Fast Implementation of the Niblack Binarization Algorithm for Microscope Image Segmentation[1,2]

O. A. Samorodova and A. V. Samorodov

Chair for Biomedical Technical Systems, Faculty of Biomedical Technique, Bauman Moscow State Technical University, Moscow, Russia

e-mail: artyukhova@bmstu.ru

Abstract—A fast way to implement the Niblack binarization algorithm is described. It uses not only the integral image for the local mean values calculation, but also the second order integral image for the local variance calculation. Following the proposed approach the time of segmentation has been significantly reduced providing the possibility of its use in practice. The generalization of integral image representation, called 'k-order integral image' could be used for fast calculation of higher order local statistics. An example of algorithm for the segmentation of cells and Chlamydial inclusions on microscope images, containing the steps for color deconvolution and fast adaptive local binarization is presented.

Keywords: k-order integral image, fast algorithm, Niblack binarization, microscope image segmentation

DOI: 10.1134/S1054661816030020

INTRODUCTION

The problem of speeding up the local image statistics calculation is not new. In [1] a box-filtering technique was proposed for fast calculation of image local mean value. In [2] another idea of fast local mean value calculation was published, and later it was popularized under the name of integral image representation in [3].

In [4] the box-filtering technique was extended for fast calculation of local statistics of different orders for N-dimensional images. Box-filtering technique compared to integral image representation is less memory consuming, but is dependent on the size of local window which is intended to be used for local statistics calculation. If local statistics with different local window sizes have to be calculated for one image, the box-filtering technique seems not to be the best choice, regarding the calculations time. Absence of simple and fast algorithm for calculation of local standard deviation resulted, e.g., in attempts to replace local standard deviation in Niblack algorithm by local intensity mean deviation [5].

In this paper we continue to develop a simple and fast approach for local standard deviation calculation based on the expansion of integral image representation, initially presented in [6] under the name of 'integral squared image.

SECOND ORDER INTEGRAL IMAGE REPRESENTATION

The idea of Niblack method is to calculate the binarization threshold $T(x,y)$ by local mean $m(x,y)$ and standard deviation $s(x,y)$ of the neighbouring pixels intensity values:

$$T(x,y) = m(x,y) + k \cdot s(x,y), \tag{1}$$

where k is the coefficient, determined experimentally. For $k = 0$ only local average is used for calculation of the threshold value.

The significant size of the window often used for the calculation of local statistics leads to significant time loss for direct implementation of the Niblack algorithm and impossibility of its use in practice.

Local mean value in (1) could be effectively calculated using integral image representation. Integral image $I(x_0, y_0)$ is calculated from the initial image $f(x,y)$ as follows [3]:

$$I(x_0, y_0) = \sum_{x=1}^{x_0} \sum_{y=1}^{y_0} f(x,y). \tag{2}$$

With integral image the local mean value inside rectangle area, multiplied by the number of area pixels, can be calculated using only tree addition operations:

[1] This paper uses the materials of the report submitted at the 9th Open German-Russian Workshop on Pattern Recognition and Image Understanding, held in Koblenz, December 1–5, 2014 (OGRW-9-2014).

[2] The article is published in the original.

Received November 10, 2015

Table 1. Computational complexity of calculation of local standard deviation directly and using the proposed integral representation

Operation	Number of operations	
	by direct calculation	using integral representation I_2
Addition/ subtraction	$1 + (2N - 1)N_1N_2$	$1 + 3N_1N_2 - N_1 - N_2$
Mmultiplication/ squaring	$1 + (N + 1)N_1N_2$	$1 + 4N_1N_2$
Division	1	1
Square rooting	N_1N_2	N_1N_2

$$m(x_0, y_0) \cdot N = I(x_0 + \Delta x, y_0 + \Delta y) + I(x_0 - \Delta x - 1, y_0 - \Delta y - 1) - I(x_0 - \Delta x - 1, y_0 + \Delta y) - I(x_0 + \Delta x, y_0 - \Delta y - 1), \quad (3)$$

where $m(x_0, y_0)$—local mean value inside rectangle area with center coordinates (x_0, y_0), upper left and the lower right corners coordinates $(x_0 - \Delta x, y_0 - \Delta y)$ and $(x_0 + \Delta x, y_0 + \Delta y)$ respectively, $N = (2\Delta x + 1)(2\Delta y + 1)$—the number of rectangle area pixels.

