Tomasz MARKIEWICZ[1,2]

Warsaw University of Technology (1), Military Institute of Medicine (2)

Comparison of single and sequential thresholding applied to segmentation of a cell cluster at Ki-67 staining

Abstract. *The paper presents the comparison of different approaches to segmentation of a cell cluster in microscopic images at Ki-67 stained tissues. In automatic quantitative analysis appropriate cell segmentation is the most important task. The main problem is a cell bordering or overlapping, which makes different cell clusters very difficult to separate. Different types of cells based on all nuclei can be recognized. Among them there exists bordering with lighted area between nuclei, the bordering without that area but with significant narrow space between nuclei and in the bad case without even this narrow space. In the first case the optimal selection of threshold value, also in the sequential thresholding manner, can be a successful approach. For nuclei bordering with significant narrow border the appropriate distance transformation and the watershed transform is the most common method to separate these cells. The advantages and disadvantages of different approaches to segmentation will be presented and discussed in the study.*

Streszczenie. *W pracy przedstawione jest porównanie różnych metod segmentacji klastrów komórek w obrazach mikroskopowych przy odczynie Ki-67. Jest to obok klasyfikacji komórek główny problem występujący w automatycznej analizie ilościowej preparatów tkanek. Klastry komórek są obiektami obrazu bardzo trudnymi do właściwej segmentacji i związane są ze stykaniem lub nakładaniem się profili jąder komórek. Można odróżnić kilka rodzajów klastrów: z jaśniejszym obszarem pomiędzy jądrami komórek, bez takiego obszaru ale z przewężeniem powierzchni oraz w najtrudniejszym przypadku bez przejaśnień i przewężeń. W pierwszym z wymienionych przypadków skutecznym podejściem jest odpowiedni wybór progu do progowania, w tym podejście oparte na progowaniu sekwencyjnym. W drugim przypadku dobre efekty daje metoda działów wodnych w połączeniu z właściwą transformacją odległościową. Skuteczność rozważanych w pracy podejść segmentacji klastrów komórek zostanie zaprezentowana i skomentowana.* ***(Porównanie metod pojedynczego i sekwencyjnego progowania w zastosowaniu do segmentacji klastrów komórek w obrazach mikroskopowych przy odczynie Ki-67)***

Keywords: thresholding, segmentation, image processing.
Słowa kluczowe: progowanie, segmentacja, przetwarzanie obrazów.

Introduction

The automatic quantitative evaluation of the medical images is a very intensive research area in recent years. In the case of the pathological microscopic images the main task is the appropriate recognition of the cells. This step is necessary for quantitative evaluation of many antibodies, marked with the specific staining. The appropriate segmentation of the images is a very important step in the image analysis process [1,2].

In practice, different space configuration of the cells can be observed in the image. Some cells are distributed separately without bordering or overlapping with other ones. These cells can be recognized more easy that the cells which are overlapped to others or have very close neighbours. The space configuration of the cells is called the cluster and produces many problems for any segmentation algorithms. The watershed transform [1] is the most common method to separate these cells. However, the efficiency of the watershed segmentation depends mainly on the thresholding transformation, being the basic process in cell recognition. The appropriate thresholding has significant influence on the cells segmentation.

This paper presents different approaches to the cell cluster segmentation. The single thresholding approaches with different measures of the optimal threshold value [3] and sequential thresholding [2] will be presented and compared.

Methods of thresholding

In the presented research the thresholding is the most common tool for recognition of the cell areas from the background of the image. Its practical realization can be performed in the different manners. The most common approach is based on the single thresholding transformation with one threshold value. The pixels of the values less than threshold are recognized as background and the pixels with higher values represent the cell areas. The main problems of that approach is the selection of the optimal threshold value, giving the most adequate segmentation without loosing any cells. There exist different measures applied at the optimal threshold value determination.

