Retinal Image Registration Through Simultaneous Camera Pose and Eye Shape Estimation

Carlos Hernandez-Matas[1,2], Xenophon Zabulis[1] and Antonis A. Argyros[1,2]

Abstract— **In this paper, a retinal image registration method is proposed. The approach utilizes keypoint correspondences and assumes that the human eye has a spherical or ellipsoidal shape. The image registration problem amounts to solving a camera 3D pose estimation problem and, simultaneously, an eye 3D shape estimation problem. The camera pose estimation problem is solved by estimating the relative pose between the views from which the images were acquired. The eye shape estimation problem parameterizes the shape and orientation of an ellipsoidal model for the eye. Experimental evaluation shows 17.91% reduction of registration error and 47.52% reduction of the error standard deviation over state of the art methods.**

I. INTRODUCTION

Assessment of small vessels in vivo can promote the diagnosis and monitor the evolution of diseases that present strong vasculopathy, such as diabetes or hypertension [1]. The eye, and the retina in particular, allows for non-invasive observation of the microvascular circulation via fundoscopy [2].

Image registration can assist greatly in that direction. It aims at warping a test image to the coordinate frame of a reference image, so that corresponding points are imaged at the same locations. For images acquired during the same session, if they present small overlap, it can be utilized for creating mosaics imaging larger areas of the retina [3], [4], [5]. If the overlap is large, the images can be combined to images of higher resolution and definition [6], [7], [8], promoting more accurate measurements. Images acquired at different sessions allow for longitudinal studies of the retina [9], [10], which enable monitoring disease progression.

Besides being a useful clinical tool, retinal image registration is also a challenging problem, as images acquired at different times or from different viewpoints can present illumination, color, and contrast changes as well as potentially small overlapping areas. The support of medical diagnoses requires precise measurements. Therefore, the requirements on registration accuracy are very high.

II. RELATED WORK

Image registration methods utilize the parts of the observed scene that are commonly visible in the image pair to be registered. This information extraction is performed either globally or locally or using a mixture of both. Global methods are based on similarity of intensities, with retinal registration methods usually relying on mutual information [11], [12]. Local methods extract information relying on localized features, such as keypoint correspondences [8], [13], [14], [15], [16], [17], vessel trees [18] and bifurcations [4], [19], [20], [21]. Recently, hybrid methods are gaining traction [22], [23].

The transformation of the images can be estimated on the basis of either 2D or 3D models. 2D methods do not explicitly account for perspective, but overcome this by utilizing non-linear transformations [11], [13], [14], [23]. These transformations do not account for the shape and size of the eye. 3D models enable metric measurements in 3D that lack perspective distortion. Simple eye models have proved to provide accurate registration [16], [17].

In this work, we propose an accurate and robust retinal image registration method that is local and utilizes a 3D transformation model. The main improvement over [16], [17] is the utilization of an ellipsoidal model whose shape parameters are calculated simultaneously with the pose estimate that enables image registration. Other improvements include the utilization of SIFT [24] keypoints instead of SURF [25] and introduction of a pose estimation initialization.

III. METHOD

The proposed method (Figure 1) registers the reference (F_0) and test (F_t) images by simultaneously estimating the relative pose of the cameras that acquired the images, as well as the 3D shape and 3D orientation of an ellipsoidal eye model. The eye model has semi-axes $[a, b, c]$ and rotations along said semi-axes $[r_a, r_b, r_c]$ leading to surface $\mathcal{E}$. If a static camera is assumed, the pose estimate can be calculated as the pose transformation of the retina between the two frames. The eye model is centered at $\mathbf{c}_s = [0, 0, 0]^T$. A calibrated camera for F_0 is located at $\mathbf{c}_c = [0, 0, -\delta]^T$. K_c and K_t are the intrinsic camera matrices for F_0 and F_t. Point correspondences between the images are utilized to achieve this registration. An initial pose estimate is calculated utilizing RANSAC and a spherical model. Subsequently, Particle Swarm Optimization is utilized to refine this pose, as well as to estimate the lengths of the semi-axes of the ellipsoidal model and their rotation. Three variants of the eye model are formulated and experimentally validated.

