

УДК 681.325

Д.В. Гринев

Харьковский университет Воздушных Сил им. И. Кожедуба, Харьков

РАСПОЗНАВАНИЕ ТЕКСТА ПО СТРУКТУРНЫМ КРИТИЧЕСКИМ ТОЧКАМ

Предлагается подход к распознаванию изображений печатного и рукописного текста, основанный на определении признаков классов объектов распознавания по структурным критическим точкам, позволяющий уменьшить влияние деформационных искажений в их контурах на результат распознавания.

распознавание образов, структура контурного изображения, деформационные искажения, критические точки

Введение

В последние годы распознавание образов находит все большее применение в повседневной жизни. Распознавание речи и рукописного текста значительно упрощает взаимодействие человека с компьютером, распознавание печатного текста используется для перевода документов в электронную форму. Особенно актуальной является проблема распознавания текста с высокой вероятностью правильной классификации в режиме реального времени [1 – 3].

Стандартный подход к проблеме распознавания заключается в сведении задачи распознавания к задаче классификации некоторого набора признаков.

В системе существует некоторое количество эталонов, которые обладают определенным набором признаков. Символ, который нужно распознать сравнивается попеременно с каждым из имеющихся образцов. Затем выбирается эталон, с которым набор признаков имеющегося символа совпадает лучше всего, и на основе этого делается вывод о принадлежности этого символа к тому или иному классу. Такой подход по сути своей не позволяет добиться высокого качества распознавания, как бы он не был усовершенствован [2, 3].

Главный его недостаток заключается в том, что в любом случае в наборе признаков содержится не вся информация об изображении и вариантах изображения того или иного символа будет больше чем эталонов. Тогда, как только система сталкивается с деформационными искажениями в структуре объекта распознавания (нестандартным написанием буквы или цифры), возникает проблема соотнесения исследуемого изображения к определенному классу объектов распознавания.

Поэтому одной из важных проблем, с которыми приходится сталкиваться при распознавании контурных изображений рукописного текста, являются различные деформационные искажения их структур, влияющие на результат распознавания [2].

Целью данной статьи является рассмотрение подхода к процессу распознавания печатного и рукописного текста, устраняющего влияния различных деформационных искажений на результат распознавания.

Результаты исследований

В настоящее время для распознавания изображений объектов в режиме реального времени наиболее перспективно применять методы, основанные на структурном и структурно-лингвистическом подходе к распознаванию, позволяющих построить инвариантные признаки распознавания к различным деформационным искажениям структуры [3 – 6].

К деформационным искажениям относятся:

- добавление новых структурных элементов различного уровня общности по сравнению с эталонными (полученными в процессе обучения) структурами;
- изменение параметров структурных элементов;
- отсутствие в структуре распознаваемого изображения определенного класса распознавания структурных элементов, которые присутствуют в структурах эталонных (используемых в процессе обучения) изображений.

Для определения структуры распознаваемого контурного изображения, зададим базисные направления ориентации X_1, X_2, X_3, X_4 в системе координат X . Данный ортонормированный базис разбивает пространство E^2 на четыре квадранта. В каждом квадранте задается множество векторов

$$\bar{x}_1^\sigma, \bar{x}_2^\sigma, \dots, \bar{x}_n^\sigma.$$

В процессе определения производных структурных элементов контура распознаваемого изображения при правом (левом) направлении обхода контура, начало координат O совмещается с началом структурного элемента изображения. В общем случае данные построения определяют ортогональное параллельное проектирование (Pr) структурных

элементов a_{ij}^n структуры z_i^n распознаваемого изображения I_i на вектора \bar{x}_ξ^σ векторных пучков \bar{X}^σ , задающих γ -е направления ориентации g_γ , $\gamma = (\overline{1, 4 \times n})$ в системе направлений ориентации Ξ при последовательной нумерации направлений ориентации данной системы в направлении обхода контура изображения [4]:

$$Pr : a_{ij}^n \rightarrow \bar{x}_\xi^\sigma.$$

Значение функции $Inv(a_{ij}^n) = g_\gamma$, определяющей отображение $a_{ij}^n \rightarrow g_\gamma$, называется структурным вариантом элемента a_{ij}^n структуры z_i^n в системе Ξ .

