2014

3D papillary image capturing by the stereo fundus camera system for clinical diagnosis on retina and optic nerve

Proceedings of SPIE,Bellingham : International Society for Optical Engineering - SPIE,v. 8936, p. 893614-1-893614-8, 2014

http://www.producao.usp.br/handle/BDPI/50196

3D papillary image capturing by the stereo fundus camera system for clinical diagnosis on retina and optic nerve

Danilo A Motta[a],André Serillo[a], Luciana de Matos[a,b], Fatima M M Yasuoka[a,b], Vanderlei Salvador Bagnato[b], Luis A V Carvalho[a,b]
Wavetek Technologies Industry LTDA, São Carlos, SP, Brazil[a]
IFSC – Universidade de São Paulo (USP), São Carlos, SP, Brazil[b]

ABSTRACT

Glaucoma is the second main cause of the blindness in the world and there is a tendency to increase this number due to the lifetime expectation raise of the population. Glaucoma is related to the eye conditions, which leads the damage to the optic nerve. This nerve carries visual information from eye to brain, then, if it has damage, it compromises the visual quality of the patient. In the majority cases the damage of the optic nerve is irreversible and it happens due to increase of intraocular pressure. One of main challenge for the diagnosis is to find out this disease, because any symptoms are not present in the initial stage. When is detected, it is already in the advanced stage. Currently the evaluation of the optic disc is made by sophisticated fundus camera, which is inaccessible for the majority of Brazilian population. The purpose of this project is to develop a specific fundus camera without fluorescein angiography and red-free system to accomplish 3D image of optic disc region. The innovation is the new simplified design of a stereo-optical system, in order to make capable the 3D image capture and in the same time quantitative measurements of excavation and topography of optic nerve; something the traditional fundus cameras do not do. The dedicated hardware and software is developed for this ophthalmic instrument, in order to permit quick capture and print of high resolution 3D image and videos of optic disc region (20° field-of-view) in the mydriatic and nonmydriatic mode.

Keywords: Glaucoma, optic nerve, retina. 3D papillary image, stereo fundus camera, mydriatic and nonmydriatic mode.

1. INTRODUCTION

According to Quigley and Broman, Glaucoma is the second main cause of blindness worldwide and is characterized by an increasing number of cases due to the raise of lifetime expectation of the world population [1]. Glaucoma happens due the damage of optic nerve and in this condition compromises the quality visual of the patient. In the most of cases the damage of optic nerve is irreversible and it occurs due to the intraocular pressure increase. One of the problems of this disease is that in its initial stage is asymptomatic [2]. The majority of the people do not know that they have the disease and when find out it; it is already in the advanced stage. For this reason is fundamental the previous diagnosis to preserve the quality of vision.

The diagnosis is made by several exams; one of the principal exams is the detailed fundus eye, which the optic nerve is evaluated. By meticulous exam of the optic disc and also the nerve fiber of the retina, related to the others exams, such as intraocular pressure and visual field measurement, it is possible to diagnose previously the glaucoma, without the significative visual lost occurs.

Currently there are three techniques for the optic nerve evaluation: retinography (digitalized images of the optic disc), optic disc topography (three-dimensional scanning laser image of optic disc) and OCT (Optical Coherence Tomography that provides the high resolution image of the transversal section of retina, including optic nerve and macula). But all the instruments, that realize this type of the exam, contain the modules as fluorescein angiography, indocyanine green angiography, red free, and others, which require a sophisticated optical system; hence these instruments become very expensive.

The objective of this work is to develop the specific and simplified fundus camera in order to realized three-dimensional image of the papillary region (optic disc), without fluorescein and indocyanine green system or red free filter. The main innovation of this fundus camera is the utilization of the stereoscopic optical system, which enables the capture of three-dimensional image and in the same time quantitative measurements of the excavation and topography of the optic nerve, something the traditional fundus cameras do not do. This is innovator equipment intended for the general practice of

Design and Quality for Biomedical Technologies VII, edited by Ramesh Raghavachari, Rongguang Liang,
Proc. of SPIE Vol. 8936, 893614 · © 2014 SPIE · CCC code: 1605-7422/14/$18 · doi: 10.1117/12.2038435

ophthalmology in the Brazilian public health system. Actually the majority of the fundus camera disposable in the market uses sophisticated optical system and has several characteristics as angular length, angiography, red free and others; which make them very expensive for the general ophthalmology practice. For this reason, it is important to develop instruments that are quick, efficient and economics in order to attain the huge public of the ophthalmology centers, mainly of Brazilian public health system.

