A modified technique for smart textural feature selection to extract retinal regions of interest using image pre-processing

N Yu Ilyasova\textsuperscript{1,2}, A S Shirokanev\textsuperscript{1,2}, R A Paringer\textsuperscript{1,2}, A V Kupriyanov\textsuperscript{1,2} and A V Zolotarev\textsuperscript{3}

\textsuperscript{1}Samara National Research University, Moskovskoe shosse st. 34, Samara, Russia, 443086
\textsuperscript{2}Image Processing Systems Institute - Branch of the Federal Scientific Research Centre “Crystallography and Photonics” of Russian Academy of Sciences, Molodogvardeyskaya str. 151, Samara, Russia, 443001
\textsuperscript{3}Samara State Medical University, Chapayevskaya st. 89, Samara Russia, 443066

e-mail: sanek-shirokanev@yandex.ru

Abstract. To increase the efficiency of the laser coagulation surgery the problem of the most accurate segmentation of fundus images is especially relevant. Fundus image segmentation is carried out with high accuracy using effective features and the minimum number of parameters for segmentation of a single image fragment. This paper describes a modified technique for smart textural feature selection to extract retinal regions of interest using image preprocessing algorithms. Preprocessing algorithms significantly influence the selected features which provide a minimum error of object recognition. In addition image preprocessing algorithms provide a more precise object selection. The informativeness of the obtained feature space is studied using discriminant data analysis. The best fragmentation block size segmentation and feature sets provides the necessary accuracy to identify regions of interest. Those regions are determined by the analysis of the following 4 classes of fundus images: exudates, thick, thin vessels and healthy areas. The advantages and disadvantages of the considered preprocessing algorithms were identified.

1. Introduction
To treat various pathologies of the fundus in time, an early diagnosis is required. Ophthalmoscopy is used to detect the vast majority of eye diseases and pathologies. Diabetes mellitus (DM) is one of most severe medical problems of the modern world. One of the symptoms of the disease is a change in the blood vessels of the retina. It is the result of a disruption in the supply of retinal vessels with oxygen. That state of the visual system leads to the appearance of diabetic retinopathy (Figure 1). Laser coagulation of retina is a gold standard in the treatment of DRP [1,2,3]. In the course of laser treatment, the macular edema zone is exposed to a series of dosed microburns or laser coagulates. The coagulates are applied either one at a time or as a series of coagulates arranged in a given regular pattern, or based on a preliminary planning of the coagulate pattern using a real-time retina image [4] (Figure 2). It is preferable to arrange coagulates in the zone of macular edema in an optimal way, possibly applying them at equal distance and avoiding blood vessels. The development of a feature selection technique for fundus image segmentation, enabling coagulated areas to be automatically
arranged in the macular oedema, including different locations of blood vessels, presents a relevant problem. In medical problems classification methods are applied for diagnosis [5, 6, 7]. Proposed technique is focused on fundus image segmentation.

Image preprocessing methods can play a key role in image segmentation algorithms. Preprocessing methods can provide a better visual separation of classified objects [8]. In this paper, we propose to use preprocessing methods to process fundus image fragments immediately before calculating clustering errors.

\[
\text{(a)} \quad \text{(b)}
\]

**Figure 1.** The example of a fundus diagnostic image without pathology (a) and with pathology detected (b).

In order to calculate an informative feature space, the image was preliminarily divided into fragments containing specific regions of interest, with four classes of objects found there: exudates, thin vessels, thick vessels, and intact areas. The area of the macular edema is characterized by the accumulation of exudate zones. When conducting laser treatment, it is forbidden to apply coagulates on thick vessels and recommended to avoid affecting intact areas and thin vessels, thus enhancing the effectiveness of the laser surgery.

\[
\text{(a)} \quad \text{(b)}
\]

**Figure 2.** Examples of laser coagulation of retina (a) and pattern examples of the software NAVILAS (b).

2. Region of interest extraction technique based on the textural analysis of biomedical images in various subspaces of RGB

To automate the laser coagulation procedure the image needs to be segmented into the specified regions of interest. Highly accurate image segmentation is often provided by using image preprocessing algorithms. Preprocessing algorithms allow to improve the legibility of recognized objects and to select novel effective features that increase the accuracy of object recognition. To be able to conduct a more accurate fragmentation a technique for generating a set of effective features based on the analysis of textural properties of the above-said image classes using discriminant analysis was described in [9]. In this paper we propose a modified technique based on the textural feature selection using image preprocessing algorithms. The regions of interest are extracted by making the decision whether an image fragment belonged to one of the four above-described object classes:
exudates, thick vessels, thin vessels, and intact areas. For fragmentation the image is divided into square-shaped blocks, which were then classified based on the technique described in [10].

