

Обзор интеллектуальных систем для оценки каллиграфии

77-48211/478895

09, сентябрь 2012

Демин А. А.

УДК. 81:002

Россия, МГТУ им. Н.Э. Баумана

fizmail@inbox.ru

Введение

Распознавание рукописного ввода — это способность компьютера получать и интерпретировать рукописные текстовые документы. В настоящее время все больше информации хранится в ЭВМ и вводится непосредственно с клавиатуры и мыши. Но одновременно с этим высока доля обмена рукописными текстами, поэтому актуальным является разработка устойчивых систем, позволяющих выполнять оперативное распознавание и интерпретацию рукописных данных путем считывания как и движения кончика ручки (онлайн метод), так и уже написанного на бумаге текста (оффлайн метод).

Эти задачи реализованы такими разработчиками, как Paragon Software group (система Pen Reader), iRex Technologies (система MyScript Notes), ABBYY (система Fine Reader). У каждого из выпущенного ими продукта своя область применения. Например, приложение «Pen Reader» работает только с динамическим вводом рукописного текста, приложение «MyScript Notes» хоть и является более функциональным решением в области распознавания рукописного текста, чем предыдущее, но напротив не распознает текст в режиме реального времени, а лишь конвертирует ранее введенный текст. Пакет «Fine Reader» работает только с печатным текстом высокого разрешения и в общем случае не распознает рукописный текст ни в статическом, ни в динамическом виде [1-8].

На сегодняшний день стоит задача разработки системы, ориентированной одновременно и на свободный динамический, и на статический ввод рукописного текста, обеспечивающей высокую точность распознавания текста с различными дефектами (склейка, разрывы) затрудняющими процесс обработки данных.

Целью работы является анализ методов оценки статистических отклонений вводимых символов от идеального начертания для предложения упражнений по коррекции почерка.

На основе методов оценки статистических отклонений вводимых символов от идеального начертания создан программно-аппаратный комплекс для распознавания рукописного текста. Решаемые при этом задачи включают анализ математических

алгоритмов классификации графических объектов; разработку алгоритмов распознавания образов, разработку структуры базы данных начертания символов, разработку методики составления индивидуального курса обучения.

Программно-аппаратный комплекс, в котором реализованы вышеописанные задачи, позволяет вводить рукописный текст с планшета, КПК и других сенсорных панелей, распознавать введенные символы, анализировать отклонения их написания от эталона и предлагать рекомендации по улучшению почерка и скорости письма. Разработанные методы анализа рукописного текста позволят оцифровывать рукописные документы, архивы, рукописи, усовершенствовать ввод информации в мобильные устройства без клавиатуры, создать рукописное заполнение бланков в электронном виде.

Сенсорная система оценки каллиграфии применима для обучения каллиграфическому почерку, иностранным языкам, предоставляет возможность дистанционного обучения правописанию. Также в связи с используемыми в системе методами оценки каллиграфии, её применение возможно в почерковедении, реабилитации, а также в частных прикладных задачах, например при автоматизированной экспертизе почерка.

1 Выявление основных методов и алгоритмов распознавания образов.

Задача распознавания (точнее, классификации) объекта ставится следующим образом. Имеется некоторый способ кодирования объектов (например, рукописных букв), принадлежащих заранее известному конечному множеству классов $S = \{C_1, \dots, C_q\}$, и некоторое конечное множество объектов (обучающее множество), про каждый из которых известно, какому классу он принадлежит. Нужно построить алгоритм, который по любому входному объекту, не обязательно принадлежащему обучающему множеству, решает, какому классу этот объект принадлежит [9].

Качество распознавания оценивается как вероятность (т.е. частота) ошибки классификации на другом конечном множестве объектов с заранее известными ответами (тестовом множестве) [9].

Типичная система оценки каллиграфии [9] включает извлечение признаков, распознавание объекта, принятие решения (рис. 1).

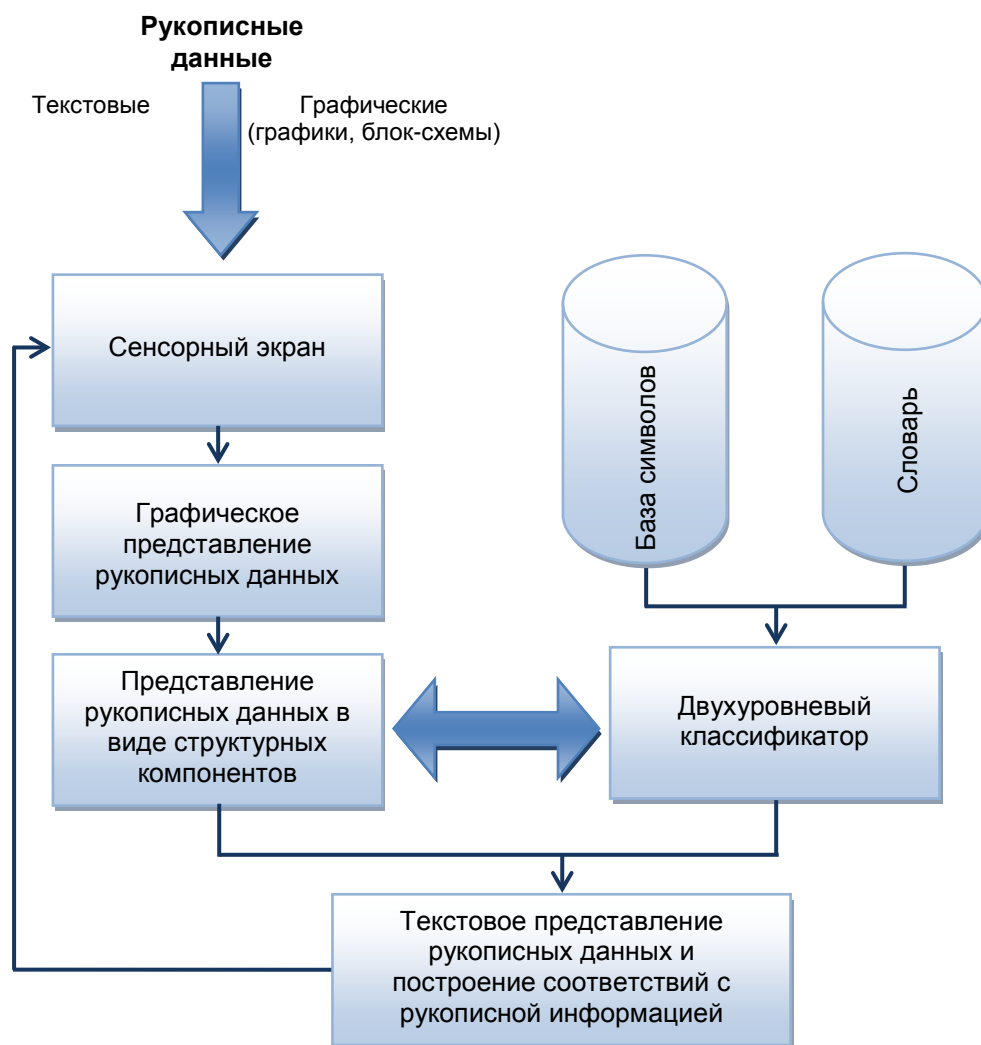


Рисунок 1 – Архитектура системы оценки каллиграфии

Извлечение признаков – это преобразование входных объектов к единообразному, компактному и удобному виду с исключением из модели подавляющей части содержащейся в объекте информации, слабо влияющей на классификацию. Наиболее удачным является представление объекта точкой евклидова пространства R^d , принадлежащей некоторому фиксированному компакту (кубу, шару, сфере, ...). Размерность d должна быть достаточно большой для успешного (в смысле качества) распознавания и достаточно малой для успешного (в смысле скорости) распознавания - реально это порядка нескольких десятков. Способ извлечения признаков зависит от природы и исходной кодировки объектов и подбирается вручную.

