Hourglass Shapes in Rank Grey-Level Hit-or-miss Transform for Membrane Segmentation in HER2/neu Images

Marek Wdowiak[1], Tomasz Markiewicz[1,2(✉)], Stanislaw Osowski[1,3], Zaneta Swiderska[1], Janusz Patera[2], and Wojciech Kozlowski[2]

[1] Warsaw University of Technology
[2] Military Institute of Medicine
[3] Military University of Technology, Warsaw, Poland
{wdowiakm,markiewt,sto,swidersz}@iem.pw.edu.pl,
{jpatera,wkozlowski}@wim.mil.pl

Abstract. The paper presents an automatic approach to the analysis of images of breast cancer tissue stained with HER2 antibody. It applies the advanced morphological tools to build the system for recognition of the cell nuclei and the membrane localizations. The final results of image processing is the computerized method of estimation of the membrane staining continuity. The important point in this approach is application of the hourglass shapes in rank grey-level hit-or-miss transform of the image. The experimental results performed on 15 cases have shown high accuracy of the nuclei and membrane localizations. The mean absolute error of continuity estimation of the stained membrane between the expert and our system results was 6.1% at standard deviation of 3.2%. These results confirm high efficiency of the proposed solution.

Keywords: Image segmentation · Object recognition · HER2/neu images

1 Introduction

Segmentation of the thin, non-continuous and highly variable objects is a difficult task in mathematical morphology application. To such problems belongs the membrane segmentation of the cells in histopathology Human Epidermal Growth Factor Receptor 2 (HER2/neu) images, related to breast cancer. The histopathological evaluation of the set of immunochemistry stains is the most common task for pathologists.

The HER2/neu biomarker is recognized as a diagnostic, prognostic and predictive factor not only, but especially in the case of breast cancer [1]. It is indicated as aid in an assessment of the breast and gastric cancers for patients for whom trastuzumab treatment is being considered. An over-expression of HER2 protein connected with HER2 gene amplification are diagnosed in approximately 20% of the analyzed breast cancer cases. For such patients, the trastuzumab treatment should be considered. Clinically, trastuzumab binds to the domain IV of the extracellular segment of HER2/neu receptor. Cells treated with this monoclonal antibody undergo arrest in the G1 phase of the cell cycle, so the reduced proliferation is observed. The combination of this

J.A. Benediktsson et al. (Eds.): ISMM 2015, LNCS 9082, pp. 3–14, 2015.
DOI: 10.1007/978-3-319-18720-4_1

antibody with chemotherapy has shown an increased survival and response rate, however, it increases the serious heart problems. It means, that inclusion of a patient to such therapy must be preceded by a reliable histological diagnosis.

The immunohistochemical HER2/neu stain is regarded as a basic step in pathomorphological evaluation of breast cancer. This semi-quantitative examination, performed on the immunostained paraffin section, needs determination of the presence, intensity, and continuity of membrane staining in the tumor cells. Four categories in grade scale are recognized: 0 (no membrane staining is observed or membrane staining is observed in less than 10% of the tumor cells), 1+ (a barely perceptible membrane staining is detected in more than 10% of tumor cells, the cells exhibit incomplete membrane staining), 2+ (a weak to moderate complete membrane staining is observed in more than 10% of tumor cells), and 3+ (a strong complete membrane staining is observed in more than 10% of tumor cells). The case 0 or 1+ indicates no HER2 gene amplification. On the other side grade 3+ indicates immediate HER2 gene amplification. The case 2+ means dubious (undecided) result and directs the case to the fluorescence in situ hybridization (FISH) based quantification of the level of HER2 gene amplification [2]. As shown, the categorization criteria are very subjective and may lead to significant differences in assessment of the particular cases. Especially the distinction between weak and strong, and the continuity of a membrane staining may result in a large variation in their practical interpretation.

