

# Распознавание символов на базе цепи Маркова

**Г. А. Юрьев,**

*программист кафедры прикладной информатики факультета информационных технологий Московского городского психолого-педагогического университета*

Рассматривается принципиально новый алгоритм распознавания символов, опирающийся на возможности цепей Маркова [1; 5] – марковских моделей с дискретными состояниями и дискретным временем. Отмечается, что используемый аппарат широко применяется при решении задач распознавания, но традиционно используется несколько иначе. Преимуществом метода является высокая скорость настройки (обучения), возможность задать произвольную, требуемую надежность результата и изменять ее в процессе работы программной системы. Алгоритм успешно реализован в системе распознавания и озвучивания плоскочечных текстов EI-Reader [2; 3; 4]. Обращается внимание, что его основным преимуществом является устойчивость к изменению гарнитуры шрифта. Описаны предпосылки создания, дано теоретическое обоснование, а также описание алгоритма, использованного в программной реализации. Дана точная статистическая оценка надежности распознавания при заданных параметрах. Подчеркивается, что алгоритм имеет ряд серьезных преимуществ перед традиционными подходами, в случае работы с искаженными изображениями.

**Ключевые слова:** цепи Маркова, распознавание, вейвлет-преобразование, анализ изображения.

## Введение

В процессе разработки системы распознавания и озвучивания плоскочечных текстов EI-Reader [2; 3; 4] возникла необходимость в использовании алгоритма, устойчивого к незначительным изменениям шрифтов и искажениям, возникающим в процессе выделения и обработки изображения. Анализ существующих подходов по-

казал, что их реализация на базе основного инструмента разработки не будет обладать достаточным быстродействием (специфика задачи заключалась в том, что необходимо было обрабатывать порядка 20 символов за секунду). Было принято решение разработать алгоритм, который сможет принимать решения на основании минимального объема информации о символе.

---

\* grinch89@mail.ru

Также оказалось необходимым обеспечить вероятностный подход в силу требований, предъявленных к гибкости алгоритма. Для реализации этих функций был выбран аппарат марковских цепей [2; 5].

### **Предварительная обработка изображения**

Введенное изображение преобразуется в монохромное черно-белое представление.

Важной операцией, необходимой для корректной реконструкции связанного текста, является выделение из изображения символьных строк. Задача усложняется тем, что камера, как правило, удерживается в руке и, вследствие этого, строки попадают в кадр повернутыми. Для повышения надежности распознавания при достаточно больших углах поворота необходимо восстанавливать горизонтальное положение строк. Для этого на основе анализа плотностей черных и белых точек на нескольких вертикальных срезах изображения определяют местоположение начал и концов строк, что позволяет оценить угол, на который следует повернуть графическое представление для получения удобного для анализа состояния. После такого поворота границы текстовых строк определяются повторно. На основе информации об этих границах фиксируются занимаемые символами прямоугольные фрагменты изображения.

По окончании выделения указанных прямоугольных фрагментов для всех символов в строке соответствующий ей участок вырезается из кадра, просматриваемого веб-камерой. При этом малые фрагменты, не превышающие заданный порог, удаляются из строки. Оставшиеся фрагменты рассматриваются как области, содержащие распознаваемые символы произвольного размера. Указанная процедура повторяется для всех строк из кадра, просматриваемого веб-камерой, причем информация о геометрических характеристиках как строк, так и прямоугольных фраг-

ментов изображения, включающих в себя символы, сохраняется для последующего распознавания в порядке расположения фрагментов в кадре.

