Fundus Image Mosaicking for Information Augmentation in Computer-Assisted Slit-Lamp Imaging

Rogério Richa*, Rodrigo Linhares, Eros Comunello, Aldo von Wangenheim, Jean-Yves Schnitzler, Benjamin Wassmer, Claire Guillemot, Gilles Thuret, Philippe Gain, Gregory Hager, and Russell Taylor

***Abstract*—Laser photocoagulation is currently the standard treatment for sight-threatening diseases worldwide, namely diabetic retinopathy and retinal vein occlusions. The slit lamp biomicroscope is the most commonly used device for this procedure, specially for the treatment of the eye periphery. However, only a small portion of the retina can be visualized through the biomicroscope, complicating the task of localizing and identifying surgical targets, increasing treatment duration and patient discomfort. In order to assist surgeons, we propose a method for creating intraoperative retina maps for view expansion using a slit-lamp device. Based on the mosaicking method described by Richa *et al.*, 2012, the proposed method is a combination of direct and feature-based methods, suitable for the textured nature of the human retina. In this paper, we describe three major enhancements to the original formulation. The first is a visual tracking method using local illumination compensation to cope with the challenging visualization conditions. The second is an efficient pixel selection scheme for increased computational efficiency. The third is an entropy-based mosaic update method to dynamically improve the retina map during exploration. To evaluate the performance of the proposed method, we conducted several experiments on human subjects with a computer-assisted slit-lamp prototype. We also demonstrate the practical value of the system for photo documentation, diagnosis and intraoperative navigation.**

***Index Terms*—Biomicroscope, fundus, mosaicking, slit-lamp.**

Manuscript received December 20, 2013; revised February 10, 2014; accepted February 25, 2014. Date of publication March 03, 2014; date of current version May 29, 2014. This work was supported in part by the Brazilian National Council for Scientific Research and Development, by the National Institutes of Health (NIH) Bioengineering Research Partnership Grant NIH 1R01 EB007969 and in part by Johns Hopkins University internal funds. *Asterisk indicates corresponding author.*

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org, provided by the authors.

*R. Richa is with the Computer Science, Universidade Federal de Santa Catarina, 88040-900 Florianopolis, Brazil
(e-mail: richa@incod.ufsc.br).

R. Linhares, E. Comunello, and A. von Wangenheim are with the Universidade Federal de Santa Catarina, 88040-900 Florianopolis, Brazil.

J.-Y. Schnitzler and B. Wassmer are with the Quantel Medical, 63039 Clermont Ferrand, France.

C. Guillemot and P. Gain are with the Department of Ophthalmology, CHU Saint-Etienne, 42055 Saint-Étienne, France, and also with the Corneal Graft Biology, Engineering and Imaging Laboratory, Jean Monnet University, 42100 Saint-Etienne, France .

G. Thuret is with the Department of Ophthalmology, CHU Saint-Etienne, 42055 Saint-Étienne, France, and also with the Institut Universitaire de France, 75005 Paris, France .

G. Hager and R. Taylor are with the Department of Computer Science, Johns Hopkins University, Baltimore, MD 21218 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2014.2309440

I. Introduction

LASER panretinal photocoagulation (PRP) is the standard treatment for sight-threatening diseases worldwide, namely diabetic retinopathy and retinal vein occlusions, which have in common the development of large areas of irreversible retinal ischemia [31]. Ischemic retinal cells produce vascular growth factors that trigger the proliferation of abnormal neovascularization responsible for the most severe complications: bleeding and retinal detachments. PRP prevents blindness by destroying the ischemic areas comprised between the macula and the periphery. Complementarily, selective laser photocoagulation in the macular area is also an efficient treatment for macular edema that also frequently occurs in both retinal diseases cited above. Naturally, with worldwide aging of the population and increasing rates of diabetes, there is a growing interest in the improvement of laser treatment techniques for the prevention of blindness.

The detection of ischemic areas is most often performed using fundus color photography and/or fluorescent angiography, which requires surgeons to mentally register the aforementioned preoperative images and live images of the retina, acquired during the procedure. The realization of the PRP (conventional or short pulse) is strictly codified and must respect several principles: spots must cover all the ischemic areas, spots density may vary according to the severity of the ischemia, spots should not be confluent or placed over large caliber vessels and, for security reasons, they must mandatorily respect the central foveal area and stop at the temporal vascular arcades. Considering the increasing number of patients which require treatment and the strict requirements for the realization of the PRP, there is a clear opportunity for computer assistance for increasing safety and efficiency of the procedure.

In this context, the slit-lamp biomicroscope in conjunction with a wide angle contact lens is the most commonly used device for laser delivery. Even though the visible area of the retina is smaller in slit-lamp images compared to those acquired by fundus cameras or scanning laser ophthalmoscopes [9], the magnification, stereoscopy and control offered by the slit-lamp makes it a very popular choice in the clinical environment [2].

To overcome some of the technical limitations associated with the slit-lamp device, methods for creating intraoperative retina mosaics have been developed. Their primary goal is to

aid surgeons in tasks such as preoperative planning, intraoperative navigation, view augmentation, and photodocumentation. The first efforts toward this goal have been reported by Markov *et al.* [20] and Barrett *et al.* [4], which proposed the first retinal tracking systems for laser surgery. The concept of retina mosaicking in slit lamp images was later proposed by Berger *et al.* [8]. Even though the preliminary results reported by the previous authors were very encouraging, the techniques employed in [4], [8], and [20] were not suitable to the illumination variations and patient motions verified in practice.

To meet these challenges, we have adapted and extended our previously developed method for mosaicking intraoperative retinal ophthalmic microscope images [24]. The method is a combination of feature-based and direct visual tracking methods, suitable for nature of retinal images. In this paper, we describe major improvements to the original formulation proposed in [24]. First, we have replaced the original visual tracking method based on the sum of conditional variance (SCV) with a sum of squared differences (SSD) method using a local illumination compensation model [26]. The previous SCV-based tracking method could only cope with global illumination variations, which caused tracking loss or drift in certain occasions. The incorporation of a local illumination compensation method enables tracking in more challenging conditions, with significantly less tracking drift. We have also incorporated the efficient pixel selection method proposed in [21] for improving computational efficiency and extended it to tracking in color images. We have incorporated an entropy-based mosaic update method to dynamically improve the retina map during variable visualization conditions. Finally, we have implemented a blending method for minimizing the visual impact of illumination variations and small distortions during the mosaicking process. The proposed method was deployed on a computer-assisted slit-lamp prototype and tested on several human subjects by two different practitioners. Results attest the practical value of the system for photo documentation, diagnosis and intraoperative navigation.

