Retinal Image Registration through 3D Eye Modelling and Pose Estimation

CARLOS HERNANDEZ MATAS

Institute of Computer Science – FORTH
Computer Science Department – University of Crete

PHD DEFENSE

JUNE 26, 2017

HERAKLION, CRETE

Structure

I – Introduction

II – Technical contributions

III – Discussion

INSTITUTE OF COMPUTER SCIENCE

Part I

Introduction

I – Introduction

What is image registration?

Why is it important?

What type of images?

Which are its challenges?

Related work

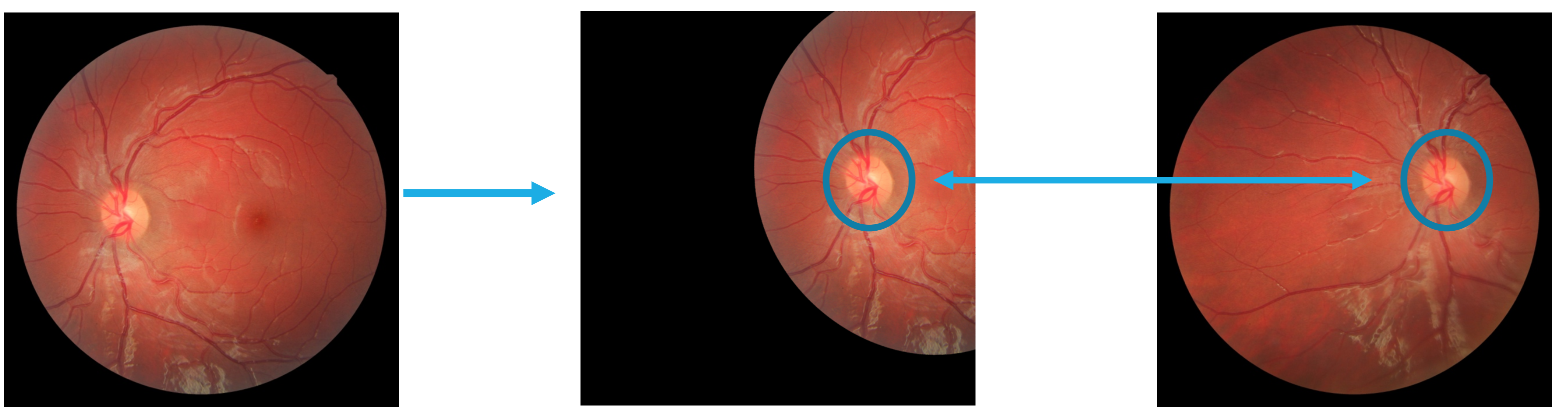
What is image registration?

Image registration involves **a pair of images**: reference and test.

It consists on transforming the **test** image so both images are aligned in the **reference** image **frame.**

This is done by utilizing the **common information** in both images**.**

Retinal image registration consists on the registration of **retinal images**.



Test image | Transformed test image | Reference image

I – Introduction

What is image registration?

Why is it important?

What type of images?

Which are its challenges?

Related work

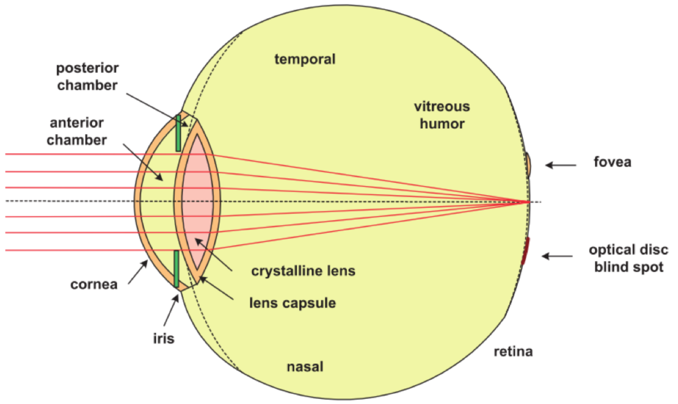
Why is it important?

Analyzing **small vessels** promotes **diagnosis** and **disease monitoring**

The **retina** allows to **directly** observe the microvasculature of the eye

Comparative **analysis** is **facilitated** by retinal image **registration**

Additionally, registration **supports** a range of **applications**



[Gross 2008]

INSTITUTE OF COMPUTER SCIENCE

I – Introduction

What is image registration?

Why is it important?

What type of images?

Which are its challenges?

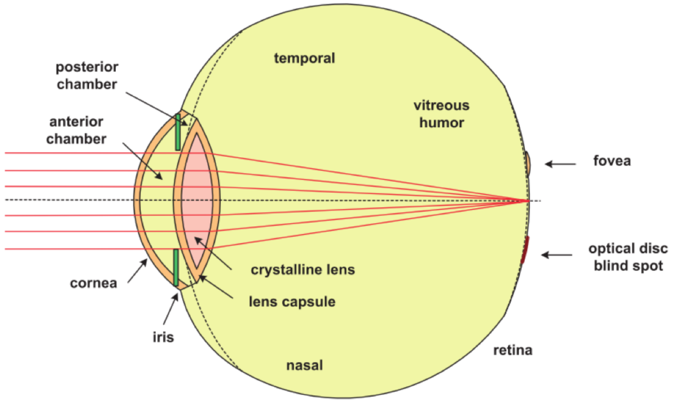
Related work

What type of images?

The eye's general **shape** is quasi-**spherical**

Doctors use **2D images** of the retina

These images are acquired using a fundus camera



[Gross 2008]

I – Introduction

What is image registration?

Why is it important?

What type of images?

Which are its challenges?

Related work

Which are its challenges?

- **Different viewpoints**
  - Image different areas of the retina.
  - **Small overlapping (little information)**.
- **Projection distortion**
  - Projecting the quasi-**spherical retina** onto a **flat image** introduces **radial distortion**.
- **Anatomical changes due to retinopathy**
  - The **information** about the same eye area in two different images **may be different**.
- **High accuracy**
  - **Required** to facilitate **analysis** by **doctors** and clinicians.

I – Introduction

What is image registration?

Why is it important?

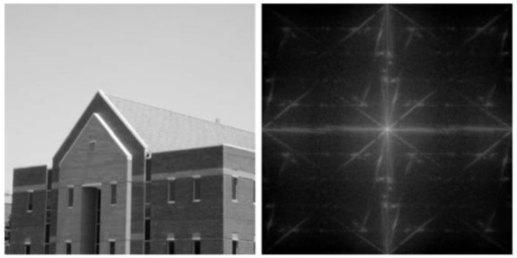
What type of images?

Which are its challenges?

