Superpixel and Stereo Matching Optic Disc Segmentation for the Diagnosis of Glaucoma

Mohammad Norouzifard[1], Arpita Dawda[1], Anmar Abdul-Rahman[2], Hamid GholamHosseini[1], and Reinhard Klette[1]

[1] School of Engineering, Computer, and Mathematical Sciences
Auckland University of Technology, Auckland, New Zealand
Mohammad.Norouzifard@aut.ac.nz

[2] Department of Ophthalmology, Counties Manukau DHB, Auckland, New Zealand
anmar.abdul-rahman@middlemore.co.nz

***Abstract*—Glaucoma is an optic neuropathy resulting in progressive vision loss. It is the leading cause of global irreversible blindness. The reported prevalence among the population in New Zealand is 2% over the age of 40 years. About 10 % of those over 70 years are diagnosed with this disease.**

Population-based studies report high rates of undiagnosed glaucoma with over 50% of the population with glaucoma living in developed countries remaining undiagnosed and unaware of their disease. Clinical diagnosis rests on the detection of the characteristic optic disc signs. Stereo optic-disc imaging improves intra and inter-observer agreement in the detection of optic disc abnormalities. We propose a robust method to detect some abnormalities in stereo optic-disc images using stereo vision and superpixel segmentation concepts. A stereo vision system produces a disparity map for the input stereo images of the retina in which abnormalities are more distinguishable. The produced disparity map is then segmented using two different superpixel segmentation algorithms (simple linear-iterative clustering and simple non-iterative clustering) to detect abnormalities. The original stereo images are also segmented using the same concept; results are compared with the segmented disparity map.

***Index Terms*—Glaucoma detection, abnormalities, stereo vision, superpixel segmentation**

I. Introduction

Glaucoma is an optic neuropathy resulting in progressive vision loss [1]. It is the leading cause of global irreversible blindness [2]. In 2013, the population aged 40–80 years with glaucoma worldwide was estimated to be 64.3 million; this is projected to increase to 76.0 million in 2020 and 111.8 million in 2040 [3].

In New Zealand, glaucoma is responsible for 7% (1,192 patients) of cases of bilateral blindness, ranking the third most common reported etiology [1]. The reported prevalence among the population in New Zealand is 2% over the age of 40 years. About 10% of those over 70 years are diagnosed with this disease [1]. Early detection is vital to reduce the burden of unnecessary blindness due to glaucoma. The Royal Australian and New Zealand College of Ophthalmologists recommends a biannual ophthalmic examination, and the New Zealand Association of Optometrists recommends a regular examination every 2–5 years for healthy adults. However, like other developed countries, New Zealand does not have a formal screening program for glaucoma [4]. Diagnosis and subtype classification is based on intraocular pressure measurement, gonioscopy, and the presence of both structural and functional evidence of glaucomatous optic neuropathy [5], [6]. Intra and inter-observer settlement in the detection of *optic disc* (OD) abnormalities could be improved with stereo fundus images [7]. See Fig. 1 for examples of such *stereo OD images* (SODIs).

The general definition of abnormalities for glaucoma aspects is as follows:

1) increased cup-to-disc ratio (i.e. the ratio of the optic cup in the central optic disc to the margin of the optic disc),
2) notching of the neuroretinal rim (focal loss of the margin of the optic disc margin),
3) symmetry of optic disc cupping, and
4) loss of retinal nerve fiber layer.

