписи методами математической морфологии.

Данная методика предварительной обработки изображений подписей была протестирована на данных из общедоступных баз CEDAR, BHSig260-Bengali, BHSig260-Hindi и TUIT, содержащих в совокупности 19880 изображений рукописных подписей, которые были выполнены людьми разных национальностей, с разными особенностями исполнения подписей. Все изображения подписей были преобразованы в шаблон единого размера. Такая нормализация подписи делает ее представление не зависящим от цвета, толщины линий, размера и ориентации на плоскости, сохраняя все геометрические особенности подписи человека, которые далее используются для вычисления инвариантных признаков.

Основные результаты, представленные в данной главе, опубликованы в работах [2–А, 3–А, 6–А – 11–А, 14–А – 16–А].

В **третьей главе** приведено описание двух новых признаков, позволяющих описывать локальные особенности структуры подписи человека независимо от ее исполнения.

В настоящей работе идея вычисления локальных бинарных шаблонов (LBP), применяемых к вычислению текстурных признаков на полутоновых изображениях, адаптирована к вычислению структурных признаков контурно-бинарного представления подписи фиксированного размера. При вычислении локального шаблона контура код 2^i присваивается i -му соседнему (из восьми) пикселу при условии, что он и центральный пиксел имеют черный цвет (значение 0) (рисунок 3). Полученный код описывает локальную структуру подписи в окрестности 3×3 , т. е. ее изгибы, либо прямолинейные фрагменты [3–А, 14–А]. Каждому пикселу изображения ставится в соответствие число от 0 до 255, определяющее его локальный шаблон. После этого строится гистограмма вычисленных LBP-признаков для 256 интервалов. Она записывается в виде массива из 256 элементов. Из этого массива отбрасываются первый и последний элементы, соответствующие вариантам, когда все восемь соседних пикселов имеют либо белые, либо черные значения. Получившийся набор из 254 чисел нормализуется путем деления на сумму всех элементов. В работе он назван LBP-кодом подписи и является новым многомерным признаком, описывающим частотные распределения локальных структур контура подписи независимо от цвета, толщины линий, исходных

размеров и ориентации оригинальной подписи, выполненной на бумаге. Такие коды подписей можно сравнивать между собой. Данный признак предлагается в первые.

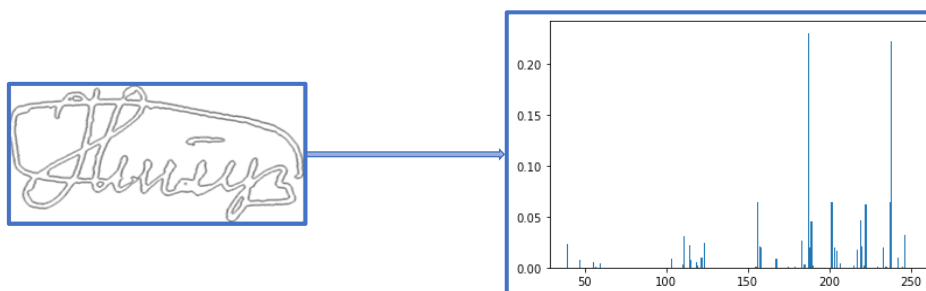


Рисунок 3 – Контурное представление подписи и гистограмма ее LBP-признаков

В качестве второго типа признаков подписи предложено использовать значения локальной кривизны кривых, аппроксимирующих растровое представление подписи в каждом пикселе ее контурного представления. После вычисления значений кривизны в каждом пикселе строится нормализованная гистограмма этих значений для 40 интервалов. Она записывается в виде 40-мерного кода локальной кривизны подписи.

