IOP Conference Series: Earth and Environmental Science

PAPER • **OPEN ACCESS**

Improving the processing of machine vision images of robotic systems in the Arctic

To cite this article: N Milyaev *et al* 2019 *IOP Conf. Ser.: Earth Environ. Sci.* **302** 012061

Improving the processing of machine vision images of robotic systems in the Arctic

N Milyaev[1], A Kaznin[1], O Sushko[1] and O Skotarenko[2]

[1]Northern (Arctic) Federal University named after M.V. Lomonosov
[2]Murmansk Arctic state university, Murmansk, Russia

ni_milyaev@mail.ru, a.kaznin@narfu.ru, o.sushko@narfu.ru, ksen12@mail.ru

Abstract. The transformation of the development of the Arctic is due to modern robotic systems. The use of unmanned vehicles in many industries in the Arctic provides an array of photo and video information. Accuracy of image analysis and pattern recognition is enhanced by image preprocessing. However, the existing binarization algorithms are not universal for images with different distortions and loss of information. The accuracy of binarization algorithms depends on many factors, such as shadows, uneven lighting, low contrast, noise, etc. Images with different characteristics of light and noise are simulated in order to model various lighting conditions on information from digital cameras of robotic systems. The paper investigates global and adaptive image binarization algorithms. The binarized images were obtained using these algorithms and the results of binarized images recognition are compared by an optical character recognition system. An analysis of the comparison results showed that for images made in poor lighting conditions or images with low contrast, or images with high noise levels, adaptive binarization algorithms are better suited. However, in most cases it is not possible to obtain fully correctly recognized images.The paper proposes a new binarization method based on artificial neural networks. The process of creating an artificial neural network is shown, include the parameters for determining the class of a pixel, the adjustment of weights, the architecture of an artificial neural network. A comparison of the proposed artificial neural network with existing image binarization algorithms demonstrate that in most cases the artificial neural network has the result of image processing at the level of adaptive algorithms or higher. The proposed method of images binarization based on the image color characteristics analysis allows to solve image recognition tasks by robotized systems.

1. Introduction

It is impossible to imagine the development of the Arctic without modern robotic systems. Robotics today occupies a considerable share in the cognitive-entertaining sphere, in training, in everyday life. Robots are used in such socio-economic areas as surgical and restorative medicine, in providing security, in the spheres of transport, energy production and space. The use of robotics in harsh climatic conditions is especially relevant. Robotics should find use in the exploration, development and extraction of natural resources, which is especially important for the Arctic, where temperatures are negative throughout the year, and there is little sunlight for most of the year. In addition to the difficult and harsh climatic conditions, the use of robotics and the application of robotic systems are due to the large distances with the most of the territory unpopulated, sparse population and underdeveloped infrastructure.

Because of the inaccessibility of the Arctic territories, unmanned vehicles are used [1]. Unmanned vehicles are not only those in which the operator is physically absent, but also the ones which are not controlled by the operator remotely. Such devices must perform the largest possible number of operations autonomously. One of the most important data used for information processing and spatial orientation of robotic systems is images and video data from digital cameras. Since robotic systems perform tasks autonomously, they make decisions on their own, based also on data from digital cameras (data can be text characters, special characters, QR codes and other graphic images that require separation of the necessary information from the image background). The solution to the computer vision problem is often confronted with the fact that the image may contain noise, flare and other graphical artifacts [2], which leads to incorrect recognition of data and, as a result, wrong decisions by robotic systems. To solve this problem, image preprocessing is often used [3]. However, the existing binarization algorithms are not universal for images with different distortions and loss of information.

To correctly evaluate the results of image binarization, it is advisable to take the original image containing the text and test the operation of the binarization algorithms, including the method proposed in the article based on artificial neural networks, by "scrolling" binarized images through the optical character recognition system

2. Image binarization

To separate the characters from the background at the binarization stage, the color image is converted to black and white. Qualitatively conducted binarization can significantly increase the number of correctly recognized characters. Another advantage of image binarization is significant reduction in the amount of memory needed to store them.

