Superimposition of eye fundus images for longitudinal analysis from large public health databases

Guillaume Noyel, Rebecca Thomas, Gavin Bhakta, Andrew Crowder, David Owens, Peter Boyle

Superimposition of eye fundus images for longitudinal analysis from large public health databases

G Noyel[1], R Thomas[2], G Bhakta[3], A Crowder[3], D Owens[2] and P Boyle[1, 4]

[1]International Prevention Research Institute, 95 Cours Lafayette, Lyon, 69006 France

[2]Diabetes Research Group, Institute of Life Sciences, College of Medicine, Swansea University, SA2 8PP, Wales, UK

[3]Diabetic Eye Screening Wales, 1 Fairway Court, Tonteg Road, Upper Boat, Treforest, Pontypridd CF37 5UA, Wales, UK

[4]Strathclyde Institute of Global Public Health, University of Strathclyde, Glasgow, UK

e-mail: guillaume.noyel@i-pri.org, r.l.thomas@swansea.ac.uk, owensDR@cardiff.ac.uk, gavin.bhakta@drssw.wales.nhs.uk, andrew.crowder@drssw.wales.nhs.uk, peter.boyle@i-pri.org

Abstract

In this paper, a method for superimposition (i.e. registration) of eye fundus images from persons with diabetes screened over many years for Diabetic Retinopathy is presented. The method is fully automatic and robust to camera changes and colour variations across the images both in space and time. All the stages of the process are designed for longitudinal analysis of cohort public health databases. The method relies on a model correcting two radial distortions and an affine transformation between pairs of images which is robustly fitted on salient points. Each stage involves linear estimators followed by non-linear optimisation. The model of image warping is also invertible for fast computation. The method has been validated 1) on a simulated montage and 2) on public health databases with 69 patients with high quality images (with 271 pairs acquired mostly with different type of cameras and with 268 pairs acquired with the same type of cameras) with a success rates of 96 % and 98 % and 5 patients (with 20 pairs) with low quality images with a success rate of 100%. Compared to two state of the art methods, ours gives better results.

Keywords: eye fundus images, image registration, invertible model, longitudinal analysis, public health databases, radial distortion

Mathematics Subject Classification (MSC):
68U10 Image processing
65D19 Computational issues in computer and robotic vision

Submitted to: Physics in Medicine and Biology

1 Introduction

Diabetic Retinopathy (DR) is one of the major causes of visual impairment in the world and therefore represents a major public health challenge. It is a complication of both types of diabetes mellitus, which affects the light perception part of the eye (retina). DR may lead to the development of sight threatening lesions and without adequate and timely treatment the patient could lose their sight and eventually become blind (International Diabetes Federation and The Fred Hollows Foundation, 2015; Scanlon *et al.*, 2009). DR is often asymptomatic until an advanced stage, thereby screening to detect sight threatening DR at an early stage is essential which has resulted in the introduction of DR Screening services in many countries such as UK (Harding *et al.*, 2003), USA, the Netherlands, France, etc. The commonest screening method involves acquiring eye fundus images on an annual or biennial basis.

As these DR screening programs have been in existence over several years, performing longitudinal analysis of the eye fundus images of the same patient is now possible. However, in order to accurately compare the evolution of DR over time, the images must be perfectly super-imposed.

The direct superimposition of two images of the same patient never gives good results (see Figure 1). Indeed, for two separate photographic-eye examinations the patient is never in exactly in same position and also the camera may differ. Therefore, the super-imposition method has to take into account the different causes of the deformation such as:

- The position of the patient: by taking into account rotation, translation and scaling.
- The change of the camera: by using scaling.
- The projection of a 3D scene assimilated to a sphere (the retina of the eye) onto the 2D plane of the sensor of the camera: by using a radial correction process.
- The radial deformation due to the optics of the camera: by using a radial correction process.
- The colour variability between images due to the light intensity and sensor.

To perform a superimposition – also named registration – two stages are regarded as necessary: a model of deformation and a matching criterion to fit the model. There are several models in existence to allow super-imposition between pairs of eye fundus images. The earliest methods relied on fluorescein images and based on a composition of translation, rotation and scaling - i.e. an affine transformation model (Zana and Klein, 1999a, b). The bifurcations of the vessels were used to match the points and fit the model. Another matching criterion consists in the minimisation of image intensity differences (Cideciyan, 1995; Matsopoulos *et al.*, 1999; Ritter *et al.*, 1999; Adal *et al.*, 2014).

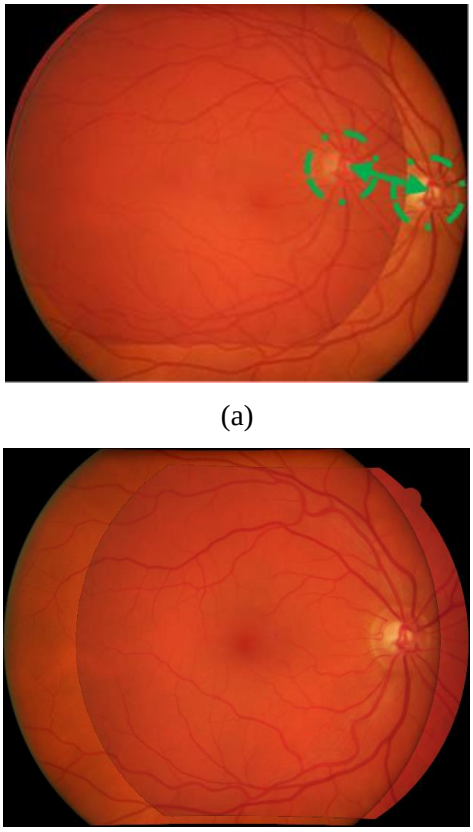
Other methods are based on a similarity (i.e. a rotation and a translation) (Matsopoulos *et al.*, 1999; Lloret *et al.*, 2000) and an elastic model of deformation (You *et al.*, 2005; Fang and Tang, 2006). In Jian *et al.* (2010); Ghassabi *et al.* (2013), new descriptors PIIFD (Partial Intensity Invariant Feature Descriptor) have been introduced for multimodal image superimposition between auto-fluorescence, infrared and red-free images (i.e. the green component of a colour eye fundus images). They have used an affine and a quadratic model. This method, powerful for multimodal image registration, has not been designed and tested for the superimposition of colour eye fundus images.

More recently, it has been shown that a quadratic model gives better registration results (Adal *et al.*, 2014; Can *et al.*, 2002; Stewart *et al.*, 2003; Chanwimaluang *et al.*, 2006; Ryan *et al.*, 2004). The difficulty inherent in these models is to estimate their parameters. To overcome such a limitation, a radial distortion model has been introduced by Lee *et al.* (2007) and compared to previous methods in Lee *et al.* (2010). It consists of adding a radial model to the affine transformation in order to correct the effects of radial distortion due to the geometry of the camera and of the eye.

Superimposition of eye fundus images have also been performed in the tri-dimensional space using a model of near planar surface (Yuping and Medioni, 2008) or an ellipsoid model (Hernandez-Matas *et al.*, 2016).

However, for all these methods the superimposition of eye fundus images is performed with colour images acquired by the same camera during the same examination. However, superimposing colour images acquired at different times by different cameras in large databases still remains a challenging problem. Two new issues are appearing:

1. the radial distortion must be corrected for the two different lenses.
2. the colour of the eye has changed because the source of light has changed and the anatomy of the eye may have possibly evolved (e.g. cataract removal, evolution of the retinopathy to a more severe stage, etc.).