Ослабить искажения в анализируемой части изображения позволяет алгоритм коррекции локального контраста, а именно, уменьшить эффект, возникающий в связи с неравномерным освещением. При переводе в черно-белый формат невозможно установить общий уровень контраста для всего изображения, так как либо часть символов распадется и распознать их будет невозможно, либо некоторые окажутся закрашены черным целиком. Суть метода в следующем: на выделенных прямоугольных фрагментах перебираются все черные точки, из числа которых удаляются слабосвязанные (имеющие малое число одноцветных соседей). В то же время области, в которых предполагается исчезновение черных точек из-за нарушения освещенности или иных условий сканирования, заливаются черным цветом<sup>1</sup>. Указанный прием улучшает качество распознавания примерно на 10 %, практически не влияя на скорость работы программы.

Изображение, вводимое с веб-камеры, обычно затемнено с одного или нескольких краев и, несмотря на все использованные алгоритмы, с таким уровнем затемнения бороться бессмысленно. Эффект обусловлен непараллельностью поверхности текста и линзы, а также оптическими искажениями и недостаточным освещением. В случае значительных искажений анализируется только центральная часть графического представления.

### **Описание алгоритма**

Каждый прямоугольник с символом переводится в контурный рисунок с помощью непрореженного вейвлет-преобразования на базе биортогонального вейвлета, прямоугольные фрагменты с контурными изображениями последовательно преобразуются в два числовых вектора. Один из них описывает распределение точек кон-

<sup>1</sup>Примером может служить одна белая точка, все соседи которой в некоторой окрестности являются черными.

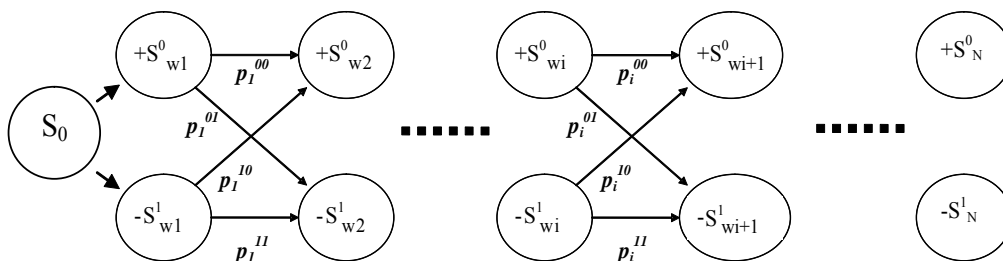
тура вдоль оси абсцисс, другой вдоль оси ординат. Таким образом, мы, теряя часть информации, получаем простой инвариант длиной в 64 числа (символы отмасштабированы для окна 32x32), такой вектор может быть проанализирован с удовлетворяющей для задачи скоростью. Полученный вектор подвергается быстрому вейвлет-преобразованию с использованием вейвлетов Добеши 4-го порядка. Полученные подмножества последовательностей вейвлет-коэффициентов  $W=\{w_{ij}|i=1,\dots,N\}$  рассматриваются как представления символов и анализируются с помощью расположенной цепи Маркова, структура которой представлена на рисунке. Процедура распознавания основана на анализе знаков вейвлет-коэффициентов из указанной последовательности.

Система, описываемая данной цепью, функционирует  $N$  моментов времени, где  $N$  – длина анализируемой последовательности вейвлет-коэффициентов. В начальный момент времени система находится в состоянии  $S_0$ . В  $i$ -й момент времени, где  $i$  изменяется от 1 до  $N$ , система может быть или в состоянии  $+S_i^0$ , или в состоянии  $-S_i^1$ , а именно: в состоянии  $+S_i^0$  – в случае неотрицательности  $i$ -го элемента последовательности  $W$  и в состоянии  $-S_i^1$  – в случае его отрицательности. Система является нестационарной: изменяется матри-

ца  $P_i=\|p_{i,j}\|$  вероятностей переходов между состояниями, такая, что  $p_i=P_i p_{i-1}$ , где  $p_i$  – вектор-столбец вероятностей пребывания в состояниях цепи в  $i$ -й момент времени, причем все ее элементы равны нулю, кроме вероятностей переходов между парами состояний  $S_{i-1}^1$  и  $S_i^1$ ,  $S_{i-1}^0$  и  $S_i^0$ ,  $S_{i-1}^1$  и  $S_i^0$  и  $S_{i-1}^0$  и  $S_i^1$ , которые далее обозначены, соответственно, как  $p_i^{11}$ ,  $p_i^{01}$ ,  $p_i^{10}$  и  $p_i^{00}$ . В каждый из моментов времени  $i$  выполняются следующие нормировочные условия:  $p_i^{11}+p_i^{10}=1$  и  $p_i^{01}+p_i^{00}=1$ .

