

БАТЫРКАНОВ Ж. И., КУДАКЕЕВА Г. М., СУБАНКУЛОВА Ж. Ж.
ПОДХОД К РАСПОЗНАВАНИЮ ЗРИТЕЛЬНЫХ ОБРАЗОВ
НА ОСНОВЕ ЭТАЛОННЫХ ОБРАЗОВ И ОБУЧЕНИЯ

Аннотация. Рассматривается и решается задача распознавания зрительных образов. При этом под образом понимается конкретное проявление распознаваемого объекта. Предлагается подход к распознаванию зрительных образов на основе использования базы эталонных образов и процедуры обучения. Предлагаемый подход можно применять для распознавания рукописного текста, идентификации личности по биометрическим данным и решения других подобных задач.

Ключевые слова: образ, объект распознавания, цифровой образ, эталонный образ, считывающее устройство.

BATYRKANOV ZH. I., KUDAKEEVA G. M., SUBANKULOVA ZH. ZH.
RECOGNITION OF VISUAL IMAGES:
AN APPROACH BASED ON REFERENCE IMAGES AND TRAINING

Abstract. The problem of recognition of visual images is considered and solved. The image is thought of as the concrete manifestation of the recognized object. The authors suggest an approach to recognizing visual images based on the use of reference image base and training procedure. This approach can be used to recognize handwritten text, to identify a person by biometric data, etc.

Keywords: image, object of recognition, digital image, reference image, reader.

Введение. В практике сегодняшнего дня достаточно часто возникает потребность в разработке методов, алгоритмов и программ для распознавания зрительных образов. Они применяются для распознавания рукописного текста, природных и техногенных катастрофических явлений по аэрофотоснимкам, при разведке полезных ископаемых, для идентификации личностей по биометрическим данным и пр.

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 подход к решению задач распознавания рукописного текста и идентификации личности по биометрическим данным.

Подход на основе базы эталонных образов и обучения. Подход состоит в последовательном выполнении двух этапов:

- 1) зрительный образ предъявляется считывающему устройству;
- 2) считывающее устройство переводит зрительный образ в двоичный компьютерный образ.

В качестве считывающего устройства может служить сканирующее устройство, цифровой фотоаппарат, смартфон, рецепторная матрица и другие специальные устройства. Алгоритм распознавания заключается в последовательном сравнении поступившего двоичного образа со всеми двоичными образами из базы эталонных образов. При сравнении вычисляются количественные оценки близости, например, по евклидовой разности сравниваемых двоичных образов. Тот эталонный двоичный образ, для которого вычисляемая величина евклидовой разности минимальна, определит класс объекта, к которому принадлежит предъявленный для распознавания образ.

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 (существующей базе данных). Алгоритм распознавания личности по биометрическим данным осуществляется по аналогичной схеме.

Процедура обучения при применении данного подхода к распознаванию состоит в корректировке алгоритма распознавания на этапе разработки и обучения системы. Рассмотрим как это делается на примере распознавания зрительных образов.

Процедура подготовки и обучения системы. В самом начале последовательно на рецепторную матрицу считывающего устройства подается определенное количество образов первого объекта A . 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

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

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

Далее предъявляется образ A^2 объекта A и так до последнего возможного образа A^L объекта A . Совокупность образов A^1, A^2, \dots, A^L объекта A записывается в память компьютера как база данных эталонных образов объекта A . Эта 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. отнесения предъявленного образа к тому или иному объекту происходит по следующей схеме. При предъявлении рецепторной матрице образа X в компьютере образуется числовой образ в виде матрицы:

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

Чтобы отнести образ X к тому или иному классу объектов А, В... С, осуществляется процедура последовательного вычисления нормы от 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 А, В, ... С. Норму вычислять удобно в виде евклидовой 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 В выходного слоя. Это говорит о том, что предъявленный образ X относится к объекту В.

Описанный алгоритм классификации (расознавания) работает не только в случаях, когда матрица X полностью совпадает с матрицами A^1, \dots, C^M , но и в тех случаях, когда X не совпадает ни с одной из них. При этом используется «принцип близости», который в теории распознавания заключается в том, что образы для одного и того же объекта в пространстве признаков должны располагаться близко друг к другу.

2-й вариант алгоритма распознавания.

При использовании этого алгоритма все матрицы $A^1, A^2, \dots, A^L, B^1, \dots, C^1, \dots, C^M, X$ переводятся в векторную форму:

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

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 векторы 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.

Таким образом, нами предложен подход распознавания зрительных образов с процедурой обучения и использования базы эталонных образов. Для более точной и достоверной работы распознающего устройства необходимо увеличивать количество ячеек рецепторной матрицы, а также базу эталонных образов распознаваемых объектов.

ЛИТЕРАТУРА

1. Хант Э. Искусственный интеллект. – М.: Мир, 1978. – 558 с.
2. Хайкин С. Нейронные сети: полный курс Neural Networks: A Comprehensive Foundation. – 2-е изд. – М.: Вильямс, 2006. – 1104 с.
3. Журавлев Ю. И. Избранные научные труды. – М.: Магистр, 2002. – 420 с.
4. Батырканов Ж. И. Системы искусственного интеллекта. – Бишкек: ИЦ «Техник», 2013. – 138 с.
5. Батырканов Ж. И., Кудакеева Г. М. Подход к распознаванию зрительных образов на основе эталонов и обучения // Известия Кыргызск. гос. техн. ун-та им. И. Раззакова. – 2015. – №1 (34). – С. 11–13.