This paper is organized as follows. In the next section, we discuss related works in the literature. In Section III, we describe in details the computer assistant for laser photo-coagulation, as well as the extended tracking and mapping method adapted to biomicroscope images. In Section IV, we discuss the results of experiments conducted on human subjects. Finally, we conclude the paper in Section V.

II. Background

Recently, several solutions for intraoperative information augmentation in minimally invasive surgery (MIS) have been proposed in the literature. Examples are the field expansion system described by Totz *et al.* [32] and the augmented reality system proposed by Scarzanella *et al.* [34]. Although the retina geometry is much simpler compared to the MIS scenario, additional challenges are imposed, such as highly variable illumination (light is directly injected inside the eye, which create reflections and glare), partial and full occlusions, focus blur due to narrow depth of field of microscopes and distortions caused by the eye lens.

Similar mapping techniques have been proposed in the field of functional imaging, such as the mosaicking technique with super-resolution described in Hu *et al.* [14] and the deformation compensation methods for mosaicking proposed by Loewke *et al.* [18] and Vercauteren *et al.* [33]. Although direct visual tracking techniques are also employed in our work, the significantly lower image quality and acquisition speeds in functional imaging impose a different set of problems.

Slit-lamp biomicroscopes and fundus cameras offer significantly higher quality images compared to invasive retinal interventions such as vitreoretinal surgery [24], since disturbances caused by infusion liquids, blood or distortions caused the manipulation of the eye are avoided. Recently, a similar system for computer-assisted laser photocoagulation based on the fundus camera, the Navilas system [1], has been made commercially available. There is also a vast literature on fundus image registration [11], [28] for generating pre/post-operative panoramic views of the retina for diagnosing. Most techniques focus on detecting and matching vessel branches among images [12], [13]. These types of techniques are not ideally suited to real-time mapping using the slit-lamp device because they require a large field of view and high quality images for matching. In addition, with the exception of [9], image registration is not required on-the-fly and the proposed methods are not computationally efficient.

To meet the specific visualization conditions in slit-lamp biomicroscopic imaging, we adopted the hybrid tracking and mosaicking technique proposed in [24], developed for intraoperative retinal mapping during vitreoretinal surgery [24]. The method is a combination of feature-based and direct based methods, suitable for coping with images where salient visual features are not predominant. In the following section, we describe the method in detail and present extensions to the original formulation, which include a local illumination compensation method, an efficient pixel selection method for achieving higher computational efficiency and an entropy based map update method for improving the retinal map quality.

III. Materials and Methods

A. Computer-Assisted Slit-Lamp Biomicroscope

The setup for computer-assisted slit-lamp biomicroscope is illustrated in Fig. 1. Its main components are highlighted in the picture. A wide-angle Volk Super Quad 160 contact lens (Volk Optical, Mentor, OH, USA) is placed on top of the patient's cornea, which allows the visualization of the extreme periphery of the eye. A high definition 720×1280 pixel camera with an Aptina (San Jose, CA, USA) CMOS sensor acquiring at 60 fps is coupled to the biomicroscope for providing a view of similar quality to the one visualized directly through the biomicroscope. Images are grabbed on the computer by a Blackmagic Decklink acquisition board connected to the cameras through an HDMI link. The camera sensor was rotated to best fit the visible part of the retina. The life video feed from the camera is shown on the HD display next to the operator.

Some of the advantages of the video display over the standard eye-piece are the superior ergonomics and absence of dazzling reflections from cornea and lenses, which causes considerable

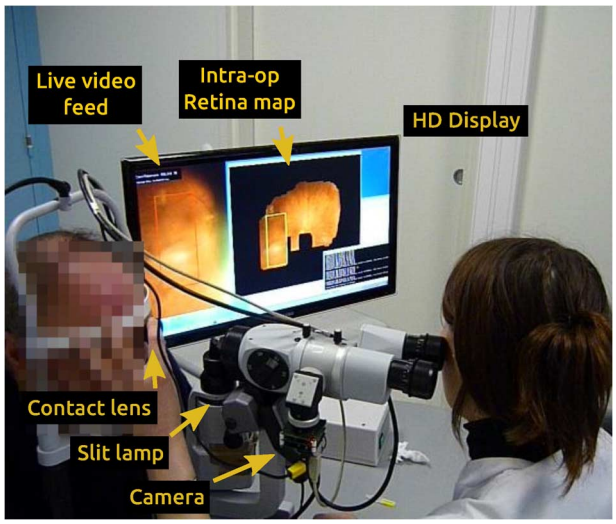

Fig. 1. Major components of the computer-assisted slit-lamp device used to develop and test the proposed mosaicking method.

discomfort to the operator. The use of the video display also opens the possibility of reducing photo toxicity by decreasing the quantity of blue light injected into the patient's eye. With a simple automatic calibration step, as proposed in [30], the video display can visually compensate for this reduction.

B. Mosaicking the Retina

A schematic overview of the proposed hybrid tracking and mosaicking method is given in Fig. 2. Notice that only a small portion of the retina is visible through the slit-lamp. To start mapping, an initial reference image of the retina is selected. Usually, the optical nerve is used as reference point on the retina. The center of the first reference image represents the origin of the retina map. Fig. 2 (top left) shows a template from the retina map being tracked. If the distance between the center of the tracked template and the center of the slit grows, a new template is incorporated to the retina map. Since the slit is narrow, the templates on the retina map overlap significantly (the center of templates on the retina map are indicated as rectangles in Fig. 2, right). At a given moment, the template closest to the current view of the retina is tracked using the direct tracking method detailed in Section III-B2.

1) Preprocessing: During examination, only a small rectangular window of the retina is visible. The size of the visible area is manually adjusted by the operator, which manages a tradeoff between patient comfort and view clarity. Due to the considerable amount of reflections from the contact lens and cornea, a preprocessing step is required to extract the visible part of the retina. The preprocessing step starts with thresholding of the input image from the microscope to generate a binary mask (i.e., "1" indicates a pixel belonging to the retina). We opted for a simple thresholding technique due to the small computational requirement. This process is illustrated in Fig. 3.

