A new method for segmentation of microscopic blood cell images by using histogram based automatic thresholding

RESUL COTELI[a,*], TURKER TUNCER[b], ENGIN AVCI[c], MEHMET USTUNDAG[d], ESIN DOGANTEKIN[e], AKIF DOGANTEKIN[f]

[a]*Firat University, Faculty, Energy Systems Engineering, 23119, Elazig, Turkey*
[b]*Firat University, Technology Faculty, Software Engineering, 23119, Elazig, Turkey*
[c]*Firat University, Technology Faculty, Digital Forensic Engineering, 23119, Elazig, Turkey*
[d]*Bıngol University, architecture-Engineering Faculty, Electrıcal-electronics Engineering, 12000, Bıngol, Turkey*
[e]*Ministry of Health, 27100, Gaziantep, Turkey*
[f]*Emek Hospital, 27100, Gaziantep, Turkey*

In this study, a novel thresholding algorithm based on histogram shape (TAHS) is proposed for determining the multiple thresholding. In the proposed thresholding method, histogram is divided into smaller windows with same size. In the proposed method, the biggest difference between pixel intensities is found. The proposed method is applied to blood cell images and is compared with Otsu method for a better validation. Advantage of the proposed method is that one or more thresholding points can be obtained. Therefore, objects with different spatial feature can be detected successfully. The experimental studies show that TAHS method gives satisfactory thresholding results.

(Received February 4, 2017; accepted August 9, 2017)

Keywords: Image segmentation, Automatic image thresholding method, Microscopic blood cell images, Multiple thresholding method, Histogram shape

1. Introduction

Image segmentation is one of the most important image-processing task. It is widely used in many image-processing applications such as object recognition, classification. A variety of segmentation methods is proposed in literature. Image thresholding is one of the most important step in image processing. Finding correct thresholding value is the main problem in the image thresholding. In addition, thresholding value should be automatically determined.

Automatic thresholding methods can be divided into the six groups based on information the algorithm manipulates [1]. In the shape-based automatic thresholding methods: the peaks, valleys and curvatures of the smoothed histogram are analysed and one or more thresholding values are calculated properly. In cluste-ring - based automatic thresholding methods, grey-level pixels are classified to two classes as background and foreground (object). In entropy - based automatic thresholding methods, the entropy values of the foreground and background areas and the cross-entropy between the original and binary image are used. In object attribute-based automatic thresholding methods, similarity between the grey-level and the binary images is determined. In spatial-based automatic thresholding methods, higher-order probability distribution and/or correlation between pixels are used. In local-based automatic thresholding methods, original image is divided into several sub regions, and various thresholds are chosen for each sub region.

Otsu's thresholding method has been widely used. In this method, discriminant analysis is utilized to find the maximum margin of classes [2]. The automatic thresholding can be used to extract objects from the background. In image thresholding literature, many studies have been carried out on different applications of image thresholding such as cell images, thermal images, document image, etc. [1]. The output of the thresholding operation is a binary image whose grey level of 0 (black) will indicate a pixel belonging to a print, legend, drawing, or target and a grey level of 1 (white) will indicate the background.

