THE IMPACT OF FILTERS ON THE QUALITY OF BINARIZATION FOR HANDWRITING IMAGES

Marcin Adamski, Krzysztof Pryzmont

Faculty of Computer Science, Bialystok University of Technology, Białystok, Poland

Abstract: Filtration and binarization techniques are often used in handwriting recognition systems. These operations are performed as part of a stage called preprocessing, the result of which is passed to feature extraction and classification processes. Operations performed as part of the preprocessing are important because their result affects the outcome of the entire system. This paper focuses on the assessment of filtration techniques influence on the binarization of handwriting images. In the experiments, four filtration methods were tested with seven thresholding algorithms for various combinations of filtration and binarization parameters. The experiments were conducted on handwriting images selected from Document Binarization Competitions (DIBCO) datasets with ground truth images for the assessment of binarization correctness. The final evaluation was conducted based on the average of quality measures: F-measure, Accuracy, Relative Foreground Area Error and Region Non-uniformity.

Keywords: image filtering, binarization, handwriting

1. Introduction

Automatic text recognition from images of scanned documents is an important research area that has led to many successful applications. Whilst there are many available solutions for typed documents such as OCR tools (Optical Character Recognition), analysis of handwritten texts may still poses a challenge, especially when handwritten symbols are not well separated but form a continuous curve. Another source of difficulty when dealing with handwritten documents is the quality of input data. In many cases one has to work with low quality images that may be acquired from stained and damaged documents. Artifacts may be also introduced during acquisition or storage due to low scanning DPI and lossy compression.

Three important stages can be distinguished in handwriting recognition systems: preprocessing, feature extraction and classification (Fig. 1). The purpose of the initial

Advances in Computer Science Research, vol. 14, pp. 17-30, 2018. DOI: 10.24427/acsr-2018-vol14-0002

preprocessing is to facilitate the feature extraction by separation of text from background information that may be discarded. At this stage, one of the most commonly used techniques is image binarization. There are many methods for binarizing images, and new algorithms are still being developed. An example of this activity can be seen in Document Image Binarization Competition (DIBCO) [10,11] that is organized since 2009. The input to the binarization process can be a raw image that

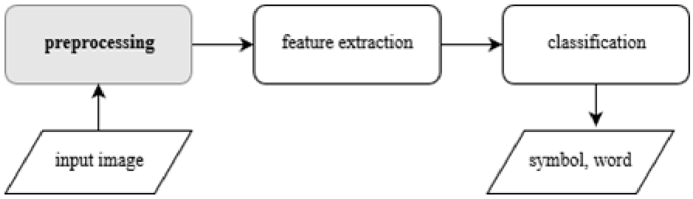

Fig. 1. Stages of text recognition system

was acquired by the scanning process. However, initial filtering may be performed to reduce noise and enhance image quality. Examples of techniques that may be used for this task include: median filtering, average and Gaussian blur [3], anisotropic processing such as Perona-Malik diffusion [9].

Whilst there are many works in the literature on binarization algorithms, the effects of filtration methods executed prior binarization process on its result have not been widely studied. Some results related to this subject can be found in [4] and [15]. The authors of [4] investigated the effect of preprocessing and postprocesing on binarization. The preprocessing was performed using nose reduction technique based on wavelet transform applied to images of typed script. However, according to [4], this approach did not led to satisfactory results. In work [15] selection of filtering techniques were investigated. The results showed that none of the applied methods was best in all cases. The experiments were carried out on a mixed set of images that contained both handwritten and typed texts. In our investigation we focused on handwritten texts due to their distinct characteristics and studied filtering techniques such as Gaussian, Kuwahara and Perona-Malik that were not included in [15].

The aim of this investigation was to assess the impact of selected filtering techniques executed before binarization process on its result. The research was carried out on exemplary handwriting images from DIBCO competition that contain various types of artifacts. To assess the results the ground truth images were used together with various quality measures.

The article begins with a short review of selected filtering and binarization techniques. The next section contains description of the methodology used during experiments and is followed by presentation of the results and conclusions.

