SEGMENTATION OF THERMAL IMAGES

Ondrej Bostik
Doctoral Degree Programme (3), FEEC BUT
E-mail: bostik@feec.vutbr.cz

Supervised by: Karel Horak
E-mail: horak@feec.vutbr.cz

Abstract:

This paper presents our ongoing work focused on segmentation of thermal images from the process of traverse wedge rolling. The goal of this work is to evaluate some of the available methods. We mainly focused on a demonstration of simple methods without using machine learning methods. Part of the work is to present the dataset we create for testing.

Keywords: segmentation, thermal images, dataset, Otsu's segmentation, adaptive thresholding, active contours

1 INTRODUCTION

Thermal images dataset was created as part of our work to make dimension measurement and quality assessment of transverse wedge rolling in the production of automotive parts. The first goal is to observe the production and sort out bad parts. The second goal is to make exact measurements of parts and store them for later use. The third goal is to help workers with prototyping new forming rollers for a better understanding of the process.

The initial stage of the project is to choose a reliable technique for component segmentation. We need to create the dataset of thermal images with and without object from forming process, annotate them with human participants and study known techniques for the best case.

2 BASIC SEGMENTATION ALGORITHMS OVERVIEW

The first method used in this experiment is Otsu's segmentation method introduced in [1]. This method could be considered as outdated, but the simplicity of the method and it's speed made this method ideal for initial testing and for method comparison. The method proposed by Otsu is based on a search of single segmentation threshold searched not in the image itself, but in the histogram of the image. The criterion is to maximize the separability of the resultant classes in gray levels.

The second method considered for use is adaptive thresholding. The method introduced in [2] was selected for the experiment. Instead of a single threshold level for the whole image, the method produces a binary image where decision level is set based on running average of a square window around the current pixel. To speed up the whole process, the presented method use integral image [3] to sum the value of the surrounding pixels.

3 DATASET

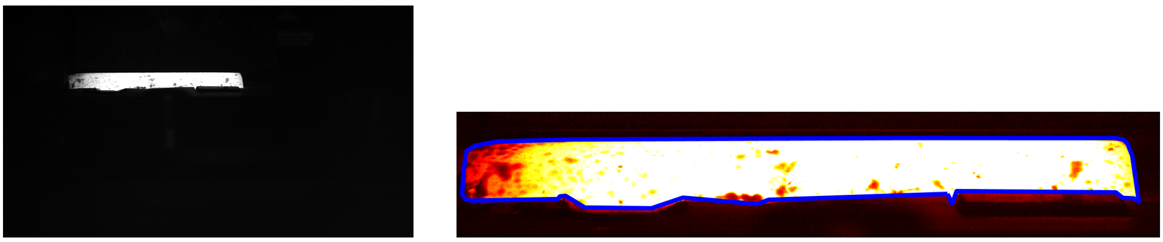
During the first stage of our project we record 7 videos of the above-explained process. Every video was taken with ImagingSource camera type DMK 33UP500 with IR filter. The resulting video has 30 frames per second with a resolution of 1920x1080px and saved to AVI file with Y800 video codec

(8 bits grayscale). The only deviation is dataset B, which has 2592x2048px. This dataset was added for speed comparison. Every video was broken into individual images, which was later annotated with human afford. The resulting annotation is a structure containing the file names and sets of points defining a polygonal region of interest. Every dataset is stored in a folder containing all images and Matlab binary data file witch annotation. Every processing was done in MATLAB R2018a with Image Processing Toolbox. Overview of our dataset is depicted in table 1. Note that not all resulting images contain the object.

Dataset	Images count	Images with object	Images without object
Dataset A	928	634	294
Dataset B	1144	1058	86
Dataset C	871	736	135
Dataset D	1183	658	525
Dataset E	1069	448	621
Dataset F	1103	710	393
Dataset G	313	210	103
Total	6611	4454	2157

Table 1: Dataset overview

Sample images are shown in figure 1 and 2. In the figure 1(a) is shown sample image taken from Dataset A without any processing. The second figure 1(b) show detailed view of the image with blue annotated border of target object. This images was colored with pseudo-colors only for visualization purposes.



(a) (b)