Тогда

$$Inv(z_i^n) = G_i^n; G_i^n = \langle g, r, B \rangle,$$

где G_i^n – структура инвариантов n -го структурного уровня в системе Ξ ; g – множество направлений ориентации в данной системе; r – бинарные отношения, в которых находятся элементы множества A_i^n ; B – аксиомы структуры, условиям которых удовлетворяют данные отношения [6].

Выделение структуры инвариантов G_i^n на структуре z_i^n позволяет осуществить устранение влияния аффинных преобразований типа гомотетий и поворотов на процесс распознавания изображения I_i .

Для устранения влияния деформационных искажений на процесс распознавания необходимо определить разнородные структурные критические точки, определяющие начало формирования новой подструктуры или продолжение развития уже выявленной подструктуры.

Определим, что каждый структурный элемент a_{ij}^2 является отрезком, соединяющим две соседние точки контура изображения. Для определения рода точки, являющейся конечной для структурного элемента a_{ij}^2 и начальной для следующего структурного элемента $a_{i,j+1}^2$ сопоставим структурный элемент a_{ij}^2 с вектором \bar{x}_1^1 , которому ставится в соответствие единичное направление ориентации $g_\gamma = 1$ в относительной системе координат. На следующем шаге анализируется ориентация обобщенного направления развития структуры изображения $g_v^{об}$ на структурных элементах a_{ij}^2 и $a_{i,j+1}^2$ относительно $g_\gamma = 1$. В зависимости от расположения $g_v^{об}$ в пра-

вой либо левой полуплоскости системы координат Ξ' относительно \bar{x}_1^1 , определяются структурные критические точки 1-го и 2-го рода.

При правом обходе контура изображения точка пересечения структурного элемента a_{ij}^2 , имеющего направление $\bar{x}_1^1 \rightarrow g_\gamma = 1$, и структурного элемента $a_{i,j+1}^2$, имеющего направление $\bar{x}_i^\sigma \rightarrow g_\mu$, называется точкой 1-го рода t_j^1 , если вектор обобщенного направления развития структуры $g_v^{об}$ на данных структурных элементах лежит в 1 или 2 квадранте системы координат Ξ' относительно \bar{x}_1^1 .

Точки 1-го рода характеризуют продолжение развития текущей подструктуры на рассматриваемом участке контура изображения и определяют выпуклость данной подструктуры. Процедура выделения структурных критических точек 1-го рода представлена на рис. 1, а.

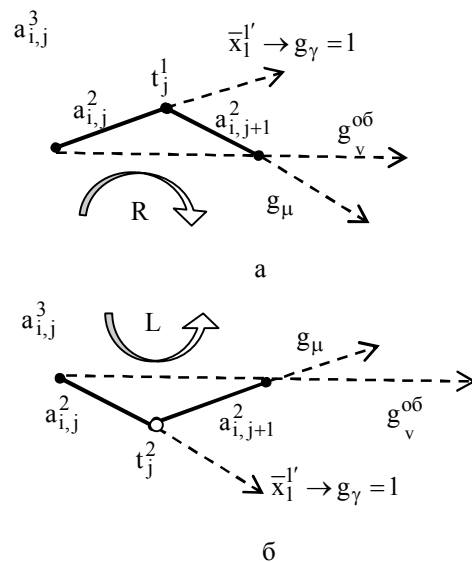


Рис. 1. Процедура выделения структурных критических точек 1-го и 2-го рода

При правом обходе контура изображения точка пересечения структурного элемента a_{ij}^2 , имеющего направление $\bar{x}_1^1 \rightarrow g_\gamma = 1$, и структурного элемента $a_{i,j+1}^2$, имеющего направление $\bar{x}_i^\sigma \rightarrow g_\mu$, называется точкой 2-го рода t_j^2 , если вектор обобщенного направления развития структуры изображения $g_v^{об}$ на данных структурных элементах лежит в 3 или 4 квадранте системы координат Ξ' относительно \bar{x}_1^1 .

Точки 2-го рода характеризуют начало развития новой подструктуры на рассматриваемом участке

контура изображения и определяют вогнутость подструктур. Процедура выделения структурных критических точек 2-го рода представлена на рис. 1, б.

Точки 1-го и 2-го рода описывают формирование подструктур различного уровня вложенности.

Для выделения подструктур 1-го уровня вложенности необходимо определить:

- структурные критические точки, в которых подструктура начинает свое развитие (точки максимальной вогнутости структуры);

- структурные критические точки, в которых подструктура имеет пик своего развития (точки максимальной выпуклости структуры).