2. Image filtering

The image obtained from scanner or camera devices may contain noise negatively affecting subsequent stages of the text recognition. There are many available techniques that can be used for noise reduction. During selection of methods one should take into account the type of noise present in the image, as well as the characteristics of objects which should be preserved after filtration. The main problem is that, in addition to noise reduction, the filtering operation may remove information that is important at later stages of image processing. In case of images containing text this may include blurring text edges, merging separated structures or removing parts of script. In this work we selected widely used methods based on image filtering: median, Gaussian, Kuwahara and Perona-Malik filters. The selection was based on author's experience from earlier investigations related to preprocessing and verification of handwritten signatures [1] and results of other related works ([4,15]). The brief description of selected filters is given below.

Median filter [3] is a simple operation which for each pixel computes new value that is a median of pixels intensities covered by the median filter mask. The mask is usually a square region centered at the input pixel that is currently processed. This type of filter can be applied to remove "salt and pepper" type of noise. This operation usually preserves sharp edges but may also remove thin structures when the size of the mask exceeds their width. An example output of median filter can be seen in Fig. 3b.

Gaussian filter [3] is performed by convolving input image with filter mask that approximates 2D Gaussian function (1). This results in smoothing input image which may reduce image noise but also blurs object edges which is disadvantageous. The extent of smoothing is controlled by the sigma parameter and the mask size. Example output in presented in Fig. 3c.

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{1}$$

Kuwahara filter [8] is a nonlinear smoothing filter that allows to preserve object edges. In order to calculate the output value for selected input pixel, the Kuwahara filter computes mean and standard deviation in four overlapping regions (Fig. 2) in the neighborhood of selected pixel. The output value of the filter is the mean value from the region with smallest standard deviation (2). Example is given in Fig. 3d.

$$I_{out}(x,y) = \mu_k, \;\; k = \arg min_j \sigma_j(x,y) \tag{2}$$

where μ_i and $\sigma_i(x,y)$ are the mean and standard deviation of pixel intensities in i-th region.

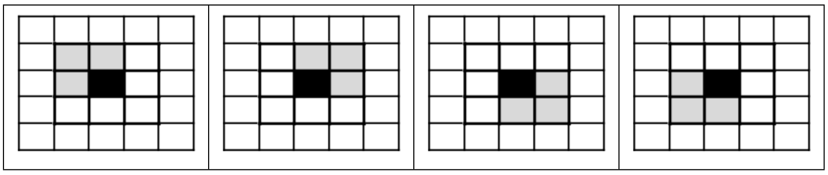

Fig. 2. Pixel's neighborhood regions analyzed in 3x3 Kuwahara filter.

Perona-Malik diffusion [9] is anisotropic smoothing filter that suppresses its effect near boundaries, therefore preserves sharp edges. The diffusion operation is based on modified heat equation (3) where I denotes image, t is time controlled by iteration process, c is a function that controls rate of diffusion at particular point at position (x,y).

$$\frac{\delta I}{\delta t} = c(x,y,t)\Delta I + \nabla c \nabla I \tag{3}$$

In order to reduce image smoothing near object boundaries the rate of diffusion may be constrained based on image gradient. In this work diffusion rate was controlled using function given in (4). The value of K parameter scales the gradient strength.

$$c(\|\nabla I\|) = e^{-(\|\nabla I\|/K)^2} \tag{4}$$

3. Image binarization

Image binarization is a basic segmentation technique used in many areas of image processing. The input to the binarization method is a grayscale image and the output is a binary map, where one of the values represents the background and the other is used for the segmented object. In this study the object (handwriting) is assigned