(a)

(b)

Figure 1. Superimposition of eye fundus images. (a) Naïve superimposition of images. (b) Perfect superimposition of images by using a model.

Currently, for colour eye fundus images, no existing method has been designed to address the problem of the change of camera lenses, camera and light for retinal examination used for DR screening (i.e. at a year of interval). In addition, in the existing methods, even when one radial distortion is corrected by a radial distortion correction or by a second order model, the registration criterion requires a similar and uniform intensity (i.e. colour) for the images, particularly if the extraction of anatomical features such as the vessels is needed. These methods are working in laboratory conditions (several acquisitions on the same day with the same equipment) but failed when they are used on public health databases constituted by routine screening across years of patients with Diabetes.

In this paper, our contribution has been to address this challenge by presenting a robust superimposition method designed for longitudinal screening of large public health image based databases. In particular, it takes into account the variation of colour, two radial distortions and the registration criterion is based on a standard salient point extraction, which does not rely on the extraction of anatomical features. In addition, our method is designed for a fast computation which is necessary for large databases. Indeed, the model is invertible, which is useful for image deformation (i.e. image warping) and the parameters of our model can be estimated by linear equations which are introduced in this paper.

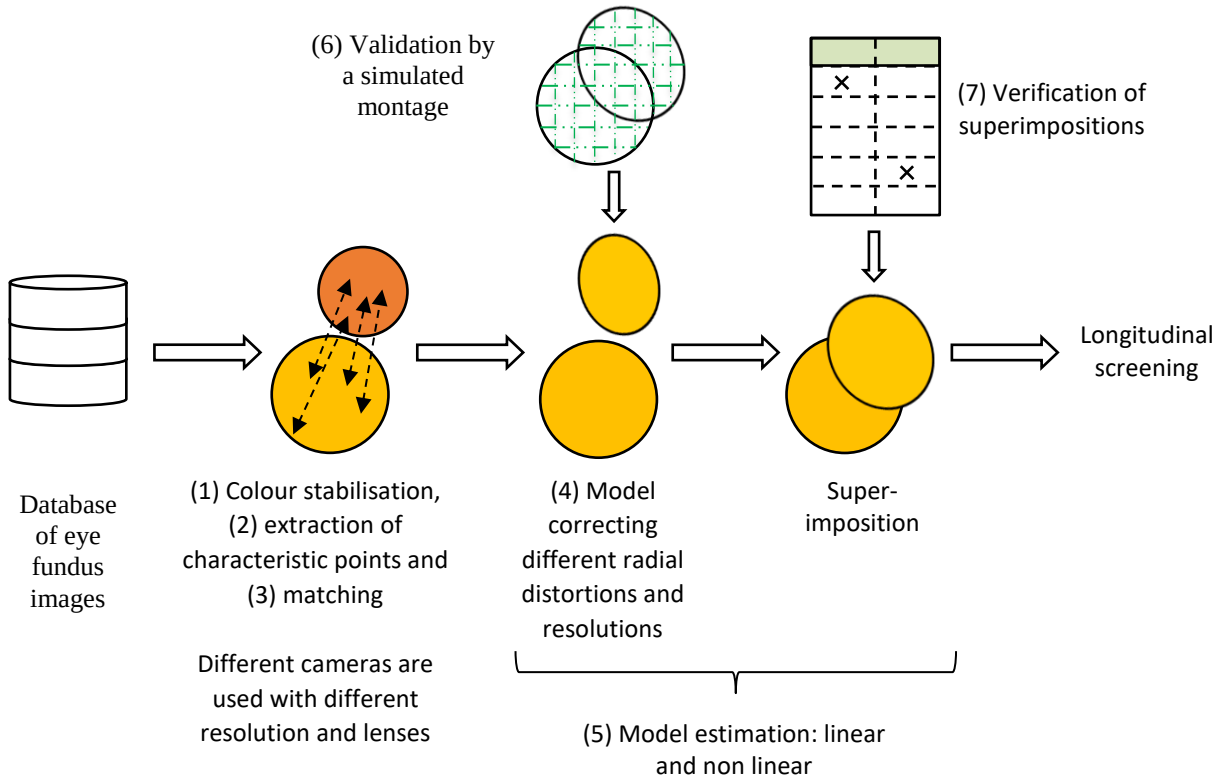


Figure 2. Framework of eye fundus superimposition of images for longitudinal screening. After colour stabilization (pre-processing) and extraction of characteristic points, the points are matched and the model is estimated. The differences are corrected in term of radial distortions and resolutions between the images.

The paper is organized as follows: after presenting a complete method to superimpose pairs of images, we will present a quality check of the registration and finally, validation of our methodology using different patient databases.

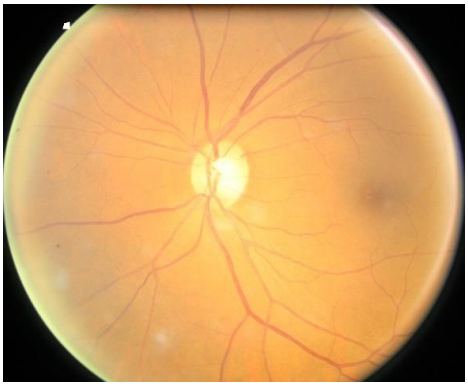
2 Methods

During a photographic eye examination, eye fundus (retina) two images of both eyes are acquired i.e. a 45 degree "nasal" and "macular" field (Figure 3). The aim of this study was to develop robust algorithms for superimposition of images in the same positions while being captured during two different exams and often with different cameras and resolutions. Our aim was not to develop large mosaics of eye fundus images acquired during the same examination with the same camera (Chanwimaluang *et al.*, 2006), but to propose a robust method for longitudinal studies involving large databases with images acquired in heterogeneous conditions, i.e. different cameras, different sources of light, different lenses, retinal examinations with screening interval of a year. Therefore, we have paid particular attention to the development of robust and fast algorithms for longitudinal screening in large databases.

Our method is based on a pre-processing stage consisting of (1) normalizing the colour of the eye fundus image (Noyel *et al.*, 2015), (2) the extraction of characteristic points in pairs of images, (3) a matching procedure, (4) the use of a model correcting radial distortion of both images and (5) the estimation of the parameters of the model by a robust optimization. The method is validated (6) on a simulated montage and the superimpositions of the image of the database are verified (7).

A schematic description of the study is represented in Figure 2. The different stages have been designed to provide efficient solutions to the superimposition of images acquired for practical screening. Between two examinations, the camera might have been changed producing differences in colour, resolution and radial distortion between images. Moreover, we will show that our method is efficient both on high quality images acquired following pupillary dilation and low quality images without dilatation of the pupils prior to photography.

(a)



(b)

Figure 3. (a) Macular and (b) nasal view of an eye fundus acquired without pupil dilation and in harsh conditions.

2.1 Extraction of characteristic points

The brightness of eye fundus images is non-uniform due to various reasons: disease such as cataract, motion of the patient, acquisition conditions and differences in absorptions of the light in the eye (Walter, 2003; Walter and Klein, 2005). Some parts of the images appear as bright while others are dark. Moreover, the possible change of the eye fundus camera between two separate examinations may contribute to a change in the colour between two images of the same eye (Figure 4).