A. Eye Models

Three models are utilized in this work. Baseline model is spherical, as utilized in our previous works [16], [17].

[1]Institute of Computer Science, Foundation for Research and Technology – Hellas (FORTH), Heraklion, Greece.
[2]Computer Science Department, University of Crete, Heraklion, Greece.
{carlos, zabulis, argyros} at ics.forth.gr

978-1-4577-0220-4/16/$31.00 ©2016 IEEE

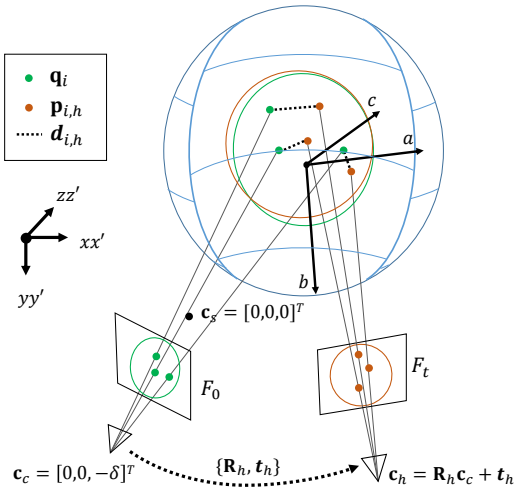

Fig. 1. Geometry of the proposed registration method.

1) Sphere: The sphere consists of an ellipsoid with equal axes, which also allows for ignoring their rotation. Thus $[r_a, r_b, r_c] = [0, 0, 0]$ and $[a, b, c] = [\rho, \rho, \rho]$. Value $\rho = 12\,mm$ is chosen as per the Navarro eye model [26].

2) Fixed orientation ellipsoid: The fixed orientation ellipsoid consists of an ellipsoid with a defined rotation that fixes semi-axes a, b and c to be parallel to the x, y and z axes respectively. Thus $[r_a, r_b, r_c] = [0, 0, 0]$.

3) Ellipsoid: An ellipsoid has three major axes of different lengths. These axes can exhibit any rotation with relation to the camera coordinate system.

B. Solution formulation

A solution $\mathbf{S} = \{\mathbf{R}, \mathbf{t}, \mathbf{A}, \mathbf{Q}\}$ consists of relative camera pose $\{\mathbf{R}, \mathbf{t}\}$ and ellipsoid $\{\mathbf{A}, \mathbf{Q}\}$. The 12 parameters to optimize correspond to the relative rotation $\mathbf{R}$ and translation $\mathbf{t}$ between the cameras, as well as the length of semi-axes $\mathbf{A}$ and the rotation $\mathbf{Q}$ of the ellipsoidal model. $\mathbf{t} = [t_x\, t_y\, t_z]^T$, while $\mathbf{R} = \mathbf{R}_x(r_\theta) \cdot \mathbf{R}_y(r_\phi) \cdot \mathbf{R}_z(r_\omega)$, $\mathbf{A}$ is the diagonal matrix with a^{-2}, b^{-2} and c^{-2} as elements and $\mathbf{Q} = \mathbf{R}_a(r_a) \cdot \mathbf{R}_b(r_b) \cdot \mathbf{R}_c(r_c)$. The search space around the initial pose is denoted as $[\gamma_\theta, \gamma_\phi, \gamma_\omega, \mu_x, \mu_y, \mu_z, \mu_a, \mu_b, \mu_c, \gamma_a, \gamma_b, \gamma_c]$, with $r_\theta \in [-\gamma_\theta/2, \gamma_\theta/2]$ and correspondingly for the rest of dimensions.

