STUDIA INFORMATICA 2016
Volume 37 Number 3B (126)

Damian BORYS, Magdalena SZEREMET, Krzysztof PSIUK-MAKSYMOWICZ,
Mariusz FRACKIEWICZ
Silesian University of Technology, Institute of Automatic Control
Ziemowit OSTROWSKI
Silesian University of Technology, Institute of Thermal Technology

STUDY AND ADJUSTMENT OF DERMOSCOPIC IMAGE PROCESSING ALGORITHMS FOR LESION BORDER DETECTION[1]

Summary. The goal of our work was an initial preprocessing of dermoscopic images towards accurate lesion border detection. Four algorithms were proposed and analyzed: MS – algorithm using mean shift clustering, HE – algorithm using histogram equalization, TTH – algorithm using the top-hat transform, PCA – algorithm using principal component analysis. Those algorithms were tested on PH2 images database that contains 200 dermoscopic images, each with a mask of the lesion. Those algorithms were optimized using lesion mask from database and Jaccard index as a measure of similarity of both sets. Simple statistical analysis of indexes was used to compare proposed algorithms in term of their accuracy.

Keywords: dermoscopy, contouring, image preprocessing

BADANIE I OPTYMALIZACJA ALGORYTMÓW PRZETWARZANIA OBRAZÓW DERMOSKOPOWYCH W CELU ZNALEZIENIA KONTURU ZNAMIENIA

Streszczenie. W artykule poruszono problem wstępnego przetwarzania obrazów dermatoskopowych w celu znalezienia konturu znamienia. Zaproponowano

[1]This work was supported by the Institute of Automatic Control under Grant No BK-213/RAu1/2016/3. Calculations were carried out using the computer cluster Ziemowit (http://www.ziemowit.hpc.polsl.pl) in the Computational Biology and Bioinformatics Laboratory of the Biotechnology Centre at the Silesian University of Technology, funded by the Silesian BIO-FARMA project No. POIG.02.01.00-00-166/08 and expanded in the POIG.02.03.01-00-040/13 project. This paper is an extended version of an International Conference on Computer Systems in Medicine and Health 2016 referred article.

i porównano cztery algorytmy: MS – wykorzystujący klasteryzację 'mean shift', HE – wykorzystujący wyrównywanie histogramu, TTH – wykorzystujący transformację 'top-hat', PCA – wykorzystujący metodę analizy głównych składowych. Algorytmy przetestowano z wykorzystaniem obrazów z bazy PH2, zawierającej 200 obrazów wraz z obrysem ręcznym, a ich parametry dobrano optymalizując indeks Jaccarda. Proste statystyki wyników pozwoliły na porównanie proponowanych algorytmów.

Słowa kluczowe: dermatoskopia, konturowanie, przetwarzanie wstępne

1. Introduction

The most frequent malignant skin cancer is melanoma [1]. This type of cancer can be characterized by an aggressive growth and early, numerous metastases. When metastases occur, they are tough in pharmacological treatment and have rather a poor prognosis. 90% of all melanomas is localized in the skin region, most frequently in the area of pigmental change, but can also occur in the potentially healthy skin. Due to poor prognosis in the two last stages of melanoma evolution, early diagnosis, followed by the immediate surgery procedure of nevus removal, is crucial and allows to decrease mortality of patients with melanomas significantly. Simple visual inspection of the skin moles with a dermatoscopy is the easiest way to perform a diagnostic study. However, a video dermatoscopy is now in a frequent usage [2]. The latter allows to acquire and store a digital image of selected lesion, and this can be an input for a computer-aided diagnosis system. In the case of the disease risk, identification a surgery procedure is planned for removal of the suspicious mole with a safe margin. Histopathological examination is then performed to characterize the probe.

Image processing is an important part of the dermatoscopy diagnosis of malignant melanoma. It allows to diagnose the changes in nevus quickly and allows to determine the probability of melanoma. Some features that can be specified in dermatoscopy image are helpful in melanoma recognition. For example, the "ABCD" (or "ABCDE") [3] rule can be useful, where A is asymmetry, B – border, C – color, D – differential structures and sometimes E is added as evolution or elevation. All of those features require a lesion contouring. To calculate all needed features accurately, as well as to determine lesion border the preprocessing step has to be done. With this step, it is required to prepare dermoscopic images to further processing by removing all unwanted contaminations in the image like hairs, air bubbles (as a result of immersion liquid usage) [4], etc.

Hairs registered during dermatoscopy, partially shade the main region of interest, that's why it needs special treatment [5,6]. In the literature one can find a lot of works, that uses some processing methods, for example, special filtration methods, thresholding and morphological operations [7]. In other, authors use morphological top-hat operation or filtration together with morphological operations and lesion filling step [8,9]. In this work, we have proposed four algorithms for border detection. Each of them consists of preprocessing step with hair removal as a substantial step of preprocessing of the dermoscopic images.