The computerized assessment of pathological cases plays an increasingly important role in image analysis, especially when the quantitative result is required for diagnosis. Development of the slide scanners has introduced a lot of software used not only to perform the scanning process, but also for the quantitative analysis of the images. At the same time large data produced by this form of image acquisition has created new challenges for the image analysis algorithms. Whereas a lot of algorithms for the nuclear reactions have been developed [3,4], the HER2/neu membrane reaction is still treated manually or in a very rough way, not taking into account the separate cells. In the last years some new approaches to solve this problem have been proposed. To such methods belong the application of the real-time quantitative polymerase chain reaction (PCR) using LightCycler [5], application of support vector machine [6], or fuzzy decision tree by using Mamdani inference rules [7]. In spite of existing methods new approaches are needed, because of the high variability of a membrane reaction and it's frequent overlapping with a cytoplasm. They create difficult problems justifying for searching the other methods which are more effective to deal with this particular analysis problem.

In this paper novel tools of mathematical morphology for a membrane staining segmentation in HER2/neu images are proposed. Our propositions of the new hourglass shapes expand the traditional and common family of structuring elements, already discussed in mathematical morphology transformations. We used them in the rank grey-level hit-or-miss transform with some modifications, offering the supporting tool for segmentation of immunostained cell membranes. We point some analogies, which connect our propositions with the unconstrained hit-or-miss transform, rank hit-or-miss transform [8] and the grey-level hit-or-miss transform [9], where the inclusion and exclusion of regions in a structural shape are defined. We show the new shapes with fuzzy inclusion and exclusion criteria increase an efficiency of the membrane recognition. Moreover, our ideas can be adapted to other image processing problems with thin and non-continuous distinctive objects.

2 Problem Statement

An automatic evaluation of the HER2/neu membrane staining aims at the recognition of tumor cell nuclei and area of positive membrane reaction. It should specify which parts of recognized membrane come from the specific cells and finally graduate the reaction. Each of these steps requires different algorithms based on various criteria. The cell nuclei detection can be performed as a task of segmentation of the blue rounded and generally non-touching objects. The main problem is a weak staining of the nuclei by the blue hematoxyline. The additional problem is their partial coverage with the brown chromogenic substrate.

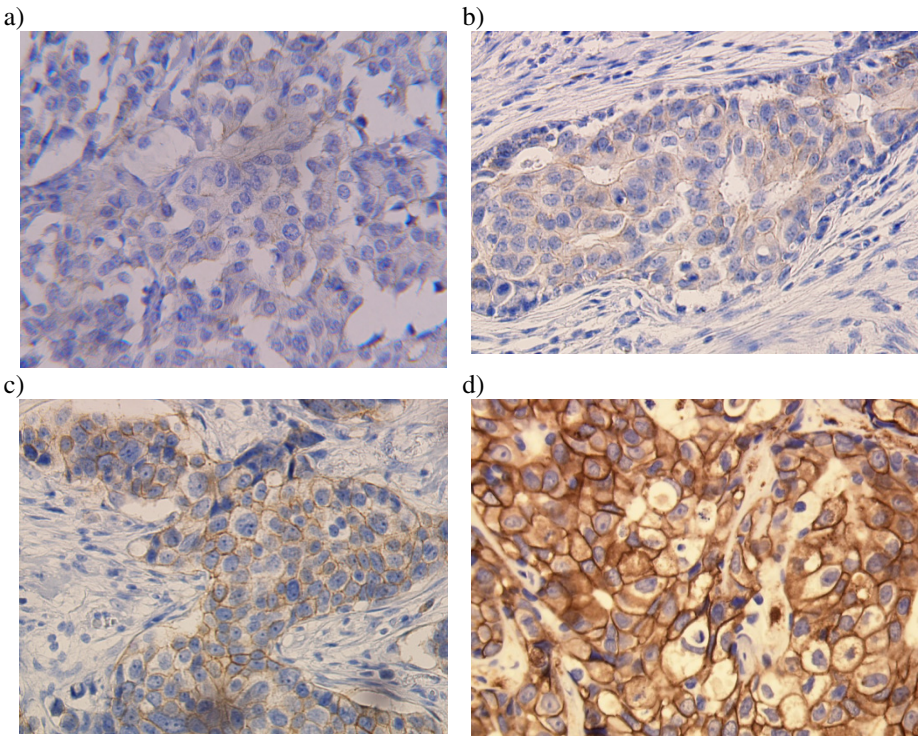

Fig. 1. Four typical cases of HER2/neu images representing the subsequent grades: a) 0, b) 1+, c) 2+, and d) 3+

The important task in this evaluation is recognition of area with a positive membrane reaction, especially when the brown chromogenic substrate is located not only in the cell membranes, but also partially in the cell cytoplasm's. In such cases, there is an identification problem of an appropriate membrane area with staining inside the brown marked regions. An algorithm for assigning parts of recognized membranes to the specific cells must be designed and applied.

