

ПОДХОД К РАЗРАБОТКЕ СИСТЕМЫ РАСПОЗНАВАНИЯ РУКОПИСНЫХ МАТЕМАТИЧЕСКИХ ВЫРАЖЕНИЙ ВВОДИМЫХ В ЭВМ В РЕЖИМЕ РЕАЛЬНОГО ВРЕМЕНИ

Предложен подход к структурному анализу рукописных математических выражений, основанный на синтаксических правилах и пространственных отношениях между объектами в выражении. Предложена динамическая база эвристических правил. Проведены экспериментальные исследования, позволяющие оценить качество распознавания рукописных математических выражений вводимых в ЭВМ в режиме реального времени.

An approach to structural analysis of handwritten mathematical expressions, based on syntactic rules and spatial relations between objects in an expression. An experimental study to assess the quality of heuristic rules-based recognition of handwritten mathematical expressions in real time.

Введение

Сегодняшнее начало эры пост-ПК, указывает на наличие глобальной тенденции роста использования вычислительных устройств с сенсорными экранами. Появляется необходимость ввода данных в ЭВМ без использования традиционных устройств ввода, таких как клавиатура и мышь. Возможность взаимодействия пользователя с ПК с помощью сенсорной функциональности станет безусловно самой естественной, удобной и быстрой альтернативой для ввода математических выражений в ЭВМ. Центральной задачей проблемы онлайн распознавания рукописных математических выражений является создание эффективного метода классификации рукописных математических символов, цифр и букв. Высокая вариативность в написании одних и тех же символов, и сложность определения структуры математических выражений затрудняют процесс распознавания.

Сегодня наиболее распространенными подходами для классификации рукописных символов являются: алгоритмы сопоставления с образцом [1], скрытые Марковские модели [2] [3], алгоритмы на основе искусственных нейронных сетей [4]. В последнее время начинает активно использоваться аппарат искусственных нейронных сетей с целью решения различных задач классификации. Преимуществом искусственных нейронных сетей является их способность к обобщению полученной информации, таким образом обученная на ограниченном множестве выборочных данных нейронная сеть может вернуть верный результат применительно к данным, которые не участвовали в процессе обучения. Поскольку решение об идентификации рукописных символов принимается системой распознавания в условиях неполной и неточной

информации, можно считать подходящим разработку метода распознавания на основе нечеткого нейронного классификатора, сочетающего в себе преимущества нейронных сетей и нечетких систем логического вывода.

Целью данной статьи является рассмотрение подхода к структурному анализу математических выражений, разработанного с целью повышения качества распознавания рукописных математических выражений вводимых в ЭВМ в режиме реального времени.

Обзор процесса распознавания рукописных математических выражений

Для того чтобы правильно распознать математическое выражение необходимо точно знать значение каждого из символов, которые в него вошли, и определить расположение этих символов относительно друг друга.

Следовательно, задачу распознавания рукописных математических выражений будет целесообразным рассматриваться с точки зрения следующих двух процессов, которые были описаны в [5]:

1. Распознавание символов;
2. Определение пространственных отношений между символами.

В предлагаемой системе распознавания рукописных математических выражений, используется набор символов, который содержит строчные буквы латинского алфавита, наиболее часто встречаемые в формулах строчные буквы греческого алфавита, цифры и специальные математические символы. Отобранные символы позволяют писать тригонометрические и логарифмические функции, интегралы, арифметические корни различной степени и др. Все символы были разделены на две группы:

1. В первую группу вошли символы, которые могут быть написаны без отрыва пера, т.е. с помощью одного отрезка, как 0 и 1. Эти символы участвуют в этапе распознавания.

2. Во вторую группу вошли символы, которые могут быть получены из комбинации символов первой группы. Такие символы будут реконструированы на этапе структурного анализа из нескольких распознанных отрезков.

Для решения задачи распознавания рукописных символов был выбран нечеткий классификатор NEFCLASS [6], поскольку решение об идентификации рукописных символов принимается системой распознавания в условиях неполной и неточной информации. Преимуществом искусственных нейронных сетей является их способность к обобщению полученной информации, таким образом обученная на ограниченном множестве выборочных данных нейронная сеть может вернуть верный результат применительно к данным, которые не участвовали в процессе обучения. Подробно предлагаемый подход к решению задачи распознавания символов был рассмотрен в [7].

Структурный анализ математических выражений

После этапа распознавания символов главной задачей становится определение пространственных отношений между составляющими математического выражения. Предлагается подход к структурному анализу математических выражений, состоящий из трех этапов и позволяющий определить расположение распознанных нейронной сетью символов относительно друг друга. Подход включает в себя этап размещения, этап реконструкции и этап группировки символов, как показано на Рис.1. Написанное пользователем математическое выражение просматривается после каждого внесенного изменения, что дает возможность записывать составляющие математического выражения в любом порядке, а также вносить изменения в уже написанное выражение.

