Landmark Detection for Fusion of Fundus and MRI Toward a Patient-Specific Multimodal Eye Model

Sandro I. De Zanet*, Carlos Ciller, Tobias Rudolph, Philippe Maeder, Francis Munier, Aubin Balmer, Meritxell Bach Cuadra, and Jens H. Kowal

***Abstract*—Ophthalmologists typically acquire different image modalities to diagnose eye pathologies. They comprise, e.g., Fundus photography, optical coherence tomography, computed tomography, and magnetic resonance imaging (MRI). Yet, these images are often complementary and do express the same pathologies in a different way. Some pathologies are only visible in a particular modality. Thus, it is beneficial for the ophthalmologist to have these modalities fused into a single patient-specific model. The goal of this paper is a fusion of Fundus photography with segmented MRI volumes. This adds information to MRI that was not visible before like vessels and the macula. This paper contributions include automatic detection of the optic disc, the fovea, the optic axis, and an automatic segmentation of the vitreous humor of the eye.**

***Index Terms*—Fundus photography, MRI, ophthalmology, patient-specific model, registration.**

I. Introduction

RETINOBLASTOMA is the most frequent eye cancer, which occurs on the retina and affects almost exclusively children. Ninety five percent of these children are under the age of 5 years [24], [22]. In Britain, one in 20 000 children is affected and 40 cases are diagnosed each year [24]. The malignant tumor grows from the retina into the vitreous humor and if left untreated, it gradually fills the bulbus. In the worst case, it can grow along the optic nerve toward the brain. It is hereditary in 40% of the cases and accounts for 5% of childhood blindness [13].

Routinely, ophthalmologists inspect the retina using Fundus photography. This enables straightforward detection of the main vessel branches, the optic disc, and the macula as well as the relative position of the tumor to the critical parts of the eye. Tumors can be delineated with a high resolution but the measurement is restricted to 2-D. Additionally, magnetic resonance images (MRI) are acquired to identify the tumor in the context of the retina, the sclera, the optic nerve, and the lens [16], where the tumor can be delineated in 3-D, but at a lower resolution.

While enucleation is necessary in an advanced tumor, a local treatment is preferred to preserve the remaining vision. External beam radiation therapy is performed as a local treatment method. Prior to the treatment, the procedure has to be thoroughly planned to prevent damaging as much healthy tissue as possible. Since Fundus and MRI are registered manually, large safety margins are added and more healthy tissue is damaged than necessary. Radiation can inflict new tumors in genetically predisposed subjects. As the affected subjects are mostly children, chemotherapy is preferred. Other tissue can be spared by injecting the drug directly into the ophthalmic nerve. In both treatments, it is important to monitor the progression of the tumor over time to assess its effectiveness.

Proper treatment planning is time consuming and error prone due to the high work load for the radiotherapist. The analysis of the MRI is a tedious task in a 3-D volume. By only examining one slice at a time, the global context of the tumor is not addressed properly and small tumors can be missed. Smaller structures like vessels or the macula and especially small tumors cannot be seen in the MRI, which further complicates the task. Additionally, monitoring disease progression becomes difficult with different modalities. Hence, an automatic segmentation of the retinal surface could greatly improve the physician's diagnosis. By projecting the Fundus photography in the correct location on the segmented retina more critical structures can be avoided for radiation treatment. Furthermore, tumors that are only visible in the Fundus can be localized in the MRI.

This paper is situated in the greater scope of the generation of a multimodal, patient-specific eye model for tumor treatment. Existing solutions that are clinically used for tumor treatment planning commonly use spherical or elliptical models for sclera and lens (EYEPLAN [25], OCTOPUS [12]). Based on EYEPLAN, Daftari *et al.* project a Fundus photography to the model [8]. More sophisticated and patient-specific methods are

Manuscript received March 17, 2014; revised June 27, 2014 and August 20, 2014; accepted September 9, 2014. Date of publication September 22, 2014; date of current version January 16, 2015. This work was supported by the Swiss National Science Foundation (205321_144312), the Centre d'Imagerie BioMédicale of the UNIL, UNIGE, HUG, CHUV, EPFL, and the Leenaards and Jeantet Foundations and the Swiss Cancer League under Grant KFS-2937-02-2012. *Asterisk indicates corresponding author.*

*S. I. De Zanet is with the ARTORG Center for Biomedical Engineering Research, the University of Bern, 3010 Bern, Switzerland, and also with the Department of Ophthalmology, University Hospital Bern, 3010 Bern, Switzerland (e-mail: sandro.dezanet@artorg.unibe.ch).

C. Ciller and M. B. Cuadra are with the Centre Hospitalier Universitaire Vaudois and the Department of Radiology, University of Lausanne, 1011 Lausanne, Switzerland, and also with the Centre d'Imagerie Biomedicale (CIBM), Signal Processing Core, Institute of Electrical Engineering, the Ecole Polytechnique Federale de Lausanne, 1015 Lausanne, Switzerland (e-mail: carlos.cillerruiz@unil.ch; meritxell.bachcuadra@unil.ch).

T. Rudolph and J. H. Kowal are with the ARTORG Center for Biomedical Engineering Research, University of Bern, 3010 Bern, Switzerland, and also with the Department of Ophthalmology, University Hospital Bern, 3010 Bern, Switzerland (e-mail: tobias.rudolph@artorg.unibe.ch; jens.kowal@artorg.unibe.ch).

F. Munier and A. Balmer are with the Jules Gonin Eye Hospital, Unit of Pediatric Ocular Oncology, 1004 Lausanne, Switzerland (e-mail: francis.munier@fa2.ch; aubin.balmer@fa2.ch).

P. Maeder is with the CHUV/MIAL, Department of Radiology, University of Lausanne, 1015 Lausanne, Switzerland (e-mail: philippe.maeder@chuv.ch).

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/TBME.2014.2359676

available, however. Eye segmentation has been performed in computed tomography (CT) images by Rüegsegger *et al.* based on a statistical shape model and subsequent fitting with an active shape model [29]. The sclera and the lens were automatically segmented. Cuadra *et al.* segmented CT images by using an active contour framework based on a parametrical model [7]. Additionally, an ultrasound image was fused into the gained model. Similarly, Bondiau *et al.* [2] detect the eye and lens based on a spherical Hough transform, and the optic nerve using a cylindrical parameterization. Fundus photography was, then, orthographically projected onto the hemisphere representing the retina without considering the geometrical distortion. For brain tumor treatment with radiotherapy structures at risk were automatically segmented using an atlas-based approach on CT images by D'Haese *et al.* [10]. Their approach is able to approximately segment the eye and the optic nerve.