We have used a method (Noyel *et al.*, 2015) to correct the variations of colour contrast between the images. Results can be seen in Figure 4. This is then followed by the extraction of several salient points (Figure 5) using the Scale-Invariant Feature Transform - SIFT - algorithm (Lowe, 2004; Vedaldi and Fulkerson, 2008). The SIFT algorithm has been designed to be robust to the variation of observation angle and to some variations in lightning. Briefly SIFT consists of extracting key points based on a multiscale analysis. Then series of descriptors are computed for each salient point. These descriptors are used for point matching

A similar detector, the SURF - Speeded Up Robust Features - detector, was previously used for eye fundus image superimposition (Cattin *et al.*, 2006).

(a)

(b)

(c)

(d)

Figure 4. (a), (c) Eye fundus images of the same patient acquired during two exams with 1.5 years of interval. (b), (d) Colour stabilisation of eye fundus images (a) and (c) with a low quality.

2.2 Point matching

As point matching with Lowe's method (Lowe, 2004) is not robust enough to estimate the rotation on our database, we have created a three-step procedure:

A first matching by Lowe's method followed by a refined selection of the correspondence vectors according to their size and orientation

An estimate of the homography using the algorithm of section 2.4.1. The position of the key points is modified according to the homography.

Step (a) is applied a second time using the transformed points.

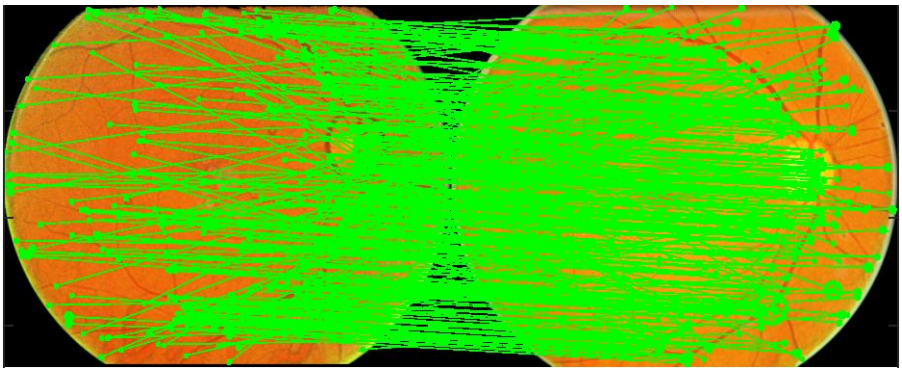
After this first matching, one image is put on the right while the other on the left after resizing and padding to have the same size (Figure 5). As some points are incorrectly matched, some correspondences vectors v between matched points are inconsistent. Using the two images stitched, a two-step selection is performed on the lengths l and orientations θ of the vectors:

only the vectors v whose length l_v and orientation θ_v is in the interval $\{|l_v - E\{l_v\}| \leq \sigma\{l_v\}\ and\ |\theta_v - E\{\theta_v\}| \leq 5°\}$. $E\{\}$ is the mean and $\sigma\{\}$ the standard deviation of a variable.

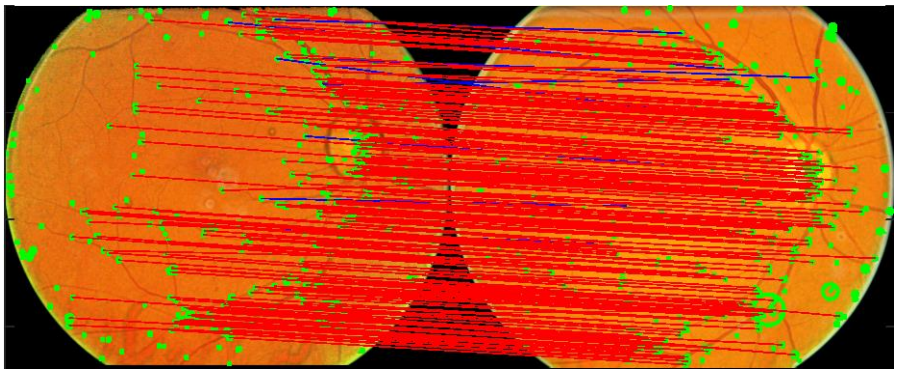
Among the selected vectors $\tilde{v}$, only the vectors whose length $l_{\tilde{v}}$ and orientation $\theta_{\tilde{v}}$ is in the interval $\{|l_{\tilde{v}} - E\{l_{\tilde{v}}\}| \leq \max(3\sigma\{l_{\tilde{v}}\}, 5\% \times ysize)\ and\ |\theta_{\tilde{v}} - E\{\theta_{\tilde{v}}\}| \leq \max(5°, \sigma\{\theta_{\tilde{v}}\})\}$. $ysize$ is the number of lines in the image.

Remarks: the value 5° is a tolerance on a relative difference of orientations. Therefore, it must be chosen close to zero. 5° has been empirically selected and perfectly works for all the processed images. Even if there are some incorrect matchings, they will be discarded during the estimation process of the homography (see section 2.4.1. Indeed, the homographies estimated with these incorrect matchings will be discarded.

After matching the points, a model of transformation is estimated.



(a)



(b)

Figure 5. Point extraction and matching between images. (a) Initial matching. (b) Matching after first simplification (in red and in blue) and after second simplification (in blue).

2.3 Model of deformation

The model of deformation ensures a correct superimposition between the images. Several deformations are taken into account: (i) the difference in terms of positions of the eye between a pair of images will be corrected by an affine transformation (i.e. and homography) and (ii) the radial deformations due to the projection of the eye into the camera and due to the optics of the camera (Hartley and Zisserman, 2004) will be corrected using a radial transformation.

Lee *et al.* (2007); (Lee *et al.*, 2008) have proposed a model coupling a unique radial transformation for both images and a homography. Lee *et al.* (2010) have made the comparison with two other second order models.

In this paper, as we were interested in analyzing images of patients acquired during exams with an approximate one year interval, we extended their approach by defining a model with one homography H and two radial distortions, i.e. one for each image. Indeed, the camera may have changed between screening exams on a large number of patients.

The affine homography H is defined as:

$$H = \begin{bmatrix} A & T \\ O^T & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ is an affine transformation, $\forall i,j \in [1 \dots 2],\ a_{ij} \in \mathbb{R}$, and $T = \begin{bmatrix} t_x \\ t_y \end{bmatrix}$, $t_x, t_y \in \mathbb{R}$, is a translation.

The radial distortion due the background of the sphere surface of the eye and of the radial distortion of the camera was modelled by a *division model* (Fitzgibbon, 2001) in the following way:

$$\bar{P}^d = \left(1 + k\left(r^d\right)^2\right).\bar{P}^u \quad (2)$$

- $P^d \in \mathbb{R}^2$ are the distorted coordinates in the original (i.e. distorted) image
- $P^u \in \mathbb{R}^2$ are the undistorted coordinates in the undistorted image
- $\bar{P}^d \in \mathbb{R}^2$ are the distorted coordinates centred on the image centre c: $\bar{P}^d = P^d - c$
- $\bar{P}^u \in \mathbb{R}^2$ are the undistorted coordinates centred on the image centre c: $\bar{P}^u = P^u - c$
- $r^d = \left\|P^d - c\right\| = \left\|\bar{P}^d\right\| \in \mathbb{R}$, is the distance of the deformed coordinates P^d from the optic centre c (i.e. the image centre).
- k is a real parameter of distortion in $[-0.2\ ;\ 0.2]$, after normalisation by $(1 + \|c\|)^2$.

