

ОДНОКЛАССОВЫЙ КЛАССИФИКАТОР ДЛЯ РАСПОЗНАВАНИЯ ПОДДЕЛЬНЫХ ПОДПИСЕЙ ПРИ МИНИМАЛЬНОМ ЧИСЛЕ ОБРАЗЦОВ

В. В. Старовойтов

*Объединённый институт проблем информатики НАН Беларуси,
Минск, Беларусь, valerystar@mail.ru*

В работе описывается метод построения классификатора, предназначенного для верификации статических подписей человека, который строится на малом числе N (от 2 до 11) образцов подлинных подписей этого человека. Метод базируется на сравнении гистограмм LBP-признаков полутоновых изображений подписей. Вычисляются значения ранговой корреляции Спирмена между парами гистограмм всех образцов подлинных подписей, они сравниваются с аналогичными значениями, вычисленными для N пар, образованных исследуемой подписью и заданными образцами. Предложенный метод верификации на базе CEDAR (2640 подписей 55 человек, половина поддельных) при использовании всего двух образцов подлинных подписей человека достиг 100% точности корректного распознавания подлинных и поддельных подписей.

Ключевые слова: одноклассовый классификатор; статическая подпись; цифровое изображение.

ONE-CLASS CLASSIFIER FOR FORGED SIGNATURE RECOGNITION WITH MINIMAL NUMBER OF SAMPLES

V. V. Starovoirov

*The United Institute of Informatics Problems,
Minsk, Belarus, valerystar@mail.ru*

The paper describes a method for constructing a classifier designed to verify offline signatures, which is built on a small number N (from 2 to 11) of samples of genuine signatures of this person. The method is based on comparing histograms of LBP features of grayscale signature images. The values of the Spearman rank correlation between pairs of histograms of all samples of genuine signatures are calculated, they are compared with similar values calculated for N pairs formed by the signature under study and the genuine samples. The proposed verification method on dataset CEDAR (2640 signatures of 55 people, half of them are fake) using only two samples of genuine human signatures achieved 100% accuracy in correct recognition genuine and fake signatures.

Keywords: one-class classifier; static signature; digital image.

1. Введение

Подпись – это вид рукописи, имеющий удостоверительное назначение. Она состоит из нескольких букв, условных знаков и безбуквенных штрихов. Подпись, выполненная на бумаге, называется статической или offline подписью. Она является одним из реквизитов документа и ее иногда подделывают. Подделки могут иметь разную степень сходства с оригиналом. Задача исследования подлинности подписи называется верификацией, ее, как правило, решает квалифицированный эксперт. Автоматизация процесса верификации рукописной подписи является актуальной задачей [1]. У каждого человека имеются свои особенности выполнения подписи. Совокупность подлинных подписей, выполненных одним человеком, представляет собой единственный класс данных, к которому или относятся верифицируемые подписи, или нет, в этом случае это подделки либо просто подписи других людей.

2. Постановка задачи верификации

Дано N подлинных подписей человека и еще одна подлинность которой нужно оценить посредством сравнения с истинно подлинными подписями. N обычно не превышает двух десятков.

В отличие от задачи распознавания символов, где требуется символ отнести к одному из заданных классов, все подлинные подписи, выполненные одним человеком, имеют визуальные различия (рис. 1), а поддельная подпись, наоборот, может быть очень похожа на подлинную (рис. 2).



Рис. 1. Подлинные подписи одного человека.
Все подписи имеют заметные визуальные отличия

