International Science Group

ISG-KONF.COM

I

INTERNATIONAL SCIENTIFIC AND PRACTICAL CONFERENCE "PROBLEMS OF MODERN SCIENCE AND PRACTICE"

Boston, USA

September 21 – 24, 2021

ISBN 978-1-63972-061-3

DOI 10.46299/ISG.2021.II.I

PROBLEMS OF MODERN SCIENCE AND PRACTICE

Abstracts of I International Scientific and Practical Conference

Boston, USA
September 21 – 24, 2021

Library of Congress Cataloging-in-Publication Data

UDC 01.1
The I International Science Conference «Problems of modern science and practice», September 21 – 24, 2021, Boston, USA. 487 p.

ISBN - 978-1-63972-061-3
DOI - 10.46299/ISG.2021.II.I

EDITORIAL BOARD

ON THE FEATURES OF METHODS OF PROCESSING AND RECOGNITION OF HANDWRITTEN TEXT

Tvoroshenko Iryna,
Ph.D., Associate Professor
Kharkiv National University of Radio Electronics

Bielinskyi Yaroslav,
Master in informatics
Kharkiv National University of Radio Electronics

Today, handwritten text remains a means of communication and information gathering in our daily life, despite the rapid development of information technology [1-5]. Many archival documents are still stored in a paper format inaccessible to all. Techniques for converting images into digital text have evolved at an extremely fast pace lately. An example of this is printed text recognition systems, which have become quite effective today [6, 7]. However, the available technologies are still at a limited level to recognize text images of various styles and languages of handwriting.

Handwriting recognition usually consists of the following steps:

– Image pre-processing [8];
– Segmentation and normalization [9];
– Construction of signs [10];
– Classification [11];
– Processing of results [12].

At the stage of image pre-processing, such methods as filtering, noise absorption and others are used, their purpose is to improve the image quality [13]. To date, the task of handwriting recognition is insufficiently solved. Improving the quality of the image when recognizing handwritten text is extremely important, because the better the image quality, the more convenient it will be to work with it in the next stages of processing.

Improving image quality includes converting the image to grayscale, removing image defects, and separating text from the background [14].

To simplify further work with the image, the image is converted to grayscale. A brightness value is calculated for each individual pixel, which is measured in the range from 0 to 255. Black corresponds to 0 brightness level, and white – 255 brightness level.

$$I = 0.299\,R + 0.587\,G + 0.114\,B, \quad (1)$$

where R, G, B are the values of the red, green and blue channels, respectively.

Defects are removed by standard image processing methods. The most commonly used to remove noise is a Gaussian filter to suppress high-frequency noise and a median filter to remove noise “salt and pepper”. A promising filter is a filter based on wavelet transform. You can use the Gaussian function to construct a convolution matrix that

helps you calculate the weighted average of neighboring pixels for each pixel in the image [15].

Thus, by using a Gaussian filter, the noise will be suppressed, because all the noisy pixels (the brightness of which is very different from the brightness of neighboring pixels) will take an average value, resulting in the contours of objects will be emphasized, which has a positive effect on image recognition (handwritten text) on digital images.

The median filter is based on the concept of median. The most common way to implement this filter is to organize the pixel brightness values using an odd-radius window and then replace the pixel brightness values with the median of the set.

Since the image can be represented as a discrete signal, it is possible to use filters based on frequency division in the discrete region. Wavelet analysis is a very promising way to analyze data. The signal can be represented as follows:

$$s(t) = f(t) + \sigma e(t), \tag{2}$$

where $f(t)$ – useful signal;

$e(t)$ – noise;

σ – noise level;

$s(t)$ – the signal under research.

Accordingly, the wavelet transform allows you to remove noise in 4 steps:

– Decomposition of the signal on the basis of wavelets;

– Selection of the noise threshold for each of the decomposition levels;

– Threshold filtering of detail coefficients;

– Restoration of the signal.

This method of filtering works best on smooth signals, i.e. on signals in the decomposition of which only a small number of detail factors differ significantly from zero. The selection of the wavelet and the depth of decomposition depend on the properties of the object being filtered by the signal. Criteria that minimize the quadratic loss function for the selected noise model are used to select the noise threshold. These filters are used less often than the median, because wavelets lead to additional parameterization of the program and slowdown, because you need to calculate additional data sets.

Separating text from the background is a special case of selecting an object in an image. The problem is to get a binary image *B* from the image of text *A*, such that

$$B(i, j) = \begin{cases} 1, P(i, j) \in T_A, \\ 0, P(i, j) \notin T_A, \end{cases} \tag{3}$$

where $P(i, j)$ – pixel (i, j);

T_A – text on the image *A*.

This transformation allows you to further use the analysis of connecting components, contours, skeletons. The most commonly used method of separating text from the background – threshold binarization.