Распознавание объекта – это алгоритм, разбивающий пространство признаков на части, соответствующие заданным классам C_1, \dots, C_q . Выделим три основных метода распознавания (или типа классификаторов): шаблонные, признаковые и структурные [1].

Принятие решения – это процедура интерпретации вычисленных вероятностей, которая строится вручную и не зависит ни от природы входных объектов, ни от

пространства признаков, ни от обучающих данных. В современных интеллектуальных системах для принятия решения при распознавании рукописного текста применяются нейронные сети [1].

2 Сравнительный анализ основных методов распознавания

2.1 Шаблонные методы

Шаблонные методы преобразуют изображение отдельного символа в растровое представление, сравнивают его со всеми шаблонами, имеющимися в базе и выбирают шаблон с наименьшим количеством точек, отличных от входного изображения (рис. 2) [5].

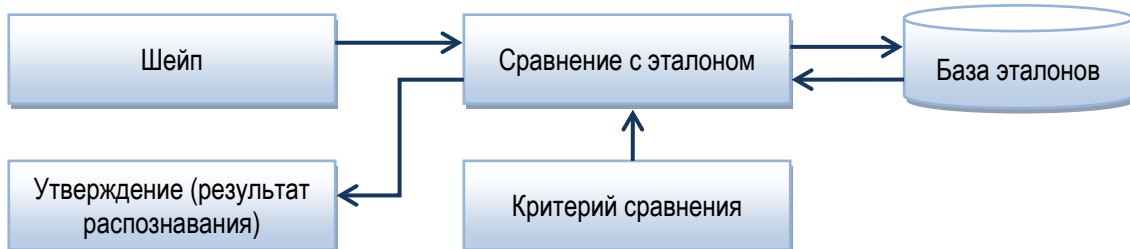


Рисунок 2 – Обобщенный алгоритм работы шаблонных методов

Рассмотрим методологию распознавания на конкретном примере шаблонного метода, который был разработан еще в 60-е годы прошлого века [2] и применялся при создании читающего устройства «РУТА 701», он также используется и в настоящее время [2].

Мерой подобия в данном методе выбран коэффициент подобия изображения символа с обобщенным образом 5-го класса, выражаемый следующей формулой:

$$R_s = \sum_{j=1}^n \left(\ln \frac{P_{js}}{1 - P_{js}} \right) x_j + \sum_{j=1}^n \ln(1 - P_{js})$$

где R_s - коэффициент подобия опознаваемого символа к эталонному изображению S -го класса символов; P_{js} – вероятность появления черного цвета в j -м элементе эталонного изображения S -го класса (выделяют три интервала вероятностей P_{js} : $0,00 \div 0,25$; $0,25 \div 0,75$; $0,75 \div 1,00$); x_j - значение интенсивности, соответствующей j -му элементу распознаваемого символа. Изображение символа отождествляется с эталонным классом, давшим максимальный коэффициент подобия R среди всех R_s .

2.2 Структурные методы

Структурные методы представляют объект как граф, узлами которого являются элементы входного объекта, а дугами – пространственные отношения между ними. Методы, реализующие подобный подход, обычно работают с векторными изображениями. Структурными элементами являются составляющие символ линии. Так, для буквы "Ф" – это вертикальный отрезок и дуга. Распознаваемый символ подвергается процедуре скелетизации (утонению) (рис. 3). Каждый полученный контур скелетного представления описывается в виде последовательного набора особых точек и «цепного» кода, состоящего из точки привязки, числа кодов и массива направлений из текущей точки к следующей [7].

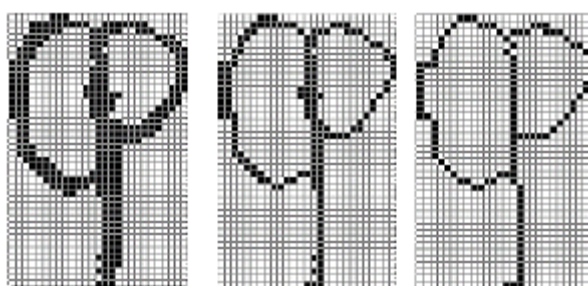


Рисунок 3 – Процесс скелетизации образа при структурных методах

Для каждой особой точки скелетного образа вычисляются следующие признаки:

- нормированные координаты особой точки;
- длина ребра до следующей вершины;
- нормированное направление из данной точки в следующую;
- нормированное направление входа в точку и выхода из точки;
- кривизна дуги, соединяющая особую точку со следующей вершиной.

На рисунке 4 условно показаны некоторые из топологических признаков. Граф имеет пять особых точек – a_0, a_1, a_2, a_3, a_4 . При обходе графа по маршруту $a_0-a_1-a_2...$ в вершине a_1 условно показаны следующие признаки: вектор R_1 – направление входа в точку, вектор R_2 – направление выхода из точки, вектор R_3 – глобальное направление на следующую особую точку. Двунаправленный вектор h показывает величину «левого» отклонения дуги (a_1, a_2) от прямой; «правое» отклонение равно нулю.

Как видно из приведенного описания, число признаков равняется восьмикратному числу вершин. Оно различается для разных топологических кодов, и признаки с одинаковым номером для разных топологических кодов могут иметь разный смысл [13].

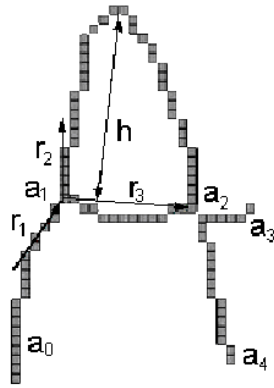


Рисунок 4 – Выделение признаков образа

a_0, a_1, a_2, a_3, a_4 - особые точки; r_1 - направление входа в точку; r_2 - направление выхода из точки; r_3 - глобальное направление на следующую особую точку; h - величина отклонения дуги (a_1, a_2) влево от прямой

Обучение метода состоит в построении деревьев распознавания для каждого из определенных заранее (вручную или автоматически) топологических кодов. Рассмотрим простую процедуру построения деревьев распознавания, приносящую неплохие результаты.

Для каждого топологического кода в обучающем множестве проводится отбор всех имен символов, имеющих достаточно большое представительство. Для каждого имени проводится анализ имеющихся значений признаков $p[i], i=1, \dots, N$, где N – число признаков для текущего топологического кода. Обозначим A_i – множество имеющихся значений для признака i для символов с именем A . Тогда для каждого $i, 0 < i \leq N, A_i$ представляется в виде

$$A_i = A_{i1} \cup A_{i2} \cup \dots \cup A_{im}, \text{ где } A_{ij} = \{x \in R : S_{ij} \leq x \leq e_{ij}\},$$

im – различно для каждого i и для каждого A .

Далее производится поиск конфликтов. Если для некоторых символов A, B значения признаков пересекаются, т.е. $A_i \cap B_i \neq \emptyset$ для $\forall i, i = 1 \dots N$, тогда проводится попытка разрешить конфликт. Делается попытка найти некоторый наилучший для разбиения признак o , выбрать точку деления этого признака k и разбить множество A на два непересекающихся подмножества A', A'' таким образом, что

$$A = A' \cup A'',$$

$$A_j = A_j' \cup A_j'',$$

где $A_j' = A_{j1} \cup A_{j2} \cup \dots \cup A_{jk}, A_j'' = A_{jk+1} \cup A_{jk+2} \cup \dots \cup A_{jm}, 0 < k, m_j$.

Затем процедура повторяется, т.е. для каждого A', A'' проводится построение областей значений признаков и поиск конфликтов с разноименными символами с возможной дальнейшей разбивкой множеств A', A'' и т.д. Конечно, все конфликты разрешить удастся не всегда, поэтому при распознавании в ряде случаев будет

выдаваться не одна альтернатива, а несколько. Оценки результирующих альтернатив будут зависеть как от значений признаков (топологический и не топологических), так и от представительности конфликтующих символов в обучающем множестве.

Таким образом, распознавание является древовидным, текущее дерево распознавания выбирается с помощью топологического кода.