The global and local thresholding techniques have been proposed in literature. In the global thresholding techniques, one thresholding value is applied to the entire image. In contrast, local thresholding methods uses different thresholding values for different regions of the image [3]. Ref. [4] proposed for measuring the "goodness" of a thresholded image, one based on a busyness criterion and the other based on a discrepancy or error criterion. In Ref. [5], several methods of the global thresholding are compared. In Ref. [6], different adaptive binarization techniques are compared with global methods. Ref. [7] presented results of 19 segmentation methods. In Ref. [8], performance of nine thresholding algorithms are presented comparatively. Ref. [9] proposed a clustering-based image thresholding method to reduce a grey level image to a binary image. Ref. [10] introduced Sobel operator to all three planes in the RGB space. A local adaptive thresholding method based on contrast of image is proposed in Ref. [11]. Ref. [12] presented an automatic thresholding method based on entropy of the grey-level histogram. Ref. [13] applied the Sobel operator to each component of the HSI space. Ref. [14] presented an entropy based thresholding method. In Ref. [15], a local-variance-based automatic thresholding method is presented. Ref. [16] suggested a novel multiresolution image segmentation algorithm using Markov random fields. Ref. [17] proposes a novel automatic thresholding method based on sliding rectangular window over the grey level image. Ref. [18] introduced an unsupervised method for colour image segmentation. Ref. [19] proposes an automatic iterative thresholding technique for determining an optimum threshold value. Ref. [20] proposed random walk based image segmentation method. Ref. [21] introduced a multi-scale binarization, framework working with the adaptive methods directly. Ref. [22] suggested an image segmentation method based on the modified edge-following scheme. Ref. [23] proposed an automatic histogram threshold approach based on a fuzziness measure. Ref. [24] suggested a novel technique for the validation of document binarization algorithms. Ref. [25] proposed a new method based on balanced histogram for automatic image thresholding. Ref. [26] suggested a new image segmentation method based on grey graph cut. Ref. [27] suggested a new combining image segmentation technique based on multilevel threshold and data fusion techniques to get a more reliable and accurate segmentation effect. Ref. [28] presented an adaptive image de-noising technique based on thresholding and obtained experimental results are given in terms of peak-to-signal ratio (PSNR). The wavelet based thresholding methods are also given in Ref. [29]. Ref. [30] proposes a method for modelling the background using per-pixel time-adaptive Gaussian mixtures in the combined input space of pixel colour and pixel neighbourhood. Ref. [31] analyses different image segmentation techniques and obtained results are given. Ref. [32] presents deep fully convolutional neural network architecture for semantic pixel-wise segmentation. Ref. [33] suggests parallel segmentation algorithm with regional growth and support vector machine (SVM). Ref. [34] introduces a factorization-based method that efficiently segments textured images. In this study, local spectral histograms are use as features.

In this study, a novel method based on histogram shape is proposed for image thresholding. In the suggested method, histogram is divided into smaller windows with same size and standard deviation for each window is calculated. Then, the biggest difference between pixel intensities is found. If pixel intensity difference is greater than standard deviation of the window, this point is accepted as a valley point. To find midpoint of the valley, a constant value is added to pixel intensity difference. This constant value is inversely proportional to number of window. Then, automatic thresholding process is realized by using these valley points. The proposed method is applied to blood cell images and the proposed method is compared with Otsu method for a better validation. When compared with previous study, advantage of the proposed method is that one or more thresholding points are obtained. Therefore, objects with different spatial feature can be detected successfully.

2. Image thresholding

The simplest way for image thresholding is to convert to grey level and take a threshold point. The mathematical formulation for a basic image thresholding can be given as follows.

$$\begin{aligned} &If\ f(x,y) > T, then\ \ f(x,y) = 0 \\ &If\ f(x,y) < T, then\ \ f(x,y) = 1 \end{aligned} \tag{1}$$

In Eq. (1), T is thresholding value. In addition, mean of the *f(x, y)* is grey level of the image. According to Eq.1, *f(x, y)* equals to zero for bigger value than *T* thresholding value whereas remaining *f(x, y)* values to unity. This method is the most primitive thresholding method. Different thresholding point should be selected for each image in this method. Therefore, thresholding value cannot be automatically chosen which is one the most important drawbacks for such thresholding methods. To eliminate drawbacks of manual thresholding methods, automatic methods should be used. There have been many studies on selection of the thresholding point automatically. Histogram based methods are one of automatic thresholding method. The peak and valley points should be determined in automatic thresholding method based on the histogram. Thresholding method to be used should be carefully selected depending on application because thresholding methods cannot give satisfactory result for each of different images. The image processing algorithms have been widely used because of many advantages such as good detection of adherent or overlap blood cells.

3. Previous thresholding methods

The automatic image thresholding methods can be divided into six main groups [1]. These methods are summarized as below.

- **Automatic Image Thresholding Methods based on Histogram Shape:** In this method, thresholding points of image are automatically calculated by using histogram shape. These points can be obtained by means of convex or peak-and-valley points in the histogram of image [1].

- **Automatic Image Thresholding Methods based on Clustering in Measure Space:** These thresholding methods can be divided into two groups. The first group aims for image thresholding by making several statistical analysis whereas second group aims separation of object from background. Otsu method is well known among clustering based thresholding methods and it has been widely used due to its simple structure and high thresholding performance. In addition, clustering based thresholding methods such as iterative, minimum error and fuzzy clustering have been utilized in automatic thresholding [1].

