Multimodal Registration Procedure for the Initial Spatial Alignment of a Retinal Video Sequence to a Retinal Composite Image

A. Martina Broehan, Christoph Tappeiner, Simon P. Rothenbuehler, Tobias Rudolph, Christoph A. Amstutz, and Jens H. Kowal*

Abstract—Accurate placement of lesions is crucial for the effectiveness and safety of a retinal laser photocoagulation treatment. Computer assistance provides the capability for improvements to treatment accuracy and execution time. The idea is to use video frames acquired from a scanning digital ophthalmoscope (SDO) to compensate for retinal motion during laser treatment. This paper presents a method for the multimodal registration of the initial frame from an SDO retinal video sequence to a retinal composite image, which may contain a treatment plan. The retinal registration procedure comprises the following steps: 1) detection of vessel centerline points and identification of the optic disc; 2) prealignment of the video frame and the composite image based on optic disc parameters; and 3) iterative matching of the detected vessel centerline points in expanding matching regions. This registration algorithm was designed for the initialization of a real-time registration procedure that registers the subsequent video frames to the composite image. The algorithm demonstrated its capability to register various pairs of SDO video frames and composite images acquired from patients.

Index Terms—Biomedical image processing, feature extraction, retinal image registration, retinal video sequences.

I. INTRODUCTION

LASERS have developed into an ubiquitous tool in ophthalmology [1]. Retinal laser photocoagulation was the first medical laser application [2] and is today an established and important treatment method for an extensive variety of retinal diseases such as diabetic retinopathy and age-related macular degeneration, both of which are the leading causes of adult blindness in the developed world [3], [4].

Computer assistance offers numerous potential improvements to retinal laser treatment such as increased accuracy, reduced execution time, improved safety, and therefore, ultimately increased treatment efficacy. In addition, it has the potential to improve diagnostic capability and allow for treatment planning and documentation. Image registration is an essential aspect of a computer-assisted retinal laser treatment system. It is used to align images from different modalities for diagnostic purposes and to continuously register the patients' anatomy (retina) to a reference image during treatment execution. A possible realization of such a real-time registration is based on the acquisition of a digital video stream, whose individual frames need to be registered to a reference image to compensate for retinal motion during the treatment.

Retinal image registration is an established and ongoing research field. Mostly the intention is to register single retinal image pairs. These are either retinal images of different modalities (e.g., a fundus image and an angiogram) or images of the same modality taken at different points of time (temporal registration). Predominantly, algorithms used for retinal image registration are based on a segmentation of the vessel tree and/or extraction of significant vessel features (e.g., vessel bifurcations). Therefore, a substantial amount of research effort has been directed toward retinal vessel extraction [5]–[8], and vessel landmark-based retinal image registration [9]–[15]. Study has also been done on methods that use intensity-based similarity metrics [16]–[18].

Computer assistance for retinal laser photocoagulation has only been approached by a few groups. Welch [19] introduced a computer-controlled retinal surgery system that included retinal tracking and automatic laser beam deviation. This study was continued by the group of Markow *et al.* [20], Barrett *et al.* [21], and Wright *et al.* [22]. They proposed a computer-assisted laser delivery system that would incorporate two different tracking (digital and analog) methods, a lesion depth control system and laser beam deviation via galvanometer-controlled mirrors. Solouma *et al.* [23] reported on the development of a real-time movement tracking system for image-guided laser treatment. The system provides real-time alignment of a preoperative treatment plan with acquired video images. The group of Becker *et al.* [24] and Shen *et al.* [25] developed variants of their registration techniques that allowed for real-time

A. M. Broehan and T. Rudolph are with the artificial organ (ARTORG) Center for Biomedical Engineering Research, University of Bern, Bern 3014, Switzerland, and also with the Department of Ophthalmology, University Hospital Bern, Bern 3010, Switzerland.

C. Tappeiner and S. P. Rothenbuehler are with the Department of Ophthalmology, University Hospital Bern, Bern 3010, Switzerland.

C. A. Amstutz is with the ARTORG Center for Biomedical Engineering Research, University of Bern, Bern 3014, Switzerland, and also with the Department of Ophthalmology, University Hospital Zurich, CH-8091 Zurich, Switzerland.

*J. H. Kowal is with the ARTORG Center for Biomedical Engineering Research, University of Bern, Bern 3014, Switzerland, and also with the Department of Ophthalmology, University Hospital Bern, Bern 3010, Switzerland (e-mail: jens.kowal@artorg.unibe.ch).

registration of retinal images with potential to be incorporated into an automatic laser photocoagulation system. Our approach toward a computer-assisted retinal laser photocoagulation system is based on a retina composite image that provides a large field of view (FOV) to aid laser treatment planning and video frames captured from a scanning digital ophthalmoscope (SDO), which can be equipped with a laser unit and provides a color video stream. The approaches presented in [20], [21], [23], [24] use either incremental tracking methods or use a reference image and video frames acquired with a fundus camera, which reduces the registration task to a monomodal problem with mostly translational or rigid transformations. In this paper, we present a new multimodal registration procedure for the spatial alignment of the first frame from an SDO retinal video sequence to a retinal composite image using retinal features and a nonlinear transformation model. The procedure is particularly challenging because substantial differences in image quality (image noise, nonuniform illumination, and motion blur), FOV, image scale, and spatial resolution may have to be taken into account. The method would initialize a real-time retinal registration procedure as contribution to a complete computer-assisted laser photocoagulation system.

The proposed multimodal registration procedure identifies points on the vessel centerline and the optic disc in the composite image and the video frame. The detected optic discs are used to establish a prealignment. Subsequently, the vessel centerline points of the SDO video frame are iteratively matched to the segmented composite image to refine the prealignment and obtain a global nonlinear registration. For the extraction of vessel points on the composite image, an iterative vessel tracing method as proposed in [5] is used, but adjusted with respect to new tracing direction selection, directional template application range, and local edge strength thresholding, to better meet the demands of our specific image modalities.