The generalization of integral image representation is k-order integral image:

$$I_k(x_0, y_0) = \sum_{x=1}^{x_0} \sum_{y=1}^{y_0} [f(x, y)]^k. \quad (4)$$

The representation (4) can be effectively used for the calculation of k-order local statistics in the same manner as local mean value (3).

The main time loss during Niblack binarization is associated with the calculation of local standard deviation of image pixel intensities. To accelerate these computations a second order integral image representation $I_2(x_0, y_0)$ should be used:

$$I_2(x_0, y_0) = \sum_{x=1}^{x_0} \sum_{y=1}^{y_0} [f(x, y)]^2. \quad (5)$$

Considering the well-known equation in statistics where the variance is calculated as the difference between the mean square value and squared mean value, local variance $s^2(x, y)$, multiplied by $(N - 1)$, could be calculated using second order integral image representation (5) as follows:

$$s^2(x_0, y_0) \cdot (N - 1) = I_2(x_0 + \Delta x, y_0 + \Delta y) + I_2(x_0 - \Delta x - 1, y_0 - \Delta y - 1) - I_2(x_0 - \Delta x - 1, y_0 + \Delta y) - I_2(x_0 + \Delta x, y_0 - \Delta y - 1) - \frac{1}{N} \cdot [m(x_0, y_0) \cdot N]^2. \quad (6)$$

In Table 1 computation complexity of direct calculation of local standard deviation $s(x_0, y_0)$ for an image region of size $N_1 \times N_2$, and of its calculation using the proposed integral representation is given. It could be clearly seen that the second order integral image representation provides significant reduction in calculation complexity and independence on the local window area N.

To calculate the Niblack local binarization threshold for one image pixel by direct calculation of the local mean and standard deviation values inside local window with area N, approximately $(3N - 1)$ addition operations, $(N + 3)$ multiplication and squaring operations, and one square rooting are required. The use of integral representations (2) and (5) provides independence from local window area and reduction in the number of mathematical operations to about twelve additions, four multiplications, two squaring and one square rooting per each local threshold value calculation.

APPLICATION OF FAST NIBLACK BINARIZATION ALGORITHM FOR MICROSCOPE IMAGE SEGMENTATION

The proposed fast binarization algorithm was used for segmentation of cells and Chlamydial inclusions in fluorescence microscope images of cell cultures infected with Chlamydia. Cultural method is the "gold standard" of Chlamydial infection diagnostics, the reference method for assessing the effectiveness of antibiotic treatment, and the only method allowing analysis of Chlamydia resistance to antibiotics in the task of therapy selection and the development of new antibacterial agents. Actual research demonstrate the necessity of at least inclusions' sizes assessment [7, 8], which could be reliably done during automated image analysis.

The segmentation of the images under consideration is the most important stage of their analysis, and its aim is to allocate the areas of cells, Chlamydial inclusions and background. In the fluorescence images Chlamydial inclusions appear bright green on the reddish-brown background of stained cells. For the segmentation of color images the methods of thresholding, clustering, region growing, construction of physical models of imaging and some others are commonly used. Each of these methods can use different color models.

To select a color representation model for the considered images providing their easiest and highest quality segmentation, a quantitative analysis of color differences between Chlamydial inclusions, cells and background was carried out in channels of the following color spaces: RGB, CMY, HSV, Lab, YCbCr, YIQ.