1. Between class variances

This method is based on statistic of pixel intensity values and was introduced by Otsu [4] to determine an optimal threshold to segment the image into nearly uniform regions. According to Otsu the mean, probability of both classes and between class variance (BCV) should be calculated

(1) $\sigma_B^2 = \omega_1 \omega_2 (\mu_2 - \mu_1)^2$

were ω_1, ω_2 represent the class probabilities and μ_2, μ_1 are mean values. The optimum threshold value corresponds to maximum of BCV.

There exist also some modifications (Kurita et al [5], Reddi et al [6]) that offer some advantages. In the study [5] the authors proposed another criteria function $Q(t)$

(2) $$Q(t) = \log(\sigma_W(t)) - \sum_{j=1}^{2} p_j \log(p_j)$$

In this definition σ_W is the within-class variance, t - threshold and p_j - the probability of the *j*th class. This modified version of BCV was proposed for unbalanced populations (thanks to the second term).

2. Minimum error thresholding

Minimum error thresholding (MET) is based on the minimalization of the Kullback information distance [7]

(3) $$J = \sum_{i=1}^{N} p(i) \log\left[\frac{p(i)}{f(i)}\right]$$

where *p(i)* is the probability of *i* class and *f* is the unknown mixture distribution. For Gaussian distribution it is transformed to the minimization of the objective function defined in the form

(4) $$H(t) = p_1(t)\log(\frac{\sigma_1}{p_1(t)}) + p_2(t)\log(\frac{\sigma_2}{p_2(t)}) + const.$$

called minimum error function.

3. Entropy-based thresholding

The entropy-based thresholding [3] bases only on the global information derived from the gray-level histogram. The effective version of threshold selection used normalized gray-level frequencies p_i observed in the image pixels, formed the gray-level distribution. The two probability distribution for object and background with threshold value *s* can be write as

(5) $$H_b(s) = \ln(P_s) + \frac{H_s}{P_s},$$
$$H_w(s) = \ln(1 - P_s) + \frac{H_T - H_s}{1 - P_s}$$

where $H_s = -\sum_{i=1}^{s} p_i \ln p_i$ is the entropy. The maximizing of the sum of functions (5) gives the optimal threshold value according to entropy-based approach.

4. Sequential thresholding technique

The opposite approach to single threshold methods is the sequential thresholding technique. It is based on the sequential changes of the threshold value from minimum to maximum. In any iteration only the objects with area in the specified range are recognized as the cells separated from the background and other cells. To this manner, most cells can be recognized without necessity to find the one optimal threshold value. This feature can be also very helpful in the cell cluster segmentation. The details of this method were presented in the paper [2].

Problem of the cell cluster segmentation

In the segmentation process, based on single or sequential thresholding, some cells can't be efficiently segmented. In the single thresholding approach the cell clusters can be effects the higher concentration of the bordered cells with the cytoplasms darker stained than normally or cell overlapping. In the sequential thresholding the first type of clusters will be eliminated in the iterative thresholding process and the clusters will be created only at overlapping of the cells. These clusters require additional tools for appropriate segmentation.

For cell cluster segmentation very popular is the watershed method. This method can extract any cells which have significant narrow space between cells. However, when such border line isn't observed, the cells couldn't be segmented and the cluster is painted as a big cells.

Results

For the evaluation of optimality of the single thresholding with different measures, sequential thresholding and other cluster thresholding approaches the microscopic images of the neuroblastoma tumour as Ki-67 staining were used. This tumour is characterized by the small, round cells with frequent cell bordering or overlapping in the acquired image. Fig. 1. presents the sample field of view with intensive cell concentration.

The segmentation process based only on a single thresholding technique with different measures of the optimal threshold value and sequential thresholding have produced significant differences in the recognition of cell areas, that is shown in the fig. 2.

