

УДК.:044.93'1

РАСПОЗНАВАНИЕ ЗРИТЕЛЬНЫХ ОБРАЗОВ

Батырканов Жениш Исакунович, доктор технических наук, профессор Кыргызского государственного технического университета им. И.Раззаков, 720044, Кыргызская Республика, г. Бишкек, пр. Ч. Айтматова 66, ORCID 0000-0002-5619-7178, e-mail: bjenish@mail.ru

Кудакеева Гулида Маданбековна, старший преподаватель Кыргызского государственного технического университета им. И.Раззаков, 720044, Кыргызская Республика, г. Бишкек, пр. Ч. Айтматова 66, ORCID 0000-0002-3727-3534, e-mail: gulida87_87@mail.ru

Асиев Абай Турусбекович к.т.н., заведующий отделом науки Кыргызского государственного технического университета им. И.Раззаков, 720044, Кыргызская Республика, г. Бишкек, пр. Ч. Айтматова 66, ORCID 0000-0002-2013-6314, e-mail: asievat@gmail.com

Проблема распознавания является ключевой проблемой при построении большинства интеллектуальных систем. Это при построении интеллектуальных роботов. Это при создании различного рода идентификационных систем.

В последнее время большое внимание уделяется построению систем мониторинга и прогнозирования природных явлений и ресурсов по аэрофотоснимкам. Во всех вышеперечисленных случаях решается проблема распознавания зрительных образов.

В данной работе предлагается один подход к решению задачи распознавания зрительных образов.

При этом под образом понимается конкретное проявление распознаваемого объекта. В работе предлагается подход распознавания зрительных образов на основе использования базы эталонных образов и процедуры обучения. Предлагаемый нами подход можно применять для распознавания рукописного текста, для идентификации личности по биометрическим данным, для мониторинга окружающей среды и для других задач. Предложена структура системы распознавания. Рассмотрены модельные примеры на распознавание зрительных образов предлагаемым подходом.

Ключевые слова: Образ, объект распознавания, двоичный образ, зрительный образ, эталонный образ, считывающее устройство, алгоритм распознавания, Эвклидова норма, скалярное произведение.

RECOGNITION OF VISUAL IMAGES

Batyrkanov Zhenish I., Professor, Doctor of Technical Science, Kyrgyz technical University named after I. Razzakov, 66 Ch. Aitmatova Avenue, Bishkek, 720044, Kyrgyz Republic, ORCID 0000-0002-5619-7178, e-mail: bjenish@mail.ru

Kudakeeva Gulida. M., senior teacher, Kyrgyz technical University named after I. Razzakov, 66 Ch. Aitmatova Avenue, Bishkek, 720044, Kyrgyz Republic, ORCID 0000-0002-3727-3534, e-mail: gulida87_87@mail.ru

Asiev Abai T., KSTU. I.Razzakova, Bishkek, Kyrgyz Republic, 720044 Kyrgyzstan, Bishkek city, Chingiz Aytmatova Ave. 66, Cand.Sci. (Techn.), Head of the Department of Science, ORCID 0000-0002-2013-6314, E-mail: asievat@gmail.com

The problem of recognition is a key problem in the construction of most intelligent systems.

This is when building intelligent robots. This is when creating a different kind of identification systems.

Recently, much attention has been paid to the construction of systems for monitoring and forecasting natural phenomena and resources from aerial photographs. In all the above cases, the problem of recognizing visual images is solved.

In this paper, we propose one approach to solving the problem of recognizing visual images..

In this case, the image is understood as the concrete manifestation of the recognized object. In this paper, the approach of recognizing visual images based on the use of the reference image base and the training procedure is proposed. The approach we offer can be used to recognize handwritten text, to identify a person by biometric data, to monitor the environment and for other tasks. The structure of the recognition system is proposed. Model examples for the recognition of visual images by the proposed approach are considered.

Keywords: Image, object of recognition, binary image, reference image, reader, recognition algorithm, Euclidean norm, scalar product.

Введение. На сегодняшний день на практике, во многих областях ведутся работы по разработке устройств, методов, алгоритмов и программ для распознавания зрительных образов. Это системы: автоматизации; робототехники; системы связи; компьютерные сети; интеллектуальные системы и т. д.