The procedure of the specimen preparation includes staining with a fluorescent dye FITC and Evans blue. So, FITC and Evans blue concentration

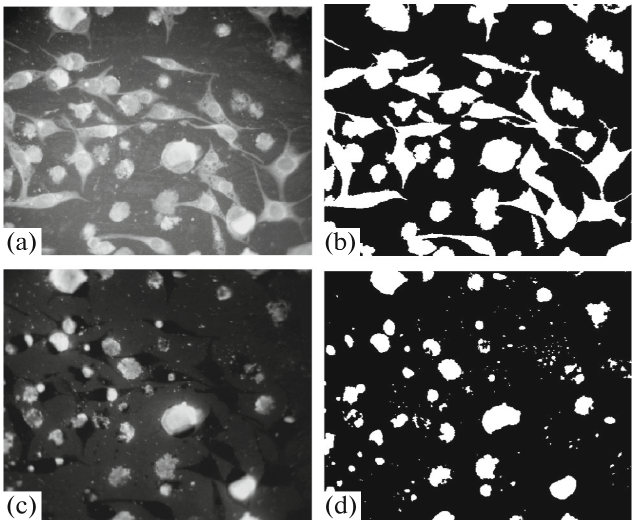

Fig. 1. The stages of image segmentation: (a) grayscale image in the color channel V of HSV color space, (b) the result of cells segmentation, (c) FITC concentration profile obtained by color deconvolution, (d) the result of Chlamydial inclusions segmentation.

profiles obtained as the result of color deconvolution method proposed in [9] and adapted for fluorescence microscopy in [6] were also used as the color channels.

Table 2. Results of different color channels comparison

Color space	Color channel	EER P_1	EER P_2
RGB	R	0.27	0.43
	G	0.48	0.11
	B	0.31	0.18
CMY	C	0.28	0.45
	M	0.37	0.17
	Y	0.44	0.16
HSV	H	0.36	0.17
	S	0.48	0.13
	V	0.25	0.34
Lab	L	0.28	0.17
	a	0.31	0.23
	b	0.32	0.16
YCbCr	Y	0.30	0.15
	Cb	0.39	0.11
	Cr	0.30	0.29
YIQ	Y	0.30	0.15
	I	0.29	0.40
	Q	0.30	0.14
The results of color deconvolution	FITC concentration profile	0.48	0.03
	Evans blue concentration profile	0.34	0.18

To quantify the comparison of color channels we used equal error rate (EER) of the pixel-wise threshold classification. On 20 microscopic images three classes of regions: Chlamydial inclusions, cells, and background—were labeled manually. EER was calculated for two classification problems: "cells and Chlamydial inclusions against the background" (P_1) and "Chlamydial inclusions against the cells and the background" (P_2). The results are shown in Table 2. The images used in this and subsequent studies were provided by the Laboratory of Chlamydial infections of Gamaleya R&D Institute of Epidemiology and Microbiology.

The FITC concentration profile allows to reduce significantly the EER of pixels belonging to Chlamydial inclusions threshold classification (to an average of 3%), but Evans blue concentration profile does not provide a similar result in the classification of pixels belonging to cells. This is consistent with the peculiarities of the specimens preparation, as Evans blue is often associated not only with cellular structures, but also with the extracellular matrix. Thus, for cells segmentation it is expedient to use channel V of the color space HSV, since it has a minimum EER value.

Despite the small EER value for pixels of Chlamydial inclusions classification, obtained for images used, often during the preparation of cell culture specimens the background fluorescence caused by the presence of a fluorescent dye not only in the microbiological organisms being studied, but in general on the entire surface of the preparation, is not completely suppressed. For this reason, for the segmentation of cells, as well as of Chlamydial inclusions in the selected color channels it is expedient to use local adaptive binarization methods, the most famous of which is the Niblack method.

The parameters of binarization algorithm that provide the best quality of segmentation were determined experimentally. The size of a square window used for the calculation of threshold value is about three times the average linear size of objects to be segmented. This equals to 700 px for the segmentation of cells and 500 px for the segmentation of Chlamydial inclusions for images captured with pixel size 0.14 μm (image size 3072×2304 px). Coefficient k in (1) is equal to -1 and 2 for the segmentation of cells and Chlamydial inclusions respectively. Image segmentation stages are shown in Fig. 1.

The use of second order integral image representation in Niblack algorithm provided the reduction of the average time of one image automatic analysis from few days (when direct calculation of the local image standard deviation was used) to 9 s (on computer with processor Intel Core i5, 2.27 GHz, RAM 4.00 Gb and 64-bit OS). This provides the possibility to use the proposed segmentation algorithm in practice.

The assessment of the proposed segmentation algorithm was performed on 50 images of Chlamydial inclusions with substantially different sizes, and 30 images of cell cultures each containing about 200 cells. These images were segmented both manually and automatically.

