OPTIMUM BINARIZATION OF TECHNICAL DOCUMENT IMAGES

Jorge Sánchez Valverde[1], *Rolf-Rainer Grigat*[2]

[1]Ericsson, Madrid, Spain
[2]Technical University Hamburg-Harburg, Germany
http://www.ti1.tu-harburg.de

ABSTRACT

We present a high quality solution to the binarization problem of technical document images. The method is based on locally adaptive methods without the need of manual interaction or tuning. The algorithm includes Niblack's binarization and two validation steps based on morphological image processing and gradient based decisions. Our result has the following advantages. Alphanumeric labels are restored, even if they are hardly readable in the source image. Parallel, multiple lines with small distance are reconstructed. Even detail of very low contrast in the vicinity of strong contrast image areas is reconstructed (e.g. faint stamp lines). Variable background intensity is suppressed, but texture and large font bold face letters are restored.

1. INTRODUCTION

The optimal binarization of technical document images used e.g. for the construction of plants, vehicles or ships is the prerequisite for efficient digital storage and transmission or for high quality printout and copies. In typical today's copy machines the user achieves some global optimization by trial and error. For large image sizes of 30 inches width or more this is not feasible as lots of large format paper would be spoiled. Typically global binarization is unsatisfactory, and the threshold has to be optimized locally, automatically and fast. Especially for low quality source images the binarization problem is not adequately solved yet.

For digital processing the paper documents have to be scanned. The resulting gray valued images are quantized into a binary representation to restore optimal quality and to reduce the amount of data. We developed a method on the subsymbolic level, i.e. only pixel values (in a local neighborhood) are used. Therefore our method is of general use and not restricted to some given semantic context. Our solution can cope with severe image degradations like preliminary bad quality copy steps, very low contrast, manual corrections in printings, rubber stamp imprints, pencil comments, metallic labels or coffee spots. The main building blocks are Niblack's binarization [6] and two validation steps based on morphological image processing and gradient based decisions.

A combination of global and local adaptability can help in cases of textured background that should be removed, and in cases of inverted images, i.e. white on black.

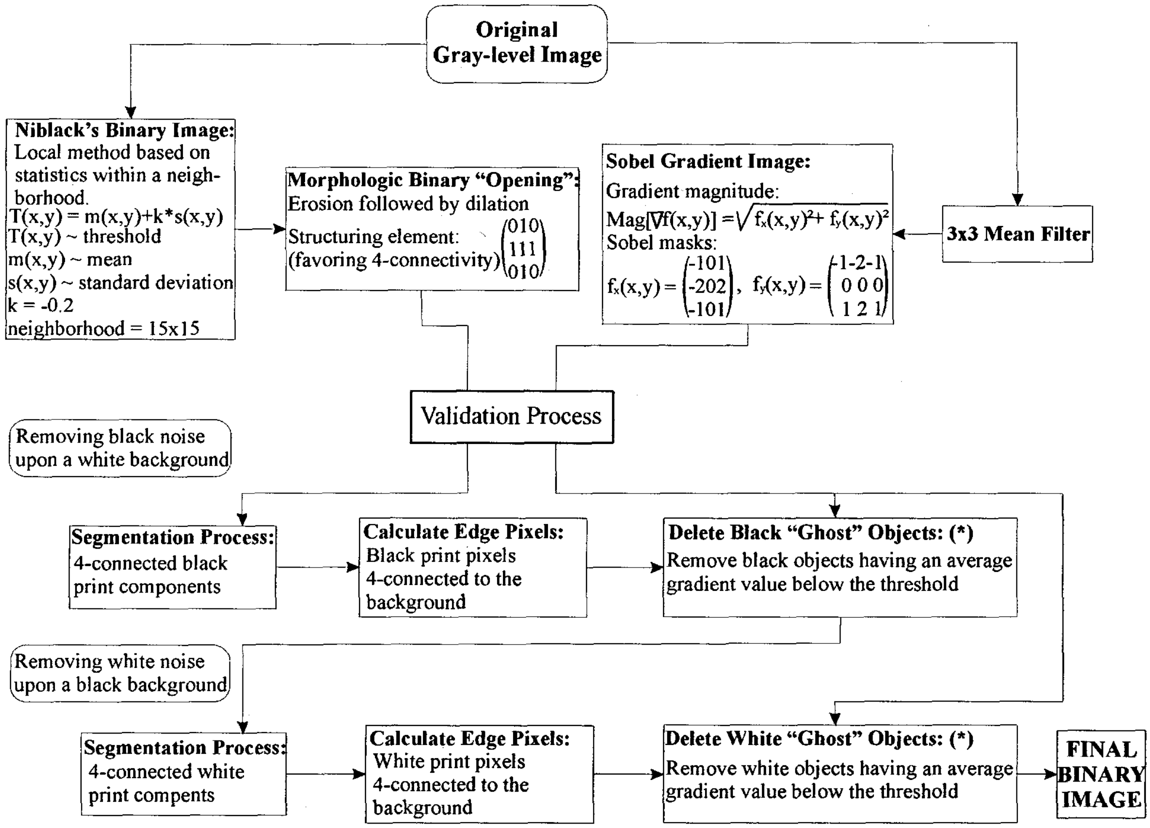
In the second section our method is described, in the third section the computational speed is discussed and in the fourth section the most relevant results of our practical evaluations are demonstrated.