In this study, a system capable of reconstructing the quantitatively and three-dimensionally images of the optic disc, quickly and efficiently, is being developed. The method utilized is a cross-correlation algorithm applied to stereoscopic images of three objects simulating a human retina, with different diameters and depths. Depth information was retrieved using the triangulation method (see Figure 6).

The measurements in central sections of the three-dimensional reconstruction of the three objects showed absolute maximum depth error between peaks < 1.58mm (see Figure 7).

The method applied to a real case shows great precision in the three-dimensional reconstruction and computational efficiency (≤ 60s).

2. METHODOLOGY

2.1 Optical system of stereoscopic fundus camera for the 3d pupillary image capture

The optical system for the stereoscopic image capture is being developed, using software of optical design, which is illustrated by the Figure 1.

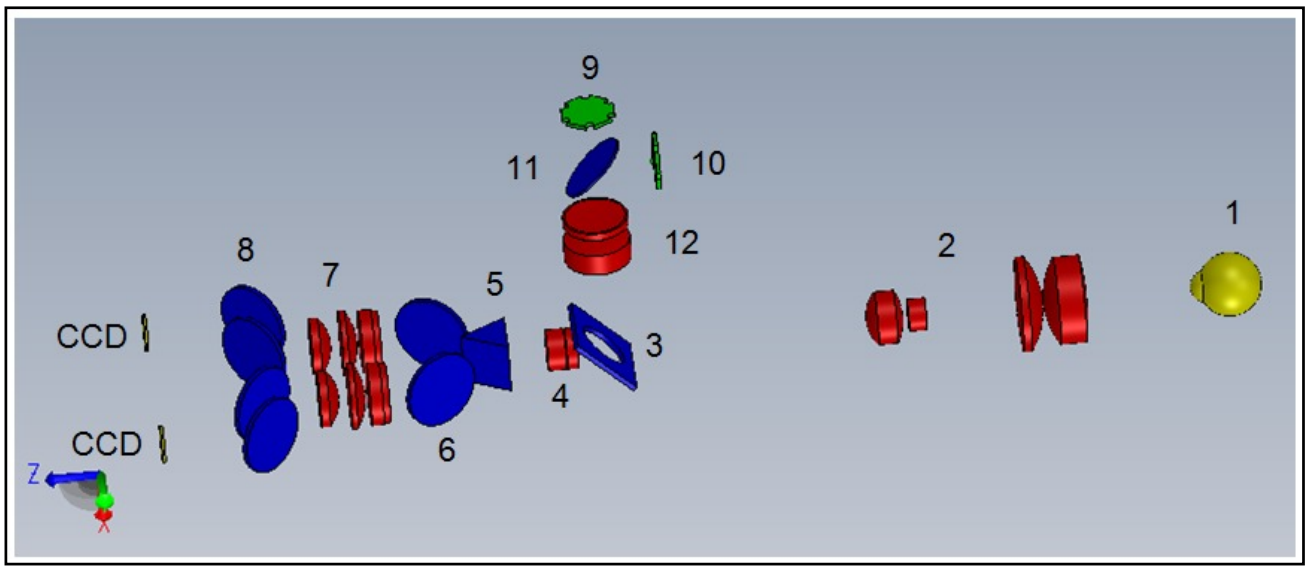

Figure 1: Optical system for the stereoscopic 3D image capture.