The threshold is computed based on the green and red color intensities. A pixel is deemed belonging to the retina if $r - 0.7g > 0$, where r and g are the red and green intensities for a given pixel. For refinement, a closing morphological operation (dilation followed by erosion with a 8×8 and 30×30 circular kernels, respectively) is applied to the binary mask to remove smaller regions. The center of mass of the resulting binary mask [Fig. 3(b)] is also computed to provide a reference to the retina tracking, which is discussed next. More details on this preprocessing step can be found in Appendix B.

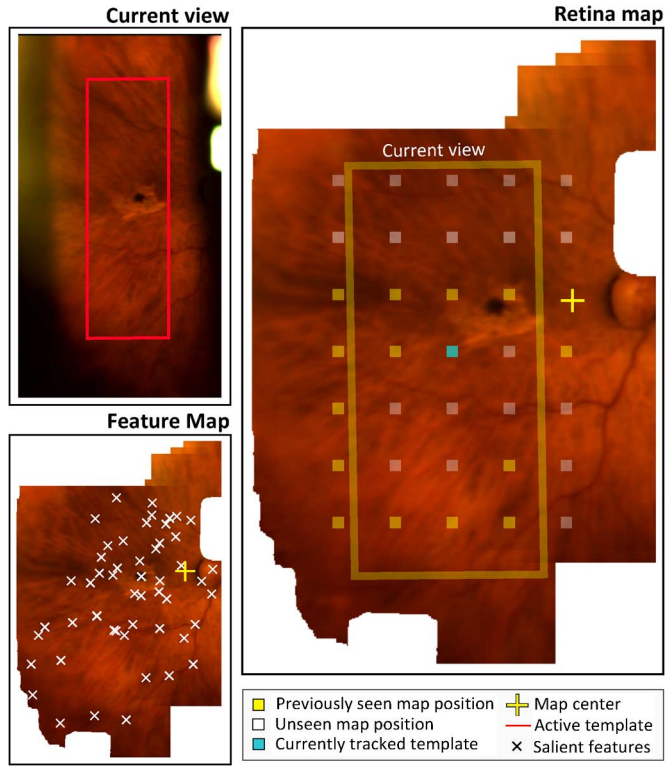

Fig. 2. Retinal mosaicking method. Direct visual tracking method (top left) is combined with a salient feature map (bottom left) for coping with full occlusions. Result is the retina map shown on the right.

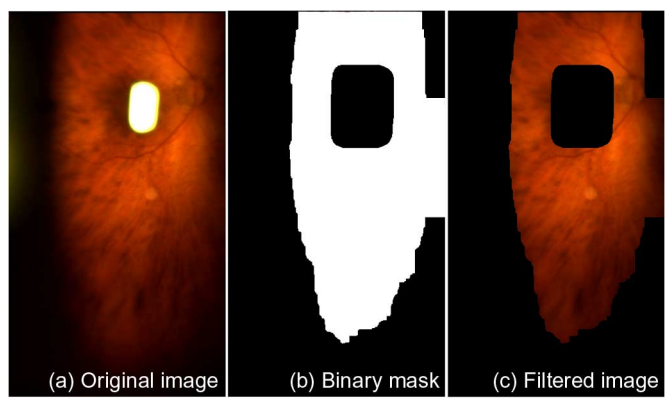

Fig. 3. Preprocessing step for detecting the visible part of the retina during examination. (a) Original image from the microscope. (b) Binary mask obtained by thresholding the color intensities. (c) Region of interest in the microscope.

2) Tracking: As discussed earlier, additional templates are incorporated to the retina map as the surgeon explores the retina. At a given moment, the template closest to the current view of the retina is tracked.

In the original tracking and mapping formulation for vitreoretinal surgery [24], the SCV [23] is used for tracking the retina under illumination variations. Compared to other robust image similarity measures from the medical imaging domain such as

mutual information (MI) and the normalized cross correlation (NCC), the SCV shows a good trade-off between robustness and convergence radius. Nevertheless, all previously cited similarity measures, including the SCV are not easily adaptable for tracking under local illumination variations [25]. For this reason, we have adopted the nonrigid compensation method proposed by Silveira *et al.* [27] to cope with the complex illumination variations in slit-lamp imaging.

By definition, tracking is formulated as an optimization problem. Given pixel positions $\mathbf{x}$ on the reference image T and $w(\mathbf{x}, \mathbf{p})$ the transformation function of parameters $\mathbf{p}$ which map every pixel $\mathbf{x}$ to the current image I, the tracking problem consists in estimating the parameters $\mathbf{p}$ which minimize the following dissimilarity function between the template T and current image $I(w(\mathbf{x}, \mathbf{p}))$

$$\min_{\mathbf{p},\mu,\beta} \sum_{\mathbf{c}} \sum_{\mathbf{x}} \left({}^{c}I^{*}\left(w(\mathbf{x}, \mathbf{p}), \mu, \beta\right) - {}^{c}T(\mathbf{x})\right)^{2} \tag{1}$$

where c indicates a specific color channel and μ and β are illumination parameters. Since patients use a head stand during examination, scale variations can be considered negligible in retinal images acquired with the biomicroscope. Slight rotations caused by head tilts occasionally occur. Therefore, the transformation function $w(.)$ is chosen to be a rotation and translation model. A good survey on rigid transformation models for direct visual tracking can be found in [3]. A more detailed description of the rotation and translation model can be found in Appendix A.

Since the retina absorbs most of the blue light component from the slit-lamp light source, the blue color channel is ignored for tracking. Hence, tracking is performed using red and green channels only $(c = R, G)$. Please refer to Appendix B for more details.

The nonrigid illumination compensation method consists in modeling local variations in contrast using a deformable parametric surface $g(\mathbf{x}, \mu)$ such as the thin-plate spline [17]

$$I^{*}\left(w(\mathbf{x}, \mathbf{p}), \mu, \beta\right) = g(\mathbf{x}, \mu) I\left(w(\mathbf{x}, \mathbf{p})\right) + \beta. \tag{2}$$

Notice that $g(.)$ is a function of pixel positions $\mathbf{x}$ and the control point vector μ, following the parametrization proposed by Lim *et al.* [17]. The parameter β is a variable for estimating global brightness variations.

a) Efficient pixel selection scheme: Direct visual tracking techniques rarely meet frame-rate requirements when the size of tracked templates are above 100 000 pixels. Alternatives such as the use of a graphical processing unit (GPU) or pixel selection schemes [3], [6], [21] are often required in order to achieve higher tracking speeds. In our specific context, we have adopted the pixel selection technique described by Meilland *et al.* [21] for two main reasons. First, it takes into consideration the transformation model used for tracking, and not only pixel intensity as in [3]. Second, it does not require an expensive preprocessing step as in [6], which is too restrictive in our context since templates are continuously being incorporated to the retinal map.