Кривизна линии в точке (x, y) в работе называется локальной кривизной. Она равна $K = 1/R$, где R – радиус окружности, касающейся линии в этой точке. Известно, что координаты центра (x_c, y_c) и радиус окружности можно точно вычислить по координатам трех точек, лежащих на окружности. Окружность, проходящая через точки (x_i, y_i) , описывается уравнением (1). Подставив в последнее уравнение координаты n последовательных пикселей контура, получим систему линейных уравнений относительно трех неизвестных p_i (2).

$$(x_i - x_c)^2 + (y_i - y_c)^2 = R^2, \quad (1)$$

$$x_i^2 - 2x_i x_c + x_c^2 + y_i^2 - 2y_i y_c + y_c^2 - R^2 = 0,$$

$$(x_i^2 + y_i^2) - 2(x_i x_c + y_i y_c) + x_c^2 + y_c^2 - R^2 = 0,$$

$$(x_i^2 + y_i^2) + p_1 x_i + p_2 y_i + p_3 = 0. \quad (2)$$

Решая систему уравнений, можно вычислить приближенные параметры окружности.

На растре три последовательных 8-связных пиксела контура имеют всего три комбинации (рисунок 4). Поворот любой из них на 45° не меняет радиус касательной окружности и, соответственно, значение кривизны.



Рисунок 4 – Варианты соседства пикселей, порождающих разные значения кривизны

Малое число вариантов касательных окружностей очень ограничивает возможности использования значений кривизны в качестве признаков подписи и не отражает истинной локальной кривизны линий, представленных на растре. При выборе пяти последовательных 8-связных пикселей контура возможны 30 разных значений локальной кривизны. Разнообразие значений увеличивается только с увеличением числа пикселей, используемых для вычисления значения кривизны (рисунок 5).

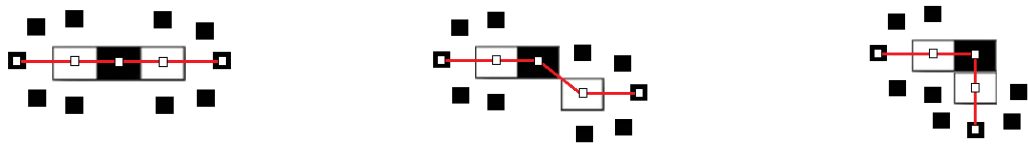


Рисунок 5 – Варианты расположения пяти соседних пикселей контура, проходящего через темный пиксел

Для вычисления кривизны окружности, касающейся темных пикселей, показанных на рисунке 5, требуется вычислить параметры касательной окружности с минимальной погрешностью.

Пусть дано $N > 3$ последовательных точек контура (x_i, y_i) , $i = 1, \dots, N$. Для вычисления кривизны в точке $(\frac{x_{N+1}}{2}, \frac{y_{N+1}}{2})$ требуется решить уравнение относительно неизвестных p_i :

$$x^2 + y^2 + p_1x + p_2y + p_3 = 0. \quad (3)$$

Подставляя в уравнение (3) координаты N последовательных точек кривой и представляя его в матричном виде, получаем

$$A = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ \dots & \dots & \dots \\ x_N & y_N & 1 \end{bmatrix}, B = - \begin{bmatrix} x_1^2 & y_1^2 \\ x_2^2 & y_2^2 \\ \dots & \dots \\ x_N^2 & y_N^2 \end{bmatrix}.$$

Решая уравнение (3) в матричном виде $B + PA = 0$ или $P = -B \setminus A$ относительно неизвестного вектора P , имеем параметры касательной окружности и величину кривизны K с минимальной среднеквадратичной ошибкой:

$$x_c = -0,5 p_1, \quad y_c = -0,5 p_2,$$

$$R = \frac{\sqrt{(p_1^2 + p_2^2)}}{4 - p_3}, \quad K = \frac{1}{R}. \quad (4)$$

При увеличении числа пикселей максимальное значение кривизны уменьшается. Поскольку подписи даже одного человека в нормализованном растровом представлении имеют разное число пикселей, гистограмму значений локальной кривизны требуется нормализовать путем деления на число точек, в которых кривизна вычислялась. Нормализованные гистограммы записываются в виде массива, который называется кодом локальной кривизны подписи. Он является ее многомерным вектор-признаком, описывающим индивидуальные особенности почерка человека. Такие коды подписей можно сравнивать между собой. Данный признак предлагается впервые.