In contrast, this study presents a fully automatic segmentation and fusion of two commonly used diagnostic image modalities in Retinoblastoma: Fundus photography and MRI volumes [26]. In the eye, landmarks such as the optic disc, the optic axis, and the macula in both the MRI volume and the Fundus photography are extracted automatically. These landmarks are used to fuse the two modalities into a single comprehensive model. Instead of using an orthographic projection, the surface of the retina is taken into consideration, especially in case of a tumor. For external beam radiation therapy, a smaller safety margin can be envisioned using the mapped fundus to spare more healthy tissue.

This paper consists of several detection steps for landmark registration which comprise eye detection, bulbus segmentation, optic disc and optic axis detection, and fovea position estimation.

II. Methods

The two modalities, Fundus photography and MRI, have few common landmarks. On the one hand, it is possible to discern the optic disc, the retina surface, the sclera, the lens, and the cornea in the MRI. On the other hand, the vessels, the optic disc, and the macula/fovea are clearly visible in Fundus photography. Hence, one obvious landmark is the optic disc. The optic disc is the circular, approximately 2 mm in diameter, entrance point of the optic nerve into the retina. It can be seen as a bright disc in the Fundus and deduced from the MRI data by intersecting the nerve with the retinal surface. A second landmark is needed for the scaling and rotation of the 2-D registration. Bondiau *et al.* [2] assume the fovea to be on the incidence of optical axis of the lens and the retina. However, it is known from the literature that the axis from the lens to the fovea (the visual axis) and the optical axis show a deviation of 4° [30]. In this paper, to get a better estimation of the position of the fovea, a statistical approximation based on the distance from the optic disc and the mentioned axial angle difference is used.

Finally, eye detection, surface extraction, and landmark detection are performed automatically. The individual steps are explained in more detail in the next sections and an overview is illustrated in Fig. 1.

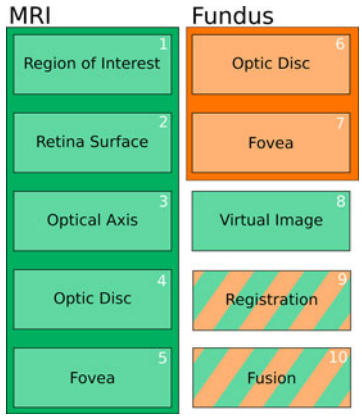


Fig. 1. Overview of all the steps for the automatic registration of Fundus photography (orange) and MRI (green). Step 1: The region of the eye is cropped around the automatically detected center. Step 2: The retinal surface is segmented. Step 3: The optical axis is automatically detected. Step 4: The optic disc position is automatically detected. Step 5: The Fovea position is estimated. Step 6 and 7: The optic disc and the fovea are automatically detected in the Fundus. Step 8: A virtual Fundus image is generated from the MRI. Step 9: Virtual and Fundus image are registered. Step 10: Fusion of the two modalities is achieved by back projecting the registered image to the retinal surface.

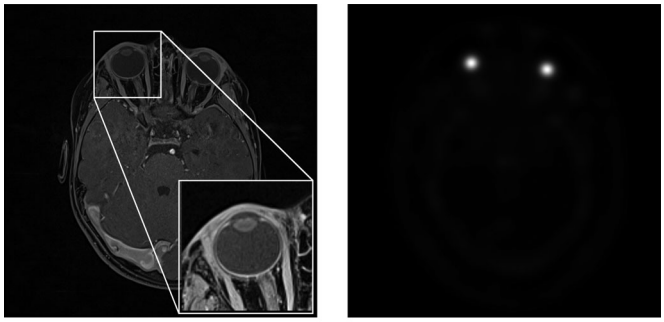
A. Data Acquisition

MRI volumes have been acquired from infants with unilateral and bilateral tumors by trained radiologists. The full dataset comprises 13 patients. The infants are aged 3.2 ± 1.7 years and had eyes in various developmental stages, mostly differing in size. All the patient data were acquired using a gadolinium enhanced T1-weighted GE VIBE (TR/TE, 20/3.91 ms) sequence [32] according to de Graaf *et al.* [16]. The patients were anesthetized beforehand, reducing motion artifacts. Acquisitions of low quality or motion artifacts were rejected and then repeated by the radiologists, as the physicians need a clear understanding of the tumor size and position. This study is based on the images the physicians commonly use. The used MRI volumes have been recorded at slightly differing resolutions with nonisotropic spacing ($0.416 \times 0.416 \times 0.390$ mm and $0.480 \times 0.480 \times 0.499$ mm). Fundus photographs have been acquired using a RetCam, a camera with interchangeable frontal lens that is applied directly on the cornea of the anesthetized patients. All images were anonymized prior to its analysis. This study was approved by the Cantonal Research Ethics Committee (Vaud).

B. Detection of Eye Region of Interest in the MRI Volume

Typically, the acquired MRI volumes encompass the whole head and are only cropped below the nose and above the forehead. An example can be seen in Fig. 2(a). Therefore, to reduce computational cost and increase robustness, both eye centers have to be detected automatically and the MRI data cropped accordingly. The most prominent feature of eyes compared to other anatomical structures is their sphericity with a diameter of 25 mm in adults, and smaller for children [17].

[1]Retcam 2, "Clarity Medical Systems," Pleasanton, CA, USA.



(a) (b)

Fig. 2. In (a), a slice of the original can be seen. The same slice transformed with FRST3D is shown in (b). The two eyes generate a distinct response seen as two high intensity points. Their positions can be extracted directly by finding local maxima. (a) Full volume. (b) FRST3D.

Fast radial symmetry transform (FRST) by Loy and Zelinski [23] is an algorithm which is used in computer vision to detect radial symmetry. Its main advantages over other methods like the Hough transform [19] is its low computational cost and ease of implementation [23]. Additionally, it does not rely on an edge detector as it is only dependent on the local gradients in the original image.

