

Алгоритм выделения экстремальных точек применительно к задаче биометрической верификации рукописной подписи

Колядин Д.В. (dvkmail2000@mail.ru), Петров И.Б.

Московский физико-технический институт
(государственный университет)

Введение

В научной литературе принято отсчитывать начало активных исследований в области верификации рукописной подписи с работы [1], в которой были сформулированные основные направления и подходы к решению данной проблемы. Хотя работы в области применения ЭВМ к криминалистической экспертизе почерка проводились и ранее, но развитие компьютерных технологий сделало возможным проведение анализа динамики начертательных движений возникающих при исполнении рукописной подписи. Специальные графические планшеты позволяют отслеживать траекторию движения пера, а также угол наклона ручки и оказываемое на перо давление. В англоязычной литературе этот подход носит название «on-line» метода распознавания, в отличие от анализа отсканированных изображений или «off-line» метода распознавания. Именно возможность измерения динамики делает привлекательным использование рукописной подписи как биометрического параметра, удобного для применения в широком спектре информационных систем. В настоящий момент проводятся исследования по данной проблематике различными исследовательскими группами. Первый международный конкурс программ по верификации рукописной подписи: SVC 2004: First International Signature Verification Competition состоялся в начале 2004 года. В этом конкурсе принимали участие команды из девяти стран: Австралии, Германии, Испании, Китая, Кореи, Сингапура, США, Турции, Франции.

Развиваемый авторами подход относится к категории функциональных методов распознавания и базируется на процедуре эластичного сопоставления методом динамического программирования. В данной статье обсуждается этап предварительной обработки «on-line» кривых, присутствующий в разработанном авторами программном продукте.

Постановка задачи

Под динамической (или «on-line») кривой подразумевается параметрически заданная кривая $\Gamma : \{x(t), y(t), t \in [0, T]\}$. Информация, содержащаяся в данной временной последовательности, отражает динамику мускульных движений руки и поэтому может быть использована как биометрическая характеристика человека. В настоящее время общепринятой является такая схема построения систем доступа, в которых пользователь указывает свой идентификатор, подтверждаемый затем рукописной подписью. Предлагаемый подход позволяет пользователю предоставлять системе только свою подпись, по которой программа находит эталон (проводит идентификацию), а затем проверяет подлинность предъявленной подписи (осуществляет верификацию) [2]. Такой метод биометрической идентификации, очевидно, обладает большей функциональностью и имеет более широкую область применения. Для решения задачи поиска эталона в данной работе предлагается описывать динамическую кривую набором экстремальных точек. Строго говоря, динамические кривые могут быть параметризованы различными способами. Поскольку подавляющее большинство исследователей занимается только

задачей верификации, то наиболее распространенной является временная параметризация. Однако помимо временной параметризации, возможна параметризация по длине траектории: $\left| \frac{\partial \Gamma}{\partial s} \right| = 1$, что означает постоянство скорости перемещения вдоль траектории.

Также допустима аффинная параметризация по длине траектории, удовлетворяющая условию: $\left| \frac{\partial \Gamma}{\partial s} \times \frac{\partial^2 \Gamma}{\partial s^2} \right| = 1$, что означает постоянство площади параллелограмма

построенного на векторах $\frac{\partial \Gamma}{\partial s}$ и $\frac{\partial^2 \Gamma}{\partial s^2}$, где параметр s находится из условия:

$s(t) = \int_0^t \left| \frac{\partial \Gamma}{\partial \tau} \times \frac{\partial^2 \Gamma}{\partial \tau^2} \right|^{\frac{1}{3}} d\tau$ [3]. Вообще говоря, динамические кривые содержат информацию,

как о динамике движения пера, так и о временной последовательности точек, составляющих траекторию. Поскольку в предлагаемом подходе на этапе идентификации (поиска эталона) нас интересует только форма кривой, то далее под термином динамической кривой договоримся понимать параметризованную по длине кривую

$\Gamma : \{x(s), y(s), s \in [0, S_\Gamma]\}$, где параметр s находится из условия: $s(t) = \int_0^t \left\| \frac{\partial \Gamma}{\partial \tau} \right\| d\tau$.

