35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

3D Optic Disc Reconstruction via a Global Fundus Stereo Algorithm

M. Bansal[1], M. Sizintsev[1], J. Eledath[1], H. Sawhney[1], D. J. Pearson[2] and R. A. Stone[2]

***Abstract*— This paper presents a novel method to recover 3D structure of the optic disc in the retina from two uncalibrated fundus images. Retinal images are commonly uncalibrated when acquired clinically, creating rectification challenges as well as significant radiometric and blur differences within the stereo pair. By exploiting structural peculiarities of the retina, we modified the Graph Cuts computational stereo method (one of current state-of-the-art methods) to yield a high-quality algorithm for fundus stereo reconstruction. Extensive qualitative and quantitative experimental evaluation (where OCT scans are used as 3D ground truth) on our and publicly available datasets shows the superiority of the proposed method in comparison to other alternatives.**

I. Introduction

A. Motivation

Degenerative eye diseases that affect the optic nerve, like glaucoma, are leading causes of vision loss and blindness. For diagnosing glaucoma, a key indicator is the 3D conformation of the optic disc [1]. Consequently, easy and reliable 3D structure recovery technology for the optic disc has long been sought for prompt and accurate glaucoma diagnosis.

To avoid the subjective challenges to human observers in assessing optic disc structure from conventional clinical photographs [1], scanning laser techniques are being applied to the optic disc and/or retina, including various approaches to optical coherence tomography (OCT). These instruments are yet to achieve convincingly superior performance for glaucoma diagnosis over the qualitative evaluation of stereoscopic optic disc photographs by experienced clinicians. These technologies are continuously evolving, but with vendor specific image formats and software, instrument specific analytical capabilities, high cost and other limitations.

With the continuing improved resolution of fundus cameras that acquire conventional images of the optic disc, several groups are pursuing the appealing alternative of recovering 3D structure from optic disc images captured from different viewpoints to create a stereo image pair.

The general idea is depicted in Fig. 1 where two images of the fundus taken from slightly different viewpoints (referred to as "left" and "right" images in the stereo pair) are used to recover the 3D geometry via a computational method that closely achieves the 3D structure recovery quality of a "ground truth" method, here assumed to be OCT.

Supported by NIH Bioengineering Research Partnership grant R01-EY017299, the Paul and Evanina Bell Mackall Foundation Trust (RAS) and Research to Prevent Blindness (RAS).

[1]SRI International, 201 Washington Rd, Princeton, NJ, USA mayank.bansal@sri.com

[2]University of Pennsylvania, Philadelphia, PA, USA stone@mail.med.upenn.edu

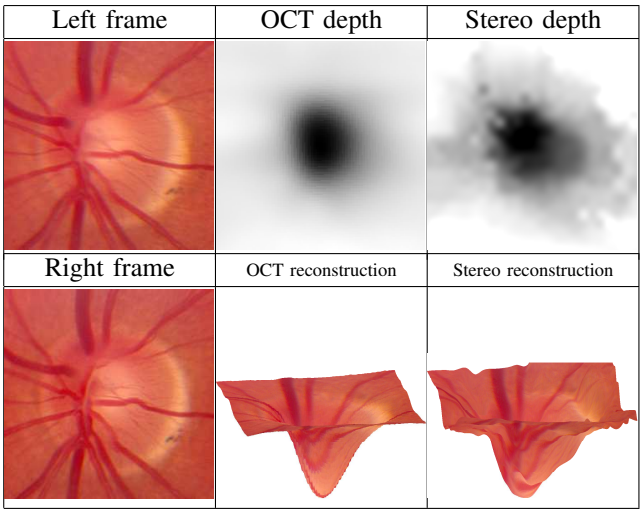

Fig. 1. 3D reconstruction of fundus from two spatially displaced views is comparable to the depth measurements taken by the OCT method. From left to right: Two retinal images taken from different viewing angles referred to as "left" and "right" images of the stereo pair; OCT scan depth data and corresponding 3D reconstruction used as the "Ground Truth"; stereo disparity map obtained by the proposed algorithm and the corresponding 3D reconstruction are a very good correspondence with the ground truth.

B. Previous Research

Computational stereo has been extensively researched for several decades, and remarkable results have been achieved in a number of applications such as 3D modeling, aid-in driving, augmented reality, etc. [13], [3]. At the same time, perfunctory application of existing stereo solutions may not provide the best results for retinal and optic disc images for a number of reasons: significant change in viewing angle results in noticeably different intensity properties of a region, violating the Lambertian surface assumption as well as introducing geometric distortions; camera focus may be significantly different from one image to another as the capture process is not synchronized (an issue infrequently addressed by the stereo community); images taken at different times may have local artefacts from ocular media opacities (e.g., vitreous floaters); images taken manually require a good automatic rectification procedure as prior calibration is not an option when manual camera displacement is needed to induce image disparity.