К задачам распознавания зрительных образов можно отнести задачи по распознаванию рукописного текста, распознавание природных и техногенных катастрофических явлений по аэрофотоснимкам, разведку полезных ископаемых, идентификацию личностей по биометрическим данным и др. Все вышеперечисленные работы относятся к задачам распознавания зрительных образов. При этом под образом понимается конкретное проявление изучаемого объекта.

Существуют различные подходы и методы распознавания зрительных образов: распознавание на основе оценок определяющих признаков; распознавание на основе аналогий; распознавание на основе обучения и использования эталонных образов. Среди этих подходов наиболее универсальным и эффективным, на наш взгляд, является подход на основе использования эталонных образов и обучения.

Ниже предлагается следующая структура системы распознавания зрительных образов (рис. 1.)

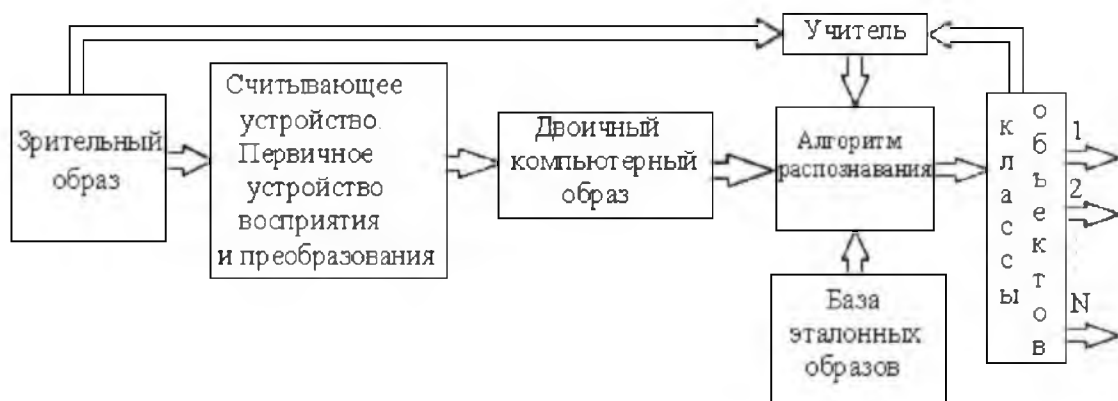


Рисунок 1. Структурная схема системы распознавания зрительных образов

Система распознавания работает следующим образом: зрительный образ предъявляется считывающему устройству, считывающее устройство переводит зрительный образ в двоичный компьютерный образ.

В качестве считывающего устройства могут служить: сканирующее устройство, цифровой фотоаппарат, смартфон, рецепторная матрица и другие специальные устройства. Двоичный образ далее последовательно сравнивает со всеми двоичными эталонными. При этом, при сравнении вычисляются количественные оценки близости, например: по евклидовой разности сравниваемых двоичных образов. Тот эталонный двоичный образ, для которого получилась минимальная величина евклидовой разности, и определяет тот класс объекта, к которому принадлежит предъявленный образ для распознавания.

Рассмотрим работу системы распознавания на задаче распознавания рукописного текста.

При распознавании рукописного текста алгоритм распознавания состоит в распознавании букв предъявленного слова, а затем в сборке распознанных букв в слово, которое затем ищется в словаре.

Алгоритм распознавания, как для рукописного текста, так и для распознавания личности по биометрическим данным осуществляется по структуре, которая представлена выше.

Процедура обучения в данном подходе состоит в корректировке алгоритма распознавания на этапе разработки и обучения системы. Рассмотрим, конкретно, подход к распознаванию рукописных букв.

Процедура подготовки и обучения системы.

В самом начале последовательно на рецепторную матрицу считывающее устройство подается определенное количество образов первого объекта «А». При предъявлении конкретного образа A^1 , активируются определенные ячейки рецепторной матрицы. Активацию ячеек матрицы, можно представить в виде конкретной числовой матрицы, при этом размер числовой матрицы определяется размером рецепторной матрицы. Имеем числовую матрицу типа

$$A = \begin{pmatrix} a_{11}^1 & a_{12}^1, \dots, a_{1N}^1 \\ a_{21}^1 & a_{22}^1, \dots, a_{2N}^1 \\ \dots & \dots & \dots \\ a_{M1}^1 & a_{M2}^1, \dots, a_{MN}^1 \end{pmatrix}, \quad (1)$$