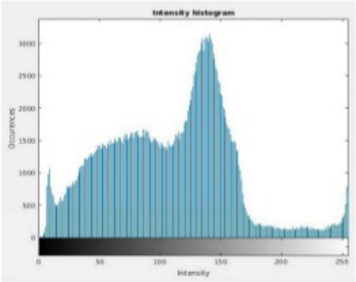
Related work

Registration methods classification

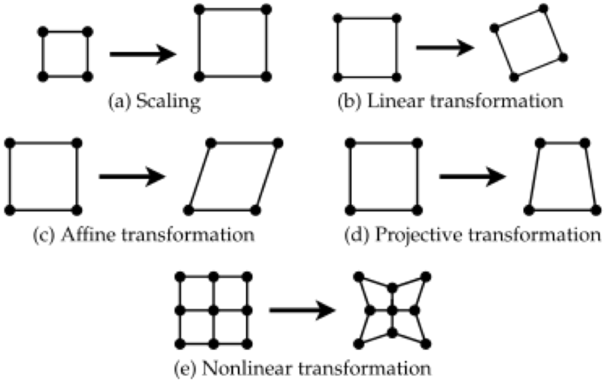
- **Spatial** or frequency domain



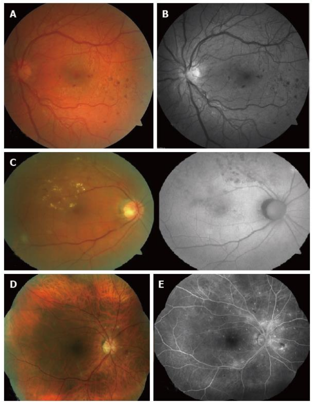
- Global or **local**



- Transformation model used



- **Intra-** or. cross-modal



Related work I

	Freq	Spa	Local	Global	Intra	Cross	Linear	Affine	Projective	Quadratic
Peli (1987)		X	X		X		X			
Cideciyan (1992)	X			X	X		X			
Matsopoulos (1999)		X	X		X	X		X	X	
Laliberte (2003)		X	X		X	X	X	X		X
Stewart (2003)		X	X		X	X	X	X		X
Ryan (2004)		X	X		X	X		X		X
Matsopoulos (2004)		X	X		X	X		X		
Chanwimaluang (2006)		X	X		X			X		X
Choe (2006)		X	X		X	X		X		
Yang (2007)		X	X		X	X	X	X		X
Lin (2008)		X	X		X	X			X	
Chaudhry (2008)		X	X		X			X		
Tsai (2010)		X	X		X	X	X	X		X
Chen (2010)		X	X		X	X	X	X		X
Deng (2010)		X	X		X	X				X

Related work II

	Freq	Spa	Local	Global	Intra	Cross	Linear	Affine	Projective	Quadratic
Perez-Rovira (2011)		X	X		X					X
Zheng (2011)		X	X		X		X	X	X	X
Troglio (2011)		X	X		X			X		
Gharabaghi (2012)		X	X		X			X		
Ghassabi (2013)		X	X		X	X				X
Reel (2013)		X		X	X	X		X		
Legg (2013)		X		X	X	X	X			
Bathina (2013)		X	X		X	X		X	X	X
Chen (2014)		X	X		X		X	X		
Adal (2014)		X		X	X		X	X		X
Lee (2015)		X	X		X	X		X		
Wang (2015)		X	X		X	X	X	X		X
Ghassabi (2015)		X	X		X		X	X		X
Liu (2016)		X	X		X	X		X		
Saha (2016)		X	X		X					X

INSTITUTE OF COMPUTER SCIENCE

Part II

Technical Contributions

II – Contributions

Dataset

Registration framework and experiments

Registration applications

Retinal image datasets

Plenty of retinal image **datasets** for **diverse purposes**

- **Segmentation**: CHASEDB1, DRIONS-DB, Drishti-GS, DRIVE, HRF, MESSIDOR, ONHSD and REVIEW
- **Diagnosis**: DIARETDB0, DIARETDB1, e-ophtha, INSPIRE-AVR, ROC, STARE and VICAVR
- **User authentication**: VARIA
- **Retinal image registration**: RODREP

Hard to find datasets suitable for retinal image registration:

- Most datasets have **no image pairs for the same eye**
- **No** datasets with **ground truth** to evaluate registration

	Images	FOV	Resolution	Pairs	Large overlap	Small overlap	Anatomical differences	Ground truth
e-ophtha	463	45^o	2544x1696	144	Yes	Yes	No	No
RODREP	1120	45^o	2000x1312	1400	Yes	Yes	No	No
VARIA	233	20^o	768x584	154	Yes	No	No	No

FIRE dataset

Fundus Image Registration Dataset:

- **129 real** retinal images acquired with a fundus camera
- **134** image **pairs**
- **3 categories**
- Some image pairs with **anatomical differences**
- **Evaluation** via manually annotated **control points**

	Images	FOV	Resolution	Pairs	Large overlap	Small overlap	Anatomical differences	Ground truth
e-ophtha	463	45^o	2544x1696	144	Yes	Yes	No	No
RODREP	1120	45^o	2000x1312	1400	Yes	Yes	No	No
VARIA	233	20^o	768x584	154	Yes	No	No	No
FIRE	129	45^o	2912x2912	134	Yes	Yes	Yes	Yes

Publicly available at http://www.ics.forth.gr/cvrl/fire/

S category

S is for "**Similar**"

71 image pairs

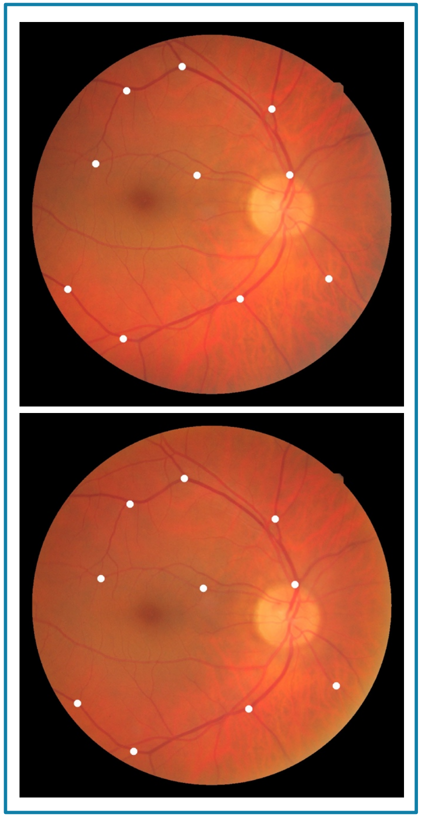
Significant overlap

No anatomic changes

A lot of potential information to use for registration

Least challenging category for registration

Image pairs can be used for super resolution



P category

P is for “**Pose difference**”

49 image pairs

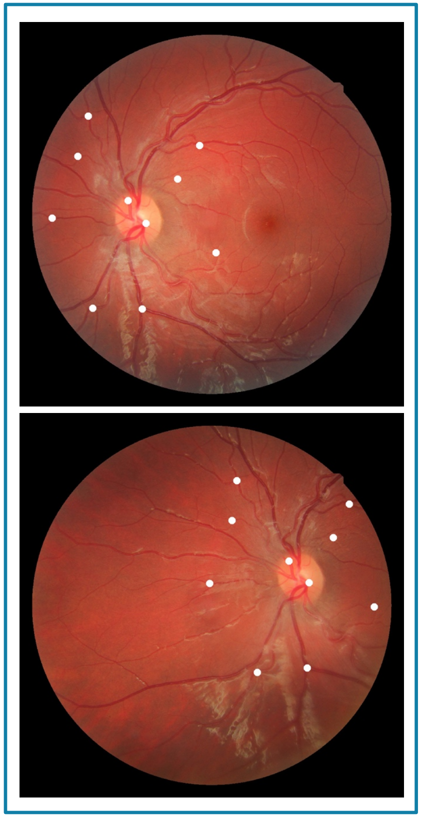
Minor overlap

No anatomic changes

Little potential information to use for registration

More challenging than previous category

Image pairs can be used for creating mosaics of the retina



A category

A is for "**Anatomic difference**"