Previous stereo approaches to ocular fundus images attempted to alleviate the aforementioned issues by performing stereo estimation only around blood vessels which can provide distinctive structure (e.g., [8]), developing a special-purpose multi-scale matching strategy (e.g., [16]), attempting to explicitly reduce noise prior to stereo matching (e.g., [12]) or using other specific assumptions such as near-planar

978-1-4577-0216-7/13/$26.00 ©2013 IEEE

surfaces of the retina way from the optic disc (e.g., [9]).

C. Contributions

In the light of the problem motivation and previous research, this paper suggests the following contributions described in detail in Sec. II. First, we perform automatic rectification of uncalibrated stereo images. Then, we propose a global stereo algorithm that derives its smoothness connected grid structure on the optic disc location and solves it efficiently via Graph Cuts method [2], [7]. We also apply simple yet efficient global blur compensation to image stereo pairs. In Sec. III, the performance of the proposed advances is demonstrated on our fundus dataset and quantitatively compared to a number of alternative stereo methods on a publicly available dataset with OCT ground truth (University of Iowa, UoI; see [16]). The paper concludes with final remarks in Sec. IV.

II. Technical Approach

Binocular stereo relies heavily on the epipolar constraint that is described by the Fundamental matrix [3], [5]. Once the fundamental matrix is determined, stereo images are rectified [11], i.e. warped such that the epipolar lines (which constrain the location of the matched points) align with horizontal image scanlines, making the subsequent stereo computation efficient and more reliable. Robust normalized 8-point algorithm is a de facto standard for Fundamental matrix estimation and subsequent rectification [5]. Nevertheless, the facts that the retinal surface is close to planar away from the optic disc and the lower number of reliable matches impose a stricter requirement on the feature point selection for Fundamental matrix computation. In this work, we adapted a method [8] where two sets of SIFT feature [10] correspondences are used: one taken from the optic disc and the other taken from the surrounding retinal area, with 2D homography being first estimated for the surrounding feature matches. The cup center is detected by the method described in [18] and the image is subsequently segmented into cup/background regions.

In the following, we will denote the rectified left and right images by $\mathcal{I}_L$ and $\mathcal{I}_R$ respectively. After rectification, the corresponding pixels in the two images lie on the corresponding scan-lines, i.e. for any pixel (x, y) in $\mathcal{I}_L$, the corresponding pixel in $\mathcal{I}_R$ will be $(x + d, y)$ where d is an offset (called the *disparity*) which depends on the 3D depth of the structure imaged at pixel (x, y). This disparity is what needs to be determined for each pixel in $\mathcal{I}_L$. The widely-agreed taxonomy of computational stereo algorithms [13] classifies all solutions between two extremes: local and global methods. Local stereo algorithms independently compare each pixel (x, y) in the reference image to all pixels $(x + d, y)$ in the other image using some similarity measure $\delta(x, y, d)$ computed over locally aggregated patches and choose the disparity value $d^\star(x, y)$ corresponding to the best match, i.e., $d^\star(x, y) = \arg\max_d\{\delta(x, y, d)\}$. This solution is fast and easy to implement, but does not always produce good results in practice since it does not enforce depth continuity between neighboring pixels. Smoothness of depth between pixels is usually enforced in global stereo algorithms that formulate the problem as minimization of an energy term E over disparity map d:

$$E(d) = E_{data}(d) + \lambda_s E_{smooth}(d). \tag{1}$$

The total energy is minimized simultaneously for all pixels in the image and this tends to produce disparity maps of significantly higher quality. In this paper, we propose a novel global algorithm that specifically handles the challenges in matching fundus images.

A. Correlation volume computation

We will refer to the collection of values $\delta(x, y, d)$ over all pixels (x, y) and all potential disparities d by the name "correlation volume". Typically, this volume is computed for each pixel (x, y) in $\mathcal{I}_L$ by taking a neighborhood of pixels Ω_W around it ($W \times W$ sized patch) and measuring its correlation with a corresponding patch centered at location $(x + d, y)$ in $\mathcal{I}_R$. We use the zero-mean normalized-cross-correlation (NCC) measure which is defined as follows:

$$\delta_W(x, y, d) = \tag{2}$$
$$\frac{\sum_{(x',y')\in\Omega_W} \left(\mathcal{I}_L(x', y') - \overline{\mathcal{I}_L}\right)\left(\mathcal{I}_R(x' + d, y') - \overline{\mathcal{I}_R}\right)}{\sqrt{\sum_{\Omega_W}\left(\mathcal{I}_L(x', y') - \overline{\mathcal{I}_L}\right))^2 \sum_{\Omega_W}\left(\mathcal{I}_R(x' + d, y') - \overline{\mathcal{I}_R}\right))^2}},$$
$$\overline{\mathcal{I}_L} = \frac{1}{W^2}\sum_{(x',y')\in\Omega_W} \mathcal{I}_L(x', y'),$$
$$\overline{\mathcal{I}_R} = \frac{1}{W^2}\sum_{(x',y')\in\Omega_W} \mathcal{I}_R(x' + d, y'),$$

where the subscript W makes explicit the dependence of the correlation score on the window size chosen. This measure handles bias differences between the images well [3].