A set of typical HER2/neu images with various grades of HER2 status are presented in Fig 1. As we can see the membrane reaction can vary from lower or higher

intensity in the location only in a cell membrane (a thin line) to a very high intensity exactly in the membrane location surrounded by slightly colored cytoplasm's of the touching cells. The aim of the presented study is to create method that will be able to identify any membrane positive reactions, irrespective of their intensity, different localization and character of brown chromogenic substrate.

3 Material and Methods

The materials used in experiments come from the archive of the Pathomorphology Department in the Military Institute of Medicine, Warsaw, Poland. Twenty cases of the breast cancer of HER2/neu preparations without any artefacts representing 1+, 2+, and 3+ grades were selected. The paraffin embedded breast tissue were stained in a standard way according to the Ventana PATHWAY anti-HER-2/neu (4B5) Rabbit Monoclonal Primary Antibody protocol [10]. The specimen images were registered on the Olympus BX-61 microscope with the DP-72 color camera under the magnification 400x and resolution 1024x768 pixels. The image processing algorithm was composed in the following steps: a) enhancement of an image contrast, b) creating a set of different color space representations, c) classification of image pixels into three classes: the nuclei, membrane with positive reaction and the remaining regions, d) segmentation of the cell nuclei, e) recognition of the stained cell membranes based on the grey level hit-or-miss transform with the new proposed shape patterns, f) allocation of the parts of membranes to the separate cells and g) calculation of the indicators of the stain intensity and continuity in each cell. This processing pathway applies a lot of methods, however, in this paper we focus mainly on the application of new hourglass shape patterns and fuzzy criteria in membrane localization.

3.1 Image Preprocessing, Colour Spaces and Pixel Classification

The first step of the HER2/neu evaluation algorithm is an image preprocessing intended to enhance the contrast and colors. It was realized by applying an automatically computed contrast stretch and normalizing the color components to fulfill their ranges. In the next step nine sub-image samples (see Fig. 2) were prepared and manually segmented into the blue nuclei (class 1), reactive brown membrane (class 2), and the remaining areas (class 3) to establish the most adequate pixel color components for a classifier. To create an efficient classification system we have to select proper diagnostic features among the possible image representations in the form of pixel intensities. We take into account the following color spaces: RGB, CMYK, HSV, YCbCr, CIE Lab, CIE Lch, CIE uvL, CIE XYZ [11]. The ability of pixel intensity representation to differentiate specific class by the support vector machine (SVM) classifier [12] was evaluated using an area (AUC) under Receiver Operating Characteristics (ROC) curve [13]. In the next step the cross-correlation between the candidate features was studied in order to select not only the best ones, but also of the least correlation between them. The best two features for each of three classes was found and these six features formed the feature vector for a classifier. The classification was

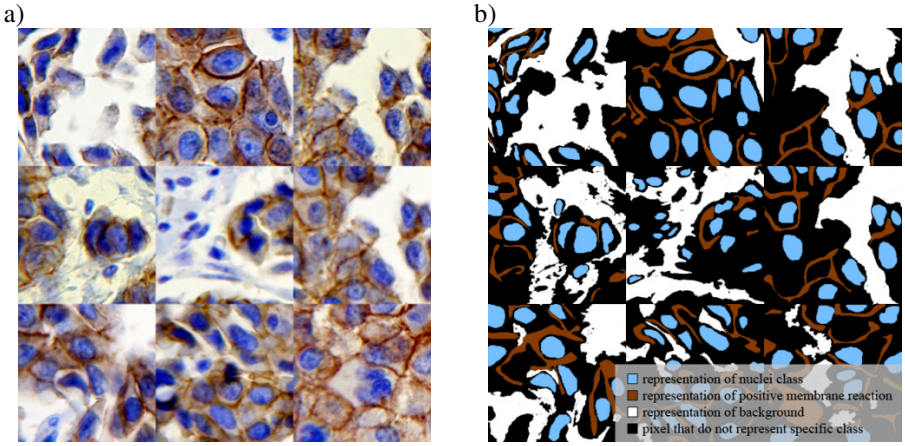

Fig. 2. The representatives of nine samples of image for establishing three classes of data for a classifier: (a) original images, (b) manually marked classes

performed using the SVM with a Gaussian kernel function. After this stage of image processing we obtain three masks representing the nuclei, positive HER2/neu reaction in cell membrane and cytoplasm's, and the remaining parts of image.