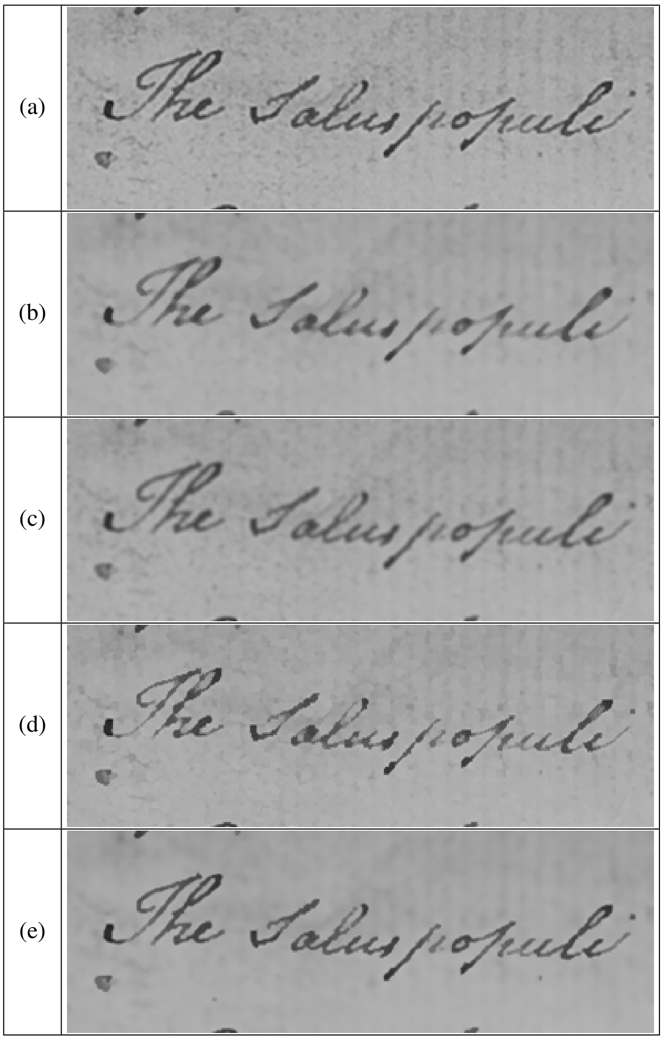

Fig. 3. Result of filtering methods applied to sample image from DIBCO dataset (a): median filter (b), Gaussian filter (c), Kuwahara filter (d), Perona-Malik diffusion (e)

a value of zero (shown as black) and the background pixels have the value of one (shown as white). In general, the binarization can be described as a thresholding operation that assigns a value of 1 or 0 to each input pixel by comparing it to the threshold value (5).

$$I_{out}(x,y) = \begin{cases} 1, & if\, I(x,y) > T \\ 0, & otherwise \end{cases} \tag{5}$$

The threshold value can be global – the same value is used for the whole image, or it may vary depending on location. This leads to two main categories of binarization algorithms: global and local methods. Many approaches to binarization have been proposed in the image processing literature. Comprehensive review that describes over 40 methods can be found in [14]. Example of image binarization is presented in Fig. 4. In this work we used seven approaches, three of which were global (Kapur-Sahoo-Wong, Otsu, Ridley-Calvard) and four local (Bradley, Niblack, Sauvola, White-Rohrer). They are briefly described in this section.

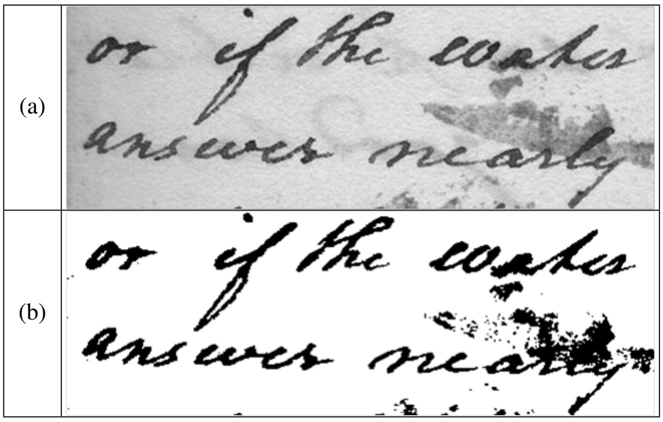

Fig. 4. Grayscale image (a) and its binarized version using Otsu metod (b)

Kapur-Sahoo-Wong method [5] is global technique that computes threshold value maximizing the sum of objects (foreground) and background distribution entropies. The distribution of foreground p_f and background p_b are computed based on normalized histogram values $p(i)$ as given by equation (6).

$$p_f = \sum_{i=0}^{t} p(i), \;\; p_b = \sum_{i=t+1}^{255} p(i) \tag{6}$$

Entropies of foreground and background are computed using equations (7) and (8), respectively.

$$H_f = -\sum_{i=0}^{t} \frac{p(i)}{p_f} \ln \frac{p(i)}{p_f}) \tag{7}$$

$$H_b = -\sum_{i=t+1}^{255} \frac{p(i)}{p_b} \ln \frac{p(i)}{p_b} \tag{8}$$

The global threshold is selected as the value that maximizes sum of entropies (9).