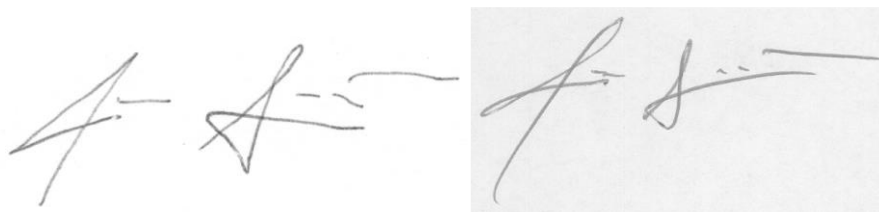


Рис. 2. Слева поддельная подпись, справа – подлинная этого же человека

Искусственные нейронные сети не могут быть использованы для решения задачи автоматической верификации подписи человека поскольку 1) эксперту, как правило, предоставлено мало подлинных подписей, 2) у эксперта не может быть образцов поддельных подписей, 3) для обучения сети необходим минимум еще один класс, т.е. сети не строят одноклассовые классификаторы. Многие исследователи в качестве альтернативного класса используют случайно выбранные подлинные подписи других людей, называя их случайными подделками. Однако это не корректно. Случайно выбранные подписи других людей легко распознаются.

В работе [2] нами был предложен метод построения одноклассового классификатора на базе машины опорных векторов. В качестве признаков использовались локальная кривизна и локальный бинарный код, вычисленные по контуру бинарного представления нормализованного изображения подписи.

3. LBP код цифрового изображения

В данной работе представлен новый метод верификации, базирующийся на вычислении локального бинарного шаблона (LBP, Local Binary Pattern) [3], вычисляемого на полутоновом изображении подписи.

Локальные бинарные шаблоны были предложены для описания текстуры изображения в работе [3]. Обычно этот код вычисляется на полутоновых изображениях в окрестности каждого пиксела. На рис. 3. представлена идея вычисления LBP кода пиксела с яркостью 54.

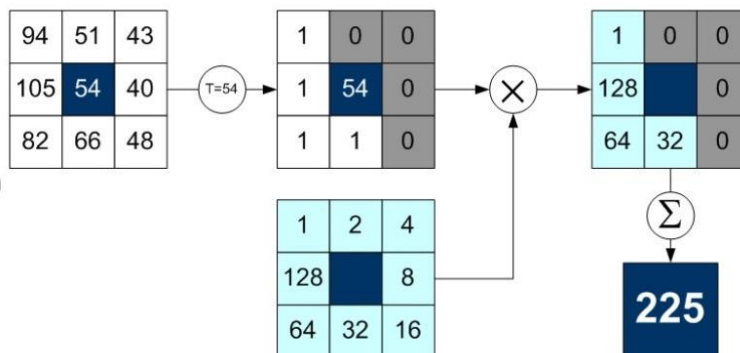


Рис. 3. Идея вычисления LBP-признака пиксела полутонового изображения [4]

В данном примере выбирается 8 соседних пикселей, расположенных от центрального на расстоянии 1; сравниваются их яркости; тем соседним пикселям у которых яркость выше чем у центрального присваивается 1, остальным – 0; затем выполняется свертка с маской, состоящей из степеней двойки и суммирование. Значение 225 является LBP кодом, присваиваемым пикселу с яркостью 54. Возможны варианты вычисления кода LBP: радиус может равняться 2 или 3; окрестности, в которых выбираются соседние пиксели могут иметь разную форму; число соседей, участвующих в вычислении кода может варьироваться от 4 до 24. По матрице LBP кодов, вычисленных для каждого пикселя изображения, строится гистограмма распределения этих кодов на изображении, которая часто используется как текстурный признак изображений разных классов.

4. Инвариантность LBP кодирования к яркости изображения

Из правила построения LBP кода для отдельного пикселя вытекает инвариантность кода к линейным и нелинейным изменениям яркости исходного изображения. На рис. 4 представлено изображение подлинной подписи с рис. 2 и три варианта его яркостной модификации: гамма-коррекция с параметрами 1,5 и 0,75, а на крайнем правом изображении все значения яркости уменьшены на 25%.



Рис. 4. Исходное изображение подписи и три варианта его модификации

При построении гистограммы LBP кодов этих изображений с малым числом интервалов (бинов), форма гистограммы может незначительно изменяться. На рис. 5 представлены гистограммы LBP кодов изображений, представленных на рис. 4, а также гистограмма поддельного изображения с рис.2 инвертированная для более наглядного представления отличий ее формы. Ранговая корреляция между парами LBP гистограмм, вычисленных для изображений подлинных подписей на рис. 4, составляет от 0,9421 до 0,9959, а для пары изображений подлинная и поддельная – 0,0920.