- **Automatic Image Thresholding Methods based on Entropy:** In such thresholding methods, the entropy of the distribution of the grey levels is calculated. If entropy of thresholded image is maximum, it means maximum information transfer. Different entropy based thresholding methods have been proposed in literature. More information that is detailed can be founded in Ref. [1].

- **Object Attribute-Based Automatic Thresholding Methods:** In these thresholding methods, thresholding value is determined by using some attribute quality or similarity measure between the original image and the binary image.

- **Automatic Image Thresholding Methods based on Spatial Properties:** Such thresholding methods use spatial information of object and background pixels. These features can include context probabilities, co-occurrence probabilities, and local linear dependence models of pixels.

- **The Automatic Image Thresholding Methods based on Local Properties:** A threshold that is calculated at each pixel characterizes this class of algorithms. The value of the threshold depends upon some local statistics like range, variance, and surface fitting parameters or their logical. In such methods, local adaptation is computed for different thresholding value of each pixel depending on local properties of image. This method is depended on regional location. In this method, some parameters are used to adjust the threshold value as local.

4. The proposed method

This paper proposes a novel thresholding method based on histogram shape (TAHS). The basic principle of the TAHS is based on automatically finding valley points in histogram. Histogram is windowed into smaller regions. The TAHS algorithm divides the histogram of image into equal parts according to precision of image. Each separated part has same size and different statistical attributes. To find valley points, standard deviation for each window is calculated. Then, the biggest difference between pixel intensities is found. If pixel intensity difference is greater than standard deviation of the window, this point is accepted as a valley point. Finding midpoint of the valley is another important subject in the proposed method. For this aim, a constant value is added to pixel intensity difference. This constant value is inversely proportional to number of window. Then, automatic thresholding process is realized by using these valley points. TAHS is based on statistical features of the related image. The realization steps of the suggested method are summarized in below.

Step 1: Obtain the histogram of related image.
Step 2: Divide into smaller parts obtained histogram. In this study, histogram is separated into eight parts.
Step 3: Calculate standard deviation of each part of histogram.
Step 4: Find the maximum difference between the elements for each of eight parts.
Step 5: Determine valley points.
Step 6: Calculate thresholding values by shifting valley points until a selected constant value.

The mathematical formulation of the TAHS algorithm can be given as below. Maximum difference (*MAD*) for histogram with '*n*' regions can be computed as follows:

$$MAD = \sum_{i=1}^{256/n} \sum_{j=i+1}^{256/n} \max(|x_i - x_i|) \tag{2}$$

In Eq.2; $i = \{1,2,....,256/n\}$ and $j = \{2,3,....,256/n\}$ x_i and x_j are *i*th and *j*th elements of the histogram, respectively. Mean of histogram can be found by means of Eq. (3).

$$x' = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{3}$$

Where x' is arithmetic mean of histogram. Standard deviation of histogram can be also calculated as follows:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - x')} \tag{4}$$

Where σ is standard deviation of histogram. Threshold value of any window can be determined depending on maximum difference and standard deviation as Eq. (5).

$$If\ MAD > \sigma,\ \ Threshold = j + k \tag{5}$$

Selection of the constant value is carried out by means of pixel values. For example, if there is an 8-bit grey level image, this image has 256 pixels values changing between 0 and 255. Square root of this value is added to valley value. The proposed algorithm is tested on a set of 30 different blood cell images. The images are 8-bit grey level images. These images have a size of 512 × 512. In Figs. 1-6, grey scale, thresholded images and histograms of images are given the results of TAHS and Otsu methods for different blood cell images.