C. Feature correspondences

The proposed method utilizes keypoint correspondences between the input pair of images F_0 and F_t to obtain the common information that will be utilized to register them. Our implementation utilizes SIFT [24] keypoint features. They exhibit significant invariance to geometric transformations (translation, scaling and rotation) as well as to illumination changes and have been successfully applied to retinal image registration [14], [27], [17]. SURF [25] features were also considered. They are faster to compute, nevertheless SIFT outperforms SURF in terms of feature localization accuracy [28], [17].

D. Initialization

An initial candidate solution $\mathbf{S}_0$ is estimated by solving the Perspective-n-Point (PnP) problem utilizing Random Sample Consensus (RANSAC) [29]. A 3D pose $\{\mathbf{R}_0, \mathbf{t}_0\}$ of an object is obtained from a set of 2D-3D correspondences and the camera projection matrix P. This pose is calculated through the minimization of the projection error between the 2D and the projected 3D points. As the matched keypoints provide us only with 2D-2D correspondences, and we have no knowledge regarding $\{\mathbf{A}, \mathbf{Q}\}$, we utilize the spherical model and the calibration of the cameras to retrieve the 3D location of the points utilizing Equation 1.

E. Optimization

The estimation of the relative pose of two cameras, as well as the length and rotation of the semi-axes of the ellipsoid, is formulated as the solution of an optimization problem whose goal is the minimization of an objective function. Given a camera pose, the keypoints of the image are traced to 3D locations on the ellipsoidal model surface $\mathcal{E}$. This is performed by calculating the line traversing from the camera optical center to the keypoint in the image. The intersection of line and $\mathcal{E}$ indicates the 3D position of the point.

Locating the coordinate origin at the center of $\mathcal{E}$, we obtain the ellipsoid equation $\mathbf{x}^T\mathbf{Q}^T\mathbf{A}\mathbf{Q}\mathbf{x} = 1$ where $\mathbf{x}$ is a point on $\mathcal{E}$. To estimate $\mathbf{x}$, the line equation from the camera center $\mathbf{c}$ ($\mathbf{c}_c$ or $\mathbf{c}_h$) and through pixel $\mathbf{u}$ is solved for λ:

$$\mathbf{x} = P^+\mathbf{u} + \lambda\mathbf{c}, \tag{1}$$

where $P^+ = P^T(PP^T)^{-1}$ (see Equation 6.13, [30, p. 162]).

Points $\mathbf{q}_i$ are the 3D locations on $\mathcal{E}$ of keypoints from F_0. Points $\mathbf{p}_{i,h}$ are the 3D locations on $\mathcal{E}$ of keypoints from F_t calculated from $\mathbf{S}_h$ (see Figure 1). The 3D distances of corresponding keypoints on $\mathcal{E}$ are $d_{i,h} = |\mathbf{q}_i - \mathbf{p}_{i,h}|$. The minimization of these distances forms the basis of the defined objective function to be minimized. To increase robustness to spurious matches on $o(\cdot)$, a percentile of accumulated distances $d_{i,h}$ is used:

$$o(\mathbf{S}_h) = \sum_j d_{j,h}, \tag{2}$$

where j enumerates the smallest 80% values of $d_{i,h}$.

F. Particle Swarm Optimization

A grid-based search of the 12D space of hypotheses to optimize Equation 2 is computationally prohibitive. Besides, that would create a discretized version of the problem. Thus, the objective function of Equation 2 is optimized via Particle Swarm Optimization (PSO) [31], a stochastic, derivative-free optimization method successfully employed for pose estimation not only in this [16], [17] but also in other domains [32], [33].

PSO performs the optimization via a set of n_p particles that evolve during n_g generations to explore the search space. In our particular case, the 12D search space consists of a hypercube centered around $\mathbf{S}_0$. When several PSO stages are executed in succession, $\mathbf{S}_h$ from the previous execution is utilized as the initial pose.

