



Ю.Е. Чуманкин, П.Е. Овчинников

## ВЫБОР ПРИЗНАКОВ ПРИ РАСПОЗНАВАНИИ СИМВОЛОВ В ФОРМУЛАХ

(Нижегородский государственный университет им. Н.И. Лобачевского)

В настоящее время широко распространены системы оптического распознавания символов. Они успешно справляются с распознаванием печатного текста. Но встречаются ситуации, в которых необходимо распознать не обычный текст, а текст, структурированный особым образом, в частности отсканированные формулы. Существуют коммерческие приложения, работающие в этой области, но они обладают существенными недостатками. Эти приложения либо не работают с формулами, обладающими сложной структурой, либо работают с изображениями высокого разрешения (~300 dpi).

Для того чтобы распознавать формулы предложен общий алгоритм, определяющий порядок действий приложения:

1. Определить области расположения отдельных символов.
2. Распознать отдельные символы.
3. Распознать структуру формулы.

В настоящей работе более подробно рассмотрен этап распознавания символов с учетом следующих особенностей, которые возникают при распознавании формул:

1. В индексах встречаются символы очень малых размеров.
2. Символы не сгруппированы в четко определенные строки и их имеется малое количество, поэтому исключается возможность определения угла поворота страницы, основанная на ориентации строк.

Целью работы является реализация распознавания символов с использованием различных признаков. Будут рассмотрены признаки, использующие расстояние Хэмминга и признаки, основанные на вычислении разложения изображения по ортогональному базису (базис Цернике, цилиндрический базис Уолша).

Для распознавания символов необходимо выделить признаки и принять решение о значении символа. Для верного распознавания необходимо выделять признаки, инвариантные относительно сдвига, масштабирования, поворота.

Инвариантность относительно сдвига и масштабирования для рассмотренных типов признаков достигается единым методом. Инвариантность относительно сдвига достигается переходом в систему центра масс изображения символа. Для достижения инвариантности относительно масштабирования необходимо преобразовать размеры распознаваемого символа и размеры символов эталонов таким образом, чтобы они стали одинаковыми.

Чтобы слабые искажения краев символа не вносили существенного изменения в конечное изображение, необходимо масштабировать не только символ, но и некоторую область пустого фона прилегающего к нему. Распознаваемые символы могут быть повернуты, поэтому захватываемая область масштабиро-



вания должна сохраняться при повороте. Такая область представляет собой круг. Для ослабления действия искажений формы символа радиус круга следует выбирать, основываясь на статистических моментах рассматриваемого изображения.

Предложена следующая формула для выбора радиуса захватываемого круга (1).

$$R_0 = \bar{\rho} + \alpha \sqrt{\sigma_\rho} \quad (1)$$

Где  $\bar{\rho}$  и  $\sigma_\rho$  это математическое ожидание и дисперсия расстояния от центра масс до текущего пиксела, а  $\alpha$  это параметр алгоритма, подбираемый экспериментально.

После выполнения всех указанных расчетов и изменения масштаба методы распознавания на основе расстояния Хэмминга и разложения изображения по ортогональному базису работают по-разному.

Расстояние Хэмминга для бинарных сигналов это количество бит, в которых сигналы различны.

Метод заключается в том, чтобы каждому классу поставить в соответствие бинарный сигнал, описывающий типичного представителя каждого класса, это и будет вектор признаков. В качестве такого сигнала можно выбрать сжатое и бинаризованное изображение типичного символа из класса, т.е. изображение не искаженного символа. Принятие решения осуществляется по методу ближайшего соседа [1].

Достоинством описанного метода является простота. Отсутствие затратных математических преобразований при выделении признаков делает метод производительным. Так как метод работает с бинарными изображениями, здесь нет необходимости в вычислениях с плавающей точкой, что еще сильнее ускоряет процесс вычислений.

К недостаткам метода относится отсутствие возможности распознавания повернутых символов.

Моментом порядка  $(n, m)$  называется величина, рассчитанная по формуле (2)

$$A_{nm} = \int V_{nm}(x, y) I(x, y) dx dy \quad (2)$$

Где  $V_{nm}$  является базисной функцией. В работе рассмотрено два базиса: базис Цернике (3) и цилиндрический базис Уолша (4).

$$V_{nm}(x, y) = V_{nm}(\rho, \theta) = \sqrt{\frac{n+1}{\pi}} \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)!}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!} \rho^{n-2s} \exp(jm\theta) \quad (3)$$

$$V_{nm}(x, y) = V_{nm}(\rho, \theta) = \sqrt{\frac{1}{\pi}} W_n(\rho) \exp(jm\theta) \quad (4)$$

Здесь  $n$  и  $m$  это целые числа,  $j$  – мнимая единица,  $W_n$  функция Уолша порядка  $n$ .

Функции определены на круге единичного радиуса и образуют полную систему ортогональных функций. Поэтому, по ним возможно разложить любую функцию, определенную на круге единичного радиуса.



За счет ортогональности коэффициенты разложения слабо коррелированы между собой. Это значит, что в небольшом количестве коэффициентов несетя большая часть информации о раскладываемой функции. В рассматриваемой задаче раскладывается функция интенсивности изображения  $I$ .