5. Инвариантность LBP кодирования к повороту изображения

Рассмотрим изображение подлинной подписи с рис. 2 и его повернутые копии (рис. 6). Отметим, что при повороте изображений используется интерполяция яркости ряда пикселей, а изображение становится больше размером.

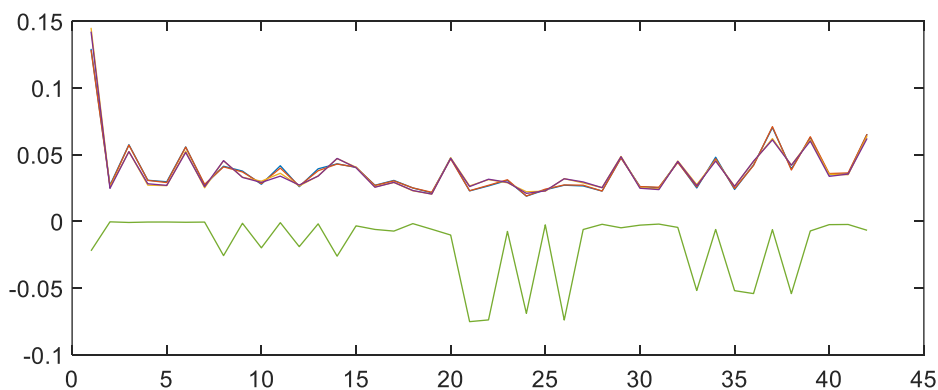


Рис. 5. Сверху – гистограммы LBP кодов четырех изображений с рис. 4, внизу представлена инвертированная гистограмма поддельного изображения с рис. 2

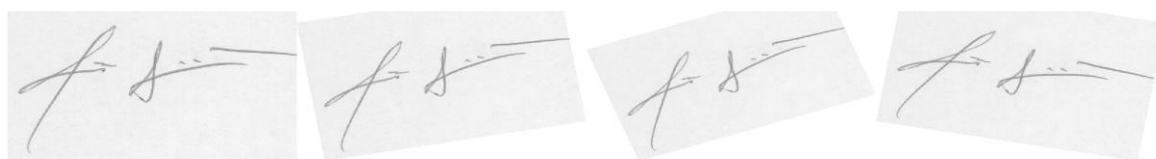


Рис. 6. Слева направо: исходное изображение, повернутое на $+10^\circ$, повернутое на $+20^\circ$, повернутое на -10°

На рис. 7 приведены LBP-гистограммы исходного изображения подлинной подписи и трех ее вариантов, повернутых на $+10$, $+20$, -10 градусов. Гистограммы подписи, повернутой на $+10^\circ$ и -10° почти совпадают. Ранговая корреляция между ними равна 0.9880, корреляция гистограмм исходного изображения подписи и повернутых равна 0,9690, 0,9298, 0,9757, соответственно, т.е. достаточно высока.