В четвертой главе описаны эксперименты по двум вариантам построения классификаторов и результаты их применения для решения задачи верификации изображений подписей. Исследованы классификаторы на базе искусственной нейронной сети и на базе машины опорных векторов.

Основной проблемой при решении поставленной задачи является ограниченное число подлинных подписей, сделанных человеком до постановки задачи верификации его подписи. В экспертной практике для анализа подлинности подписи используется ее сравнение с подлинными подписями человека числом от 5 до 15. Этого количества образцов недостаточно для обучения нейронной сети. Следует отметить, что стандартные процедуры аугментации обучающих данных не создают принципиально иных образов подписи человека.

Лучшая точность корректной верификации подписей в экспериментах с применением сверточной нейронной сети была достигнута на изображениях размером 250×150 пикселей для подписей из баз TUIT и CEDAR и равнялась 90,04 и 94,38 %, соответственно. При тестировании в основном использовались подписи тех людей, для которых выполнялось обучение. Вопрос, каков будет результат верификации подписей, на которых сеть не обучалась, остается открытым.

В настоящей работе вместо классического представления образа объекта (в данном исследовании подписи) в виде точки в многомерном пространстве впервые предлагается признаковое пространство нового типа, в котором в качестве признаков представлена близость пар подписей, оцениваемая коэффициентами корреляции между их кодами. Возможно использование и других функций вычисления близости, однако коэффициенты корреляции в

отличие от других функций имеют ограниченный диапазон значений $[-1; +1]$ и просты в вычислении. Далее строятся образы близости всех возможных пар анализируемых подписей, а признаками служат коэффициенты корреляции между парами «верифицируемая подпись и подлинная» и «подлинная подпись и другая подлинная подпись» для LBP-кодов и кодов их кривизны. Такое корреляционное представление признаков позволяет вместо пространства исходных LBP-признаков и признаков кривизны каждой подписи суммарной размерностью около 294 использовать наглядное двумерное признаковое пространство попарного сходства подписей. Оба корреляционных признака независимы и имеют одинаковый диапазон значений $[-1; +1]$. При попарном сравнении подписей число образов в признаковом пространстве как верифицируемой подписи, так и подлинных подписей увеличивается (таблица 1). Однако новые образы – это образы не одной подписи, а парного сходства двух подписей.

В таблице 1 показано, как нелинейно увеличивается число признаков при обучении модели классификатора на N подлинных подписях человека. В новом признаковом пространстве анализируемая на подлинность подпись сравнивается с каждой подлинной подписью.

Классификация в исходном многомерном пространстве при малом числе образов подписей менее точна и чувствительна к выбросу одного или нескольких признаков, поскольку их диапазоны значений неравны. В новом признаковом пространстве этот недостаток нивелируется, поскольку корреляция имеет ограниченный диапазон значений.

Таблица 1 – Число образов для разного числа подлинных подписей

Число образов при стандартной верификации			Число образов при попарной верификации	
№	Верифицируемая подпись	Подлинные подписи, N штук	Образы с верифицируемой подписью, N штук	Образы подлинных подписей, $N(N-1)/2$ штук
1	1	5	5	10
2	1	7	7	21
3	1	9	9	36
4	1	11	11	55
5	1	13	13	78
6	1	15	15	105

По результатам экспериментальных исследований для верификации подписи предложено использовать модель не нейронной сети, а одноклассовой машины опорных векторов (SVM), обученной на образах близости пар N подлинных подписей человека, где $5 \leq N \leq 15$. Образы близости

описываются коэффициентами корреляции между LBP-кодами пар подлинных подписей и корреляции между кодами локальной кривизны этих же подписей. Исходные коды имеют размерность 254 и 40, но после вычисления коэффициентов корреляции между парами кодов подписей размерность признакового пространства снижается до двух. В то же время число образов в этом пространстве возрастает при обучении SVM на N подлинных подписях до $N(N-1)/2$ штук. В частности, при $N = 15$ машина обучается не на 15, а на 105 образах близости пар подписей. Модель SVM, обученная в этом пространстве, оценивает, образуют ли выбросы образы пар верифицируемой подписи с N подлинными относительно класса пар подлинных подписей человека, на котором она обучена. Некоторые подлинные подписи могут существенно отличаться и породить выбросы в паре с верифицируемой подписью, поэтому окончательный результат верификации определяется голосованием. Если выбросов больше $N/2$, подпись считается поддельной, и наоборот (рисунки 6 и 7).