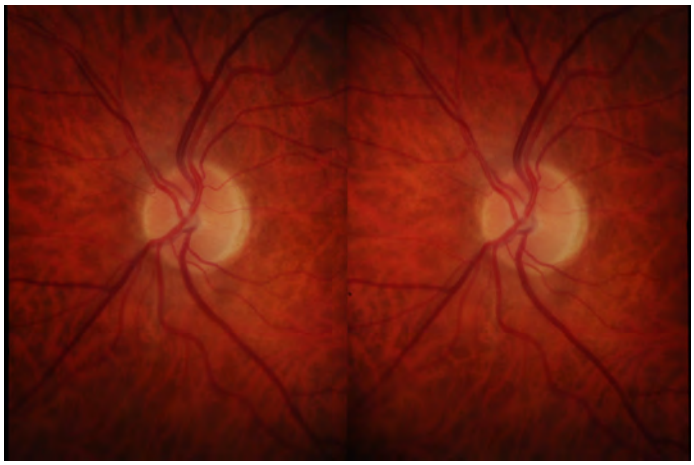
We assessed optic discs from patients with different subtypes of glaucoma defined by abnormal SODIs. The RIM-ONE, release three, is a public SODIs dataset that consists of 85 healthy and 74 glaucoma stereo images. It illustrates the accurate *optic nerve head* (ONH) gold standard for professionals in this field. Figure 1a shows a stereo image pair of normal optic discs, and Fig. 1b shows optic discs with glaucomatous optic neuropathy. Glaucoma manifests as ONH cupping; clinically this is most easily recognised at the superior and inferior poles of the optic disc as well as a focal and diffuse retinal nerve fiber layer loss [5], [6].

Stereo vision extracts 3D information from multiple 2D views of a scene. Stereo vision works by using a stereo-matching algorithm which finds corresponding points in a stereo image pair and produces a disparity map. This map encodes the difference in horizontal coordinates of corresponding image points. The values in the disparity map are inversely proportional to the scene depth at the corresponding pixel location [8]. Up-to-date stereo matching is robust (due to progress in matching algorithm design) and fast (as it only processes one stereo image pair). Many modern applications, such as advanced driving assistance system or robot navigation, work on principles of stereo vision to estimate the actual distance

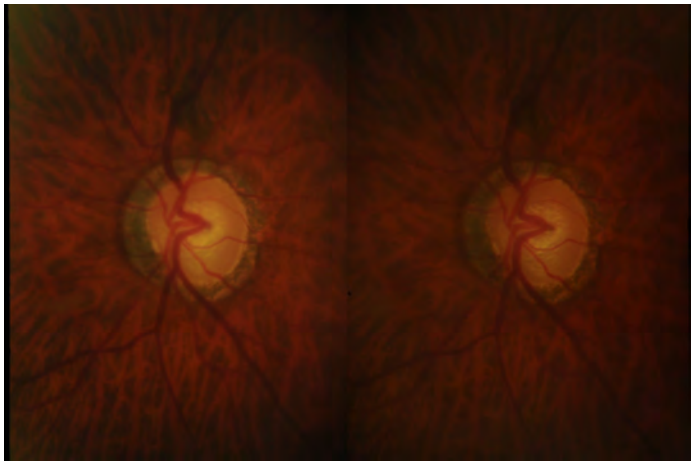
or the range of objects of interest relatively to the camera [9]. In this paper, we apply a stereo-matching algorithm to our input SODIs to assess the difference in depth values for abnormalities compared to OD. We use a commercial stereo-vision system SP1 which produces disparity maps for input SODIs.

The process of partitioning a digital image into multiple segments is called *image segmentation*. These multiple segments are sets of pixels which are also known as *superpixels*. Pixels that share certain characteristics such as similar colors, intensities or gray-levels are grouped together to form superpixels [14]. By doing so, image segmentation simplifies the representation of an image and makes it easier to analyze [15]. The *simple linear-iterative clustering* (SLIC) algorithm is one of the widely used algorithms. It efficiently generates compact, nearly uniform-sized superpixels by clustering pixels in a combined five-dimensional color and image coordinate space [16]: "SLIC produces superpixels at a lower computational cost while achieving a segmentation quality equal to or greater than four state-of-the-art methods, as measured by boundary recall and under-segmentation error [17]." In this paper, we apply the SLIC algorithm to divide input SODIs and disparity maps into distinct areas to estimate the abnormalities on OD. We also apply the *simple non-iterative clustering* (SNIC) algorithm on SDFIs and disparity maps; SNIC is an improved version of SLIC. Results of SLIC and SNIC algorithms are compared at the end.