The process of image binarization is traditionally divided into two main areas: global and local [3, 5, 6]. The main difference between these groups is the method of calculating the binarization threshold, relative to which it is determined which image pixels should be classified as "useful" and which as "background".

2.1. Global algorithms

The work of global algorithms is based on the determination of a single binarization threshold for all image pixels. The advantage of this group of algorithms lies in the simplicity of implementation and speed with respect to local algorithms [6]. Global binarization algorithms include: Otsu's method [7], upper and lower threshold methods [8].

2.2. Adaptive algorithms

In adaptive algorithms, the binarization threshold is calculated not for the entire image, but for each individual pixel or a part of the image. This group of algorithms shows a higher-quality result of binarization compared to global algorithms for images containing light shifts, blur, low resolution. The adaptive binarization algorithms include: Niblack's binarization method [9], Sauvola's binarization method [10].

3. Comparison of algorithms

To conduct the study comparing the operation of binarization algorithms, the following algorithms, most often used in character recognition tasks, were chosen: Otsu's method, Niblack's method, Sauvola's method [3, 5, 6].

All of the above algorithms are adaptive, except for Otsu's method, which is a global one, but is recognized as one of the most efficient binarization algorithms in its group. To test the algorithms, a sample was made of 6 images containing different lighting conditions, contrast, and noise (Fig. 1).

IOP Conf. Series: Earth and Environmental Science **302** (2019) 012061 doi:10.1088/1755-1315/302/1/012061

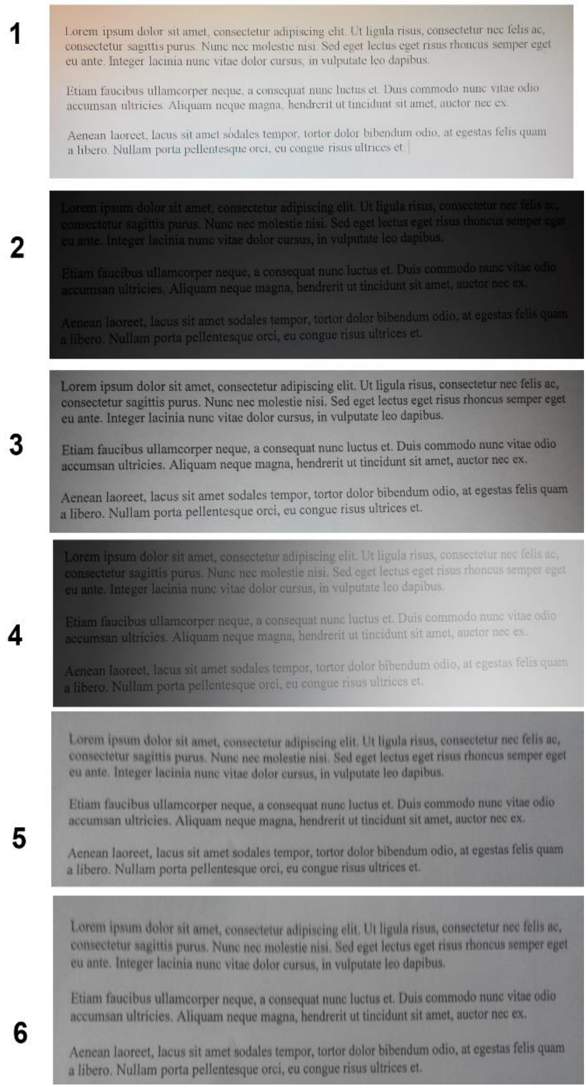

Figure 1. Pictures with different conditions
1 – Display; 2 – Paper (low contrast) 3 – Paper (gradient lighting) 4 – Paper (blur) 5 – Display (glare)

The quality of the OCR systems [11] is estimated by the number of correctly recognized characters. If only one of the hundred characters is recognized incorrectly, then the recognition system accuracy is 99%, if one of the thousand, then the accuracy is 99.9% [4].

The measurement is made by comparing the source text and the text obtained as a result of the operation of the OCR system.