Отметим, что скорость написания рукописных кривых (т.е. иными словами «динамика» движения руки) тесно связана с формой траектории. Действительно, при выполнении закруглений скорость пера ниже, чем при выполнении прямолинейных движений. Также темп выполнения элементов подписи варьируется от одного исполнения к другому, что иногда сильно осложняет проверку подлинности. Исходя из подобных соображений В. Налва в своей работе [4], одним из немногих отказался от параметризации по времени, и принципиально рассматривает только форму траектории, что приводит к отказу от рассмотрения таких характеристик как, например, нажим на перо или наклон пишущей ручки. Поэтому в силу специфичности постановки задачи и приведенных доводов мы отказались от временной параметризации, тем более что, как показано в работе [2], она не имеет явных преимуществ перед другими видами параметризации, а иногда и проигрывает им.

Сопоставление траекторий на основе экстремальных точек

Распознавание динамических кривых очень близко к тематике распознавания слитной человеческой речи. Данные, получаемые с графического планшета, похожи на звуковой сигнал. Как и фонемы при распознавании речи, объекты распознавания рукописного текста (буквы и символы) четко заданы. Форма рукописных знаков зависит от положения в слове, что идейно близко к вариативности фонем при произнесении слов в зависимости от их окружения. Все это позволило адаптировать разработанные методы из области распознавания речи к распознаванию рукописного текста [5,6]. Тем не менее, прямое применение этих методов к распознаванию слитного текста сопряжено с определенными трудностями, связанными с

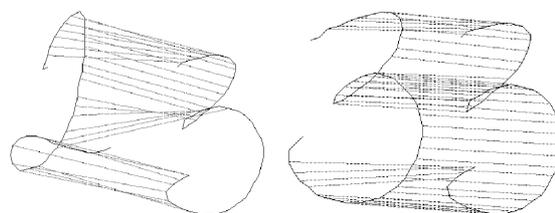


Рис. 1. Пример сопоставления двух рукописных цифр методом dynamic time warping [13].

нахождением диакритических знаков [7]. В тоже время слова в письменной речи часто удаётся достаточно легко отделять одно от другого, в отличие от звуков в слитной речи.

Основными методами распознавания рукописной подписи являются: статистический подход, основанный на аппарате скрытых марковских моделей, подход с использованием нейронных сетей, а также сопоставление динамических кривых методом динамической трансформации временной шкалы (dynamic time warping). В зависимости от принципов сопоставления элементов траекторий в литературе встречаются различные вариации последнего метода:

- Symmetric Dynamic Time Warping [8],
- Continuous Dynamic Time Warping [9],
- Piecewise Aggregate Approximation Dynamic Time Warping [10],
- Iterative Deepening Dynamic Time Warping [11],

и т.д.

В большинстве работ сопоставлению подвергаются все точки траектории.

Например, в работе Ф. Грисса с коллегами [12] был применен метод сравнения траекторий, апробированный ранее в работах С. Коннелла при распознавании рукописных символов [13] (см. рис.1). Тем не менее, как было отмечено в работе [14], метод DTW обладает двумя основными недостатками: высокой вычислительной стоимостью и приведением поддельной подписи в соответствие с эталоном. Для устранения указанных недостатков в работе [14] был предложен метод сравнения подписей на основе поиска соответствия экстремальных точек (extreme points warping, EPW). Исходный сигнал авторы цитируемой работы представляют как череду «пиков» и «впадин» (см. рис. 2). Поскольку при таком подходе соответствие необходимо находить только между точками соответствующего типа, то процесс сравнения существенно ускоряется. На основе образцов подписей 25 пользователей Ф.Хао и К.Чан сделали эмпирическое обобщение, что ошибки сопоставления бывают трех типов: из-за пропуска сегмента в начале или в конце одной из подписей, а также из-за «зубчиков» – случайной пары «пик-впадина» с небольшим перепадом по высоте. Для исключения из рассмотрения подобных точек вводится эвристическое правило: точка помечается как экстремальная, если выполнены одновременно два условия:

$$r \geq h_0, d \geq h_0$$

где r – перепад высот между последовательными «впадиной» и «пиком», d – перепад высот между «пиком» и последующей «впадиной», h_0 – пороговое значение (равное одному пикселю). Полученные экстремальные точки сопоставлялись методом динамического программирования [13,14]. На рис.3. показана схема прямого хода, когда путь из точки (i,j) может быть продолжен в ту из трех точек $(i+1,j+1)$, $(i+1,j+3)$, $(i+3,j+1)$ которая дает минимальный штраф за несоответствие. Прямой проход позволяет найти

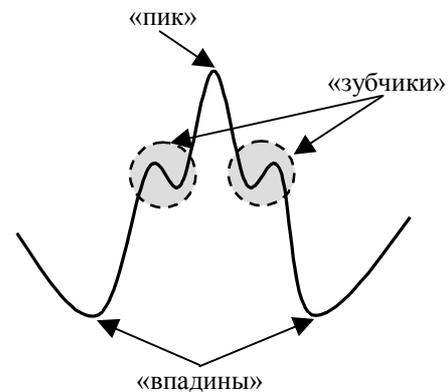


Рис. 2. Вид неустойчивых экстремальных точек в соответствии с работой

Рис. 2. Вид неустойчивых экстремальных точек в соответствии с работой

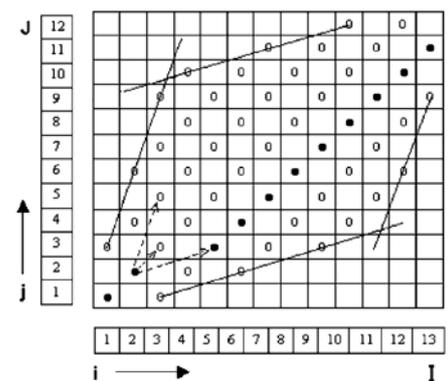


Рис.3. Реализация EPW [6].

меру различия двух кривых, при этом последовательность точек (i, j) задает путь, каждая точка которого обозначает локально-оптимальное соответствие между i -той точкой эталона (точки эталона занумерованы по оси абсцисс) и j -той точкой образца (точки образца занумерованы по оси ординат). Правила возможного продолжения пути следуют из факта чередования экстремумов, при этом выбирается такое продолжение, которое дает минимальный штраф за несоответствие. По окончании счета в правом верхнем углу таблицы будет содержаться мера различия двух кривых. В случае необходимости обратный проход по уже установленному пути позволяет найти оптимальное соответствие двух кривых.

Аналогичный способ сопоставления кривых был ранее предложен в работе [15]. Однако невозможность учесть некоторые вариативные особенности почерков, а также накопление ошибок в ходе последовательного сопоставления хорд, образуемых соответствующими экстремальными точками, потребовало пересмотреть процедуру отбора и сопоставления экстремальных точек. Ниже мы излагаем улучшенный вариант нахождения экстремальных точек по сравнению с работами [4,14,15].

Выделение экстремальных точек

В той или иной степени многие подходы к верификации рукописной подписи основываются на процедуре нахождения характерных или экстремальных точек. Одной из первых работ в области верификации статической рукописной подписи в Советском Союзе была проделана совместно сотрудниками Вычислительного центра Ленинградского университета и Литовского института судебных экспертиз [16]. В этой работе точки выбирались вручную, после чего их координаты вводились в ЭВМ для обработки. Применение сенсорных панелей и графических планшетов позволяет автоматизировать данную процедуру. Но какие точки кривой являются наиболее информативными? Как правило, выбор падает на точки со значением кривизны больше некоторого порогового значения. Вычисление кривизны по формуле дифференциальной геометрии требует существования второй производной по времени, что накладывает ограничение на гладкость исследуемой траектории.