Рисунок 6 – Верхний ряд – изображения для верификации из базы CEDAR слева – поддельная подпись, справа – подлинная; нижний ряд – семь подлинных подписей этого же человека

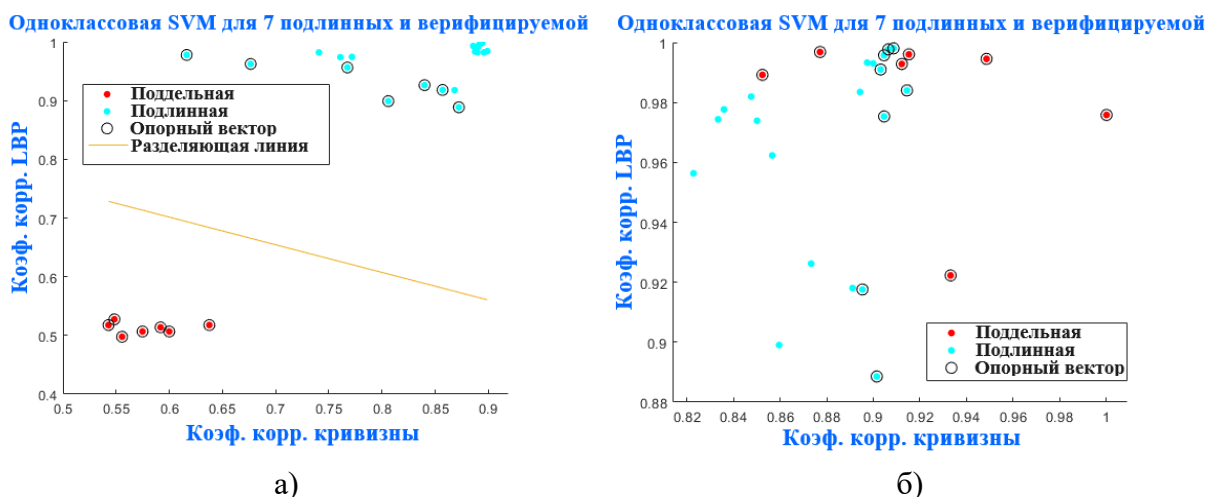


Рисунок 7 – Пример верификации подписи при сравнении с семью подлинными: (а) подпись Ver_f.png можно отделить от подлинных подписей, она определена как поддельная, (б) подпись Ver_t.png невозможно отделить от других подписей, она подлинная

Для проведения экспериментальных исследований использовался персональный компьютер со следующими характеристиками:

- процессор: AMD Ryzen 5 3600 6-Core Processor, количество ядер 6, тактовая частота 3600 МГц;
- ОЗУ: 16 Гб;
- видеокарта: NVIDIA Quadro RTX 4000, 6 Гб;
- операционная система: Windows 11 Pro, разрядность 64.

Верификация выполнялась с помощью одноклассовой SVM посредством сравнения анализируемой подписи человека с N подлинными подписями этого же человека, где N изменялось от 5 до 15. В таблице 2 представлены результаты верификации всех 2640 подписей из общедоступной базы CEDAR. Корреляция вычислялась двумя методами: по Пирсону и по Спирмену. TP , TN , FP , FN – параметры матрицы ошибок, $Accuracy$ – точность верной классификации в процентах. Нижняя строка выделена жирным шрифтом, в ней представлены лучшие результаты верификации.