14 image pairs

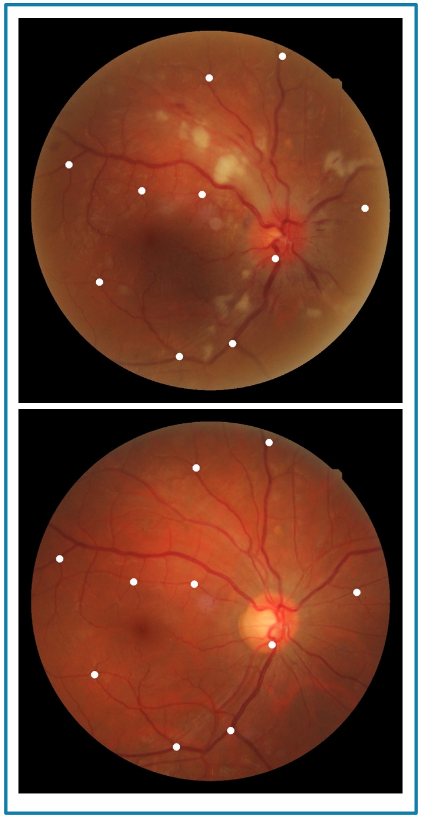
Significant overlap

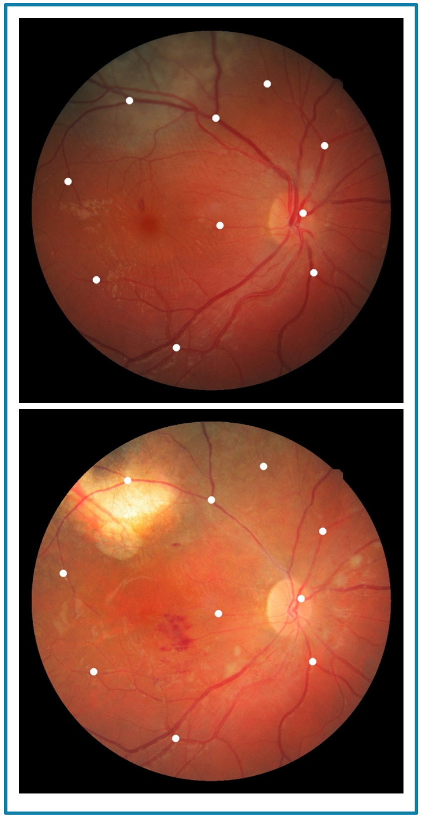
Anatomic changes

Corresponding points may look different due to retinopathy

More challenging than S category

Image pairs can be used for longitudinal studies





II – Contributions

Datasets

Registration framework and experiments

Registration applications

Framework

Geometrical approach

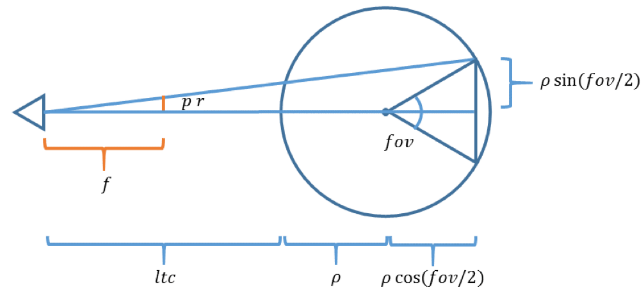
Simultaneously estimate relative **camera pose** and **eye shape** and **orientation**

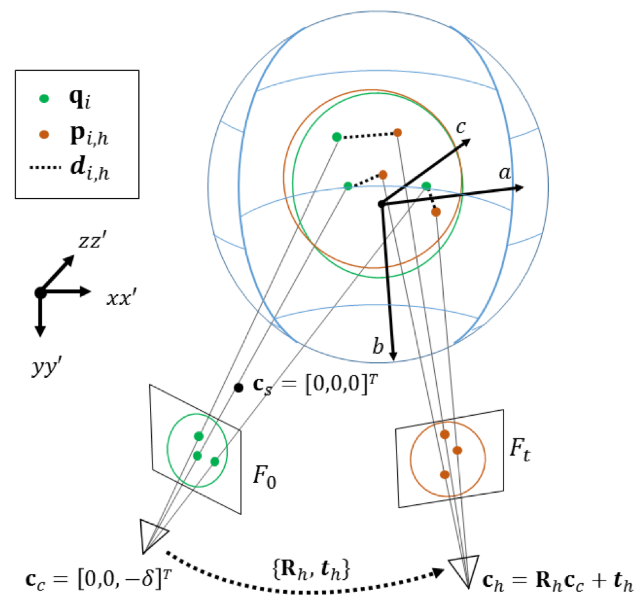
Project points from hypothetical cameras to an eye **model**

Distance of points in model should be **0**

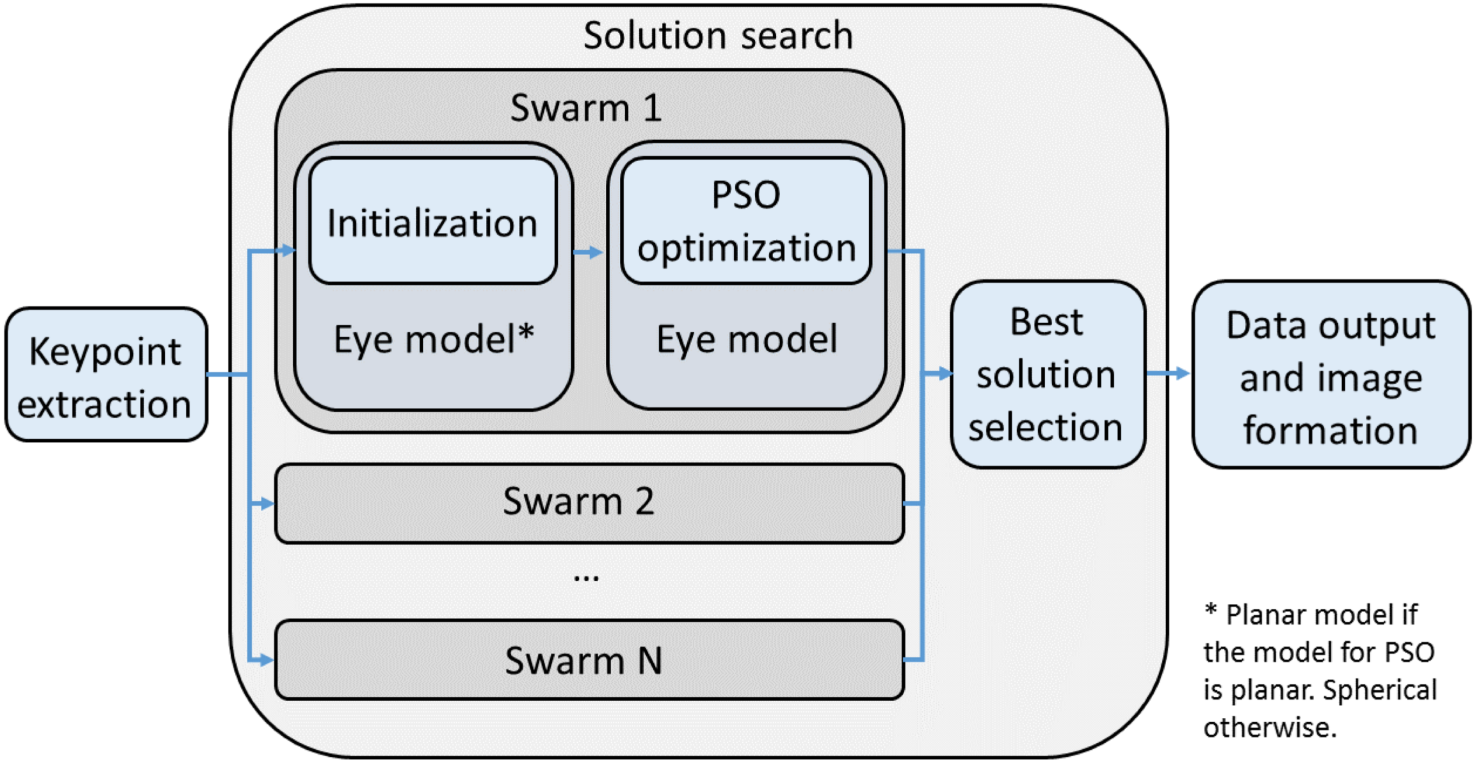
Solution is represented via 12 parameters, divided in 4 groups:

- **R:** $[r_x, r_y, r_z]$
- **t:** $[t_x, t_y, t_z]$
- **A:** $[a, b, c]$
- **Q:** $[r_a, r_b, r_c]$