The reminder of this paper is structured as follows: Section II presents the used public dataset and the SP1 stereo vision system that is used in this research. In Section III, all steps of the methodology are explained such as stereo matching and superpixel segmentation. Section IV presents the results of stereo matching and superpixel segmentation. Section V concludes.



(a) Stereo fundus images of a healthy person



(b) Stereo fundus images of a glaucoma patient

Fig. 1: Stereo fundus images; courtesy of [12], [13].

II. Materials

The SP1 stereo vision system has been used in this research; it is a product of Nerian Vision Technologies (NVCom); see [11]. It performs stereo matching in real time using a semi-global matching algorithm. It is a stand-alone processing system with an integrated *field programmable gate array* (FPGA) which produces a dense disparity map. The system can be configured through a web interface using its IP address. A gigabit ethernet connection is established between the SP1 and a client computer for the transmission of images. The NVCom software is required to send input stereo images to the SP1 and to display and write the received images from the SP1. The system can process the images from two industrial USB cameras in real time, or it can also generate disparity maps for input stereo images transmitted from a client computer. The SP1 can process images with a resolution from 320×240 pixels up to $1,440 \times 1,440$ pixels; the number of calculated disparities ranges from 32 to 256 pixels. It can also reconstruct the 3D location of corresponding scene points from the disparity map [11].

The third version of the RIM-ONE [12] dataset has been used in this research. It contains 159 SODIs with a resolution of $2,144 \times 1,424$ pixels. Three hospitals, namely the Hospital Clìnico San Carlos, the Hospital Universitario Miguel Servet, and the Hospital Universitario de Canarias, had contributed to the development of this database [13].

III. Methodology

The flow chart in Fig. 2 represents our approach for the detection of abnormalities on ODs in SODIs. At first, input stereo fundus images are separated in left and right images. Then, the left and right images are cropped in order to have a region of interest (ROI) accurately. These two images are given to the SP1 system to calculate the disparity map. Afterwards, in two stages, the disparity map and the left image are entered into SNIC and SLIC algorithms to achieve the segmented masks. Finally, SNIC and SLIC masks are overlaid on the input left image as outputs; see Fig. 6. Current segmentation studies, as reviewed in [10], do not yet cover such types of superpixel segmentations.

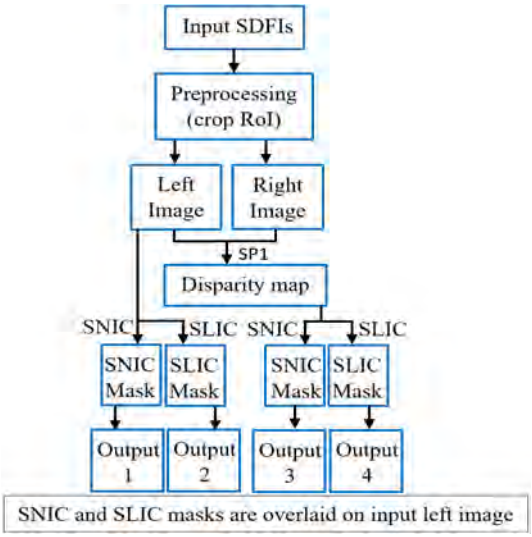


Fig. 2: Overall approach for segmentation of abnormalities on optic discs.

The concepts of stereo vision and superpixel segmentation for glaucoma detection are discussed below in the rest of this section; processes applying this flow chart are explained in the following sections.

A. Stereo Vision

We cite [18]: "Stereopsis is a term that is most commonly used to refer to the perception of depth and 3-dimensional (3D) structure obtained on the basis of the visual information deriving from two eyes by individuals with normally developed binocular vision"; the difference in the relative horizontal position of objects in the two images is referred to as binocular disparity. The visual cortex of the brain processes disparities to yield depth perception.

Stereo vision is a well-known ranging method because it resembles the basic mechanism of the human eye. Computer stereo-vision systems use the same principle by replacing eyes with two CCD cameras. They are displaced horizontally to obtain two different views. A disparity map, which encodes the difference in horizontal coordinates of corresponding image points, can be obtained by comparing these two slightly different images. The values in the disparity map are inversely proportional to the scene depth at the corresponding pixel location [8].