2.3 Признаковые методы

Признаковые методы базируются на том, что изображению ставится в соответствие N -мерный вектор признаков. Распознавание заключается в сравнении его с набором эталонных векторов той же размерности. Задача распознавания, принятия решения о принадлежности образа тому или иному классу, на основании анализа вычисленных признаков, имеет целый ряд строгих математических решений в рамках детерминистического и вероятностного подходов [1; 6]. В системах распознавания символов чаще всего используется классификация, основанная на подсчете евклидова расстояния между вектором признаков распознаваемого символа и векторами признаков эталонного описания. Тип и количество признаков в немалой степени определяют качество распознавания. Формирование вектора производится во время анализа предварительно подготовленного изображения. Данный процесс называют извлечением признаков. Эталон для каждого класса получают путем аналогичной обработки символов обучающей выборки.

Основные достоинства признаковых методов – простота реализации, хорошая обобщающая способность, хорошая устойчивость к изменениям формы символов, низкое число отказов от распознавания, высокое быстродействие. Наиболее серьезный недостаток этих методов – неустойчивость к различным дефектам изображения. Кроме того, признаковые методы обладают другим серьезным недостатком – на этапе извлечения признаков происходит необратимая потеря части информации о символе. Извлечение признаков ведется независимо, поэтому информация о взаимном расположении элементов символа утрачивается.

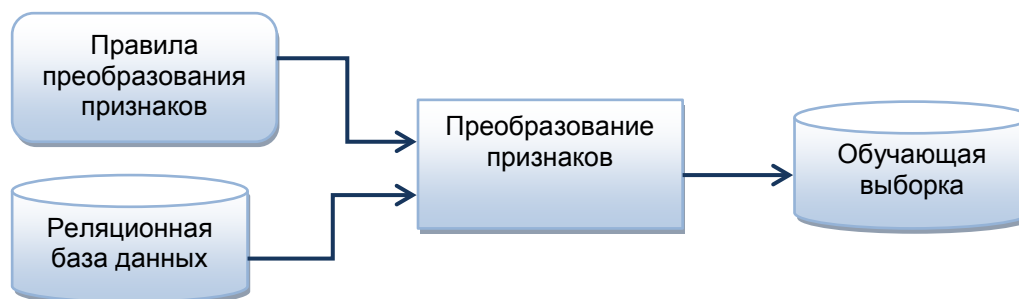


Рисунок 5 – Обобщенный алгоритм работы признаковых методов

Рассмотрим несколько методов создания наборов признаков, использующихся в настоящее время.

В качестве признаков изображений, например, можно использовать последовательности m_k , предложенные Е.С. Абрамовым [1], которые строятся по описанному ниже алгоритму.

Пусть на изображении A имеется объект O . Рассмотрим N радиус векторов r_k с началом в центре изображения C и концом на границе изображения, отстоящих друг от друга на угловое смещение $2\pi/N$. Вдоль каждого вектора существуют точки изображения, принадлежащие объекту O , и точки, не принадлежащие этому объекту. Эти точки формируют отрезки B_{kl} , $1 \leq l \leq N_k$, которые представляют собой «куски» объекта O вдоль вектора. Для каждого из этих отрезков мы можем определить длину L_{kl} и его расстояние от центра изображения d_{kl} . Величину, определяемую выражением $m_r = \sum_{l=1}^{N_k} d_{kl} L_{kl}$ будем называть моментом изображения вдоль вектора r_k . Проведя нормировку $m_k = \frac{m_k}{\sum_k m_k}$ можно обеспечить инвариантность момента изображения к масштабированию. Это позволяет обеспечить распознавание объектов на изображениях независимо от их масштаба.

Рассмотрим основную идею формирования признаков системы распознавания на основе функции инвариантных к сдвигу, вращению, предложенную Д.В. Горловым [15]. Для каждого символа вычисляются параметры (высота, ширина, количество точек) и признаки.

Процедура вычисления признаков состоит из пяти этапов.

1. Для контура текущего символа вычисляется значение центра тяжести.
2. Для текущей точки окружности u с центром в центре тяжести и радиуса $r = \text{const}$ вычисляется значение $\Phi_Q = P_Q(\theta^2(x, u))$, где x – текущая точка контура символа; Q – количество признаков; P_Q – полином Лежандра; $\theta(x, u) = p(x, u)/r$; $\eta = p^2(x, u)$; η – евклидово расстояние на плоскости.
3. Для всех точек контура символа вычисляется значение $\Phi'_{Q-1} \dots \Phi'_1$.
4. Для всех точек окружности находится характерная точка, в которой значение Φ'_Q максимально.
5. Для этой точки окружности вычисляются значения $\Phi'_{Q-1} \dots \Phi'_1$.

В качестве признаков используются значения вида $\Phi_Q = \frac{1}{k} \sum_{i=1}^k P_Q(\theta^2(x_i, u))$, где k – количество точек контура символа. Множитель $1/k$ введен с целью уменьшения влияния дискретности изображения. В качестве признаков выбираются ортогональные полиномы Лежандра:

$$P_1(\eta) = \frac{2}{b-a}(\eta - a) - 1, P_2(\eta) = \frac{1}{2}(3P_1^2 - 1), P_3(\eta) = \frac{1}{2}(5P_1^3 - 3P_1).$$

Из выкладок в [5] следует, что $a = 0$, $b = 2$.

Распознавание осуществляется следующим образом. Для контура символа вычисляются параметры (высота, ширина, количество точек) и признаки.

Последовательно контур символа сравнивается с объектами, полученными при обучении. При этом вычисляется мера близости в виде

$$L = \sum_{i=1}^Q |\Phi_i'' - \Phi_i|$$

где Φ_i'' – признак объекта, полученный при обучении; Φ_i – признак объекта, полученный при распознавании; Q – количество признаков.

Шаблонные, признаковые и структурные методы распознавания имеют как свои преимущества, так и недостатки. Сравнительный анализ этих методов приведен в таблице 1.

Таблица 1 - Сравнительный анализ основных методов распознавания

Методы	Достоинства	Недостатки
ШАБЛОННЫЕ	<ul style="list-style-type: none"> • высокая скорость распознавания; • простая реализация; • высокая точность распознавания дефектных символов. 	<ul style="list-style-type: none"> • необходимость настройки системы на типы и размеры шрифтов; • не может использоваться для описания объектов с высокой степенью изменчивости • может приниматься для распознавания только печатных символов • надежно распознают только те шрифты, шаблоны которых им "известны" • если распознаваемый шрифт хоть немного отличается от эталонного, шаблонные системы могут делать ошибки даже при обработке качественных изображений. • невозможность распознать шрифт, хоть немного отличающийся от заложенного в систему (размером, наклоном или начертанием).
СТРУКТУРНЫЕ	<ul style="list-style-type: none"> • применяется для рукописных шрифтов, имеющих множество вариантов начертания; • данные могут быть представлены в графовой форме, что обеспечивает инвариантность. 	<ul style="list-style-type: none"> • высокую чувствительность к дефектам изображения, нарушающим составляющие элементы. • для этих систем до сих пор не созданы эффективные автоматизированные процедуры обучения. • трудность распознавания дефектных символов и медленная работа. • векторизация может добавить дополнительные дефекты. • как только вы представите «разорванную» из-за дефектов печати букву, она уже не подойдет под свое описание.
ПРИЗНАКОВЫЕ	<ul style="list-style-type: none"> • простота реализации; • хорошая обобщающая способность; • хорошая устойчивость к изменениям формы символов; • высокое быстродействие. 	<ul style="list-style-type: none"> • неустойчивость к дефектам изображения; • при вычислении признаков теряется существенная часть информации • трудно гарантировать, что к данному классу удастся отнести только объекты этого класса.