G. Method variants

Several method variants were considered:

- **PSO Pairing Dimensions and Sphere (PDS):** This variant is the method proposed in [16]. It utilizes the spherical model, no RANSAC initialization, and PSO is executed three times. First execution is the baseline method described in Section III-F. In the second execution the search will be along μ_y and γ_θ and in the third execution along μ_x and γ_ϕ, as in small scale they can produce transformations with similar geometric effect.
- **RANSAC-PSO and Sphere (RPS):** This variant is the method proposed in [17]. It utilizes the RANSAC-based initialization described in Section III-D before performing a PSO search.
- **RANSAC-PSO and Fixed Ellipsoid (RPFE):** Same as RPS but utilizing the fixed rotation ellipsoid model.
- **RANSAC-PSO and Ellipsoid (RPE):** Same as RPS but utilizing the ellipsoid model.

H. Multiple process execution

Both RANSAC and PSO are of stochastic nature, so the proposed method is non-deterministic, which can lead to suboptimal results. Due to this, the process comprised by RANSAC and PSO is executed n_s times and the parameters that gave rise to the best score (objective function minimization) is selected as the final solution. We adopt $n_s = 10$ as suggested in [16], [17]. Execution of multiple swarms leads to an increase on the computational cost, but this solution offers increased accuracy, robustness and reliability.

I. Image formation

When the solution $\mathbf{S}$ is chosen, Equation 1 is utilized to estimate the locations of the pixels from F_t on the surface of the ellipsoid. Then, these locations are projected to the reference camera utilizing the projection matrix P and the image reconstructed by bilinear interpolation.

IV. EXPERIMENTS AND RESULTS

One goal of the conducted experiments was to investigate whether modeling the eye as an ellipsoid, allows for a better approximation of its real shape and leads to increased accuracy in the registration of retinal images. This is achieved by comparing the registration accuracy of the proposed variants that use eye models of different complexity and degrees of freedom. A second goal was to perform a quantitative and comparative evaluation of the proposed approach to state of the art methods in retinal image registration that utilize quadratic image transformation models.

A. Datasets

To support the performed experiments, images were acquired with a Nidek AFC-210 fundus camera. This camera has a resolution of 2912×2912 pixels, and a Field of View (FOV) of 45^o both in the horizontal and vertical dimensions.

A collection of 123 image pairs has been classified into three datasets, regarding parameters of each pair such as the time between image acquisition, the amount of overlap and the presence of large anatomical differences. Each image pair is member of only one dataset. The characteristics of these three dataset are summarized in Table I and example image pairs are in Figure 2.

Dataset	1	2	3
# Image pairs	71	44	8
Examination session	Same	Same	Different
Overlap	$> 75\%$	$< 75\ \%$	$> 75\ \%$
Anatomical changes	No	No	Yes
Indicative application	Super resolution	Mosaicing	Longitudinal study

TABLE I

THE CHARACTERISTICS OF THE EMPLOYED REAL IMAGE DATASETS.

For evaluating the registration accuracy, the reprojection error for the control points is calculated. These control points were obtained by a computational method and verified by a human supervisor. Given that the proposed method utilizes SIFT features, alternative features were chosen to maintain independence from the registration method. SURF features were utilized due to their reliability and accuracy for retinal images [8], [16], [17].

Additionally, the success of a method relative to a competitor method is measured for every individual image pair. A method is considered successful in registering an image pair if the registration error was at least 2.5% lower compared to the competitor. Otherwise, the result is considered a tie.

B. Variant comparison

This experiment compares the four variants proposed in Section III-G so as to identify which performs best in terms of registration accuracy. The computational cost, in n_p per generation, increases with the degrees of freedom.

Table II shows the mean and standard deviation of the error of the control points for the datasets utilizing the four variants of the presented method. It is shown that the registration error decreases as the degrees of freedom increase. This is the case for all three datasets and it is attributed to the better approximation of the retina's shape and pose.