Using this method of quality assessment, the operation of various binarization algorithms can be compared when processing the images presented in Figure 1.

For the recognition of characters on binarized images, Microsoft computer vision API was used. This API is a part of Microsoft Cognitive Services (formerly Project Oxford) [12] and allows developers and researchers to add machine learning functions to their applications. At the moment, the set includes tools such as the detection of emotions in photos and videos, face recognition, voice recognition, OCR and others. As a result of character recognition on binarized images, the results presented in Figure 2 were obtained.

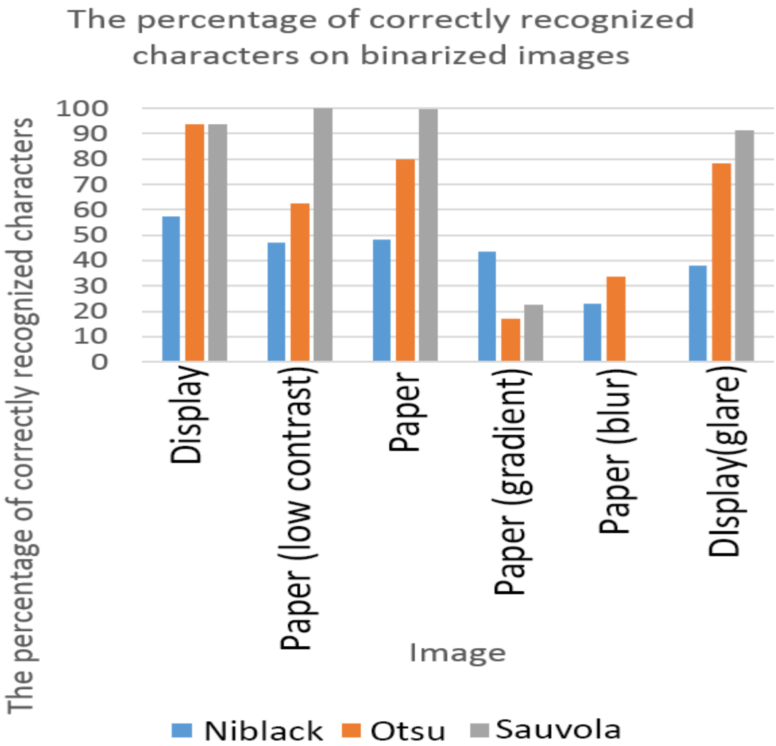

Figure 2**.** The percentage of correctly recognized characters on binarized images

From this analysis, it follows that adaptive algorithms are better suited for images with complex lighting conditions, low contrast and high noise levels, but they cannot cope with all the possible drawbacks found in the images from which text should be extracted.

4. Application of artificial neural networks for image binarization

Image binarization can be viewed as the task of classifying pixels into "useful" and "background" ones.

It is possible to separate objects from the background on an image only on the basis of a certain probability, since image pixels with the same characteristics can belong both to the class of "useful" information and to the class of "background" information. An example is shown in Figure 3. The red color in the figure highlights the pixels that have the same brightness level on the original image.

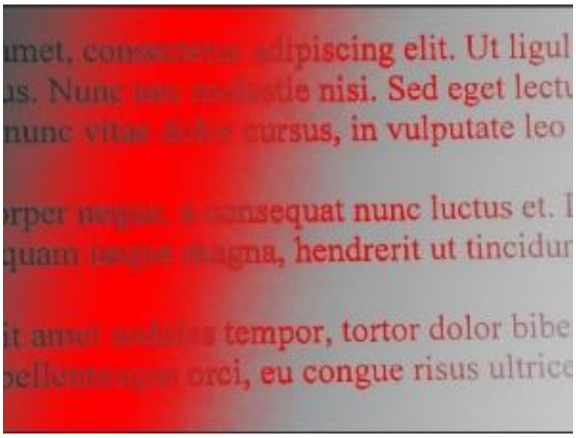

Figure 3. Text and background area with the same level of pixel brightness

Artificial neural network (ANN) with direct connection have become a proven means of approximation of functions and classification. Neural networks effectively cope with the task of classification due to the fact that in the process of work they generate a large number of regression models used in the classification using traditional statistical methods.