The model was named division model because the distorted coordinates were divided by the radial distortion $\bar{P}^d / \left(1 + k\left(r^d\right)^2\right) = \bar{P}^u$. The distorted image corresponds to the original image and the undistorted image is the image after the correction of radial distortion.

Given P_1^d and P_2^d the coordinates of the points in the original (i.e. deformed) images 1 and 2, k_1 and k_2 the distortion parameters, $c_1$and c_2 the image centres, the model mapping image 1 into image 2 is defined as follows:

$$\frac{\bar{P}_2^d}{\left(1 + k_2\left(r_2^d\right)^2\right)} + c_2 = H\left[\frac{\bar{P}_1^d}{\left(1 + k_1\left(r_1^d\right)^2\right)} + c_1\right] \quad (3)$$

If the camera used to acquire both images is the same, the distortion parameters are equal, $k_1 = k_2$, and the model corresponds to the model of Lee *et al.* (2007).

The model is estimated after having extracted and matched the points in the pair of original images (target and reference). Therefore, the radial distortion correction is performed after the detection of the feature correspondence points in the original (i.e. distorted) image. It has been programmed in the subsequent way.

2.4 Estimation of the model parameters

The parameters of the model are estimated by a different method of (Lee *et al.*, 2007; Abramoff *et al.*, 2012). Indeed, the radial distortion is estimated after the homography without needing any preliminary estimation by a calibration of the camera (Hartley and Zisserman, 2004). Moreover, we have conceived linear initializers at each step of the optimization of the model (Figure 6).

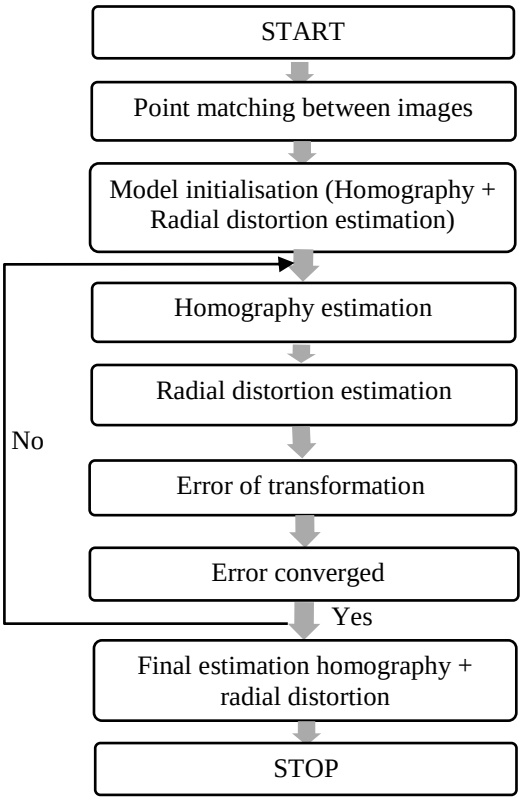


Figure 6. Flowchart of the superimposition method based on homography and radial distortion.

2.4.1 Estimation of the homography.

An affine homography can be decomposed according to the following scheme (Hartley and Zisserman, 2004):

$$A = R(\theta)R(-\phi)DR(\phi) \tag{4}$$

$R(\theta)$ and $R(\phi)$ are rotation matrices of angle θ and ϕ respectively and D is a diagonal matrix:

$$D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \tag{5}$$

λ_1 and λ_2 are two scaling values. The matrix A is a composition of a rotation by ϕ, a scaling by λ_1(respectively λ_2) in x (respectively y) direction, a rotation by $-\phi$ and then another rotation by θ. The decomposition is obtained using the Singular Value Decomposition (SVD) method.

The estimation of an homography is basically performed using the "gold standard" algorithm of Hartley and Zisserman (2004). This algorithm estimates around 100 homographies on randomly selected 4 pairs of points. Then the homography with the minimum error when transforming the matched points is selected. Therefore, the incorrect matchings are removed. However, as there is also a radial distortion in the image, the deformation is not entirely modelled by a homography. Therefore, some homographies must be discarded. In particular, those with a scaling factor on the x and y axis with a relative difference greater than 1%, because the sensor resolution is almost the same in x and in y. For this purpose, several estimates (until 50) using the gold standard algorithm are performed if the relative difference between the scaling factors is greater than 1%. If the value of 1% is never reached, then the homography with the smallest relative difference between the scaling factors is kept.

2.4.2 Estimation of the model with one radial distortion.

When the camera is the same for both images, only one radial distortion needs to be estimated. In the flowchart of Figure 6, the convergence criteria becomes:

$$convergence = [err_n < \varepsilon]\ and \left[\frac{err_n - err_{n-1}}{err_{n-1}} < tol\right]\ and\ [n < MaxIter]$$

err is the mean error at iteration n, $\varepsilon = 0.01$ is the tolerance on the error, $tol = 0.01$ is a tolerance on the relative error between iterations and $MaxIter = 100$ is the maximum number of iterations.

Linear estimators are used at each step of the parameter estimation. The final optimisation is performed with a linear estimator followed by a non-linear optimiser such as Levenberg-Marquardt (More, 1977; Bonnans *et al.*, 2006). We will now present the linear estimators.

2.4.2.1 Linear estimator of the radial distortion parameter.

Using equations (1) and (3), the following equation is obtained:

$$\frac{\bar{P}_2^d}{\left(1 + k\left(r_2^d\right)^2\right)} + c_2 = A\left[\frac{\bar{P}_1^d}{\left(1 + k\left(r_1^d\right)^2\right)} + c_1\right] + T \tag{6}$$

A is the matrix of the affine transformation and T is the vector of translation. Equation (6) implies that:

$$k^2\left[\left(r_1^d r_2^d\right)^2 . d\right] + k\left[\left(r_1^{d\,2} + r_2^{d\,2}\right) . d + r_1^{d\,2} . \bar{P}_2^d - r_2^{d\,2} . A\bar{P}_1^d\right] = -\left[\bar{P}_2^d + d - A\bar{P}_1^d\right] \tag{7}$$

with $d = c_2 - Ac_1 - T$. Equation (7) is a linear equation in k when H (i.e. A and T) is known. k is determined using least squares algorithm.

2.4.2.2 Linear estimator of the homography H and of the radial distortion parameter k.

Using equation (6) the following equation is determined:

$$k^2 d\left[\left(r_1^d r_2^d\right)^2\right] + kd\left[r_1^{d\,2} + r_2^{d\,2}\right] + k\left[r_1^{d\,2} . \bar{P}_2^d\right] - kA\left[r_2^{d\,2} . \bar{P}_1^d\right] - A\left[\bar{P}_1^d\right] + d[1] = -\left[\bar{P}_2^d\right] \tag{8}$$

Equation (8) is a linear equation with its variables in the brackets. k and H (with the intermediate of A and d) are estimated using least squares.

2.4.3 Estimation of the model with two radial distortions.

When a different camera is used for each image of a pair, two radial distortions must be estimated.

The convergence criteria (Figure 6) becomes:

$$convergence = [err_n < \varepsilon]\ and \left[\frac{err_n - err_{n-1}}{err_{n-1}} < tol\right]\ and\ \left[k_1 and\ k_2 \in [-0.2\ ;\ 0.2]\right]\ and\ [n < MaxIter]$$

As for a single distortion, the same parameters $\varepsilon = 0.01$, $tol = 0.01$, $MaxIter = 100$ are used. For each estimate, the radial distortion parameters k_1, k_2 must be in the interval $[-0.2\ ;\ 0.2]$. If not, the algorithm stops and the model estimate with the smallest error is selected.