$$T_{opt} = \max_{t=0,\,\ldots,\,255} H_f(t) + H_b(t) \tag{9}$$

Otsu method [7] is one of the most widely used binarization techniques. The threshold value is computed globally for the whole input image based on optimization procedure that maximizes intra-class variance between two classes of pixels defined by threshold criterion (1). The intra-class variance σ_B^2 is computed using equation (10).

$$\sigma_B^2 = P_o P_1 (\mu_0 - \mu_1)^2 \tag{10}$$

where μ_i is the mean value of input pixel values that are classified as foreground (i=0) and as background (i=1), and P_i is the sum of normalized histogram values for the levels that belong to a particular class. The value of global threshold is given by (11).

$$T_{opt} = \arg max_{0<t<L-1} \left(\sigma_B^2(t) \right) \tag{11}$$

Ridley-Calvard method [12] represent iterative approach to computing global threshold value. At each iteration i the threshold value calculated as the average of mean intensities of foreground and background pixels (12). Initial means can be computed as the mean of corner pixels (background) and the rest of the image (foreground). After computing new threshold, the mean values are updated based on pixel classification obtained using new threshold. The process continues until the threshold value does not change.

$$T(i) = \frac{\mu_b(i-1) + \mu_f(i-1)}{2} \quad i = 1,\, 2,\, \ldots \tag{12}$$

In **Bradley method** [2] the threshold value is computed locally based on the formula (13):

$$T(x,y) = \mu(x,y) \left(\frac{100-t}{100} \right) \tag{13}$$

The mean value $\mu(x,y)$ is computed using a window of fixed size. The parameter t is selected heuristically. To compute the mean value, this method uses integration image which allows to perform computations in linear time.

Niblack method [6] is a simple local binarization technique where the threshold value is computed for each pixel based on mean and standard deviation values in a window centered at pixel's location (14).

$$t(x,y) = \mu(x,y) + k * \sigma(x,y), \;\; k < 0 \tag{14}$$

where (x,y) is pixel position; $\mu(x,y)$, $\sigma(x,y)$ are mean and standard deviation computed in local neighborhood window centered at (x,y), k is user defined parameter.

Sauvola algoritm [13] is modification of Niblack method where additional scaling factors are included to reduce sensitivity to background noise (15).

$$T(x,y) = \mu(x,y)\left[1 + k\left(\frac{\sigma(x,y)}{R} - 1\right)\right], \; k > 0 \tag{15}$$

where R is dynamic range of standard deviation and k is user defined parameter. The size of the window and parameter k are selected heuristically by the analysis of results obtained for particular class of images. The parameter R can be computed separately for each image as a maximum value of standard deviation $\sigma(x,y)$.

White-Rohrer method [16] computes local threshold value using formula (16). The method was initially used for text images based on assumption that intensity of pixels comprising text differ significantly from pixel values of neighboring background.

$$I_{out}(x,y) = \begin{cases} 1, & I(x,y) \geq \frac{\mu(x,y)}{k}, \; k > 1 \\ 0, & otherwise \end{cases} \tag{16}$$

4. Evaluation procedure

To assess the impact of filtration methods on the results of binarization algorithms we conducted several experiments using images from the Document Image Binarization Competition (DIBCO) databases. Each sample image from DIBICO dataset has its ground truth version – a binary image with “ideal” binarization. By comparing ground truth image with the result of particular method its performance can be assessed. In this study we used eight images from DIBCO datasets (Fig. 5). Selected samples represent different types of distortions that may occur in scanned images of handwritten documents.