Практика показывает, что достаточную для распознавания информацию несут моменты порядка  $n \leq 12$  для моментов Цернике и  $n \leq 16$  для моментов Уолша.

Описанные выше системы полиномов были выбраны из-за необычного свойства получаемых моментов. Можно показать, что модули моментов не зависят от угла поворота исходного изображения [2]. А значит, распознавание с использованием модулей моментов будет одинаково успешным при различных углах поворота эталона и распознаваемого изображения. Для принятия решения используется метод наименьших квадратов в пространстве модулей моментов [2].

Рассмотренный метод принятия решения никаким образом не учитывает информацию о фазе моментов. Это не позволяет различить между собой символы, различающиеся лишь углом поворота друг относительно друга (такие символы как «p» и «d» или «q» и «b»).

Достоинством описанного метода является возможность распознавания повернутых изображений. При этом фаза момента пропорциональна углу поворота изображения, что позволяет не только распознавать повернутые изображения, но и находить угол поворота. Недостатком метода являются сложные вычисления с плавающей точкой, требующие затрат компьютерного времени.

Для тестирования предложенных методов было разработано приложение, способное распознавать изображения букв. Для составления выборки эталонов использовались изображения латинских строчных и прописных букв набранных шрифтами «Times New Roman» и «Cambria». Для тестирования приложения было составлено 9 выборок. Каждая из которых, состоит из 50 бинарных изображений строчных и прописных букв латинского алфавита, набранных шрифтами «Times New Roman» и «Cambria».

В каждой тройке выборок использовался свой диапазон высот символов. Выборки сгруппированы тройками для оценки погрешности определения точности. Если в выборки входят символы одинаковых размеров и шрифтов, то эксперименты над каждой выборкой можно считать независимыми. А вероятность верного распознавания символа, полученную в каждом эксперименте, можно усреднить и вычислить ее погрешность. При всех дальнейших расчетах использовалась доверительная вероятность  $\gamma=0.95$ . Обработка результатов эксперимента проводилась по методике, изложенной в [3]. Результаты экспериментов приведены в таблице 1.

Все рассмотренные методы справляются с распознаванием искаженных символов. Тестирование проводилось в широком диапазоне размеров символов. Системы признаков остаются достаточно информативными даже при очень малых размерах символа, меньше 16 пикселей. Что важно, при распознавании формул. По полученным результатам видно, что наилучшие результаты распо-



знавания показывают методы с использованием моментов Цернике и расстояния Хэмминга. При этом метод с использованием расстояния Хэмминга делает это несколько точнее. Более высокую точность метода Хэмминга можно объяснить учетом информации о повороте. При распознавании не повернутых изображений предпочтение следует отдавать методу с использованием расстояния Хэмминга. Но при распознавании повернутых изображений, как в случае распознавания символов в формулах, этот метод не применим, в отличие от метода с использованием моментов Цернике.

Таблица 1. Вероятности верного распознавания символа

	Более 40 пикселей	От 16 до 40 пикселей	Менее 16 пикселей
Моменты Цернике	[0.94;0.99]	[0.90;0.98]	[0.85;0.93]
Моменты Уолша	[0.90;98]	[0.85;0.95]	[0.78;86]
Расстояние Хэмминга	[0.94;0,99]	[0.93;0,99]	[0.90;0.98]

### Литература

1. Lippmann R.P. An Introduction to Computing with Neural Nets // IEEE ASSP magazine, 1987. Pp. 4-22.
2. Khotanzad A., Hua Hong Y., Invariant Image Recognition by Zernike Moments. // IEEE Trans. Pattern Anal. Machine Intell. Vol. 12, no. 5, 1990. Pp. 489-497.
3. Фаддеев М.А. Элементарная обработка результатов эксперимента. – Н. Новгород: Нижегородский государственный университет, 2002 – 108 с.

Р.А. Шаталин, П.Е. Овчинников

### АЛГОРИТМ ОБНАРУЖЕНИЯ НЕШТАТНЫХ СИТУАЦИЙ В ЗАДАЧАХ ВИДЕОНАБЛЮДЕНИЯ НА ОСНОВЕ МЕТОДА ГЛАВНЫХ КОМПОНЕНТ

(Нижегородский государственный университет им. Н.И. Лобачевского)

Видеонаблюдение приобретает все большее значение в нашей повседневной жизни, предоставляя важную информацию для сохранения безопасности, как в общественных местах, так и на частных территориях. Современные системы видеонаблюдения позволяют свести на несколько мониторов данные со всех камер, записать и просмотреть их в случае необходимости. Из этого следует, что такие системы создают огромный поток данных, хотя в реальности потенциально опасные ситуации возникают достаточно редко. Более того, человек оказывается неспособен удерживать внимание на видеокадрах, где многие события являются обыденными и повторяющимися, что в свою очередь снижает шанс обнаружения нештатных ситуаций. Эту ситуацию можно исправить, если сосредотачивать внимание оператора только на существенных изменениях в сцене и необычном поведении людей. Основными требованиями к таким си-