To compute the transform, the local image gradient is used to vote for the symmetry of a pixel in a predefined range of radii r. For each pixel p, the pixel at the position

$$p_{\pm ve}(p) = p \pm \text{round}\left(\frac{g(p)}{\|g(p)\|} r\right) \tag{1}$$

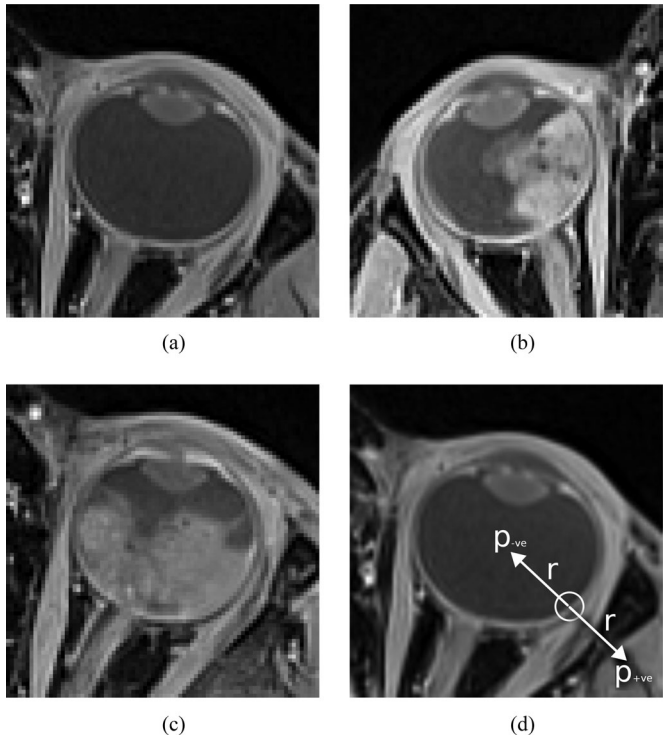
is accumulated by the gradient value in the so-called orientation projection image O and ± 1 in the so-called magnitude projection image M [see (2)]. An example of the two vectors $p_{\pm ve}$ is shown in Fig. 3(d)

$$\Delta O(p_{\pm ve}) = \pm 1, \quad \Delta M(p_{\pm ve}) = \pm \|g(p)\|. \tag{2}$$

The position in the direction of the gradient is accumulated with a positive sign, while in the opposite direction it is accumulated with a negative sign. The final transformation is $S = F * A$, where

$$F(p) = \|O'(p)\|^{\alpha} \cdot M'(p) \tag{3}$$

(O', M' being the normalized O, M) and A a 2-D Gaussian kernel ($A(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$) with a standard deviation σ linearly dependent on the radius. α determines the radial strictness to enforce or relax the circularity (sphericity) of structures. If a structure is radially symmetrical, the center of the symmetry will accumulate the most votes in S; hence, a local maximum search will yield centers of symmetry. The original algorithm was conceived for the 2-D space. For our purposes, we adapted it to the 3-D space which is a straightforward process. Once the volume is transformed, the values with the highest intensity correspond to the points with highest symmetry and, hence, the centers of the eye. An example of the transformation can be seen in Fig. 2.



(a) (b) (c) (d)

Fig. 3. Resulting regions of interest after detection are robustly found in healthy and pathological cases. The slices are centered on the detected eye center. The gradient at the point p points to the outside at the fixed radius r. Only the vector p_{-ve} was used for eye center detection. (a) Healthy eye. (b) Eye with Retinoblastoma. (c) Eye with Retinoblastoma filling the whole bulbus. (d) Schematic of FRST3D computation. The gradient at the point p points to the outside at the fixed radius r. Only the vector p_{-ve} was used for eye center detection.

As can be seen in Fig. 3(d), the gradients point *out* of the eye at the border retina/vitreous humor. Therefore, for a robust detection, only the vectors pointing in the opposite direction of the gradient (p_{-ve}) were used. This corresponds to the so-called *dark* portion of the FRST algorithm [23]. With this method, the two eyes are detected by finding the two local maxima in the filtered image $\arg\max \text{FRST}(V)$. Finally, the volume is cropped around the eye center at a predefined width for further processing. The cropping width is set to 40 mm (an average adult eye size has a length of 24 mm) to include eye details and a part of the optic nerve. No preliminary filtering is needed to find the eyes robustly.

C. Segmentation of Retinal Surface

Once the eye centers are detected and the MRI data is cropped, the retinal surface has to be segmented to provide a surface for Fundus projection. Manual segmentation, especially in a volumetric setting, is time consuming and error prone. Hence, a fully automatic segmentation approach is chosen.

As a preprocessing step, anisotropic filtering is applied to the image to remove noise. Since the volume is cropped to the region of the eye, the computation is not excessively high. The image is, thus, greatly improved as shown in Fig. 5(a). The noise

in the vitreous humor is completely removed, while preserving the edge between the retinal surface and the vitreous humor.

The first segmentation step consists of finding a set of points that lie on the retinal surface. The previously computed eye detection step finds the (symmetry) center of the eye. By casting rays from this center in all directions, radial profiles are generated. The first high gradient in the profile is most likely a point on the vitreous humor border. These points are collected into a cloud of points. To ensure an even distribution of the points across the sphere, the method of Vogel is used [31]. This algorithm generates an approximate uniform distribution of points on a sphere using spirals.

A number of outliers remain in the point cloud, mostly occurring around the zonules and the ciliary body, situated at the edges of the lens. These outliers do not lie in the vicinity of inliers and are sparsely distributed. DBSCAN [14] clusters point clouds based on their vicinity. Thus, by selecting the largest cluster the outliers can be removed, given the fact that most points are inliers. The outlier removal, however, leads to holes in the surface. If a mesh were to be generated from these points, the holes and unfiltered outliers lead to an erroneous mesh.

To address this problem, an approach is used that ensures a global solution with a smoothness constraint. The final mesh is generated by using a graph cut classification, more specifically quadratic pseudo-Boolean optimization [3], [27]. The input is a six-connected volume graph with every voxel as a graph node. The graph cut algorithm then finds the optimal cut through the volume based on costs on the edges. In this case, the computed point cloud is transformed to a distance map, with each voxel indicating the distance to the nearest point. The distance map can be directly used as a cost for the graph. Additionally, it is important to constrain the labels for outside (labeled as 0) and inside (labeled as 1).