где $a_{ij}^1 = \begin{cases} 0, \\ 1. \end{cases}$

Далее, предъявляется образ A^2 объекта «А» и так до последнего возможного образа A^L объекта «А». Вся это совокупность образов A^1, A^2, \dots, A^L объекта «А» записывается при помощи учителя в память компьютера, как база данных эталонных образов объекта «А». Вся вышеописанная процедура повторяется для всех объектов В, ... ,С. В итоге, в компьютере создается база данных эталонных образов объектов А, В, ... ,С

$$\begin{aligned} A &= \{A^1, A^2, \dots, A^L\}, \\ B &= \{B^1, B^2, \dots, B^P\}, \\ C &= \{C^1, C^2, \dots, C^M\}. \end{aligned} \quad (2)$$

После этапа обучения система готова к распознаванию предъявленных образов.

Алгоритмы распознавания.

Предлагается два варианта алгоритма распознавания.

1-й вариант алгоритма распознавания.

Процедура классификации, т.е. отнесение предъявленного образа к тому или иному объектам происходит по следующей схеме.

При предъявлении рецепторной матрице, например, образа «Х» в компьютере образуется числовой образ в виде матрицы

$$x = \begin{pmatrix} X_{11} & X_{12}, \dots, X_{1N} \\ X_{21} & X_{22}, \dots, X_{2N} \\ \dots & \dots \\ X_{M1} & X_{M2}, \dots, X_{MN} \end{pmatrix} \quad (3)$$

Чтобы отнести образ «X» к тому или иному классу объектов A, B, ... C, осуществляется процедура последовательного вычисления нормы разностей матрицы «X» со всеми матрицами

$A^1, A^2, \dots, A^L, B^1, \dots, C^1, \dots, C^M$, которые хранятся в базе эталонных образов объектов A, B, ... C. Удобно норму вычислять в виде Эвклидовой нормы разностей матрицы (1), (2), (3) в виде

$$\|X - A^k\| = \left(\sum_i \sum_j (X_{ij} - a_{ij}^k)^2 \right)^{1/2} \quad (4)$$

Среди этого множества (4) существует одна минимальная норма, пусть, например минимальная норма разностей будет для образа B^s

$$\|X - B^s\| = \left(\sum_i \sum_j (X_{ij} - B_{ij}^s)^2 \right)^{1/2} \quad (5)$$

Тогда компьютер активирует выход «B» выходного слоя, это говорит о том, что предъявленный образ «X» относится к объекту «B».

Описанный алгоритм классификации (распознавания) работает не только в случаях, когда матрица «X» совпадает с одним из A^1, \dots, C^M но и работает в тех случаях когда «X» не совпадает ни с одним из них. Но и здесь, согласно «принципа близости», классификация проходит успешно. Суть «принципа близости» в теории распознавания заключается в том, что образы для одного и того же объекта в пространстве признаков расположены близко друг к другу.

2-й вариант алгоритма распознавания.

Для этого алгоритма все матрицы $A^1, A^2, \dots, A^L, B^1, \dots, C^1, \dots, C^M, X$ переводятся в векторную форму, например, в виде

$$\bar{X} = (X_{11}, X_{12} \dots X_{1N}, X_{21}, \dots, X_{2N}, \dots, X_{M1}, \dots, X_{MN})^T.$$

Далее осуществляется вычисление скалярных произведений вектора \bar{X} со всеми векторами $\bar{A}^1, \dots, \bar{C}^M$.

Скалярное произведение вектора изображения \bar{X} на вектора эталонов $\bar{A}^1, \dots, \bar{C}^M$, характеризует их сходство, при этом изображение «X» относят к тому классу эталонов, для которого одно скалярное произведение среди $(\bar{X}, \bar{A}^1), \dots, (\bar{X}, \bar{C}^M)$ дало максимальное значение.

Для более точной достоверной работы распознающего устройства необходимо увеличивать количество ячеек рецепторной матрицы, а также увеличивать базу эталонных образов распознаваемых объектов. Для более точной достоверной работы распознающего устройства необходимо увеличивать количество ячеек рецепторной матрицы, а также увеличивать базу эталонных образов распознаваемых объектов. Это необходимо, более точно отображать двоичный цифровой образ распознаваемых объектов. При распознавании рукописного текста алгоритм заключается в последовательном распознавании букв представленного слова, а затем в сборке.