To create a neural network capable of binarizing images, a number of steps should be performed:

- determine the parameters of the input vector;
- create a learning sample;
- select the network architecture.

Modern digital cameras and scanners, as a rule, work on the basis of the RGB color scheme, which makes it one of the most common ones [13]. Channels of colors in this model range from 0 to 255. Different shades are obtained by combining different combinations of original channels: red, green, and blue.

It is these parameters of each pixel that can be used as the main characteristics with which the pixel class is determined.

To configure the significance of the network, a learning sample of text photographs was created with various conditions of brightness and contrast. A binarized image was created for each image from the sample. The process of creating pairs of input data and the expected output of the network consisted in applying Sauvola's method and further correcting graphical artifacts using Photoshop CC.

To classify image pixels represented in the RGB color scheme, one neuron for each color channel must be placed in the input network layer. There will be one neuron in the output layer, since all the pixels are divided into two classes: "useful" and "background" ones. A hidden layer with neurons, the number of which is chosen empirically in the process of network learning is placed between the input and output layers.

The network diagram for pixel classification is presented in Figure 4.

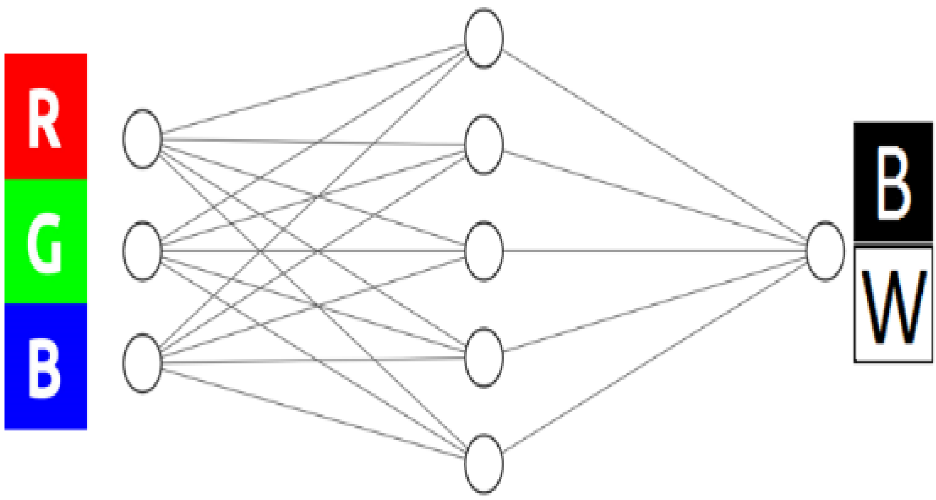

Figure 4. Architecture of neural network for image binarization

A multilayer perceptron, consisting of input, hidden, output layers containing 3, 5 and 1 neuron, respectively was chosen as the network architecture. The neural network was implemented and learnt using the software developed on the basis of FANN library [14] and the learning algorithm RPROP [15]. The results of binarization by existing algorithms and using the proposed neural network are presented in Figure 5.