As opposed to numerous other retinal image registration methods [9], [24]–[26], the retinal registration procedure presented in this paper does not rely on the identification of vessel bifurcations. For our application, the accurate extraction of well-distributed corresponding vessel bifurcations in both image modalities, which is necessary for a good registration result, failed in numerous cases. We make use of the optic disc to establish a prealignment, which is a unique feature on the retina, and therefore, eliminates the need to look for correspondences. In addition, comparing the center and radius of two optic discs provides information about translation and differences in scale. Optic disc detection in retinal images is a topic that has been tackled in a number of publications [27]–[31]. We present a new method for the optic disc detection in retinal images that incorporates techniques from existing methods such as mean intensity variance computation [28] for identifying the high intensity variations between vessels and background in the optic disc region and Hough transform application [31]–[33] for calculating the center and radius. In our developed method, we constrain intensity variance calculations to regions at detected vessel points and limit the Hough transform calculation to the established region of interest (ROI) for the optic disc, which makes the method more efficient and reliable.
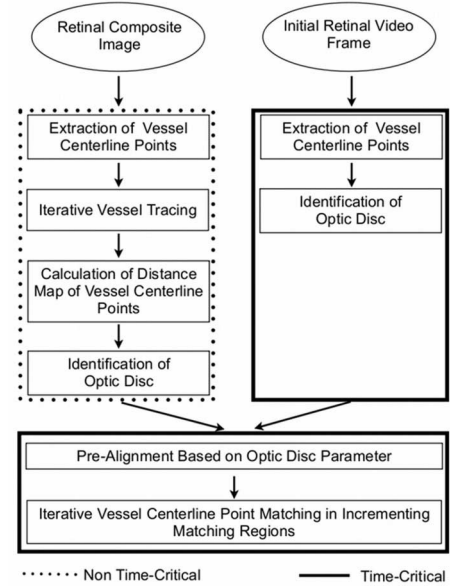

Fig. 1.   Outline of the proposed registration procedure.

II. Methods

Fig. 1 provides an overview of the proposed registration procedure and gives the scheduling information of the algorithm. All processing steps on the composite image such as vessel centerline point and optic disc extraction, as well as the computation of a distance map used for the iterative point matching, can be performed offline (before a laser intervention starts). On the composite image, which serves as a reference, a tracing of the vasculature is performed to obtain an extensive segmentation of the vessel centerline. On the video frame, only vessel centerline points along a fine grid of vertical and horizontal lines are detected to reduce the computational load. All color composite images and video frames are converted to grayscale images using the following weighting: $I = 0.299R + 0.587G + 0.114B$, where R, G, and B are the red, green, and blue color channel.

A. Detecting Vessel Centerline Points

In both retinal image modalities, vessels appear darker than the background. In both modalities, points on the vessel centerline are detected by analyzing the images along a grid of vertical and horizontal lines and searching for local intensity minima bounded by two edges [5]. Intensity values on each line are low-pass filtered using a discrete approximation to a 1-D Gaussian kernel with a standard deviation set to 1. At each pixel location along the grid lines, the derivative value $I'(x, y)$ is computed. For a horizontal line, $I'(x, y)$ is calculated as [34]

$$I'(x,y) = -I(x-2,y) - 2I(x-1,y) + 2I(x+1,y) + I(x+2,y). \quad (1)$$

Derivative values along a vertical line are computed accordingly. A linear search along the grid lines looks for derivatives with the greatest local magnitude with opposite signs that lay within a given vessel width range. Based on measurements from numerous images, the vessel width range is set between 4–20 pixels
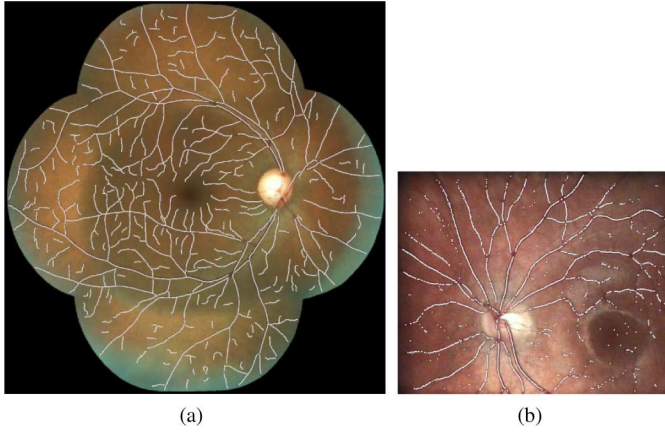
Fig. 2. Extracted vessel centerline points. (a) Retinal composite image with vessel centerline points extracted through iterative vessel tracing. Seed points were detected with a grid spacing of 30 pixels. (b) Retinal video frame with extracted vessel centerline points, which were detected with a grid spacing of two pixels and no subsequent vessel tracing.

for composite images and between 4–16 pixels for video frames. The intensity value at the location between two maximal derivative magnitudes, where the derivative value changes sign from negative to positive, is considered a local intensity minimum. False detection of intensity minima and vessel edges is minimized by establishing appropriate intensity and edge strength thresholding as proposed by Fritzsche *et al.* [34]. Thresholding is applied locally to cope with nonuniform illumination frequently occurring in retinal images.

Local intensity minima are not unique to vessels and may also occur on retinal background caused by image noise or pathologies. Therefore, only intensity minima below a threshold t_I are considered. The intensity values within a rectangular region w_I are denoted I_{w_I}. The intensity threshold $t_I(I_{w_I})$ is expressed as

$$t_I(I_{w_I}) = m(I_{w_I}) - s(I_{w_I}) \tag{2}$$

where $m(I_{w_I})$ is the median and $s(I_{w_I})$ is the median of the absolute deviations of all intensity values I_{w_I} from $m(I_{w_I})$. The edge strength threshold represents the cutoff between what is considered a vessel boundary and what is considered background noise. At each pixel location within a rectangular neighborhood $w_{I'}$ of a possible vessel edge location, a set of derivatives in four directions ($0°, 45°, 90°,$ and $135°$, respectively) is calculated. This derivative set is referred to as $I'_{w_{I'}}$. The mean $\mu(I'_{w_{I'}})$ and standard deviation $\sigma(I'_{w_{I'}})$ describe the local contrast and are used to set up an edge strength threshold $t_E(I'_{w_{I'}})$ as follows:

$$t_E(I'_{w_{I'}}) = \mu(I'_{w_{I'}}) + 0.5\sigma(I'_{w_{I'}}). \tag{3}$$

The size of w_I and $w_{I'}$ was set to 25×25 pixels for video frames and 50×50 pixels for composite images. If both opposite signed derivative magnitudes are above the edge strength threshold, the enclosed local intensity minimum is considered a point on a vessel. Fig. 2 illustrates a video frame with detected vessel points.

B. Iteratively Tracing the Vasculature

On the composite image, an iterative tracing of the vasculature is performed. The vessel tracing method proposed by Can *et al.* [5] is applied with a few variations. We modified the calculation of the new tracing direction, limited the range for applying the templates in each direction, and introduced local edge strength thresholding. The main idea is to use a set of 16 left and right 2-D directional templates that detects the parallel edges of a vessel in an iterative manner. The previously detected points on the vessel centerline are used as seed points. At each seed point, the iterative tracing procedure is initiated in two opposing directions.