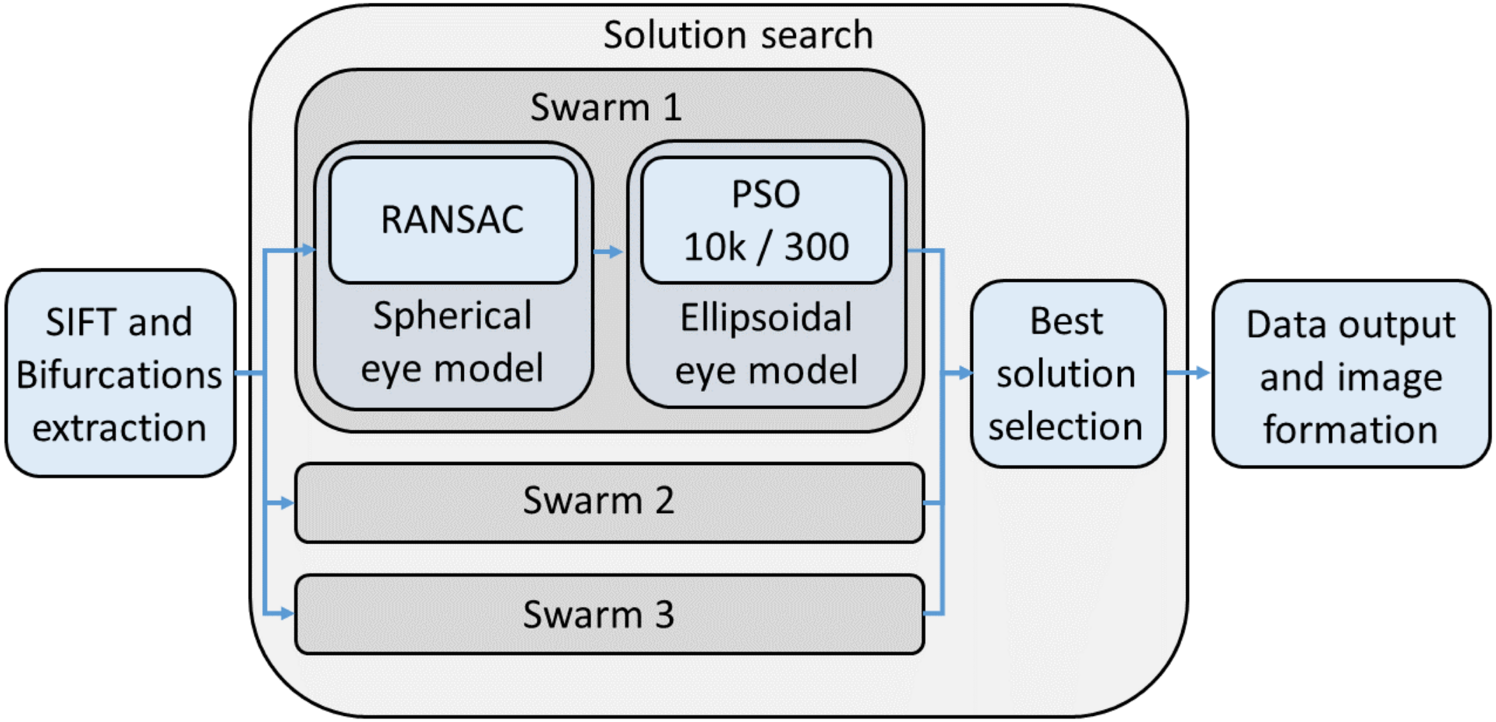


Framework



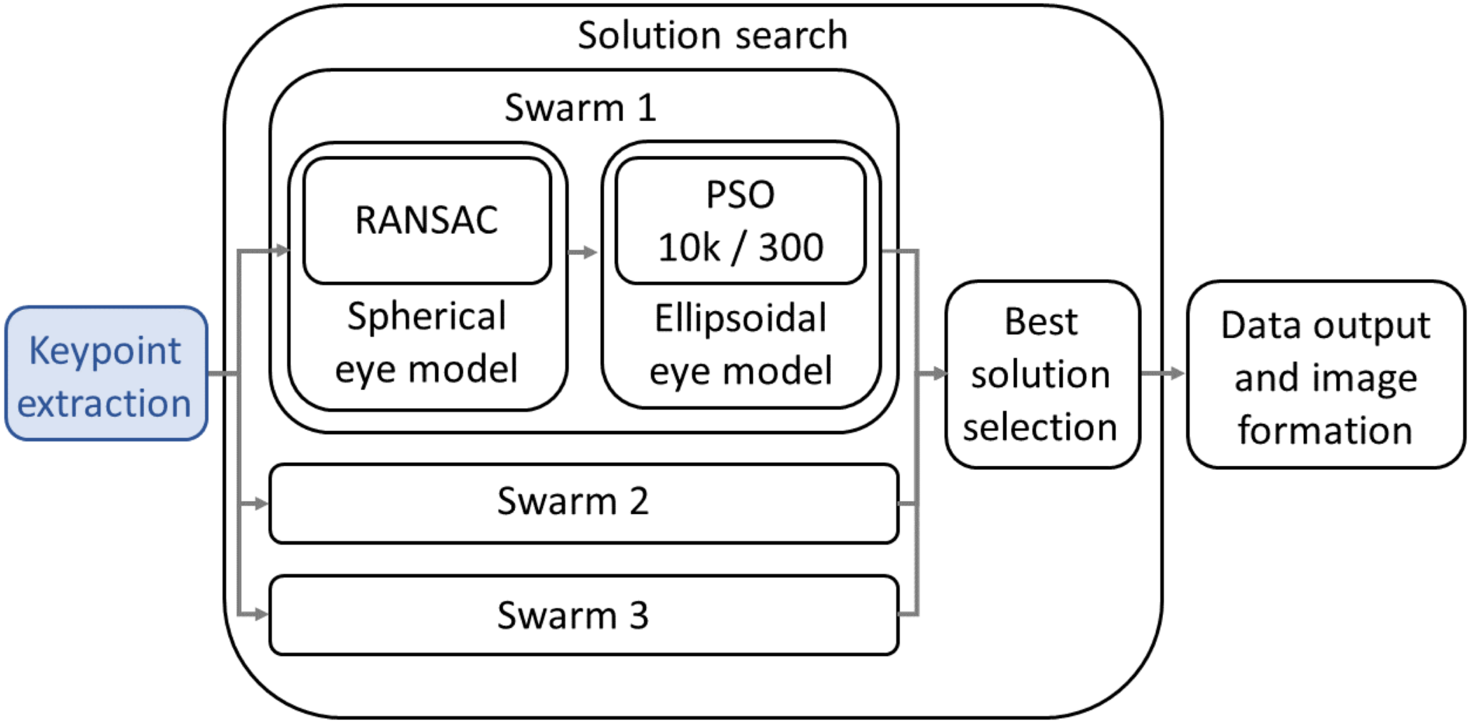
* Planar model if the model for PSO is planar. Spherical otherwise.

Method



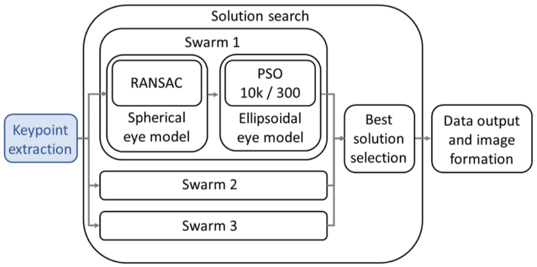
Solution search
Swarm 1
RANSAC
Spherical eye model
PSO
10k / 300
Ellipsoidal eye model
SIFT and Bifurcations extraction
Swarm 2
Swarm 3
Best solution selection
Data output and image formation

Keypoints



Solution search
Swarm 1
RANSAC
Spherical eye model
PSO
10k / 300
Ellipsoidal eye model
Keypoint extraction
Best solution selection
Data output and image formation
Swarm 2
Swarm 3