The absence of quantitative biomedical studies so far makes it impossible to set substantiated requirements for segmentation accuracy beforehand. The quality of the segmentation of cells and Chlamydial inclusions was assessed by the relative error of their areas determination. These errors for cells and Chlamydial inclusions did not exceed 5.2 and 3.0%, respectively. The proposed segmentation algorithm enables for the first time to carry out quantitative microscope studies of cells infected with Chlamydia (see, e.g., [8]), in which substantiated requirements for the segmentation error of the considered images will also be established, allowing to confirm adequacy of the proposed segmentation algorithm or to make a reasonable decision on the need of its improvement.

CONCLUSION

This study presents the k-order integral image representation, which is the generalization of well-known integral image. The use of the second order integral image provides fast variance (and standard deviation) calculation. Higher order local statistics can be further calculated by means of higher order integral image representations.

The method proposed was used to implement fast Niblack binarization algorithm for microscope image segmentation of cell cultures infected with Chlamydia. The expanded integral image approach provides the significant reduction in segmentation time which opens the possibility of practical use of the developed automatic segmentation algorithm.

REFERENCES

1. M. J. McDonnell, "Box-filtering techniques," Comput. Graph. Image Processing **17**, 65–70 (1981).
2. F. C. Crow, "Summed-area tables for texture mapping," Comput. Graph. **18** (3), 207–212 (1984).
3. P. Viola and M. J. Jones, "Robust real-time face detection," Int. J. Comput. Vision **57** (2), 137–154 (2004).
4. C. Sun, "Fast algorithm for local statistics calculation for N-dimensional images," Real-Time Imag. **7**, 519–527 (2001).
5. T. Romen Singh, Sudipta Roy, O. Imocha Singh, Tejmani Sinam, and Kh. Manglem Singh, "A new local adaptive thresholding technique in binarization," Int. J. Comput. Sci. **8** (6), 271–277 (2011).
6. O. A. Artyukhova and A. V. Samorodov, "Elaboration of an automatic segmentation algorithm for fluorescence microscopic images of cell cultures preparations for the problems of microbiology," Sci. Edu.: Electron. Sci. Tech. Publ., No. 6 (2013). http://technomag.edu.ru/doc/574140.html [in Russian].
7. N. A. Zigangirova, E. S. Zayakin, L. N. Kapotina, E. A. Kost, L. V. Didenko, D. Y. Davydova, J. P. Rumyanceva, and A. L. Gintsburg, "Development of chlamydial type III secretion system inhibitors for suppression of acute and chronic forms of Chlamydial infection," Acta Naturae **4** (2 (13)), 87–97 (2012).
8. E. A. Kost, "The effect of an inhibitor of type III secretion system on the development of experimental infection caused by Chlamydia trachomatis," PhD Thesis (Moscow, 2014) [in Russian].
9. A. C. Ruifrok and D. A. Johnston, "Quantification of histochemical staining by color deconvolution," Anal. Quant. Cytol. Histol. **23** (4), 291–299 (2001).

Ol'ga Aleksandrovna Samorodova. Born. 1986. Graduated from Bauman Moscow State Technical University in 2009. Received candidate (PhD) degree in engineering sciences at Bauman Moscow State Technical University in 2013. Associate professor, Chair for Biomedical Technical Systems, Bauman Moscow State Technical University. Fields of research priorities: automated microscopy of biomedical preparations, biomedical image processing and analysis. biomedical statistics, computer-aided decision-making systems in medicine. Author of 60 scientific publications, including 13 journal papers.

Andrei Vladimorovich Samorodov. Born 1975. Graduated from Bauman Moscow State Technical University in 1999. Received candidate (PhD) degree in engineering sciences at Bauman Moscow State Technical University in 2002. Head of the Chair for Biomedical Technical Systems, Bauman Moscow State Technical University. Fields of research priorities: methods and algorithms of pattern recognition and multiclassification, automated microscopy of biomedical preparations.computer-aided decision-making systems in medicine, methods and technique for biomedical images and signals recognition Author of more than 200 scientific publications (including 1 collective monograph and more than 30 journal papers).