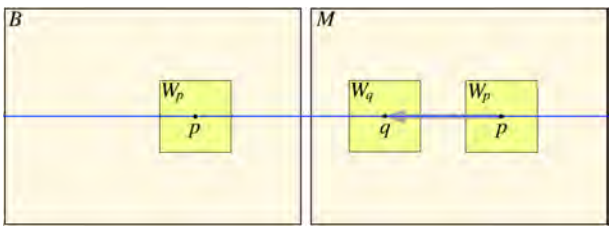


Fig. 3: Left and right camera image for canonical stereo geometry [19].

As stated above, two CCD cameras replace human eyes in computer stereo vision system. So, they should be as identical as possible for avoiding difficulties. In *canonical stereo geometry*, both the cameras have identical effective focal lengths and they should be arranged in such a way that they have parallel optic axes. If the cameras are not arranged in canonical stereo geometry, then both images should be rectified before performing stereo matching [19].

Figure 3 illustrates matching in canonical stereo geometry; both the images have collinear rows y, defining the *epipolar lines*. Here, the left image is the base image indicated by B, and the right image is the match image indicated by M. The projection of a 3D world point P in the base image is displayed by pixel $p = (x, y)$. Now, we must search for a corresponding pixel on the same epipolar line in the match image M. The two pixels are corresponding if they are projections of the same point P in the scene.

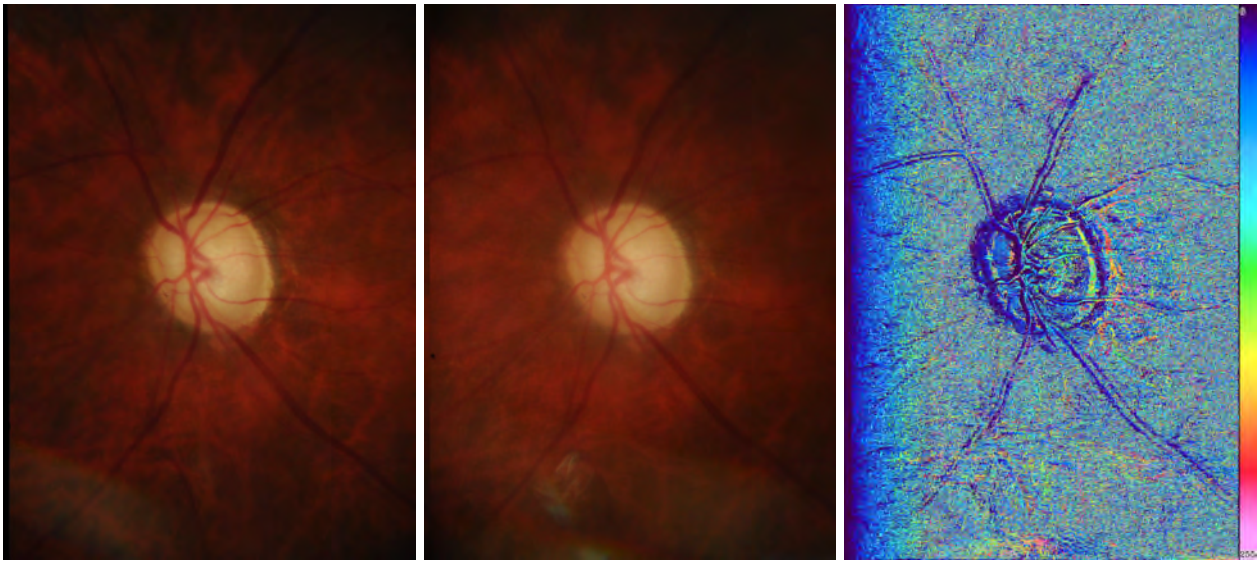
Let pixel $q = (x + d, y)$ be the corresponding pixel of pixel p. Here, d is the value of the disparity which is plotted in a disparity map. A disparity map encodes the difference in horizontal coordinates of corresponding image points. We can generate the 3D world coordinates for each pixel by processing the disparity map [19]. Figure 4 shows an input stereo pair and a resulting disparity map; we use a color key for visualising integer disparities.