Максимальными структурными критическими точками 1-го рода $t_{i \max}^1 = \max(t_j^1)$ называются такие точки 1-го рода, которые являются потенциальными точками захвата контура изображения при сканировании бинарной матрицы значений сигналов прибора с зарядовой связью.

В данных точках структура максимально выпуклая и имеет пик развития подструктуры 1-го уровня вложенности при правом обходе контура изображения. Следующие структурные критические точки 1-го рода будут определять уменьшение развития данной выпуклой подструктуры относительно точки захвата до возникновения структурообразующей критической точки 2-го рода (точки максимальной вогнутости структуры), после которой начнется развитие следующей выпуклой подструктуры.

Таким образом, первой выделенной максимальной структурной критической точкой 1-го рода при обходе контура изображения является точка захвата контура изображения $t_{i \max}^1 = \max(t_j^1)$. Для нахождения следующей структурной критической точки 1-го рода, каждый структурный элемент $a_{i,j}^2$, имеющий начало в уже выделенной структурной критической точке сопоставляется с вектором \bar{x}_1^1 единичного направления $g_\gamma = 1$ в относительной системе координат Ξ' .

Строится обобщенное направление развития структуры от уже выявленной максимальной структурной критической точкой 1-го рода до каждой точки 1-го рода по направлению обхода контура. Если существуют обобщенные направления ориентации лежащие как в левой полуплоскости (3 и 4 квадрант), так и в правой полуплоскости (1 и 2 квадрант) системы координат Ξ' , то направление ориентации с минимальным значением для 3 и 4 квадранта полуплоскости будет определять следующую максимальную структурную критическую точку 1-го рода.

Если обобщенное направление ориентации лежит только в правой полуплоскости, то направление

ориентации с минимальным значением для 1 или 2 квадрантов полуплоскости будет определять следующую максимальную структурную критическую точку 1-го рода.

Если при обходе контура не определено ни одного вектора в 3 или 4 квадранте, то все выделенные структурные критические точки данной подструктуры являются максимальными, что определяет выпуклость данной подструктуры.

Для обнаружения возможных вогнутых подструктур данная процедура повторяется для всей последовательности выделенных структурных критических точек 1-го рода.

После выделения максимальных структурных критических точек 1-го рода определяются точки максимальной вогнутости структуры, определяющие переходы между выпуклыми подструктурами. Данные структурные критические точки являются минимальными структурными критическими точками 2-го рода. Для их определения осуществляется последовательное соединение, относительно точки захвата, выделенных максимальных структурных критических точек 1-го рода отрезками, направление ориентации которых совпадает с обобщенным направлением развития данных подструктур (вогнутых подструктур, находящихся между двумя соседними максимальными критическими точками 1-го рода). Данные отрезки представляют собой новые структурные элементы $a_{i,j}^m$ m-го структурного уровня.

Таким образом, строится виртуальный описанный вокруг контура изображения выпуклый многоугольник, ребра которого соответствуют структурным элементам $a_{i,j}^m$. Каждый элемент $a_{i,j}^m$ соединяет две выпуклые подструктуры 1-го уровня вложенности вогнутой структуры контура изображения, для которых необходимо найти точку максимальной вогнутости (критическую точку структурного перехода). Данные структурные критические точки будут являться минимальными критическими точками 2-го рода.

Минимальными структурными критическими точками 2-го рода $t_{i \min}^2 = \min(t_j^2)$ называются такие точки 2-го рода структуры контурного изображения, для которых перпендикуляры p_j , опущенные от соответствующих структурных элементов $a_{i,j}^m$ на данные структурные критические точки, будут иметь максимальную длину для соответствующих выделенных подструктур.

Построение структуры высшего уровня общности по структурным критическим точкам 1-го и 2-го рода, представленное на рис. 2, имеет преимущество по сравнению с известными структурными метода-

ми распознавания контурных изображений, основанных на построении скелетов путем разложения их контуров на ленты за счет того, что определение критических точек в предложенном подходе не связано с нахождением экстремумов кривизны, для отыскания которых необходимо явное задание кривой, что не всегда возможно [2].