In the right side of the Figure 1, the patient eye is indicates for number 1, the image capture system of fundus eye is 2, the mirror with hole is 3, afocal doublet is 4, the beamsplitter prism is 5, the stereoscopic system of image focalization on CCDs is 7. After the beamsplitter prism 5, the optical system is divided in two parts in order to obtain the stereoscopic image, which is focalized in two different CCDs. The mirror with hole 3 has the purpose to reflect the beam coming from the LED 9 at 45 degree in order to illuminate the fundus eye by the optical system 2, and transmit by the hole the image captured of fundus eye to focalize in optical system 7. There is two LEDs, one of them covers the visible region 9, that is responsible for the illumination of fundus eye and the other, infrared LED 10 to help in the alignment of entire system related to the patient's eye. The beamsplitter 11 is responsible to transmit the beam of LED 9 in the visible region and reflect at 45° the beam of infrared LED 10. The infrared LED (850 nm) is used to help during the alignment of the system related to patient's eye, because it is less inconvenient to eye. The illumination system 12 helps on the

focalization of the beam 9 to illuminate the fundus eye, as well as the LED 10 helps in the alignment. The mirrors 6 and 8 are used to deflect the stereoscopic beams in order to facilitate the mechanical design of the equipment.

2.2 Stereoscopic image processing

The first step of this work is the acquisition of the stereoscopic image by the fundus camera developed (see the section 2.1). This system obtain synchronous form of the stereoscopic image of the pappilary region, then the stereoscopic pair is simultaneously obtained, in order to avoid problems, like as the patient or camera moving, which damage the quality of the results.

These images are result of 2D image mapped from de 3D scene. This map is a transformation of several points of the scene to the same pixel of the image. Stereoscopic image techniques can be used to recover the information of the depth lost during the mapping process of the image.

In the general reconstruction processes, it is necessary firstly the preprocessing step to remove the difference between the two images [3], in this method fortunately it is not necessary. It is used the adaptation of an algorithm of stereoscopic correspondence. Basically this algorithm begins with a pixel in an image to find out the correspondent pixel in the other image and the distance between theses pixels provides information, which is necessary for the depth reconstruction. In this adaptation the quantity of pixels related to the images is decreased in order to obtain a viable computational method to be executed to attend the ophthalmological protocol. The depth in each pixel is calculated from the map of correspondence, using the triangulation method (see Figure 2), which considers only the map calculated and the parameters of the cameras calibration [4]. For simplicity it is adopted a stereoscopic system where the plane of image of the cameras is parallel, which does not accept rotation or vertical translation between the cameras. Hence the epipolar lines cut the images in lines perfectly horizontal and the projection centers of the cameras are positioned in the same height. An approach to find out the depth is by triangulation method. Be the parallel stereoscopic system illustrate for the Figure 2, the depth recovery is given by:

$$Depth = \frac{L\ x\ tan(\theta-\beta_L)\ x\ tan(\theta+\beta_R)}{tan(\theta-\beta_R)+tan(\theta+\beta_R)} \quad (1)$$

$$\beta_L = \arctan\left(x_L\ x\ tan\left(\frac{\alpha}{2} x \frac{\pi}{180}\right) x \frac{2}{W}\right. \quad (2)$$

$$\beta_R = \arctan\left(x_R\ x\ tan\left(\frac{\alpha}{2} x \frac{\pi}{180}\right) x \frac{2}{W}\right. \quad (3)$$

$$x_R = x_L + disparity \quad (4)$$

where x_L and x_R are the horizontal coordinates of the corresponding points on the left and right images, respectively; the origin of the coordinate systems is the right and left points on the optical axis; α is visual angle of the cameras; θ is the angle of optical axis; β is the angle between the corresponding point position and the optical axis; W is the image width and L is the baseline length. The baseline is the distance between the two cameras. The disparity is the propriety that measure the distance between two projection points of the cameras in the Euclidian space.

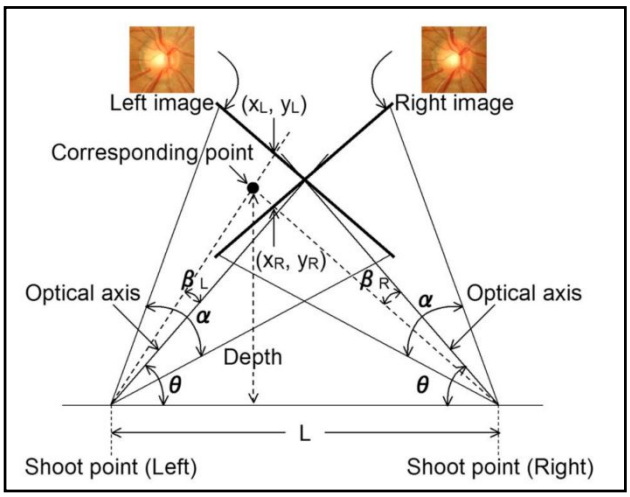

Figure 2: Parallel stereoscopic system [4].