Таблица 2 – Результаты верификации рукописных подписей при разных параметрах

Коэффициент корреляции	N	TP	TN	FP	FN	$Accuracy$, %
Пирсона	5	1185	465	135	855	62,50
Пирсона	7	1267	300	53	1020	59,36
Пирсона	9	1296	242	24	1078	58,26
Пирсона	11	1307	200	13	1120	57,08
Пирсона	13	1311	288	9	1032	60,59
Пирсона	15	1313	321	7	999	61,89
Спирмена	5	1131	1320	189	0	92,84
Спирмена	7	1237	1320	83	0	96,86
Спирмена	9	1279	1320	41	0	98,45
Спирмена	11	1293	1320	27	0	98,98
Спирмена	13	1307	1320	13	0	99,51
Спирмена	15	1314	1320	6	0	99,77

Комплекс разработанных алгоритмов, двух новых признаков и метод верификации рукописных подписей на базе одноклассовой машины опорных векторов позволили при использовании 15 подлинных подписей человека достичь точности верификации всех 2640 подписей, равной 99,77 % на общедоступной базе CEDAR. Лучший из опубликованных результатов верификации подписей этой базы составляет 94,40 %.

ЗАКЛЮЧЕНИЕ

В результате выполненных исследований разработана методика предварительной обработки изображений подписи, новые признаки и методика верификации подписи на базе одноклассовой машины опорных векторов, которые позволили при наличии 15 подлинных подписей человека достичь 99,77 % точности верификации подписи на общедоступной базе CEDAR, состоящей из 2640 подписей.

Основные научные результаты диссертации

1) Разработана методика предварительной обработки цифровых изображений рукописных подписей, отличающаяся от известных единообразием и независимостью от исполнения (размера, цвета, типа ручки, наклона и ориентации подписи) и позволяющая представить подписи в виде бинарного контурного изображения фиксированного размера [1–А, 2–А, 4–А, 6–А – 16–А].

2) Предложен новый признак описания локальных структурных особенностей подписи, присущих конкретному человеку, который в отличие от известных передает частотное распределение кодов локальных бинарных шаблонов (LBP) контуров подписи [3–А, 14–А – 16–А].

3) Предложен признак контуров подписи, описывающий, в отличие от известных, нормализованное частотное распределение значений локальной кривизны контуров подписи [5–А, 17–А].

4) Разработана и экспериментально проверена методика верификации подписи на базе модели одноклассовой машины опорных векторов, построенной для N ($5 \leq N \leq 15$) подлинных подписей человека в двумерном признаковом пространстве. Она отличается от известных тем, что обучение классификатора выполняется в пространстве, где представлены образы пар подписей в виде коэффициентов корреляции между описанными выше признаками для всех пар подлинных подписей. Это позволяет увеличить число образов для обучения с N до $N(N-1)/2$, что приближает точность процедуры верификации к 100 % [5–А, 17–А].

Рекомендации по практическому использованию результатов

Результаты, полученные в диссертации, могут быть использованы при создании систем контроля подлинности рукописной подписи, при экспертизе документов специалистами банков, Государственного комитета судебных экспертиз Республики Беларусь и других заинтересованных организаций.

В приложении приведены документы об использовании результатов диссертационной работы.

СПИСОК ПУБЛИКАЦИЙ СОИСКАТЕЛЯ

Статьи в рецензируемых научных журналах

1–А Ахунджанов У. Ю., Старовойтов В. В. Offline верификация рукописной подписи с применением сверточной нейронной сети // Системный анализ и прикладная информатика. – 2022. – № 1. – С. 12–18.

2–А Ахунджанов У. Ю., Старовойтов В. В. Предварительная обработка изображений рукописных подписей для последующего распознавания // Системный анализ и прикладная информатика. – 2022. – № 2. – С. 4–9.

3–А Старовойтов В. В., Ахунджанов У. Ю. Новый признак для описания изображений рукописной подписи на базе локальных бинарных шаблонов // Информатика. – 2022. – Т. 19, № 3. – С. 62–73.

4–А Ахунджанов У. Ю., Старовойтов В. В. Об инвариантности цифрового описания рукописной подписи // Системный анализ и прикладная информатика. – 2022. – № 4. – С. 47–55.

5–А Старовойтов В. В., Ахунджанов У. Ю. Распределение значений локальной кривизны как структурный признак для offline верификации рукописной подписи // Системный анализ и прикладная информатика. – 2023. – № 2. – С. 49–58.