Individual comparison of the losing variants against RPE are shown in Table III (PDS), Table IV (RPS) and Table V (RPFE). In all cases, RPE is shown to outperform the other variants in the majority of image pairs. An analysis of the statistical significance of RPE with relation to the rest of the variants utilizing Student's t-test [34] indicates the results to be statistically highly significant ($P < 0.001$) in all 3 cases.

This experiment proves that independently from the characteristics of an image pair, a model that is able to better approximate the actual shape of the eye assists in performing more accurate registration. Figure 3 shows examples of registration results for each dataset with the RPE variant. It also shows details close to the periphery of the image, for being challenging areas.

C. Comparison with state of the art

This experiment compares the registration accuracy of the proposed variant to GDB-ICP [27], a general registration

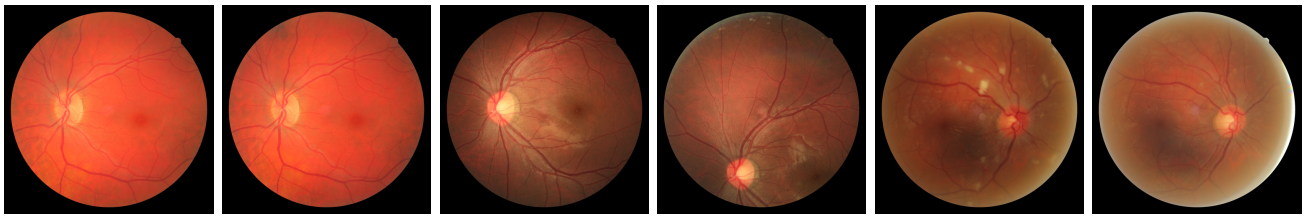

Fig. 2. Fundus images pairs acquired in the same examination (left, middle) and from examinations 1 year apart (right).

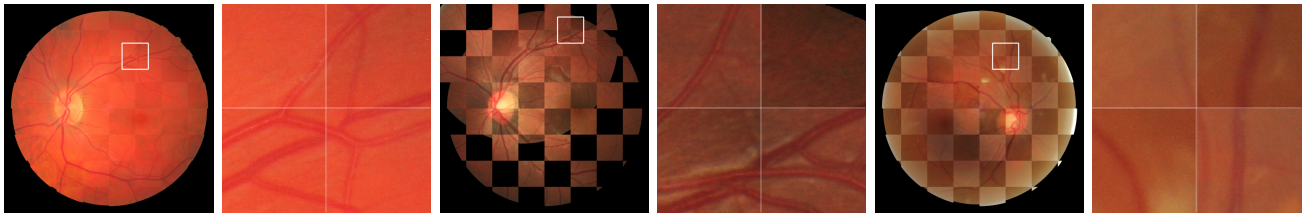

Fig. 3. Registration results for the image pairs of Fig. 2, respectively. The collages show, alternatively, the reference and registered images. The marked region with solid white line indicates the magnified image detail shown on the right.

Dataset	1	2	3	Total
Error PDS	0.590 (0.317)	1.630 (1.391)	1.333 (0.620)	0.709 (0.627)
Error RPS	0.566 (0.284)	1.356 (1.20)	1.257 (0.517)	0.658 (0.532)
Error RPFE	0.557 (0.273)	1.111 (0.787)	1.115 (0.478)	0.623 (0.405)
Error RPE	**0.533 (0.252)**	**0.920 (0.560)**	**1.060 (0.392)**	**0.582 (0.329)**

TABLE II

VARIANT COMPARISON. MEAN (AND STD) REGISTRATION ERROR (IN PIXELS).

Dataset	1	2	3	Total
RPE is better	**60.56%**	**65.90%**	**62.50%**	**62.60%**
PDS is better	19.71%	22.72%	25.00%	21.13%
Tie	19.71%	11.36%	12.50%	16.26%

TABLE III

RPE VS PDS. PERCENTAGE OF IMAGE PAIRS IN WHICH THE PROPOSED METHOD PERFORMED MORE ACCURATELY THAN PDS.