While this approach is sufficient for scenes where the object scales are similar, it leads to erroneous results for fundus images since the scale of the structure at each pixel can be quite variable. Choosing a small window size everywhere will lead to insufficient matching regions for low-texture and thick vessels. Similarly, choosing a large window size everywhere will lead to an overall smooth disparity map where the high-frequency variation of depth is lost. To resolve this, we propose a window-selection approach that determines the best window size at each pixel and show that the recovered size estimates agree with the scale of the structure at that point. In addition, by choosing the optimal window size at each pixel, we are able to use the corresponding correlation volume for the subsequent optimization process.

For each pixel, we compute the correlation volume $\delta_W(x, y, d)$ for a range of window sizes $W \in \{11, 21, 31, 41, 51\}$ pixels. Now, for each pixel, we can compare the maximum correlation magnitudes achieved by each window size and pick the window which gives the best correlation as the selected window for that pixel. However, this approach can lead to erroneous results since the correlation magnitudes across different window sizes are not directly

comparable. In addition, we prefer a correlation volume that has a more well defined peak to one with a higher but shallow peak. We found this to be particularly true for fundus images where, due to lack of texture in a smaller window, the correlation magnitudes tend to be high but using a larger window can include features like vessel boundaries which lead to a more sharply peaked correlation volume. Therefore, we use the correlation confidence measure from [19] that captures the distinctiveness of the best peak in the correlation volume for each pixel. If $s_1(x, y, W)$ and $s_2(x, y, W)$ are the best and second best correlation scores, respectively, for two distinct peaks in the correlation volume for a pixel (x, y) at window-size W, its correlation confidence is defined as:

$$c(x, y, W) = \left| \frac{s_1(x, y, W) - s_2(x, y, W)}{1 + s_2(x, y, W)} \right|. \quad (3)$$

The best window for pixel (x, y) is then selected as $W^\star(x, y) = \arg\max_W \{c(x, y, W)\}$. The correlation volume $\delta_{W^\star}(x, y, d)$ is chosen for subsequent optimization. Note that the $W^\star$ will, in general, be different for each pixel. In the following, we will drop the subscript and refer to the chosen correlation volume for each pixel simply as $\delta(x, y, d)$.

Once the correlation volume is computed, the standard approach to global stereo is to represent the input image as a 4-connected grid of pixels and set the data and smoothness terms as follows [13]. The term E_{data} is a unary data term that captures the match energy cost between associated pixels at each possible disparity and is thus set using the correlation volume $\delta(x, y, d)$. The term E_{smooth} is a binary term that penalizes nearby pixels if they achieve different disparities (usually defined via the Potts model, assigning 0 when disparities are equal and 1 when they are not [3]), thus, enforcing overall smoothness in the disparity map. For tractability, this smoothness is only enforced between pairs of pixels that are 4-connected in the image.

$$E_{data}(x, y, d) = 1 - \delta(x, y, d), \quad (4)$$

$$E_{smooth}(d_1, d_2) = \min(|d_1 - d_2|, 1) \quad (5)$$

The total energy $E(d)$ can then be optimized by techniques like Belief Propagation, Dynamic Programming, etc. In this paper, we use the Graph-Cuts α-expansion algorithm [2].

In the following, we describe our formulation where by adding domain specific components to the data and smoothness terms above, we are able to obtain a much more accurate solution to the fundus structure recovery problem.

B. Data Term

In our formulation, this term is computed as a composition of three sub-terms:

$$E_{data}(x, y, d) = \lambda_{conf} \cdot \lambda_{disc} \cdot E^{match}_{data}(x, y, d) \quad (6)$$

where,

$$E^{match}_{data}(x, y, d) = 1 - \delta(x, y, d), \quad (7)$$

$$\lambda_{conf}(x, y) = 0.5 e^{\sigma_c c(x, y, W^\star)}, \quad (8)$$

$$\lambda_{disc}(x, y) = e^{(3 - \frac{4r}{R})}, \quad (9)$$

where σ_c is an empirically determined constant ($\sigma_c = 10$ in all our experiments), r is the pixel distance of pixel (x, y) from the detected optic-disc center (k_x, k_y) and R is the width of the image in pixels.