The energy function is defined as

$$E(x) = \theta_{\text{const}} + \sum_{p \in \mathcal{V}} \theta_p(x_p) + \sum_{(p,q) \in \mathcal{E}} \theta_{pq}(x_p, x_q) \quad (4)$$

where the costs are

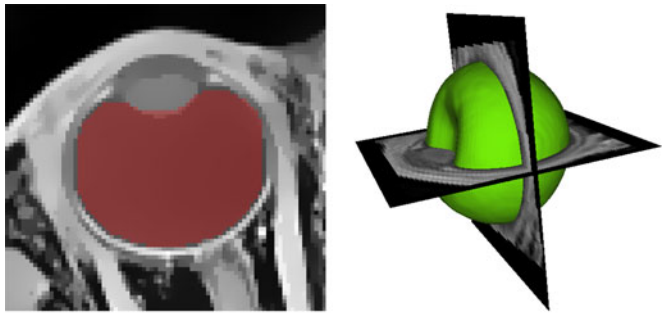
$$\theta_{\text{const}} = 0 \quad (5)$$

$$\theta_p(0) = M - \|c - p\|, \quad \theta_p(1) = \|c - p\| \quad (6)$$

$$\theta_{pq}(0,0) = \theta_{pq}(1,1) = 0 \quad (7)$$

$$\theta_{pq}(0,1) = d(p), \quad \theta_{pq}(1,0) = d(q). \quad (8)$$

The definition takes the form of $\theta_p(\text{label for } p)$ for vertices and $\theta_{pq}(\text{label for } p, \text{label for } q)$ for edges. The first equation in 6 makes sure that labels far from the center c are more likely to be labeled *outside*. Conversely, the second equation labels points close to the center as *inside*. To ensure a connected shape, the cost for same-label edges are kept to zero (7). Finally, the most important constraint is given by (8). The cost for a label transition from *outside* to *inside* is lowest where the value in distance map d is lowest. By minimizing the energy function in (4), each voxel is labeled as either *outside* or *inside*. In Fig. 4(a), an example of this labeling is shown in a volume slice. The red region depicts the *inside* labeled voxels.



(a) (b)

Fig. 4. (a) MRI slice with the segmentation labels superimposed. The red pixels designate the *inside* label. (b) Transition between *inside* and *outside* labels is used to define the surface of the vitreous humor (green). Two corresponding MRI slices are intersected with the surface. (a) Graph cut segmentation. (b) Segmentation.

Finally, the desired surface is located on the edge between the *inside* and *outside* label. Using this information, the mesh surface is extracted as seen in Fig. 4(b). We perform a general smoothing of the surface to reduce staircase artifacts.

D. Detection of Optical Axis

The optical axis is defined as the axis going through the lens center and the cornea curvature center. A first point on the axis is detected in the lens. The lens is an ellipsoidal, biconvex shape and, thus, radially symmetric to some extent. By providing the FRST3D algorithm with a set of radii from 2 (half axial length) to 5 mm (half diameter), the symmetry point of the lens is highlighted. The cornea contributes to this result by being circular itself and of the same size, so the maximum will not be exactly in the lens center. However, it is sufficient, since the detected center still lies on the optical axis.

The optical *axis* is initialized by connecting the lens center and the eye center. This axis is a good estimation but can be inaccurate in case of a slight error in the eye center detection. The optical axis has to be the symmetry axis for the lens. Thus, the initialized optic axis is optimized to yield a maximal symmetry in the vicinity of the lens. The energy function to be minimized is

$$E(n) = \sum_{\alpha, \lambda_1, \lambda_2} V[p + \lambda_1 n + \lambda_2 m_\alpha] - V[p + \lambda_1 n - \lambda_2 m_\alpha] \quad (9)$$

where n and p denote the normalized axis direction and the lens position, respectively. m is a vector perpendicular to n at an angle α. V denotes the MRI volume. The energy is optimized over n to yield the direction with the highest symmetry. This direction is the most refined orientation for the optic axis. In Fig. 5(b), an unoptimized (red) and optimized (green) axis can be seen with sampling points in one slice.

E. Detection of the Optic Nerve Head

The optic nerve head in the MRI is found at the entrance of the optic nerve into the retina. A small indentation of 100–300 μm exists where the vessels enter the bulbus through the

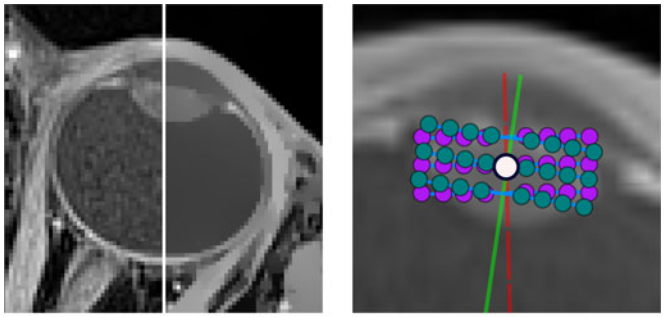


Fig. 5. Example of anisotropic filtering in (a) removes the prevalent noise and facilitates further segmentation by enhancing the gradients. Optimization of the optical axis in (b). (a) Anisotropic filtering. (b) Example of an optimized axis (green) with green sample points and its initialization in red.

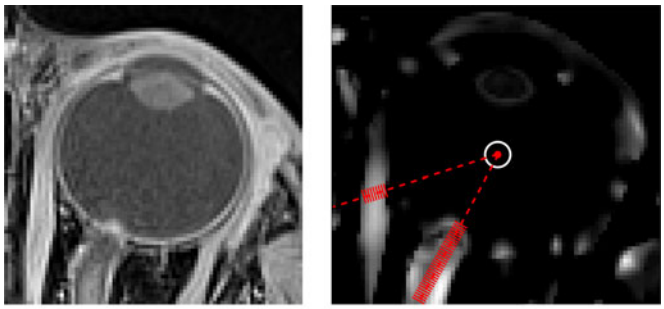


Fig. 6. Original and Frangi-filtered image. In the filtered image (b), the optic nerve is visible but also the muscles are highlighted. In red the rays cast out from center are depicted. They visualize the difference in accumulation between optic nerve and muscle. (a) Original. (b) Filtered image.

optic nerve. However, for the available resolution of around 400 μm, this indentation cannot be used as a landmark [see Fig. 6(a)]. The optic disc constitutes the intersection between the optic nerve and choroid/retina. This landmark is more easily discernible because the optic nerve is identifiable in the MRI volume as a bright tubular structure. Therefore, in this paper, a vesselness filter [15] is used, parameterized to the diameter of the optic nerve. It is based on the eigenvalues of the hessian, which represent the curvatures along the axes. If one of the three curvatures is near zero and two others are nonzero, then the point lies on a tubular structure.