Linear estimators are used at each step of the parameter estimation. The final optimisation is followed by a non-linear optimiser such as the "trust region method" (Moré, 1983; Bonnans *et al.*, 2006) with bounds $[-0.2\ ;\ 0.2]$ for the radial distortion parameters k_1, k_2.

2.4.3.1 Linear estimator of the radial distortion parameters k_1 and k_2.

Equation (3) gives the following equation:

$$\frac{\bar{P}_2^d}{\left(1 + k_2\left(r_2^d\right)^2\right)} + c_2 = A\left[\frac{\bar{P}_1^d}{\left(1 + k_1\left(r_1^d\right)^2\right)} + c_1\right] + T \tag{9}$$

Equation (9) implies that:

$$k_1 k_2\left[\left(r_1^d r_2^d\right)^2 . d\right] + k_1\left[{r_1^d}^2 . d + {r_1^d}^2 . \bar{P}_2^d\right] + k_2\left[{r_2^d}^2 . d - {r_2^d}^2 . A\bar{P}_1^d\right] = -\left[\bar{P}_2^d + d - A\bar{P}_1^d\right] \tag{10}$$

with $d = c_2 - Ac_1 - T$.

Equation (10) is a linear equation in k_1 and k_2 when H (composed of A and d) is known. k_1 and k_2 are estimated using least squares method.

2.4.3.2 Linear estimator of the radial distortion parameters k_1 and k_2 and the homography.

From equation (3):

$$k_1 k_2 d\left[\left(r_1^d r_2^d\right)^2\right] + k_1 d\left[{r_1^d}^2\right] + k_1\left[{r_1^d}^2 . \bar{P}_2^d\right] + k_2 d\left[{r_2^d}^2\right] - k_2 A\left[{r_2^d}^2 . \bar{P}_1^d\right] - A\left[\bar{P}_1^d\right] + d[1] = -\left[\bar{P}_2^d\right] \tag{11}$$

Equation (11) is a linear equation with its variables in the bracket. k_1, k_2 and H (with the intermediate of A and d) are estimated using least squares.

2.5 Image warping

In order to analyse a large database, a fast algorithm of image warping is needed. Forward warping is time consuming and so we therefore use inverse warping. However, the registration model needs to be invertible (Wolberg, 1990).

The radial distortion is modelled in equation (3) by a *division model* (Fitzgibbon, 2001). Wonpil (2003) and Park *et al.* (2009) have computed an approximate transformation for a standard distortion method. Here, we compute the exact inversion of the *division model*.

Given $r^u = \|P^u - c\| = \|\bar{P}^u\| \in \mathbb{R}$, the distance of the undistorted coordinates P^c from the optic centre c, using equation (2), we have:

$$r^d = \left(1 + k\left(r^d\right)^2\right) . r^u \tag{12}$$

Equations (2) and (12), implies that:

$$\bar{P}^u = \frac{r^u}{r^d}\bar{P}^d = W^{-1}(\bar{P}^d) \tag{13}$$

In order to use invert warping, it is necessary to determine W^{-1} transforming the distorted points P^d the undistorted points P^u. From equation (13), it is equivalent to determine r^u knowing r^d.

Equation (12) is equivalent to: $kr^u r^{d\,2} - r^d + r^u = 0$, a second order equation in r^d. Its discriminant is equal to: $\Delta = 1 - 4kr^{u\,2}$ with $\Delta > 0$. Its roots are $r^d = \frac{1 \pm \sqrt{1-4kr^{u\,2}}}{2kr^u}$.

The inverse transformation W^{-1} corresponds to the root:

$$r^d = W^{-1}(r^u) = \frac{1 + \sqrt{1 - 4kr^{u\,2}}}{2kr^u} \tag{14}$$

Therefore, the transformation used is invertible. An invertible image warping method compared to a non-invertible method reduces the time from about 10 minutes to a few seconds on a standard computer using Matlab (16Go RAM, processor Intel i7-4702HQ, 2.20GHz).

In Figure 2 and in Figure 7, the results of superimposition with the radial distortion model are shown. One can notice the good quality of the superimposition. In the next section we will evaluate the quality of superimposition.

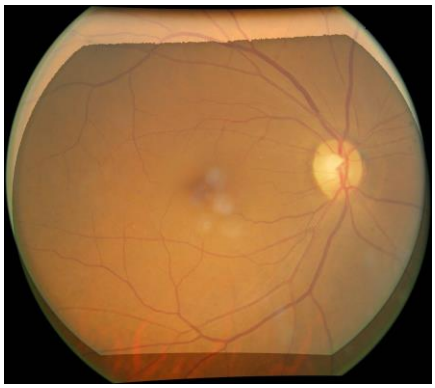


Figure 7. Superimposition of a pair of eye fundus images with correction of two radial distortions.

3 Results

The quality of image superimposition was evaluated through a simulated montage and using a database of patients. This latest validation is important as our method has been designed to analyse large public health image databases.

3.1 Validation by a simulated montage

We have created a montage by superimposing two eye fundus images and deforming them according to the methods presented by Lee *et al.* (2010). We have taken real images registered with an overlap percentage of 80% corresponding to the case that we have in a longitudinal database. No modification of colour was done to the images. After adding equally spaced landmarks, we have cut and deformed the images according to the model of Lee *et al.* (2010).

An affine transformation has been used, with rotation scaling and shearing. Then the image has been modified by a projective distortion. The radius of the eye ball is equal to the ratio between the radius of the disk of the image divided by the observation angle of the camera (45 degrees).

Then we have registered the images and measured the error between the landmarks after registration and their true position. With a single distortion, we have obtained a mean registration error of 0.86 pixels (standard deviation 1.75 pixels) in images of size 1568 x 2352 pixels (Figure 8) with vessels of maximum diameter greater than of 30 pixels. The relative error respective to the image is 0.03%, and respective to the vessels is 2.9 %. With two distortions, the mean registration error is of 0.92 pixels (standard deviation 1.94 pixels) and the relative error is 0.03 % respective to the image and 3.1 % respective to the vessels (see Table 1). The error is mainly located on the external part of the superimposed image.

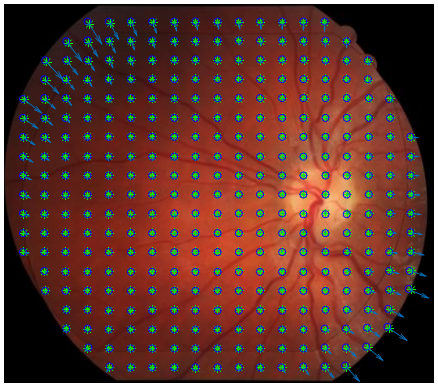


Figure 8. Validation of the superimposition model by registering a pair of images previously deformed. The green points corresponds to the points of the reference image and the blue points to the points of the current image. The arrows represent the registration errors between the two images.

We have compared our method to another state of the art method - whose software is publicly available - "gdbicp" developed by Yang *et al.* (2007). It superimposes eye fundus images with a quadratic model (Stewart *et al.*, 2003). Using the simulated montage, the mean registration error is of 2.44 pixels (standard deviation 1.64 pixels). Therefore, our model has a smaller error than the quadratic one of "gdbicp" (see Table 1). This confirms the results of Lee, et al. [16] who previously showed that quadratic model have a greater error than a homography and a radial distortion model. We remind that the model of deformation proposed in this paper with a single radial distortion corresponds to the model of Lee *et al.* (2007), however, the fitting method used to estimate the parameters is different.