2. Materials and Methods

All algorithms were tested on real dermoscopic images from an open-access database PH2 [10] which is one of the results of the ADDI project. This database consists of 200 images in RGB with, manually created by a specialist, region of interest as a binary mask in a separate image. The database also contains all necessary information about each lesion (clinical diagnosis, asymmetry, color, pigment network, etc.), which could be useful for build or test classification systems. Examples of images with their mask of the lesion are shown in Fig. 1.

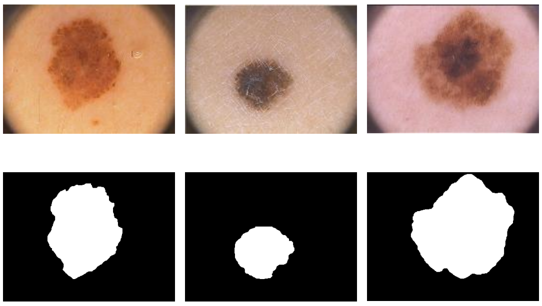

Fig. 1. Images from PH2 [10] database (upper row) with manually created lesion (lower row)
Rys. 1. Przykładowe obrazy znamion z bazy PH2 [10] (górny rząd) wraz z ręcznie obrysowanymi konturami (dolny rząd)

2.1. Algorithms description

Four algorithms have been proposed, implemented and tested using PH2 database. Proposed algorithms have been created iteratively using the probe and error method. All algorithms are multi-step procedures, and each step is presented in Figs 2 – 5. Each step of the algorithm has its parameters, like gamma correction level or size of the mask, etc. Due to that fact, one optimal parameter set has to be chosen. For each algorithm, we have created a discrete space of parameters, depending on the algorithm, and we have performed the contouring procedure. Then, the results of each iteration have been compared with the mask from the database using the Jaccard index (range from 0 to 1 – where 1 is a full similarity of sets). The parameter set for the algorithm with the best results of the Jaccard index was taken to comparison with other algorithms.

The first algorithm was the one that uses white top-hat transform (TTH). Within this following steps have been used: Gauss filtration of a color input image, RGB to grayscale conversion, gamma correction, negative Laplacian filtering added to image, white top-hat transform added to image (for the small elements enhancement), gauss filtration, histogram equalization, Otsu binarization, morphologic operations (dilation, clean, fill, bridge, majority, closing), black frame removal (to get rid of the black border from the image), image inversion, black and white gaps removal. Fig. 2 presents the algorithm flow with an example of the input image.

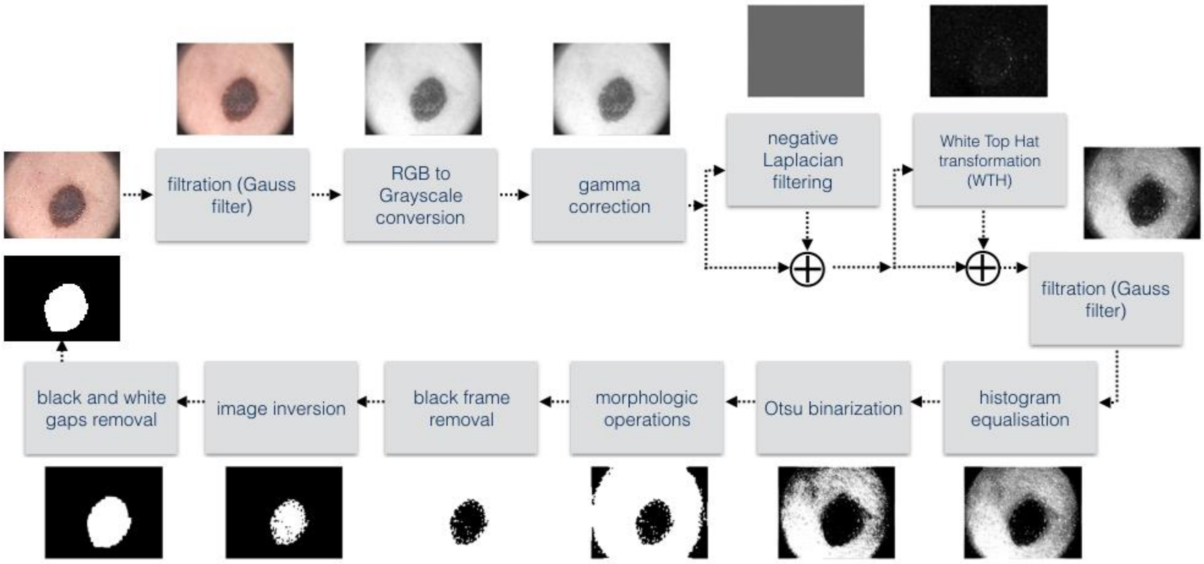

Fig. 2. Algorithm TTH : schema of the algorithm with visualization of each step using one image from PH2 database

Rys. 2. Algorytm TTH: schemat algorytmu wraz z wizualizacją przejścia każdego kroku na przykładzie obrazu z bazy PH2

The second algorithm was the one that uses mean-shift clustering technique (MS). Within this following steps have been used: RGB to grayscale conversion, gamma correction, mean-shift clustering technique, Canny edge detector added to image to enhance the edges in the image after Otsu binarization, mean-shift clustering technique for the contrast enhancement, black frame removal, morphologic operations (clean, fill, erosion, dilation), image inversion and white and black gaps removal. Fig. 3 presents the algorithm flow with an example of the input image.