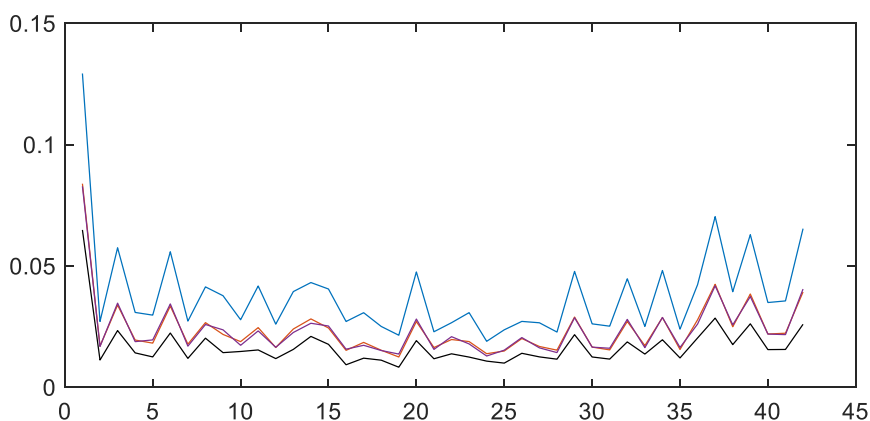


Рис. 7. LBP гистограммы исходного изображения подписи и трех повернутых на $+10$, $+20$, -10 градусов

Таким образом, небольшие различия в яркости изображений подписи и его поворот до 10° несущественно уменьшает корреляционный признак подобия LBP гистограмм изображений подписей.

Обозначения на рис. 8: model – коэффициенты корреляции между парами восьми подлинных подписей, использованных для построения модели классификатора подписей человека (рис. 1); real – корреляция между верифицируемой подлинной подписью и восьмью модельными; fake – корреляция между верифицируемой поддельной подписью и восьмью модельными, увеличенная на +0.8 для удобства представления.

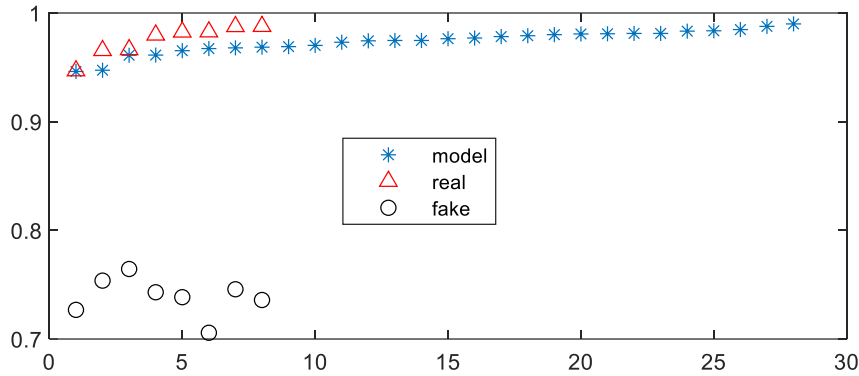


Рис. 8. Коэффициенты корреляции между LBP гистограммами изображений подлинной и поддельной подписей, представленных на рис. 2

Несмотря на видимые различия подлинных подписей и сходство с ними поддельной, коэффициенты корреляции, вычисленные для верифицируемой подлинной и поддельной подписи, легко разделяются по порогу.

6. Новый метод верификации подписи

Метод верификации рукописной подписи состоит из нескольких этапов:

1. Вычислить LBP-коды N подлинных подписей и верифицируемой подписи.
2. Построить гистограммы их распределения.
3. Вычислить $N \times (N - 1) / 2$ коэффициентов ранговой корреляции между парами всех гистограмм заданных подлинных подписей (массив $M1$).
4. Вычислить N коэффициентов ранговой корреляции между парами, образованными верифицируемой подписью и N заданных подлинных (массив $M2$).
5. Оценить: медианное значение массива $M2$ больше наименьшего числа в массиве $M1$? Если да, верифицируемая подпись считается подлинной, иначе – поддельной.

7. Тестирование предложенного метода

Тестирование предложенного метода было выполнено на общедоступной и часто используемой для сравнительных тестов базе CEDAR [5]. Она содержит изображения подписей 55 человек, по 24 подлинных и 24

поддельных каждого. Подписи оцифрованы с разрешением 300 dpi и их полутоновые представления записаны в формате png. Общее число подписей равно 2640.

Эксперименты показали, что корректное распознавание подлинных подписей из базы CEDAR возможно даже при построении классификатора на базе двух-трех подлинных модельных подписей. При значениях N от 2 до 11 точность верификации составила 100%. На рис. 9 приведены примеры восьми подлинных подписей из этой базы имеющие видимые различия.