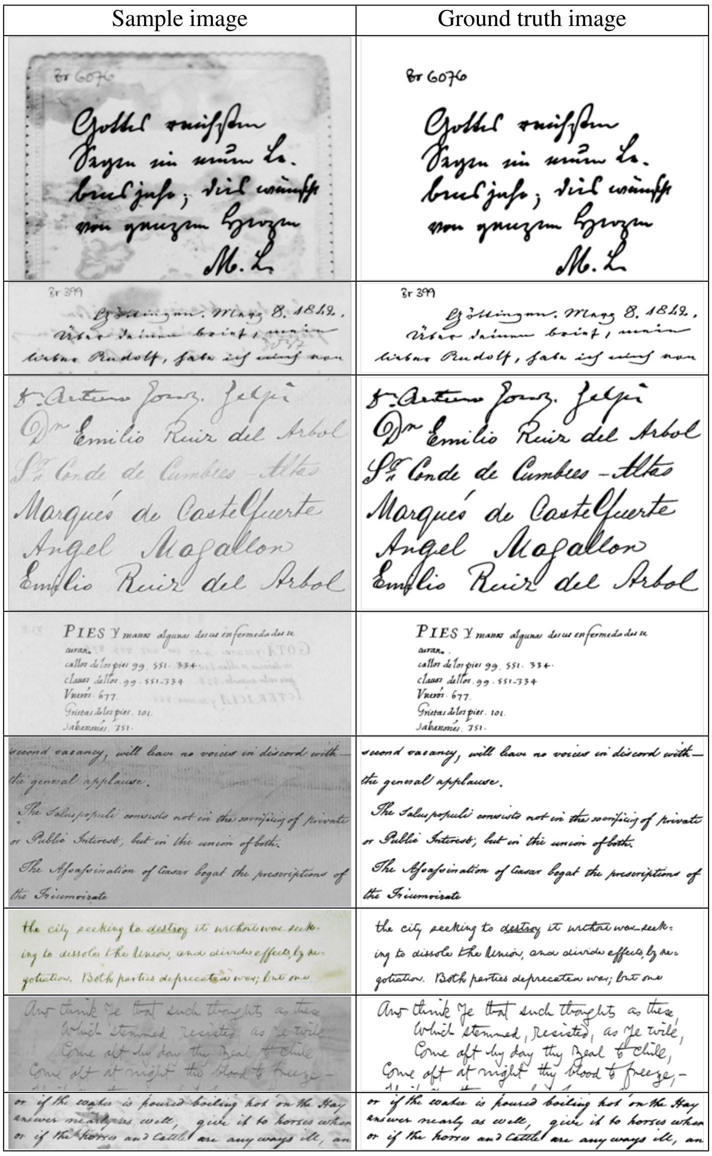

Fig. 5. Sample images and their ground truth versions

4.1 Measures

There are several measures that may be used to compare ground truth image with the output of binarization method. In our investigation we used evaluation functions proposed in works [11,14].

F-measure (17) is harmonic mean of recall and precision measures (18) . Its value is in range <0,1> where 1 represents perfect binarization.

$$FM = \frac{2 * Recall * Precision}{Recall + Precision} \tag{17}$$

$$Recall = \frac{TP}{TP+FN}, \;\; Precision = \frac{TP}{TP+FP} \tag{18}$$

where TP, FP, FN denote True Positive, False Positive and False Negative.

Accuracy is a ratio of correctly binarized pixels to total number of pixels in the image (19). Its value is in range <0,1> , 1 represents perfect binarization.

$$AC = \frac{TP+TN}{TP+TN+FN+FP} \tag{19}$$

where TN is True Negative.

Relative foreground area error represents relative accuracy of foreground pixels classification using notion of area (20). Its value is in range <0,1>, 0 represents perfect binarization.

$$RAE = \begin{cases} \frac{A_O - A_T}{A_O}, A_T < A_O \\ \frac{A_T - A_O}{A_T}, A_T \geq A_O \end{cases} \tag{20}$$

where A_O is the area of objects in ground truth image, A_T is the area of objects in binarized (test) image.

Region nonuniformity is computed using only grayscale image and its binarized version (21), without using ground truth reference. Well segmented image should have value of NU close to 0, where the value close to 1 means that foreground and background are hardly distinguishable.

$$NU = \frac{|F_T|}{|F_T + B_T|} \frac{\sigma_f^2}{\sigma^2} \tag{21}$$

where σ^2 and σ_f^2 represent standard deviations of pixel intensities for the whole image and foreground, respectively, F_T is the number of foreground pixels, B_T is the number of background pixels in binarized image.

In order to summarize all measures using one statistic we calculated arithmetic average. Due to different semantics of measure limits, RAE and NU values had to be reflected in final expression (22).

$$AV = \frac{FM + AC + (1 - RAE) + (1 - NU)}{4} \tag{22}$$

4.2 Method parametrization

The results of selected filtering and binarization methods depend on their parametrization. For certain methods one may find recommended values in the literature, however, the parametrization is usually chosen heuristically and may be suboptimal if the same setting is applied to images with different characteristics. To take this into account we conducted experiments with several possible configurations. For each image various combinations of filters and binarization methods parametrizations were verified. For final evaluation we selected only the best setting per image obtained for every filter and binarization algorithm pair.