Как видно из приведенного обзора, для всех трех методов свойственна неполнота и ограниченность условий применения. Каждый из описанных методов сам по себе имеет специализированную область применения: шаблонные методы эффективнее использовать для распознавания печатных шрифтов, структурные - рукописных при оффлайн-распознавании, признаковые - рукописных при онлайн-распознавании.

В современных системах распознавания обычно используются все три типа классификаторов, но основным является структурный. Два других служат для ускорения и повышения качества распознавания. Комбинация различных методов распознавания приводит к наилучшим результатам, примером может служить метод структурно-пятенных эталонов компании АBBYY.

2.4 Структурно-пятенные эталонные методы

Технология распознавания с помощью структурно-пятенных эталонов имеет также название «фонтанное преобразование» (от английского font - шрифт) [12]. Данный способ обработки символьной информации разработан программистами компании АBBYY [12] и используется в системе оптического распознавания текста "FineReader Рукопись" [12] (рис. 6).

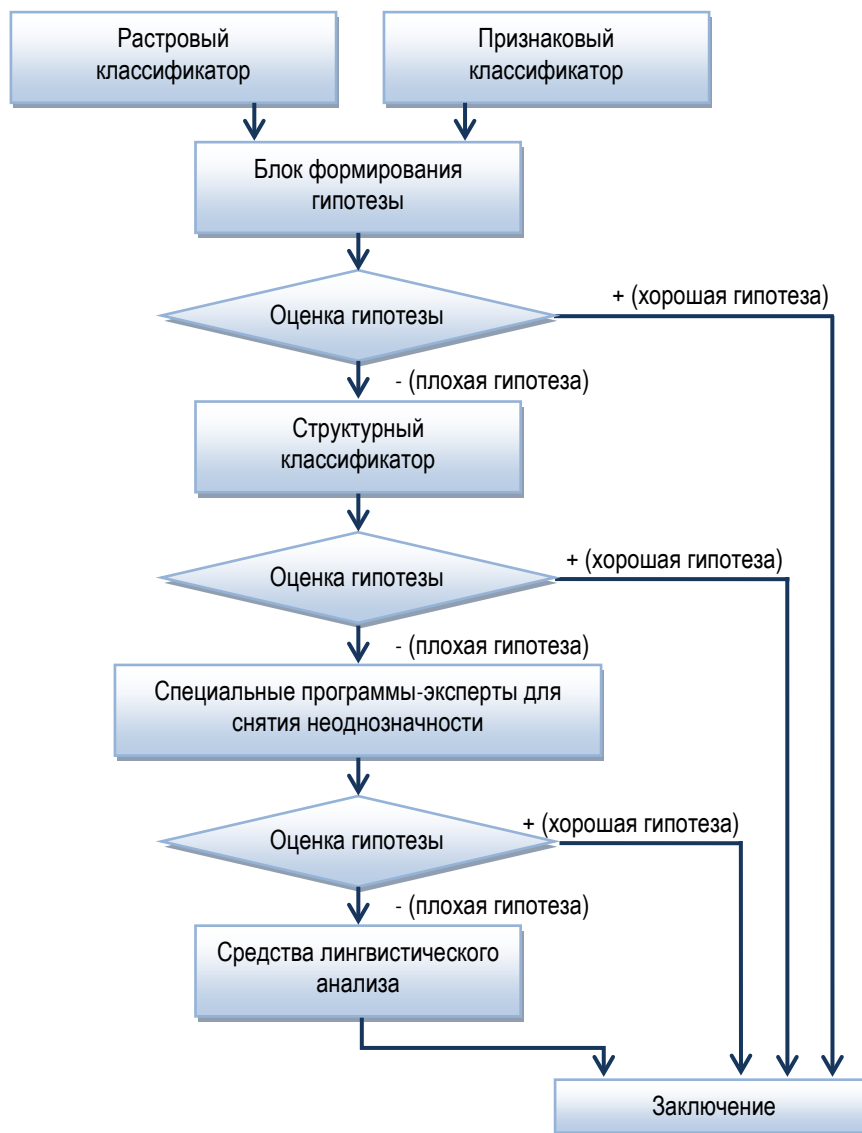


Рисунок 6 - Последовательность работы системы распознавания на основе метода фонтанного преобразования.

Структурно-пятенный эталон (рис. 7) позволяет представить изображения в виде набора пятен, связанных между собой n -арными отношениями, задающими структуру символа. Эти отношения (то есть расположение пятен друг относительно друга) образуют структурные элементы, составляющие символ. Так, например, отрезок - это один тип n -арных отношений между пятнами, эллипс - другой, дуга - третий. Другие отношения задают пространственное расположение образующих символ элементов.

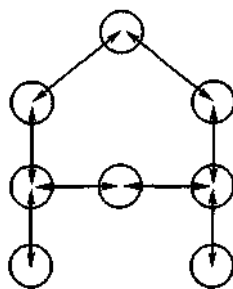


Рисунок 7 - Структурно-пятенный эталон

При задании структурных элементов и отношений используются конкретизирующие параметры [14], позволяющие доопределить структурный элемент или отношение при использовании этого элемента в эталоне конкретного класса. Для структурных элементов конкретизирующими могут являться, например, параметры, задающие диапазон допустимой ориентации отрезка, а для отношений - параметры, задающие предельное допустимое расстояние между характерными точками структурных элементов в отношении «соприкасается». В качестве отношений используются связи между структурными элементами, которые определяются либо метрическими характеристиками этих элементов (например, «длина больше»), либо их взаимным расположением на изображении (например, «правее», «соприкасается»). Таким образом в эталоне задаются:

- имя;
- обязательные, запрещающие и необязательные структурные элементы;
- отношения между структурными элементами;
- отношения, связывающие структурные элементы с описывающим прямоугольником символа;
- атрибуты, используемые для выделения структурных элементов;
- атрибуты, используемые для проверки отношений между элементами;
- атрибуты, используемые для оценки качества элементов и отношений;
- позиция, с которой начинается выделение элемента (отношения локализации элементов).

Структурные элементы, выделяемые для класса изображений, могут быть исходными и составными. Исходные структурные элементы - это пятна, составные - отрезок, дуга, кольцо, точка. В качестве составных структурных элементов, в принципе, могут быть взяты любые объекты, описанные в эталоне. Кроме того, они могут быть описаны как через исходные, так и через другие составные структурные элементы.

Например, для распознавания японских иероглифов (слоговое письмо) составными элементами для описания слога являются описания отдельных букв (но не отдельные элементы букв). В итоге, использование составных структурных элементов позволяет строить иерархические описания классов распознаваемых объектов [9] (рис. 8).

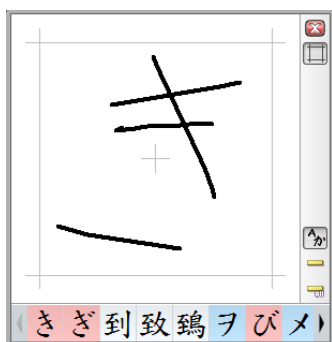


Рисунок 8 – Пример распознавания японских иероглифов

Распознавание с использованием структурно-пятенного эталона происходит следующим образом. Эталон накладывается на изображение, и отношения между выделенными на изображении пятнами сравниваются с отношениями пятен в эталоне. Если выделенные на изображении пятна и отношения между ними удовлетворяют эталону некоторого символа, то данный символ добавляется в список гипотез о результате распознавания входного изображения.

Данная разработка фирмы АВВУУ[12] первоначально использовалась для распознавания рукописного текста, а затем была успешно применена и для обработки печатных символов. При этом обеспечиваются все достоинства шаблонного и структурного классификаторов. Также данное представление нечувствительно к различным начертаниям и дефектам символов.

2.5 Нейросетевые методы

Для принятия решения при распознавании рукописного текста в современных интеллектуальных системах (PenReader, Calligrapher и др.) применяются нейронные сети, которые решают задачу классификации объектов [14-18].

В результате обучения на примерах строятся математические решающие функции (передаточные функции или функции активации), которые определяют зависимости между входными (i x) и выходными (j y) признаками (сигналами) (рис. 9).