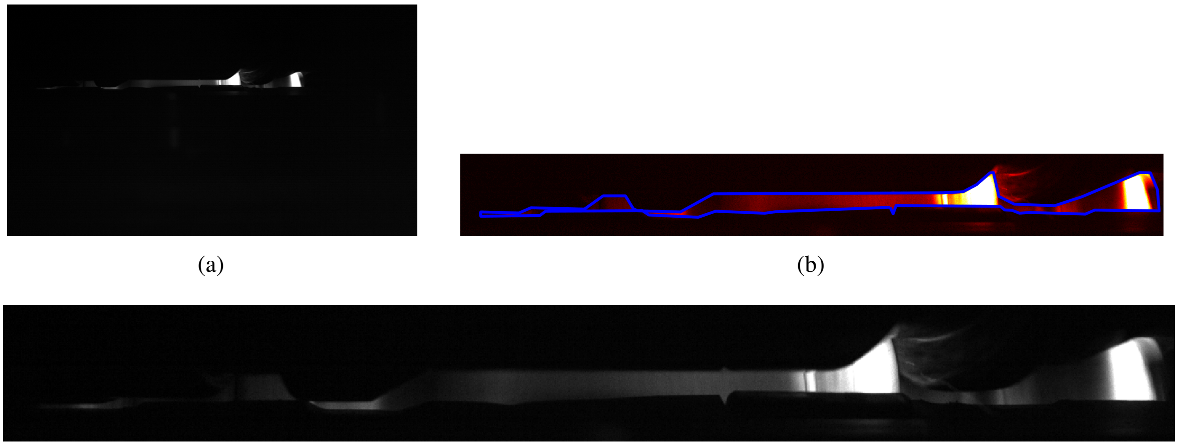
Figure 1: Example image from our thermal images dataset (a) original gray-scale image (b) detail of rolled object colored with pseudo-colors for visualization and a blue border made by annotators

The second example of the dataset (figure 2) displays one of the challenging images. Original image is shown in figure 2(a), the pseudo-colored image with annotated borders is depicted in 2(b). Detailed gray-scale figure 2(c) was added for the reader to better understand the complexity of the segmentation task.

4 EXPERIMENTS

Before the experiment itself has started, the evaluation framework was prepared. This framework takes all available datasets and compares region created during the annotation with either second region or with a binary image. For every image, we obtain error matrix [4] based on classification results. The main parameters we compare are Accuracy, Sensitivity, and Specificity.

In binary classification, accuracy can be explained as for how well the classification algorithm identifies the object or foreground. Accuracy is a portion of correctly segmented pixels over the total



(a) (b)

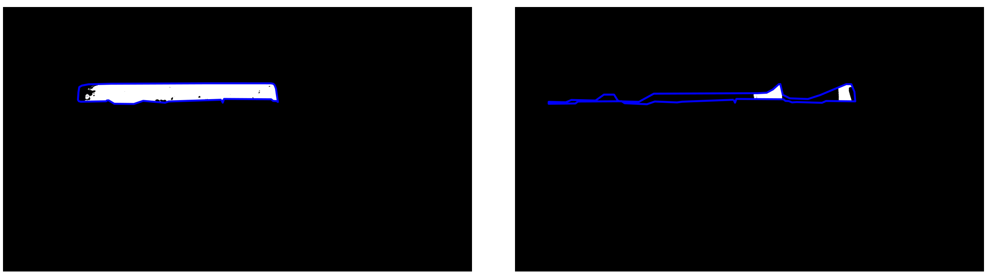
(c)

Figure 2: Example image of one of the challenging images (a) original gray-scale image (b) detail of rolled object colored with pseudo-colors for visualization and a blue border made by annotators (c) enlarged detail of rolled object in gray-scale

number of pixels. As the area of the segmented object is the only small portion of the whole image, this number can be very high even the object is not properly segmented. Sensitivity, also known as a probability of detection, is defined as a portion of correctly detected pixels with a presumed object to all pixels with an object from the annotation. This statistical information is much more informative to our case than the other two. The third parameter is specificity which measures a portion of correctly identified foreground pixels to actual foreground pixels.

The first experiment with the dataset utilized the Otsu's segmentation method explained above. Illustrative figure 3 depicted two cases. Left image 3(a) is almost correctly segmented because the metal is hot and the segmentation is therefore easy. In the right figure 3(b) we can observe picture from later part of the process where parts of metal are not as hot but clearly distinguishable for a human (compare to fig. 2(c)).

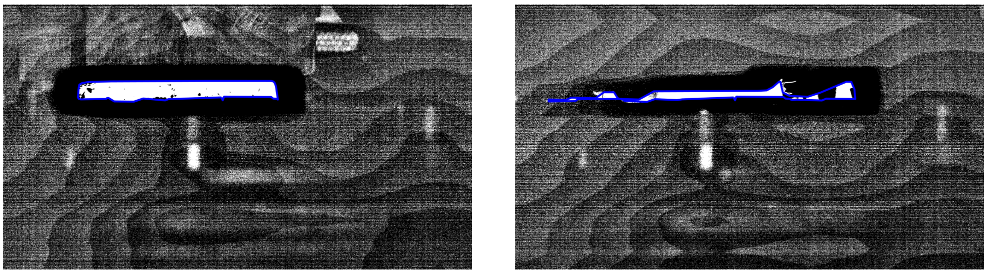
Otsu's method clearly fails for this kind of pictures. Results are depicted in table 3. As can be seen, the method is quickest of all. As we expected, accuracy is high at 91.40%. But the important parameter sensitivity is only 57.47%, which means that in the average image almost half of the image is not correctly labeled.