2. THE BINARIZATION AND VALIDATION METHODS

The block diagram of our system is given in figure 1. It is based on a binarization step and two subsequent validation steps. For the first binarization step it was immediately clear that we had to focus on local binarization methods, see [3-8]. According to discussions by Trier and Jain [1] as well as Trier and Taxt [2] Niblack's method [6] has the best ratio of performance/speed in terms of error rate and reject rate. In Niblack's method a local threshold $T(x,y)=m(x,y)+k\, s(x,y)$ is used, where m and s are the sample mean and standard deviation values of a neighborhood of (x,y). The neighborhood has to be large enough to remove noise and small enough to conserve relevant structures. We found a neighborhood 15x15 and $k= -0{,}2$ to be a good choice. An example is given in figure 2. It is obvious, that some post-processing is required to remove "ghost" objects. Nevertheless, even the most faded objects are well segmented and separated from other structures.

Noise removal comprises two phases. Using morphological opening operations [9-11] the noise attached to valid structures is isolated. In a subsequent segmentation step the connectivity is tested with a 4-

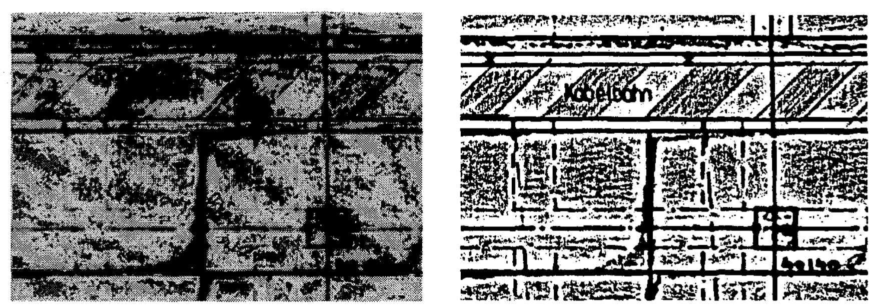
0-7803-6297-7/00/$10.00 © 2000 IEEE



(*) We divide the edge pixels in groups to calculate the gradient average. These groups as well as the various thresholds have been chosen depending on the object size. These criteria are a consequence of the type of images we are dealing with and a "trial and error" work.

Figure 1: Flow chart of the binarization algorithm

Figure 2: Binarization result of Niblack's operator. Even the most faded objects are well segmented and separated from other structures. Remaining ghost objects have to be removed in a post-processing step.



connected structuring element, as lines in technical documents are mainly oriented horizontally or vertically. The connected structures are then validated by using edge gradient information, i.e. taking full information of the input gray valued image into account. By using gradients we become independent on the background amplitude. We found the Sobel operator to be the most adequate gradient operator cause to its low noise sensitivity [12]. Black and white 4-connected "ghost" objects are removed by calculating the average gradient of connected areas. If the average gradient is below a certain threshold, the object is deleted.

The problem of the reconstruction of parallel, faint lines close to one another could in principle be solved by a combination of thinning [9], directional close and dilation.

Unfortunately this would be very specific to a certain pattern, while we need an algorithm to work globally in all kind of features. Such directional methods also proved to be too slow for our speed requirements.
Finally, in our application the 4-connectivity was adequate.