3.2 Segmentation of Cell Nuclei

The binary nuclei mask created by the classifier is used for the cell nuclei segmentation. The morphological closing operation was applied to reduce a noise. For separation of the connected nuclei the distance map was build and each extended regional maxima [8] with the selected h value was recognized as the center of cell nucleus. To define area of reaction the series of morphological operations, such as erosion, dilation, closing and hole filling were applied on the binary mask, which represented the positive HER2/neu reaction in cell membrane. Finally, the restriction to a single cell nucleus area was applied in order to eliminate non-cancer cell masks.

3.3 Intensity Map for Membrane Segmentation

To detect the immunoreactive cell membrane the most differentiated intensity map should be defined. It can be done for a single color channel, e.g. luminance, yellow component, or composed from the set of color channels. The selection of them is based on statistical analysis of the training images from Fig. 2. In the following section we will present the results concerning the analysis of ROC curves [13] and correlation between image descriptors to establish the most useful diagnostic features in classification process. The sample intensity map of the image corresponding to Fig. 2 is presented in Fig. 3. It is created as an element-wise product of Y channel from CMYK and u channel form CIE uvL colour space.

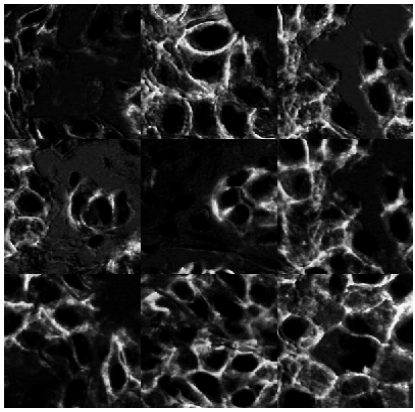

Fig. 3. The sample intensity map of the images from Fig. 2 based on Y (CMYK) and u (CIE uvL) components

3.4 Novel Hourglass Shapes for Discovering the Line Structures

Detection of immunoreactive cell membrane is a complex task due to the significant variability of cell shapes and spatial configuration, stain intensity and continuity, as well as non-specificity of staining. However, one characteristic point can be observed in each case: local domination of the stain intensity in the real cell membrane. This observation stands the basis for the proposition of the novel hourglass and half-hourglass shapes as the composite structuring elements in a detection of the sections of immunoreactive cell membrane (Fig. 4) by using the modified rank grey-level hit-or-miss approach. Thus, we define the horizontal and vertical hourglasses (the first two shapes) and four diagonal half-hourglass shapes as presented in Fig. 4. The central region (marked on red) treated as a foreground (FG) and the hourglass shape marked in dark grey (HG), represent the demanded location of the immunoreactive cell membrane. The other areas in the pattern (marked white in the figure) represent background (BG).

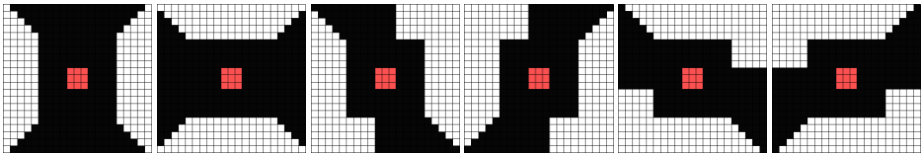

Fig. 4. The hourglass shapes as the proposed structuring elements in immunoreactive membrane recognition

The membrane curve can be treated as a line. The hourglass shapes offer some tolerance for the real line structure. Two cases might happen. In the first one the line structure is mainly included in HG region and only few pixels of the higher intensity values, comparable with values in FG region, are placed in BG area. In the second

case a line structure does not match HG region and there are significantly more pixels with a higher intensity value in BG region. The following formula of the rank grey-level hit-or-miss transform based on rank filter ζ [8] has been modified:

$$HMT_{\mathbf{B},k_{BG},h}(X) = \zeta_{B_{BG},k_{BG}}(X < \mu_{FG} - h) \tag{1}$$

where μ_{FG} is the mean value of the image X in FG region, k_{BG} is the rank value, and h is the assumed threshold value. The positive results represent the recognized line structure. The above formula is applied for each pixel of the image and for the sets of masks of Fig. 4 in the line structure detection. Each mask detects the line feature in a given direction. Final result is composed as the logical OR operation of the first two hourglass shape masks or the last four from Fig. 4.