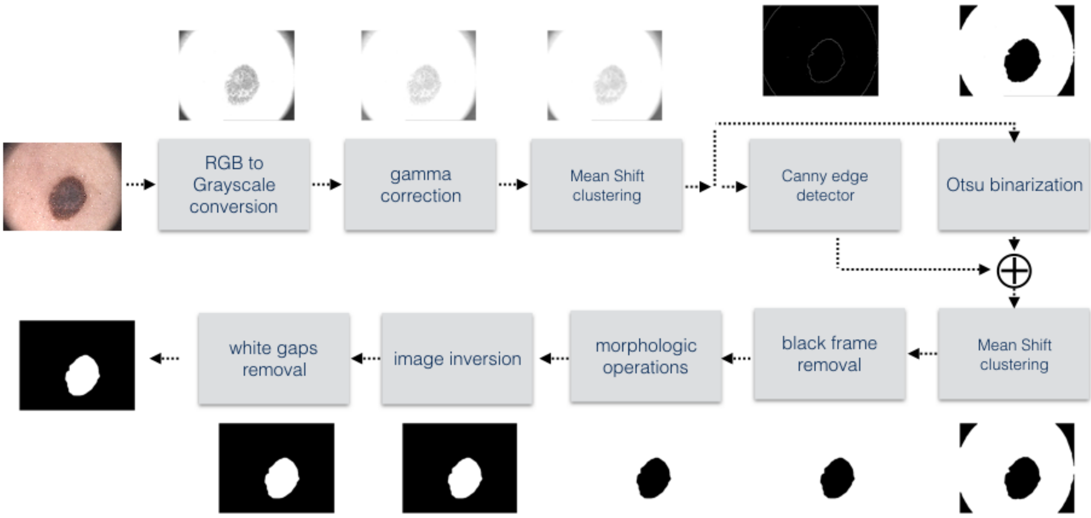

Fig. 3. Algorithm MS : schema of the algorithm with visualization of each step using one image from PH2 database

Rys. 3. Algorytm MS: schemat algorytmu wraz z wizualizacją przejścia każdego kroku na przykładzie obrazu z bazy PH2

The third algorithm was the one that uses histogram equalization technique (HE). Within this following steps have been used: RGB to grayscale conversion, gamma correction, contrast enhancement using histogram equalization technique, Canny edge detector added to image to enhance the edges in the image, Otsu binarization, black frame removal, morphologic operations (close, clean, fill, bridge, erosion, dilation), image inversion and removal of white gaps. Fig. 4 presents the algorithm flow with an example of the input image.

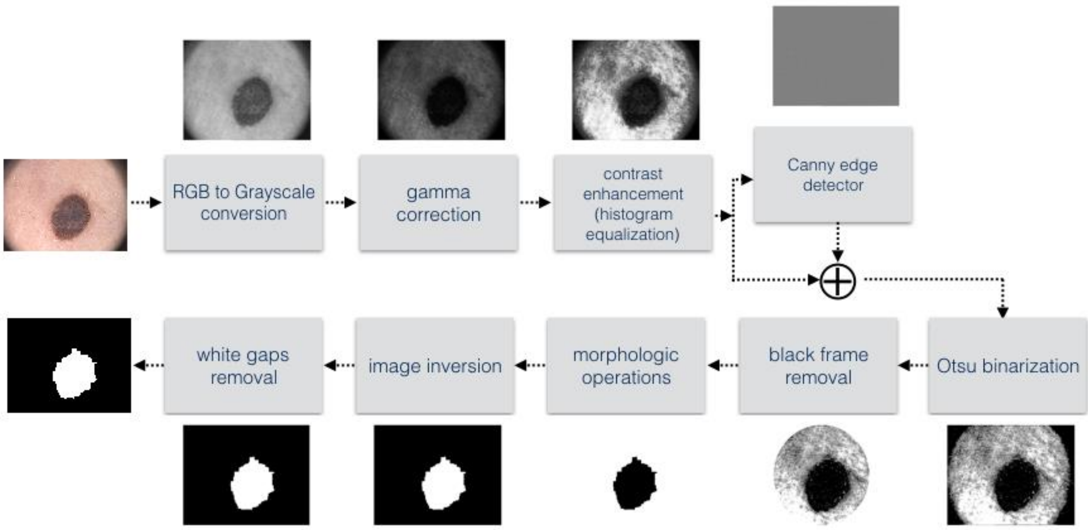

Fig. 4. Algorithm HE: schema of the algorithm with visualization of each step using one image from PH2 database