The following settings were investigated during experiments:

- median, Gauss and Kuwahara filter size : 3x3, 5x5,
- number of iterations Perona-Malik filter: 5, 10,
- window size for local binarization techniques: 9, 15, 25, 45
- Bradley binarization t parameter: 7, 10, 15, 20
- White'a-Rohrer binarization k parameter: 1.2, 1.5, 2.0, 2.3
- Niblack binarization k parameter: -0.25, -0.5, -1.0, -1.5
- Sauvola binarization k parameter: 0.15, 0.3, 0.5, 0.7

As a result we verified 603 configurations per image giving the total number of evaluations 4824.

5. Results

In accordance with the adopted assumptions seven binarization algorithms were tested in combination with four filtration techniques. For comparison the images were also binarized without prior filtration. The experiments were carried out on 8 images from DIBICO datasets. To evaluate the effectiveness of analyzed methods their output was compared to ground truth images using the AV measure based on 4 separate coefficients. Table 1 shows obtained results. Each row of the table represents particular binarization algorithm, whilst separate filtration techniques are collected into columns. Each value was computed as an arithmetic average of results obtained for all images, where for each image only the best result among all parameterizations was used. The last row shows mean values for given filtration technique. The highest (best) values obtained for each binarization are given in bold font whilst the lowest (worst) results are grayed out. As can be seen from the Tab. 1, the worst results were obtained for the Kuwahara filtration method. This technique preserved the sharpness of the edges, however, it also narrowed handwriting lines – some of the pixels

Table 1. Evaluation of binarization and filtration techniques using AV measure

	Filtering				
Binarization	**Raw**	**Median**	**Gaussian**	**Kuwahara**	**Perona-Malik**
Kapur-S-H	0,8580	**0,8669**	0,8599	0,8424	**0,8669**
Otsu	0,9002	0,9015	**0,9057**	0,8856	0,8993
Ridley-Calvard	0,9021	0,9028	**0,9063**	0,8860	0,8995
Bradley	0,9115	0,9149	0,9172	0,8976	**0,9180**
Niblack	0,8644	0,8647	0,8622	0,8702	**0,8735**
Savola	0,9136	0,9121	0,9129	0,9008	**0,9143**
White-Rohrer	**0,8981**	0,8956	0,8966	0,8896	0,8979
Average	0,8926	0,8941	0,8944	0,8817	**0,8956**

comprising script were incorrectly marked as a background. which had a negative impact on the final assessment. This is the effect of the Kuwahara algorithm, where each pixel is replaced with mean value of the neighbourhood section with smallest variance. As a result, the writing curve was partially replaced by background mean, which is characterized by lower variability. The Malik-Perona diffusion algorithm performed best on average, but the improvement did not occur for every binarization method. The local thesholding techniques: Bradley, Niblack and Savola obtained best results when applied to the output of Malik-Perona filtration. On the other hand, the global methods: Otsu and Ridley-Calvard did not achieved their highest performance. This may be due to the fact that the Malik-Perona algorithm also preserves artifacts edges, and therefore makes it more difficult to separate them from true script using one global threshold. In the case of local binarization methods, the threshold adapts to the surroundings and it is easier to separate the artifacts even if their edges become more pronounced. Otsu and Ridley-Calvard global methods were better suited to Gauss filtration. Kapur performed best with both median and Malik-Perona, however the performance it achieved was smallest among all of the analyzed algorithms. For the White-Rohrer method, the best result was obtained without the use of filtration at all. This indicates that the filtering image before binarization is not always helpful.

6. Conclusions

Based on the obtained results it can be concluded that the decision whether to use the initial filtering or which filtration algorithm should be selected depends on the binarization algorithm. Initial filtration, if well chosen for the binarization method, may improve the result, but if it is not appropriate, the effect may be opposite of what is expected. As part of further research, it is planned to examine other techniques of filtration and binarization. We also planning to increase the number of analyzed images and conduct separate experiments with images containing different types of artifacts.

References

[1] Adamski, M. and Saeed, K.: Signature verification by only single genuine sample in offline and online systems, AIP Conference Proceedings, vol. 1738, no. 1, 2016.