$L(x, y, s)$ and $R(x, y, s)$ denote the left and right template responses; in other words, the correlations between the image data and the left and right templates in direction s (as defined in [5]). Direction s is represented by an index value $\{0, 1, \ldots, 15\}$ corresponding to one of the 16 discrete directions. $R_{\max}(s)$ and $L_{\max}(s)$ denote the maximal right and left template responses in direction s. Starting from a point p_k and a direction s_k, the directional templates are applied to estimate the next point p_{k+1} on the vessel and its orientation s_{k+1}. There are only small changes in vessel width expected when tracing a single vessel segment; therefore, the maximal range for applying the templates in a direction s is set to $(d_k/2) \pm 2$, where d_k is the detected vessel diameter at point p_k. As opposed to [5], where the new direction s_{k+1} has been chosen as the direction $s \in \{s_k - 1, s_k, s_k + 1\}$ with the maximal $R_{\max}(s)$ or $L_{\max}(s)$ value ("following the strongest edge"), we choose the new tracing direction as the direction, where the sum of the maximal left and right template responses is maximum ("following the strongest edge pair"). Mathematically

$$s_{k+1} = \max_{s \in s_k - 1, s_k, s_k + 1}\{R_{\max}(s) + L_{\max}(s)\}. \tag{4}$$

The next point p_{k+1} on the vessel centerline is estimated as in [5]

$$p_{k+1} = p_k + lu_{k+1} + \beta_{k+1} \tag{5}$$

where k is the iteration number, l is the step size, β_{k+1} is a lateral refinement vector for adjusting the location of the point to the vessel centerline (see [5]), and u_{k+1} is a unit vector defined as

$$u_{k+1} = \left[\cos\left(\frac{2\pi s_{k+1}}{16}\right), \sin\left(\frac{2\pi s_{k+1}}{16}\right)\right]. \tag{6}$$

Local edge strength thresholding, similar to that described in (3), is applied at the locations of the maximal template responses, $R_{\max}$ and $L_{\max}$, to assure that both detected edges are strong enough to be considered a vessel edge. Thus, we established the following criterion:

$$R_{\max} > t_E \text{ and } L_{\max} > t_E. \tag{7}$$

A tracing step is terminated, if the criterion in (7) is not fulfilled three times in a row, if the new vessel centerline point p_{k+1} has already been traced, or if the trace meets the image border. The result of an iterative vessel tracing on a composite image is shown in Fig. 2.

C. Extracting the Optic Disc

We developed a new method for the optic disc detection in retinal images that is efficient with respect to run-time speed and flexible enough to detect the optic disc in large composite images as well as small SDO video frames. In addition, it was reasonable to make use of the detected vessel points, which are already available for later matching. The optic disc is a bright circular structure, where the major blood vessels converge. Our proposed optic disc detection procedure exploits these characteristics, uses the previously extracted vessel centerline points, and exploits the fact that particularly high-intensity variations between vessels and background occur in the area of the optic disc. We employed ideas from existing methods, such as mean intensity variance computation [28] and Hough transform application [31]–[33], but we constrained the intensity variance calculations to regions at detected vessel points and established a ROI for calculating the Hough transform in favor of computation time and reliability.

In a first step, a local intensity variance value is calculated at each vessel centerline point, average values of these variances are computed, and the location of the highest average variance value is determined. A ROI is defined around the location of the highest average intensity variance. In a second step, the bright circular structure of the optic disc is identified within the previously defined ROI by applying the circular Hough transform on intensity derivatives. The grayscale composite image and initial video frame are downsampled by a factor of two using linear interpolation. Gaussian filtering is applied prior to resampling.

1) Defining the ROI for the Optic Disc: Let V be a set of intensity values v in a neighborhood of a vessel centerline point. The intensity variance value $\sigma_I{}^2(V)$ is calculated as

$$\sigma_I{}^2(V) = \frac{1}{N_V} \sum_{v \in V} (v - \mu(V))^2 \tag{8}$$

where N_V is the number of values in V and $\mu(V)$ is the mean. Single high-intensity variance values may also occur due to noise, rapid illumination changes, and pathologies. Calculating the average intensity variance effectively suppresses these single peaks. This effect is demonstrated in Fig. 3. Let U be a set of intensity variance values $\sigma_I{}^2$ in a neighborhood of a centerline point. The average intensity variance $\bar{\sigma}_I{}^2(U)$ is given by

$$\bar{\sigma}_I{}^2(U) = \frac{1}{N_U} \sum_{\sigma_I{}^2 \in U} \sigma_I{}^2 \tag{9}$$

where N_U is the number of values in U. The size of the square neighborhood for variance and mean variance calculation is defined by the expected diameter of the optic disc, which has been empirically determined as 160 pixels for a composite image and 120 pixels for a video frame, respectively. The location of the highest average intensity variance value defines the center of the ROI and its dimensions are $s \times s$, where $s = 3 \times$ "expected optic disc diameter". This ensures that a sufficiently large area around the optic disc is covered for Hough transform application because the detected center of the ROI is not necessarily close to the optic disc center, but may be anywhere in the optic disc area.

(a)

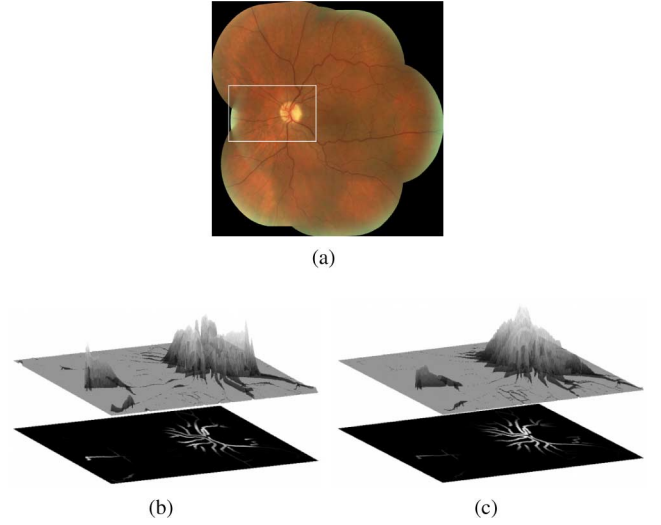(b)   (c)

Fig. 3. Illustration of the intensity variance and average intensity variance at vessel centerline points. (a) Retinal composite image. The region identified is depicted in (b) and (c). (b) Intensity variance at vessel centerline points. High variance values are visible in the area of the optic disc, but in the region of bright illumination near the image border as well. (c) Average intensity variance at vessel centerline points. High average variance values occur only in the region of the optic disc. Any high variance values in the region of bright illumination are effectively suppressed.

2) Identifying the Optic Disc: The optic disc has an approximately circular shape. The circular Hough transform is, therefore, a suitable method for identifying its center and radius in the previously defined ROI. A circle in Hough space is represented by

$$(x - a)^2 + (y - b)^2 = r^2 \tag{10}$$

where (a, b) is the center and r is the radius. The Hough transform is directly applied to derivative values within the ROI. The directional information of the intensity changes can thereby be taken into account. It allows us to specifically look for a bright circle on a darker background. The x-derivative values X are obtained as follows:

$$X = I(x + 1, y) - I(x, y). \tag{11}$$

The values for the y-derivative Y are obtained in an analogous way. When the transitions from a dark vessel to the retinal background yield higher derivative values than the transitions from the optic disc to the retinal background, a circle may falsely be fitted between converging vessels during Hough transform calculation. To prevent this, a binary vessel mask is generated from the vessel centerline points. Each vessel centerline point is expanded by replacing the single point by a square region with the size of the detected vessel width at that point. Derivative values at locations covered by the binary vessel mask are excluded for Hough transform calculation and, therefore, not taken into account as part of a possible optic disc border. Furthermore, the ROI boundary of the composite image is removed by defining a background threshold.