Rys. 4. Algorytm HE: schemat algorytmu wraz z wizualizacją przejścia każdego kroku na przykładzie obrazu z bazy PH2

The fourth algorithm was the one that uses Principal Component Analysis (PCA) technique for Grayscale conversion. Within this following steps have been used: RGB to grayscale conversion using PCA, gamma correction, noise reduction using Wiener filter, Otsu binarization, black frame removal, Canny edge detector added to image to enhance the edges in the image, image inversion, morphologic operations (fill, clean, erosion, dilation) and removal of white gaps. Fig. 5 presents the algorithm flow with an example of the input image.

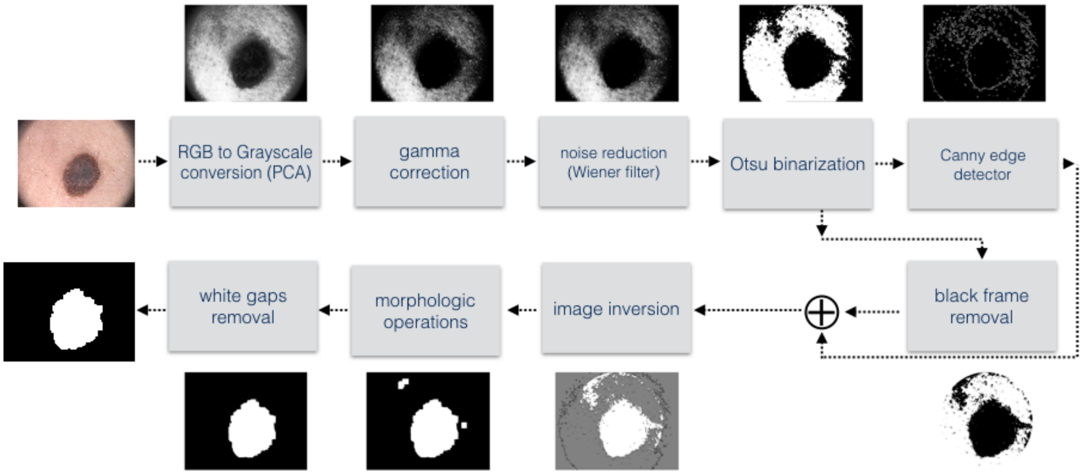

Fig. 5. Algorithm PCA: schema of the algorithm with visualization of each step using one image from PH2 database

Rys. 5. Algorytm PCA: schemat algorytmu wraz z wizualizacją przejścia każdego kroku na przykładzie obrazu z bazy PH2

2.1.1. *Choosing an optimal parameter set for each method*

For each algorithm we have tested multi dimensional parameter space. Depending on the algorithm it was : for TTH – Gauss filter kernel size, gamma correction coefficient, filter kernel size, color to grayscale method; for MS - color to grayscale method, gamma correction coefficient, kernel size; for HE - color to grayscale method, gamma correction coefficient; for PCA – Wiener filter kernel size, gamma correction coefficient. In case of color to grayscale method we have tested three methods: RGB2GS_1 - weighted sum of the *R*, *G*, and *B* components: (0.2989 * R + 0.5870 * G + 0.1140 * B); RGB2GS_2 – normalized B channel after Gaussian blurring; RGB2GS_3 – V channel after RGB to HSV conversion. After obtaining results for all parameter combination for each method we have calculated elementary statistics for each image. For the best Jaccard index values we have found the parameters combination that was the most frequent in each method. Those parameters have been accepted to the final comparison of all method.

2.1.2. Comparison of algorithms

Having an optimal parameter set for each algorithm results have been investigated and compared using the Jaccard index elementary statistics. Outliers have been detected using box plot technique and the mean value as the range of the index was calculated.

3. Results

Optimal sets of parameters have been found for each of the method. Tables 1-8 presents those parameters sets for each method. Figures 6-9 shows histograms and box plots of the Jaccard index. RGB to grayscale were : RGB2GS_1 – weighted mean of RGB channels, RGB2GS_2 – B channel, RGB2GS_3 – V channel in HSV color space.