Поэтому на практике кривизну определяют по скорости пера (для этого требуется только первая производная) как коррелированной величины [8], при этом в качестве порогового значения выбирается значение, составляющее 15-18% от средней скорости написания (см. рис.4). Либо заменяют предел $c = \lim_{\Delta s \rightarrow 0} \frac{\Delta \alpha}{\Delta s}$, где $\Delta \alpha$ – изменение угла

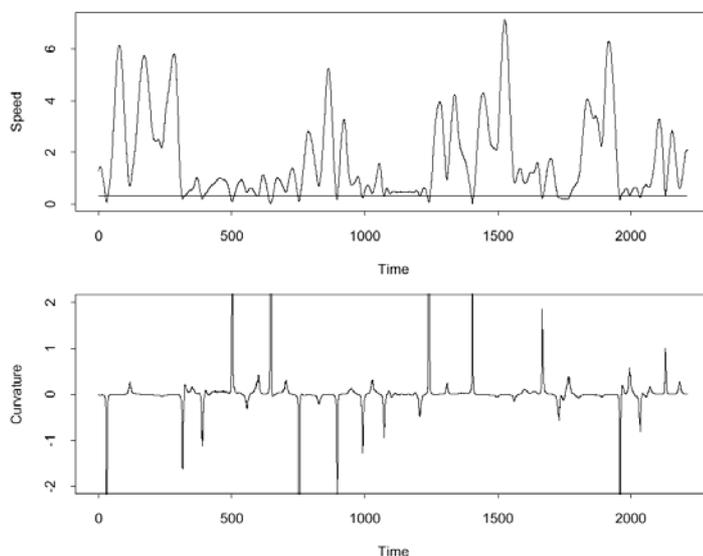


Рис.4. Зависимости кривизны траектории и значения модуля скорости от времени [8].

наклона касательной на протяжении участка траектории от s до Δs , на конечную разность. Кривизна траектории в i -той точке семейства точек $P_{i-2}, P_{i-1}, P_i, P_{i+1}, P_{i+2}$ определяется по углу между векторами $\overrightarrow{P_{i-2}, P_i}$ и $\overrightarrow{P_i, P_{i+2}}$. Как вариант, допустимо заменить функцию кривизны ее дискретным аналогом в виде временной последовательности изменения углов [17].

Но если в случае установления подлинности подписи достаточно сравнить две временные последовательности точек, то в случае распознавания подписи на основе формы и структуры кривых становится необходимым проведение тщательного отбора наиболее информативных точек. Для последующего описания введем определения экстремальных точек, несколько отличные от терминологии предыдущего раздела.

Определение 1. Вертикальным экстремумом называется точка $\{x(t^*), y(t^*)\}$, для которой $\left. \frac{\partial y(t)}{\partial t} \right|_{t=t^*} = 0$. ■

Определение 2. Горизонтальным экстремумом называется точка $\{x(t^*), y(t^*)\}$, для которой $\left. \frac{\partial x(t)}{\partial t} \right|_{t=t^*} = 0$. ■

При этом каждый вертикальный экстремум определяет локальный максимум или минимум траектории. Под экстремальными точками по кривизне будем понимать точки, локальная кривизна траектории в которых больше порогового значения. Можно

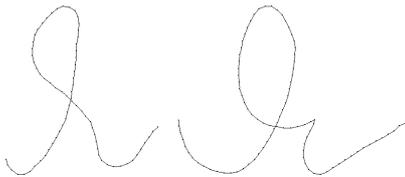


Рис.5. Русские буквы «Я».