By applying the filter, all tubular structures are highlighted. An example can be seen in Fig. 6(b). Not only the optic nerve is visible but also the rectus muscles and to some extent the eye lid. However, the muscles run along the surface of the sclera, in contrast to the optic nerve which enters it radially at an approximately right angle. This characteristic is used to discern the optic nerve from the muscles. Rays are cast from the already known eye center outwards [31]. For each point on the ray, the value of the filtered image is accumulated. The highest probability for the direction of the optic nerve is, therefore, the one with the highest accumulated value. Muscles only contribute on their diameter, while the optic nerve contributes all along the axis.

F. Estimation of the Fovea Position

The macula position in the MRI can be estimated by using previously detected landmarks. The visual axis, the axis through the lens and the fovea, deviates from the optical axis by angle κ of about $3.93 \pm 2.68^\circ$ in the left eye and $3.91 \pm 2.73^\circ$ in the right eye [30]. This constrains the position of the fovea to the cone of angle κ from the lens intersected with the retina. A second constraint is the distance from the optic disc. Previous studies [9] have shown that the distance between fovea and optic disc is 4.4 ± 0.4 mm in infant eyes. The two constraints define a joint probability

$$P(x) = f_{\mathrm{OD}}(\|x - x_{\mathrm{OD}}\|) \cdot f_\kappa(\phi(n_o, c_{\mathrm{eye}} - c_{\mathrm{lens}})) \quad (10)$$

where $f_{\mathrm{OD}} \sim N(\mu_{\mathrm{OD}}, \sigma_{\mathrm{OD}})$, $f_\kappa \sim N(\mu_\kappa, \sigma_\kappa)$, and $\phi(x_1, x_2)$ the angle between the two vectors x_1 and x_2. c_{eye} and c_{lens} denote the centers of the eye and the lens, respectively. As they roughly delineate two intersecting circles, two possible fovea positions are probable. The position closer to the optic axis incidence with the retina (the principal point) is chosen to be the fovea.

G. Detection of Optic Disc and Macula in the Fundus

In the previous sections, the automatic landmark detection for the MRI was described. The same two landmarks, the optic disc and the fovea, must be found in the Fundus photography in order to fuse the two modalities. A wide range of methods have been used to detect and segment the optic disc in Fundus photography such as template matching, active contours, level sets or intensity thresholding [1], [20], [33] or methods based on vessel density/direction [18], [34]. Our detection is based on the work of Budai *et al.* [5], [6]. They capitalize on the radial symmetry of the optic disc, using the FRST in its 2-D variant [23]. As the gradients at the rim of the optic disc point to the center, the positive portion p_{+ve} is used (*bright*). By selecting the maximum in the filtered image, the optic disc is found. However, the strong border gradients of the image causes false positives; therefore, the border is masked out. The mask is generated by a threshold on the intensity.

A similar approach is used (based on Budai [5]) to find the macula. The macula is a dark spot at approximately one optic disc diameter away from the optic disc in a temporal direction. In contrast to the optic disc, it follows a slow negative gradient. Hence, the negative portion p_{-ve} is used at a continuous range of radii (*dark*). The filter performs better with the same mask used in the optic nerve case. Additionally, the vessels are masked out [4], as their stronger gradients add noise and cause false positives to the filtered image. By selection of the minimum, the macula is detected. To increase robustness, the search is confined to the approximate distance from the optic disc. Examples of the filtered images can be seen in Fig. 7.

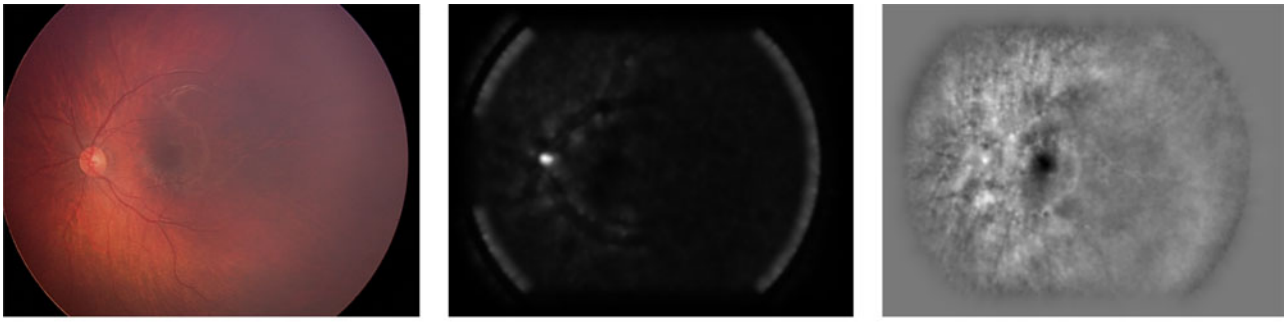


Fig. 7. Landmark detection in Fundus photography. Optic disc and macula are both automatically detected using the FRST algorithm and removing vessels and border gradients. (a) Original Fundus photography. (b) Image filtered with bright FRST2D filter to detect the optic disc. (c) Image filtered with dark FRST2D filter to detect the macula.

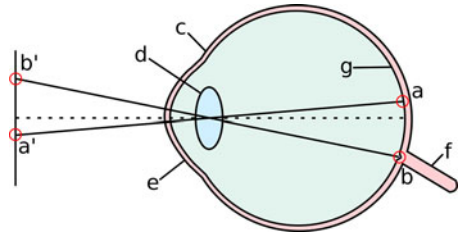


Fig. 8. Schematic of the mapping of the fovea (a) and the optic disc (b) to a virtual image plane in front of the eye. The two points (a') and (b') are the images of the landmarks. Additional anatomical structures are indicated (c) sclera, (d) lens, (e) cornea, (f) optic nerve, and (g) retina.