Зрительным образом выступает написанный вручную текст, который набирается на считывающем устройстве (рис.2.).

СПАСИБО

Рисунок 2. Зрительный образ

Основной составляющей системы алгоритма распознавания является база эталонных образов. Своего рода огромная библиотека, состоящая из всех букв алфавита. Т.к. в мире существует огромное количество подчерков, не говоря уже о количестве языков, эта база должна хранить в себе достаточно большое количество всевозможных вариантов написания каждой буквы. Машина не обладает способностью распознавать подчерк, как это делает человек на интуитивном уровне и также, основываясь на свой огромный опыт. Поэтому и существует эта база, которая помогает распознавать текст или слово, разбивая его на буквы и рассматривая каждую букву отдельно.

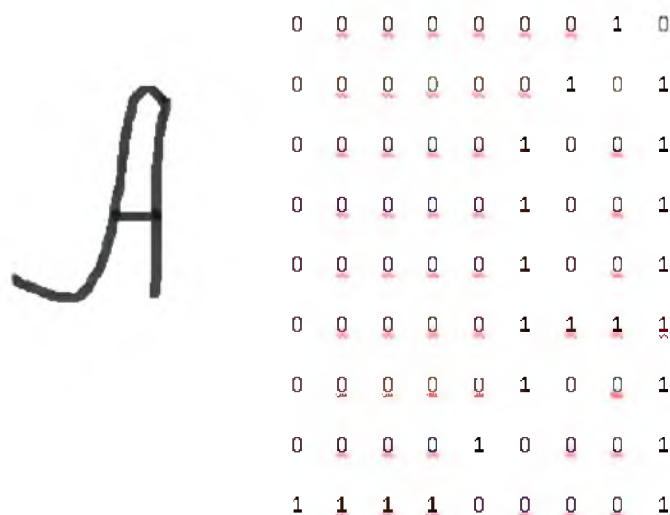
Также база эталонных образов способна постоянно пополнять эталоны, бесконечно запоминая написания букв различных подчерков, чем совершенствует процесс распознавания.

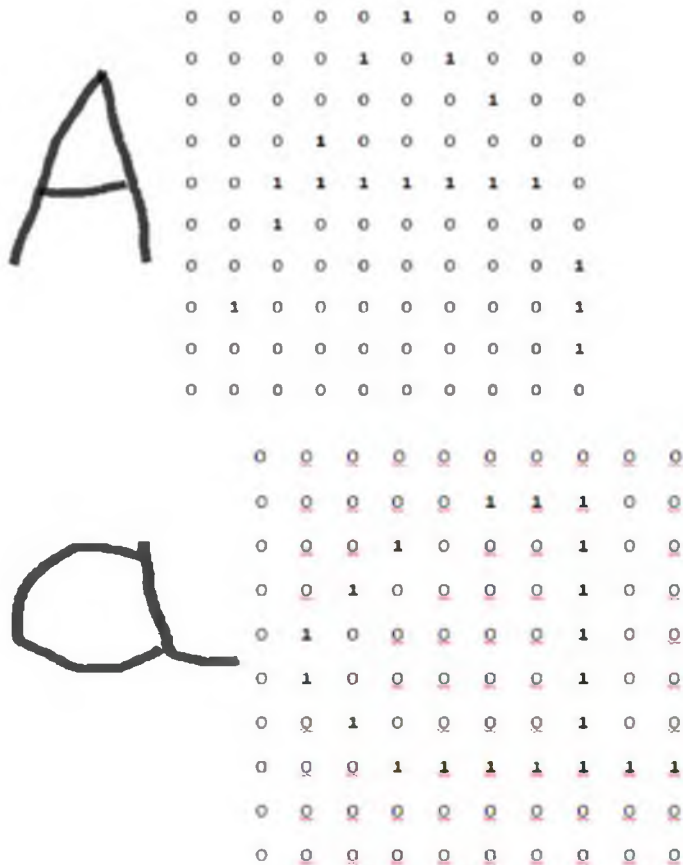
Эталоны хранятся в базе в виде матриц, состоящих из 1и 0. Размерность матриц для примера была взята 10x10.