утверждать, что выше описанные методы нахождения экстремальных точек не всегда оптимальны, а применительно к обсуждаемому подходу и малоэффективны. Действительно, очень часто вертикальные и горизонтальные экстремумы находятся вблизи от экстремальных точек по кривизне, что делает необходимым привлечение процедуры фильтрации. В тоже

время, задание неправильного порогового значения для данной конкретной подписи может привести к пропуску важных точек, не обладающих большой кривизной. На рис. 5

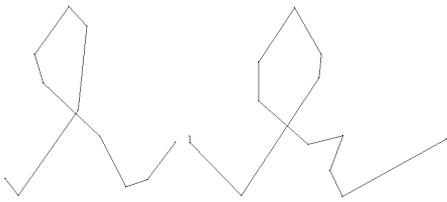


Рис.6. Пример нахождения экстремальных точек.

приводится пример двух русских букв «Я» отличающихся формой траектории. Из-за пропуска точки, характеризующей выпуклость на ниспадающем участке, первая буква может быть ошибочно классифицирована как буква «е». Разработанный способ нахождения экстремальных точек позволяет отслеживать изменения траектории указанного типа, что особенно важно при распознавании рукописных букв и слов,

например, при графическом вводе текста на карманных компьютерах. На рис. 6. мы приводим «скелетное» представление тех же самых букв, полученное путем соединения экстремальных точек. Из этого рисунка видно, что вся необходимая информация о структуре букв была сохранена и передана гораздо меньшим количеством точек. Как было установлено в работе [15] сопоставление всех точек сегментов по евклидовой мере с последующим суммирование всех локальных штрафов может приводить к ошибкам распознавания, когда рукописное слово «Лена» сопоставляется со словом «Леша», т.к. штраф за пропуск лишнего элемента оказывается меньше итогового штрафа за несоответствие двух вариантов слова «Лена».

На рис.7. приведена последовательность шагов, выполняемых для получения набора экстремальных точек. Первые две процедуры относятся к этапу предобработки,

когда исходная траектория сглаживается и производится повторная выборка точек методом интерполяции (resampling) для получения эквидистантного представления кривой. Поскольку вследствие шумов в оцифровке, а также дрожания пера точки располагаются неровно вдоль траектории, то желательно производить сглаживание кривой. Для этой цели был использован достаточно простой, но эффективный метод сглаживания по координатам соседних точек. По сравнению с более распространенным методом сглаживания «по Гауссу» он обладает меньшей вычислительной стоимостью при практически тех же результатах. Усреднение проводилось по двум соседям, отдельно для каждой из координат, по формулам:

$$X_{smoothed}(i) = \frac{1}{4}X(i-1) + \frac{1}{2}X(i) + \frac{1}{4}X(i+1)$$

$$Y_{smoothed}(i) = \frac{1}{4}Y(i-1) + \frac{1}{2}Y(i) + \frac{1}{4}Y(i+1)$$

Полученное представление кривой использовалось для нахождения критических точек. Здесь и далее условимся использовать термин «критические точки» для обозначения предварительно найденных экстремальных точек, т.к. затем часть из этих точек будет отфильтрована (см. рис.7). В отличие от работы [6] мы не исключали из рассмотрения критические точки с малым перепадом по высоте, поскольку последовательность таких точек может образовывать так называемое «плато» (см. рис.8). Аналогичная ситуация возможна и с нахождением горизонтальных критических точек. Поэтому на шаге