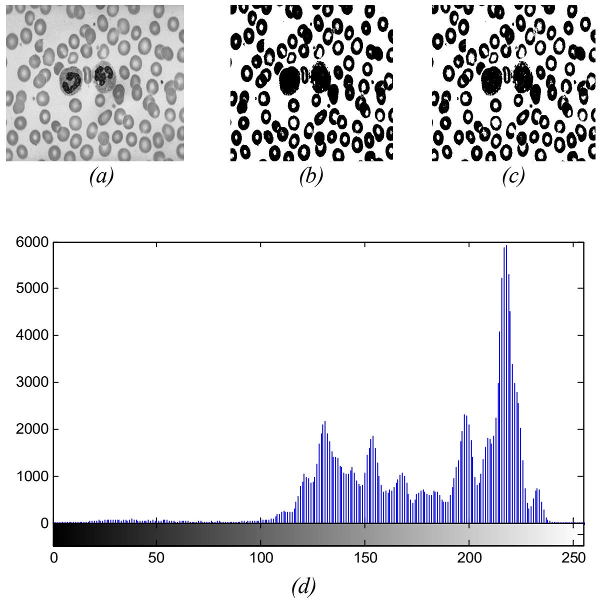


Fig. 1. a) Original image, b) Thresholded image by using Otsu method, b) Thresholded image by using TAHS method, d) Histogram of orginal image

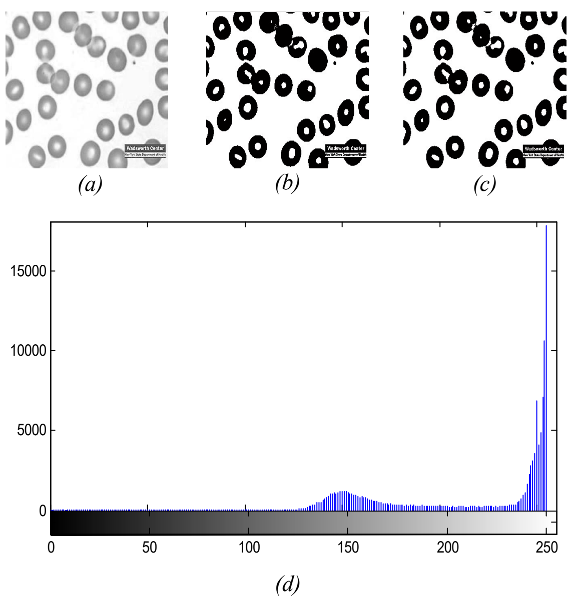


Fig. 2. a) K2 original image, b) Thresholded image by using Otsu method, b) Thresholded image by using TAHS method, d) Histogram of K2 image

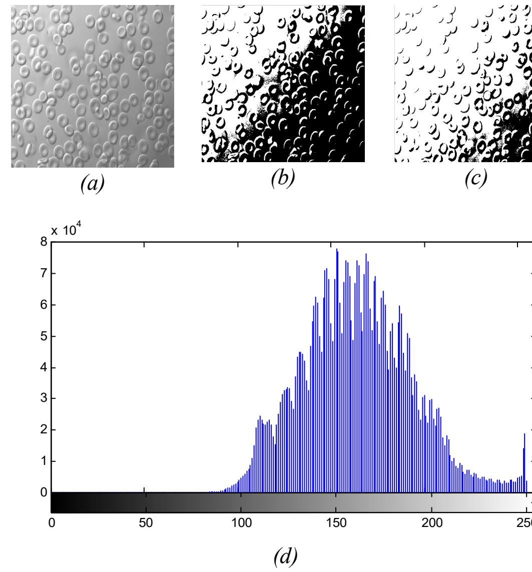


Fig. 3. a) K3 original image, b) Thresholded image by using Otsu method, b) Thresholded image by using TAHS method, d) Histogram of K3 image

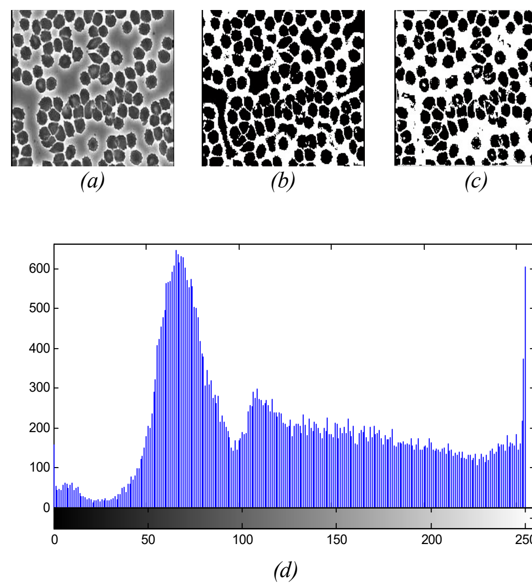
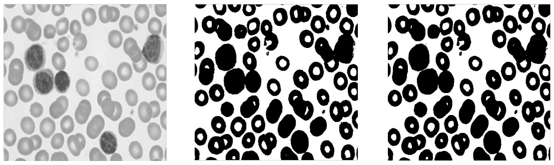
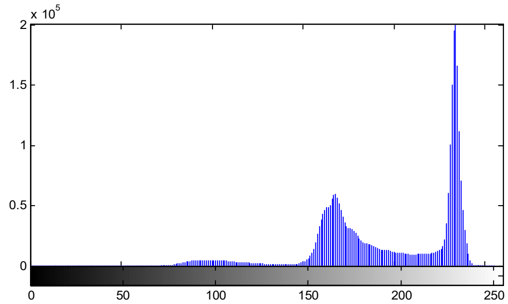