Статьи в сборниках научных трудов, материалов конференций и семинаров

6–А Ахунджанов У. Ю. Актуальные задачи в области статического распознавания рукописных подписей // Актуальные проблемы и инновационные технологии в области естественных наук : Междунар. науч.-практ. online конф., Ташкент, 20-21 нояб. 2020 г. / Таш-й гос. техн. ун-т. имени И. Каримова ; редкол.: С. М. Турабджанов [и др.]. – Ташкент, 2020. – С. 482–484.

7–А Ахунджанов У. Ю. Разработка методики и алгоритма верификации рукописной подписи в режиме offline // Наука и инновации : сб. науч. тр. по материалам Междунар. конф., Ташкент, 26 нояб. 2020 г. / Министерство инновационного развития Респ. Узбекистан, Центр передовых технологий ; редкол.: Ш. У. Турдикулова [и др.]. – Ташкент, 2020. – Ч. 2. – С. 476–478.

8–А Akhundjanov U. Yu. Recognition of static (offline) handwritten

signatures based on machine learning // Наука и инновации : сб. науч. тр. по материалам Междунар. конф., Ташкент, 25 нояб. 2021 г. / Министерство инновационного развития Респ. Узбекистан, Центр передовых технологий ; редкол.: Ш. У. Турдикулова [и др.]. –Ташкент, 2021. – С. 512–514.

9–А Ахунджанов У. Ю. Проблемы рукописных подписей в системах доступа // Наука и инновации : сб. науч. тр. по материалам Междунар. конф., Ташкент, 25 нояб. 2021 г. / Министерство инновационного развития Респ. Узбекистан, Центр передовых технологий ; редкол.: Ш. У. Турдикулова [и др.]. – Ташкент, 2021. – С. 550–552.

10–А Akhundjanov U. Yu., Starovoitov V. V. Problems of biometric identification in access systems // САПР и моделирование в современной электронике : сб. науч. тр. V Междунар. науч.-практ. конф., Брянск, 21–22 окт. 2021 г. / под общ. ред. А. Ю. Дракина. – Брянск, 2021. – С. 69–72.

11–А Ахунджанов У. Ю., Старовойтов В. В. Предварительная обработка изображений рукописных подписей для систем распознавания // 58-я науч. конф. аспирантов, магистрантов и студентов БГУИР, Минск, 18–22 апр. 2022 г. / Белорус. гос. ун-т информатики и радиоэлектроники. – Минск, 2022. – С. 201–202.

12–А Ахунджанов У. Ю. Нейросетевой подход для верификации рукописной подписи в режиме offline // Компьютерные технологии и анализ данных (СТДА'2022) : материалы III Междунар. науч.-практ. конф., Минск, 21–22 апр. 2022 г. / Белорус. гос. ун-т ; редкол.: В. В. Скакун [и др.]. – Минск, 2022. – С. 18–21.

13–А Akhundjanov U. Yu., Starovoitov V. V. Static signature verification based on machine learning // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сб. научных статей VIII Междунар. науч.-практ. конф., Минск, 11–12 мая 2022 г. / редкол.: В. А. Богуш [и др.]. – Минск, 2022. – С. 44–50.

14–А Ахунджанов У. Ю., Старовойтов В. В. Новый тип признаков для описания изображений рукописной подписи на базе локальных бинарных шаблонов // Технические средства защиты информации : тез. докл. XX Белорус.-Рос. науч.-техн. конф., Минск, 7 июня 2022 г. / Белорус. гос. ун-т информатики и радиоэлектроники ; редкол.: Т. В. Борботько [и др.]. – Минск, 2022. – С. 18–19.

15–А Ахунджанов У. Ю., Старовойтов В. В. Экспериментальное исследование инвариантного представления рукописной подписи // Информационные технологии в образовании, науке и производстве : материалы науч.-техн. интернет-конф., Минск, 21–22 ноября 2022 г. / Белорус. нац. техн. ун-т ; редкол.: И. А. Сатиков [и др.]. – Минск, 2022. – С. 464–468.