3.5 Allocation of Membranes in Separate Cells

As a result of these steps of algorithm the cell nuclei mask and the immunoreactive membrane regions are recognized in an independent way. The next step is to connect them into one compact system. To realize this task, a watershed algorithm is performed [14]. Its aim is to determine the potential cell membrane location.

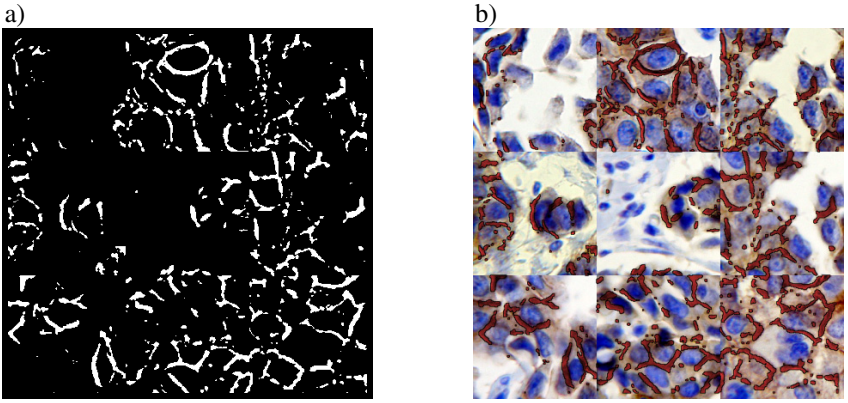

Fig. 5. Final result of application of hourglass shapes for line structure detection of immunoreactive membrane (a) and an input image with the annotated lines of immunoreactive membranes (b)

Series of image filtering operations are performed to reduce the number of false minima. They include (in sequence): the morphological operations of dilation using the cross structuring element of the length equal 3, the gamma correction of γ=0.7 to reduce the nonlinearity of luminance, the morphological operations of closing using the square 5×5 structuring element to smooth the edges of the objects, and finally h-minima transformation of h=0.04 to equalize the values of edge pixels. Watershed segmentation results related to the image of Fig. 3 are presented in Fig. 5. Each watershed region is considered as a separated cell and the region contour as full membrane. Logical AND

operation on dilated contour region and detected line structures presented in Fig. 5a were performed to allocate immunoreactive membrane in separated cells.

Final step of image analysis is to calculate the stain continuity in each recognized cell, e.g., the cell with segmented nucleus. The continuity of membrane staining is calculated as a percentage of the recognized stained parts of cell membrane to a cell perimeter length. The exemplary results are presented in the Fig. 6.

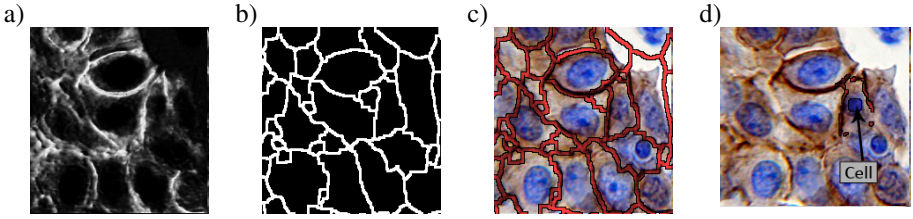

Fig. 6. The immunoreactive membrane representation (a), watershed segmentation results (b), watershed segmentation results presented on the input image (c), allocated membrane for one cell with the calculated stain continuity of 68% (d)

4 Results

Segmentation of thin, non-continuous and highly variable objects is a difficult task in automatic image processing. This section will present the numerical results concerning different stages of image processing, which lead finally to recognition of the membrane. The basic difficulty is recognition of three classes of objects (nuclei, reactive membrane, and remaining areas of the image). To solve this problem we have to select the color mapping of the image, which represents the highest ability of class discrimination (so called diagnostic features). The potential features under selection represent the image pixels intensity in different color representations. Fig. 7 presents the exemplary results concerning the uvL color space in the form ROC curve (Fig. 7a) and correlation family between investigated features (Fig. 7b). The higher AUC the better is the diagnostic significance of the feature.