Keypoints

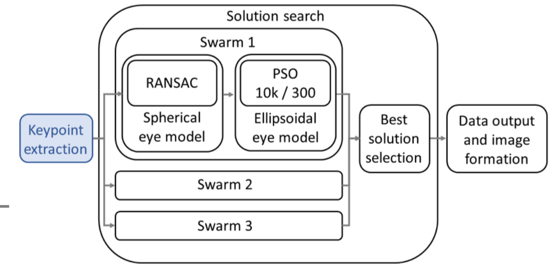


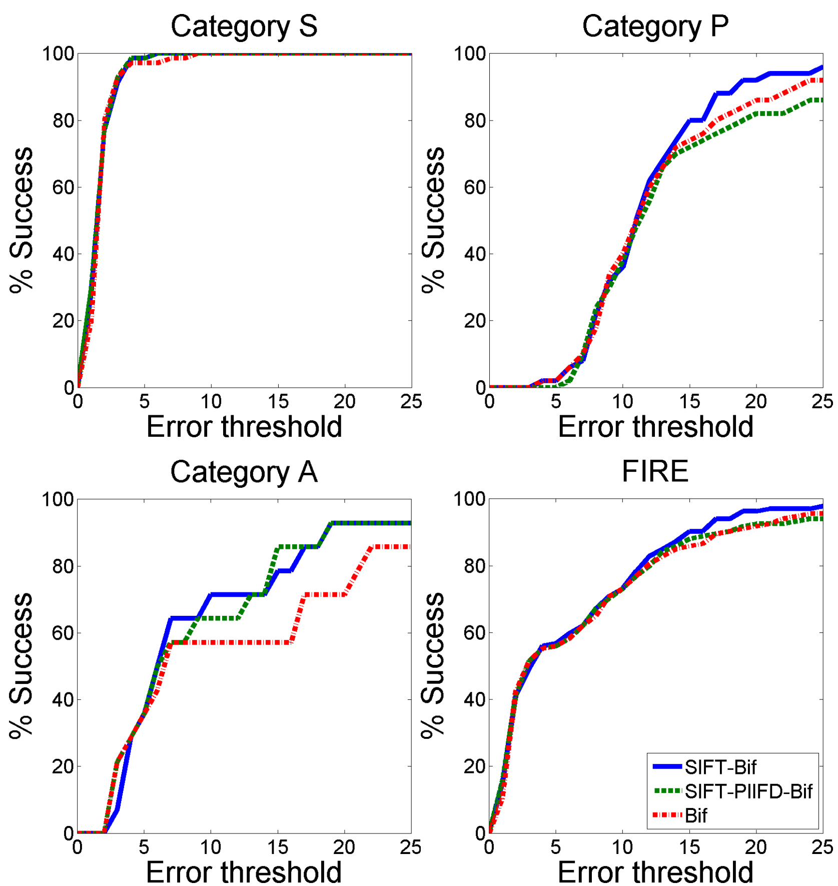
Keypoints are used to identify common points in the pair of images

Four types of keypoints and their combinations are studied

- **SIFT** [Lowe 2004]: milestone method in extracting characteristic points in images.
- **SURF** [Bay 2008]: is another widely used method to detect and represent keypoints.
- **Harris-PIIFD** [Chen 2010]: Partial Intensity Invariant Feature Detector on Harris corners. Developed specifically for retinal image registration.
- **Vessel bifurcations**: Bifurcations on the vessel tree are extracted.

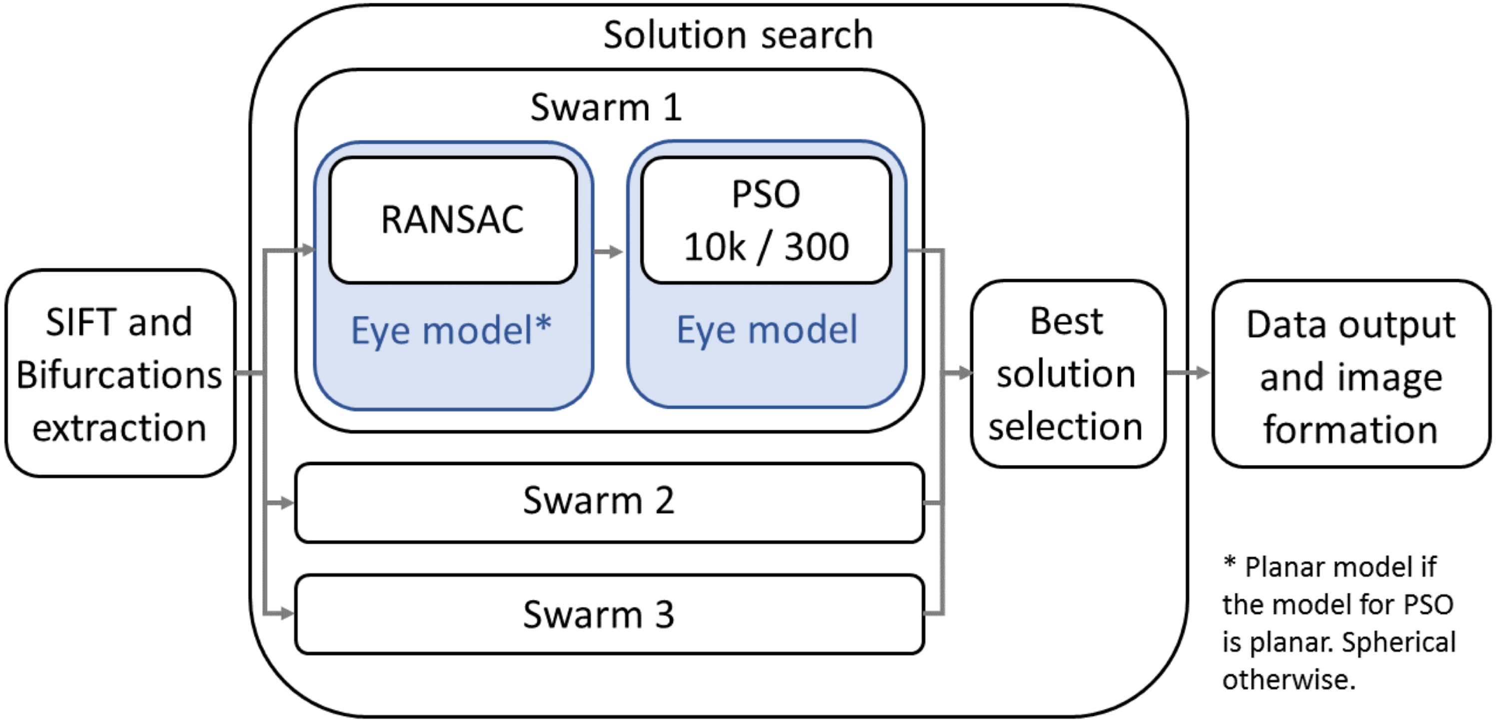
Keypoints



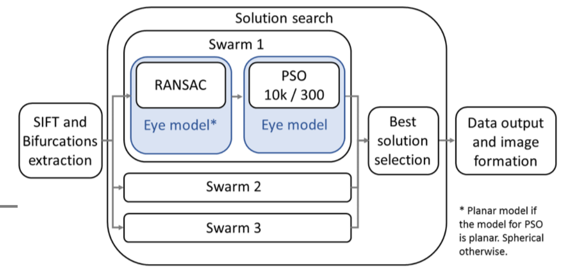


SIFT	SURF	PIIFD	Bif	S	P	A	FIRE
X				0.945	0.443	0.577	0.721
	X			0.947	0.348	0.466	0.675
		X		0.846	0.134	0.429	0.538
			X	0.953	0.516	0.563	0.751
X	X			0.953	0.423	0.526	0.712
X		X		0.951	0.396	0.503	0.699
X			**X**	0.958	**0.541**	**0.660**	**0.773**
	X	X		0.940	0.264	0.426	0.636
	X		X	0.956	0.404	0.489	0.703
		X	X	0.954	0.472	0.563	0.736
X	X	X		0.952	0.333	0.491	0.674
X	X		X	0.956	0.435	0.480	0.713
X		X	X	**0.959**	0.490	0.657	0.754
	X	X	X	0.954	0.400	0.474	0.699
X	X	X	X	0.956	0.409	0.514	0.707