The sub-term $\lambda_{conf}(x, y)$ uses the correlation confidence measure $c(x, y, W^\star)$ obtained from the correlation volume at the selected window-size $W^\star$ to modulate the weight of the data-term. A sharply peaked correlation profile implies a good confidence match and forces the correlation profile to have higher influence in the optimization. Otherwise, the lower magnitude of $\lambda_{conf}(x, y)$ reduces the influence of the correlation profile so that low-confidence points can be moved around in disparity space more readily. As an example, ignoring the factor λ_{disc}, for $\sigma_c = 10$ we will get $E_{data} = 0.5\, E^{match}_{data}$ for pixels with confidence 0, and $E_{data} = 1.35\, E^{match}_{data}$ for pixels with confidence 0.1.

In this work, we propose a novel "disc-dependent" term $\lambda_{disc}(x, y)$ that controls the contribution of the data term for each pixel according to its distance from the optic disc center. For pixels which are close to the disc center, we would like to avoid any smoothness based regularization and instead base their disparity computation on the data term as much as possible (still modulated appropriately by the confidence term λ_{conf}). In contrast, for pixels outside the disc, we would like to capture the context of the surrounding pixels and use that to regularize the computed disparity. Thus, the term $\lambda_{disc}(x, y)$ is a decreasing function of the distance of the pixel from the disc center.

C. Smoothness Term

Since we work with a 4-connected image graph, we will define smoothness only between pairs of pixels (x_1, y_1) and (x_2, y_2) such that they are either horizontal or vertical neighbors. Further, instead of the standard Potts model, we will employ the truncated quadratic model along with pixel-dependent weights to modulate the smoothness locally.

$$E_{smooth}(x_1, y_1, d_1, x_2, y_2, d_2) = \lambda_{int} \cdot E^{match}_{smooth}(d_1, d_2) \quad (10)$$

where,

$$E^{match}_{smooth}(d_1, d_2) = \min(|d_1 - d_2|^2, S_{max}) \quad (11)$$

$$\lambda_{int}(x_1, y_1, x_2, y_2) = e^{-|\mathcal{I}_L(x_1, y_1) - \mathcal{I}_L(x_2, y_2)|} \quad (12)$$

The term $\lambda_{int}(d)$ is typically employed in global stereo formulations and modulates the smoothness penalty as a function of the distance between the neighboring pixels in the intensity space. This modulation decreases the influence of the smoothness term for pixel neighbors with large intensity difference, e.g., when one pixel is on a retinal vessel and its neighbor is on the background retina.

In our approach, we use a truncated quadratic model to penalize neighboring disparity differences. Since, sharp depth discontinuties accompanied by half-occlusions are not typical for fundus images, we set $S_{max} = 32^2 = 1024$, which is noticeably higher than for regular stereo. At the same time, it is important to use the robust version of quadratic penalty function to avoid gross matching errors that are possible in fundus stereo as described in Sec. I-B.

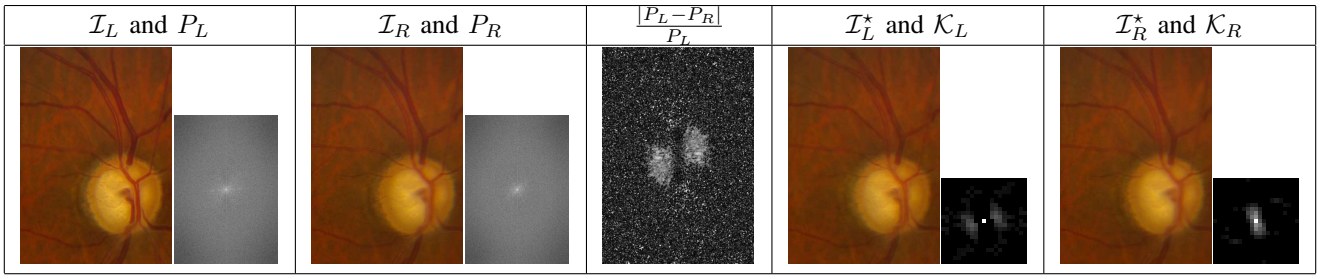

Fig. 2. Reference and matching images have a difference in overall blur properties (an example from *UoI Dataset* [16]). From left to right: Left image and its magnitude power spectrum; Right image and its magnitude power spectrum; Pointwise relative differences between left and right power spectra $\frac{|P_L - P_R|}{P_L}$; Blur-compensated left image with corresponding convolution kernel; Blur-compensated right image with corresponding convolution kernel.