The first step in our experiments is to perform stereo matching. We use the SP1 stereo-vision system to produce a disparity map for the input fundus images. The calculated disparity map describes a mapping of image points from the left camera image to corresponding image points in the right camera image (i.e. the SP1 produces a disparity map from the perspective of the left camera image). Corresponding pixels in both images should only differ in their horizontal coordinates, corresponding to canonical stereo geometry. The produced disparity map encodes a horizontal coordinate difference [20].

The input image data is transmitted from a client computer to the SP1 for stereo matching. The system performs stereo matching using *semi-global matching* (SGM) algorithm and produces a disparity map with a bit depth of 8 bits per pixel. The NVCom application is used to receive and display disparity maps from the SP1 [20]. Results for tested 3D scenes verified the accuracy of the SP1; the followed SGM algorithm has been developed for accurate, pixel-wise matching at low run time [21]. It combines concepts of global and local stereo methods [22].

B. Simple Linear Iterative Clustering Algorithm

Superpixel algorithms group pixels into perceptually meaningful atomic regions. These pixels should adhere well to image boundaries [17]. The SLIC algorithm produces a desired number of regular, compact and nearly uniform superpixels with a low computational overhead. The SLIC algorithm is very simple, fast and efficient, which makes it extremely easy to use. [16] documented that this algorithm achieved a segmentation quality better than four state-of-art methods (at that time).



(a) Input left image (b) Input right image (c) Disparity map

Fig. 4: Input stereo pair and resulting disparity map (used color code is shown on the right).

SLIC adapts a *k*-means clustering approach to efficiently generate superpixels [23]. There are two important distinctions compared to general *k*-means clustering:

1) The search space for the number of distance calculations in the optimization is limited to a region proportional to the superpixel size.
2) The color and spatial proximity are combined by a weighted distance measure. It also provides control over the size and compactness of the superpixels [17].

SLIC has only one parameter k which is the desired number of approximately equally sized superpixels.

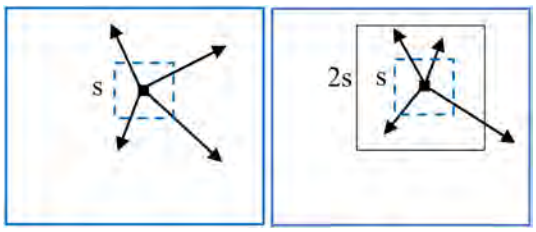


Fig. 5: Search regions of *k*-means and SLIC algorithms [17].

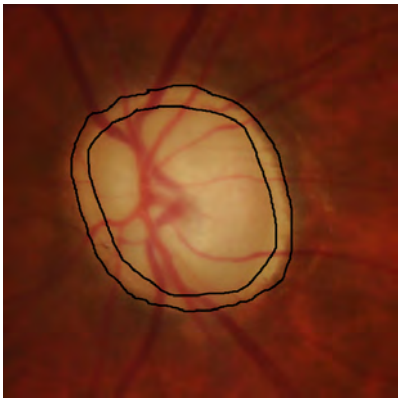
For color images in the CIELAB color space, the first step of the clustering procedure is to sample k initial cluster centers $\mathbf{c}_i$=[l_i, a_i, b_i, x_i, y_i]$^\top$ on a regular grid, spaced s pixels apart. Here, $s = \sqrt{\frac{N}{k}}$ for roughly equally sized superpixels. To avoid centering a superpixel on an edge, the centers are transferred to seed locations corresponding to the lowest gradient position in a 3×3 neighborhood. The next step is to assign each pixel at location p to the nearest cluster center whose search region overlaps its location. As depicted in Fig. 5, the size of the search region is limited to reduce the number of distance calculations, in contrast to general *k*-means clustering. Since the expected spatial extent of a superpixel is a region of approximate size $s \times s$, the search for similar pixels is done in a region $2s \times 2s$ around the superpixel's center. The last step is to adjust the cluster centers to be the mean vector $[l, a, b, x, y]^\top$ of all the pixels belonging to the cluster. The residual error E between the new and previous cluster centers is calculated. The last two steps of algorithm are repeated iteratively until the error converges [17]. Finally, a post processing step reassigns isolated pixels to nearby superpixels to enforce connectivity [17], [25].