Such results demonstrate that our method gives a superimposition without noticeable difference. Therefore, this approach is suitable to perform an analysis in a large database.

Table 1. Errors with the simulated montage

Model	Mean (pixels)	Standard Deviation (pixels)	Mean relative error respective to the image (%)	Mean relative error respective to the vessels (%)
Homography and 1 radial distortion	0.86	1.75	0.03 %	2.9 %
Homography and 2 radial distortions	0.92	1.94	0.03 %	3.1 %
"gdbicp" quadratic	2.44	1.64	0.08 %	8.1 %

3.2 Validation with a public health database

In order to assess the evolution of Diabetic Retinopathy several screening programs in the world are in existence.

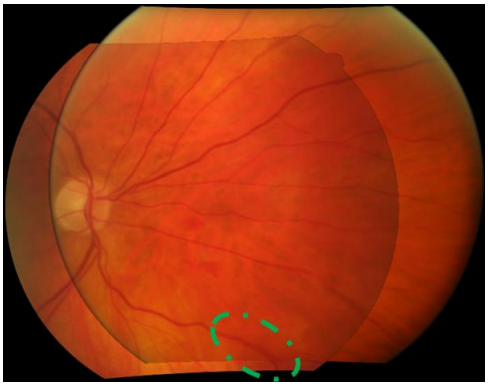
Among them, in the United Kingdom, in Wales, the Diabetic Eye Screening Wales (DESW) has developed a program to screen the whole population with diabetes over the age of 12 years old. The programme has been in existence for just over 10 years at a national level and several thousands of patients have been screened annually for five or more years (Thomas *et al.*, 2012).

We have performed trials of a database of 69 patients coming from the DESW. For each patient we have kept two series of two examinations with an approximate screening interval of one year between the examination events. For each event exam, four images are available, two positions (nasal and macular) for each eye. There were two series of images acquired for different years, with the first series are made up of 271 pairs of sufficient image quality (63% are acquired with different cameras) and the second series included of 268 pairs (9% are acquired with different cameras). For each position, we have performed the superimposition of the images between the two different examinations. For all pairs of images the superimposition has been visually checked by an expert. The classification has been done in two categories: 1) no noticeable difference (correct) and 2) noticeable difference (incorrect).

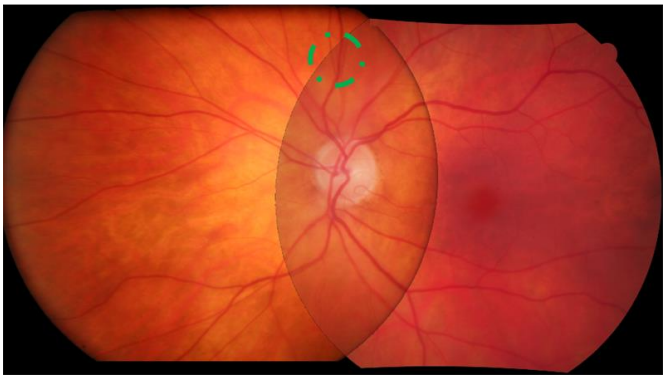
The category “incorrect” includes differences of the size of a small diameter vessel, differences of the size of a large diameter vessel or even larger. The results are presented in table 2.

In the first series of 271 pairs, 2 pairs have small differences in the external part of the superimposition. These differences are of the size of the diameter of a vessel (Figure 9). When the percentage of overlapping surface is low (around 30 %) compared to the surface of the superimposed image, we have noticed differences of the size of 1 vessel on 8 pairs of images. Therefore, the superimposition was successful for 96 % of the pairs and 99% if we consider the pairs in the same position.

(a)



(b)



(c)

Figure 9. Examples of superimposition. (a) Good superimposition. (b) Superimposition with small differences. (c) Superimposition with small differences (in the green circle) and a small overlap.

In the second series of 268 pairs of images acquired a few years later, there are 6 images with a small difference (i.e. less than 1 vessel diameter). The superimposition was successful for 98 % of the pairs. However, in these images the central part was perfectly superimposed. We have developed another algorithm using, in addition to the matched points, the distance between the retinal vessels. This algorithm similar to those described by Can *et al.* (2002) and

Lee *et al.* (2010) solves the registration problem for the images with a small overlapping area. These findings will be presented in a future paper.

In the pairs with a sufficient overlap, no noticeable difference has been perceived between them. This means that our method is suitable to be applied to analyse large databases.

We have compared our method to the state of the art method “gdbicp” (Stewart *et al.*, 2003; Yang *et al.*, 2007) which can make superimposition of images with a quadratic model (Table 2). In the first series with 271 pairs (respectively second series with 268 pairs), their method has found a superimposition solution for 77% of pairs (resp. 61%) while our method gives a correct result for 96% of the pairs (resp. 98%). In addition, when using with “gdbicp” a homography and two radial distortions using a multiplicative model - which is therefore not invertible - less pairs are registered (48% in the first series and 33% in the second series). Therefore, our method has better results than those obtained with “gdbicp”. In addition, the method of “gdbicp” with a quadratic model gives better results for eye fundus images than “gdbicp” with radial distortions. Indeed, the method “gdbicp” with a quadratic model has been designed to superimpose eye fundus images in Stewart *et al.* (2003).

We have also compared our model to the one of Lee *et al.* (2007) by producing results using our fitting approach with an homography and a single radial distortion for all the image pairs without colour stabilisation (Table 2). In the first series with 271 pairs (respectively second series with 268 pairs), the method with a single radial distortion has found a superimposition solution for 92% of pairs (respectively 95%) while our method gives a correct result for 96% of the pairs (respectively 98%). Therefore, our method has better results than the superimposition model of Lee *et al.* (2007). To facilitate the comparison of both models, we have performed the model fitting with our method, using the characteristics points extracted by SIFT, while they use the centreline of the vessels and the branch centres.

As a second validation test, we have performed superimposition of images of low/poor quality for 5 patients (with the same size of images) (table 2). The acquisitions conditions were significantly harsher compared to the high quality images and the quality of images was quite heterogeneous in part due to the lack of pupillary dilatation prior to photography. The quality of image superimposition has been checked for the 20 pairs of images and in each case, there was no noticeable difference in the superimposition, with our method, even on the border of the images. The single radial distortion method (similar to the one of Lee et al.) gives a correct superimposition for 95% of the pairs, “gdbicp” with a quadratic model 40% and “gdbicp” with a homography and two radial distortions 25%. In this second validation, our method gives better result than the others.

Finally, in order to show the influence on the superimposition of the colour stabilisation on public health databases, we have applied the “gdbicp” method with a quadratic model on images with colour stabilisation (i.e. normalised). The results obtained with colour stabilisation are much better than without: in the first series with 271 pairs, 92% vs. 48%, in the second series with 268 pairs, 90% vs. 33% and in the third series with 20 low quality pairs of 100 % vs. 25%. This shows the importance of contrast stabilisation for image superimposition in public health databases. Importantly, even with using normalised images with “gdbicp”, our method remains better.

Table 2. Percentage of correct superimpositions of image pairs of eye fundus in public health databases.