This method is advisable because does not depends of the several parameters, solving the same problem stylishly. Related to the cameras is important to observe that is essential to watch out its calibration. The focus system must be adjusted to the focal point in order to be positioned between the camera and the scenery. So that the stereoscopic pair possesses more disparity in sceneries with less depth and sceneries with more depth have less disparity.
The approach used in this work is based on the Nakagawa's method [4], although the main modifications is introduced according to can be seen in the following sections.

2.3 Section of interesting region

In this subsection it is presents two algorithms for the section of interesting region in the fundus eye images. One is developed in this work and the other presented by Nakagawa et al [4]. The algorithm developed is simple but effective. The objective is to cut the specific region of the image, and then the algorithm is made considering the fundus eye image as the entrance and the optic disc as the target. This algorithm is showed in the Figure 3.
In this algorithm, the calling to the bounding boxes is a reference to obtain the involved region of each group of the pixels, which have interface each other. Obviously the greater agglutinated region corresponds to the optic disc. Results of this approach can be seen in the Figure 4.
The algorithm showed by Nakagawa et al. is described in the Figure 5. Where the calling to P-tile [7] is an algorithm that tries to obtain a threshold to the image, so that 50% of the pixels are divided using a distribution of the frequency intensities. That it, the image is divided in two significative groups. But to perform this task is necessary a preprocessing step to remove the blood vases of the image, which increases the complexity of the problem as seen in [4].

Data: Fundus Image I
Result: Bounding box of the interest region
1 $R \leftarrow$ **RedChannel**(I);
2 $t \leftarrow$ **OtsuThreshold**(R);
3 $IBinary \leftarrow R > t$;
4 $boxes \leftarrow$ **BoundingBoxes**$(IBinary)$;
5 $i \leftarrow greater(area(boxes))$;
6 **return** $boxes(i)$;

Figure 3: Algorithm 1 developed in this work.

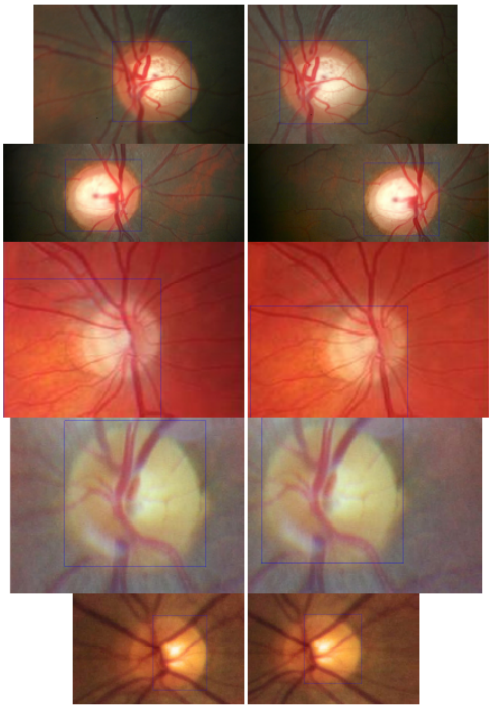

Figure 4: Section of the interesting region.

Data: Fundus image after a preprocessing step to remove the blood vases I'
Result: Bounding box of the interest region
$R \leftarrow$ RedChannel(I');
$G \leftarrow$ GreenChannel(I');
$B \leftarrow$ BlueChannel(I');
$R_{p-tile} \leftarrow$ P-tile(R);
$G_{p-tile} \leftarrow$ P-tile(G);
$B_{p-tile} \leftarrow$ P-tile(B);
$boxes \leftarrow (R_{p-tile} \ \& \ G_{p-tile}) \| (G_{p-tile} \ \& \ B_{p-tile}) \| (R_{p-tile} \ \& \ B_{p-tile})$;
$i \leftarrow greater(area(boxes))$;
return $boxes(i)$;

Figure 5: Algorithm 2 of the section of interesting region by Nakagawa et al [4].