Superpixel partitioning provides accurate boundaries around different tissues which makes the extraction of image features easier [24]. In this paper, superpixel segmentation is used to reduce the negative impacts of vessels on abnormalities detection. Another advantage of using superpixel segmentation is that it separates a background region in an image from the foreground. Moreover, it groups the abnormalities on OD in one or more clusters which helps in glaucoma detection.

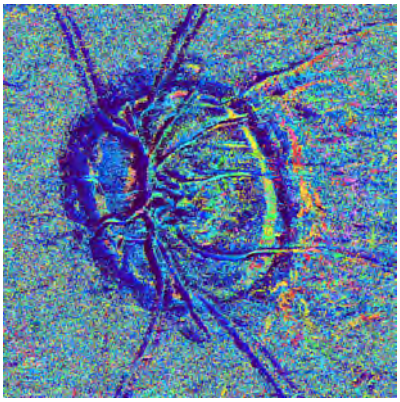
IV. Experiments and Results

The first step of this experiment is to perform stereo matching on to the input dataset. The SODIs are cropped to ensure that the left and right camera images are of identical size. The size of the left and right camera images is set to $1,056 \times 1,424$ pixels. The stereo fundus images of glaucoma patients are then transferred from a client computer to the SP1 through gigabit ethernet. The SP1 computes the disparity map and transmits it back to a client computer [20].

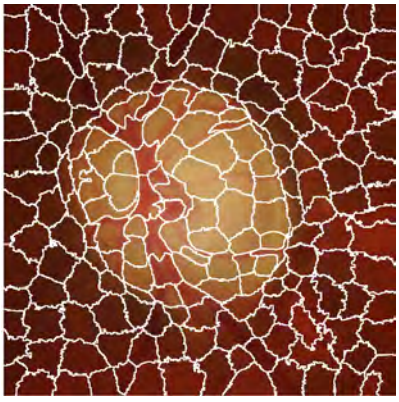
For computing the disparity map, the value of maximum number of disparities is set to 256. For generating disparity map, the SP1 uses intrinsic and extrinsic calibration parameters. The value of reprojection error is inversely proportional to the accuracy of the disparity map. For our experiments here, the value of the reprojection error was 0.06. The SP1 produces a disparity map from the perspective of the left camera image.