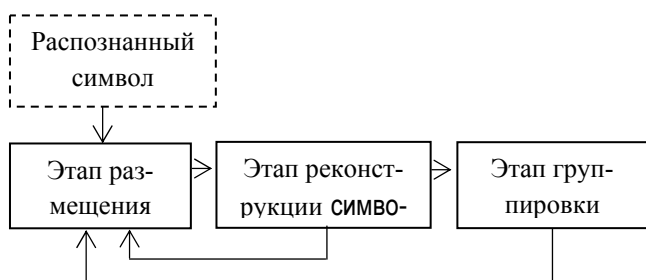


Рис.1 Схема структурного анализа математических выражений

Символы разделены на группы, в зависимости от допустимых и обязательных позиций размещения других символов относительно них. На Рисунке 2 схематически изображены все восемь возможных позиций. Под допустимой позицией расположения символов понимается возможная, но не обязательная к заполнению символом позиция. Под обязательной позицией понимается позиция обязательная к заполнению символом.

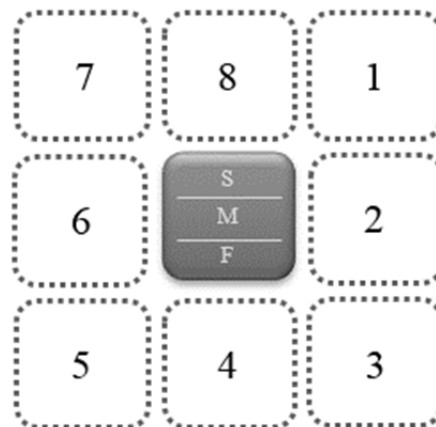


Рис.2 Схематическое изображение допустимых и обязательных позиций размещения символов относительно текущего рассматриваемого символа

На этапе размещения символов наилучшая позиция для размещения символа определяется по формуле:

$$NP = P * k$$

где P – процент попадания символа
k – коэффициент позиции, принимающий следующие значения:

- 0-для недопустимых позиций,
- 1-для допустимых позиций,
- 1,5- для обязательных позиций.

Этап размещения символов работает по следующему принципу:

- Если наилучшая позиция размещения для данного символа принимает значение ноль, то найти наилучшую позицию размещения для соседних символов относительно данного.
- Если ноль, то поместить данный символ на месте ближайшей допустимой или обязательной позиции.

На Рис.3 показан пример определения наилучшей позиции для размещения символов «2» и «3».

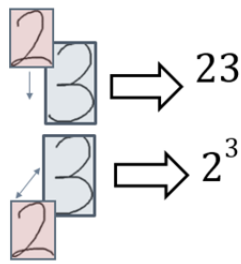


Рис. 3 Пример определения наилучшей позиции для размещения символов

Этап реконструкции символов решает следующие задачи:

1. Преобразование распознанных последовательных отрезков в соответствующий символ. Например, если отрезок «\» пересекает отрезок «/» в позиции M, то отрезки преобра-

зуются в символ «x».

2. Преобразование распознанных последовательных отрезков в соответствующую комбинацию символов, имеющую семантическое значение.

3. Коррекция ошибочно распознанных отрезков, путем нахождения семантического значения комбинации символов в которую они входят.

Для решения поставленных задач используется динамическая база эвристических правил, пример которой приведен на Рис.4. Данная таблица основана на последовательности записи каждого последующего символа, и расположении символов относительно друг друга.

Текущий символ i	Символ S _{i-1}	Позиция s _i относительно S _{i-1}	Символ S _{i-2}	Позиция s _{i-1} относительно S _{i-2}	Результат
-	v	M			∇
-	+	4			±
.	l	8			i
-	l	S			t
-	f	S			f
.	j	8			j
/	\	M			x
\	/	M			x
>	-	2 or F			→
/	-	S or M	c	M	∉
-	C	M			∈
-	Z	M			z
o	O	7	/	7	%
	-	F or M		S	π
l	-	F or M	j	S	π
l	~	F or M	j	S	π
n	1	6	s	6	sin
s	0	6	c	6	cos
m	l	6	l	6	lim

Рис. 4 Пример базы эвристических правил

Рассмотрим этап группировки. Некоторые символы позволяют сгруппировать несколько отдельных символов в одну группу. К таким символам относятся:

- Различного вида скобки,
- Математические аббревиатуры,
- Дробная черта,
- Сумма Σ ,
- Произведение Π ,
- Интеграл,
- Арифметический корень,
- Точка, запятая.

Группа символов рассматривается как единое объединение, для нее существуют допустимые и обязательные позиции размещения соседних символов. На Рис. 5 приведен пример группировки символов с помощью круглых скобок.