The goal of the pixel selection scheme is to select the "best" M pixels from the template image T for tracking. Usually, M is chosen between 5% and 25% of N (the original number of pixels in T), without any loss in tracking accuracy. It can be summarized in three main steps.

1) For each color channel, a $N \times 3$ Jacobian matrix $\mathbf{J}(\mathbf{p})$ is computed for $\mathbf{p} = \mathbf{p}_0$ (i.e., the original template configuration, where $I(w(\mathbf{x}, \mathbf{p}_0) = T(\mathbf{x}))$.
2) For each column $j \in \{1, 2, 3\}$, the absolute values on each row of $\mathbf{J}$, denoted here $|\mathbf{J}_j|$, are sorted in decreasing order of magnitude.
3) The list of "best" pixels is recursively filled by taking the pixel positions from one of the three columns of $\mathbf{J}$ at a time, carefully not incorporating the same pixel twice. This ensures that pixels chosen by this method equally represent each degree-of-freedom in the transformation estimation process.

Intuitively, the algorithm above selects pixels with the strongest gradients in parameter space, instead of image space. Details on the computation of $\mathbf{J}$ can be found in [7]. The final Jacobian matrix $\mathbf{J}'$ is obtained by stacking the Jacobian matrices generated by the selected pixels from each color component.

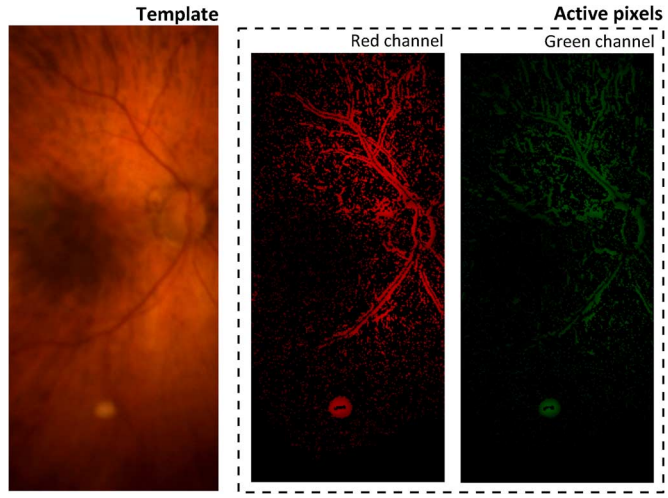

Fig. 4. Pixel selection scheme for tracking. From left to right, an image extracted from an *in vivo* experiment with a human subject and the selected pixels on red and green channels. Only 20% of the original template pixels are active in this example.

Typical results of the pixel selection scheme are displayed in Fig. 4. As shown in the figure, even though the red and green color components contain different structural information, the majority of selected pixels on each channel is redundant. For this reason, instead of selecting M pixels on each color channel separately, we have opted to select the "strongest" M pixels among both color channels. The result is a drastic improvement in tracking efficiency, since we are able to track with equally high accuracy with a significantly smaller set of pixels.

To solve the minimization problem in (1), the efficient second-order minimization (ESM) [7] is used. Weights are incorporated to its formulation for removing pixels that do not belong to the retina from the estimation of tracking parameters, using the binary mask described in Section III-B1.

3) Detection: If the template tracking confidence drops below a given threshold λ, tracking is deemed unreliable and is suspended. Tracking confidence is measured as the average NCC between the reference image ${}^{c}T$ and the compensated

warp ${}^{c}I^{*}(w(\mathbf{x},\mathbf{p}))$ [(2)] over red and green color channels c. The NCC was chosen as confidence measure since it is a bounded global measure (it varies between $[-1, 1]$), making it easier to define tracking confidence thresholds. The parameter λ was set as 0.85, after several experiments with our database. Although this parameter is empirically defined, it was kept constant for all videos on the database. For more details on the derivation of the NCC, please refer to [22].

For recovering tracking in case of full occlusions, a map of salient visual features on the retina is also created and maintained. For every new template incorporated in the map, the set of visual features within the new template is included in the retina feature map. Due to the overlap between templates, new features are matched against the existing features in the map. If the distance (in pixels) between old and new features on the map is small, two features are merged by taking the average of their positions and descriptor vectors.

In the original work [24], a combination of SURF (speeded-up robust features) [5] feature detection and description was used in conjunction with FLANN (fast library for approximate nearest neighbors) and RANSAC (random sample consensus) for feature matching. However, we have found through experiments on our database of retina videos that SIFT (scale invariant feature transform) [19] features offer a slightly better trade-off between computational complexity and robustness than SURF. Therefore, we replaced SURF by SIFT for this specific task.

4) Retina Map Update: As the platform operator explores the retina, new templates are incorporated to the retina map when tracking confidence is high (i.e., over a given threshold ϵ). Similar to the detection threshold, the inclusion threshold ϵ was empirically set to 0.97, which represents a very high tracking confidence.

a) Entropy-based map update: Due to the variable visualization conditions during the exam, templates already included in the map must be updated when visualization conditions improve. To this purpose, we use entropy as a measure of visual information. Entropy can be used as a measure of histogram dispersion and therefore it is useful to differentiate between sharpness levels of a pair of images. Although frequency domain techniques provide more accurate measures of contrast [15], the low computational requirement of entropy computation makes it an interesting solution in our context.

In our mapping method, we measure the entropy from the current tracked template $T(\mathbf{x})$ from the retina map, as well as the current warped image $I(w(\mathbf{x},\mathbf{p}))$, for every tracked frame from the video feed. These two images are supposed to be as similar as possible, and represent the same region of the retina. If the template image from the retina map is "less sharp" than the current warp, we replace the template for the warp, improving the sharpness of the retina map.

The total entropy h for a given image is computed as the entropy sum for all color channels ($h = \sum {}^{c}h$)

$$ {}^{c}h = -\sum_{i=0}^{i=b} {}^{c}p_i \log {}^{c}p_i \tag{3}$$

where b represents the number of intensity bins (i.e., for 8-bit images, $b = 256$), ${}^{c}p_i$ is the value of the ith bin from cth channel image histogram p. If the entropy of the current warped image is higher than the current template entropy and tracking confidence is higher than ϵ, then the template is updated. Examples of templates updated using the entropy-based method are shown in Fig. 5.