[2] Bradley, D. and Roth, G.: Adaptive Thresholding using the Integral Image, Journal of Graphics Tools, vol. 12, no. 2, pp. 13-21, 2007.

[3] Gonzalez, R. C. and Woods, R. E.: Digital Image Processing, 3rd Edition, Prentice-Hall, Inc., 2007.

[4] Gupta, M. R., Jacobson, N. P., and Garcia, E. K.: OCR binarization and image pre-processing for searching historical documents, Pattern Recognition, vol. 40, no. 2, pp. 389-397, 2007.

[5] Kapur, J. N., Sahoo, P. K., and Wong, A. K. C.: A new method for gray-level picture thresholding using the entropy of the histogram, Computer Vision, Graphics, and Image Processing, vol. 29, no. 3, pp. 273-285, 1985.

[6] Niblack, W.: An introduction to image processing, Prentice-Hall, 1986.

[7] Otsu, N.: A Threshold Selection Method from Gray-Level Histograms, IEEE Transactions on Systems, Man, and Cybernetics, vol. 9, no. 1, pp. 62-66, 1979.

[8] Papari, G., Petkov, N., and Campisi, P.: Artistic Edge and Corner Enhancing Smoothing, IEEE Transactions on Image Processing, vol. 16, no. 10, pp. 2449-2462, 2007.

[9] Perona, P. and Malik, J.: Scale-space and edge detection using anisotropic diffusion, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no. 7, pp. 629-639, 1990.

[10] Pratikakis, I., Zagoris, K., Barlas, G., and Gatos, B.: ICDAR2017 Competition on Document Image Binarization (DIBCO 2017), 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), pp. 1395-1403, 2017.

[11] Pratikakis, I., Zagoris, K., Barlas, G., and Gatos, B.: ICFHR2016 Handwritten Document Image Binarization Contest (H-DIBCO 2016), 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), pp. 619-623, 2016.

[12] Ridler, T. W. and Calvard, S.: Picture Thresholding Using an Iterative Selection Method, IEEE Transactions on Systems, Man, and Cybernetics, vol. 8, no. 8, pp. 630-632, 1978.

[13] Sauvola, J. and Pietikäinen, M.: Adaptive document image binarization, Pattern Recognition, vol. 33, no. 2, pp. 225-236, 2000.
[14] Sezgin, M. and Sankur, B.: Survey over image thresholding techniques and quantitative performance evaluation, Journal of Electronic Imaging, vol. 13, no. 1, pp. 13-20, 2004.
[15] Smith, E. H. B., Likforman-Sulem, L., and Darbon, J.: Effect of pre-processing on binarization, Document Recognition and Retrieval XVII, SPIE Electronic Imaging Symposium, vol. 7534, 2010.
[16] White, J. M. and Rohrer, G. D.: Image Thresholding for Optical Character Recognition and Other Applications Requiring Character Image Extraction, IBM Journal of Research and Development, vol. 27, no. 4, pp. 400-411, 1983.

WPŁW FILTRÓW NA JAKOŚĆ BINARYZACJI OBRAZÓW PISMA ODRĘCZNEGO

Streszczenie Filtracja i binaryzacja są często stosowanymi technikami w systemach rozpoznawania pisma odręcznego. Operacje te są wykonywane w ramach etapu zwanego wstępnym przetwarzaniem , którego rezultat jest przekazywany do kolejnych etapów: ekstrakcji cech i klasyfikacji. Operacje wykonywane w ramach wstępnego przetwarzania są istotne ponieważ ich wynik wpływają na poprawność pracy całego systemu. W niniejszej pracy skupiono się nad oceną wpływu wyboru metody filtracji obrazu na efekt procesu binaryzacji dla obrazów z pismem odręcznym. W eksperymentach zbadano 4 metody filtracji w połączeniu z 7 metodami progowania dla różnych kombinacji parametrów tych metod. Do eksperymentów użyto wybrane obrazy z pismem odręcznym z baz konkursów binaryzacji dokumentów DIBCO oraz obrazy referencyjne do oceny poprawności binaryzacji. Ocenę wykonano na bazie średniej z miar F-measure, Accuracy, Relative Foreground Area Error, Region nonuniformity.

Słowa kluczowe: filtracja obrazu, binaryzacja obrazu, pismo odręczne

Artykuł zrealizowano w ramach pracy badawczej S/WI/2/2018.