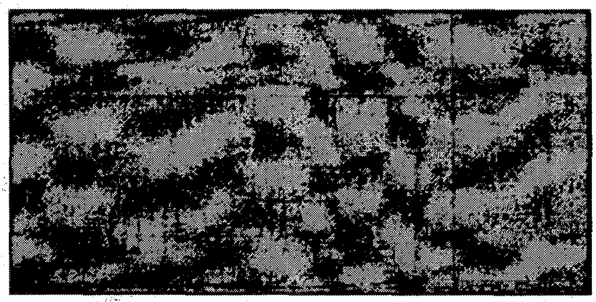
3. IMPLEMENTATION

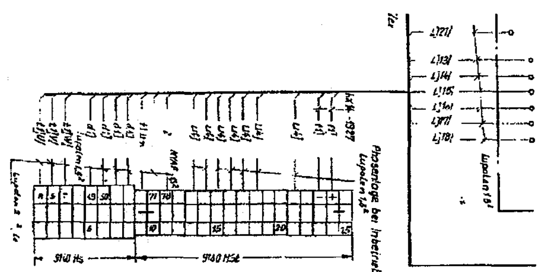
The algorithm has been evaluated on an AMDK6-2 400 processor with 128 MByte of RAM. The first steps have been implemented in MATLAB. By porting to C and some optimizations of the algorithms an acceleration of factor 100 was achieved.

The calculation time is now about 32 µs per pixel, i.e. 7 min 24 sec for a 4486x3400 image, which will be further improved.

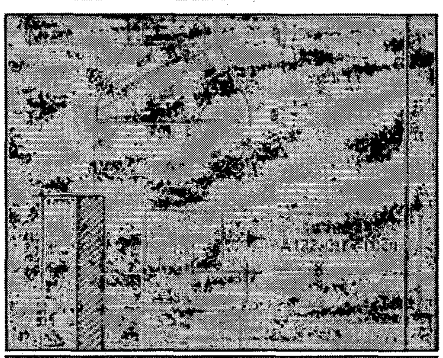
4. RESULTS

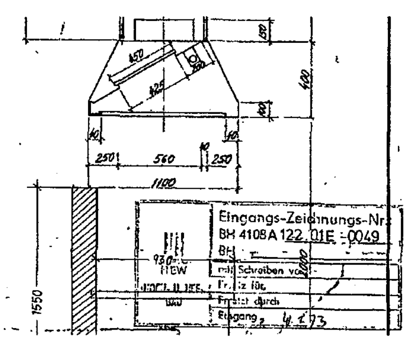
In this section the most difficult examples in our test set are presented. In those situations other binarization methods failed, which was the motivation for our investigations. It is important to point out that all the results documented here have been obtained with the same parameter setting, i.e. without any human trimming of the program behavior. The following five give an idea of challenging local details of some large original images of size 17x24 square inches or larger. The left image of each pair is always the original copy, as it is produced by the scanner. The right image is our binarization result.
In the first example a very low contrast image is restored. The feature contrast is comparable to the background texture contrast.



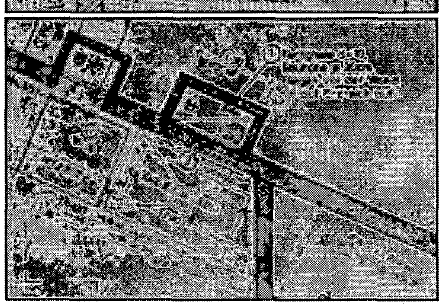


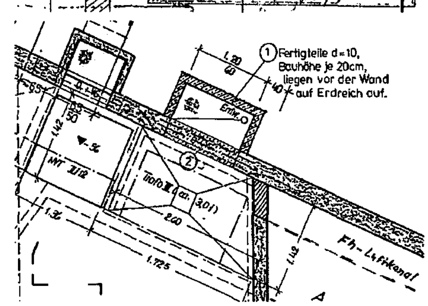
Faint contrast detail, as e.g. the stamp, is restored in the vicinity of high contrast detail. The readability is mostly better than in the gray valued original image.