Fig. 4. a) K4 original image, b) Thresholded image by using Otsu method, b) Thresholded image by using TAHS method, d) Histogram of K4 image

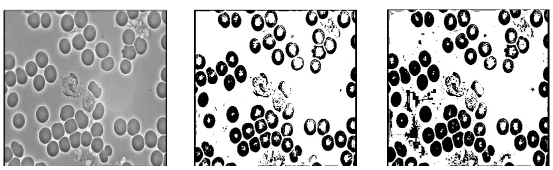


(a) *(b)* *(c)*

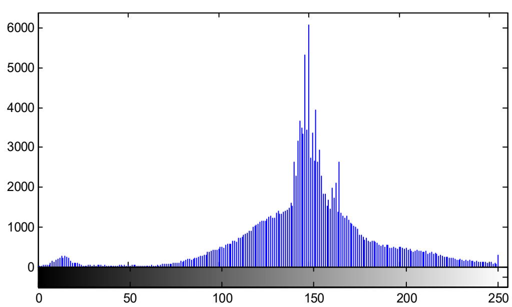


(d)

Fig. 5. a) K5 original image, b) Thresholded image by using Otsu method, b) Thresholded image by using TAHS method, d) Histogram of K5 image



(a) *(b)* *(c)*



(d)

Fig. 6. a) K6 original image, b) Thresholded image by using Otsu method, b) Thresholded image by using TAHS method, d) Histogram of K6 image

The obtained results show that the proposed method extracts objects in blood cell images more clearly than Otsu method. The obtained thresholding values for sample blood cell images are given in Table 1. It is concluded that the proposed method can exactly find the valley point. This fact increases thresholding performance of the suggested method.

Table 1. Obtained thresholding values from TAHS algorithm for example images

Image	Th-1	Th-2	Th-3	Th-4	Th-5	Th-6	Th-7
K1	20	39	68	119	133	167	200
K2	3	36	85	110	139	167	198
K3	3	55	81	107	131	165	195
K4	56	81	92	110	-	-	-
K5	74	103	147	158	190	-	-
K6	9	35	81	113	144	167	207

Thresholding processing can be carried out more sensitively if histogram is divided into more parts. This state can be explained by means of Eq. (6).

$$S = n + \frac{n(n-1)}{2} \tag{6}$$

Where S and n is the number of the thresholding images and points, respectively. Since there are total of n thresholding points, total of n images are obtained by using the thresholding with single point. However, $n(n-1)/2$ expresses the thresholding images with multiple points. The best automatic thresholding results can be obtained by optimizing total of S thresholding images.

5. Conclusion

This paper proposed an automatic multi-level thresholding method based on histogram statistics for grey-level images. The proposed method determines multi threshold values depending on standard deviation inside of a group. If maximum difference in the group is greater than standard deviation, that point is assigned to valley point. Thus, thresholding values are chosen by using statistical attributes of the histogram, and maximizing the between-class standard deviation. Further, the proposed method has a simple structure and it can be easily realized. The method is applied to blood cell images tested over set on a set of 30 different blood cell images. Otsu method has been widely used for automatic thresholding. Therefore, obtained results of the suggested method are compared with Otsu method. The results shows that the proposed method can detected blood cell images more clearly than Otsu method.

Acknowledgment

This work is supported by Scientific and Technological Research Council of Turkey (TUBITAK) with project numbered 113E155.

References

[1] B. Sankur, M. Sezgin, Pattern Recognition **34**(2), 1573 (2001).