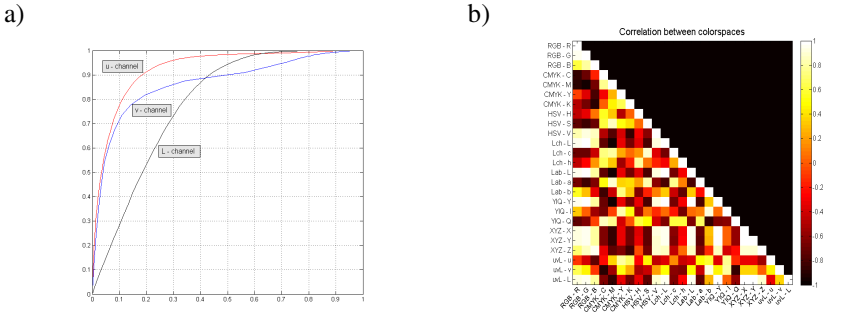

Fig. 7. The ROC curve for uvL colour representation (a) and correlation map between pairs of all 25 features representing different colour spaces (b)

We have applied the significance measure of the feature in the form of difference between the actual AUC and the value 0.5, which corresponds to the random classifier. Fig. 8 presents this significance measure for all features, arranged from the highest to the lowest recognition ability of the class 1 (Fig, 8a), class 2 (Fig, 8b) and class 3 (Fig, 8c).

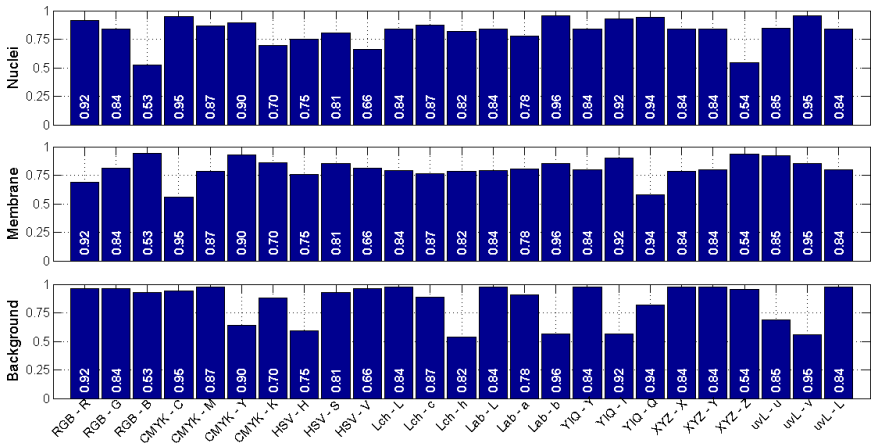

Fig. 8. The diagrams presenting the significance of the succeeding colour features for recognition of a) cell nuclei, b) immunoreactive cell membrane and c) the remaining part of the image (stroma and background)

To limit the number of features used in the class recognition we have chosen 2 highest discriminative ability features for each class. On the basis of the presented results the following features have been selected: R and B components from RGB space, u and L components from uvL representation, b component from CIE Lab and Y component from CMYK. Next, a set of learning data corresponding to nine parts of the image presented in the Fig. 2 was used in learning the SVM classifier of the Gaussian kernel. Finally, based on our experience and the introductory experiments, the threshold value h=0.2, used to recognize the stain intensity gradation was selected.

To identify the areas of immunoreactive cell membranes a set of proposed hourglass structuring shapes presented in Fig. 4 was used and evaluated. Based on the preliminary experiments the following parameters for full hourglass shape were adjusted: FG size 3×3 pixels, the distance between FG and BG equal 4 pixels, mask size 21×21 pixels. The second shape, half-hourglass of the structures presented in the Fig. 4, used the same parameters as the full hourglass shape. Moreover, the rank value k_{BG} in (1) was selected in such a way that at least 80% of BG pixels of the image match the relation. The normalized h value has been settled as 0.2 for both shapes.