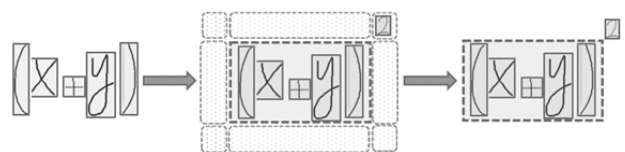


Рис. 5 Пример группировки символов с помощью круглых скобок

Для проверки эффективности работы предложенных подходов были проведены экспериментальные исследования, позволяющие оценить качество распознавания рукописных математических выражений. Программная часть системы реализована на языке C# на базе платформы Microsoft .NET. Преимущество использования платформы .NET Framework заключается в возможности реализации исполняемого кода независимого от аппаратной части, поскольку исходный код переводится компилятором в промежуточный байт-код Common Intermediate Language (CIL) и затем код выполняется виртуальной машиной Common Language Runtime (CLR), которая с помощью встроенного в неё JIT-компилятор преобразует промежуточный байт-код в машинные коды нужного процессора. Кроме того CLR заботится о базовой безопасности, управлении памятью и системе исключений. Необходимо отметить, что использование современной технологии динамической компиляции позволяет достигнуть высокого уровня быстродействия. Результат работы можно сохранить в формате изображения .png или в виде кода языка TEX [8].

Для обучения нечеткого классификатора NEFCLASS была использована выборка, состоящая из 260 рукописных математических выражений. Для генерации базы правил каждому образцу ставился в соответствие подходящий рукописный символ. После чего нейронная сеть была обучена с помощью генетического алгоритма обучения. Таким образом, были найдены значения весовых коэффициентов необходимые для инициализации нейронной сети.

Тестирование проводилось путем рукописного ввода математических выражений в режиме реального времени в планшетный ПК с помощью стилуса. Общая выборка для тестирования составила 140 математических выражений. Среднее время распознавания каждого символа составило 150 миллисекунд на компьютере со следующими характеристиками: CPU Intel Core 2 Quad с частотой 2.4 ГГц, RAM 4 Гб. Результаты тестирования отражены на Рис.6 и отражают

качество распознавания и структурирования математических выражений.

Распознавание отрезков	Реконструкция символов (для правильно распознанных отрезков)	Структурный анализ математических выражений
91,54%	95,32%	71,29%

Рис.6 Результаты тестирования

Поле «Распознавание отрезков» показывает процент правильно классифицированных символов, написанных без отрыва пера с помощью одного штриха, и позволяет оценить эффективность распознавания. Поле «Реконструкция символов» отображает процент верно реконструированных символов из распознанных правильно отрезков. Поле «Структурный анализ математических выражений» показывает процент правильно структурированных выражений.

Выводы

В статье был описан предлагаемый подход к структурному анализу рукописных математических выражений вводимых в ЭВМ в режиме реального времени. Рассмотрена разработанная динамическая база эвристических правил, основанная на знаниях о порядке записи и пространственных отношениях между символами, позволяющая проводить реконструкцию символов, коррекцию неправильно распознанных символов в последовательности путем нахождения ее семантического значения. Для проверки эффективности работы предложенного подхода были проведены экспериментальные исследования, позволяющие оценить качество распознавания рукописных математических выражений. По результатам тестирования были получены следующие результаты: процент правильно классифицированных символов, написанных без отрыва пера с помощью одного штриха составил 91,54%, процент правильно структурированных выражений составил 71,29%.

Список литературы

1. Nakayama Y. A prototype pen-input mathematical formula editor. – EDMEDIA, 1993. - pp. 400-407.
2. Winkler H.-J., Fahrner H., Lang M. A soft-decision approach for structural analysis of handwritten mathematical expressions. – ICASSP, 1995. - pp. 2459-2462.
3. Kosmala A., Rigoll G. On-line handwritten formula recognition using statistical methods. - International Conference on Pattern Recognition, 1998. – pp. 1306-1308.

4. Dimitriadis Y. A new interactive mathematical editor, using on-line handwritten symbol recognition, and error detection-correction with an attribute grammar. – ICDAR, 1991. - pp. 885-893.
5. Dorothea Blostein. General Diagram-Recognition Methodologies. - Queen's University, Canada, 1995. – pp. 1-13.
6. Зайченко Ю.П. Нечёткие модели и методы в интеллектуальных системах. Учебное пособие для студентов высших учебных заведений. – К.: «Издательский Дом «Слово»», 2008. – с. 344.
7. Вісник НТУУ «КПІ» Інформатика, управління та обчислювальна техніка: Зб. наук. пр. – К.:Век+, - 2012. - №57. – с.124-128.
8. Кнут Д. Э. Все про ТЕХ. / Пер. с англ. М. В. Лисиной — Протвино: РДТЕХ, 1993. – с. 592