In the Hough transform, each derivative value contributes a circle of radius r to the accumulation space. The accumulation space is a set of accumulators A_r each representing a circle
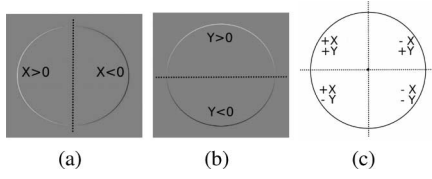
Fig. 4. Illustration of the (a) *x*-derivative and (b) *y*-derivative of a bright circle on darker background. (c) It shows how the *x*- and *y*-derivative values contribute to a circle in accumulator space.
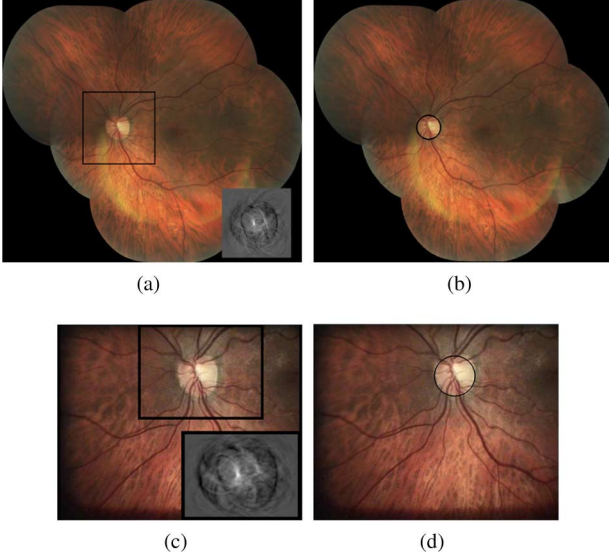

Fig. 5. Optic disc detection. (a) Retinal composite image with the labeled ROI and the corresponding accumulator revealing the peak value. (b) Segmented optic disc of a retinal composite image. (c) Retinal video frame with the labeled ROI and the corresponding accumulator revealing the peak value. (d) Segmented optic disc of a retinal video frame.

radius r. The accumulation space covers an optic disc radius range of $r = 40 \pm 5$ pixels for the downsampled composite images and $r = 30 \pm 5$ pixels for the downsampled video frames. To explicitly look for a bright circle, the circle is divided in four regions according to [35]. In each region, the derivative values contribute differently. For the X values, the circle is divided into left and right regions. On the left side, high X values are needed, since we are looking for a rising edge. On the right side, low X values are required that represent a falling edge. For Y values, the circle is divided into upper and lower regions (see Fig. 4). The accumulators reveal a peak, where the contributory circles overlap. Thus, the location of the highest value $a_{\max}$ from all accumulators is assumed to be the optic disc's center (a, b). The accumulator A_r, which contains $a_{\max}$, defines the optic disc's radius r. Detected optic discs are presented in Fig. 5.

D. Determining the Prealignment Matrix

After the optic disc is located in both modalities, a prealignment step is performed. From the center (a, b) and radius r of the optic discs, an initial scale and translation estimate is computed.
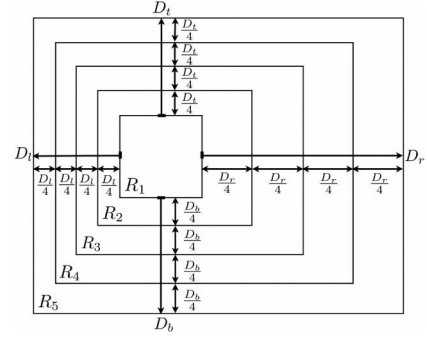

Fig. 6. Outline of the five incrementing matching regions for vessel point matching.

The scale estimate is given by

$$S = \frac{r_R}{r_F} \tag{12}$$

where r_R is the radius of the optic disc in the composite image and r_F is the radius of the optic disc in the video frame. The translation vector is computed by

$$T = \begin{pmatrix} a_R \\ b_R \end{pmatrix} - \begin{pmatrix} a_F \\ b_F \end{pmatrix} \tag{13}$$

where (a_R, b_R) are the coordinates of the optic disc's center in the composite image and (a_F, b_F) are the coordinates of the optic disc's center in the video frame. From the scale and translation estimates, a prealignment matrix M_{pre} is constructed as

$$M_{\mathrm{pre}} = \begin{pmatrix} S & 0 \\ 0 & S \end{pmatrix} T = \begin{pmatrix} S & 0 & t_x \\ 0 & S & t_y \end{pmatrix}. \tag{14}$$

This matrix is applied to the initial video frame to prealign it to the composite image.

E. Matching of Vessel Centerline Points

The prealignment is refined by iteratively minimizing distances between vessel points of both images. This matching procedure requires robust outlier handling to cope with distance measures from missing or incorrect vessel points. This can be realized by using the robust error function of an M-estimator, which assigns appropriate weights to the distances during the matching process. Vessel point matching is performed in expanding matching regions until the entire video frame is spatially aligned. As the matching region expands, the transformation model changes from affine to quadratic (second-order polynomial). Expanding the matching region for vessel point matching and refining the transformation model is a concept presented in [10]. Overall, we define five matching regions $R_\kappa, \kappa = 1, \ldots, 5$, of increasing area. The initial region R_1 is formed around the optic disc center in the composite image, which is already aligned with the optic disc center in the video frame. The size of R_1 is set to $2r_R \times 2r_R$. The regions grow linearly up to R_5, which covers the entire video frame. Fig. 6 outlines the matching regions. The initial transformation matrix $\hat{M}_\kappa$ for each matching region R_κ is the final transformation

estimate from the previous region. For R_1, the initial transformation matrix is $\hat{M}_1 = M_{\mathrm{pre}}$. In matching region R_3, where the transformation parameter estimate represents a good approximation to the global estimate, the transformation model changes from affine to quadratic.