TABLE I
RESULTS FOR SEGMENTATION AND LANDMARK DETECTION

	DICE	Jaccard	OD [mm]	EC-H [mm]	EC-T [mm]
Mean	0.967	0.936	0.963	0.715	0.833
StDev	0.006	0.010	0.358	0.350	0.380

From left to right, the columns represent the DICE coefficient and the Jaccard index of the retinal surface, the optic disc detection error (OD), the eye center detection error in healthy eyes (EC-H), and pathological eyes (EC-T) for 13 patients.

H. Registration and Mapping

In this final step, the found landmarks and the segmentation are used to fuse MRI and Fundus photography into one 3-D model. The fusion is performed in three steps. First, a virtual image from the landmarks in the MRI is generated. Subsequently, the landmarks in the Fundus are registered to the virtual image. In the last step, the Fundus photography is projected onto the previously segmented retinal surface.

To generate a virtual image, the fovea and optic disc are back-projected through the lens onto a plane in front of the eye (see Fig. 8). The plane is oriented in a way that the plane normally coincides with the optical axis. The virtual image is transformed to the same 2-D plane as the Fundus photography. The two images are then registered as these two landmarks are in both the virtual and the Fundus photography. A similarity transform is chosen for the registration.

Finally, the registered Fundus is projected back to the plane in the front of the eye and from there through the lens onto the segmented retinal surface. Unlike Bondiau *et al.* [2], which used an orthographic projection, this method takes into consideration the geometrical distortion induced by the nonplanar surface of the retina.

III. RESULTS

The different steps in the algorithm were evaluated separately, since there was no ground truth to confirm the correctness of the Fundus photography mapping onto the retinal surface. However, all the intermediate steps can be evaluated based on the comparison of the results with the manual segmentation. In the following sections, these steps are evaluated.

A. Eye Detection in MRI

The automatic detection of the centers of the eyes in the MRI was evaluated on 13 eyes. Eye centers were segmented automatically and compared to a manually segmented ground truth. For each eye, the Euclidean distance was computed. To compare healthy and tumor-affected eyes, the two were evaluated separately. Three observers segmented the eye centers manually. In Table I, the error for each subject to the average observer can be seen for healthy (EC-H) and pathological (EC-T) eyes. The average error of 0.715 ± 0.350 mm in healthy eyes is comparable to the average error in pathological eyes of 0.832 ± 0.379 mm. A two-sided student's unpaired test revealed no statistically significant difference between detection of healthy and pathological eyes (with a p-value of 0.21).

B. Segmentation of Retinal Surface

For the evaluation of the automatic vitreous humor segmentation, 13 volumes were both manually and automatically segmented. To assess the accuracy, the commonly used DICE similarity coefficient [11] was measured between the manual and the automatic segmentation computed with the graph cut approach. Similarly, the Jaccard index [28] was computed. The

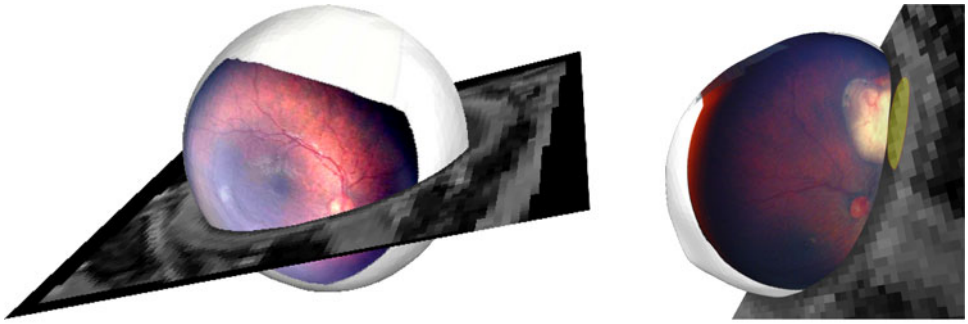


Fig. 9. Fusion of Fundus and retinal surface from MRI from different perspectives, without (a) and with (b) a tumor. (a) Exterior view with the optic nerve entering in the place of the optic disc of the Fundus. (b) Mapping with an incidence of a Retinoblastoma (yellow).

DICE coefficient and Jaccard index are defined, respectively as

$$\text{DICE} = \frac{2|A \cap B|}{|A| + |B|} \tag{11}$$

$$\text{Jaccard} = \frac{|A \cap B|}{|A \cup B|}. \tag{12}$$

Overall, the DICE coefficient resulted in 0.967 ± 0.006. Deviations from the manual segmentation were mostly found in the anterior part of the eye. The segmentation takes less than 20 s on an average laptop[2].

C. Detection of Optic Nerve Head in MRI

The optic nerve head was found in all healthy eyes. The automatically found position by the algorithm A was compared to a manual segmentation of two clinically trained observers O_1 and O_2. The error between the algorithm and the observers was measured as the euclidean distance d between the two points: $d(O_1, A) = 0.977 \pm 0.318$ mm, $d(O_2, A) = 1.062 \pm 0.343$ mm. The interobserver error was $d(O_1, O_2) = 0.525 \pm 0.350$ mm. By averaging the coordinates from O_1 and O_2, the average error results in $d(O, A) = 0.963 \pm 0.358$ mm.

D. Detection of Optic Disc and Macula in the Fundus

A similar approach to Budai *et al.* [5] was used. In this paper, the same modality is used and the algorithms do not differ. Budai *et al.* report an optic disc diameter distance error of 0.2 optic disc diameters. This results into a typical distance of 0.4 mm which is accurate enough for the given MRI volume resolution of 0.4–0.5 mm.

E. Results of Mapping

In Fig. 9(a), an example of a Fundus photography mapped on the posterior part of the eye is shown. Vessels, which were not previously detectable, are now mapped on the surface and can be seen in the context of the full MRI. No ground truth is available for the verification of this fusion. However, visual inspection by clinicians showed a plausible result. The optic disc of the Fundus corresponds to the position where the optic nerve enters the eye. Vessels arch in the temporal direction, which is to be expected. The fovea lies next to the optical axis incidence with the retina, which is another indication that the result is plausible. In Fig. 9(b), a Fundus is mapped that depicts a Retinoblastoma. The tumor in the Fundus and the MRI coincide even though only the fovea and optic disc were used as landmarks.