The technique is based on clustering methods with the use of texture features. Analysis of the regions of interest (ROI) of the source images has shown their texture properties to be fairly different. It is worth noting that texture features were earlier shown to provide good results in biomedical image recognition and their subsequent use in diagnostics [11, 12, 13, 14]. To calculate the texture features, we used the MaZda library [15, 16]. A similar problem of the identification of a disease using blood cell images was handled via analyzing texture features of different classes of original images in different color subspaces in Ref. [17].

Thus, this work is an attempt to improve the earlier proposed technique via forming a fragment sampling not only in the initial color space but also in separate components of color spaces RGB and HSL, as well as forming a gray-level image fragment sampling. Comparative experimental studies intended to identify more informative color subspaces, as well as more informative texture features of different color subspaces were conducted. Thus, we formed seven initial samplings of feature vectors, which were calculated in the images in different combinations of color components of spaces RGB and HSL. Besides, a decision was made to carry out an additional procedure of rotating learning image fragments, thus obtaining more objective results when forming a set of effective features.

In this work, the general and pair-wise feature selection was conducted. With the general approach to feature selection, the feature space is composed of features which are the best in terms of the discriminant analysis' separation criteria [9, 18, 19, 20, 21], which is calculated for all four classes of the objects under study. We utilized a pairwise feature selection in which features characterized by the largest separation criteria were selected, while separating each pair of classes and subsequently uniting them in an integrated set. To estimate the quality of the resulting sets the clustering error was calculated for all fragmentation block sizes under study. Clustering was conducted using a K-means method, with the Euclidean and Mahalanobis distance used as a similarity measure [20].

The modernized technique of effective features selection contains the following global processing steps: 1) Selection of image fragments and classification based on medical assessment; 2) Improving the image informativeness using image preprocessing algorithms; 3) Forming images in RGB and HSV subspaces with and without rotation; 4) Forming features per fragment in the MaZda library; 5) Choice of the feature selection technique (pairwise or general); 6) Assessment of features efficiency for the visibility of classes; 7) Selection of a set of informative features to provide the required classification accuracy; 8) Forming new effective features space and selection of the best mask dimension; 9) Image segmentation.

3. Texture features
Texture features have shown themselves to be well-suited for biomedical image recognition and subsequent diagnostics. There have been a large number of texture features. The software “MaZda” [15, 16] utilized in this work is able to calculate the following groups of texture features: a) those based on statistical characteristics (histogram-based features); b) those based on gradient parameters; c) those based on the co-occurrence matrix; and d) those based on run-length matrix (see MaZda user manual, http://www.eletel.p.lodz.pl/mazda/download/mazda_manual.pdf for further details) [14]. Below, we give the most frequently found features at different fragmentation block sizes and different selection techniques used in the proposed technique for informative feature selection.
analyzing the relationships revealed in the course of the experiments we choose the smallest fragmentation block at which the clustering error and separation criteria show a qualitative leap.

Figure 3 shows the value of the group separation criteria versus the fragmentation block size at a varying number of chosen features (from 1 to 30), which are characterized by the maximal individual separation criteria. Figure 4 shows that the optimal number of features equals 13, with their further increase not resulting in an essential growth of the general separation criteria. Note that from the plot the qualitative leap is seen to occur at the block size of 12, signifying a possibly low clustering error at this point. This is confirmed by the plot of the clustering error versus the fragmentation block size at different similarity measures (Figure 4). A qualitative leap of the clustering error occurs at the same size of the fragmentation block of 12 (marked with a dashed line), enabling an acceptable error to be attained.

Dependencies of the general group separation criteria and the fragmentation of block size were obtained for a different number of selected features. However a pairwise selection method allows to significantly increase the separation criteria by using the similar value of minimum window size – 12. Experiments have shown that pairwise selection technique allows to achieve a smaller clustering error compared to the general selection technique. Figure 5 shows the clustering error compared to the fragmentation block size for RGB subspaces when using Mahalanobis similarity measurement for a set of 5 features with the maximal separation criteria for a pairwise feature selection technique. Data analysis of the plot suggests that the smallest clustering error and the minimal-size fragmentation block with the acceptable error rate (<2.5%) is reached for 12-pixel fragments when using features selected pairwise in green subspace.