Dataset	1	2	3	Total
RPE is better	**59.15%**	**54.54%**	**62.50%**	**57.72%**
RPS is better	22.53%	34.09%	25.00%	26.82%
Tie	18.30%	11.36%	12.50%	15.44%

TABLE IV

RPE VS RPS. PERCENTAGE OF IMAGE PAIRS IN WHICH THE PROPOSED METHOD PERFORMED MORE ACCURATELY THAN RPS.

Dataset	1	2	3	Total
RPE is better	**50.70%**	**61.36%**	**62.50%**	**55.28%**
RPFE is better	22.53%	25.00%	25.00%	23.57%
Tie	26.76%	13.63%	12.50%	21.13%

TABLE V

RPE VS RPFE. PERCENTAGE OF IMAGE PAIRS IN WHICH THE PROPOSED METHOD PERFORMED MORE ACCURATELY THAN RPFE.

method widely employed with retinal images [13], [19], [35], [36]. GDB-ICP when configured for registering retinal images utilizes a quadratic transformation model, thus, sharing a similar approximation to our method.

The employed GDB-ICP implementation[1] may not yield a result if it is unable to find a satisfactory registration for a given image pair. This is the case in 33.33% of the proposed image pairs, as opposed to the proposed method which always yields a result. Comparing the results of the proposed method between Tables II and VI it is observed that when the error of the proposed method is calculated in the whole dataset, it increases, thus showing that the pairs missed by GDB-ICP present increased registration difficulty.

Table VI summarizes the results for this experiment. The proposed method provides a lower average registration error across all datasets. Additionally, the proposed method is shown to outperform GDB-ICP in the majority of image pairs. The errors displayed in this experiment are lower than for Table II; this is due to the error being computed over less image pairs (i.e., the ones for which GDB-ICP provided a result). This indicates that the proposed method is more robust, as it copes with these challenging cases, with just a marginal increase of error compared to less demanding cases. Statistical significance analysis of RPE with relation to GDB-ICP indicates the results to be statistically significant ($P = 0.0024$).

V. CONCLUSION

A retinal image registration method for fundus images is proposed. The method is based on the simultaneous estimation of (a) the relative 3D pose of the cameras that aquired

[1]We utilized the implementation provided by the authors at http://www.vision.cs.rpi.edu/gdbicp/exec/

Dataset	1	2	3	Total
Error RPE	**0.526 (0.245)**	**0.790 (0.414)**	**1.049 (0.396)**	**0.555 (0.281)**
Error GDB-ICP	0.544 (0.254)	0.887 (0.402)	1.249 (0.461)	0.581 (0.300)
RPE is better	**46.47%**	**18.18%**	**50.00%**	**36.58%**
GDB-ICP is better	22.53%	11.36%	0.00%	17.07%
Tie	16.90%	9.09%	0.00%	13.00%
No GDB-ICP Solution	14.08%	61.36%	50.00%	33.33%

TABLE VI

RPE VS GDB-ICP. TOP: TABLE INTERPRETATION IN TABLE II. BOTTOM: TABLE INTERPRETATION IN TABLE III; SEE TEXT FOR BOTTOM, ADDITIONAL ROW.

the images to be registered and, (b) the parameters and the orientation of an ellipsoidal model of the eye. The performed experimental evaluation shows that the better the model approximates the eye, the more accurate the registration. Additionally, the proposed method is shown to perform better in terms of robustness and accuracy than GDB-ICP, a widely employed method for retinal image registration, which also utilizes a quadratic model for the shape of the retina.