D. Blur compensation

Since fundus images are often acquired with camera with manual or autofocus but the subject's eyes can move, the overall blur properties of the images may be noticeably different due to effectively independent focal length values, motion blur and/or aberrations in the ocular media, as depicted Fig. 2. This blur discrepancy will complicate the stereo matching procedure, as the normalized correlation match measure (2) works best under the assumption that image patches differ only in global additive and multiplicative intensity constants, and not in the frequency content.

Here, we propose to minimize this discrepancy by finding the maximal magnitude power spectrum profile that is common to both images. Consider stereo images $\mathcal{I}_L$ and $\mathcal{I}_R$ having their magnitude power spectrums P_L and P_R computed via Fourier analysis, e.g., FFT as

$$P_L = \|\mathcal{F}(\mathcal{I}_L)\| \quad \text{and} \quad P_R = \|\mathcal{F}(\mathcal{I}_R)\|. \tag{13}$$

Then, the maximal power spectrum profile common to both images is just the pointwise minimum of corresponding images' power spectra computed as

$$P_{best}(x) = \min(P_L(x), P_R(x)). \tag{14}$$

Figure 2 shows corresponding magnitude power spectra and exemplifies their common differences by showing the pointtise relative difference map $\frac{|P_L - P_R|}{P_L}$. Finally, since both left and right images must have the same P_{best}, and convolution in spatial domain is equivalent to pointwise multiplication in Fourier domain, the spatial convolution kernels $\mathcal{K}_L$ and $\mathcal{K}_R$ are computed as

$$\mathcal{K}_L = \mathcal{F}^{-1}\left(\frac{P_{best}}{P_L}\right) \quad \text{and} \quad \mathcal{K}_R = \mathcal{F}^{-1}\left(\frac{P_{best}}{P_R}\right). \tag{15}$$

These kernels are then used to transform the original $\mathcal{I}_L$ and $\mathcal{I}_R$ to images $\mathcal{I}_L^\star$ and $\mathcal{I}_R^\star$ with optimal common blur properties. Specifically,

$$\mathcal{I}_L^\star = \mathcal{K}_L \otimes \mathcal{I}_L \quad \text{and} \quad \mathcal{I}_R^\star = \mathcal{K}_R \otimes \mathcal{I}_R, \tag{16}$$

where $\otimes$ denotes the convolution operator. Furthermore, for convenience, spatial kernels are bound to the 21×21 window size (majority of values farther from the center are very close to zero anyway) and normalized to have a unit energy.

E. Overall Algorithm

Once our data and smoothness terms are defined as in equations (6) and (10), we can optimize the overall energy E in equation (1) to compute the disparity map by running the graph-cuts expansion algorithm developed by [2]. In our implementation, the overall weight of the smoothness term E_{smooth} in equation (1) is empirically set to $\lambda_s = 10$.

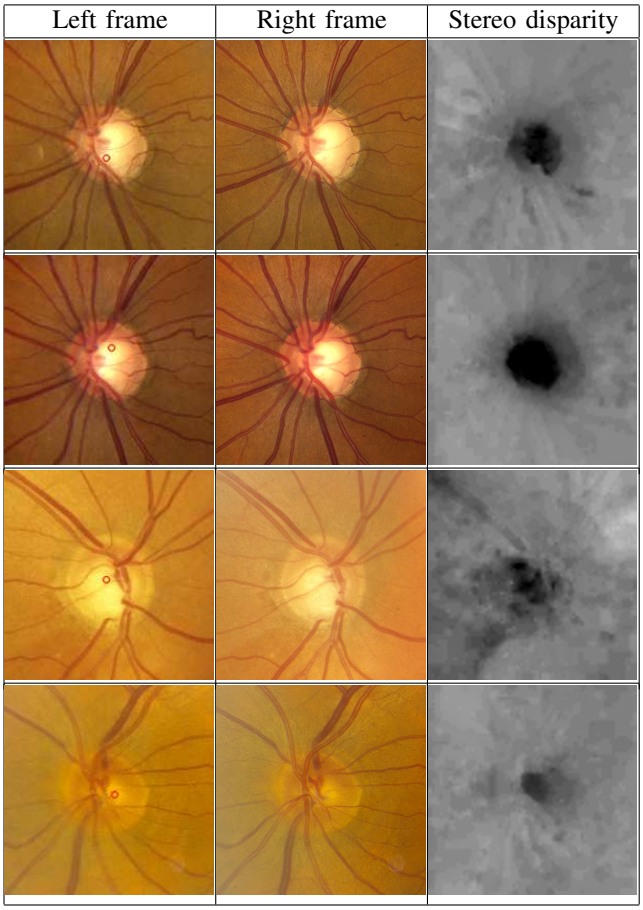

Fig. 3. Rectified stereo pairs and stereo reconstruction results for our dataset examples (see test for details). Red circle in the left image indicates the automatically detected optic disc center. In disparity maps, brighter value pixel denotes closer distance.