Let the image be given, $I(i, j)$ – pixel brightness with coordinates (i, j). Threshold image binarization is called pixel-by-pixel conversion:

$$f(i, j) = \begin{cases} 1, I(i, j) \geq d, \\ 0, I(i, j) < d, \end{cases} \quad (4)$$

where d is called the binarization threshold.

Usually, the histogram of the brightness of the text image shows two peaks: a high peak in the area of light pixels, which corresponds to the background, and a lower peak in the area of dark pixels, which corresponds to the text. Thus, the task of finding the threshold value of brightness, i.e. such that pixels with a brightness above this value (background) will be considered black, and below (text) – white (such "inversion" of color is done to simplify the application of many methods in the future), is the task of finding the optimal value between the two peaks of the histogram. To solve this problem, there is the method of Otsu and its variations.

In the Otsu method, the brightness range [0; *L*] the image is divided into two parts by the limit value *T*. The essence of the algorithm is to minimize the intraclass variance, which is defined as the weighted sum of the variances of the two classes. In the Otsu method, the minimization of intraclass variance is equivalent to the maximization of interclass variance, which is calculated as follows:

$$\sigma_b^2 = \omega_1 \omega_2 (\mu_1 - \mu_2)^2, \quad (5)$$

where σ_b – interclass variance;

ω_1 and ω_2 – probabilities of the first and second classes;

μ_1 and μ_2 – the arithmetic mean of each of the classes.

The disadvantage of this method is the sensitivity to uneven light. To solve this problem, usually obtain a lighting component by low-pass filtering *G* image using a Gaussian filter.

Another disadvantage of the Otsu method is the adhesion of adjacent areas, which can affect the further processing and recognition of the image. Therefore, there is a method of adaptive binarization, which also allows you to solve the problem of light difference.

As mentioned earlier, handwriting recognition is an important issue. At the moment, small recognition accuracy is obtained. Better performance can only be achieved using contextual and grammatical information. For example, in the process of recognition, it is easier to look for whole words in a dictionary than to try to analyze individual characters from the text. Knowledge of the grammar of the language can help determine whether a word is a verb or a noun. The forms of individual handwritten

characters may sometimes not contain enough information to accurately recognize the entire handwritten text.

Methods of automatic image recognition and their implementation in optical reading systems (OCR-systems) one of the most advanced technologies of artificial intelligence.

OCR refers to the automatic recognition of images of printed or handwritten text characters using special programs, such as input to a computer using a scanner, and converting it to a format suitable for processing by word processors, word processors, etc. OCR is sometimes understood as a device for optical character recognition or automatic text reading. Currently, such devices in industrial use process up to 100 000 documents per day. The introduction of good and medium quality documents is envisaged – for example, census forms, tax returns, statistical accounting. OCR handwriting recognition technology is used for large databases, with the level of transcription accuracy increasing every day, and it is already close to ideal (100%).

Recognition systems are implemented in the form of classifiers that use different methods: template (raster), feature, structural.

When using template methods of character recognition, the scanned image is translated into raster (current), and then it is compared with the reference templates that were generated in the database. The criterion for choosing a template is the smallest number of points other than the studied image. The template for each class is obtained by averaging the image of the symbols of the educational sample. The advantage of template methods is the high accuracy of defective character recognition. The main disadvantage is the dependence on the font that occurs in the image. The font must be known in advance; otherwise it will be impossible to correctly recognize the studied symbol.

Sign methods are the most common. The analysis in them is based on the fact that the image is matched to the *N*-dimensional vector of features. Recognition is to compare it with a set of reference vectors of the same dimension. The quality of recognition depends on the types of features and their quantities. The formation of the vector is carried out during the analysis of the prepared image; the standard for each class is obtained in a similar way of processing symbols from the training sample. The main advantages of this group are ease of implementation, good generalization, resistance to changes in character shape and speed. Among the disadvantages of these methods is instability to image defects, loss of information about the symbol at the stage of obtaining signs, the lack of clearly defined rules for the formation of signs.

Structural methods of recognition use information not about the spelling of the symbol, but about its topology – the relative position of the individual components of the symbol. The advantages of these methods include independence in font types and sizes.

The main problem of topological methods is the identification of signs that have defects (for example, line rupture or merging of adjacent lines), as well as low speed.

Along with other methods, methods that do not require prior segmentation, such as hierarchical hidden Markov models and convolutional neural networks have become relevant. With the advent of a new wave of popularity of neural network classifiers, they have become more commonly used in text recognition research. The main

advantage of using neural network technologies is good generalization ability, the ability to use contextual analysis and recognize symbols based on the surrounding symbols.

The simplest recognition algorithm is segmentation, recognition of each part. However, this method does not work effectively. It does not correct segmentation errors and is only adequate in recognizing a random sequence of independent well-separated objects. Therefore, to solve this problem, various stochastic grammars and the corresponding probabilistic methods are used, among which we can specify the following:

– Hidden Markov model;

– Graph recognition.