2.4 Stereoscopic correspondence

In the problem of stereoscopic correspondence it is necessary to find out which candidate pixels of an image corresponds to the reference pixels of the other image. The Chebyshev distance is used as a mensure between the reference pixel and correspondent marked pixel, which gives the information for the three-dimensional reconstruction. Next we describe an algorithm that finds this correspondence.
Following, it is described the algorithm of the stereoscopic correspondence reported in the reference [4] which was modified and adopted in our pipeline. The detection of the correspondent pixels is made initially selecting the interesting regions around the reference point and the candidate point separately. Two points in the right and left images, with the similar texture in its interesting regions respectively, are giving as the correspondent points. The similarity is measured using the cross-correlation coefficient that follows below:

$$r = \frac{\sum_{i=-W/2}^{W/2}\sum_{j=-H/2}^{H/2}\{L(x_L+i\,,y_L+j)-L\}\{L(x_R+i\,,y_R+j)-R\}}{\left(\left[\sum_{i=-W/2}^{W/2}\sum_{j=-H/2}^{H/2}\{(L(x_L+i\,,y_L+j)-L\}^2\right]^{1/2}\left[\sum_{i=-W/2}^{W/2}\sum_{j=-H/2}^{H/2}\{(R(x_R+i\,,y_R+j)-R\}^2\right]^{1/2}\right)} \quad (5)$$

The same rectification process, using the colors channels and Sobel filters, are used as entrance for the correlation. The pixel that has the maximum cross-correlation value is marked as the correspondent point. If this correlation value is less than a threshold value (originally set to 0.5), is assumed that the reference pixel does not have correspondent pixel. This case also occurs if the interesting region does not have so much variance of the texture. Measuring the difference between the maximum and minimum value of the pixels and testing this value with other threshold value (originally set to 10), the texture has little variance and the pixel does not have correspondent. Then, it has the similarity given by the equation:

$$Similarity = \frac{C_R \cdot r_R + C_G \cdot r_G + C_B \cdot r_B + C_{R-Sobel} \cdot r_{R-Sobel} + C_{G-Sobel} \cdot r_{G-Sobel} + C_{B-Sobel} \cdot r_{B-Sobel}}{C_R + C_G + C_B + C_{R-Sobel} + C_{G-Sobel} + C_{B-Sobel}} \quad (7)$$

$$C = 1 \; if \; r \geq 0.5 \text{ and } [\max(L) - \min(L) \geq 10.$$

In the original algorithm the author expands the image in the horizontal direction in order to have an approach of the sub-pixel, but the algorithm has an excessive processing time, making impracticable the expansion of the image in a system that needs an answer in a viable time. To improve the performance, this step was removed in our implementation. The parameters used in the algorithm can be seen in the reference [4].
Finally, for the reduction of noise are used filters in the stereoscopic correlation result. There are five interactions of median filters 5x5 followed by movable mean filter 3x3. These filters, used in the implementation of the Nakagawa's method, were substituted by the filters based on the density (see section 2.5).

2.5 Filter DBSCAN

Noises and missmatchs are common in process of stereoscopic correlation. The Nakagawa's approach uses the filters that soft the result, which is not indicated, because the result depends directly of the disparity map, and this softy induces unnecessary errors. In this work, it is assuming that do not have noise agglomerations of same value, then these noises

are removed basing on in the neighborhood density. For this it is used the cluster algorithm called DBSCAN (Densitybased spatial clustering of applications with noise) [6].
This algorithm is based on in two points:

1. Inside of a cluster the density of the points is greater than out of the clusters.
2. Noise areas do not have the density of a cluster.

For this reason it is assumed that the algorithm of correspondence produces sparse errors.

When the data set *D* is given, it is assumed that this set can be cluster and at least has a two-dimensional point *p*, that called internal. The cluster of *p* consists in all the points that can be attained, called neighborhood, in a radius of *Eps* from *p*. That is, from the point *p* are found all the neighbors in a radius of *Eps* and from these neighbors the process is repeated until the greater number of possible points can be attained, defining a cluster. Following the process is repeated for the rest of the points. A point can be called internal if has a certain minimal number of neighbors *MinPts*. The points that have neighbor between *MinPts* and a determined point *C* are classified as the border and the rest are noise. In this case $D = (i; j; d)$, where *i* and *j* are the index of the point *p* in the image and *d* is its value of stereoscopic correspondence. Removing the points classified as the noise it has more precise results.