Eye model

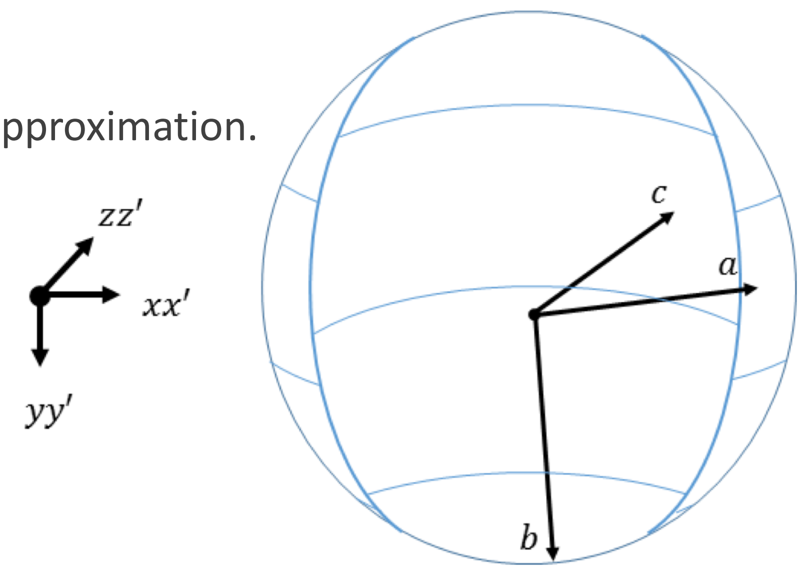

SIFT and Bifurcations extraction
Solution search
Swarm 1
RANSAC
Eye model*
PSO
10k / 300
Eye model
Swarm 2
Swarm 3
Best solution selection
Data output and image formation
* Planar model if the model for PSO is planar. Spherical otherwise.

Eye model

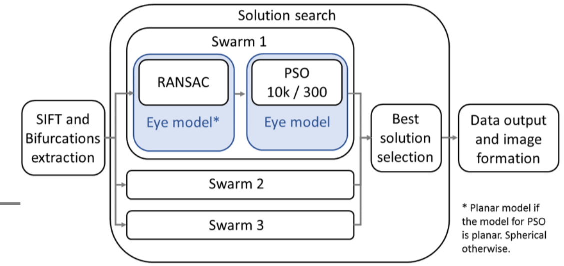


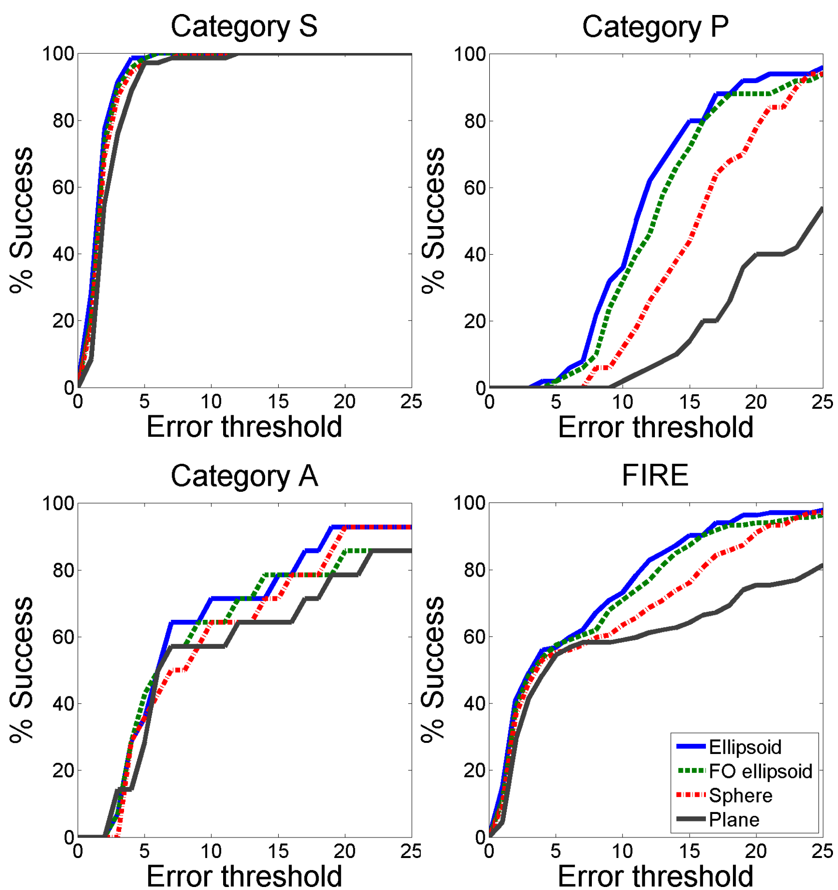
Four eye models are considered:

- **Plane**: Baseline model. Appropriate for images with a narrow Field of View.
  - A = $[a, b, c] = [10000, 10000, 1]$
  - Q = $[r_a, r_b, r_c] = [0, 0, 0]$
  - Only {R, t} are calculated (6 parameters)
- **Sphere** [Navarro 1985]: Simplest full eye approximation.
  - A = $[a, b, c] = [12, 12, 12]$
  - Q = $[r_a, r_b, r_c] = [0, 0, 0]$
  - Only {R, t} are calculated (6 parameters)
- **Fixed orientation ellipsoid**
  - A = $[a, b, c] = [\rho_a, \rho_b, \rho_c]$
  - Q = $[r_a, r_b, r_c] = [0, 0, 0]$
  - {R, t, A} are calculated (9 parameters)
- **Ellipsoid**: The most complex model used.
  - A = $[a, b, c] = [\rho_a, \rho_b, \rho_c]$
  - Q = $[r_a, r_b, r_c] = [\theta_a, \theta_b, \theta_c]$
  - {R, t, A, Q} are calculated (12 parameters)