The overall algorithm is the following:

1) Compute disc center and rectify left and right images.
2) Compensate for the global blur discrepancy between the rectified left and right images using the algorithm in section II-D.
3) Compute the optimal correlation volume $\delta(x, y, d)$ using the algorithm in section II-A.
4) Compute the initial disparity assignment for each pixel using the best correlation value for the pixel, i.e.

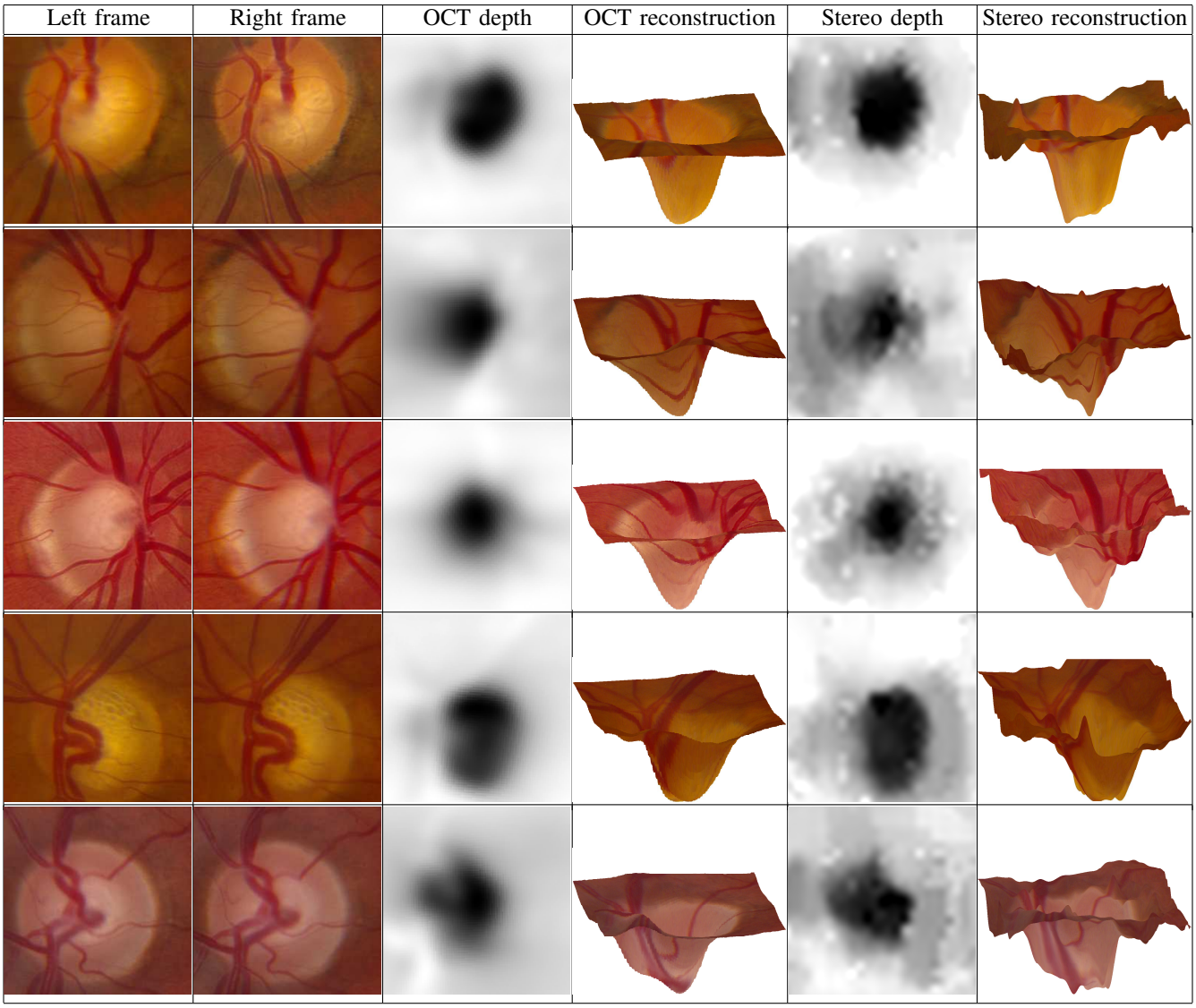

Fig. 4. More examples of stereo pairs and stereo reconstruction results for *UoI Dataset*.

$d^\star(x,y) = \arg\max_d\{\delta(x,y,d)\}$.

5) Optimize E using graph-cuts.
6) Estimated pixel disparities are the final solution.

III. Experimental Evaluation

The algorithm described in Sec. II has been implemented in software and was tested on two different datasets.