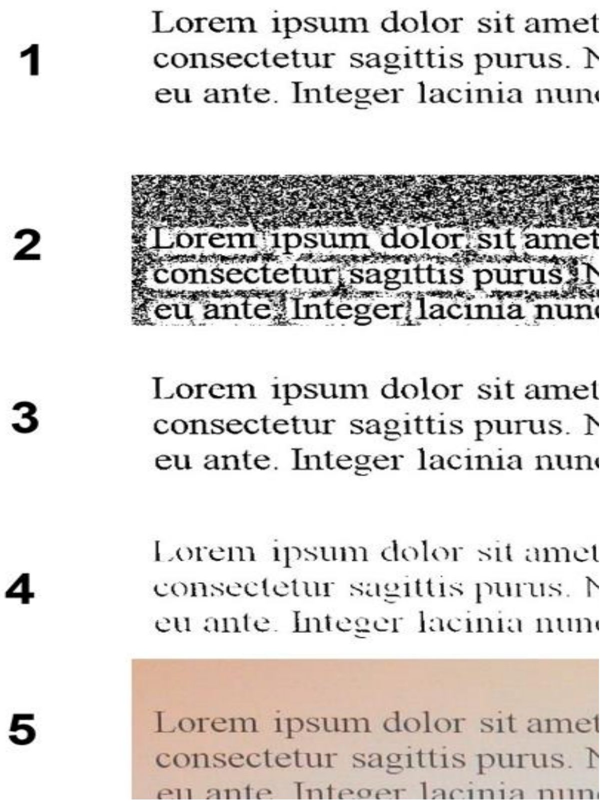

Figure 5. The results of image binarization
1 – ANN; 2 – Niblack's method; 3 – Otsu's method; 4 – Sauvola's method; 5 – Original image

All binarization algorithms showed a good result, but in the case of Sauvola's method, gaps appeared in the characters, which can lead to incorrect character recognition. The disadvantage of Niblack's method is the appearance of background noise and the loss of useful information on the image. Image processing using the Otsu's method places high demands on the contrast of the original images.

The final step in comparing the developed neural network and existing binarization algorithms is to compare the percentage of correctly recognized characters on binarized images using Microsoft Cognitive Services. The result is shown in Figure 6.

All binarization algorithms showed a good result, but in the case of Sauvola's method, gaps appeared in the characters, which can lead to incorrect character recognition. The disadvantage of Niblack's method is the appearance of background noise and the loss of useful information on the image. Image processing using the Otsu's method places high demands on the contrast of the original images.

The final step in comparing the developed neural network and existing binarization algorithms is to compare the percentage of correctly recognized characters on binarized images using Microsoft Cognitive Services. The result is shown in Figure 6.