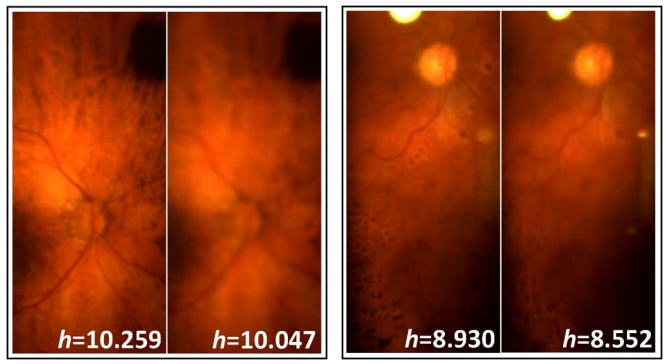

Fig. 5. An example of a retina map update based on entropy level h computed in (3). During the examination, dark, super exposed and blurry retina map templates are gradually replaced when visualization conditions improve.

b) Mosaic blending: In the original formulation [24], the resulting retina mosaic was a simply large image of stacked templates accumulated over time. Differences in lighting conditions between adjacent templates created "seams" in the image that often confused the operator during the exam. In order to circumvent this issue, we have implemented the weighted averaging blending method proposed by Szeliski *et al.* [29] to render a photo-realistic mosaic of the retina on the fly, taking advantage of the overlap between stored templates. Even though this technique is not capable of compensating for distortions created by tracking drifts, the visual quality of the resulting mosaic is satisfactory for our purposes.

IV. Experiments

The method described in the previous section was tested on more than 1 h of experiments on more than 10 different patients, performed at the University Hospital of St. Etienne, France. The examined patients had an age average of 66.8 years (ranging between 42 to 82 years of age) and the average examination duration was 30 s/eye. Exams were performed by two different practitioners.

A. Computational Aspects

The method was implemented using OpenCV and the CISST library, developed at JHU [16]. The system ran at frame-rate on an Intel Xeon 3.5 GHz computer with 16 Gb of RAM.

All tracking parameters remained fixed in all experiments (we provide them as supplementary material to this publication). The maximum number of iterations in the minimization procedure from (1) was set to 15. In average, the required number of iterations for convergence is seldom over 5 (at a cost of 1 ms per iteration). The preprocessing step described in Section III-B1 and the retina map update step in Section III-B4 have a relatively constant cost per image—an average of 3.5 and 4.0 ms per image, respectively.

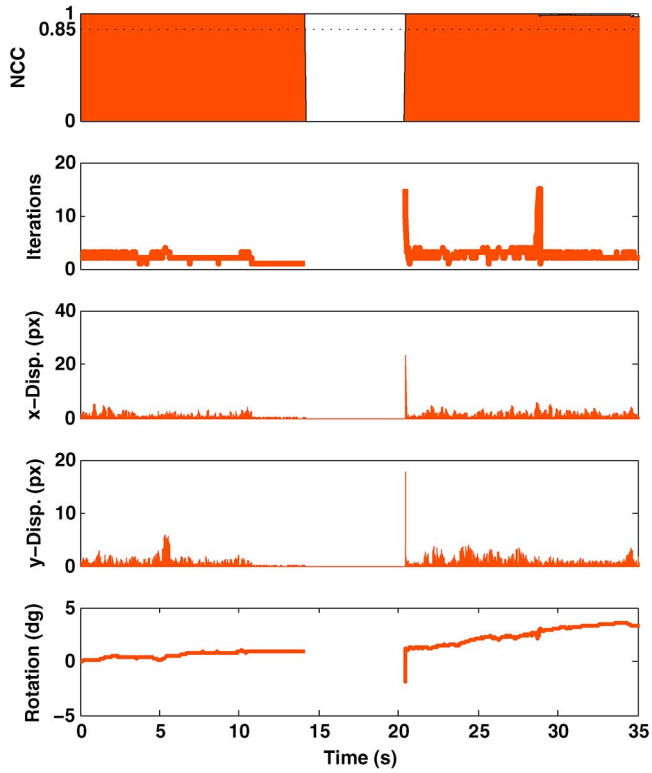

Fig. 6. Tracking statistics (from top to bottom, tracking confidence, number of iterations, horizontal and vertical displacements, and the rotational component) for exam #1.

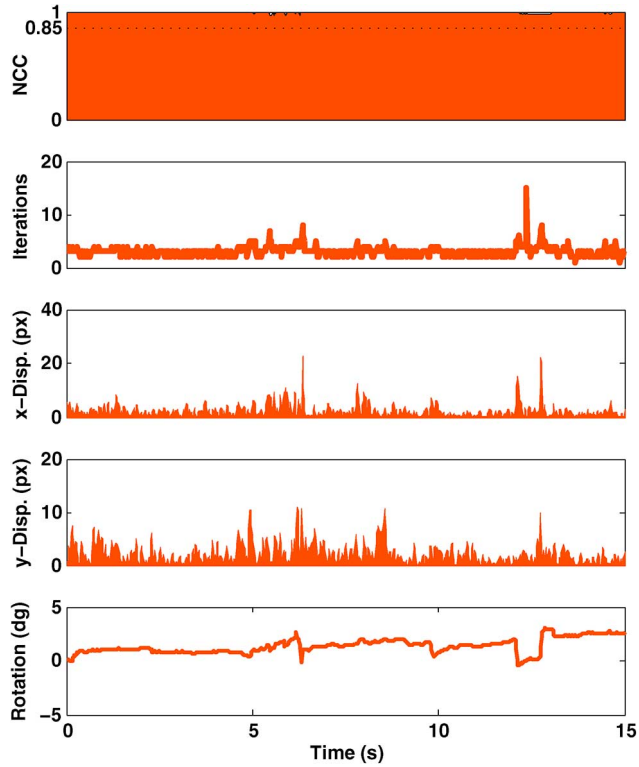

Fig. 7. Tracking statistics (from top to bottom, tracking confidence, number of iterations, horizontal and vertical displacements, and the rotational component) for exam #2.

B. Reconstructed Retina Maps

An example of a reconstructed retina map using the proposed method is shown in Fig. 8. From the images, one can easily identify a macular scar (arrow "1," highlighted in the picture). A comparison with a retina image acquired by a 30° Cannon CR-2 (New York, NY, USA) fundus camera attests the consistency of the reconstructed mosaic.

A comparison between the photo-realistic retina map (left) and the template superposition from the original formulation [24] (top right) clearly shows the ability of the refinement step described in Section III-B4 to compensate for illumination variations and small registration errors. However, mosaicking drifts can still reduce the contrast of the retina map, deteriorating certain parts of the map (as indicated by arrow "3").