ACKNOWLEDGMENTS

This research was made possible by a Marie Curie grant from the European Commission in the framework of the REVAMMAD ITN (Initial Training Research Network), Project 316990. It was also supported by the FORTH-ICS internal RTD Programme "Ambient Intelligence and Smart Environments". Authors thank Areti Triantafyllou and Panagiota Anyfanti (Hippokration General Hospital of Thessaloniki, Greece) for providing datasets for the experiments, Ruth Ibán-Arias (University of Crete, Greece) for assistance with the statistical significance analysis and the anonymous reviewers for their invaluable feedback.

REFERENCES

[1] A. Grosso, F. Veglio, M. Porta, F. Grignolo, and T. Wong, "Hypertensive retinopathy revisited: some answers, more questions." *British Journal of Ophthalmology*, vol. 89, no. 12, pp. 1646–54, 2005.

[2] M. Abramoff, M. Garvin, and M. Sonka, "Retinal imaging and image analysis." *IEEE Rev. in Biomedical Eng.*, vol. 3, pp. 169–208, 2010.

[3] A. Can *et al.*, "A feature-based technique for joint linear estimation of high-order image-to-mosaic transformations: Mosaicing the curved human retina." *IEEE PAMI*, vol. 24, no. 3, pp. 412–19, 2002.

[4] N. Ryan, C. Heneghan, and P. de Chazal, "Registration of digital retinal images using landmark correspondence by expectation maximization," *Image Vision Comp.*, vol. 22, no. 11, pp. 883–98, 2004.

[5] P. Cattin, H. Bay, L. J. V. Gool, and G. Szkely, "Retina mosaicing using local features." in *MICCAI*, vol. 4191, 2006, pp. 185–92.

[6] N. Meitav and E. Ribak, "Improving retinal image resolution with iterative weighted shift-and-add." *JOSA A*, vol. 28, no. 7, pp. 1395–402, 2011.

[7] G. Molodij, E. Ribak, M. Glanc, and G. Chenegros, "Enhancing retinal images by extracting structural information," *Optics Communications*, vol. 313, pp. 321–8, 2014.

[8] C. Hernandez-Matas and X. Zabulis, "Super resolution for fundoscopy based on 3D image registration," in *IEEE Engineering in Medicine and Biology Society Conf.*, Chicago, USA, 2014, pp. 6332–6338.

[9] H. Narasimha-Iyer *et al.*, "Integrated analysis of vascular and non-vascular changes from color retinal fundus image sequences." *IEEE Trans. on Biomedical Eng.*, vol. 54, no. 8, pp. 1436–45, Aug. 2007.

[10] G. Troglio, M. Alberti, A. Benediksson, and G. Moser, "Unsupervised Change-Detection in Retinal Images by a Multiple-Classifier Approach," in *Multiple Classifier Systems*, 2010, pp. 94–103.

[11] J. P. Pluim, J. Maintz, and M. Viergever, "Mutual-information-based registration of medical images: a survey." *IEEE Trans. on Medical Imaging*, vol. 22, no. 8, pp. 986–1004, 2003.

[12] P. Legg, P. Rosin *et al.*, "Improving accuracy and efficiency of mutual information for multi-modal retinal image registration using adaptive probability density estimation." *CMIG*, aug 2013.

[13] C. Tsai, C. Li *et al.*, "The edge-driven dual-bootstrap iterative closest point algorithm for registration of multimodal fluorescein angiogram sequence." *IEEE TMI*, vol. 29, no. 3, pp. 636–49, 2010.

[14] Y. Lin and G. Medioni, "Retinal image registration from 2D to 3D," *IEEE CVPR*, pp. 1–8, 2008.

[15] Li Tang, M. K. Garvin *et al.*, "Robust Multiscale Stereo Matching from Fundus Images with Radiometric Differences," *IEEE PAMI*, vol. 33, no. 11, pp. 2245–2258, nov 2011.