Table 1

The optimal sets of parameters for the HE method

Set of parameters No.	number of occurrences	RGB2GRAY	gamma
13	13	RGB2GS_2	gamma=1
15	13	RGB2GS_2	gamma=3

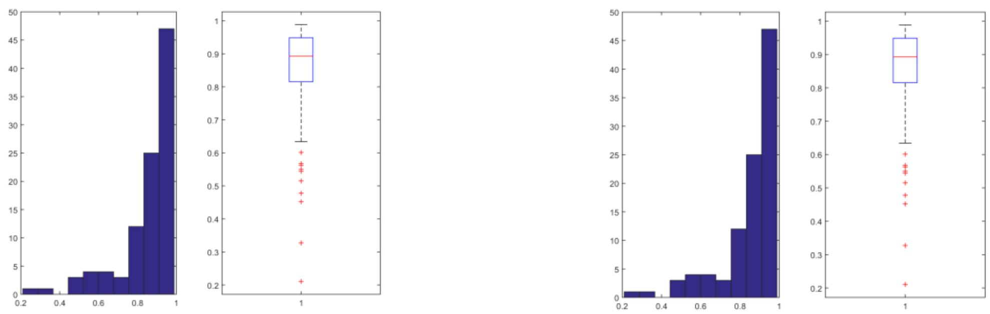

Fig. 6. Histogram and boxplot of the Jaccard index for HE method and a set of parameters No. 13 & 15

Rys. 6. Histogram oraz boxplot indeksów Jaccarda dla metody HE i zestawu parametrów 13 & 15

Table 2

The Jaccard index statistics of the optimal set of parameters for HE method

Set of parameters No.	Minimum	Maximum	Average	Median	Range
13	0.210	0.988	0.849	0.893	0.778
15	0.210	0.988	0.849	0.893	0.778

Table 3

The optimal sets of parameters for the MS method

Set of parameters No.	number of occurrences	RGB2GRAY	gamma	MeanShift mask size
55	37	RGB2GS_3	gamma=0.25	7x7

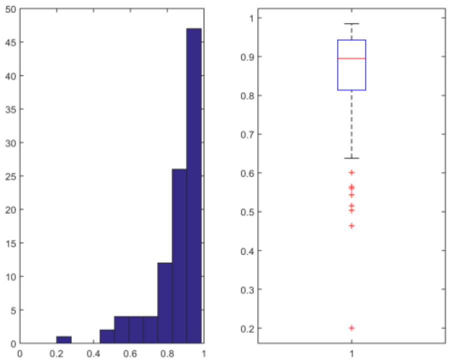

Fig. 7. Histogram and boxplot values of Jaccard index for MS method and a set of parameters No. 55

Rys. 7. Histogram oraz boxplot indeksów Jaccarda dla metody MS i zestawu parametrów 55

Table 4

The Jaccard index statistics of the optimal set of parameters for MS method

Set of parameters No.	Minimum	Maximum	Average	Median	Range
55	0.199	0.985	0.855	0.895	0.785

Table 5

The optimal sets of parameters for the TTH method

Set of parameters No.	number of occurrences	Gauss filter size	RGB2GRAY	gamma	Mask size
181	23	3x3	RGB2GS_1	gamma=1	7x7
187	23	3x3	RGB2GS_2	gamma=1	7x7
190	23	3x3	RGB2GS_3	gamma=1	7x7

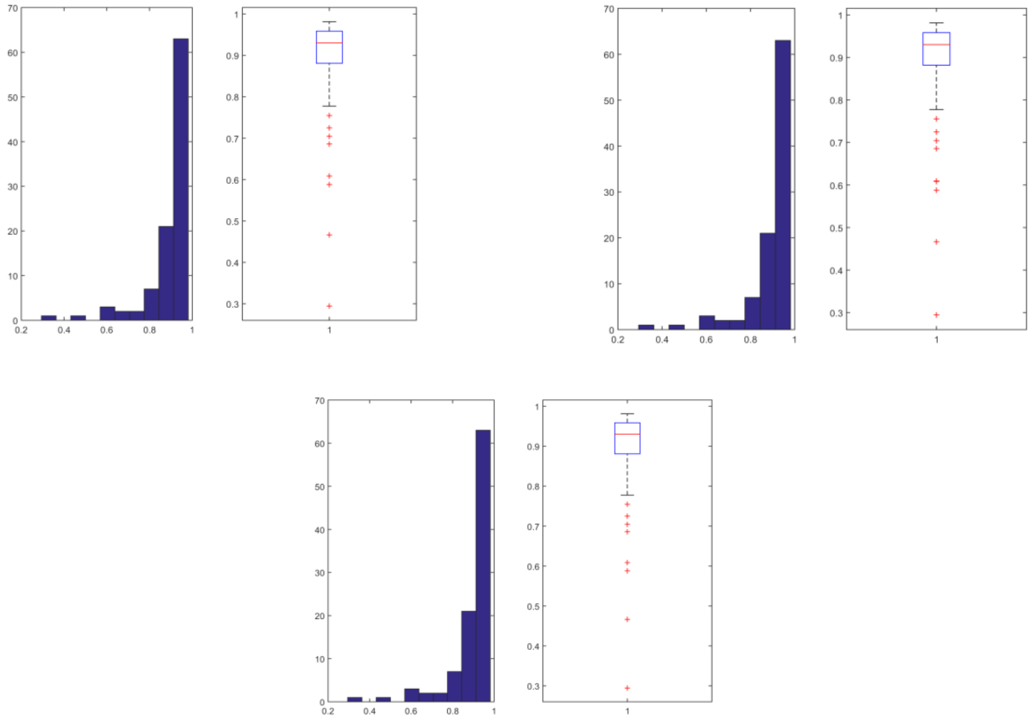

Fig. 8. Histogram and boxplot of Jaccard index for TTH method and a set of parameters No. 181, 187 & 190