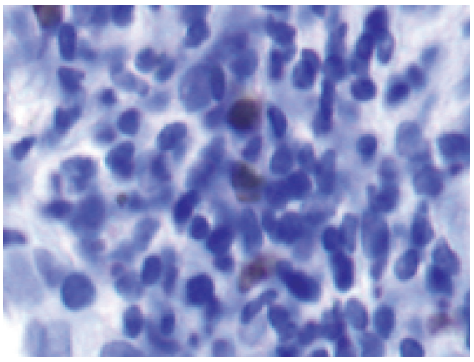

Fig.1. The neuroblastoma image (x400, Ki-67, part of view field)

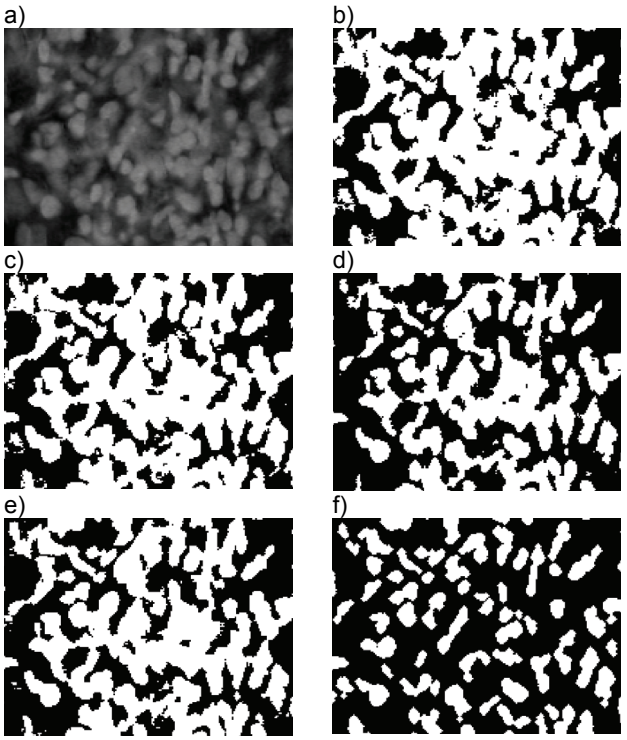

Fig.2. The cell area recognition in image in gray scale a) based on: b) BCV, c) Kurita, d) minimum error, e) entropy, f) sequential thresholding

We observe that using measures of the optimal threshold value produces the different results. For the pixel value intensity from 0 to 148, the optimal threshold values indicated by different methods were as following: BCV – 58, Kurita – 61, minimum error – 73 and entropy – 66. Any method of single thresholding approaches produces high number of cell clusters (too high for efficient segmentation of the cells). Significantly better result was obtained for sequential thresholding approach, with high rate of cell separation. However, there is still exist many cell clusters, but with smaller number of the cells forming the cluster. This result is most appropriate to perform the watershed transform to next step of segmentation. The result of the second step of cell recognition is presented in the fig. 3.

Most cells from the clusters with significant narrow line between bordered cells were well segmented. Only few clusters were not segmented. That kind of results were not possible at application of the single threshold approach. The detailed results for whole image are presented in table 1.

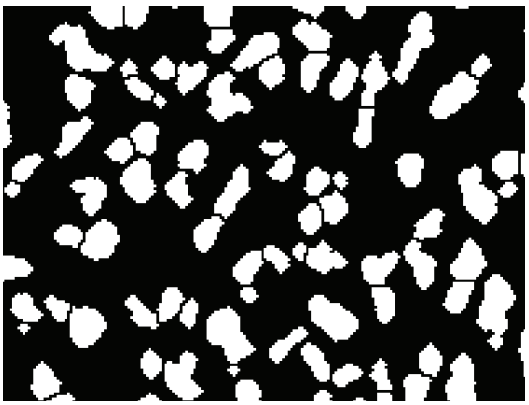

Fig.3. The results of the second step segmentation based on watershed transform.