IV. Discussion

The evaluation of the various detection and segmentation steps add a valuable contribution to automatic landmark extraction from MRI images in human eyes. A segmentation with a DICE coefficient of 97% is comparable to segmentations of the sclera in CT images by Rüegsegger *et al.* [29] and Isambert *et al.* [21] of 95% and 93%, respectively. In contrast to statistical shape models, it has the advantage that it is adaptable to pathological cases which do not follow the statistical model. The parallax distortion induced by the shape of the retina is compensated by the projection of the Fundus through the lens. Especially, in case of deformations from tumors, this reduces errors from mapping compared to an orthogonal projection.

The eye centers were detected with an error of less than 1 mm considering a resolution of 0.5 mm. All the steps for MRI landmark detection and segmentation rely on a good eye center estimation. However, the center does not have to be as exact as other landmarks. Cropping the MRI around the eyes only requires a rough estimation: large enough margins compensate for a slight error. As can be seen in the example images, the eye is not perfectly spherical. The contributions of each voxel gradient is spread with a convolution of a Gaussian. Thus, the contributions overlap in the center of the eye. To find the optic axis, the axis is initialized with a point in the lens and the center of the eye. Since the axis is optimized using local symmetry, a rough estimation is adequate. For these reasons, the achieved accuracy is sufficient for the proposed method.

Optic disc detection showed an accuracy of 0.963 ± 0.358 mm. For the available resolution of at least 0.5 mm, which corresponds to a distance of two voxels, the interobserver variability, however, is 0.5 mm smaller. The reasons for this discrepancy are mainly due to the fact that optic nerve does not enter the retina fully, radially in some cases. Therefore, the outward ray cast will not travel along the optic nerve's axis but rather diagonally through it. A method to increase accuracy might be to optimize the position locally using radial symmetry. Additionally, more

constraints could be applied for a faster computation. Based on the position of the lens, the search can be restricted to the posterior part of the eye.

The fusion of the two modalities offers a comprehensive view of small structures like vessels on the global eye context, which was the goal of this paper. Since there was no ground truth to compare the fusion with, the result relies on the accuracy of the previous steps. The next step toward a ground truth is using an eye with a mild pathology that does not affect the optic disc nor the fovea. The fusion accuracy can then be measured by comparing the tumor overlap in the virtual image. Furthermore, the same technique can be used to register pathological cases where landmarks are missing. Then the tumor outline in the Fundus and the projection of the tumor in the virtual image can be used for registration. It has to be taken into consideration that due to the limited number of only two landmarks, the accuracy far from the optic disc and fovea suffers. This is due, on the one hand, to the distortion introduced by the Fundus photography and, on the other hand, by the estimation of the fovea. However, geometrical distortions caused by the topology of the retina are minimized by using the proposed method.

V. Conclusion

In this study, a fully automatic fusion of Fundus photography and MRI volumes in infant eyes was presented. In addition to a complete fusion, preliminary steps for landmark detection and extraction were shown. They comprise eye detection from a full-volume surface extraction of the retina and detection of the optic disc. These landmarks are suitable for a range of other applications like initialization of statistical shape models or active contours and registration of further modalities. The primary application is found for tumor treatment planning and monitoring.

The fusion of Fundus photography and MRI enables other modalities to be fused with MRIs. Angiograms can be registered to the Fundus using the vessels as landmarks. OCT is already registered to an en-face view of the retina by commercially available devices. This will add retinal layer information in the μm range to a high-resolution MRI in the mm range, paving the way for a multimodal multiscale eye model. Furthermore, some tumors can be detected in OCT earlier than in Fundus. By registering the OCT to the MRI, the early stages of tumors can be mapped to the 3-D position in the eye.

References

[1] M. D. Abramoff and M. Niemeijer, "The automatic detection of the optic disc location in retinal images using optic disc location regression," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, 2006, pp. 4432–4435.

[2] P. Bondiau and G. Malandain,. (1997). Eye reconstruction and CT-retinography fusion for proton treatment planning of ocular diseases. *CVRMed-MRCAS* [Online]. Available: http://link.springer.com/chapter/10.1007/BFb0029296

[3] E. Boros and P. L. Hammer, "Pseudo-Boolean optimization," *Discrete Appl. Math.*., vol. 123, nos. 1–3, pp. 155–225, Nov. 2002.

[4] M. Broehan, T. Rudolph, C. Amstutz, and J. H. Kowal, "Real-time multimodal retinal image registration for a computer-assisted laser photocoagulation system. *IEEE Trans. Bio-Med. Eng.* vol. 58, no. 10, pp. 2816–2824, Oct. 2011.

[5] A. Budai, A. Aichert, B. Vymazal, J. Hornegger, and G. Michelson, "Optic disk localization using fast radial symmetry transform," in *Proc. Comput. Based Med. Syst.*, 2013, pp. 59–64.

[6] A. Budai, L. Laurik, J. Hornegger, G. M. Somfai, and G. Michelson, "Probability map based localization of optic disk," in *Proc. Int. Conf. Syst. Signal, Image Process.*, 2012, pp. 568–571.

[7] M. Cuadra, S. Gorthi, F. I. Karahanoglu, B. Paquier, A. Pica, H. Do, A. Balmer, F. Munier, and J.-P. Thiran, "Model-based segmentation and fusion of 3d computed tomography and 3D ultrasound of the eye for radiotherapy planning," *Comput. Vis. Med. Image Process. Comput. Methods Appl. Sci.*, vol. 19, no. 12, pp. 247–263, 2011.

[8] I. K. Daftari, K. K. Mishra, J. M. OBrien, T. Tsai, S. S. Park, M. Sheen, and T. L. Phillips, "Fundus image fusion in EYEPLAN software: An evaluation of a novel technique for ocular melanoma radiation treatment planning," *Med. Phys.*, vol. 37, no. 10, pp. 5199–5207, 2010.

[9] D. J. De Silva, K. D. Cocker, G. Lau, S. T. Clay, A. R. Fielder, and M. J. Moseley, "Optic disk size and optic disk-to-fovea distance in preterm and full-term infants," *Invest. Ophthalmol. Visual Sci.*, vol. 47, no. 11, pp. 4683–4686, Nov. 2006.