Последовательности  $W_a$ , соответствующей заданному символу  $a$ , приведенные правила ставят в соответствие определенную последовательность  $S_a$  прохождения состояний рассматриваемой цепи Маркова<sup>2</sup>.

При обучении цепи Маркова на алфавите  $A=\{a_m\}$  для каждого символа  $a_m$  формируется множество его эталонов<sup>3</sup>  $a_{mk}$ ,  $k=1,\dots,K_m$ . Каждому символу  $a_m$  ставится в соответствие цепь Маркова, вероятности переходов между состояниями которой определяются методом максимального правдоподобия. Для этого вычисляется распределение вероятностей переходов, обеспечивающее наибольшую вероятность  $P$  появления заданного множества последовательностей состояний, соответствующих эталонам данного символа:  $P=\prod_{k=1,\dots,K_m} P_k$ , где  $P_k=\prod_{i=1,\dots,N} p_i^{signWi-1,k signWi,k}$ , где  $k$  – номер эта-



Структура цепи Маркова

<sup>2</sup>Следует заметить, что рассматриваемая цепь Маркова может быть описана и в терминах конечного автомата.

<sup>3</sup>Эталоны, как правило, соответствуют различным вариантам изображения символов, принятым в разных шрифтах.

лона символа, а  $i$  – номер состояния, причем полагается, что  $\text{sign}w_{0,k}=1$ .

Указанная задача может быть решена одним из подходящих численных методов нелинейного программирования<sup>4</sup>. Часто удобно, с целью упрощения решения, перейти к критерию  $\text{In}P$ , монотонно связанному с исходным критерием  $P$ , переформулировав исходную постановку в задачу оптимизации для величин  $\text{In}p_i^{\text{sign}w_{i-1,k} \text{sign}w_{i,k}}$ , на которые наложено условие неположительности<sup>5</sup>.

Распознавание поступающих на вход системы новых символов возможно после завершения обучения. При этом для цепи Маркова, соответствующей каждому из символов алфавита  $A$ , вычисляется вероятность прохождения последовательности состояний, определяемой вейвлет-коэффициентами  $\{w_{i,\mu}\}_{i=1,\dots,N}$  распознаваемого символа, которая равна  $\prod_{i=1,\dots,N} p_i^{\text{sign}w_{i-1} \text{sign}w_i}$ . Тип символа, который соответствует цепи Маркова, имеющей наибольшую вероятность такого рода, выдается как результат распознавания.

Оценка вероятности корректного распознавания символов с помощью описанного алгоритма, полученная по результатам 1000 проб на базе асимптотической аппроксимации биномиального распределения, равна 0,791, причем нижняя и верхняя границы 95 % доверительного интервала равны 0,76 и 0,83.

Оказалось, что достаточно использовать последовательности  $W$ , состоящие из двенадцати вейвлет-коэффициентов ( $N=12$ ), начиная с третьего коэффициента разложения распределений контурных точек. Расчеты выявили неинформативность первых двух вейвлет-коэффициентов этого разложения для решения задачи распознавания.

### Результаты и выводы

Разработан и программно реализован новый алгоритм распознавания символов на базе цепей Маркова. Даны формальные оценки его эффективности. Алгоритм используется как один из ключевых компонентов системы распознавания и озвучивания плоскочечатных текстов EI-Reader. Его преимуществами является:

- устойчивость к искажениям, вызванным особенностями процесса сканирования;
- устойчивость к различным типам шрифтов;
- высокое быстродействие;
- возможность динамического уточнения текущей гипотезы за счет использования большего числа коэффициентов при анализе.