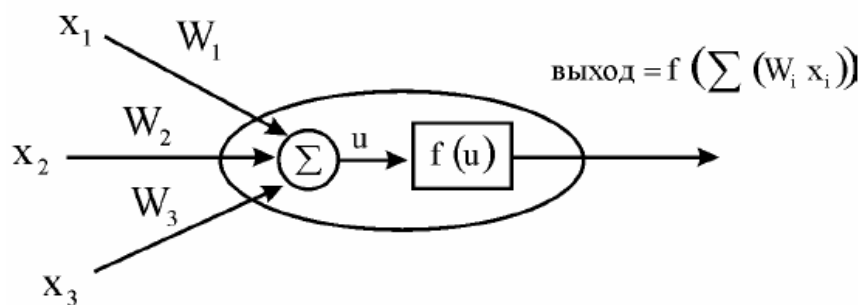


Рисунок 9 – Решающая функция – "нейрон"

Каждая такая функция, называемая по аналогии с элементарной единицей человеческого мозга – нейроном, отображает зависимость значения выходного признака ($j y$) от взвешенной суммы (U) значений входных признаков ($i x$), в которой вес входного признака ($i W$) показывает степень влияния входного признака на выходной.

Решающие функции используются в задачах классификации на основе сопоставления их значений при различных комбинациях значений входных признаков с некоторым пороговым значением [17]. В случае превышения заданного порога считается, что нейрон сработал и таким образом распознал некоторый класс ситуаций. Нейроны используются и в задачах прогнозирования, когда по значениям входных признаков после их подстановки в выражение решающей функции получается прогнозное значение выходного признака. Функциональная зависимость может быть линейной, но, как правило, используется сигмоидальная форма, которая позволяет вычленять более сложные пространства значений выходных признаков. Такая функция называется логистической (рис. 10).

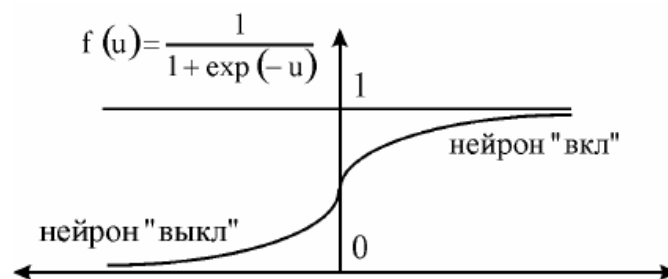


Рисунок 10 – Логистическая (сигмоидальная) функция

Нейроны могут быть связаны между собой, когда выход одного нейрона является входом другого. Таким образом, строится нейронная сеть (рис. 11), в которой нейроны, находящиеся на одном уровне, образуют слои.

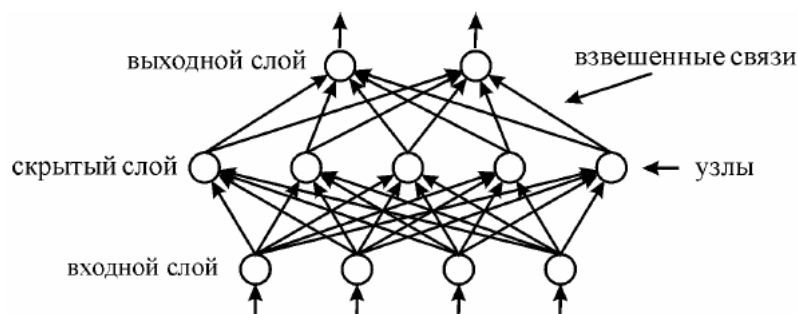


Рисунок 11 – Нейронная сеть

Обучение нейронной сети сводится к определению связей (синапсов) между нейронами и установлению силы этих связей (весовых коэффициентов). Алгоритмы обучения нейронной сети упрощенно сводятся к определению зависимости весового коэффициента связи двух нейронов от числа примеров, подтверждающих эту зависимость. Наиболее распространенным алгоритмом обучения нейронной сети является алгоритм обратного распространения ошибки. Целевая функция по этому алгоритму должна обеспечить минимизацию квадрата ошибки в обучении по всем примерам [3]:

$$\min \sum_i (T_i - y_i)^2,$$

где T_i – заданное значение выходного признака по i -му примеру; y_i – вычисленное значение выходного признака по i -му примеру. Сущность алгоритма обратного распространения ошибки сводится к следующему:

1. Задать произвольно небольшие начальные значения весов связей нейронов.
2. Для всех обучающих пар «значения входных признаков – значение выходного признака» (примеров из обучающей выборки) вычислить выход сети (Y).
3. Выполнить рекурсивный алгоритм, начиная с выходных узлов по направлению к первому скрытому слою, пока не будет достигнут минимальный уровень ошибки.

Вычислить веса на $(t + 1)$ шаге по формуле:

$$W_{i,j}(t + 1) = W_{i,j}(t) + \eta \delta_i X_i,$$

где

$W_{i,j}$ – вес связи от скрытого i -го нейрона или от входа к j -му нейрону на шаге t ;

X_i – выходное значение i -го нейрона;

η – коэффициент скорости обучения;

δ_i – ошибка для j -го нейрона.

Если j -й нейрон – выходной, то $\delta_i = y_i(1 - y_i)(T_i - y_i)$.

Если j -й нейрон находится в скрытом внутреннем слое, то

$\delta_j = X_j(1 - X_j) \sum_k \delta_k W_{j,k}$, где k – индекс всех нейронов в слое, расположенном вслед за слоем с j -м нейроном.

Выполнить шаг 2.

Достоинство нейронных сетей перед индуктивным выводом заключается в решении не только классифицирующих, но и прогнозных задач. Возможность нелинейного характера функциональной зависимости выходных и входных признаков позволяет строить более точные классификации.

Сам процесс решения задач в силу проведения матричных преобразований проводится очень быстро. Фактически имитируется параллельный процесс прохода по нейронной сети в отличие от последовательного в индуктивных системах. Нейронные сети могут быть реализованы и аппаратно в виде нейрокомпьютеров с ассоциативной памятью.

Все нейронные сети можно разделить на следующие типы [10].

Многослойные нейронные сети. Архитектура многослойной нейронной сети (МНС) состоит из последовательно соединённых слоёв, где нейрон каждого слоя своими входами связан со всеми нейронами предыдущего слоя, а выходами – следующего. НС с двумя решающими слоями может с любой точностью

аппроксимировать любую многомерную функцию. НС с одним решающим слоем способна формировать линейные разделяющие поверхности, что сильно сужает круг задач ими решаемых, в частности такая сеть не сможет решить задачу типа "исключающее или". НС с нелинейной функцией активации и двумя решающими слоями позволяет формировать любые выпуклые области в пространстве решений, а с тремя решающими слоями – области любой сложности, в том числе и невыпуклой.

При этом МНС не теряет своей обобщающей способности. Обучаются МНС при помощи алгоритма обратного распространения ошибки, являющегося методом градиентного спуска в пространстве весов с целью минимизации суммарной ошибки сети. При этом ошибки (точнее величины коррекции весов) распространяется в обратном направлении от входов к выходам, сквозь веса, соединяющие нейроны.

Нейронные сети высокого порядка. Нейронные сети высокого порядка (НСВП) отличаются от МНС тем, что у них только один слой, но на входы нейронов поступают так же термы высокого порядка, являющиеся произведением двух или более компонент входного вектора. Такие сети так же могут формировать сложные разделяющие поверхности. Особенность такой сети заключается в том, что для обучения некоторому классу достаточно предъявить его образ без вариаций масштабов и поворотов - после обучения сеть будет распознавать известные классы инвариантно к масштабу и поворотам. Такая сеть не является полносвязной, быстро обучается и работает. Отмечено существенное повышение точности классификации такой сетью повернутых и масштабированных изображений по сравнению с МНС.