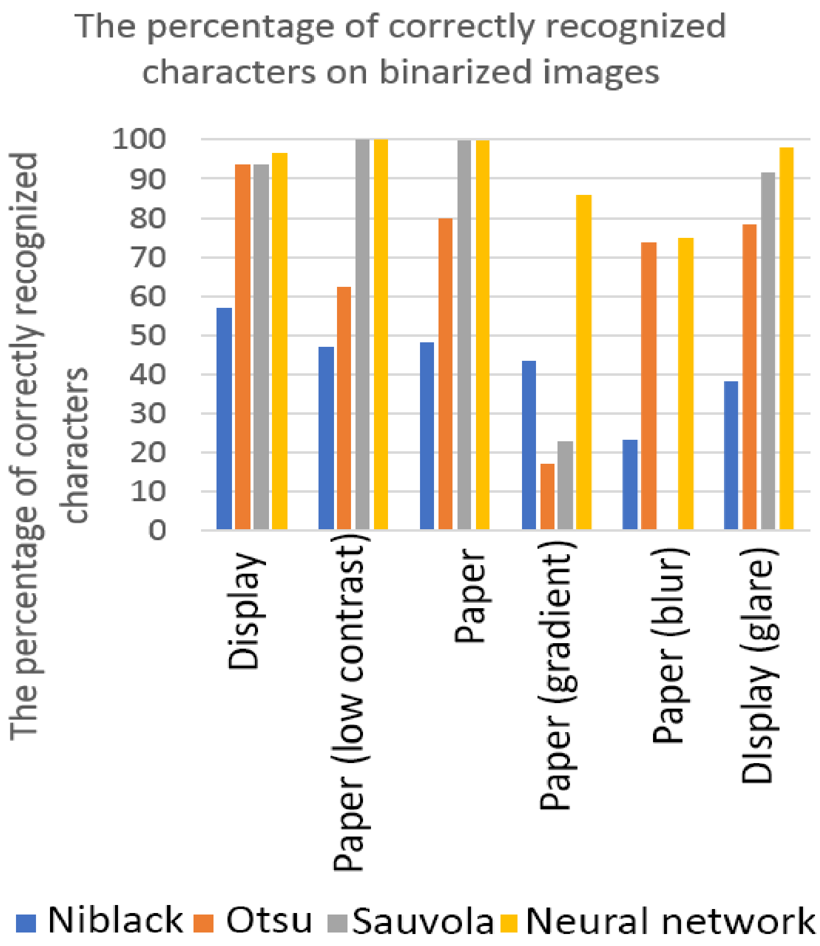

Figure 6. The percentage of correctly recognized characters on binarized images

From this comparison it is clear that in most cases the proposed neural network showed the result of image processing at the level of adaptive algorithms or higher.

5. Conclusions

The transformation of the development of the Arctic is due to the active expansion of robotic systems. Robotic technologies are of particular importance for the study and development of the Arctic, and they open up new socio-economic, political perspectives and opportunities. Robotization and automation also increase the efficiency and safety of traditional human activities, and, in difficult conditions, are able to function without human intervention. The new binarization method proposed in this paper based on the analysis of color characteristics of the image makes it possible to obtain binarized data suitable for further recognition. When creating robotic systems for the Arctic, the use of the binarization method based on the analysis of the color characteristics of the image will effectively solve recognition problems, especially text character, special characters and QR codes recognition. Further improvement of the method is possible by increasing the number of image parameters processed using a neural network.

References

[1] *Arctic - a place for robots* 2019 Available from**:** https://www.korabel.ru/news/comments/arktika_-_mesto_dlya_robotov.html [Accessed 20th March 2019]
[2] Kaznin A A *et al* 2017 *Developing the algorithm allowing business-dedicated mobile applications to read texts* (Proceedings of the 2017 International Conference "Quality Management, Transport and Information Security, Information Technologies") pp 207-214
[3] Rashmi S 2015 *Document Image Binarization Techniques, Developments and Related* Issues: A Review (International Journal of Computer Applications) **116**(7) pp 118 – 122
[4] *Improve OCR Accuracy with Advanced Image Preprocessing* 2019 Available from**:** https://docparser.com/blog/improve-ocr-accuracy [Accessed 20th March 2019]
[5] Mahajan P A 2015 *Review of Image Binarization Techniques for Ancient Degraded Documents* (International Journal of Science and Research (IJSR)) **11**(4) pp 567 – 570
[6] Chaki N A 2014 *Comprehensive Survey on Image Binarization Techniques* (Springer India) p 82 Available from**:** https://www.springer.com/gp/book/9788132219064 [Accessed 20th March 2019]
[7] Otsu N 1979 *A threshold selection method from gray-level histogram* (IEEE Transactions on Systems, Man, and Cybernetics) **9**(1) pp 62–66
[8] Bathinda P G 2013 *Binarization Techniques used for Grey Scale Images* (International Journal of Computer Applications) **71**(1) pp 55 – 59
[9] Niblack W 1985 An introduction to digital image processing (Prentice Hall, Eaglewood Cliffs) p 215
[10] Sauvola J 1997 *Adaptive document binarization* (Fourth International Conference Document Analysis and Recognition, Ulm (Germany)) pp 147-152
[11] *Optical Character Recognition* 2019 Available from**:** https://www.abbyy.com/ru-ru/finereader/what-is-ocr [Accessed 20th March 2019]
[12] *Cognitive Services* 2018 Available from**:** https://docs.microsoft.com/ru-ru/azure/cognitive-services/welcome [Accessed 10th March 2019]
[13] Nishad P M 2013 *Various color spaces and color space conversion algorithms*: [Text] (Journal of global research in computer science) **4**(1) pp 82 – 87
[14] *Fast Artificial Neural Network Library* 2019 Available from**:** http://leenissen.dk/fann/wp [Accessed 10th March 2019]
[15] Riedmiller M 2002 *A direct method for faster backpropagation learning: the RPROP algorithm* (IEEE) p 15 Available from: https://www.springer.com/gp/book/9788132219064 [Accessed 20th March 2019]