The affine transformation of a point $p = (x, y)$ with transformation matrix $M(\theta)$ is defined as

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = M(\theta, p) = \begin{pmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (15)$$

and for a quadratic transformation model, the definition is

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = M(\theta, p) = \begin{pmatrix} \theta_{11} & \theta_{12} & \theta_{13} & \theta_{14} & \theta_{15} & \theta_{16} \\ \theta_{21} & \theta_{22} & \theta_{23} & \theta_{24} & \theta_{25} & \theta_{26} \end{pmatrix}$$
$$\times \begin{pmatrix} x \\ y \\ 1 \\ x^2 \\ xy \\ y^2 \end{pmatrix}. \quad (16)$$

Let P be the detected vessel centerline points of the composite image and Q be the vessel centerline points of the initial video frame. From the points in P, a distance map is calculated by applying the method described in [36]. The robust iterative point matching procedure takes the point locations from Q that are in the current matching region R_κ, applies the current transformation parameter estimate θ_t, and finds the closest centerline point locations in P in the precalculated distance map. From the resulting set of correspondences $C_t = (p_i, q_j)$, a new transformation estimate is computed by minimizing the robust objective function

$$E(\theta_t) = \sum_{(p_i, q_j) \in C_t} \frac{\rho(d(p_i, M_t(\theta_t, q_j)))}{\hat{\sigma}} \quad (17)$$

where M_t is the matrix that maps q_j into the composite image, d is the Euclidean distance, $\hat{\sigma}$ is a robust scale estimate, and ρ is a robust loss function of an M-estimator. Iterative reweighted least squares (IRLS) is applied to minimize the objective function in (17). IRLS includes alternating steps of calculating weights ω_t using the current transformation parameter estimate θ_t and calculating new weighted least squares estimates of θ. These alternating steps are repeated until the parameter estimate converges. The weight function introduced is the Beaton–Tukey biweight function [37], which allows an aggressive rejection of outliers. The limit beyond which weights are set to zero is about $4\hat{\sigma}$ in statistics. Mathematically

$$\omega(u) = \rho(u)'/u = \begin{cases} (1 - (\frac{u}{a})^2)^2, & |u| \le c \\ 0, & |u| > c \end{cases} \quad (18)$$

where $u = (d(p_i, M(\theta, q_j)))/\hat{\sigma}$ and $c = 4$. Estimation of error scale $\hat{\sigma}$ is done for each set of correspondences. At the start of IRLS, a minimum unbiased scale estimator [38] is used to construct an unbiased scale estimate from the k smallest out

of n absolute residuals $\epsilon_j = |d(p_i, M(\theta_t, q_j))|$. As the IRLS procedure iterates, $\hat{\sigma}$ is defined by a robustly weighted rms [11] given by

$$\hat{\sigma}^2 = \frac{\sum_j \omega_j \epsilon_j^2}{\sum_j \omega_j}. \quad (19)$$

The iterative point matching is performed for each defined matching region and the final registration matrix is given by the matrix obtained from R_5. In Fig. 7, different stages of the registration procedure are visualized.

III. EXPERIMENTAL ANALYSIS

All test data used for this study were recorded from outpatient volunteers at the Ophthalmology Department coming for a regular appointment (age 24–91 years, mean 60.2 years, 66% male, 34% female). Fundus images of seven fields were acquired using a VISUCAM NM/FA nonmydriatic fundus camera (Carl Zeiss Meditec, Jena, Germany) with a $45°$ FOV. The built-in mosaicing procedure was utilized to construct the composite images. The resulting composite images vary in spatial resolution due to differences in image overlap and exclusion of single images by the mosaicing procedure. The resolution along the horizontal axis is between 1723 and 2171 pixels and between 1731 and 2027 pixels along the vertical axis, respectively, with a pixel size of approximately 9.4 μm. Video data were captured with an SDO-A scanning digital ophthalmoscope (WILD Medtec, Wien, Austria) connected to a standard digital camcorder. Each SDO video frame has a resolution of 720×576 pixels and a pixel size of approximately 12.5 μm. Most of the video sequences were acquired in a red-free illumination mode.

The proposed registration procedure was applied to 35 composite image/SDO video frame pairs covering a wide range of image qualities and retinal pathologies. For 31 of the 35 pairs, well-aligned vasculature was observed when presented in checkerboard overlay (88.5% successfully registered). The registration of the remaining image pairs failed due to errors detecting the optic disc. For three cases, the optic disc could not be correctly identified in the SDO video frame, and for one case, the optic disc could not be correctly identified in the composite image. Some registration results are shown in Fig. 8. The accuracy of the proposed registration method was quantitatively assessed for each registered composite image/SDO video frame pair. Two ophthalmologists and a researcher working in the field of ophthalmology were asked to manually select 10–15 corresponding control points spread over the overlap area of the images. The registration matrix estimated with our proposed registration procedure was applied to the control points of the SDO video frame. A registration error for each image pair was given by the mean Euclidean distance between the control points of the composite image and the transformed control points of the SDO video frame (see Fig. 9). For each observer, the mean and standard deviation of the registration errors was computed. The results are presented in Table I. The total average registration error obtained for the presented multimodal retinal registration algorithm using a quadratic transformation model was 2.68 ± 1.62 pixels (approximately $25 \pm 15 \mu$m) The interobserver variations
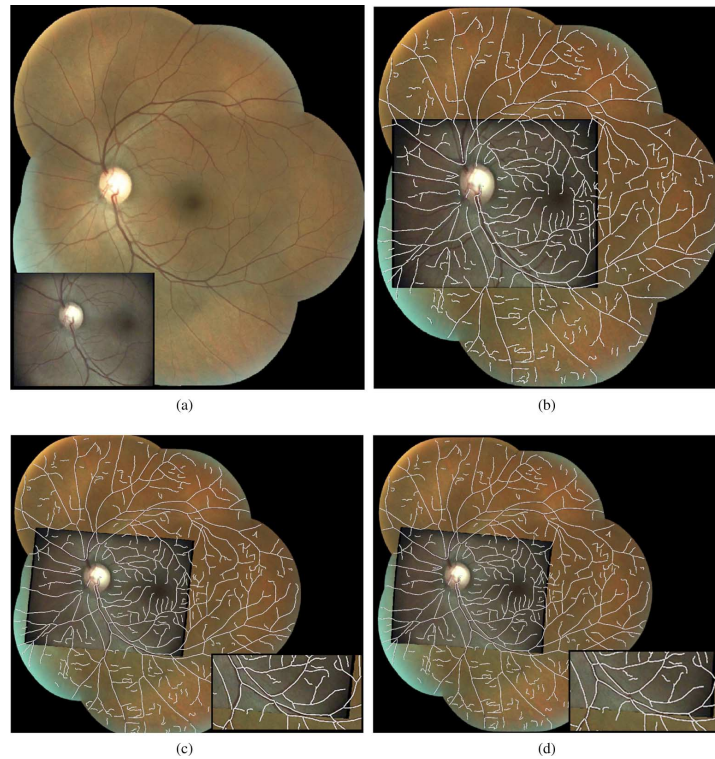
Fig. 7. Different stages of the proposed multimodal registration procedure. (a) Composite image and video frame displayed in the same coordinate system before registration. (b) Prealignment using the optic disc. The video frame is translated and scaled, but misalignment of the vasculature is visible. (c) Registration result after performing iterative vessel point matching in three matching regions with the affine transformation model. Small misalignments remain in the lower part of the overlapping region seen in magnified view. (d) Final registration result with the quadratic transformation model. The vasculature in all regions appears well aligned.
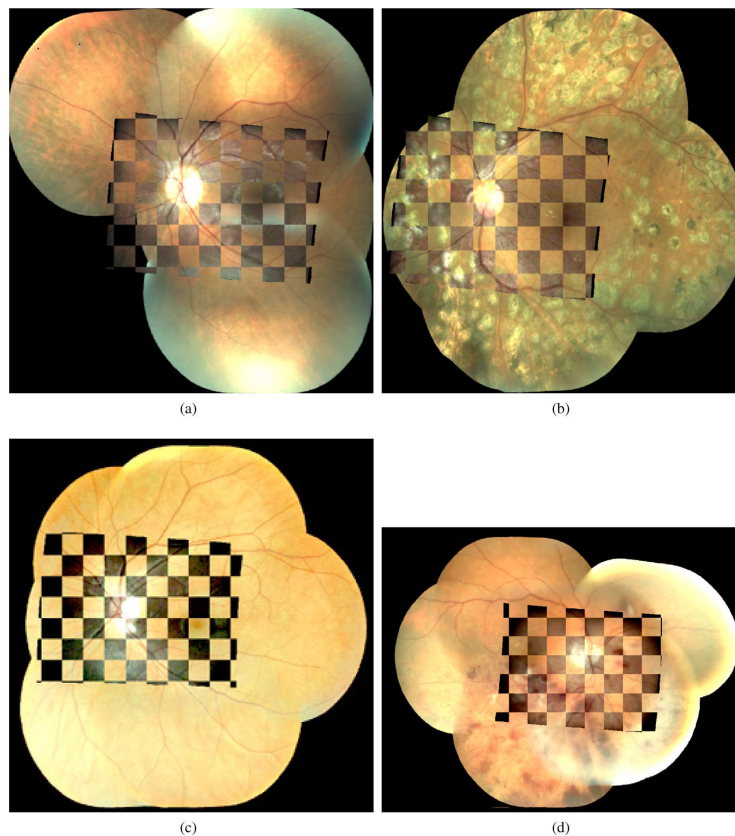