Model	1ST series 271 pairs (63% different cameras) (high quality images)	2nd series 268 pairs (9% different cameras) (high quality images)	3rd series 20 pairs (similar cameras) (low quality images)
Homography and 1 or 2 radial distortions (normalised images) (our method)	96%	98%	100%
Homography and 1 radial distortion (original images) (similar to (Lee *et al.*, 2007))	92%	95%	95%
"gdbicp" quadratic (original images)	77%	61%	40%
"gdbicp" radial (original images)	48%	33%	25%
"gdbicp" quadratic (normalised images)	92%	90%	100%

4 Interpretation & discussion

The validation and the experimental results have shown that our method of superimposition of fundus (retinal) images is efficient to analyse public health databases. In this section, we will examine the reasons of these results and how these reasons answers the superimposition challenges of the introduction part. Then we will summarise the comparison of our method with existing methods. Finally, we will present one limitation and will propose leads to be explored.

Our registration model takes into account two radial distortions (one for each image) and a rotation, a translation and two scalings. The accuracy of the superimposition methods has been validated on a simulated montage. We have shown that the superimposition error is smaller for a model with a homography and one or two radial distortions than with a quadratic model of superimposition such as the one used in "gdbcip" quadratic (Stewart *et al.*, 2003; Yang *et al.*, 2007; Adal *et al.*, 2014; Jian *et al.*, 2010).

In order to assess the efficiency of our method for public health purposes, our superimposition method has been validated on public health databases with high quality images of 69 patients (two series including 271 pairs and 268 different pairs acquired for different years) and 5 patients (20 pairs) with low quality images. In the series with 271 pairs, the majority of the pairs were acquired with two cameras while in the series of 268 pairs the majority were acquired with the same camera. In each case, there is no noticeable difference in the superimposed images if the overlap is sufficient (more than 50% about). The superposition is successful in 96%, 97% and 100 % of the cases respectively. Moreover, the interest of the superimposition is to compare the evolutions in a public health database over many years. This is only useful when the image overlap is large enough. Therefore, our method is well suited for this purpose.

Compared to the state of the art method, "gdbicp" quadratic, which uses a quadratic model, our method better register the eye fundus images - 96 % versus 77% with the first series with 271 pairs, 98% versus 61% with the second series of 268 pairs and 100% versus 40% with the third series of 20 pairs. This confirm the results obtained by Lee *et al.* (2010) who have shown that quadratic models introduces more errors than models based on a homography and a radial distortion. Indeed, with 12 parameters to be fitted (vs 8 parameters for our model – 6 for the homography and 2 for the radial distortions) quadratic models have more degree of freedom that is a source of additional errors. Therefore, our model goes further than the previous approaches with quadratic models (Can *et al.*, 2002; Lee *et al.*, 2010; Jian *et al.*, 2010; Adal *et al.*, 2014).

In addition, for all the three series of images, our method improves the results obtained with the model of Lee *et al.* (2007), made of a homography and a single distortion: in the first series 96% versus 92%, in the second series 98% versus 95% and in the third series 100% versus 95% (Table 2). This result is due to the use of two radial distortions, instead of one, when images of different size are acquired by different cameras. This shows the importance of using an adapted model to the type of cameras used to acquire a pair of images. Therefore, our method goes further than the previous one that were not taking into account the change of camera

The obtained results in the validation montage and in public health databases demonstrate that our model better corrects the errors coming from the different position of the patient during image acquisition, the change in the camera employed (resolution and optical lenses) and the projection of a 3D scene onto a plane and the variability of colour between images.

In addition, we have introduced a colour stabilisation step, which is useful when images have strong contrast variations. The comparisons of the results obtained with "gdbcip" quadratic with and without the colour stabilisation show the importance of using normalised images for colour in public health databases (Table 2). In our method, the registration model is estimated based on images after colour stabilisation (Noyel *et al.*, 2015). Then, the SIFT points are extracted and matched on the normalised images for colour contrast. Experiments have shown that using normalised images gives more robust results when extracting and matching the SIFT points, even if these points are known to be robust to colour variations (Lowe, 2004; Vedaldi and Fulkerson, 2008).

Let us summarise the contribution of our method.

The novelty of our method is to deal with eye fundus images acquired with different cameras (i.e. different resolutions and different lenses), and also with the same camera, and to be robust to strong colour variations between the images (Noyel *et al.*, 2015). The image warping is performed using a *division model* which is invertible which makes it fast to compute. We have written the linear equations to estimate the parameters of the model in a fast way. Then a refined estimate is computed by non-linear algorithms. Finally, our method has been designed to analyse large cohorts of patients' eye fundus images (i.e. examinations across time).

Our method has been tested for images acquired with a field of view (FOV) of 45° which are used for DR screening by single-field fundus photography (Williams *et al.*, 2004). In addition, our method could be useful for automatic detection of referral patients due to Diabetic Retinopathy (Fleming *et al.*, 2010; Decencière *et al.*, 2013; Abramoff *et al.*, 2013; Quellec *et al.*, 2016).

Despite the good results obtained, our method does have a limitation. For images with a smaller overlap (e.g. 30% of the surface of the mosaic image), the superimposition may present small differences on the external part. To address this issue, we have developed another algorithm using in addition to the matched points, the distance between the vessels. It will be presented in a future paper.

5 Conclusions

We have therefore successfully achieved a new method to superimpose eye fundus images coming from large public health databases. In addition to the previously existing methods, ours has been designed to deal with changes in terms of camera, lens, image resolution and colour between two exams of the same patient.

The method presented consists of fitting a registration model composed of a homography and two radial distortions on salient points extracted in images after colour stabilisation. The method is easy to use and does not require to extract intrinsic characteristics of the image such as the vessels or their branch points.

All the stages of the method have been designed to be robust and fast on heterogeneous databases. In particular, the equation of linear estimators of the parameters have been provided and an invertible model has been used to warp the images.

Our method has been validated on a montage and on public health databases of eye fundus exams of patients. Some patients had high quality images while other had images of lower quality due to differences in the conditions of acquisition. However, the results show that there is no noticeable difference between the images from two examinations with the eye in the same position (nasal or macular). The superimposition is correct in more than 96% of the cases.

In order to better superimpose the images with a smaller overlap (e.g. 30% of the surface of the mosaic image), we have developed an algorithm which integrates a matching between the vessels in addition to the matched points. This algorithm superimposing eye fundus images acquired in two different positions with a small overlap and with different cameras, will be presented in a future paper.

Conflict of interest

No conflict of interest.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

6 References

Abramoff M, Niemeijer M, Lee S and Reinhardt J 2012 Optimal registration of multiple deformed images using a physical model of the imaging distortion. In: *Google Patents,* (USA: University of Iowa Research Foundation)

Abramoff M D, Folk J C, Han D P, Walker J D, Williams D F, Russell S R, Massin P, Cochener B, Gain P, Tang L, Lamard M, Moga D C, Quellec G and Niemeijer M 2013 Automated analysis of retinal images for detection of referable diabetic retinopathy *JAMA Ophthalmol* **131** 351-7

Adal K M, Ensing R M, Couvert R, van Etten P, Martinez J P, Vermeer K A and van Vliet L J 2014 *Biomedical image registration,* ed Springer (London

Bonnans J F, Gilbert J C, Lemaréchal C and Sagastizabal C A 2006 *Numerical optimization. Theoretical and practical aspects* (Berlin Heidelberg

Can A, Stewart C V, Roysam B and Tanenbaum H L 2002 A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24** 347-64