One can also verify some residual artifacts created by unfiltered specular reflections (arrow "2"). These are usually specular reflections smaller than eight pixels, which are not detected in the preprocessing step due to the chosen size of the erosion mask. However, their effect is minimized by the blending step. One can also notice that regions occluded by specular reflections are not included in the retina map. Eventually, this leads to map holes, such as the one indicated by arrow "4." As supplementary material to this publication, we provide several examples of mosaics reconstructed using the proposed method.

C. Quantitative Analysis

For a quantitative analysis of system performance, we have recorded tracking statistics during two exams, which represent typical visualization conditions during a retina examination. They are shown in Figs. 6 and 7.

In both exams, we recorded the tracking quality (NCC score mentioned in Section III-B4), the number of ESM iterations for tracking convergence, the displacement in pixels in the vertical and horizontal directions and the rotation component. Tracking degradations such as occlusions or large specular reflections result in tracking loss during the procedure. These correspond to the intervals where the NCC coefficient drops below the predefined threshold ($\lambda = 0.85$). We recall that the value for the detection threshold λ was determined to be the typical tracking break point in experiments with our database and no fine tuning was required. In fact, any value in the range $\lambda = [0.6, 0.9]$ provides similar tracking smoothness.

In the first exam, we have simulated a long occlusion in the interval [10.0,32.5] s (see Fig. 6). In practice, this occlusion is equivalent to a sudden motion by the ophthalmologist while adjusting the contact lens or slit-lamp. Notice that an inaccurate tracking relocalization through feature detection and matching leads to a peak of iterations for convergence. However, tracking is successfully reestablished after the event. Overall, from a total of 1500 frames in this specific sequence, 85.4% (1282 frames) were tracked continuously, while tracking was lost in 14.6% of the sequence duration (218 frames). The corresponding video for the plot can be found in the supplementary materials.

In the second exam, we show the tracking performance in during head tilts and eye blinking (see Fig. 7). Analyzing the tracked angle θ, one can clearly notice sudden rotations caused by blinking. The proposed tracking method successfully copes

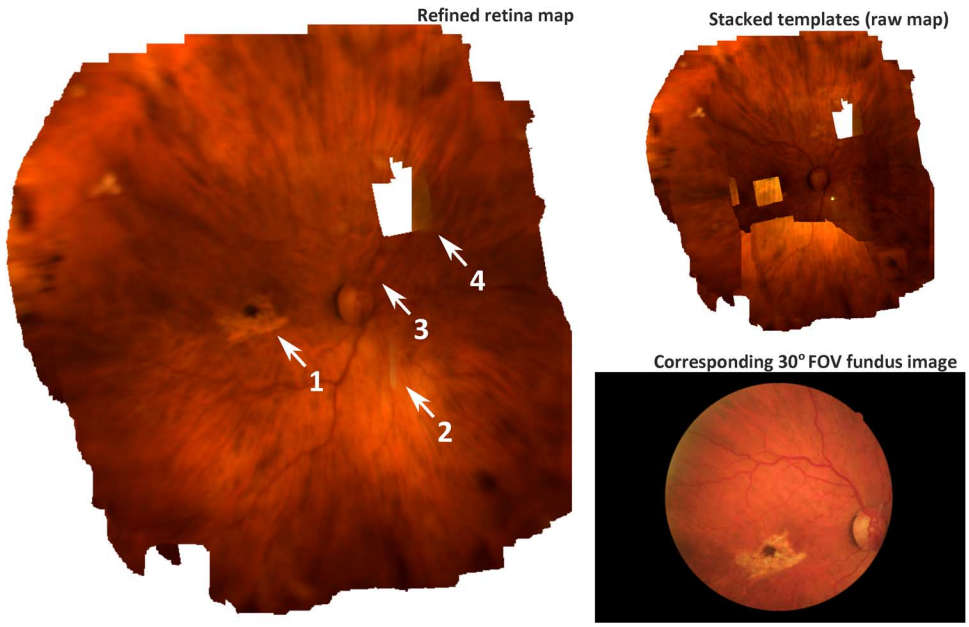

Fig. 8. Typical results of the proposed retina map reconstruction method. Comparison between the refined retina map (left) and the fundus image (bottom right) attests the quality of the proposed method. Notice also the significantly wider mosaic field of view (approximately 110°) compared to the fundus 30° field of view angle.

with such events. The corresponding video can be found in the supplementary materials.

One of the main goals of this analysis is also to validate the choice for a direct visual tracking method. Analyzing the horizontal and vertical displacements in Fig. 6 and 7, one can easily notice that the proposed tracking method can successfully track horizontal displacements of up to 20 pixels during the exam. Naturally, during periods of larger inter-frame motion, a higher number of iterations are required for convergence. Notice also that no instructions were given to the doctors regarding the system's usage and they did not have to constrain their gesture to meet the system's capabilities. This is a very positive aspect of the system.

To evaluate the impact of the pixel selection method, we have varied the number of active pixels for the sequence corresponding to the first exam and measured the number of required iterations for tracking convergence. For an active pixel percentage of 100%, 70%, 50%, and 20%, we have obtained an average of 2.25, 2.21, 2.21, and 2.15 iterations, respectively. These results indicate the practical value of the efficient pixel selection technique for reducing the computational load, while keeping equally high tracking accuracy. We have extended this experiment to five different sequences, featuring different levels of retinal texture and illumination intensity, and results systematically indicate a higher tracking efficiency when using 20% active pixels.

D. Local Illumination Compensation

The advantages of a local illumination compensation model become clear when we compare the reconstructed mosaics using the proposed method and the previous SCV-based tracking method [24]. In Fig. 9, one can clearly notice the accuracy gain using the proposed method, which consequently reduces mapping drift. We also provide videos as supplementary material to show the advantage of the local compensation method in videos with a narrow field of view (the light slit is usually set to a minimal size for photophobic patients).

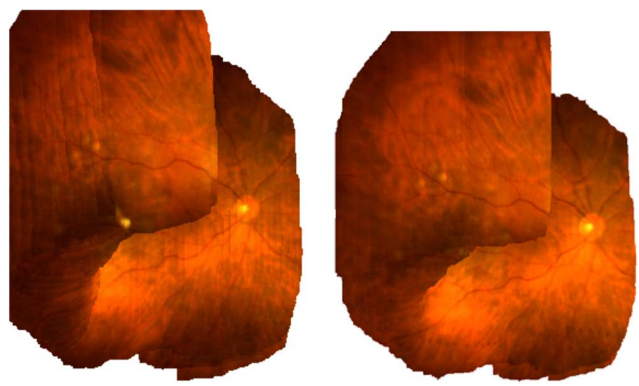

Fig. 9. Reconstructed mosaics (left) using the proposed SSD-based tracking method with local illumination compensation. (right) using an SCV-based tracking method with global illumination compensation. Blending is not used in this example to clearly show the drift caused by the integration of small tracking errors between adjacent templates.