16–А Старовойтов В. В., Ахунджанов У. Ю. Экспериментальная программа для верификации рукописной подписи в режиме offline // III Междунар. науч.-прак. конф. «Судебная экспертиза: теория и практика в современных условиях», Минск, 26–27 апреля 2023 г. / Гос. ком. судеб. экспертиз Респ. Беларусь ; редкол.: А. А. Волков [и др.]. – Минск, 2023. – С. 72–76.

17–А Верлыго Д. Э., Ахунджанов У. Ю., Старовойтов В. В. Верификация рукописной подписи с помощью классификатора OCSVM // Наука молодых – наука будущего : сборник статей V Междунар. науч.-прак. конф. Петрозаводск, 5 июня 2023 г. / МЦНП «НОВАЯ НАУКА», Российская Федерация ; редкол.: Б. Т. Аймурзина [и др.]. – Петрозаводск, 2023. – С. 246–250.



РЭЗІЮМЭ

Ахунджанаў Уміджон Юнус углі

Верыфікацыя рукапіснага подпісу ў статычным рэжыме

Ключавыя словы: подпіс, верыфікацыя подпісу, папярэдняя апрацоўка відарысаў, структурныя прыкметы подпісу, класіфікатар.

Мэта работы: аўтаматызацыя працэсу верыфікацыі рукапіснага подпісу, прадстаўленага на папяровым носьбіце.

Метады даследавання: метады лічбавай апрацоўкі відарысаў, вылічэнні прыкмет, пабудовы класіфікатара.

Атрыманья вынікі і іх навізна:

методыка папярэдняй апрацоўкі лічбавых відарысаў адвольных рукапісных подпісаў, якая не залежыць ад іх выканання (памеру, колеру, тыпу ручкі, нахілу і павароту подпісу);

новая прыкмета апісання лакальных структурных асаблівасцяў подпісу ўласцівых канкрэтнаму чалавеку ў выглядзе масіва, які апісвае частотнае размеркаванне значэнняў лакальных бінарных шаблонаў (LBP) контураў подпісу, прапанаваны ўпершыню;

новая прыкмета апісання подпісу чалавека ў выглядзе масіва, які апісвае частотнае размеркаванне значэнняў лакальнай крывізны контураў подпісу, прапанаваны ўпершыню;

новая прыкметная прастора, якая апісвае не выявы асобных подпісаў, а карэляцыю пар подпісаў на базе прапанаваных прыкмет. Такое ўяўленне прыкмет дазваляе паменшыць памернасць прыкметнай прасторы да двух і павялічыць колькасць выяў сапраўдных подпісаў да $N(N-1)/2$. У якасці класіфікатара прапанавана выкарыстоўваць аднокласавую машыну апорных вектараў (OC-SVM), навучаную на карэляцыі пар N сапраўдных подпісаў чалавека.

Комплекс прапанаваных у працы алгарытмаў, прыкмет і метадаў верыфікацыі дазволілі пры наяўнасці 15 сапраўдных подпісаў чалавека дасягнуць 99,77 % дакладнасці пры верыфікацыі ўсіх 2640 подпісаў агульнадаступнай базы CEDAR. Найлепшы апублікаваны вынік верыфікацыі подпісаў гэтай базы складае 94,40 %.

Рэкамендацыі па выкарыстанні: вынікі былі прыменены пры выкананні Дзяржаўнай праграмы навуковых даследаванняў "Лічбавыя і касмічныя тэхналогіі, бяспека чалавека, грамадства і дзяржавы", № дзяржрэгістрацыі 20210776, перыяд выканання 2021-2023 гг.

Галіна прымянення: сістэмы кантролю сапраўднасці подпісу.

РЕЗЮМЕ

Ахунджанов Умиджон Юнус угли

Верификация рукописной подписи в статическом режиме

Ключевые слова: подпись, верификация подписи, предварительная обработка изображений, структурные признаки подписи, классификатор.

Цель работы: автоматизация процесса верификации рукописной подписи, представленной на бумажном носителе.

Методы исследования: методы цифровой обработки изображений, вычисления признаков, построения классификатора.