Нейронные сети Хопфилда. НС Хопфилда (НСХ) является однослойной и полносвязной (связи нейронов на самих себя отсутствуют), её выходы связаны со входами. В отличие от МНС, НСХ является релаксационной - т.е. будучи установленной в начальное состояние, функционирует до тех пор, пока не достигнет стабильного состояния, которое и будет являться её выходным значением. НСХ применяются в качестве ассоциативной памяти и для решения оптимизационных задач. В первом случае НСХ обучается без учителя (например, по правилу Хебба), во втором случае веса между нейронами изначально кодируют решаемую задачу. НСХ бывают синхронными, когда одновременно пересчитываются все нейроны и асинхронными, когда пересчитывается случайно выбранный нейрон. Для исследования динамики функционирования НСХ используются методы Ляпунова.

Показано, что асинхронная НСХ всегда сходится к устойчивым точкам, а аттракторами синхронной НСХ являются устойчивые стационарные точки и предельные циклы длины два. Таким образом НСХ из начального состояния сходится к ближайшему локальному минимуму энергии сети, состояние нейронов в котором и будет восстановленным образом для задач распознавания, и решением - для оптимизационных задач. Для поиска глобального минимума применительно к оптимизационным задачам используют стохастические модификации НСХ.

Самоорганизующиеся нейронные сети Кохонена. Самоорганизующиеся нейронные сети Кохонена (НСК) обеспечивают топологическое упорядочивание входного пространства образов. Они позволяют топологически непрерывно отображать входное n -мерное пространство в выходное m -мерное.

Нейронная сеть с радиально-базисной функцией (НСРБФ) является дальнейшим развитием НС Кохонена, в которой после конкурентного слоя добавлен ещё один слой, обучаемый по методу обратного распространения. В отличие от НС Кохонена в НСРБФ выходами нейронов конкурентного слоя являются значения функции Гаусса с нормальным законом распределения, и обнуление не победивших нейронов не требуется. Ширина радиально-базисной функции характеризует расстояние между центром кластера, который образуется каждым нейронным элементом и его ближайшими соседями.

Когнитрон. Когнитрон своей архитектурой похож на строение зрительной коры, имеет иерархическую многослойную организацию, в которой нейроны между слоями связаны только локально. Обучается конкурентным обучением (без учителя). Каждый слой мозга реализует различные уровни обобщения; входной слой чувствителен к простым образам, таким, как линии, и их ориентации в определенных областях визуальной области, в то время как реакция других слоев является более сложной, абстрактной и независимой от позиции образа. Аналогичные функции реализованы в когнитроне путем моделирования организации зрительной коры.

Рассмотренные нейросетевые методы обеспечивают быстрое и надёжное распознавание изображений. Однако при применении этих методов к рукописному тексту возникают трудности, связанные с нарушениями каллиграфии у операторов системы. Такие ограничения обычно преодолеваются путём предъявления всевозможных вариаций изображения (различные варианты написания символов) при обучении, но построение такого обучающего набора – трудная задача. Поэтому к классифицирующим системам предъявляется требование – имея конечный репрезентативный набор вариаций образов некоторых классов, обобщить свой опыт на все остальные классы, не входившие в обучающий набор. Т.е. система должна извлечь характеристики, инвариантные к внутриклассовым изменениям и максимально репрезентативные по отношению к межклассовым изменениям. Такая задача в общем виде для систем распознавания ещё не решена, но существуют методы, которые показывают возможности решения отдельных её аспектов [19 - 21].

3 Обобщенная архитектура интеллектуальных систем оценки каллиграфии

Рассмотрим создание интеллектуальных систем для оценки каллиграфии на примере системы «Электронная пропись» российских разработчиков.

Программно – аппаратный комплекс обучающей системы может быть размещен на различные аппаратные конфигурации. Структура аппаратной подсистемы представлена на рисунке 12.

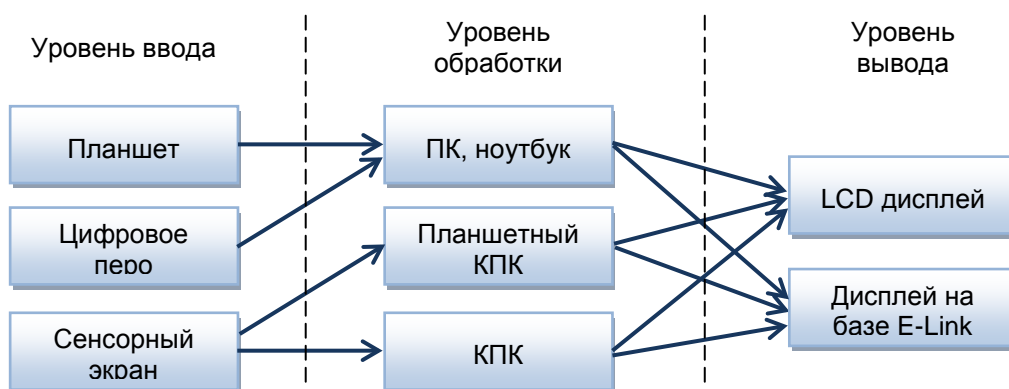


Рисунок 12 – Структура аппаратного обучающего комплекса

В качестве элементов ввода могут быть использованы планшеты, сенсорные экраны, подходящие для целей обучающей системы (в основном резистивного типа для ввода при помощи стилуса заменяющего ручку), либо цифровое перо позволяющее писать на обычной бумаге.

Вычислительным ядром может выступать персональный компьютер, ноутбук, КПК (возможно применение смартфона постепенно вытесняющего карманные компьютеры).

Для вывода информации может быть применены LCD дисплеи или «электронная бумага» E-ink, которая имеет ряд преимуществ, заключающихся в низком энергопотреблении и высокой контрастности изображения. Функциональный состав комплекса отображен на рисунке 13.



Рисунок 13 – Функциональный состав комплекса

4 Анализ экспериментальных результатов по оценке каллиграфии

Под надежностью тестовых результатов понимается характеристика, показывающая точность измерения знаний заданиями теста. Нужно отметить, что речь идет не о надежности теста, а о надежности результатов тестирования, т.к. на нее сильно влияет степень однородности различных групп обучаемых, уровень их подготовленности и ряд других факторов, связанных не с самим тестом, а с условиями проведения процесса тестирования.

Разработанная методика оценки каллиграфии рукописного текста для интеллектуальной интерактивной обучающей системы предлагает пятибалльную шкалу оценки.

Основные критерии оценки: сложность упражнения, правильность выполнения, время выполнения, процент правильно выполненных упражнений.

Алгоритм оценки:

1. распознавание символов
2. сравнение с эталоном
3. оценка скорости написания
4. определение силы нажима
5. оценка времени выполнения
6. оценка сложности упражнения
7. расчет итоговой оценки

В качестве модели получения знаний в системе используется адаптивная модель. Адаптивным называется комплекс упражнений, в котором сложность заданий меняется в зависимости от правильности выполнения предыдущих [12].

В адаптивных обучающих системах результаты применяются как средство идентификации личности для построения индивидуальной последовательности обучения. При этом можно различают три уровня оценки: предварительный (до начала обучения), текущий (во время обучения) и итоговый (после окончания обучения).

Предварительная оценка каллиграфии рукописного текста пользователя применяется для определения начального уровня сложности упражнений, выявления слабо получающихся элементов письма. На основе полученных данных формируется предварительная последовательность упражнений, которая будет максимально эффективна для конкретного пользователя системы.

Текущая оценка – это контроль или самоконтроль прогресса по отдельному элементу учебного курса. По его результатам строится последовательность отработки навыков письма, а также может осуществляться возврат к упражнениям на те элементы, которые были освоены недостаточно хорошо.

И, наконец, итоговая оценка – это оценка прогресса по курсу в целом или по совокупности курсов. По его результатам корректируется последовательность прохождения комплекса упражнений.

Основной особенностью применяемой методики оценки является, что заранее неизвестно, сколько упражнений необходимо выполнить пользователю, для определения его уровня подготовки.