Table 1. The results of cell segmentation

Method	Number of the recognized cells	Cells area [%]
BCV	731	41.5
Kurita	755	38.3
Minimum error	763	26.6
Entropy	790	33.3
Sequential thresholding	988	27.3

Comparing the single thresholding approaches the best seems to be entropy, 790 cells were recognized at its application. However, in comparison to the sequential thresholding, this result was much worse. The obtained 988 cells at sequential thresholding are very close to the manual result (1030 cells). The cell counting errors of the single thresholding approaches are higher than 23%.

Very interesting results concern cells area. The BCV and Kurita measures give too high cells area in the cell segmentation process. The minimum error criteria returns smaller cells area (26.6% of the image), but entropy produces the better result (from single thresholding methods) with cells area (33.3%). The sequential thresholding gives cells area slightly higher than minimum error method. Probably the real cells area is closes to 30% of the image area, but it is hard to measure. We estimate that the relative error of measure of cells area is lower than 10%.

For evaluation of the thresholding methods, we selected the set of 24 representative images. The results depicted in the Table 2 shown that comparing the single thresholding approaches the best seems to be Kurita modification with average 866 cells recognized in the one field of view. However, the entropy-based approach gives more appropriate cells area segmentation.

Table 2. The results of cell segmentation for N=24 images

Method	Mean number of the recognized cells	Mean cells area in image[%]
BCV	832	46.9
Kurita	866	43.6
Minimum error	739	22.4
Entropy	817	32.4
Sequential thresholding	1056	30.0

Conclusions

The presented results indicate that it isn't possible to obtain the appropriate segmentation of the cell clusters based on single thresholding approaches, also with support of the additional segmentation. The differences of the colour intensity in the practical specimen staining, mainly in the form of higher intensity of the cytoplasms in the bordered and overlapped cells give no changes to get good results of segmentation at one optimal threshold value. For the quantitative evaluation of the image, based on cell segmentation and recognition, the multithresholding approaches are suggested. The presented results have shown that sequential thresholding can reduce the error from 25% at single thresholding to less than 5%. However, it should be noted, that sequential thresholding can result in some reduction of the cells area. The degree of this reduction is related to the cell bordering and overlapping and it is very hard for exact evaluation. Based on the presented results the relative error of cells area estimation is less than 10% for the image with high density of the cell clustering.

Acknowledgments

This work was supported by the European Union in the framework of the European Social Fund through the Warsaw University of Technology Development Programme.

REFERENCES

[1] Soille P., Morphological Image Analysis, Principles and Applications. Springer, Berlin 2003

[2] Markiewicz T., Wisniewski P., Osowski S., Patera J., Kozlowski W., Koktysz R., Comparative analysis of the methods for accurate recognition of cells in the nuclei staining of the Ki-67 in neuroblastoma and ER/PR status staining in breast cancer. *Anal Quant Cytol*, 31 (2009), No. 1, 49-62

[3] Demirkaya O., Asyali M. H., Sahoo P. K., Image Processing with MATLAB: Applications in Medicine and Biology. CRC 2009

[4] Otsu N., A threshold selection method from grey-level histograms. *IEEE Trans. on Systems, Man, and Cybernetics* (1979), No. 9, 62-66

[5] Kurita T, Otsu N, Adbelmalek N: Maximum likelihood thresholding based on population mixture model. *Pattern Recogn*, 25 (1992), 1231:1240

[6] Reddi SS, Rudin FS, Keshavan HR: An optimal multiple threshold scheme for image segmentation. *IEEE Trans Syst Man Cybernetics*, SMC-14 (1984), 661-665

[7] Kittler J, Illingworth J: Minimum error thresholding. *Patters Recogn*, 19 (1986), 41-47

***Authors**: dr hab. inż. Tomasz Markiewicz, Politechnika Warszawska, Instytut Elektrotechniki Teoretycznej i Systemów Informacyjno-Pomiarowych, ul. Koszykowa 75, 00-662 Warszawa, oraz Wojskowy Instytut Medyczny, Zakład Patomorfologii, ul. Szaserów 128, 04-141 Warszawa, E-mail: markiewt@iem.pw.edu.pl.*