Figure 3. Values of the group separation criteria vs. the fragmentation size at a varying number of selected features characterized by the maximal separation criteria (general selection technique).

Figure 4. Clustering error vs. the fragmentation block size at varying similarity measures when using six features with the maximal separation criteria for a pairwise selection: a) Euclidean distance, original features; b) Mahalanobis distance, original features; c) Euclidean distance, a set of newly formed features; d) Mahalanobis distance, a set of newly formed features features.
Thus, the green subspace is most informative. The probability of feature use (selection frequency) is shown in Figure 6 for varying-size fragmentation blocks (from 10 to 50) and different rotation angles. In Fig. 10 one can see features that occur most frequently for an informative feature vector formed at varying-size fragmentation block, which turned out to be image-rotation invariant.

The experiments have shown the following features to be most informative and having highest selection probability for varying-size fragmentation blocks: (а) at the pairwise selection: G_Sigma, G_Skewness G_S(5,0)Entropy, G_S(5,5)Entropy, Perc.10%, H_Perc.99%, B_Perc.10%, B_Perc.01%, B_Perc.99%, B_GrVariance, H_S(1,0)Correlat, B_S(5,5)Entropy, H_S(0,1)Correlat, and (b) at the general selection: G_S(5,-5)Entropy, G_S(5,5)Entropy, B_Perc.99%, G_S(0,5)Entropy, G_S(4,-4)Entropy, G_S(5,0)Entropy, G_S(4,4)Entropy, G_Sigma, G_S(0,1)Entropy, G_S(1,0)Entropy, H_Perc.99%, G_S(0,1)Correlat, G_S(1,0)Correlat. At any selection technique, the following features turned out to be common: G_S(5,5)Entropy, B_Perc.99%, G_S(5,0)Entropy, G_Sigma, H_Perc.99% (the first letter denotes the color subspace). Several feature sets characterized by the maximal separation criteria are shown in Table 1 depending on a particular feature selection approach and the presence of image rotation.

Table 1. Feature sets characterized by the maximal separation criteria vs. a particular feature selection technique and the presence of image rotation.

| Without rotation | With rotation | Without rotation | With rotation |
|-----------------|--------------|-----------------|--------------|
| B_Perc.99%      | B_Perc.99%   | Perc.10%        | Perc.10%     |
| B_Perc.90%      | B_Perc.90%   | G_Skewness      | S(0,5)Entropy|
| G_S(1,0)Entropy | G_S(1,0)Entropy | G_S(5,0)Entropy | G_Skewness   |
| G_S(0,1)Entropy | G_S(0,5)Entropy | B_Perc.99%    | G_S(0,5)Entropy |
| H_Perc.90%      | G_S(0,3)Entropy | B_GrVariance   | B_GrVariance |
| G_S(0,5)Entropy | G_S(0,4)Entropy | H_Perc.90%     | H_Perc.90%   |

5. Results of preprocessing fundus image segments

Image preprocessing produces an effect on the precision of object selection in the image. Moreover, it will result in new result on the field of study. The figure 7 shows the clustering error when using different image preprocessing methods and a set of 6 features with the maximal separation criteria for a pairwise feature selection technique. We used the linear contrast, the equalization and the median filtering.

The equalization decreased the quality of the result obtained using another preprocessing methods. However the clustering error falls short when the size of the fragmentation block is too large. In figure 8, the equalization illustrates a decently small error for large fragmentation block sizes. The results
show that the equalization breaks the textural properties of the fundus image. Therefore, the equalization should be used for large fragmentation block sizes (18 and more). When using 20 features and pairwise selection technique, the linear contrast improve the results if fragmentation block sizes vary in range \([9, 10]\). The cross-mask of 3x3 was chosen for median filtration. The median filter allows to improve the source results obtained without preprocessing.

When using general selection technique, the linear contrasting gives a significant reduction of clustering errors if the fragmentation block size more 12 (Figure 9). Compared to results obtained by general selection techniques without preprocessing (Figure 9), where the error rate reached 5%, the linear contrast has an error rate of about 2%.