Поэтому при формировании курсов необходимо разнообразить количество сценариев выполнения упражнения одного уровня сложности, чтобы избежать однообразия в процессе оценки каллиграфии рукописного текста. Кроме того следует ограничить максимальное количество упражнений, которое будет предлагаться пользователю. В этом случае оценка формируется на основе полученных результатов, с учетом линии прогресса от первого - к последнему упражнению.

Первое упражнение предлагаемое пользователю соответствует среднему уровню сложности. Рост сложности следующего упражнения при успешном выполнении предыдущего происходит нелинейно, что позволяет сократить количество заданий и соответственно время оценки.

Оценка складывается исходя из следующих критериев:

1. Сложность задания.
2. Правильность выполнения.
3. Время выполнения.
4. Процент правильно выполненных упражнений.

После прохождения квалификационного комплекса упражнений, система выстраивает индивидуальную последовательность курсов направленную на улучшение каллиграфии рукописного текста пользователя.

Ниже представлен алгоритм работы системы при прохождении квалификации пользователем (рис.14).

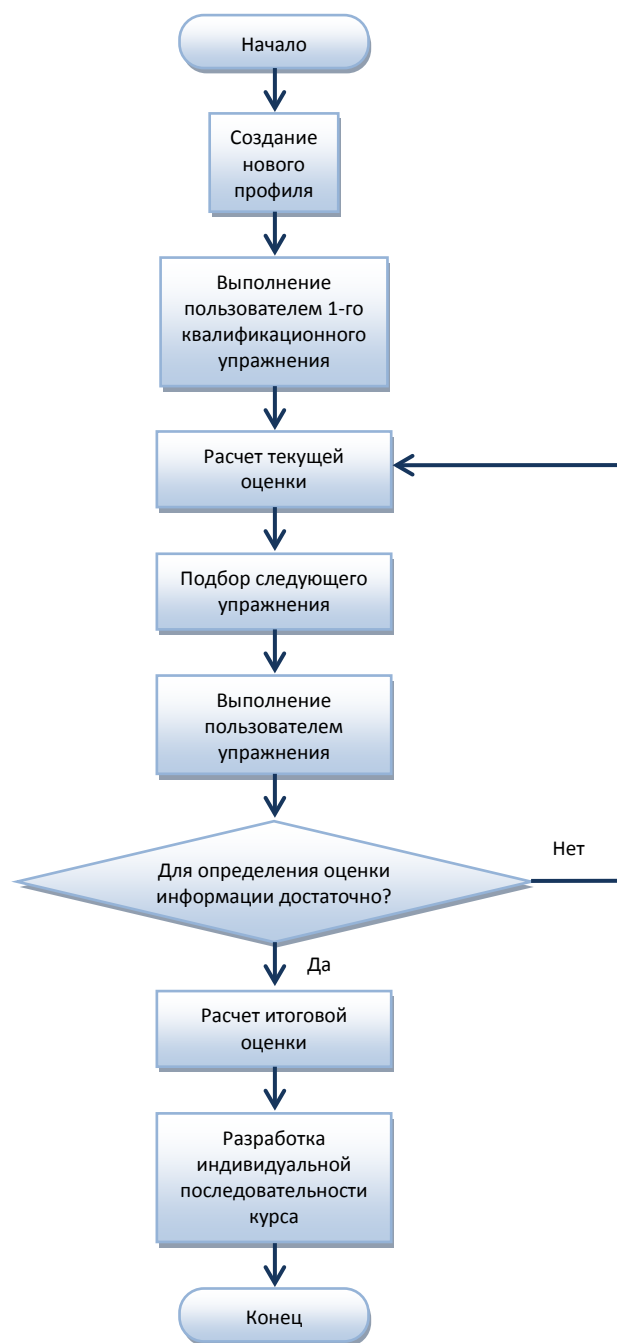


Рисунок 14 –Алгоритм работы системы при прохождении квалификации пользователем

Области применения. Сенсорная система оценки каллиграфии применима в различных областях.

Развитие моторики у детей дошкольного возраста. Актуальность данной темы заключается в том, что развитие моторики у детей дошкольного возраста позволяет сформировать координацию движений пальцев рук, развить речевую деятельность и подготовить ребёнка к школе [16] (рис. 15).

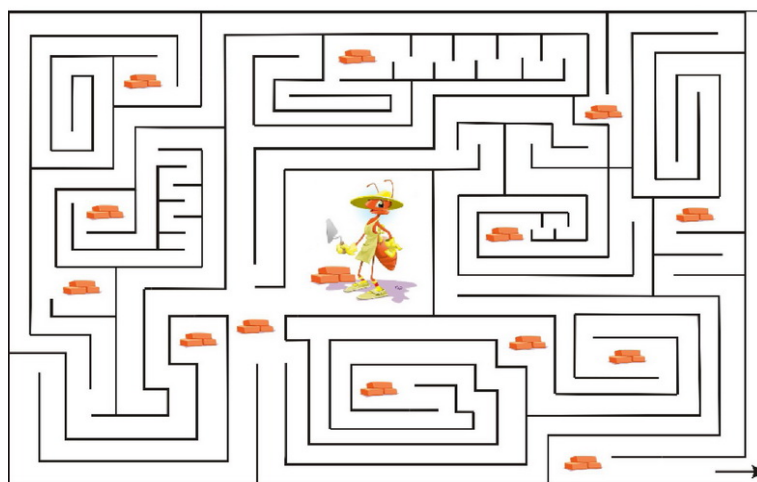


Рисунок 15 – Упражнение для дошкольников

Обучение письму школьников является трудным процессом в связи с тем, что в ходе написания включается ручная моторика, на первых порах весьма несовершенная, у некоторых детей обнаруживаются нарушения моторики мелких мышц руки, вялость, некоординированность движений, тремор и т.п. Кроме того, сам процесс письма представляет собой сложный акт, требующий включения многих анализаторов. [8] Алгоритм индивидуального обучения системы «Электронная пропись» подбирает только необходимые для конкретного ребенка упражнения, анализируя его ошибки. Благодаря этой технологии работа с упражнениями не утомляет ребенка и вызывает интерес для дальнейшего изучения материала (рис. 16).

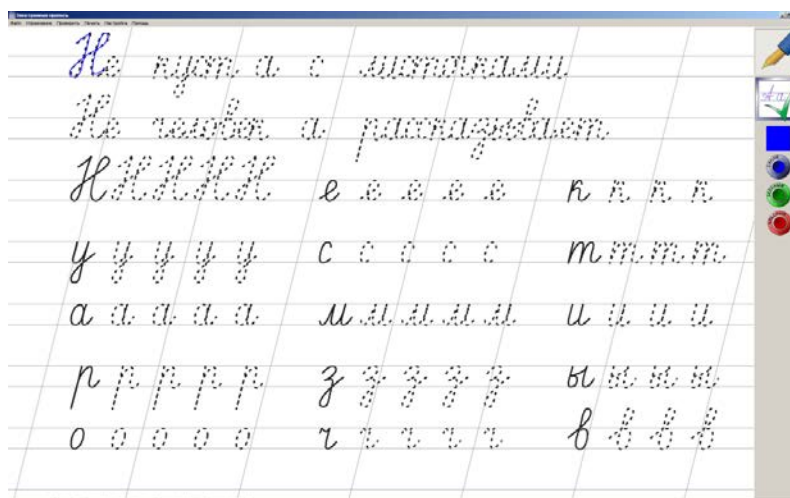


Рисунок 16 – Упражнение по русскому языку

Обучение иностранным языкам и каллиграфическому подчерку так же стоит на адаптивной модели обучения, которая позволяет индивидуально подбирать уровень сложности в зависимости от подготовки обучающегося и позволяет самостоятельно оценить результат выполненной работы (рис. 16).