Cattin P C, Bay H, Van Gool L and Székely G 2006 *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006,* ed R Larsen, *et al.* (Zurich, Zwitserland: Springer Berlin Heidelberg) pp 185-92

Chanwimaluang T, Fan G and Fransen S R 2006 Hybrid retinal image registration *IEEE Trans Inf Technol Biomed* **10** 129-42

Cideciyan A V 1995 Registration of ocular fundus images: an algorithm using cross-correlation of triple invariant image descriptors *IEEE Engineering in Medicine and Biology Magazine* **14** 52-8

Decencière E, Cazuguel G, Zhang X, Thibault G, Klein J C, Meyer F, Marcotegui B, Quellec G, Lamard M, Danno R, Elie D, Massin P, Viktor Z, Erginay A, Laÿ B and Chabouis A 2013 TeleOphta: Machine learning and image processing methods for teleophthalmology *Irbm* **34** 196-203

Fang B and Tang Y Y 2006 Elastic registration for retinal images based on reconstructed vascular trees *IEEE Trans Biomed Eng* **53** 1183-7

Fitzgibbon A W 2001 Simultaneous linear estimation of multiple view geometry and lens distortion *Computer Vision and Pattern Recognition. Proceedings of the 2001 IEEE Computer Society Conference* **1** I-125-I-32

Fleming A D, Goatman K A, Philip S, Prescott G J, Sharp P F and Olson J A 2010 Automated grading for diabetic retinopathy: a large-scale audit using arbitration by clinical experts *Br J Ophthalmol* **94** 1606-10

Ghassabi Z, Shanbehzadeh J, Sedaghat A and Fatemizadeh E 2013 An efficient approach for robust multimodal retinal image registration based on UR-SIFT features and PIIFD descriptors *EURASIP Journal on Image and Video Processing* **2013** 25

Harding S, Greenwood R, Aldington S, Gibson J, Owens D, Taylor R, Kohner E, Scanlon P and Leese G 2003 Grading and disease management in national screening for diabetic retinopathy in England and Wales *Diabet Med* **20** 965-71

Hartley R and Zisserman A 2004 *Multiple view geometry in computor vision*: Cambridge University Press)

Hernandez-Matas C, Zabulis X and Argyros A A *IEEE 38th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC),2016),* vol. Series*)*

International Diabetes Federation and The Fred Hollows Foundation 2015 Diabetes Eye Health. A guide for health professionals. (Brussels, Belgium: International Diabetes Federation and The Fred Hollows Foundation)

Jian C, Jie T, Lee N, Jian Z, Smith R T and Laine A F 2010 A Partial Intensity Invariant Feature Descriptor for Multimodal Retinal Image Registration *IEEE Transactions on Biomedical Engineering* **57** 1707-18

Lee S, Abràmoff M D and Reinhardt J M 2007 Feature-based pairwise retinal image registration by radial distortion correction *Proc of SPIE* **6512** 651220--10

Lee S, Abràmoff M D and Reinhardt J M 2008 Retinal image mosaicing using the radial distortion correction model *Proc. of SPIE* **6914** 691435--9

Lee S, Reinhardt J M, Cattin P C and Abramoff M D 2010 Objective and expert-independent validation of retinal image registration algorithms by a projective imaging distortion model *Med Image Anal* **14** 539-49

Lloret D, Serrat J, Lopez A M, Soler A and Villaneuva J J *15th International Conference on Pattern Recognition.,2000),* vol. Series 3*)*: IEEE) pp 203-6

Lowe D G 2004 Distinctive Image Features from Scale-Invariant Keypoints

Matsopoulos G K, Mouravliansky N A, Delibasis K K and Nikita K S 1999 Automatic Retinal Image Registration Scheme Using Global Optimization Techniques *IEEE Transactions on Information Technology in Biomedicine* **3** 47-60

More J J 1977 *Numerical Analysis,* ed G A Watson (United Kingdom: Springer Verlag) pp 105-16

Moré J J 1983 *Mathematical Programming The State of the Art,* ed A Bachem, *et al.*: Springer Berlin Heidelberg) pp 258-87

Noyel G, Jourlin M, Thomas R, Bhakta G, Crowder A, Owens D and Boyle P 2015 Contrast enhancement of eye fundus images. In: *IDF 2015*, (Vancouver

Park J, Byun S-C and Lee B-U 2009 Lens distortion correction using ideal image coordinates *IEEE Transactions on Consumer Electronics* **55** 987-91

Quellec G, Lamard M, Erginay A, Chabouis A, Massin P, Cochener B and Cazuguel G 2016 Automatic detection of referral patients due to retinal pathologies through data mining *Med Image Anal* **29** 47-64

Ritter N, Owens R, Cooper J, Eikelboom R H and van Saarloos P P 1999 Registration of Stereo and Temporal Images of the Retina *IEEE Trans Med Imaging* **18** 404-18

Ryan N, Heneghan C and de Chazal P 2004 Registration of digital retinal images using landmark correspondence by expectation maximization *Image and Vision Computing* **22** 883-98

Scanlon P H, Wilkinson C P, Aldington S J and Matthews D R 2009 *A practical manual of diabetic retinopathy management*: Wiley Online Library)

Stewart C V, Tsai C L and Roysam B 2003 The dual-bootstrap iterative closest point algorithm with application to retinal image registration *IEEE Trans Med Imaging* **22** 1379-94

Thomas R L, Dunstan F, Luzio S D, Roy Chowdury S, Hale S L, North R V, Gibbins R L and Owens D R 2012 Incidence of diabetic retinopathy in people with type 2 diabetes mellitus attending the Diabetic Retinopathy Screening Service for Wales: retrospective analysis *BMJ* **344** e874

Vedaldi A and Fulkerson B 2008 VLFEat: An open and portable library of computor vision algorithms. http://www.vlfeat.org/)

Walter T 2003 Application de la morphologie mathématique au diagnostic de la rétinopathie diabétique à partir d'images couleur. (Paris, France: Ecole Nationale Supérieure des Mines de Paris)

Walter T and Klein J-C 2005 *Handbook of Biomedical Image Analysis,* ed J S Suri, *et al.* pp 315-68

Williams G A, Scott I U, Haller J A, Maguire A M, Marcus D and McDonald H R 2004 Single-field fundus photography for diabetic retinopathy screening: a report by the American Academy of Ophthalmology *Ophthalmology* **111** 1055-62

Wolberg G 1990 *Digital image warping*: Wiley-IEEE Computer Society Press)

Wonpil Y 2003 An embedded camera lens distortion correction method for mobile computing applications *IEEE Transactions on Consumer Electronics* **49** 894-901

Yang G, Stewart C V, Sofka M and Tsai C L 2007 Registration of challenging image pairs: initialization, estimation, and decision *IEEE Trans Pattern Anal Mach Intell* **29** 1973-89

You X, Fang B, He Z and Tang Y Y 2005 *Pattern Recognition and Image Analysis,* ed J S Marques, *et al.* (China: Springer Berlin Heidelberg) pp 259-67

Yuping L and Medioni G *IEEE Conference onComputer Vision and Pattern Recognition,2008),* vol. Series): IEEE) pp 1-8

Zana F and Klein J C 1999a A Multimodal Registration Algorithm of Eye Fundus Images Using Vessels Detection and Hough Transform *IEEE Trans Med Imaging* **18** 419-28

Zana F and Klein J C 1999b A registration algorithm of eye fundus images using a Bayesian hough transform. In: *Seventh International Conference on Imaging Processing and its Applications*: IEE) pp 479-83