The first dataset was obtained by the authors, and 4 pairs[1] with recovered disparity maps are shown in Fig. 3. The general shape of the retina is recovered correctly, and structure of the cup is clearly assessable for subsequent diagnostics. The variable appearance between each stereo pair of images reflects the underlying differences in individual anatomy. The proposed algorithm introduces artifacts only in the periphery of the image away from the optic disc – regions that are usually characterized by severe lack of structure.

[1]All these images were acquired on a Zeiss FF4 fundus camera (Carl Zeiss Meditec, Inc., Dublin CA). The first, second and fourth stereo pairs were acquired digitally using the OIS system (Ophthalmic Imaging Systems, Sacramento CA) and extracted as .tif files. The third image pair was obtained on Fujichrome Velvia 100 slide film and scanned as described in [15].

To perform quantitative evaluation as well as direct comparison of the proposed stereo algorithm to alternative solutions, we use the UoI dataset [16]. The dataset contains 30 fundus stereo pairs of variable quality that have proved to be particularly challenging due to the lack of precise rectification as well as significant radiometric and blur differences.

Importantly, each stereo pair has an associated registered 3D depth map obtained via OCT that is assumed to be a ground truth for stereo disparity maps. The accuracy of the estimated depth map is evaluated as a root of mean squared (RMS) difference between the OCT depth map and the stereo depth map. Since stereo is metrically uncalibrated, the disparity map is converted to the depth map via finding a global linear transformation function, as described in [17].

Example frames, their associated depth ground truth, disparity maps recovered by the proposed algorithm and their corresponding 3D reconstruction visualizations are depicted in Figs. 1 and 4. Depth maps recovered by stereo are in very good correspondence with the OCT-based ground truth depth maps albeit minor high-frequency artifacts arise as a

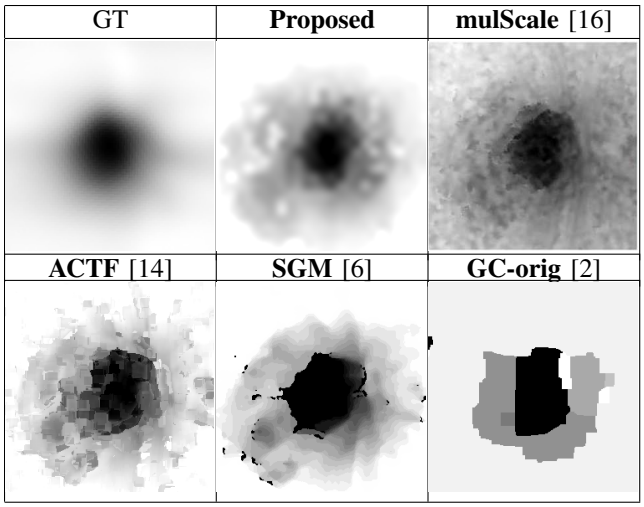

Fig. 5. Stereo pair and estimated disparity maps for *UoI Dataset* example using a set of algorithms outlined in Tab. I. Visualization of result for **mulScale** has been directly adapted from [16].

Alg.	**Proposed**	**mulScale**	**ACTF**	**SGM**	**GC-orig**
RMS	0.1469	0.1592	0.222	0.8195	3.0353
SD	0.047	0.0879	0.14	0.412	2.6591
95CI	0.12-0.16	0.13-0.19	0.17-0.27	0.66-0.97	1.93-4.13

TABLE I

ERROR RATES FOR SEVERAL ALGORITHMS. SEE TEXT FOR DETAILS.

result of violation of assumptions of computational stereo: specularities, significant local differences in blur and errors in automatic fundamental matrix estimation.

Importantly, Tab. I shows the error rate of the proposed algorithm, as well as other well-known methods used in stereo computation. To be directly comparable to [16], we compute RMS, its standard deviation (SD) and 95% confidence interval (95CI). In addition to our algorithm and methods considered in [16] (the method proposed by authors of [16] is denoted **mulScale**), we assess other effective stereo techniques: (i) adaptive coarse-to-fine stereo, being a robustified improved version of standard correlation block matching method [13] iteratively applied over image pyramid [4] and being one of the fastest and the most versatile stereo methods (**ACTF** as in [14]); (ii) semiglobal matcher, that aggregates match cost globally but makes disparity decisions locally, being also a quite fast algorithm able to display state-of-the-art results (**SGM** as in [6]); (iii) original Graph Cuts stereo, being a good example of classic global stereo method (**GC-orig** [2]) and the algorithm from which our proposed solution derives. Based on the quantitative evaluation, the proposed algorithm outperforms all other alternatives; furthermore, straightforward application of existing stereo solutions as-is proved to be insufficient in the fundus reconstruction domain. To exemplify the point, Fig. 5 shows disparity maps for all aforementioned algorithms applied to the third example of Fig. 4. Here, results shown in the bottom row represent the structure of the fundus poorly – the structure of the cup is either oversmoothed (and the surrounding retina mapped to a plane) or numerous false depth discontinuities are introduced throughout the retina.