Fig. 8. Registration results. (a)–(d) Pairs of retinal composite images and SDO retinal video frames registered with the proposed multimodal registration procedure. The registered SDO video frames are displayed as a checkerboard overlay onto the retinal composite images. The image pairs evaluated exemplify the diversity of images encountered in clinics.
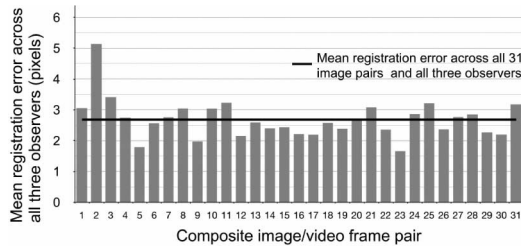
Fig. 9.   Mean registration error across all three observers for each of the 31 registered composite image/SDO video frame pairs is illustrated.

TABLE I
MEAN ± STANDARD DEVIATION OF THE REGISTRATION ERROR
IN PIXELS ACROSS 31 IMAGE PAIRS

Observer 1	2.46 ± 1.32
Observer 2	2.89 ± 1.70
Observer 3	2.70 ± 1.84
Average All Observers	2.68 ± 1.62

were small and the accuracy results consistent. The presented accuracy validation procedure is influenced by the ability of the individual observers to correctly identify corresponding control points with a software tool we provided. To quantify this influence, the average error in point placement was determined which was found to be 0.8 pixels. The average computation time required to register an SDO video frame to a segmented composite image was approximately 1.2 s on a 2.4-GHz Intel Core 2 Duo. The processing steps on the composite image are performed in 5 s on average. The actual runtime depends on the number of extracted vessel points, which relates to the density of the vasculature. The current implementation in C++ has not been optimized for speed. The algorithm was integrated into the MARVIN medical research application framework [39].

IV. DISCUSSION AND CONCLUSION

To our knowledge, this is the first time a registration procedure was specifically created for video frames acquired from an SDO. Furthermore, we believe that this is the first time that the video frame to reference registration for a computer-assisted retinal laser treatment system was considered a multimodal problem.

The average registration error in this study was 2.68 pixels, corresponding to approximately 25 μm in the composite image domain. Our method is not directly comparable to other methods, since there is no other validated registration procedure involving SDO retinal video frames available. However, we stated the registration errors of other retinal registration procedures to classify the range of our obtained accuracy. In doing so, it has to be taken into account that the image characteristics of the respective retinal images and the registration model differ significantly. Difficulties that arise in our specific registration task are scale differences, FOV differences, resolution differences, image noise, motion blur, nonuniform illumination, and a nonlinear transformation model. Furthermore, it has to be considered that the composite images are the result of a mosaicing procedure, which registers and fuses seven-field fundus images, and may introduce additional distortions and artifacts. Ryan

et al. [13] reported an accuracy of 2.9 pixels rmse for the registration of pairs of color fundus photographs and angiograms of 575×480 pixels using an affine transformation model. Can et al. [9] estimated an average centerline error of 2.47 pixels for an affine model and 0.83 pixels for a quadratic model after landmark position refinement using pairs of 1024×1024 pixels fundus photographs. Lee et al. [26] obtained a centerline error of 1.88 pixels for an affine transformation model with radial distortion correction applied to fundus image pairs with a resolution of 598×512 pixels.

A discussion about accuracy has to clearly incorporate its clinical relevance for a laser photocoagulation procedure. For panretinal laser photocoagulation, where hundreds of laser spots are placed in the retinal periphery, it is most important to achieve an equal distribution of lesions. For treatments of the central retina, accuracy and patient safety are paramount. In focal laser treatment, microaneurysms with a typical diameter of 20–100 μm need to be coagulated using a laser spot size of 50–200 μm [40]. A computer-assisted retinal photocoagulation system should allow a planned spot on a microaneurysm to be reproduced by the system in such a way that the microaneurysm would be covered by the laser beam and can be successfully coagulated. Our method fulfills this clinical demand.

Optic disc detection in retinal images is topic of extensive research, but there is no established standard for successful optic disc detection. The shape and the appearance of the optic disc in clinical retinal fundus images varies significantly. If the optic disc cannot be accurately identified, our registration procedure may fail. Potential causes of failure include: pathological deformations, very low contrast between the optic disc and background, and a very bright optic cup within the optic disc. For these cases, the Hough transform may incorrectly fit the circle. To handle difficult clinical cases, we provide the ophthalmologist, in addition to the full-automatic registration procedure, the option to semiautomatically detect the optic disc, which requires very little intervention (one mouse click to identify the optic disc center). The subsequent parts of the registration procedure are then executed automatically to perform an initial spatial alignment.

Our method is sensitive to rotations, but generally rotations in the image plane are small, since the patient's head is constrained to an upright position for image acquisition and laser treatment. Situations with large rotations have, therefore, not been evaluated. The final transformation model in the registration procedure is the quadratic. Arbitrary distortions were occasionally observed when applying this model to image pairs with only a small overlapping region as demonstrated in [9], [17], [26]. This phenomenon was not observed when testing our registration procedure, because our test pairs have high overlap.