Rys. 8. Histogram oraz boxplot indeksów Jaccarda dla metody TTH i zestawu parametrów 181, 187 & 190

Table 6

The Jaccard index statistics of the optimal set of parameters for TTH method

Set of parameters No.	Minimum	Maximum	Average	Median	Range
181	0.294	0.981	0.894	0.930	0.687
187	0.294	0.981	0.894	0.930	0.687
190	0.294	0.981	0.894	0.930	0.687

Table 7

The optimal sets of parameters for the PCA method

Set of parameters No.	number of occurrences	gamma	Wiener filter mask size
3	33	gamma=1	3x3

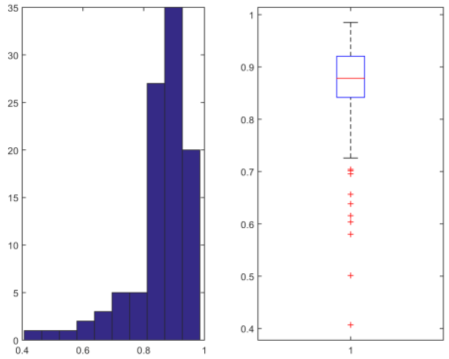

Fig. 9. Histogram and boxplot values of Jaccard index for PCA method and a set of parameters No. 3
Rys. 9. Histogram oraz boxplot indeksów Jaccarda dla metody PCA i zestawu parametrów 3

Table 8

The Jaccard index statistics of the optimal set of parameters for PCA method

Set of parameters No.	Minimum	Maximum	Average	Median	Range
3	0.406	0.985	0.856	0.878	0.578

Figures 10-13 presents result of boundary finding of each method with differential image between calculated mask and mask from the PH2 database. In Fig. 14 all results for two test images were gathered.

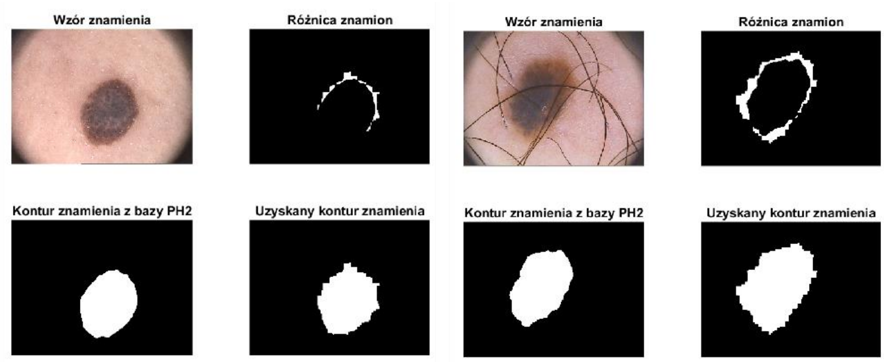

Fig. 10. Result of border finding using HE method, test image 1 and 2
Rys. 10. Wynik znajdowania konturu metodą HE dla obrazów testowych 1 i 2

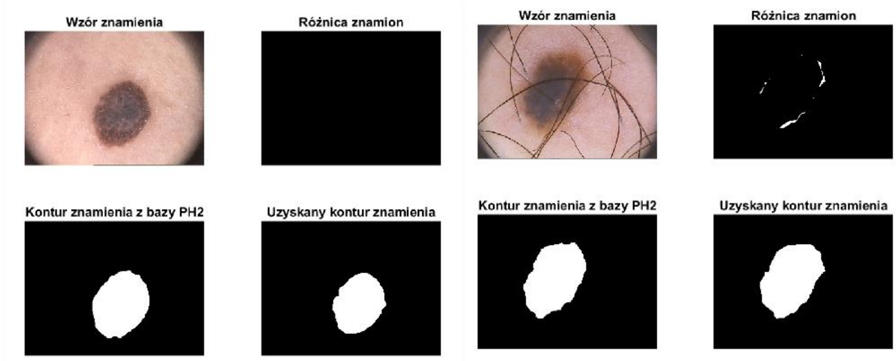

Fig. 11. Result of border finding using MS method, test image 1 and 2
Rys. 11. Wynik znajdowania konturu metodą MS dla obrazów testowych 1 i 2