(a) (b)

Figure 3: Segmentation results for Otsu's segmentation method with a blue border from the annotation process (a) result for the image shown in figure 1 (b) results for challenging image described in figure 2

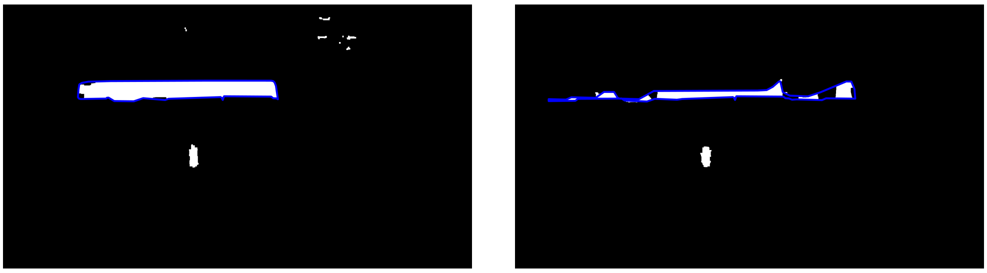
The second tested method was the use of adaptive thresholding. We utilized standard algorithm with neighborhood size set to 325x257px for dataset B and 135x241px for other datasets, local threshold compute by mean intensity and sensitivity set to 40%. As depicted in figure 4, plain adaptive thresholding works great around the object itself, but elsewhere fails due to noise. In table 3 this situation is shown by decreasing the accuracy and specificity, but the increase of sensitivity.



(a) (b)

Figure 4: Segmentation results for adaptive segmentation method with blue border from annotation process (a) results for image shown in figure 1 (b) results for challanging image described in figure 2

This method was therefore improved by use of the mathematical morphology [5]. The first added step was morphological opening with a square structural element with a size of 6x6px to reduce noise. The second step was to fill gaps in an object with morphological closing with a square element of 15x15px. Resulting images can be see in figure 5.



(a) (b)

Figure 5: Segmentation results for adaptive segmentation method augmented with mathematical morphology with blue border from annotation process (a) results for image shown in figure 1 (b) results for challanging image described in figure 2

The overall results for each dataset are depicted in table 2. Notice that average time for one image to process image with resolution 1920x1080 (datasets A, C-G) is only around 0.28s. Average processing time for image with resolution of 2592x2048px is 0.68% (2.4 times slower).

The total result in comparison with other methods are in table 3. Resulting accuracy and specificity are both increased from the other two methods. Sensitivity is 65.01% which is better only than Otsu's method.

5 CONCLUSION

This work presents the initial experiments in an ongoing project and describes some segmentation method usable for segmenting grayscale thermal images. We compare computationally simple methods, namely Otsu's segmentation method for compassion and two variations of adaptive thresholding.

Dataset	Time per image	Accuracy	Sensitivity	Specificity
Dataset A	0.29s	99.43%	85.03%	99.64%
Dataset B	0.68s	99.20%	82.61%	99.33%
Dataset C	0.28s	98.65%	62.21%	99.10%
Dataset D	0.28s	97.08%	56.65%	98.18%
Dataset E	0.28s	98.01%	58.44%	98.89%
Dataset F	0.28s	97.94%	60.80%	99.18%
Dataset G	0.29s	97.06%	59.23%	98.34%
Total	0.35s	98.47%	65.01%	99.07%

Table 2: Detailed results for adaptive segmentation method augmented with mathematical morphology

Method	Time per image	Accuracy	Sensitivity	Specificity
Otsu's segmentation method	0.19s	91.40%	57.47%	92.00%
Adaptive thresholding	0.30s	75.97%	68.37%	76.11%
Adaptive thresholding with morphology	0.35s	98.47%	65.01%	99.07%

Table 3: Summary comparison of all used methods

The best of this method is adaptive thresholding improved with some minor morphological operations. This method evaluated pixel-by-pixel has the accuracy of 98,47%. But sensitivity is still quite low at 65.05%, that means 35% of an object is not properly located. The method is very quick, processing time increase almost linearly with resolution. It will be used for initial rough region estimation for further processing and object localization.

The further work will be focused around segmentation with a convolutional neural network.

ACKNOWLEDGEMENT

The completion of this paper was made possible by the grant No. FEKT-S-17-4234 - 'Industry 4.0 in automation and cybernetics' financially supported by the Internal science fund of Brno University of Technology.

REFERENCES

[1] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Trans. Syst. Man. Cybern.*, vol. 9, no. 1, pp. 62–66, 1979.

[2] D. Bradley and G. Roth, "Adaptive Thresholding using the Integral Image," *J. Graph. Tools*, vol. 12, pp. 13–21, 2007.

[3] F. C. Crow, "Summed-area Tables for Texture Mapping," in *Proc. 11th Annu. Conf. Comput. Graph. Interact. Tech.*, SIGGRAPH '84, (New York, NY, USA), pp. 207–212, ACM, 1984.

[4] S. V. Stehman, "Selecting and interpreting measures of thematic classification accuracy," *Remote Sens. Environ.*, vol. 62, pp. 77–89, oct 1997.

[5] L. Najman and H. Talbot, "Mathematical Morphology: From Theory to Applications," *Math. Morphol. From Theory to Appl.*, 2013.