As part of a retinal laser photocoagulation system to initialize a real-time registration procedure, the method presented in this paper needs to be sufficiently fast and further optimization of the current implementation is required. Our intention is to implement parts of our algorithm in CUDA, which allows parallel computing in the graphics processing units. A first attempt to implement the vessel point detection method in CUDA revealed

a tenfold increase in speed. This confirms our opinion that significant potential for further speed improvements remain.

We have introduced a new multimodal registration method that registers the initial frame from an SDO retinal video sequence to a retinal composite image, which can potentially contain a laser treatment plan. This method was designed to be used within a computer-assisted retinal laser photocoagulation system to initialize a real-time registration procedure, which may use the segmented features from the composite image and the spatial information obtained from the initial registration. The algorithm demonstrated its capability to register various pairs of SDO video frames and composite images acquired from patients with respectable accuracy. The registration procedure has the potential to be modified for other retinal modalities. The development of a real-time registration method to register the subsequent video frames from a retinal video sequence to a reference composite image is underway.

ACKNOWLEDGMENT

REFERENCES

[1] J. M. Krauss and C. A. Puliafito, "Lasers in ophthalmology," *Lasers Surg. Med.*, vol. 17, no. 2, pp. 102–159, 1995.

[2] F. Fankhauser and S. Kwasniewska, Eds., *Lasers in Ophthalmology – Basic, Diagnostic and Surgical Aspects: A Review*. The Hague, The Netherlands: Kugler Publications, 2003.

[3] The Eye Diseases Prevalence Research Group, "The prevalence of diabetic retinopathy among adults in the united states," *Arch. Ophthalmol.*, vol. 122, no. 4, pp. 552–563, Apr. 2004.

[4] The Eye Diseases Prevalence Research Group, "Causes and prevalence of visual impairment among adults in the united states," *Arch. Ophthalmol.*, vol. 122, no. 4, pp. 477–485, Apr. 2004.

[5] A. Can, H. Shen, J. N. Turner, H. L. Tanenbaum, and B. Roysam, "Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms," *IEEE Trans. Inf. Technol. Biomed.*, vol. 3, no. 2, pp. 125–138, Jun. 1999.

[6] A. D. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Trans. Med. Imag.*, vol. 19, no. 3, pp. 203–210, Mar. 2000.

[7] C.-L. Tsai, C. V. Stewart, H. L. Tanenbaum, and B. Roysam, "Model-based method for improving the accuracy and repeatability of estimating vascular bifurcations and crossovers from retinal fundus images," *IEEE Trans. Inf. Technol. Biomed.*, vol. 8, no. 2, pp. 122–130, Jun. 2004.

[8] M. Sofka and C. V. Stewart, "Retinal vessel centerline extraction using multiscale matched filters, confidence and edge measures," *IEEE Trans. Med. Imag.*, vol. 25, no. 12, pp. 1531–1546, Dec. 2006.

[9] A. Can, C. V. Stewart, B. Roysam, and H. L. 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.

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

[11] C. V. Stewart, C.-L. Tsai, and A. Perera, "A view-based approach to registration: theory and application to vascular image registration," *Inf. Process Med. Imag.*, vol. 18, pp. 475–486, Jul. 2003.

[12] F. Laliberte, L. Gagnon, and Y. Sheng, "Registration and fusion of retinal images–an evaluation study," *IEEE Trans. Med. Imag.*, vol. 22, no. 5, pp. 661–673, May 2003.

[13] N. Ryan, C. Heneghan, and P. de Chazal, "Registration of digital retinal images using landmark correspondence by expectation maximization," *Image Vis. Comput.*, vol. 22, no. 11, pp. 883–898, Sep. 2004.

[14] F. Zana and J. C. Klein, "A multimodal registration algorithm of eye fundus images using vessels detection and hough transform," *IEEE Trans. Med. Imag.*, vol. 18, no. 5, pp. 419–428, May 1999.

[15] G. K. Matsopoulos, P. A. Asvestas, N. A. Mouravliansky, and K. K. Delibasis, "Multimodal registration of retinal images using self organizing maps," *IEEE Trans. Med. Imag.*, vol. 23, no. 12, pp. 1557–1563, Dec. 2004.

[16] N. Ritter, R. Owens, J. Cooper, R. H. Eikelboom, and P. P. Van Saarloos, "Registration of stereo and temporal images of the retina," *IEEE Trans. Med. Imag.*, vol. 18, no. 5, pp. 404–418, May 1999.

[17] T. Chanwimaluang, G. Fan, and S. R. Fransen, "Hybrid retinal image registration," *IEEE Trans. Inf. Technol. Biomed.*, vol. 1, no. 1, pp. 129–142, Jan. 2006.

[18] R. Kolar, L. Kubecka, and J. Jan, "Registration and fusion of the autofluorescent and infrared retinal images," *Int. J. Biomed. Imag.*, vol. 2008, pp. 513478-1–513478-11, 2008.

[19] A. J. Welch, "University of texas lab studies tissue optics, ablation, automation," *Biomed. Opt.: Newslett. Biomed. Opt. Soc.*, vol. 2, no. 2, May 1993.

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

[21] S. F. Barrett, C. H. G. Wright, H. Zwick, M. Wilcox, B. A. Rockwell, and E. Naess, "Efficiently tracking a moving object in two-dimensional image space," *J. Electron. Imag.*, vol. 10, no. 3, pp. 785–793, 2001.

[22] C. H. G. Wright, S. F. Barrett, and A. J. Welch, "Design and development of a computer-assisted retinal laser surgery system," *J. Biomed. Opt.*, vol. 11, no. 4, p. 041127, 2006.

[23] N. H. Solouma, A. B. M. Youssef, Y. A. Badr, and Y. M. Kadah, "A new real-time retinal tracking system for image-guided laser treatment," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 9, pp. 1059–1067, Sep. 2002.

[24] D. E. Becker, A. Can, J. N. Turner, H. L. Tanenbaum, and B. Roysam, "Image processing algorithms for retinal montage synthesis, mapping, and real-time location determination," *IEEE Trans. Biomed. Eng.*, vol. 45, no. 1, pp. 105–118, Jan. 1998.

[25] H. Shen, C. V. Stewart, B. Roysam, G. Lin, and H. L. Tanenbaum, "Frame-rate spatial referencing based on invariant indexing and alignment with application to online retinal image registration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 3, pp. 379–384, Mar. 2003.

[26] S. Lee, M. Abramoff, and J. M. Reinhardt, "Feature-based pairwise retinal image registration by radial distortion correction," presented at the SPIE Conf. Med. Imag., San Diego, CA, 2007.