E. The Transformation Model

To justify our choice for a transformation model, we have compared the performance of the rotation and translation model (3DOF model) with a translation only (2DOF) and rotation, translation and scale model (4DOF). For a fair comparison, we used the same starting point on the same sequence for all models. The videos containing tracking results using the three models above are provided as supplementary material. They clearly indicate that the 3DOF model provides the most accurate retina map. In the video example, the 2DOF model cannot cope with slight rotations, which result in a blurry mosaic. This

Fig. 10. Example of reconstructed mosaics (left) without the update step (right) dynamically updated using the entropy-based contrast update step.

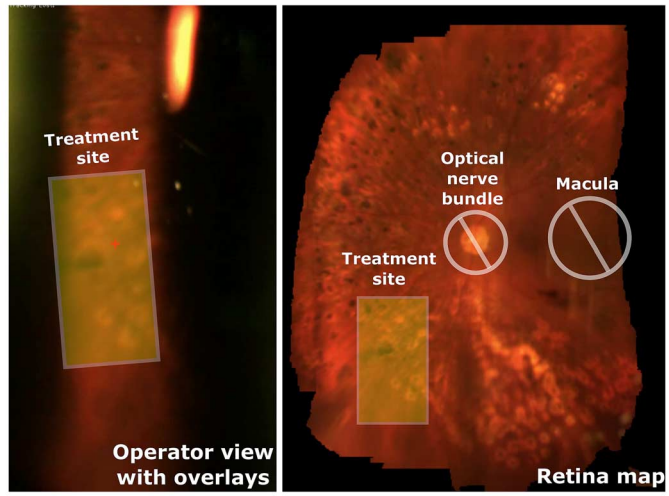

Fig. 11. Annotations made on the intraoperative retina map can be overlaid directly onto the live view, assisting surgeons in identifying targets and navigation.

indicates a large misalignment between adjacent templates on the retina map. Among all transformation models, the 4DOF has proven to be the most unstable model. This is due to the fact that the scale factor caused tracking to often get stuck in local minima and diverge.

F. Entropy-Based Mosaic Update

Since visualization conditions vary during the exam, the entropy-based mosaic update step was introduced to constantly update the mosaic and increase its contrast. To visually show the impact of this update step, we show an example in Fig. 10. Notice how the reconstructed mosaic using the entropy-based contrast update method creates sharper mosaics. This is due to the fact that the templates which compose the mosaic are dynamically updated when sharper images of the same area on the retina are available.

G. Limitations

Tracking Drift: Distortions in the imaging process caused by lens distortion induce tracking drifts, which can degrade the mosaic quality. This can induce blurring or ghosting effects, as indicated by Arrow "3" in the example from Fig. 8. To overcome this current limitation, we are currently incorporating a bundle adjustment step, similar to [18].

Clinical Cases: In patients with cataracts, the mosaicking method is likely to provide worse results due to the lack of texture on the retina images. This is one major characteristic of direct methods, which tend to "lock" on texture frequency components of highest spectral power. The presence of dense cataract, which result in blurry retina images, increases the probability of tracking lock on illumination artifacts and other disturbances. Note that patients combining dense cataract and retinal ischemia remain difficult to treat with laser, whatever the examination method and may require an association of surgery and laser.

Finally, in patients with high photosensitivity which require the slit to be very narrow, the operator is required to explore the eye very slowly, since the maximum inter-frame retina displacement supported by the proposed tracking method is usually around 10% of the slit lamp size.

H. Information Augmentation

As suggested in our previous work on vitreoretinal surgery augmentation [24], the proposed mapping system can be used for intraoperative guidance, facilitating the localization and identification of surgical targets, as illustrated in Fig. 11. This figure also shows how the proposed system could be used for mentoring apprentices. Through information overlay, a mentor could guide a novice by indicating surgical targets or other interest points on the retina, demonstrate surgical gestures or create virtual fixtures. Finally, preoperative images coupled with preoperative planning can be overlaid directly to the surgeon, provided the preoperative images and intraoperative mosaic are registered. For a more detailed demonstration of information augmentation, please see the supplementary videos.

V. Conclusion and Future Work

In this paper, we have developed a method for creating intraoperative retina maps using images from a slit-lamp biomicroscope. We propose three major improvements to the original tracking and mosaicking method described in [24]. The first is a visual tracking method with local illumination compensation to cope with challenging illumination variations. The second is an efficient pixel selection scheme for increasing tracking efficiency. The third is an entropy-based mosaic update step for coping with the variable visualization conditions during the exam. We have also incorporated a photo-realistic mosaic creation step for minimizing the visual impact of illumination variations and small local degradations. In order to validate the methods described in this paper, we have conducted more than 1 h of experiments on more than 10 different patients using a computer-assisted slit-lamp prototype. Results suggest the proposed method can provide valuable intraoperative guidance for surgeons, as well as generate high-quality panoramic views of the retina for photodocumentation and diagnosis.

We are currently working on a bundle adjustment step, similar to [18], to avoid image degradations created by tracking misalignment and lens distortions. This bundle adjustment step

will also enable us to generate higher-quality images by incorporating super-resolution techniques [14]. In the future, we plan on performing registration of preoperative angiography images and the reconstructed mosaicking for information augmentation during treatment based on the technique described by Can *et al.* [10].

REFERENCES

[1] The Navilas Laser System [Online]. Available: http://www.od-os.com

[2] J. Asmuth, B. Madjarov, P. Sajda, and J. Berger, "Mosaicking and enhancement of slit lamp biomicroscopic fundus images," *Br. J. Ophthalmol.*, vol. 85, no. 5, pp. 563–565, 2001.

[3] S. Baker and I. Matthews, Lucas-Kanade 20 years on: A unifying framework: Part 1 Robotics Inst., Carnegie Mellon Univ., Pittsburgh, PA, 2002, Tech. Rep..

[4] S. Barrett, M. Jerath, H. Rylander, and A. Welch, "Digital tracking and control of retinal images," *Opt. Eng.*, vol. 33, no. 1, pp. 150–159, 1994.