Алгоритм обладает высокой универсальностью и имеет широкие перспективы в решении задач распознавания.

<sup>4</sup> Для этого, в частности, может быть использован модуль Solver электронной таблицы Microsoft Excel.

<sup>5</sup> Поскольку  $0 \leq p_i^{\text{sign}w_{i,k} \text{sign}w_{i+1,k}} \leq 1$  для всех  $i$  и  $k$ .

### Литература

1. Куравский Л. С., Юрьев Г. А. Распознавание и озвучивание текстов для облегчения обучения людей с нарушениями зрения // Психологическая наука и образование. 2009. № 5.
2. Куравский Л. С., Юрьев Г. А. Технология распознавания и озвучивания текстов для людей с нарушениями зрения // Нейрокомпьютеры: разработка, применение. 2009. № 9.
3. Овчаров Л. А. Прикладные задачи теории массового обслуживания. М., 1969.
4. Свид. № 2009613028 об офиц. рег. прогр. для ЭВМ. EI-Reader. Программное обеспечение распознавания и озвучивания текстов для людей с нарушениями зрения / Г. А. Юрьев, Л. С. Куравский. М.: РосПатент, 2009.
5. Kuravsky L. S., Baranov S. N. Synthesis of Markov networks for forecasting fatigue failures. In: Proc. Condition Monitoring 2003, Oxford, United Kingdom, July 2003.

## Character Recognition Based on the Markov Chains

G. A. Yuriev,

*Programmer, Chair of Applied Computer Informatics, Department of Information Technologies, Moscow State University of Psychology and Education*

A fundamentally new algorithm for character recognition is described, which is based on the abilities of the Markov chains [1; 5] – the Markov models with discrete states and discrete time. The applied apparatus is widely used in solving the problems of recognition, but it is noted that traditionally it has been used somewhat differently. The advantage of this method is the high speed of the tuning (teaching), the ability to specify the arbitrary and required reliability of the result and to modify it in the work process of the program system. The algorithm is successfully implemented in the EI-Reader [2; 3; 4] – a system of recognition and vocalization of the flat bed texts. The main advantage of it is its resistance to change in the font styles. The prerequisites for the development, theoretical justification and description of the algorithm used in the software implementation are provided in the article. The accurate statistical evaluation of the reliability of recognition for the given parameters is presented. It is emphasized that the algorithm has a number of advantages over traditional approaches, in case of working with the distorted images.

**Keywords:** Markov chains, recognition, wavelet transformation, image analysis.

### References

1. Kuravskij L. S., Jur'ev G. A. Raspoznavanie i ozvuchivanie tekstov dlja oblegchenija obuchenija ljudej s narushenijami zrenija // Psihologicheskaja nauka i obrazovanie. 2009. № 5.
2. Kuravskij L. S., Jur'ev G. A. Tehnologija raspoznavanija i ozvuchivanija tekstov dlja ljudej s narushenijami zrenija // Nejrokomputery: razrabotka, primenenie. 2009. № 9.
3. Ovcharov L. A. Prikladnye zadachi teorii masovogo obsluzhivaniya. M., 1969.
4. Svid. № 2009613028 ob ofic. reg. progr. dlja JeVM. EI-Reader. Programmnoe obespechenie raspoznavanija i ozvuchivanija tekstov dlja ljudej s narushenijami zrenija / G. A. Jur'ev, L. S. Kuravskij. M.: RosPatent, 2009.
5. Kuravsky L. S., Baranov S. N. Synthesis of Markov networks for forecasting fatigue failures. In: Proc. Condition Monitoring 2003, Oxford, United Kingdom, July 2003.