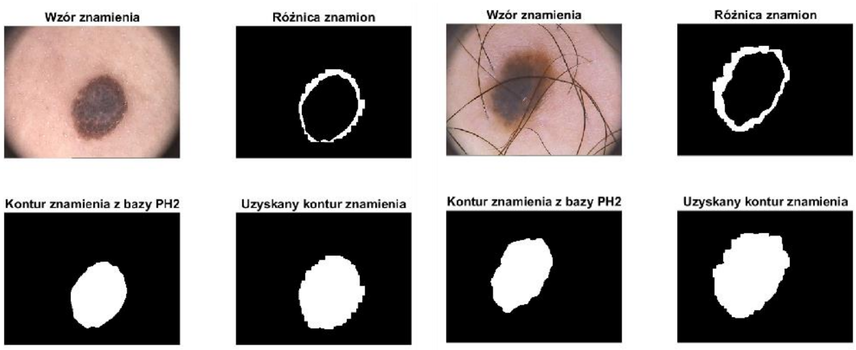

Fig. 12. Result of border finding using TTH method, test image 1 and 2
Rys. 12. Wynik znajdowania konturu metodą TTH dla obrazów testowych 1 i 2

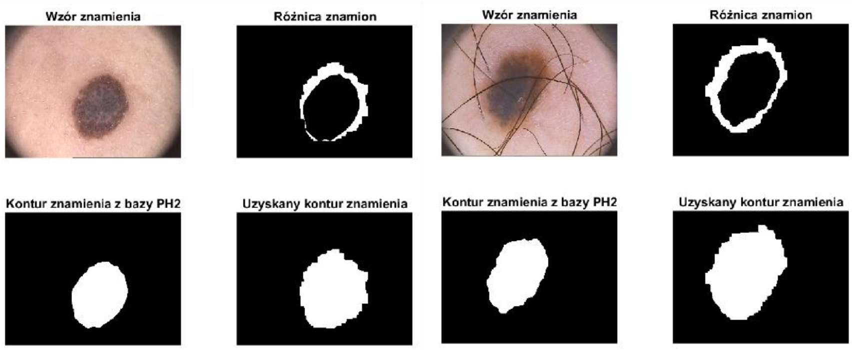

Fig. 13. Result of border finding using PCA method, test image 1 and 2
Rys. 13. Wynik znajdowania konturu metodą PCA dla obrazów testowych 1 i 2

4. Discussion and conclusions

In this paper we proposed and compared four methods for preprocessing and contouring dermatoscopy images. These methods allow to obtain lesion masks which are very similar to the expert assessment. The parameters required for these methods were found based on Jaccard index and were unchanged for the entire set of test images. The best average Jaccard index was obtained for the TTH method (comparison of the results in Table 2, Table 4, Table 6 and Table 8). Each of the presented methods has some outlier results (difficult cases of segmentation) what can be seen on boxplot on Fig. 6-9. After removing outliers Jaccard index interval is 0.60-1.00 for RH and MS methods, 0.75-1.00 for TTH and 0.70-1.00 for PCA. TTH method is statistically the best but there are cases where other methods give better results as shown in Fig. 10–13 where the smallest difference between calculated mask and mask from the PH2 database occurs for MS method.

Further analysis of the quality of the results requires comparison with the proposals of other authors which it is often difficult, mainly due to absence of detailed descriptions.

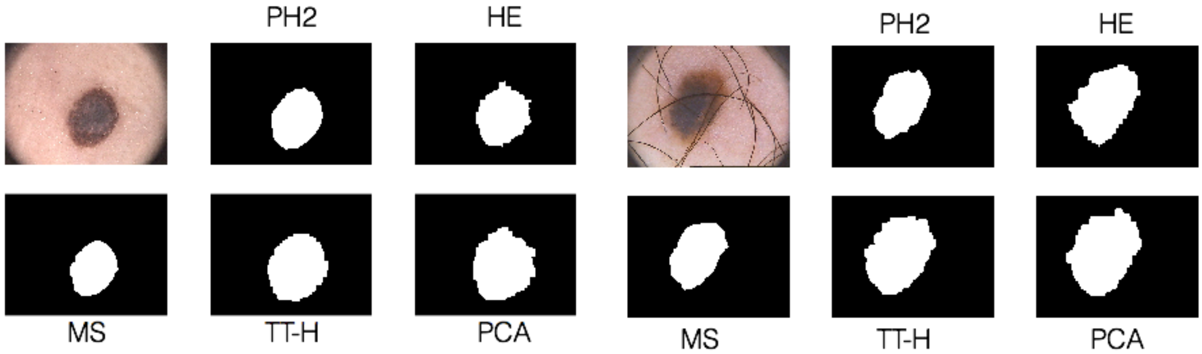

Fig. 14. Result of border finding, all methods, test image 1 and 2
Rys. 14. Wynik znajdowania konturu, wszystkie metody dla obrazów testowych 1 i 2