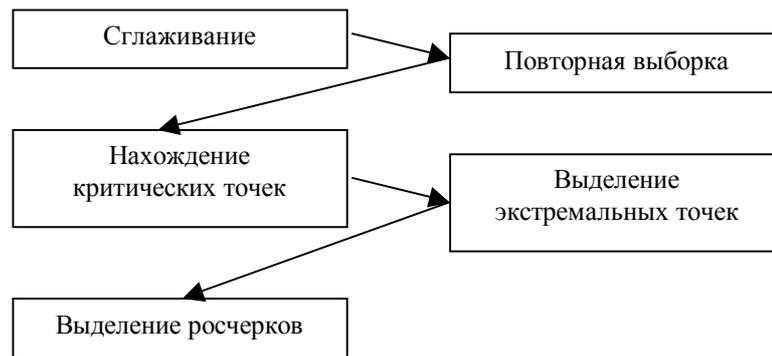


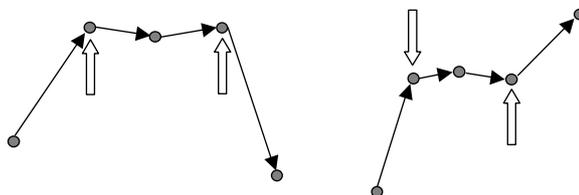
Рис.7. Последовательность этапов поиска экстремальных точек.

нахождения критических точек выделялись следующие типы точек: концевые точки сегментов кривой, вертикальные и горизонтальные экстремумы, граничные точки вертикального плато, а также точки перегиба. При этом для нахождения вертикальных экстремумов применялось простое эвристическое правило: точка P_i является экстремальной, если выполнено условие:

$$(P_{i-1}^y - P_i^y)(P_i^y - P_{i+1}^y) < 0.$$

Если для некоторой i -той точки приведенное выражение обращается в ноль, причем ординаты трех последовательных точек не совпадают, то такая точка добавлялась в набор как граничная точка плато. Аналогично находились и горизонтальные экстремумы (критические точки). Поскольку особое внимание уделялось скорости и эффективности алгоритма, точки перегиба находились не из равенства нулю второй производной, а путем анализа знака векторного произведения $sign[\overrightarrow{P_{i-2}P_i} \times \overrightarrow{P_{i+2}P_i}]$. Последнее выражение в нашем случае можно заменить на упрощенную координатную запись для плоского случая:

$$sign((P_{i-2}^x - P_i^x)(P_{i+2}^y - P_i^y) - (P_{i-2}^y - P_i^y)(P_{i+2}^x - P_i^x))$$



Таким образом, после выполнения данного шага мы получаем набор

Рис.8. Горизонтальные плато.

критических точек, который, строго говоря, является избыточным. В зависимости от параметров программы количество полученных точек в 2-2,5 раза превышает количество окончательно найденных экстремальных точек. Поэтому следующим шагом является удаление лишних точек, а также добавление новых, более точно характеризующих форму траектории. В частности, все точки перегиба подлежат удалению, т.к. располагаются, как правило, в середине «прямых» участков, что снижает их информационную значимость. Вместо этого мы добавляем в набор экстремальные точки, условно называемые экстремумами по кривизне, хотя само значение кривизны в этих точках может быть и не очень велико. В основу процедуры фильтрации положено следующее простое **Утверждение**. *Критические точки, характеризующие горизонтальное плато, выделяются парами.*

Доказательство. Рассмотрим величины $dy_{prev} = P_i^y - P_{i-1}^y$ и $dy_{next} = P_{i+1}^y - P_i^y$. i -я точка принадлежит плато тогда и только тогда, когда $dy_{prev} = 0$ или $dy_{next} = 0$. Пусть при последовательном просмотре точек кривой для i -той точки $dy_{next} = 0$, тогда для $i+1$ точки $dy_{prev} = 0$. Значит i -тая и $(i+1)$ -ая точки принадлежат одному и тому же плато, что и требовалось доказать. ■

Данное утверждение позволяет сформулировать правила отбора для точек плато P_u и P_d , где символ P_d означает, что точка может являться локальным минимумом, а символ