IV. DISCUSSION AND CONCLUSIONS

We present an algorithm to compute dense disparity maps from two images of the retina for subsequent 3D reconstruction of the optic disc. The method is based on the Graph Cuts global optimization, takes into account specific structure of the fundus imagery, and accounts for global blur differences. Based on qualitative and quantitative experimental evaluation, our algorithm exhibits better performance than all other tested computational stereo alternatives. In comparison to modern laser scanning instruments, an image-based computational approach is more clinically intuitive, can accommodate improvements in software and camera technology, is compatible with many image formats, permits use of archived images and is cost-effective. By accurately reconstructing the optic disc, this refined approach to stereo recovery can permit robust detection of optic disc stability or change over time, and it offers great promise for advancing optic nerve diagnosis in glaucoma.

REFERENCES

[1] P. Airaksinen, A. Tuulonen, and E. Werner. *The Glaucomas*, chapter 28: Clinical evaluation of the optic disc and retinal nerve fiber layer, pages 617–657. St. Louis, Mosby, 1996.
[2] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *TPAMI*, 23(11):1222–1239, 2001.
[3] M. Z. Brown, D. Burschka, and G. D. Hager. Advances in computational stereo. *TPAMI*, 25(8):993–1008, 2003.
[4] P. J. Burt and E. H. Adelson. The Laplacian pyramid as a compact image code. *Trans. Comm.*, 31(4):532–540, 1983.
[5] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision. Second Edition.* Cambridge University Press, 2004.
[6] H. Hirschmuller. Accurate and efficient stereo processing by semi-global matching and mutual information. *CVPR*, 2:807–814, 2005.
[7] V. Kolmogorov and R. Zabih. Computing visual correspondence with occlusions using graph cuts. In *ICCV*, pages 508–515, 2001.
[8] S. Kuthirummal, M. Bansal, H. S. Sawhney, J. Eledath, D. J. Pearson, and R. A. Stone. 3D alignment and change detection from uncalibrated eye images. In *HIBS*, pages 299–306, 2011.
[9] Y. Lin and G. Medioni. Retinal image registration from 2D to 3D. In *CVPR*, 2008.
[10] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
[11] J. Mallon and P. F. Whelan. Projective rectification from the fundamental matrix. *IVC*, 23(7):643–650, 2005.
[12] T. Nakagawa, T. Suzuki, Y. Hayashi, Y. Mizukusa, Y. Hatanaka, K. Ishida, T. Hara, H. Fujita, and T. Yamamoto. Quantitative depth analysis of optic nerve head using stereo retinal fundus image pair. *J. Biomed Opt.*, 13(6):064026, 2008.
[13] D. Scharstein and R. Szeliski. Taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *IJCV*, 47:7–42, 2002.
[14] M. Sizintsev, S. Kuthirummal, S. Samarasekera, R. Kumar, H. Sawhney, and A. Chaudhry. GPU accelerated realtime stereo for augmented reality. In *3DPVT*, 2010.
[15] R. A. Stone, G. s Ying, D. J. Pearson, M. Bansal, M. Puri, E. Miller, J. Alexander, J. Piltz-Seymour, W. Nyberg, M. G. Maguire, J. Eledath, and H. Sawhney. Utility of digital stereo images for optic disc evaluation. *Invest Ophthalmol Vis Sci*, 51:5667–5674, 2010.
[16] L. Tang, M. K. Garvin, K. Lee, W. L. M. Alward, Y. H. Kwon, and M. D. Abramoff. Robust multiscale stereo matching from fundus images with radiometric differences. *TPAMI*, 33(11):2245–2258, 2011.
[17] L. Tang, Y. H. Kwona, W. L. M. Alwarda, E. C. Greenleea, K. Leeb, M. K. Garvinb, and M. D. Abramoff. 3D reconstruction of the optic nerve head using stereo fundus images for computer-aided diagnosis of glaucoma. In *Proc of SPIE*, volume 7624, pages 76243D–7, 2010.
[18] D. Wong, J. Liu, J. Lim, X. Jia, F. Yin, H. Li, and T. Wong. Level-set based automatic cup-to-disc ratio determination using retinal fundus images in ARGALI. In *EMBS*, pages 2266–2269, 2008.
[19] Q. Yang, L. Wang, R. Yang, H. Stewenius, and D. Nister. Stereo matching with color-weighted correlation, hierarchical belief propagation and occlusion handling. *TPAMI*, 31:492–504, 2009.