Accordingly, today there is a great demand for digital storage of textual information, which is placed in printed, graphic or handwritten documents, for further processing, editing and analysis [16-19]. The development of digitization and scanning tools has led to the active use of methods for detecting and recognizing objects in images. This is due to the development of optical character recognition systems that allow you to automatically analyze printed or handwritten documents and prepare text data in editable formats for processing.

Thus, promising areas of further research are the study of applied capabilities of methods in the recognition of handwritten documents.

References:

1. Daradkeh Y.I., Tvoroshenko I., Gorokhovatskyi V., Latiff L.A., and Ahmad N. (2021) Development of Effective Methods for Structural Image Recognition Using the Principles of Data Granulation and Apparatus of Fuzzy Logic, *IEEE Access*, 9, pp. 13417-13428.

2. Matarneh Rami, Tvoroshenko Irina, and Lyashenko Vyacheslav (2019) Improving Fuzzy Network Models For the Analysis of Dynamic Interacting Processes in the State Space, *International Journal of Recent Technology and Engineering*, 8(4), pp. 1687-1693.

3. Кучеренко Є.І., Творошенко І.С. (2011) Оперативне оцінювання простору станів складних розподілених об’єктів з використанням нечіткої інтервальної логіки. *Штучний інтелект*, 3, С. 382-387.

4. Tvoroshenko I.S., and Kramarenko O.O. (2019) Software determination of the optimal route by geoinformation technologies, *Radio Electronics Computer Science Control*, 3, pp. 131-142.

5. Gorokhovatskyi V., Rusakova N., and Tvoroshenko I. (2020) The application of image analysis methods and predicate logic in applied problems of magnetic monitoring, *Telecommunications and Radio Engineering*, 79(20), pp. 1801-1811.

6. Daradkeh Y.I., Gorokhovatskyi V., Tvoroshenko I., Gadetska S., and Al-Dhaifallah M. (2021) Methods of Classification of Images on the Basis of the Values of Statistical Distributions for the Composition of Structural Description Components, *IEEE Access*, 9, pp. 92964-92973.

7. Sonka M., Hlavac V., and Boyle R. (2014) *Image Processing, Analysis, and Machine Vision*, Atlanta, USA: Thomson-Engineering, 920 p.

8. Tvoroshenko I.S., and Gorokhovatsky V.O. (2019) Modification of the branch and bound method to determine the extremes of membership functions in fuzzy intelligent systems, *Telecommunications and Radio Engineering*, 78(20), pp. 1857-1868.

9. Tvoroshenko I.S., and Gorokhovatsky V.O. (2020) Effective tuning of membership function parameters in fuzzy systems based on multi-valued interval logic, *Telecommunications and Radio Engineering*, 79(2), pp. 149-163.

10. Flah P. (2015) *Machine learning. The science and art of building algorithms that extract knowledge from data*, Moscow, Russia: DMK Press, 400 p., (in Russian).

11. Gorokhovatskyi V.O., Tvoroshenko I.S., and Vlasenko N.V. (2020) Using fuzzy clustering in structural methods of image classification, *Telecommunications and Radio Engineering*, 79(9), pp. 781-791.

12. Kobylin O., Gorokhovatskyi V., Tvoroshenko I., and Peredrii O. (2020) The application of non-parametric statistics methods in image classifiers based on structural description components, *Telecommunications and Radio Engineering*, 79(10), pp. 855-863.

13. Sharma G., and Schiele B. (2015) Scalable Nonlinear Embeddings for Semantic Category-based Image Retrieval, *Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 7-13.

14. Nong Ye. (2013) *Data Mining: Theories, Algorithms, and Examples*, Florida, USA: CRC Press, 349 p.

15. Peters J.F. (2017) *Foundations of computer vision: Computational Geometry, Visual Image Structures and Object Shape Detection*, Cham, Switzerland: Springer International Publisher, 417 p.

16. Yousef Ibrahim Daradkeh, and Iryna Tvoroshenko (2020) Application of an Improved Formal Model of the Hybrid Development of Ontologies in Complex Information Systems, *Applied Sciences*, 10(19). p. 6777.

17. Lyashenko V., Mustafa S.K., Tvoroshenko I., and Ahmad M.A. (2020) Methods of Using Fuzzy Interval Logic During Processing of Space States of Complex Biophysical Objects, *International Journal of Emerging Trends in Engineering Research*, 8(2), pp. 372-377.

18. Кучеренко Е.И., Корниловский А.В., Творошенко И.С. (2010) О методах настройки функций принадлежности в нечетких системах. *Системы управления, навигации и связи*, 1(13), С. 94-98.

19. Кучеренко Е.И., Творошенко И.С. (2003) Процессы принятия решений в сложных системах на основе нечетких интервальных представлений. *Вісник Національного технічного університету «ХПІ». Тематичний випуск: Системний аналіз, управління та інформаційні технології.* Х.: НТУ «ХПІ», 1(7), С. 79-86.