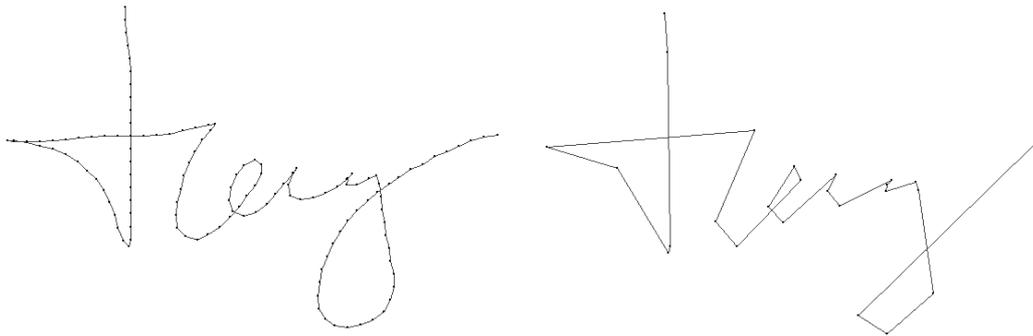


Рис.9. Пример подписи с горизонтальным штрихом.

P_u соответственно означает, что точка может являться локальным максимумом. В этих обозначениях на рис.8 представлены варианты плато $\{P_u, P_u\}$ и $\{P_u, P_d\}$. Если обозначить буквой V вертикальный экстремум, а буквой C – экстремум по кривизне, то правила отбора запишутся в следующем виде:

$\{P_u, P_u\} \rightarrow \{V_u\}$ либо $\{P_d, P_d\} \rightarrow \{V_d\}$, если размер плато меньше порогового;
 $\{P_u, P_u\} \rightarrow \{C, V_u, C\}$ либо $\{P_d, P_d\} \rightarrow \{C, V_d, C\}$, если размер плато больше порогового;
 $\{P_d, P_u\} \rightarrow \{\emptyset\}$ либо $\{P_u, P_d\} \rightarrow \{\emptyset\}$,

если размер плато меньше порогового;

$\{P_d, P_u\} \rightarrow \{C, C\}$ либо

$\{P_d, P_u\} \rightarrow \{C, C\}$, если размер плато больше порогового;

аналогичные правила можно сформулировать и для случая горизонтальных экстремумов.

В настоящий момент в основу алгоритма сопоставления кривых положена идея о нахождении

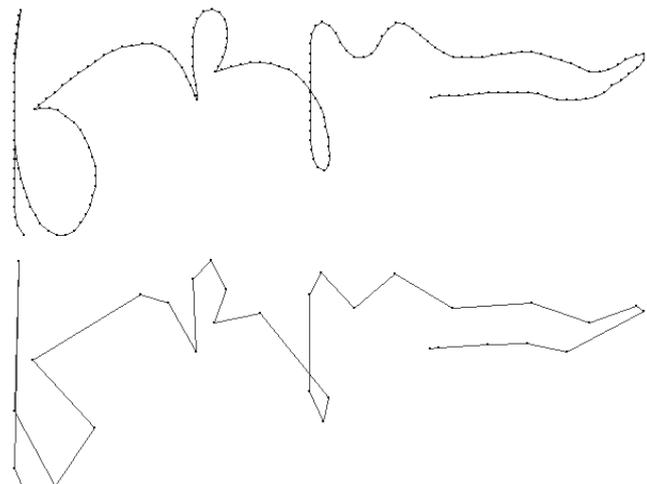


Рис. 10. Пример обработанной подписи, взятой из базы данных SVC 2004.

соответствия между вертикальными экстремальными точками. Приведенный выше алгоритм нахождения экстремальных точек вполне корректен, но он опирается на локальные особенности кривой, что может приводить к потере важной информации. На рис.9. приведен пример подписи, для которой сформированное представление является не точным, т.к. граничные точки горизонтального Т-штриха не указывают на наличие вертикального экстремума или плато (хотя это и не является ошибкой работы программы). Произошло это вследствие дрожания пера при выполнении данного штриха, а также в силу указанного локального рассмотрения окрестностей критических точек. Но это может привести к ошибке, поскольку применяемый алгоритм сопоставления ориентирован на нахождение пар вертикальных экстремумов противоположного типа. Вообще говоря, при другом исполнении этой же подписи данный горизонтальный штрих может быть классифицирован и обработан по правилу $\{Pu, Pu\} \rightarrow \{C, Vu, C\}$ (см. выше). Тем не менее, с целью избежать непредвиденных ошибок, последним этапом разметки подписи с помощью экстремальных точек является выделение горизонтальных участков. На этом же этапе помечались и заключительные горизонтальные штрихи (росчерки), если таковые в подписи присутствовали.