Eye model

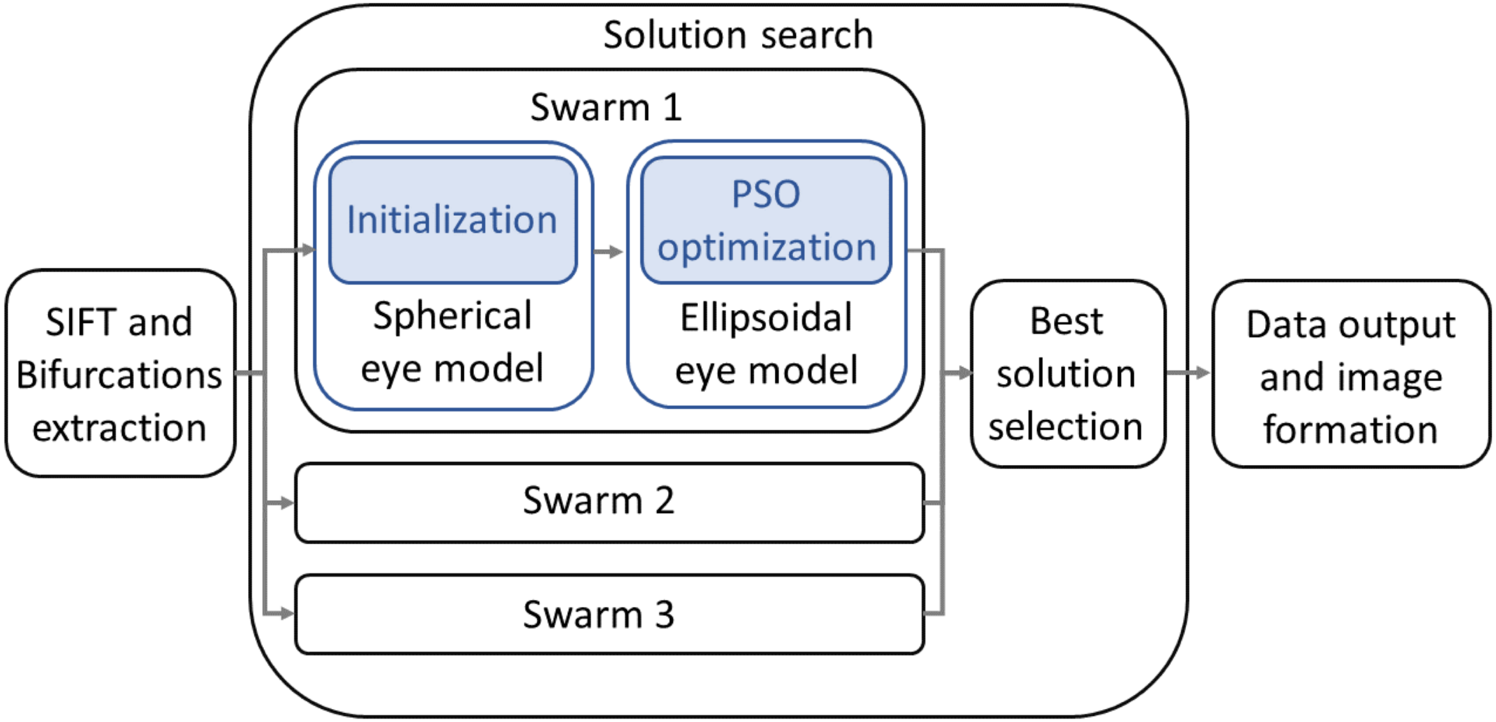




Model	S	P	A	FIRE
Plane	0.926	0.164	0.574	0.608
Sphere	0.45	0.385	0.617	0.703
FO Ellipsoid	0.951	0.498	0.626	0.750
Ellipsoid	**0.958**	**0.541**	**0.660**	**0.773**

The **higher** the order of **complexity**, the **higher** the **capability** to approximate the actual shape of the eye and the more accurate the registration

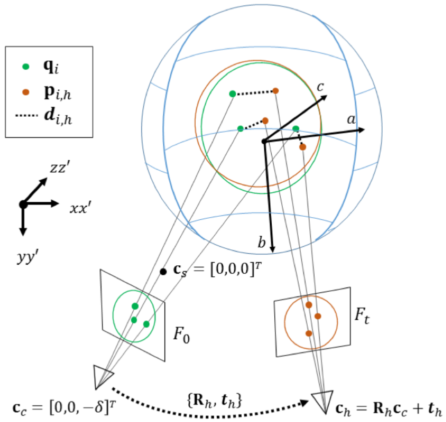
Swarm structure



Swarm structure

Initialization:

- **No initialization**: Test camera initialized to reference camera pose.
- **Random Sample Consensus (RANSAC)** [Fischler 1981]: Estimates the 3D pose of an object given **a set of 2D-3D correspondences** and the camera projection matrix. Spherical eye model

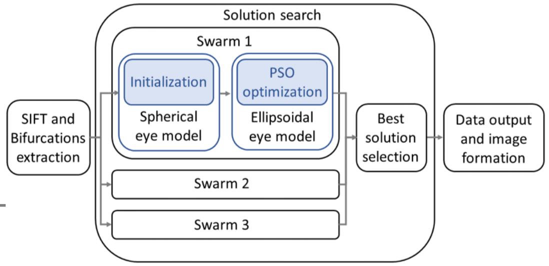


Optimization

- Attempt to **minimize objective function**. Sum of the 80% shortest distances of the corresponding points on the eye model
- We look for the solution on a search space with 12 dimensions. 1 for each solution parameter.
- **Particle Swarm Optimization (PSO)** [Kennedy 1995]:
  - Particles are given random initial position and velocity in the space.
  - Each particle represents a candidate solution (objective function evaluation)
  - Particles evolve through generations.
  - Requires few configuration parameters and no derivatives

$$o(S_h) = \sum_j |q_j - p_{j,h}|$$

Swarm structure



Coarse (C): No initialization. PSO search in a coarse space

Coarse-to-Fine (CF): No initialization. 2 PSO searches. First a coarse search followed by a fine search around the coarse solution.

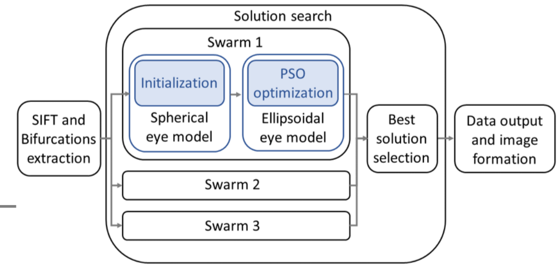
RANSAC (R): Only RANSAC initialization is performed.

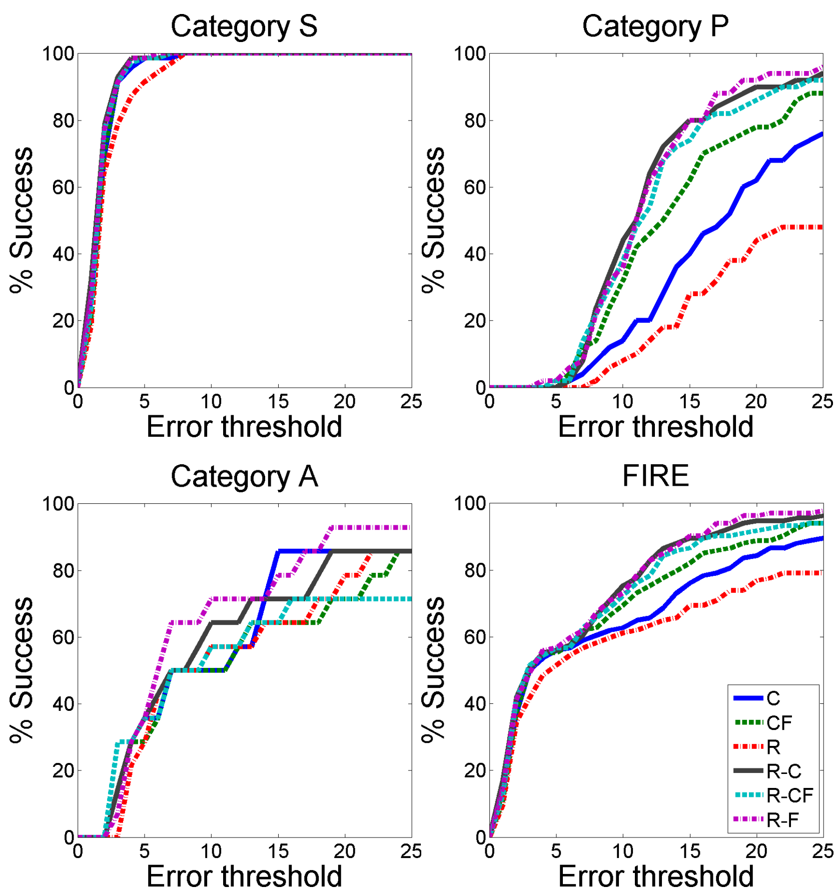
RANSAC-Coarse (R-C): As C, but with RANSAC initialization.

RANSAC-Fine (R-F): As R-C, but with a fine search instead of a coarse one

RANSAC-Coarse-to-Fine (R-CF): As CF, but with RANSAC initialization.