[27] M. Niemeijer, M. D. Abramoff, and B. van Ginneken, "Segmentation of the optic disc, macula and vascular arch in fundus photographs," *IEEE Trans. Med. Imag.*, vol. 26, no. 1, pp. 116–127, Jan. 2007.

[28] C. Sinthanayothin, J. F. Boyce, H. L. Cook, and T. H. Williamson, "Automated localization of the optic disc, fovea, and retinal blood vessels from digital color fundus images," *Br. J. Ophthalmol.*, vol. 83, no. 8, pp. 902–910, Aug. 1999.

[29] M. Lalonde, M. Beaulieu, and L. Gagnon, "Fast and robust optic disc detection using pyramidal decomposition and hausdorff-based template matching," *IEEE Trans. Med. Imag.*, vol. 20, no. 11, pp. 1193–1200, Nov. 2001.

[30] M. Foracchia, E. Grisan, and A. Ruggeri, "Detection of the optic disc in retinal images by means of a geometrical model of vessel structure," *IEEE Trans. Med. Imag.*, vol. 23, no. 10, pp. 1189–1195, Oct. 2004.

[31] S. Sekhar, W. Al-Nuaimy, and A. K. Nandi, "Automated localisation of retinal optic disk using hough transform," in *Proc. 5th IEEE Int. Symp. Biomed. Imag.*, 2008, pp. 1577–1580.

[32] P. Hough, "Method and means for recognizing complex patterns", U.S. Patent 3 069 654, 1962.

[33] R. O. Duda and P. E. Hart, "Use of the hough transformation to detect lines and curves in pictures," *Commun. Assoc. Comput. Mach.*, vol. 15, no. 1, pp. 11–15, Mar. 1972.

[34] K. Fritzsche, A. Can, H. Shen, C. Tsai, J. Turner, H. Tanenbuam, C. Stewart, and B. Roysam, "Automated model based segmentation, tracing and analysis of retinal vasculature from digital fundus images," in *State-of-The-Art Angiography, Applications and Plaque Imaging Using MR, CT, Ultrasound and X-rays*, J. S. Suri, S. Laxminarayan Eds. San Francisco, CA: Academic, 2003, pp. 225–298.

[35] F. ter Haar, "Automatic localization of the optic disc in digital colour images of the human retina," M.S. thesis, Inst. Inf. Comput. Sci., Utrecht Univ., Utrecht, The Netherlands, 2005.

[36] M. N. Kolountzakis and K. N. Kutulakos, "Fast computation of the euclidian distance maps for binary images," *Inf. Process. Lett.*, vol. 43, no. 4, pp. 181–184, Sep. 1992.

[37] P. W. Holland and R. E. Welsch, "Robust regression using iteratively reweighted least-squares," *Commun. Stat.-Theor. Meth.*, vol. A6, pp. 813–827, 1977.

[38] J. V. Miller and C. V. Stewart, *Muse: Robust Surface Fitting Using Unbiased Scale Estimates*. Los Alamitos, CA: IEEE Comput. Soc. Press, 1996, pp. 300–306.

[39] T. Rudolph, M. Puls, C. Anderegg, L. Ebert, M. Broehan, A. Rudin, and J. Kowal, "Marvin: A medical research application framework based on open source software," *Comput. Methods Programs Biomed.*, vol. 91, no. 2, pp. 165–174, Aug. 2008.

[40] A. S. Neubauer and M. W. Ulbig, "Laser treatment in diabetic retinopathy," *Ophthalmologica*, vol. 221, no. 2, pp. 95–102, 2007.

Tobias Rudolph received the M.S. degree in computer science from the University of Magdeburg, Magdeburg, Germany, in 2002, and the Ph.D. degree in biomedical engineering from the University of Bern, Bern, Switzerland, in 2008.

He is currently a Postdoctoral Research Fellow at the artificial organ (ARTORG) Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland, and the Department of Ophthalmology, University Hospital Bern, Bern, Switzerland. His research interests include medical image analysis, image registration and medical data visualization.

Dr. Rudolph is a member of the Association for Research in Vision and Ophthalmology (ARVO).

A. Martina Broehan received the M.S. degree in computer science from Berlin Free University, Berlin, Germany, in 2004. She is currently working toward the Ph.D. degree at the artificial organ (ARTORG) Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland, and the Department of Ophthalmology, University Hospital Bern, Switzerland.

Her research interests include medical image registration and analysis.

Ms. Broehan is member of the Association for Research in Vision and Ophthalmology (ARVO).

Christoph A. Amstutz received the B.Sc. and D.Sc. degrees in engineering science from the Swiss Federal School of Technology (ETH), Zurich, Switzerland, in 1991, and the M.D.-Ph.D. degree from the University of Bern, Bern, Switzerland, in 2001.

Since 2007, he has been an Assistant Medical Director at the Department of Ophthalmology, University Hospital Zurich, Zurich, Switzerland. Currently, he is a Research Fellow at the artificial organ (ARTORG) Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland.

Dr. Amstutz is a member of the Association for Research in Vision and Ophthalmology (ARVO) and the German Association of Ophthalmology (DOG). He is also a Fellow of the European Board of Ophthalmology (FEBO).

Christoph Tappeiner received the M.D. degree from the University of Innsbruck, Innsbruck, Austria, in 2004 and successfully completed the European Board of Ophthalmology examination, in 2009.

He was an Ophthalmologist in Muenster, Germany, Bruck/Mur, Austria, and is currently at the Department of Ophthalmology, University Hospital Bern, Switzerland. His research interests include uveitis, anterior and posterior eye-segment imaging, and image processing.

Dr. Tappeiner is member of the Association for Research in Vision and Ophthalmology (ARVO) and the German Association of Ophthalmology (DOG).

Jens H. Kowal received the M.S. degree in technical computer science from the Technical University Berlin, Berlin, Germany, in 1997, and the Ph.D. degree in biomedical engineering from the University of Bern, Bern, Switzerland, in 2002.

He was a Postdoctoral Research Fellow at Tufts University, Somerville, MA and the Massachusetts Eye and Ear Infirmary, Boston, MA, the M.E. Mueller Institute, Bern, Switzerland, and the University of Western Australia, Perth, Australia. From 2004 to 2008, he was the Head of the group for Smart Surgical Instrumentation at the Institute for Surgical Technology and Biomechanics, University of Bern, Bern, where he has been an Assistant Professor for ophthalmic technologies at the artificial organ (ARTORG) Center for Biomedical Engineering Research, since 2009. He is also at the Department of Ophthalmology, University Hospital Bern, Bern.

Dr. Kowal is member of the Association for Research in Vision and Ophthalmology (ARVO).

Simon P. Rothenbuehler received the M.D. degree from the University of Bern, Bern, Switzerland, in 2009.

He is currently a Research Assistant at the Department of Ophthalmology, University Hospital Bern, Bern, Switzerland. His research interests include retinal degenerations, medical imaging, and image processing.

Dr. Rothenbuehler is a member of the Association for Research in Vision and Ophthalmology (ARVO).