2.6 Reconstruction

The Figure 6 presents the depth calculated in each pixel presented in the Euclidian space. Finally the interpolated mesh of the pixels is generated, that demonstrates the depth of the images recovered. From this resultant mesh, other analyses generally in 2D sections can be made, that are used to facilitate the understanding of the obtained excavation conditions. This information helps to the ophthalmologist to evaluate the progress and treatment of the patient.

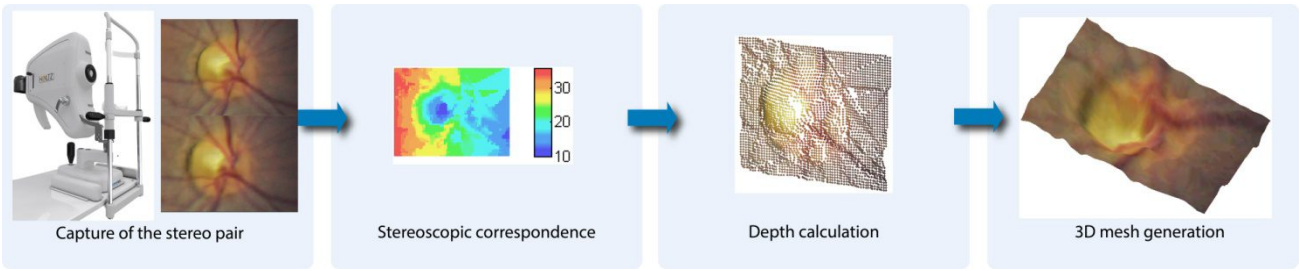

Figure 6: Flow chart of the stereoscopic method.

3. DATA

The Table 1 shows the data of errors of the results using the algorithm developed in this work.

Table 1: Data of errors of the results

Model	Mean error (mm)	Maximum error	Standard deviation	Error between peaks (mm)
1	3.1797	8.8502	2.97	1.5737
2	3.2743	9.1390	3.1976	1.1903
3	1.4072	2.6856	0.9709	0.0115

4. RESULTS

The results of the algorithm approach used are described in this section.

4.1 Reconstruction

For the quantitative tests of the reconstruction, it was developed a visual system composed by two cameras. Three objects modeled with different size of excavation and painted with simple patterns were realized in order to enable the correlation of the image.

Applying the algorithm developed in this work, it is obtained the results illustrated by the Figure 7. At left it is illustrated the molded objects with excavations of different size: deep, medium and shallows. Patterns simulating the fundus eye are painted in these models in order to have the comparison base between the stereoscopic pairs. In the center it is illustrated the mesh generated for each model and at right it is compared a section of the mesh generated (in green) and real (in blue), measured with high precision.

The results presented show that in a central section in the three-dimensional reconstruction and maximum depth error of excavation is less than 1.58 mm.

The depth calculus and the reconstruction of the three-dimensional surface of each model take less than 60 seconds, which is a viable time to the execution considering the system to help the ophthalmological diagnosis.

.