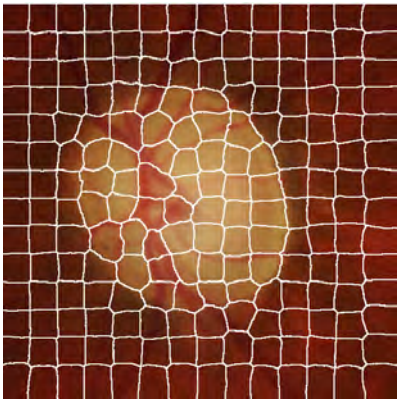
(a) ROI in input left fundus image



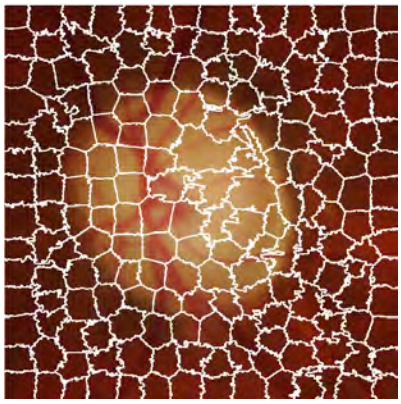
(b) ROI in disparity map



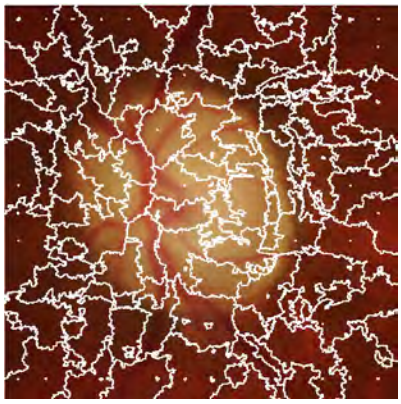
(c) Input image segmentation using SLIC



(d) Input image segmentation using SNIC



(e) Disparity map segmentation using SLIC



(f) Disparity map segmentation using SNIC

Fig. 6: Results after applying SLIC and SNIC algorithms.

A rainbow colormap is applied on the disparity map for better visualization. As we go from blue color to pink color, the disparity value increases gradually. See Fig. 4 for an example of an output disparity map for one SODI.

The original left image or the produced disparity map are now considered as inputs for the SLIC and SNIC algorithms. As the abnormalities are present on or near the OD, we crop the input left camera image and the output disparity map into a *region of interest* (ROI) around the OD.

The OD shown in Fig. 6a has advanced glaucoma as the inner line is close to the outer line indicating a loss of tissue between both lines; the tissue between the two lines is called the *neuroretinal rim*.

Figure 6b is a cropped disparity image. The size of the cropped images is uniformly 585×585 pixels. The SLIC and SNIC methods segment the images and produce boundary masks. Both algorithms are applied for the same number k (we decided for $k = 200$) of superpixels for comparison purposes.

The value of compactness in the SNIC algorithm is set to 40. If the number of superpixels increases (above 200), there are high chances that the clustering may lead to the production of regions having black holes or dots (showing non-simple region topologies).

The computed SLIC and SNIC boundary masks of the input fundus image and also of the disparity map are then overlaid on the original (cropped) input left image for a better comparative visualization.

As a result, four output images are generated for each SODI. Figures 6c and 6d show the segmented input fundus image using SLIC and SNIC algorithms, respectively. Resulting images, after overlaying the boundary masks of segmented disparity maps onto original input images, are shown in Figs. 6e and 6f. Figure 6e illustrates the result of applying the SLIC algorithm, and Fig. 6f is the result of applying SNIC algorithm.

On the processed SODIs, SLIC and SNIC algorithms grouped abnormalities on the OD into one or more superpixels without generating superpixels containing both (at significant percentage) pixels of abnormal tissue and of normal tissue.

As a result, finding abnormalities in these clustered images is easier compared to the fine structure of the pixel grid. The four types of segmented output images are suggested for being analyzed by the specialist. A subsequent large-scale study needs to evaluate the efficiency of the proposed procedure for detecting abnormalities (and, eventually, glaucoma).

V. Conclusions

An image segmentation method is proposed for SODIs combining two superpixel algorithms (SNIC and SLIC) and disparity map calculation, resulting into four kinds of segmentations of optic discs. The goal is to support the early detection of (small) abnormalities.

According to the ophthalmologist, results as shown in Figs. 6c and 6d appear to be promising due to the following factors:

1) ability to identify the margin of the optic disc,
2) some outlining of blood vessels, and
3) outlining of the optic cup (the center depression of the optic disc).

Results of the SLIC disparity mask also appear appropriate and the presented evaluations on fundus images show encouraging results. From our experiments we can conclude that disparity values for abnormalities are higher than the disparity values for normal parts (Fig. 4c). This change in disparity values can also be further studied towards abnormality segmentation.

In future work, we focus on the development of a system to eliminate the healthy segments of a disparity map image to achieve improved abnormality detection.

References

[1] New Zealand: Latest stats at a glance, see www.blindfoundation.org.nz, 2015.

[2] World health organization. “Global data on visual impairments 2010,” Geneva: World Health Organization Organization, 2012.

[3] Y. C. Tham, X. Li, T. Y. Wong, H. A. Quigley, T. Aung, and C. Y. Cheng, “Global prevalence of glaucoma and projections of glaucoma burden through 2040: A systematic review and meta-analysis,” Ophthalmology, 121(11):2081–2090, 2014.