In the testing phase of the adapted system we have analyzed 15 cases of 2+ HER2/neu status. A set of microscopic images was automatically processed and the results were evaluated in two aspects: a) how many immunoreactive membrane areas

were properly recognized and b) what was the mean value of the analyzed immunoreactive cell membrane staining continuity. The detailed numerical results of estimation of the membrane staining continuity related to the application of full hourglass and half-hourglass shapes are presented in Table 1. They represent the testing cases, not taking part in fitting parameters of the image processing. As can be seen, the half-hourglass shapes offer better than hourglass accuracy of a immunopositive membrane recognition. Thus, we recommend them to the HER2/neu image analysis. However both approaches show some little bias, lower than common pathological criteria of between experts results. In the future research larger population of more representative learning images will be used to parameter adjusting and this approach should reduce the bias.

Table 1. The numerical results of membrane staining continuity estimation made by expert and by our system at application of full hourglass and half-hourglass shapes

Case	Continuity [%]				
	Expert's result	Hourglass	Error	Half-hourglass	Error
1	40.4	38.3	-2.0	35.1	-5.3
2	31.6	36.7	5.1	33.5	1.9
3	29.3	42.3	13.1	38.8	9.6
4	35.9	53.4	17.5	47.4	11.6
5	31.6	37.1	5.5	33.2	1.6
6	30.4	44.6	14.2	40.1	9.7
7	32.2	33.0	8	28.9	-3.3
8	35.3	37.8	2.5	33.9	-1.4
9	23.2	30.9	7.7	27.9	4.7
10	32.8	45.0	12.2	40.9	8.0
11	30.8	38.0	7.2	34.3	3.6
12	19.1	29.5	10.4	27.1	8.0
13	25.6	33.9	8.3	31.4	5.8
14	29.8	43.1	13.4	39.6	9.8
15	37.3	49.0	11.6	44.7	7.4
	Mean absolute error		8.8%		**6.1%**
	Standard deviation		4.8%		**3.2%**

The results of image analysis may be also presented in the graphical form. Fig. 9 depicts the exemplary original image under analysis (Fig. 9a) and the results of recognition of the nuclei and membranes made by the expert (Fig. 9b) and by our automatic system (Fig. 8c). The close similarity of the results of image segmentation obtained by our system and expert is confirmed.

It should be noted, that our computerized system is fully automatic. Hence the recognition of the nuclei has been done according to the embedded procedure defined within the algorithm. The user does not interfere in this process. On the other hand the human expert selects the nuclei according to his professional knowledge, blind to the selection results of an automatic system. Therefore, slight differences between the recognized nuclei can be observed in both segmentation results of Fig. 9.