Полученные результаты и их новизна:

методика предварительной обработки цифровых изображений произвольных рукописных подписей, не зависящая от их исполнения (размера, цвета, типа ручки, наклона и поворота подписи);

новый признак описания локальных структурных особенностей подписи, присущих конкретному человеку, в виде массива, описывающего частотное распределение значений локальных бинарных шаблонов (LBP) контуров подписи, предложен впервые;

новый признак описания подписи человека в виде массива, описывающего частотное распределение значений локальной кривизны контуров подписи, предложен впервые;

новое признаковое пространство, описывающее не образы отдельных подписей, а корреляцию пар подписей на базе предложенных признаков. Такое представление признаков позволяет уменьшить размерность признакового пространства до двух и увеличить количество образов подлинных подписей до $N(N-1)/2$. В качестве классификатора предложено использовать одноклассовую машину опорных векторов (OC-SVM), обученную на корреляции пар N подлинных подписей человека.

Комплекс предложенных в работе алгоритмов, признаков и метод верификации позволили при наличии 15 подлинных подписей человека достичь 99,77 % точности при верификации всех 2640 подписей общедоступной базы CEDAR. Лучший опубликованный результат верификации подписей этой базы составляет 94,40 %.

Рекомендации по использованию: результаты были применены при выполнении Государственной программы научных исследований «Цифровые и космические технологии, безопасность человека, общества и государства», № госрегистрации 20210776, период выполнения 2021-2023 гг.

Область применения: системы контроля подлинности подписи.

SUMMARY

Akhundjanov Umidjon Yunus ugli

Verification of a handwritten signature in static mode

Keywords: signature, signature verification, image preprocessing, signature structural features, classifier.

Purpose of work: automation of the process of verification of a handwritten signature presented on paper.

Research methods: methods of digital image processing, feature calculation, classifier construction.

The results obtained and their novelty:

a technique for pre-processing digital images of arbitrary handwritten signatures that does not depend on their execution (size, color, pen type, inclination and rotation of the signature);

a new feature for describing local structural features of a signature inherent in a particular person in the form of an array describing the frequency distribution of the values of local binary patterns (LBP) of signature contours, proposed for the first time;

a new feature for describing a person's signature in the form of an array that describes the frequency distribution of the values of the local curvature of the signature contours, proposed for the first time;

a new feature space that describes not the images of individual signatures, but the correlation of pairs of signatures based on the proposed features. Such a representation of the features allows you to reduce the dimension of the feature space to two and increase the number of images of genuine signatures to $N(N-1)/2$. As a classifier, it is proposed to use a one-class support vector machine (OC-SVM) trained on the correlation of pairs of N authentic human signatures.

The set of algorithms, features and verification method proposed in the work made it possible, in the presence of 15 authentic human signatures, to achieve 99,77 % accuracy in verifying all 2640 signatures of the public CEDAR database. The best published signature verification result for this database is 94,40 %.

Recommendations for use: the results were applied in the implementation of the State Research Program "Digital and space technologies, security of man, society and the state", state registration number 20210776, implementation period 2021-2023.

The scope: signature authentication systems.

Научное издание

Ахунджанов Умиджон Юнус угли

**ВЕРИФИКАЦИЯ РУКОПИСНОЙ ПОДПИСИ
В СТАТИЧЕСКОМ РЕЖИМЕ**

Автореферат
диссертации на соискание ученой степени
кандидата технических наук

по специальности 05.13.01 – Системный анализ, управление и обработка
информации

Подписано в печать 12.09.2023. Формат 60×84 1/16. Бумага офсетная. Гарнитура «Таймс».
Отпечатано на ризографе. Усл. печ. л. 1,63. Уч.-изд. л. 1,4. Тираж 60 экз. Заказ 217.

Издатель и полиграфическое исполнение: учреждение образования
«Белорусский государственный университет информатики и радиоэлектроники».
Свидетельство о государственной регистрации издателя, изготовителя,
распространителя печатных изданий №1/238 от 24.03.2014,
№2/113 от 07.04.2014, №3/615 от 07.04.2014.
Ул. П. Бровки, 6, 220013, г. Минск