BIBLIOGRAPHY

1. Wojciechowska U., Didkowska J.: Zachorowania i zgony na nowotwory złośliwe w Polsce. Krajowy Rejestr Nowotworów, Centrum Onkologii – Instytut im. Marii Skłodowskiej-Curie, http://onkologia.org.pl/raporty/, accessed 30/04/2016.
2. Soyer H.P., Argenziano G., Hofmann-Wellenhof R., Zalaudek I.: Dermoscopy: The Essentials, 2nd ed. Elsevier Saunders, 2011.
3. Johr R.H.: Dermoscopy: Alternative melanocytic algorithms. The ABCD rule of dermatoscopy, Menzies scoring method, and 7-point checklist. Clin. Dermatol. 20, 2002, p. 240÷247.
4. Tasli L., Oguz O.: The role of various immersion liquids at digital dermoscopy in structural analysis. Indian J Dermatol Venereol Leprol. 77(1):110, 2011.
5. Lee T., Ng V., Gallagher R., Coldman A., McLean D.: DullRazor: A software approach to hair removal from images. Computers in Biology and Medicine 27, 1997, p. 533÷554.

6. Borys D., Kowalska P., Frackiewicz M., Ostrowski Z.: A Simple Hair Removal Algorithm from Dermoscopic Images. Bioinformatics and Biomedical Engineering, Vol. 9043 of the series Lecture Notes in Computer Science, 2015, p. 262÷273.
7. Nguyen N.H., Lee T.K., Atkins M.S.: Segmentation of light and dark hair in dermoscopic images: a hybrid approach using a universal kernel. In: Dawant B.M., Haynor D.R. (eds.) Medical Imaging 2010: Image Processing. Proc. of SPIE, vol. 7623, 2010, p. 76234N.
8. Abbas Q., Celebi M.E., Garcia I.F.: Hair removal methods: A comparative study for dermoscopy images. Biomedical Signal Processing and Control 6, 2011, p. 395÷404.
9. Abbas Q., Garcia I.F., Celebi M.E., Ahmad W.: A Feature-Preserving Hair Removal Algorithm for Dermoscopy Images. Skin Research and Technology 19, 2013, p. e27÷e36.
10. Mendoca T., et al.: PH2 – A dermoscopic image database for research and benchmarking. In: 35th Annual International Conference of the IEEE EMBS, Osaka, Japan, July 3-7, 2013.

Omówienie

Dermatoskopia jest podstawową techniką w diagnostyce złośliwego czerniaka skóry. W artykule poruszono problem wstępnego przetwarzania obrazów dermatoskopowych w celu znalezienia konturu znamienia. Obrazy znamion często zawierają poza badanym fragmentem skóry również pęcherzyki powietrza oraz włosy, które utrudniają ich przetwarzanie. Zaproponowano, zaimplementowano i przetestowano cztery algorytmy konturowania obrazów dermatoskopowych: MS – oparty na technice klasteryzacji 'mean-shift', HE – wykorzystujący wyrównywanie histogramu, TTH – wykorzystujący transformację 'top-hat', PCA – oparty na konwersji do skali szarości metodą analizy głównych składowych. Algorytmy przetestowano z wykorzystaniem obrazów z bazy PH2, zawierającej 200 obrazów wraz z obrysem ręcznym wykonanym przez specjalistę dermatologa. Obrazy te stanowiły tzw. złoty standard, który wyznaczał jakość wyznaczonego konturu. Parametry algorytmów dobrano optymalizując indeks Jaccarda, którym porównano wynik algorytmu i maskę binarną z bazy. Dla optymalnych zestawów parametrów każdej z metod porównano metody między sobą, wykorzystując podstawowe statystyki. Wyniki wskazują, iż metoda TTH pozwoliła na uzyskanie najwyższych wartości badanego wskaźnika, co wskazuje na najbliższe dopasowanie do konturu ręcznego. Z analizy usuwane były tzw. punkty odstające, których konturowanie jest utrudnione, np. ze względu na fakt, iż znamię zapełniało cały obraz.

Addresses

Damian BORYS: Silesian University of Technology, Institute of Automatic Control, ul. Akademicka 16, 44-100 Gliwice, Poland, damian.borys@polsl.pl .

Magdalena SZEREMET: Silesian University of Technology, Institute of Automatic Control, ul. Akademicka 16, 44-100 Gliwice, Poland, Magdalena Szeremet magdsze188@student.polsl.pl .

Krzysztof PSIUK-MAKSYMOWICZ: Silesian University of Technology, Institute of Automatic Control, ul. Akademicka 16, 44-100 Gliwice, Poland, krzysztof.psiuk-maksymowicz@polsl.pl.

Ziemowit OSTROWSKI: Silesian University of Technology, Institute of Thermal Technology, ul. Akademicka 16, 44-100 Gliwice, Poland, ziemowit.ostrowski@polsl.pl.

Mariusz FRACKIEWICZ: Silesian University of Technology, Institute of Automatic Control, ul. Akademicka 16, 44-100 Gliwice, Poland, mariusz.frackiewicz@polsl.pl.