[16] C. Hernandez-Matas, X. Zabulis, and A. A. Argyros, "Retinal Image Registration Based on Keypoint Correspondences, Spherical Eye Modeling and Camera Pose Estimation," in *IEEE Eng. in Medicine and Biology Society Conf.*, Milan, Italy, 2015, pp. 5650–5654.

[17] C. Hernandez-Matas, X. Zabulis, A. Triantafyllou, P. Anyfanti, and A. A. Argyros, "Retinal Image Registration under the Assumption of a Spherical Eye," *Computerized Medical Imaging and Graphics*, 2016.

[18] G. Matsopoulos, N. Mouravliansky *et al.*, "Automatic retinal image registration scheme using global optimization techniques," *IEEE Trans. on Inf. Tech. in Biomedicine*, vol. 3, no. 1, pp. 47–60, 1999.

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

[20] A. Chaudhry and J. Klein, "Ophthalmologic Image Registration based on shape-context: Application to Fundus Autofluorescence (FAF) images," *Visualization, Imaging, and Image Proc.*, pp. 1–7, 2008.

[21] G. Matsopoulos, P. Asvestas, N. Mouravliansky, and K. Delibasis, "Multimodal Registration of Retinal Images Using Self Organizing Maps," *IEEE TMI*, vol. 23, no. 12, pp. 1557–1563, dec 2004.

[22] P. Reel, L. Dooley, K. Wong, and A. Börner, "Multimodal retinal image registration using a fast principal component analysis hybrid-based similarity measure," *Int. Conf. on Image Processing*, 2013.

[23] K. M. Adal, P. G. van Etten *et al.*, "Accuracy Assessment of Intra- and Intervisit Fundus Image Registration for Diabetic Retinopathy Screening," *IOVS*, vol. 56, no. 3, pp. 1805–1812, 2015.

[24] D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[25] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded-up robust features (SURF)," *CVIU*, vol. 110, no. 3, pp. 346–59, 2008.

[26] R. Navarro, J. Santamaría, and J. Bescós, "Accommodation-dependent model of the human eye with aspherics," *Journal of the Optical Society of America A*, vol. 2, no. 8, pp. 1273–81, 1985.

[27] G. Yang, C. Stewart, M. Sofka, and C. Tsai, "Registration of challenging image pairs: Initialization, estimation, and decision," *IEEE PAMI*, vol. 29, no. 11, pp. 1973–1989, 2007.

[28] E. Oyallon and J. Rabin, "An analysis and implementation of the SURF method, and its comparison to SIFT," *Image Processing On Line*, pp. 1–31, 2013.

[29] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *CACM*, vol. 24, no. 6, pp. 381–95, 1981.

[30] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, 2004.

[31] R. Poli, J. Kennedy, and T. Blackwell, "Particle Swarm Optimization," *Swarm Intelligence*, vol. 1, no. 1, pp. 33–57, 2007.

[32] I. Oikonomidis, N. Kyriazis, and A. Argyros, "Full DOF Tracking of a Hand Interacting with an Object by Modeling Occlusions and Physical Constraints," in *Int. Conf. on Computer Vision*, 2011, pp. 2088–95.

[33] N. Kyriazis and A. Argyros, "Scalable 3D tracking of multiple interacting objects," in *IEEE CVPR*, 2014, pp. 3430–37.

[34] W. S. Gosset, "The probable error of a mean," *Biometrika*, vol. 6, no. 1, pp. 1–25, March 1908.

[35] J. Chen, J. Tian, N. Lee *et al.*, "A partial intensity invariant feature descriptor for multimodal retinal image registration." *IEEE Trans. on Biomed. Eng.*, vol. 57, no. 7, pp. 1707–18, Jul. 2010.

[36] J. Zheng, J. Tian, K. Deng, X. Dai, X. Zhang, and M. Xu, "Salient feature region: a new method for retinal image registration." *IEEE Trans. on Inf. Tech. in Biomedicine*, vol. 15, no. 2, pp. 221–32, 2011.