[4] B. R. LaHood, J. Erceg, T. H. Bevin, and G. Sanderson, “High rate of incidental glaucoma detection in New Zealand,” Ethnicity, 25, p. 7, 2016.

[5] P. J. Foster, R. Buhrmann, H. A. Quigley, and G. J. Johnson, “The definition and classification of glaucoma in prevalence surveys,” British J. Ophthalmology, 86(2):238–242, 2002.

[6] P. Sharma, P. A. Sample, L. M. Zangwill, and J. S. Schuman, “Diagnostic tools for glaucoma detection and management,” Survey Ophthalmology, 53(6):S17–S32, 2008.

[7] R. Varma, W. C. Steinmann, and I. U. Scott, “Expert agreement in evaluating the optic disc for glaucoma,” Ophthalmology, 99(2):215–221, 1992.

[8] “Stereo vision for depth estimation 2018,” The Mathworks, Inc., see www.mathworks.com/discovery/stereo-vision.html

[9] Stereo matching 2018, see www.devpy.me/what-is-stereo-matching/

[10] A. Almazroa, R. Burman, K. Raahemifar, and V. Lakshminarayanan, “Optic disc and optic cup segmentation methodologies for glaucoma image detection: A survey,” J. Ophthalmology, 42: 162–189, 2015.

[11] K. Schauwecker, “SP1 stereo vision system,” Nerian Vision Technologies, Germany, 2017.

[12] F. Fumero, S. Alayon, J. Sanchez, J. Sigut, and M. Gonzalez-Hernandez, “Rim-one: An open retinal image database for optic nerve evaluation,” Proc. IEEE Int. Symp. Computer-based Medical Systems, doi: 10.1109/CBMS.2011.5999143, 2011.

[13] Medical image analysis group, University of La Laguna, Spain, medimrg.webs.ull.es/research/retinal-imaging/rim-one/, access date: 20 August, 2018.

[14] L. Shapiro and G. Stockman, “Computer Vision,” pp. 279–325, Pearson, 2001.

[15] B. Fulkerson, A. Vedaldi, and S. Soatto, “Class segmentation and object localization with superpixel neighborhoods,” Proc. IEEE Int. Conf. Computer Vision, 2009.

[16] SLIC superpixels 2017, see ivrl.epfl.ch/research/superpixels

[17] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, ”SLIC superpixels compared to state-of-the-art superpixel methods,” IEEE Trans. Pattern Analysis Machine Intelligence, 34(11):2274–2282, 2012.

[18] I. Howard and B. Rogers, ”Binocular vision and stereopsis”, New York: Oxford University Press, 1995.

[19] R. Klette, “Concise Computer Vision”, pp. 287–330, Springer, London, 2014.

[20] “SP1: Real-time 3D stereo vision through FPGA technology, 2018,” Nerian Vision Technologies, see nerian.com/products/sp1-stereo-vision/

[21] H. Hirschmüller, ” Stereo processing by semiglobal matching and mutual information”, IEEE Trans. Pattern Analysis Machine Intelligence, 30(2): 328–341,2008.

[22] D. Scharstein and R. Szeliski, “A taxonomy and evaluation of dense two-frame stereo correspondence algorithms”, Int. J. Computer Vision, 47(1-3):7–42, 2002.

[23] C. Zitnick and S. Kang, “Stereo for image-based rendering using image over-segmentation,” Int. J. Computer Vision, 75:49–65, 2007.

[24] M. Soltaninejad, G. Yang, T. Lambrou, N. Allinson, T. L. Jones, T. R. Barrick, F. A. Howe, and X. Ye, “Automated brain tumour detection and segmentation using superpixel-based extremely randomized trees in FLAIR MRI,” Int. J. Computer Assisted Radiology Surgery, 12(2):183–203, 2017.

[25] P. Felzenszwalb and D. Huttenlocher, “Efficient graph-based image segmentation,” Int. J. Computer Vision, 59(2):167–181, 2004.