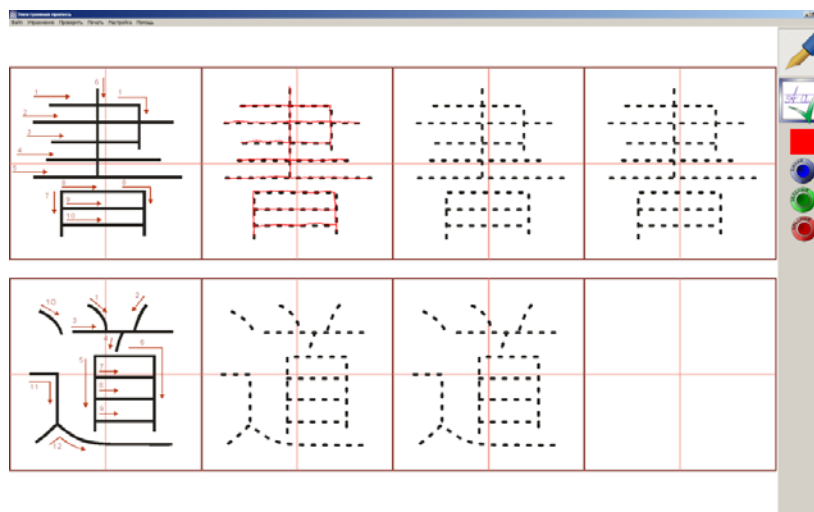


Рисунок 17 – Упражнение по японскому языку

Реабилитация неврологических больных с двигательными нарушениями. Алгоритм индивидуальной тренировки системы «Электронная пропись – Реабилитация» подбирает только необходимые упражнения, анализируя допущенные ошибки в письме (рис.18).

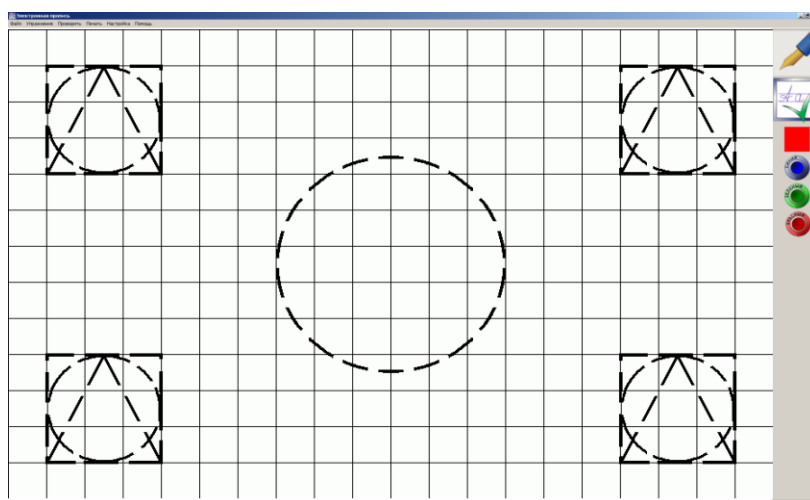


Рисунок 18 – Упражнение для реабилитации

Благодаря этой технологии работа с упражнениями не утомляет внимание и вызывает интерес для скорейшего завершения реабилитационного периода.

Заключение

В рамках данной статьи дан обзор методов получения и обработки графических образов в реальном времени, используемых в большинстве известных приложениях, описаны их основные преимущества и недостатки.

В обзоре приведены математические модели, структурные схемы и алгоритмы обработки образов рукописного текста, позволяющие анализировать векторные

представления графических образов в режиме реального времени, которые легли в основу обучающей системы. Так же проанализированы основные функциональные части современных интеллектуальных систем распознавания рукописного текста и оценки каллиграфии: извлечение признаков, распознавание, принятие решения и используемые методы классификации для решения подобных задач (шаблонный, структурный, признаковый).

Разработка программно-алгоритмического обеспечения системы распознавания образов была рассмотрена на примере системы «Электронная пропись». Предложен и разработан алгоритм оценки рукописного текста, позволяющий оценить скорость письма, качество каллиграфии, и прогресс улучшения основных показателей обучения.

Разработанная система может быть использована для изучения каллиграфии в образовательных учреждениях и для реабилитации неврологических больных в медицинских учреждениях.

Литература

1. Абрамов Е. С. Моделирование систем распознавания изображений (На примере печатных текстов) : Дис. ... канд. техн. наук : 05.13.01 СПб., 2006 – 140 с.
2. Багдонас А., Жлабис Р., Кведаравичус А., Нашлю-нас Р., Швагждис П., Яшинскас П. Читающее устройство «РУТА 701» // Автоматизация ввода письменных знаков в электронно-вычислительные машины: доклады науч.-техн. совещания, Вильнюс, 1968. – С. 96-121.
3. Бухалто А.Н., Булаев В.И., Власов А.И. и др. Нейрокомпьютеры в системах обработки изображений. – М.: Радиотехника, 2003. – 192 с.
4. Власов А.И., Кирбабин О.Е., Шепель А.С. Методология визуального проектирования как инструмент организации поддержки учебного процесса // Новые информационные технологии и менеджмент качества (NIT&MQ`2008). Материалы международного симпозиума/Редкол.: А.Н.Тихонов (пред.) и др. // ФГУ ГНИИ ИТТ «Информика». – М.: ЭГРИ, 2008. – 219 с.
5. Горлов Д.В. Распознавание изображений на основе признаков, инвариантных к сдвигу, вращению, масштабированию: Автореф. дис. канд. техн. наук: – Красноярск, 2002. – 20 с.
6. Дуда Р., Харт П. Распознавание образов и анализ сцен. М.: – Мир, 1989. – 510 с.
7. Котович Н.В., Славин О.А. Распознавание скелетных образов. Электронный ресурс. // Режим доступа: <http://ocrai.narod.ru>
8. Кудреватова Л.И. Типичные каллиграфические ошибки и их исправления // Начальная школа, 1995. – №10. – С.39-45.
9. Мерков А.Б. Основные методы, применяемые для распознавания рукописного текста // Лаборатория распознавания образов МЦНМО, 2004.
10. Морозов М. Н. Курс лекций по дисциплине «Системы искусственного интеллекта» // Марийский государственный технический университет, 2005.
11. Проненко Л.И. Каллиграфия для всех. – М.: Книга, 1990. – 115 с.
12. Телков А.Ю. Экспертные системы: Учебное пособие. – Воронеж: ИПЦ ВГУ,

2007. – 83 с.

13. Фу К. Структурные методы в распознавании образов. – М.: Мир, 1977 – 320 с.
14. Фукунага К. Введение в статистическую теорию распознавания образов. – М.: Наука, 1979. – 368 с.
15. Хорошилова Е.И. Характеристика специфических ошибок письма младших школьников // Начальная школа, 2001. – №7. – С.32–35.
16. Цыновникова Ю.Л. Автоматизация навыка письма на этапе первичного усвоения буквенных знаков // Начальная школа, 2001. – №7. – С.57-62.
17. Шахнов В.А., Власов А.И., Кузнецов А.С., Поляков Ю.А. Нейрокомпьютеры: архитектура и схемотехника. - М.: Машиностроение, 2000. – 64 с.
18. Шахнов В.А., Власов А.И., Поляков Ю.А., Кузнецов А.С. Нейрокомпьютеры: Архитектура и схемотехника // ChipNews, 2011. – №1. - С.18-21.
19. Шахнов В.А., Власов А.И., Зинченко Л.А. О методическом обеспечении инженерного образования в современных условиях // Высшее образование в России, 2012. – №3. – С.104-108.
20. Belhumeur P. N., Hespanha J. P. and Kriegman D. J. Eigenfaces vs Fisherfaces: Recognition Using Class Specific Linear Projection. IEEE Transactions on Pattern Analysis and Machine Intelligence 1997, Vol. 19, pp. 711-720.
21. Vetter T. and Poggio T. Linear Object Classes and Image Synthesis From a Single Example Image. IEEE Transactions on Pattern Analysis and Machine Intelligence 1997, Vol. 19, pp. 733-742.