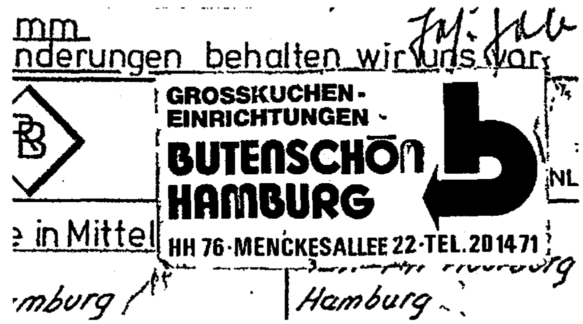


Again the valid information intensity differs significantly. The texture in the thick bar should not be removed, but the background texture should.

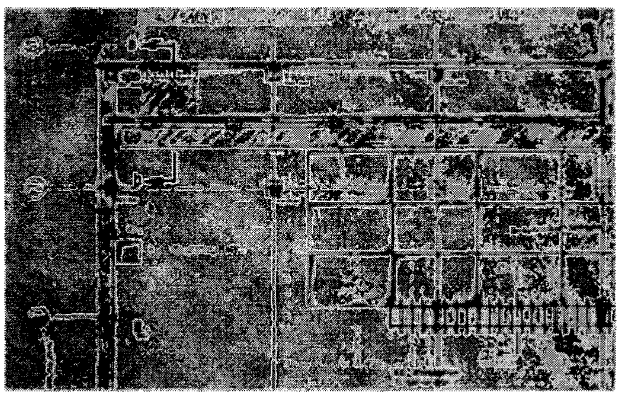


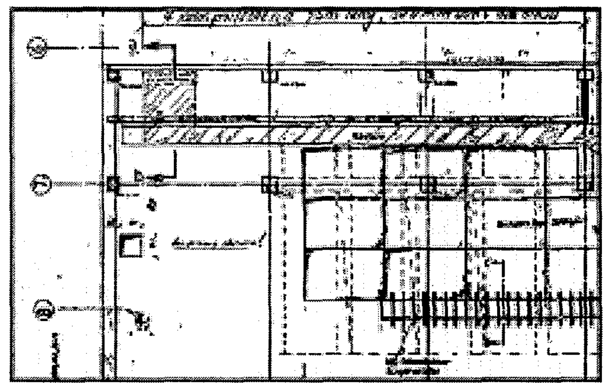


The aluminum label is reflecting the scanning light. The black areas between the bright reflections are removed, but the large area black symbol is maintained. The text is restored to a much better readability than in the original.



In the last example the text string "Kabelbahn" and also the very small numerical annotations should be compared to the gray valued original. Very low contrast lines are completely restored.





5. REFERENCES

[1] Ø. D. Trier and T. Taxt, Evaluation of binarization methods for document images. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 17, no. 3, pp. 312-315, Mar. 1995.

[2] Ø. D. Trier and A. K. Jain, Goal-Directed Evaluation of Binarization Methods. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 17, no. 12, Dec. 1995.

[3] J. Bernsen, Dynamic thresholding of gray-level images. Proc. Eighth Int'l Conf. Pattern Recognition, pp. 1,251-1,255, Paris, 1986.

[4] C. K. Chow and T. Kaneko, Automatic detection of the left ventricle from cineangiograms. Computers and Biomedical Research, vol. 5, pp. 388-410, 1972.

[5] K. V. Mardia and T. J. Hainsworth, A spatial thresholding method for image segmentation. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 10, no. 6, pp. 919-927, 1988.

[6] W. Niblack, An Introduction to Digital Image Processing, pp. 115-116. Englewood Cliffs, New Jersey. Prentice Hall, 1986.

[7] S. D. Yanowitz and A. M. Bruckstein, A new method for image segmentation. Computer Vision, Graphics and Image Processing, vol. 46, no. 1, pp. 82-95, Apr. 1989.

[8] Y. Nakagawa and A. Rosenfeld, Some experiments on variable thresholding. Pattern Recognition, vol. 11, no. 3, pp. 191-204, 1979.

[9] J. Serra, Image Analysis and Mathematical Morphology. Academic Press, London, 1982.

[10] C. R. Giardina and E. R. Dougherty, Morphological Methods in Image and Signal Processing. Prentice-Hall, Englewood Cliffs, NJ, 1988.

[11] J. Serra and P. Soille, Mathematical Morphology and Its Applications to Image Processing. Kluwer Academic Publishers, 1994.

[12] W. K. Pratt, Digital Image Processing. Second Edition, John Wiley & Sons, 1991.