**Figure 6.** The probability (frequency) of feature selection at varying-size fragmentation block obtained at pairwise feature selection.

**Figure 7.** Clustering error vs. fragmentation block size when using different image preprocessing methods and a set of 6 features with the maximal separation criteria for a feature pairwise selection technique.

To improve the accuracy of fundus image segmentation, preprocessing methods can be used. The linear contrasting is more efficient to use for a general feature selection technique. The equalization allows improving the source results by large fragmentation block size. The median filtering works better when a pairwise feature selection technique is used.

The figure 10 shows the results of the fundus image segmentation by using the presented feature selection technique. Based on the visual assessment by a medical expert, it can be concluded that the best segmentation result is presented in figure 10d, which has also been confirmed by the previously-mentioned studies.

6. Conclusion

Image preprocessing algorithms contributed to the selection of novel textural features allowed ensuring high clustering accuracy. The presented technique for selecting effective features using image preprocessing algorithms is intended for fundus image clustering and based on the use of different
colour subspaces. The technique allowed to improve the results of the intelligent feature analysis to be conducted, when extracting the regions of interest containing the four classes of objects (exudates, intact areas, thick vessels, and thin vessels) for laser coagulation surgery. The effectiveness of the set of features was estimated using a clustering procedure based on the K-means method. The Euclidean and Mahalanobis distances were used as a similarity measure. The choice of the required fragmentation block size and the similarity measure was based on the criteria of minimal clustering error among all smallest-size fragmentation blocks.

**Figure 8.** Clustering error vs. fragmentation block size when using different image preprocessing methods and a set of 20 features with the maximal separation criteria for a feature pairwise selection technique.

**Figure 9.** Clustering error vs. the fragmentation block size when using different image preprocessing methods and a set of 13 features with the maximal separation criteria for a general feature selection technique.

The experimental verification of the technique conducted on a series of a hundred fundus images (256200 fragments containing different classes of fundus elements in different colour subspaces) has made possible to identify most informative texture features (13 features for the general feature selection technique and 6 features for the pairwise selection). The fragmentation block size (12 pixels) obtained the best clustering results. It is worth noting compared to Ref. [9] the use of differently coloured subspaces and additional rotation of images in the learning samplings has made possible to reduce the fragmentation image size (from 47 to 13 pixels) while maintaining the clustering reliability within 95%, which is an important factor in laser coagulation surgery. The feature pairwise selection approach has enabled the clustering error to be reduced twice, and the number of features to be reduced more than twice (up to six). Depending on a particular feature selection approach and the presence of image rotation, several feature sets characterized by the maximal separation criteria have been obtained. Based on this study, the following recommendations for the best clustering results have been worked out: 1) the use of a 12-pixel fragmentation block; 2) pairwise feature selection; and 3) the use of the following texture features: Perc.10%, G_Skewness, G_S(5,0)Entropy, B_Perc.99%, B_GrVariance, H_Perc.90%.
It also should be noted that the proposed technique has made possible not only to extract the informative features in specific colour spaces but also to identify the most informative colour subspace for the best feature selection technique. The results of the studies using the fundus image pre-processing methods reduces the clustering error under specific conditions. The linear contrast provides a significant reduction of the clustering error when using a general feature selection technique. The median filtering allows reducing the error when using a pairwise selection technique. The equalization breaks the textural properties of the image and can only be effective in large fragmentation block sizes.

![Segmentation Results](image)

**Figure 10.** Results of fundus image segmentation obtained depending on a particular feature selection technique, the presence of image rotation and varying similarity measures: a) original image; b) c) d) e) - pairwise feature selection; f), g), h), i) - general feature selection; b), c), f), g) - without rotation; d), e), h), i) - with rotation; b), d), f), i) - Mahalanobis distance; c), e), g), h) - Euclidean distance.