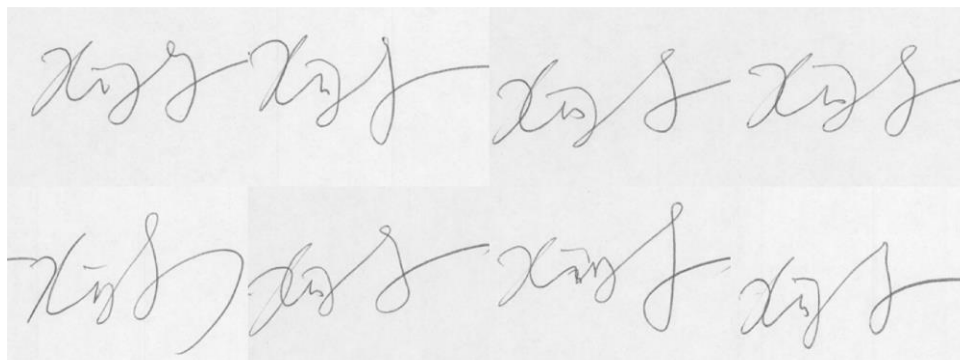


Рис. 9. Восемь подлинных подписей. Две первых использованы при построении классификатора подлинности подписей этого человека

На рис. 10 слева показано изображение поддельной подписи этого же человека, а справа – подлинной. Субъективно кажется, что справа поддельная подпись. При автоматической классификации даже при $N = 2$ все подписи этого человека были распознаны верно.

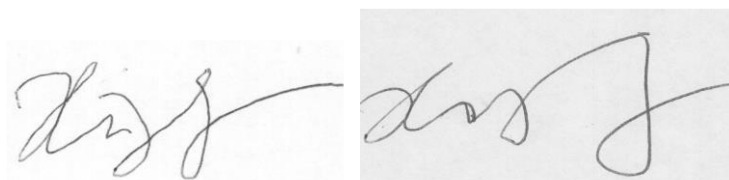


Рис. 10. Слева поддельная подпись, справа – подлинная

8. Заключение

В работе описан новый метод верификации полутоновых изображений рукописных подписей. Его особенностью является то, что классификатор строится индивидуально для распознавания подлинности подписей конкретного человека. Для его построения может использоваться минимальное число подлинных подписей этого человека (для базы CEDAR достаточно двух подписей). LBP признаки, используемые в классификаторе, инвариантны к повороту подписи на угол до 10° и к изменению яркости

фона изображения. Построенный классификатор является одноклассовым, т.е. он распознает только подписи конкретного человека, поддельные подписи и подписи других людей идентифицируются как выбросы из данного класса. Такой классификатор не зависит от языка написания подписи и не требует подписей других людей для обучения.

Классификатор, ориентированный на выявление особенностей исполнения подписи конкретным человеком, обученный на двух подлинных подписях и достигающий 100% точности при верификации всех подписей базы CEDAR, до настоящего времени в научной литературе не описывался.

Библиографические ссылки

1. *Singla A, Mittal A*. Exploring offline signature verification techniques: A survey based on methods and future directions // *Multimedia Tools and Applications*. 2025. Vol. 84, iss. 6. P. 2835–2875.

2. *Starovoitov V. V., Akhundjanov U. Y.* A writer-dependent approach to offline signature verification based on one-class support vector machine // *Pattern Recognition and Image Analysis*. 2024. Vol. 34, iss. 2. P. 340–351.

3. *Ojala T., Pietikainen V., Harwood D.* Performance evaluation of texture measures with classification based on Kullback discrimination of distributions // *Proc. 12th Int. Conf. on Pattern Recognition*. 1994. Vol. 1. P. 582–585.

4. *Prakasa E.* Texture feature extraction by using local binary pattern // *INKOM Journal*. 2016. Vol. 9, iss. 2. P. 45–48.

5. CEDAR dataset. URL: <https://cdar.buffalo.edu/NIJ/data/sintures.rar> (date of access: 05.09.2025).