Заключение

Описанный метод нахождения экстремальных точек был использован для быстрого сопоставления динамических кривых. В отличие от альтернативных существующих подходов предложенный алгоритм позволяет более качественно учитывать особенности формы траектории (см. рис. 5,6). Собственно сам метод сопоставления кривых будет описан в следующей статье: «Быстрое сопоставление рукописных динамических подписей в биометрической системе контроля доступа».

Литература

1. *R. Plamondon, G. Lorette.* Automatic signature verification and writer identification – the state of the art. // *Pattern Recognition*, 1989, vol. 22, no. 2, pp. 107–131
2. *Д.В.Колядин.* Анализ динамических кривых применительно к задаче верификации рукописной подписи // *Математические методы распознавания образов (ММРО-11)*, 2003, с. 330-332
3. *M. E. Munich, P.Perona.* Visual Signature Verification using Affine Arc-length // *Conference on Computer Vision and Pattern Recognition CVPR*, 1999, pp. 2180-2186
4. *V. S. Nalwa.* Automatic On-Line Signature Verification. // *Proceedings of the IEEE*, 1997, vol. 85, no. 2, pp.215-239
5. *T. Starner, J. Makhoul, R. Schwartz, G. Chou.* On-line cursive handwriting recognition using speech recognition methods // *IEEE Conference on Acoustics, Speech, and Signal Processing*, 1994, vol. 5, pp. 125-128
6. *L. Yang, B.K. Widjaja, R. Prasad.* Application of Hidden Markov Models for Signature Verification. // *Pattern recognition*, 1995, vol. 28, no.2, pp. 161-170
7. *G. Seni, J. Seybold.* Diacritical processing for unconstrained on-line handwriting recognition using forward search // *International Journal on Document Analysis and Recognition*, 1999, no. 1, pp. 24-29
8. *T. Hastie, E. Kishon.* A Model for Signature Verification. // *IEEE International Conference on Decision Aiding for Complex Systems*, 1991, pp. 191-196
9. *M.E. Munich, P.Perona.* Continuous Dynamic Time Warping for translation-invariant curve alignment with applications to signature verification // *California Institute of Technology International Conference on Computer Vision*, 1999
10. *E. Keogh, M. Pazzani.* Scaling up Dynamic Time Warping for Datamining Applications. // *The 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2000
11. *S.Chu, E.Keogh, D.Hart, M.Pazzani.* Iterative Deepening Dynamic Time Warping for Time Series // *IEEE International Conference on Data Mining*, Maebashi City, Japan, 2002
12. *A. K. Jain, F. D. Griess, S. D. Connell.* On-line Signature Verification. // *Pattern Recognition*, vol. 35, no. 12, pp. 2963–2972, 2002

13. *S.D. Connell, A.K. Jain*. Template-based Online Character Recognition. // *Pattern Recognition*, 2001, vol. 34, no. 1, pp.1-14
14. *F. Hao, C. W. Chan*. Online Signature Verification Using a New Extreme Points Warping Technique // *Pattern Recognition Letters*, 2003, vol. 24, no. 16, pp. 2943-2951
15. *Д.В.Колядин, А.А.Савин*. О проблеме верификации подписи в системах контроля доступа // *Обработка информации и моделирование*. – М.: МФТИ, 2002, с.81-89
16. *Р.М.Ланцман*. Кибернетика и криминалистическая экспертиза почерка. – М.: Наука, 1968, 94 с.
17. *X. Li, M. Parizeau, R. Plamondon*. Detection of Extreme Points of On-line Handwritten Scripts // *Pattern Recognition Progress in Handwriting Recognition*, A. C.Downton (ed.), S. Impedovo (ed.), 1997