7. References

[1] Doga A V, Kachalina G F, Pedanova E K and Buryakov D A 2014 Modern diagnostic and treatment aspects of diabetic macular edema *Ophthalmology, Diabetes* 4 51-59

[2] Astakhov Y S, Shadrichev F E, Krasavina M I and Grigoryeva N N 2009 Modern approaches to the treatment of a diabetic macular edema *Ophthalmologic Sheets* 4 59-69

[3] Ilyasova N Yu Methods for digital analysis of human vascular system. Literature review 2013 *Computer Optics* 37(4) 517-541

[4] Kernt M, Cheuteu R and Liegl R G 2012 Navigated focal retinal laser therapy using the NAVILAS® system for diabetic macula edema *Ophthalmologe* 109 692-700

[5] Gaidel A V and Krasheninnikov V R 2016 Feature selection for diagnosing the osteoporosis by femoral neck X-ray images *Computer Optics* 40(6) 939-946 DOI: 10.18287/2412-6179-2016-40-6-939-946
[6] Smelkina N A, Kosarev R N, Nikonorov A V, Bairikov I M, Ryabov K N, Avdeev A V and Kazanskiy N L 2017 Reconstruction of anatomical structures using statistical shape modeling Computer Optics 41(6) 897-904DOI: 10.18287/2412-6179-2017-41-6-897-904

[7] Smelkina N A, Kolsanov A V, Chaplygin S S, Zelter P M and Khramov A G 2017 Pulmonary emphysema recognition by CT scan Computer Optics 41(5) 726-731 DOI: 10.18287/2412-6179-2017-41-5-726-731

[8] Rubis A Yu, Lebedev M A, Vizilter Yu V and Vygolov O V 2016 Morphological image filtering based on guided contrasting Computer Optics 40(1) 73-79 DOI: 10.18287/2412-6179-2016-40-1-73-79

[9] Ilyasova N, Paringer R and Kupriyanov A 2016 Regions of interest in a fundus image selection technique using the discriminative analysis methods Lecture Notes in Computer Science 9972 408-417

[10] Ilyasova N, Paringer R, Kupriyanov A and Kirsh D 2017 Intelligent feature selection technique for segmentation of fundus images IEEE Xplore (INTECH) 138-143

[11] HeiShun Yu, Tischler B, Qureshi M M, Soto J A, Anderson S, Daganawala N, Li B and Buch K 2016 Using texture analyses of contrast enhanced CT to assess hepatic fibrosis European Journal of Radiology 85(3) 511-517

[12] Gentillon H, Stefańczyk L, Strzelecki M and Respondek Liberska M 2016 Parameter set for computer-assisted texture analysis of fetal brain BMC Research Notes 9 p. 496

[13] Acharya U R, Ng E Y, Tan J H, Sree S V and Ng K H 2012 An integrated index for the identification of diabetic retinopathy stages using texture parameters Journal of Medical Systems 36(3) 2011-2020

[14] Hajek M, Dezortova M, Materka A and Lerski R 2006 Texture Analysis for Magnetic Resonance Imaging (Med 4 publishing) p. 234

[15] Strzelecki M, Szczypielki P, Materka A and Klepaczko A 2013 A software tool for automatic classification and segmentation of 2D/3D medical images Nuclear Instruments and Methods In Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 702 137-140

[16] Szczypielki M, Strzelecki M, Materka A and Klepaczko A 2009 MaZda – A software package for image texture analysis Computer Methods and Programs in Biomedicine 94(1) 66-76

[17] Nikitaev V 2010 Method for recognizing the image of the texture of cells Patent RF. No.2008141633/28 (Access mode: http://www.freepatent.ru/images/patents/78/2385494/patent-2385494.pdf)

[18] Ilyasova N Yu, Kupriyanov A V and Paringer R A 2015 The Discriminant Analysis Application to Refine the Diagnostic Features of Blood Vessels Images Optical Memory and Neural Networks (Information Optics) 24(4) 309-313

[19] Kim J-A, Myuller Ch U and Klekka W R 1988 Factor, discriminant and cluster analysis (Financy i Statistica Publisher) p. 215

[20] Fukunaga K 1972 Introduction to statistical pattern recognition (New York and London: Academic Press) p. 369

[21] Ilyasova N Yu, Kupriyanov A V and Paringer R A Formation of features for improving the quality of medical diagnosis based on discriminant analysis method 2014 Computer Optics 38(4) 751-756

Acknowledgments

This work was partially supported by the Ministry of education and science of the Russian Federation in the framework of the implementation of the Program of increasing the competitiveness of Samara University among the world’s leading scientific and educational centers for 2013-2020 years; by the Russian Foundation for Basic Research grants (# 15-29-03823, # 16-41-630761, # 17-01-00972, # 18-37-00418), in the framework of the state task #0026-2018-0102 "Optoinformation technologies for obtaining and processing hyperspectral data".