[10] P.-F. D. D'Haese, V. Duay, R. Li, A. du Bois d'Aische, T. E. Merchant, A. J. Cmelak, E. F. Donnelly, K. J. Niermann, B. M. M. Macq, and B. M. Dawant, "Automatic segmentation of brain structures for radiation therapy planning," in *Proc. SPIE Med. Imag.: Image Process.*, May 2003, pp. 517–526.

[11] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecol. Soc. Amer.*, vol. 26, no. 3, pp. 297–302, 1954.

[12] B. Dobler and R. Bendl, "Precise modelling of the eye for proton therapy of intra-ocular tumours." *Phys. Med. Biol.*, vol. 47, no. 4, pp. 593–613, Feb. 2002.

[13] S. Donaldson and L. Smith, "Retinoblastoma: Biology, presentation, and current management," *Oncology*, vol. 3, no. 4, pp. 45–51, 1989.

[14] M. Ester, H. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. Data Mining, Discovery*, 1996, pp. 226–231.

[15] A. Frangi and W. Niessen, "Multiscale vessel enhancement filtering," *Med. Image Comput. Comput.-Assisted Intervention*, vol. 1496, pp. 130–137, 1998.

[16] P. D. Graaf, S. Göricke, F. Rodjan, P. Galluzzi, P. Maeder, J. A. Castlijns, and H. J. Brisse, "Guidelines for imaging retinoblastoma: Imaging principles and MRI standardization," *Pediatric Radiol.*, vol. 42, no. 1, pp. 2–14, 2012.

[17] M. J. Hogan, J. A. Alvarado, and J. E. Weddell, *Histology of the Human Eye: An Atlas and Textbook*. Philadelphia, PA, USA: Saunders, 1971.

[18] A. Hoover and M. Goldbaum, "Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels," *IEEE Trans. Med. Imag.*. vol. 22, no. 8, pp. 951–958, Aug. 2003.

[19] P. V. C. Hough, "Methods and means for recognizing complex patterns," U.S. Patent 3 069 654, Dec. 18, 1962.

[20] O. C. Huiqi Li, "Automatic location of optic disk in retinal images," in *Proc. Int. Conf. Image Process.*, 2001, vol. 2, pp. 837–840.

[21] A. Isambert, F. Dhermain, F. Bidault, O. Commowick, P.-Y. Bondiau, G. Malandain, and D. Lefkopoulos, "Evaluation of an atlas-based automatic segmentation software for the delineation of brain organs at risk in a radiation therapy clinical context," *Radiotherapy Oncol.: J. Eur. Soc. Therapeutic Radiol. Oncol.*. vol. 87, no. 1, pp. 93–99, Apr. 2008.

[22] T. Kivela, "The epidemiological challenge of the most frequent eye cancer: Retinoblastoma, an issue of birth and death," *Brit. J. Ophthalmol.*, vol. 93, pp. 1129–1131, 2009.

[23] G. Loy and A. Zelinsky. (2002). A fast radial symmetry transform for detecting points of interest. *Computer Vision–ECCV 2002*. [Online]. Available: http://link.springer.com/chapter/10.1007/3-540-47969-4_24

[24] A. Maccarthy, J. M. Birch, G. J. Draper, J. L. Hungerford, J. E. Kingston, M. E. Kroll, Z. Onadim, C. A. Stiller, T. J. Vincent, and M. F. G. Murphy, "Retinoblastoma in Great Britain 1963–2002," *Brit. J. Ophthalmol.*, vol. 93, no. 1, pp. 33–38, 2009.

[25] T. M. M.Goitein, "Planning proton therapy of the eye," *Med. Phys.*, vol. 10, no. 275, pp. 275–283, 1983.

[26] F. L. Munier, J. Verwey, A. Pica, A. Balmer, L. Zografos, A. Hana, T. Beate, G. Gudrun, and M. Raphaël, "New developments in external beam radiotherapy for retinoblastoma: From lens to normal tissue-sparing techniques," *Clin. Exp. Ophthalmol.*, vol. 36, no. 1, pp. 78–89, Jan. 2008.

[27] B. S. P. L. Hammer and P. Hansen, "Roof duality, complementation and persistency in quadratic 01 optimization," *Math. Program.*, vol. 28, pp. 121–155, 1984.

[28] PaulJaccard, "The distribution of the flora in the alpine zone," *New Phytologist*, vol. 11, no. 2, pp. 37–50, 1912.
[29] M. B. Rüegsegger, M. Bach Cuadra, A. Pica, C. A. Amstutz, T. Rudolph, D. Aebersold, and J. H. Kowal, "Statistical modeling of the eye for multimodal treatment planning for external beam radiation therapy of intraocular tumors," *Int. J. Radiation Oncol., Biol., Phys.*, vol. 84, no. 4, pp. e541–e547, Nov. 2012.
[30] F. Schaeffel, "Kappa and Hirschberg ratio measured with an automated video gaze tracker," *Optometry Vis. Sci.*, vol. 79, no. 5, pp. 329–334, 2002.
[31] H. Vogel, "A better way to construct the sunflower head," *Math. Biosci.*, vol. 44, nos. 3/4, pp. 179–189, 1979.
[32] S. G. Wetzel, G. Johnson, A. G. S. Tan, S. Cha, E. A. Knopp, V. S. Lee, D. Thomasson, and N. M. Rofsky, "Imaging of the brain with a volumetric interpolated examination," *Amer. J. Neuroradiol.*, vol. 23, pp. 995–1002, 2002.
[33] D. K. Wong, J. Liu, J. H. Lim, X. Jia, F. Yin, H. Li, and T. Y. Wong, "Level-set based automatic cup-to-disc ratio determination using retinal fundus images in ARGALI," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. Conf.*, Jan. 2008, pp. 2266–2269.
[34] A. A.-H. A.-R. Youssif, A. Z. Ghalwash, and A. A. S. A.-R. Ghoneim, "Optic disc detection from normalized digital fundus images by means of a vessels' direction matched filter," *IEEE Trans. Med. Imag.* vol. 27, no. 1, pp. 11–18, Jan. 2008.

Authors' photographs and biographies not available at the time of publication.