[5] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Comput. Vis. Image Understand.*, vol. 110, no. 3, pp. 346–359, 2008.

[6] S. Benhimane, A. Ladikos, V. Lepetit, and N. Navab, "Linear and quadratic subsets for template-based tracking," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Minneapolis, MN, 2007, pp. 1–6.

[7] S. Benhimane and E. Malis, "Homography-based 2D visual tracking and servoing," *Int. J. Robot. Res.*, vol. 26, no. 7, pp. 661–676, 2007.

[8] J. Berger and D. Shin, "Image-guided macular laser therapy: Design considerations and progress toward implementation," in *Ophtal. Technol. IX—SPIE Med. Imag.*, 1999, pp. 241–247.

[9] A. Broehan, T. Rudolph, C. Amstutz, and J. Kowal, "Real-time multimodal retinal image registration for computed-assisted laser photocoagulation system," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 10, pp. 2816–2824, Oct. 2011.

[10] A. Can, C. Stewart, B. Roysam, and H. Tanenbaum, "A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 3, pp. 347–364, Mar. 2002.

[11] P. Cattin, H. Bay, L. Van Gool, and G. Szekely, "Retina mosaicking using local features," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI'06)*. New York: Springer, 2006, pp. 185–192.

[12] T. Choe, I. Cohen, M. Lee, and G. Medioni, "Optimal global mosaic generation from retinal images," in *Proc. Int. Conf. Pattern Recognit.*, Washington, DC, 2006, pp. 681–684.

[13] K. Deng, J. Tian, J. Zheng, X. Zhang, X. Dai, and M. Xu, "Retinal fundus image registration via vascular structure graph matching," *Int. J. Biomed. Imag.*, 2010.

[14] M. Hu, G. Penney, D. Rueckert, P. Edwards, F. Bello, M. Figl, R. Casula, Y. Cen, J. Liu, Z. Miao, and D. Hawkes, "A robust mosaicing method with super-resolution for optical medical images," in *MIAR*. New York: Springer, 2010, vol. 6326, Lecture Notes Computer Science, pp. 373–382.

[15] M. Kristan, J. Pers, M. Perse, and S. Kovacic, "A Bayes-spectral-entropy-based measure of camera focus using a discrete cosine transform," *Pattern Recognit. Lett.*, vol. 27, no. 13, pp. 1431–1439, 2006.

[16] CISST Software Library, Lab. Computat. Sensing Robot., Johns Hopkins Univ. [Online]. Available: http://trac.lcsr.jhu.edu/cisst/, 2011

[17] J. Lim and M. Yang, "A direct method for modeling non-rigid motion with thin plate spline," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Washington, 2005, pp. 1196–1202.

[18] K. Loewke, D. Camarillo, W. Piyawattanameth, M. Mandella, C. Contag, S. Thrun, and J. Salisbury, "*In vivo* micro-image mosaicing," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 159–171, Jan. 2011.

[19] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[20] M. Markov, G. Rylander, and A. Welch, "Real-time algorithm for retinal tracking," *IEEE Trans. Biomed. Eng.*, vol. 40, no. 12, pp. 1269–1281, Dec. 1993.

[21] M. Meilland, A. Comport, and P. Rives, "Dense visual mapping of large scale environments for real-time localization," in *Proc. IEEE Conf. Intell. Robots Syst.*, San Francisco, CA, 2011, pp. 4242–4248.

[22] R. Richa, R. Sznitman, and G. Hager, Robust similarity measures for gradient-based direct visual tracking Johns Hopkins Univ., 2012, Tech. Rep..

[23] R. Richa, R. Sznitman, R. Taylor, and G. Hager, "Visual tracking using the sum of conditional variance," in *Proc. IEEE Conf. Intell. Robots Syst.*, San Francisco, CA, 2011, pp. 2953–2958.

[24] R. Richa, B. Vagvolgy, M. Balicki, G. Hager, and R. Taylor, "Hybrid tracking and mosaicking for information augmentation in retinal surgery," in *Proc. MICCAI'12*, Nice, France, 2012, pp. 397–404.

[25] G. Scandaroli, M. Meilland, and R. Richa, "Improving NCC-based direct visual tracking," in *Proc. Eur. Conf. Comput. Vis.*, Firenze, Italy, 2012, pp. 442–455.

[26] G. Silveira and E. Malis, "Real-time visual tracking under arbitrary illumination changes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Minneapolis, MN, 2007, pp. 1–6.

[27] G. Silveira and E. Malis, "Unified direct visual tracking of rigid and deformable surfaces under generic illumination changes in grayscale and color images," *Int. J. Comput. Vis.*, vol. 89, no. 1, pp. 84–105, 2010.

[28] C. Stewart, L. Tsai, and B. Roysam, "The dual-bootstrap iterative closest point algorithm with application to retinal image registration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1379–1394, Nov. 2003.

[29] R. Szeliski and H. Shum, "Creating full view panoramic image mosaics and environment maps," in *Proc. 24th Annu. Conf. Comput. Graph. Interactive Tech.*, New York, 1997, pp. 251–258.

[30] R. Sznitman, D. Rother, J. Handa, P. Gehlbach, G. Hager, and R. Taylor, "Adaptive multispectral illumination for retinal microsurgery," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI'10)*. London, U.K.: Springer, 2010, vol. 6363, pp. 465–472.

[31] H. Taylor and J. Keeffe, "World blindness: A 21st century perspective," *Br. J. Ophthalmol.*, vol. 85, no. 3, pp. 261–266, 2001.

[32] J. Totz, P. Mountney, D. Stoyanov, and G.-Z. Yang, "Dense Surface Reconstruction for Enhanced Navigation in MIS," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI'11)*. New York: Springer, 2011, Lecture Notes Computer Science, pp. 89–96.

[33] T. Vercauteren, A. Perchant, G. Malandain, X. Pennec, and N. Ayache, "Robust mosaicing with correction of motion distortions and tissue deformations for *in vivo* fibered microscopy," *Med. Image Anal.*, vol. 10, no. 5, pp. 673–692, 2006.

[34] M. Visentini-Scarzanella, G. P. Mylonas, D. Stoyanov, and G. Z. Yang, "i-BRUSH: A gaze-contigent virtual paintbrush for dense 3D reconstruction in robotic assisted surgery," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI '09)*. New York: Springer, 2009, Lecture Notes in Computer Science, pp. 353–360.