Ниже представлены примеры, того в каком виде хранятся эталоны в базе образов (рис.3.):



Рисунок 3. Разные варианты написания буквы А и их двоичные образы





Описание алгоритма распознавания рукописного слова

1. Слово разбивается на буквы:



Рисунок 4. Разбиение слова на буквы

2. Далее следует алгоритм распознавания каждой буквы отдельно

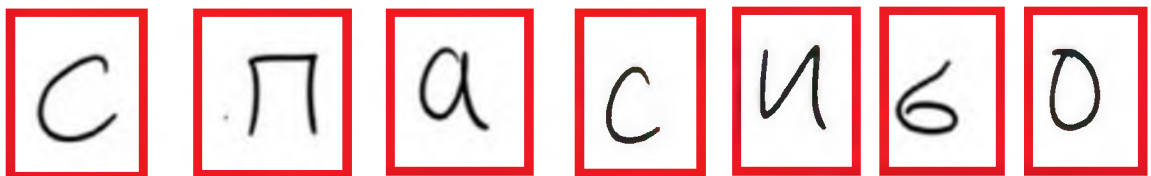


Рисунок 5. Распознавание каждой буквы отдельно

3. Алгоритм обращается к базе эталонных образов и происходит процесс сравнения со всеми эталонами согласно вышеописанной процедуре.

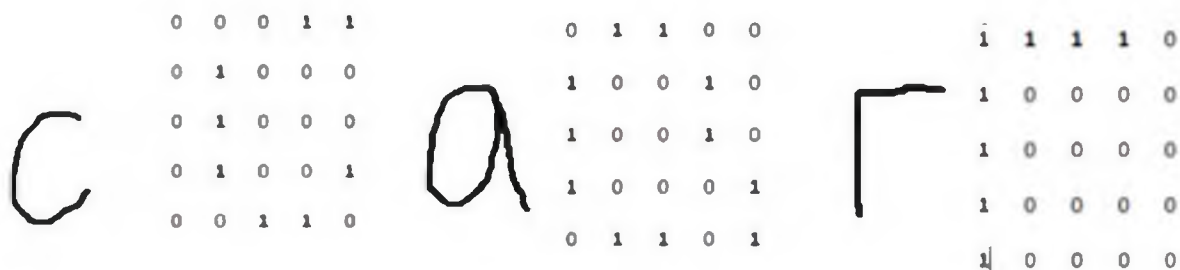


Рисунок 6. Сравнение матрицы с эталонами в базе образов

Результат распознавания в конце к слову

С П А С И Б О

Рисунок 7. Результат распознавания

Вывод: Приведенный модельный пример показал работоспособность и достоверность предложенного подхода распознавания зрительных образов.

Список литературы

1. Абраменко А. Принципы распознавания / А. Абраменко – К.: Компьютер-пресс, 1997 – 123 с.
2. Батырканов Ж. И. Системы искусственного интеллекта. – Б. : ИЦ «Техник», 2013. – 138 с.
3. Батырканов Ж. И. Кудаеева Г. М. Подход распознавания зрительных образов на основе эталонов и обучения. – Б. : ИЦ «Техник», Известия КГТУ №1 (34), 2015. – С. 11-13.
4. Батырканов Ж. И., Кудаеева Г. М., Субанкулова Ж. Ж. Подход к распознаванию зрительных образов на основе эталонных образов и обучения [Электронный ресурс] // Огарев-online. – 2017. – №15. – Режим доступа: <http://journal.mrsu.ru/arts/podxod-k-raspoznaniyu-zritelnyh-obrazov-na-osnove-etalonnyh-obrazov-i-obucheniya>
5. Гаврилов Г.П. Логический подход к искусственному интеллекту / Г.П. Гаврилов – М.: Мир, 1998 – 256 с.
6. Журавлев Ю. И. Избранные научные труды. – Изд. Магистр, 2002. – 420с.
7. Кучуганов А.В., Лапинская Г.В. Распознавание рукописных текстов / А.В. Кучуганов, Г.В. Лапинская – Ижевск.:Мир, 2006 – 514 с.
8. Саймон Хайкин. Нейронные сети: полный курс Neural Networks: A Comprehensive Foundation. – 2-е изд. – М. : «Вильямс», 2006. – 1104 с.
9. Шамис А.Л. Принципы интеллектуализации автоматического распознавания / А.Л. Шамис – К.:2000 – 312 с.
10. Хант Э. Искусственный интеллект. – М. : Мир, 1978. – 558 с.