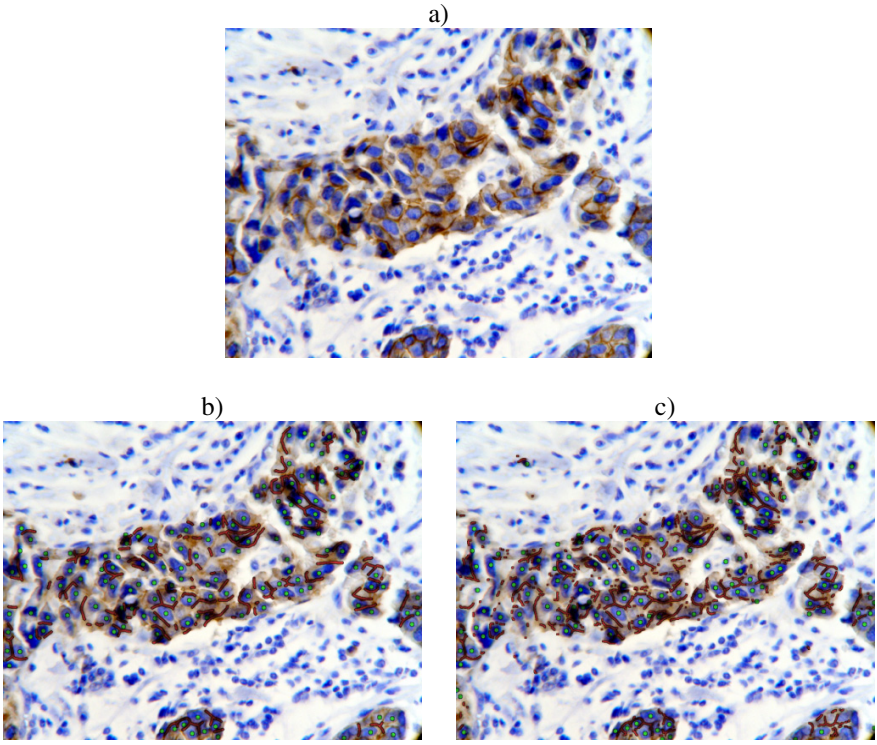

Fig. 9. The original input image (a) and its segmentation made by b) an expert, c) our automatic system with the visible membranes (brown colour lines) and cell nuclei (green points)

5 Conclusions

The paper has presented new approach to the automatic evaluation of the HER2/neu membrane staining in the breast cancer samples, applying the morphological image analysis. The developed algorithm uses few steps of analysis leading to the recognition of the cell nuclei and the membrane localizations. The paper proposed the automatic method of the membrane staining continuity estimation. The important point in this approach is the application of the hourglass shape structuring element in the rank grey-level hit-or-miss transform for the analysis of the image. The experimental results have shown high efficiency of image segmentation with respect to the nuclei and membrane localizations. The mean absolute error of continuity estimation between the expert and our system results, obtained for 15 analyzed cases, was 6.1% at standard deviation of 3.2%. In our opinion this automatic approach, after some additional research, will be able to substitute human expert in this very demanding and tedious task of image processing.

Acknowledgement. This work is partly supported by the National Centre for Research and Development (Poland) by the grant PBS2/A9/21/2013 in the years 2013-2016.

References

1. Littlejohns, P.: Trastuzumab for early breast cancer: evolution or revolution? Lancet Oncology 7(1), 22–33 (2006)
2. Les, T., Markiewicz, T., Osowski, S., Cichowicz, M., Kozlowski, W.: Automatic evaluation system of FISH images in breast cancer. In: Elmoataz, A., Lezoray, O., Nouboud, F., Mammass, D. (eds.) ICISP 2014. LNCS, vol. 8509, pp. 332–339. Springer, Heidelberg (2014)
3. Lezoray, O., Elmoataz, A., Cardot, H., Gougeon, G., Lecluse, M., Elie, H., Revenu, M.: Segmentation of colour images from serous cytology for automated cell classification. Anal. Quant. Cytol. Histol. 22, 311–322 (2000)
4. Grala, B., Markiewicz, T., Kozlowski, W., et al.: New automated image analysis method for the assessment of Ki-67 labeling index in meningiomas. Folia Histochemica et Cytobiologica 47(4), 587–592 (2009)
5. Kim, Y.R., Choi, J.R., Song, K.S., Chong, W.H., Lee, H.D.: Evaluation of HER2/neu status by real-time quantitative PCR in breast cancer. Yonsei Med. J. 43(3), 335–340 (2002)
6. Pezoa, R., Rojas-Moraleda, R., Salinas, L.: Automatic membrane segmentation of IHC images enables the analysis of cancer breast tissues. In: Third Chilean Meeting on Biomedical Engineering, pp. 1–6. JCIB, Santiago (2012)
7. Tabakov, M., Kozak, P.: Segmentation of histopathology HER2/neu images with fuzzy decision tree and Takagi-Sugeno reasoning. Comput. Biol. Med. 49, 19–29 (2014)
8. Soille, P.: Morphological Image Analysis, Principles and Applications. Springer, Berlin (2003)
9. Naegel, B., Passat, N., Ronse, C.: Grey-level hit-or-miss transforms Part I:Unified theory. Pattern Recogn. 40, 635–647 (2007)
10. PATHWAY HER-2/neu (4B5) – user manual. Ventana Medical Systems, Tucson (2009)
11. Gonzalez, R., Woods, R.: Digital image processing. Prentice Hall, New Jersey (2008)
12. Schölkopf, B., Smola, A.: Learning with kernels. MIT Press, Cambridge (2002)
13. Tan, P.N., Steinbach, M., Kumar, V.: Introduction to data mining. Pearson Education Inc., Boston (2006)
14. Matlab Image Processing Toolbox, user's guide. MathWorks, Natick (2012)