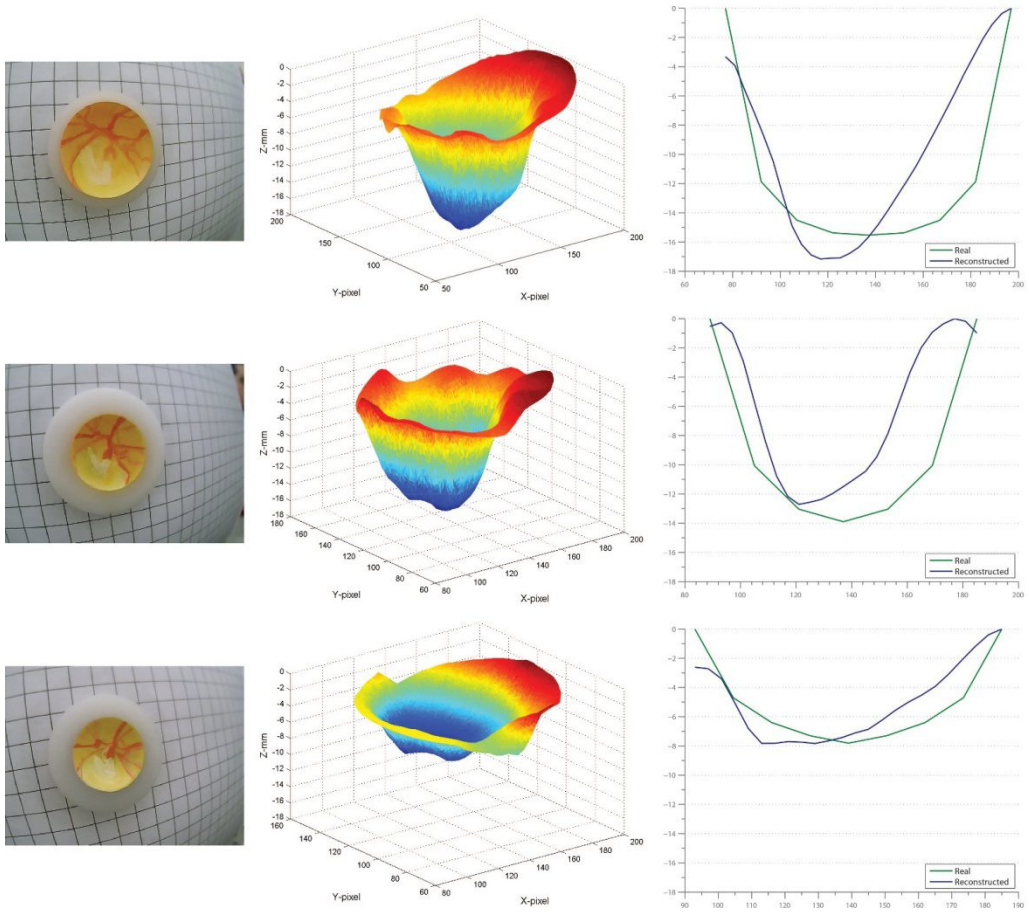

Figure 7: Mesh generated with the objects modeled in three different sizes.

5. CONCLUSIONS

The measurements in central sections of the three-dimensional reconstruction of the three different objects showed maximum depth error between < 1.58mm.

The method applied to a real case shows great precision in the three-dimensional image reconstruction and computational efficiency, that is, the time of processing is less than 60 s.

The results presented show the precision of the method developed, using the specific hardware of the fundus camera developed by Wavetek Technologies Industry and a simplified approach. In the real testes, the method reconstructs objects similar to the optic disc geometry, presenting the error in the maximum depth of the excavation less than 1.58 mm.

Therefore the method developed in this work attained its objectives; precision in the image reconstruction, viable execution time applied to the ophthalmological diagnosis and reduced cost for the entire equipment considering the necessities of Brazilian population.

Future works include the optimization of current method and the validation of the method with ophthalmological support, which is being realized with Ophthalmological Department of Universidade Federal de São Paulo.

6. REFERENCES

[1] Quigley, H. A. and Broman, A. T. "The number of people with glaucoma worldwide in 2010 and 2020," Br. J. Ophthalmol. 90(3), 262-267 (2006).

[2] Drance, S. M. "The early field defects in glaucoma," Invest. Ophthalmol. Vis. Sci. 8(1), 84-91 (1969)

[3] Zhang, Z., "Flexible camera calibration by viewing a plane from unknown orientations," Proc. of the 7th IEEE, International Conference on Computer Vision 1, 666–673 (1999).

[4] Nakagawa, T. and et al. "Quantitative depth analysis of optic nerve head using stereo retinal fundus image pair," J. Biomedical Optics, International Society for Optics and Photonics 6(13), 64026–64026 (2008).

[5] Scharstein, D.; Szeliski, R. "High-accuracy stereo depth maps using structured light," IEEE. Computer Vision and Pattern Recognition, Proc. IEEE Computer Society Conference 1(1), 195 (2003).

[6] Ester, M.; et al. "A density-based algorithm for discovering clusters in large spatial databases with noise," KDD 1(96), 226–231(1996).

[7] Glasbey, C. A. "An analysis of histogram-based thresholding algorithms," CVGIP: Graphical Models and Image Processing 1(55), 532–537(1993).

7. ACKNOWLEDGEMENT

Thanks for financial support to the project development obtained by CNPq (Process No. 455970/2012-2) and FAPESP (Process No. 2013/22940-5).