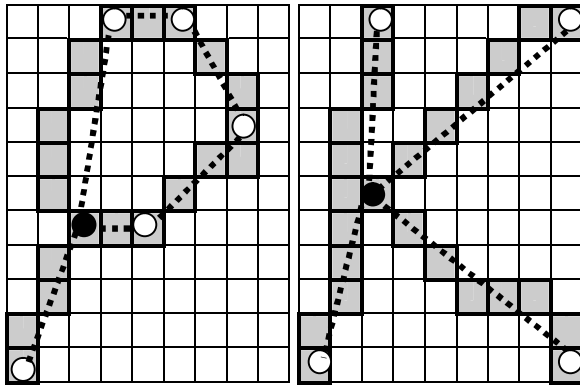


Рис. 2. Структурные критические точки
○ – 1-го рода ● – 2-го рода

Таким образом, концепт Spt распознавания изображения I_i представляет собой структуру z_i^{m+1}

$$Spt(I_i) = z_i^{m+1} = \langle A_i^{m+1}, r, B \rangle,$$

где A_i^{m+1} – множество структурных элементов $m+1$ уровня общности;

$$A_i^{m+1} = \{a_{i1}^{m+1}, a_{i2}^{m+1}, \dots, a_{ik}^{m+1}\}.$$

$$\text{Пусть } T_{\max}^1 = \{t_{1 \max}^1, t_{2 \max}^1, \dots, t_{n \max}^1\},$$

$$T_{\min}^2 = \{t_{1 \min}^2, t_{2 \min}^2, \dots, t_{m \min}^2\}.$$

При отображении h данных множеств T_{\max}^1 и T_{\min}^2 на общую линейную шкалу порядка P

$$h : (T_{\max}^1, T_{\min}^2) \rightarrow P;$$

структурные элементы a_{ij}^{m+1} образуются структурными критическими точками

$$a_{ij}^{m+1} = f(t_{k \max}^1, t_{n \min}^2),$$

где f – функция построения линейного структурного элемента по двум критическим точкам последовательно расположенным на шкале P .

$$\text{Тогда } A_i^{m+1} = F(T_{\max}^1, T_{\min}^2)$$

$$\text{и } Spt(I_i) = \langle F(T_{\max}^1, T_{\min}^2), r, B \rangle.$$

Данная структура инвариантна относительно аффинных преобразований и деформационных искажений контура распознаваемого изображения.

Таким образом, предложенный подход позволяет устранить влияние различного рода аффинных преобразований и деформационных искажений в структуре контура распознаваемого изображения на качество распознавания печатного и рукописного текста. Полученная обобщенная структура контурного изображения образует концепт распознавания, содержащий необходимые и достаточные признаки классов объектов распознавания.

Заключение

Процесс представления исходной структуры контурного изображения в виде структуры высшего уровня общности, обладающей характерными признаками класса распознавания, происходит на основе определения структурных критических точек 1-го и 2-го рода. Максимальные критические точки 1-го рода являются точками максимальной выпуклости структуры, а минимальные критические точки 2-го рода – точками максимальной вогнутости структуры.

Применение предложенного подхода к распознаванию изображений печатного и рукописного текста позволяет уменьшить вычислительные и временные затраты на распознавание и повысить качество распознавания печатного и рукописного текста за счет инвариантности концептуальных структур изображений от деформационных искажений в их контурах.

Список литературы

1. Форсайт Д., Понс Ж. Компьютерное зрение. Современный подход: Пер. с англ. – М.: Издательский дом "Вильямс", 2004. – 928 с.
2. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2006. – 1072 с.
3. Шапиро Л., Стокман Дж. Компьютерное зрение: Пер. с англ. – М.: БИНОМ. Лаборатория знаний, 2006. – 752 с.
4. Гринев Д.В. Классификация и идентификация объектов с использованием структурно-лингвистического метода // Системи обробки інформації. – Х.: ХВУ, 2004. – Вип. 11 (39). – С. 44-48.
5. Паржин Ю.В., Гринев Д.В., Онищенко В.В. Устранение влияния проективных искажений при классификации изображений объектов, полученных искусственным спутником Земли // Радіоелектронні і комп'ютерні системи. – 2006. – № 3 (15). – С. 34-37.
6. Паржин Ю.В., Ковальчук В.С., Гринев Д.В. Структурное распознавание изображений в реальном времени // Збірник наукових праць ІПМЕ ім. Пухова. – К.: ІПМЕ, 2004. – Вип. 25. – С. 143-147.

Поступила в редколлегию 15.03.2007

Рецензент: д-р техн. наук, проф. В.А. Краснобаев, Харьковский Национальный технический университет сельского хозяйства, Харьков.