[2] N. Otsu, IEEE Transactions on Systems Man Cybernet, **9**(1), 62 (1979).
[3] G. Leedham, et al., Proceedings of the Seventh International Conference on Document Analysis and Recognition **2**, 859 (2003).
[4] J. S. Weszka, A. Rosenfeld, IEEE Transactions on Systems, Man and Cybernetics **8**(8), 622 (1978).
[5] S. U. Lee, S. Y. Chung, R. H. Park, Computer Vision, Graphics and Image Processing **52**(2), 171 (1990).
[6] P. W. Palumbo, P. Swaminathan, S. N. Srihari, Document image binarization: Evaluation of algorithms. In Applications if Digital Image Processing IX. International Society for Optics and Photonics 278 (1986).
[7] O. D. Trier, A. K. Jain, IEEE Transactions on Pattern Analysis and Machine Intelligence **17**(12), 1191 (1995).
[8] P. K. Sahoo, S. A. K. C. Soltani, A. K. Wong, Computer vision, Graphics and Image Processing **41**(2), 233 (1988).
[9] N. Otsu, Automatica **11**(285-296), 23 (1975).
[10] M. Hedley, H. Yan, Journal of Electronic Imaging **1**(4), 374 (1992).
[11] J. Bernsen, In International Conference on Pattern Recognition 1251 (1986).
[12] G. Johannsen, J. Bille, In ICPR **82**, 140 (1982).
[13] T. Carron, P. Lambert, Saturation and Intensity, (1994).
[14] J. N. Kapur, P. K. Sahoo, A. K. Wong, Computer Vision, Graphics and Image Processing **29**(3), 273 (1985).
[15] J. Sauvola, M. Pietikäinen, Pattern Recognition **33**(2), 225 (2000).
[16] J. Liu, Y. Yang, IEEE Transactions on Pattern Analysis and Machine Intelligence **16**(7), 689 (1994).
[17] W. Niblack, Strandberg Publishing Company, (1985).
[18] Z. Cao, X. Zhang, X. Mei, In Intelligent Information Technology Application, 2008. IITA'08. Second International Symposium on, 2, 99-103, (2008).
[19] T. W. Ridler, S. Calvard, IEEE Transactions on Systems, Man and Cybernetics **8**(8), 630 (1978).
[20] Y. H. Lan, Y. Zhang, C. H. Li, X. F. Zhao, In Computational Intelligence and Industrial Applications **1**, 207 (2009).
[21] R. F. Moghaddam, M. Cheriet, Pattern Recognition, **43**(6), 2186 (2010).
[22] C. Yuan Been, Oscal T. C. Chen, EURASIP Journal on Image and Video Processing **2009**, 140492 (2009).
[23] N. V. Lopes, P. A. Mogadouro do Couto, H. Bustince, P. Melo-Pinto, IEEE Transactions on Image Processing **19**(1), 199 (2010).
[24] P. Stathis, E. Kavallieratou, N. Papamarkos, J. UCS **14**(18), 3011 (2008).
[25] A. Dos Anjos, H. Shahbazkia, In Biosignals **2**, 70 (2008).
[26] M. Ma, J. He, H. Guo, H. Tian, In Computational Science and Optimization (CSO), 2010 Third International Joint Conference **1**, 477 (2010).
[27] R. Harrabi, E. B. Braiek, EURASIP Journal on Image and Video Processing **2012**, 11 (2012).
[28] H. Om, M. Biswas, WSEAS Trans. on Signal Processing **10**, (2014).
[29] S. Sutha, E. J. Leavline, D. A. A. G. Singh, WSEAS Trans. Signal Process **9**, 203 (2013).
[30] K. Sundaraj, WSEAS Transactions on Computers, **7**(10), 1762 (2008).
[31] Umaa S. Mageswari, M. Sridevi, C. Mala, Procedia Engineering **64**(1), 36 (2013).
[32] V. Badrinarayanan, K. Alex, C. Roberto (2015). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. arXiv preprint arXiv:1511.00561.
[33] S. Yongqian, L. Xi, International Journal of Signal Processing, Image Processing and Pattern Recognition **8**(2), 139 (2015).
[34] J. Yuan, W. Deliang, M. C. Anil, IEEE Transactions on Image Processing **24**(11), 3488 (2015).

*Corresponding author: rcoteli@gmail.com