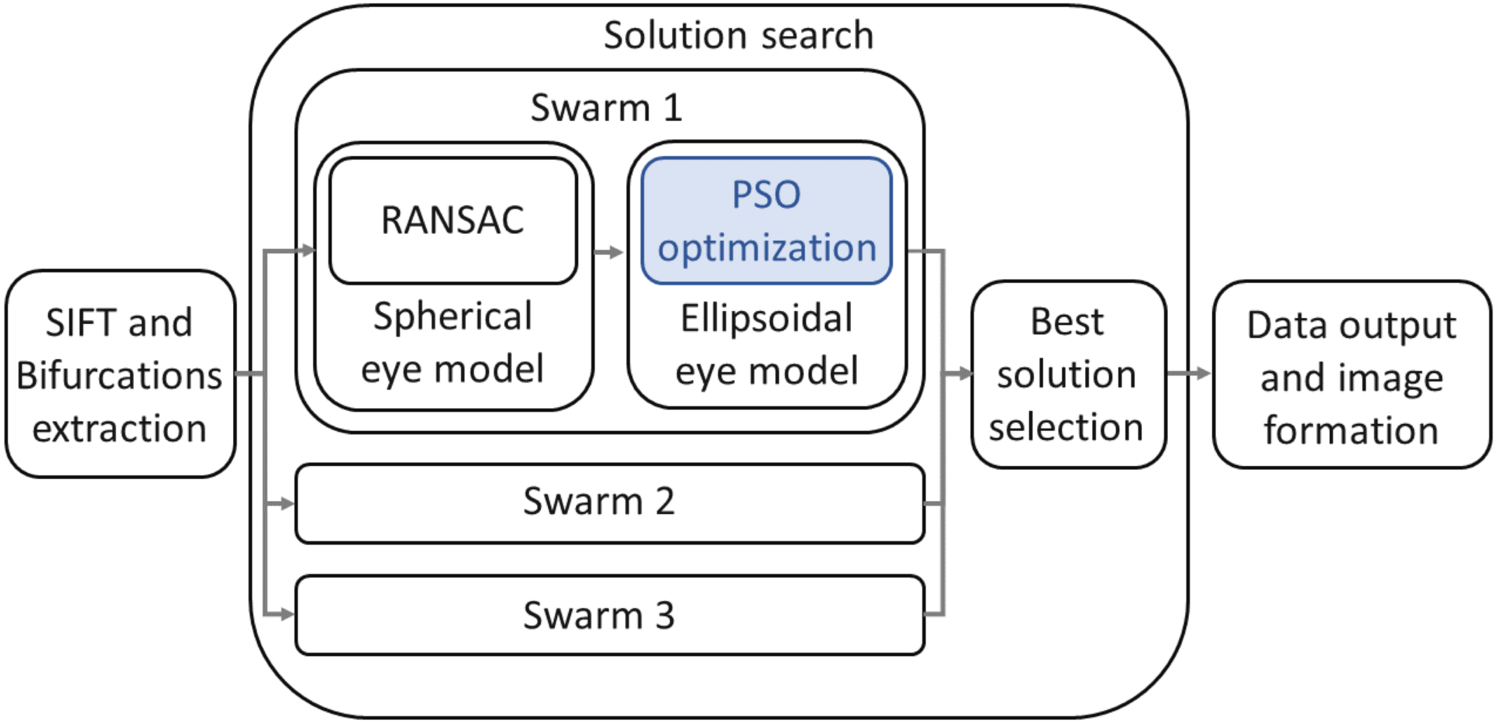
Swarm structure





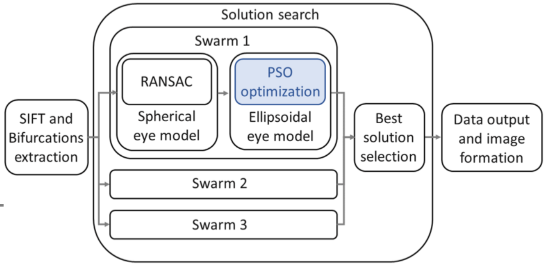
Variant	S	P	A	FIRE
C	0.951	0.324	0.597	0.682
CF	0.952	0.453	0.537	0.724
R	0.933	0.209	0.549	0.624
R-C	**0.960**	0.532	0.603	0.765
R-CF	0.955	0.516	0.543	0.750
R-F	0.958	**0.541**	**0.660**	**0.773**

PSO budget



Solution search
Swarm 1
RANSAC
Spherical eye model
PSO optimization
Ellipsoidal eye model
SIFT and Bifurcations extraction
Best solution selection
Data output and image formation
Swarm 2
Swarm 3

PSO budget



PSO performs **particles times generations** objective function evaluations. This is called ‘**budget**’ of the PSO process.

A **small budget** will terminate the process **prematurely** with a poor pose estimate

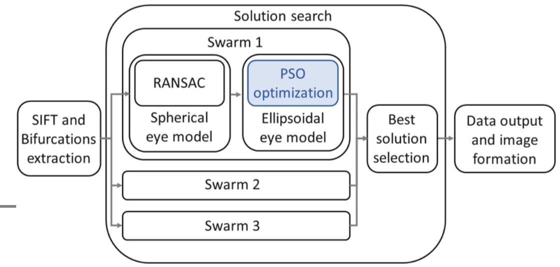
A too **large budget** will lead to extra processing time with **no** noticeable **improvements** in accuracy.

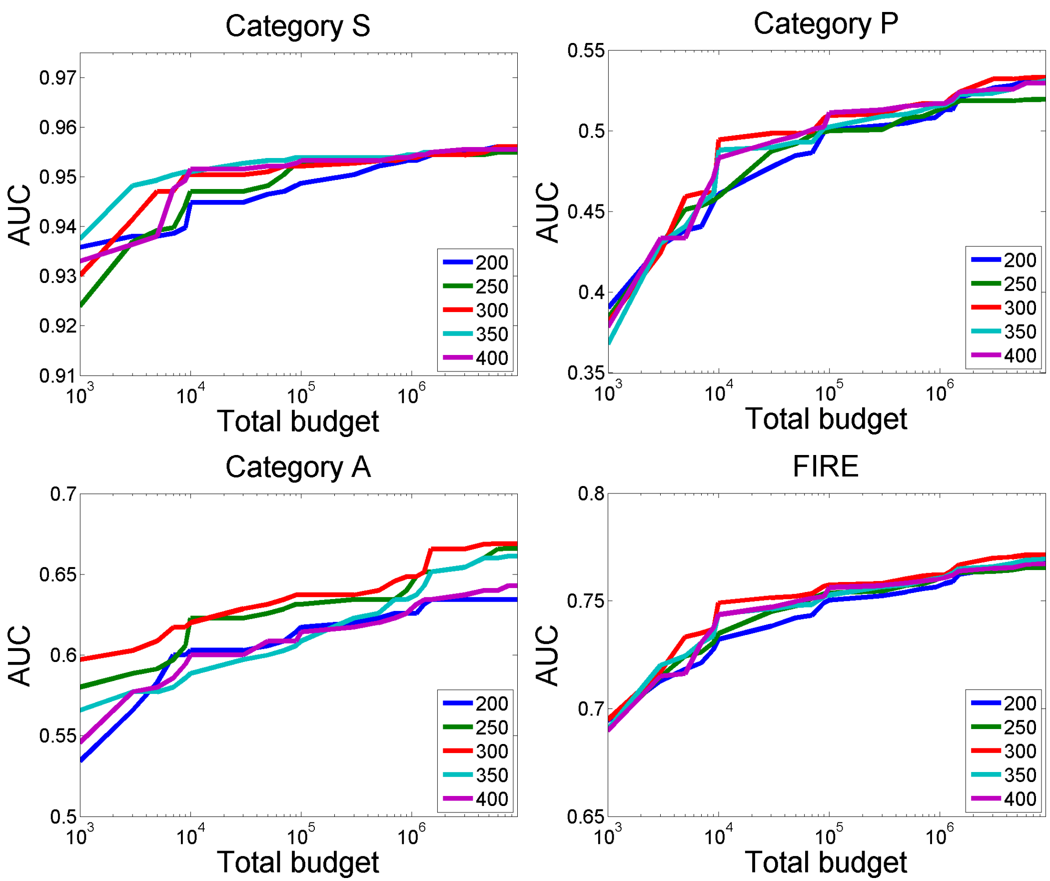
Budget selection offers a **trade-off** between the accuracy and the speed of the method.

The **distribution** of the **budget** across particles and generations is **relevant** to the final **performance** of the method.

Experiment covering budgets from 1000 to 9 million particles distributed across 200, 250, 300, 350 and 400 generations.

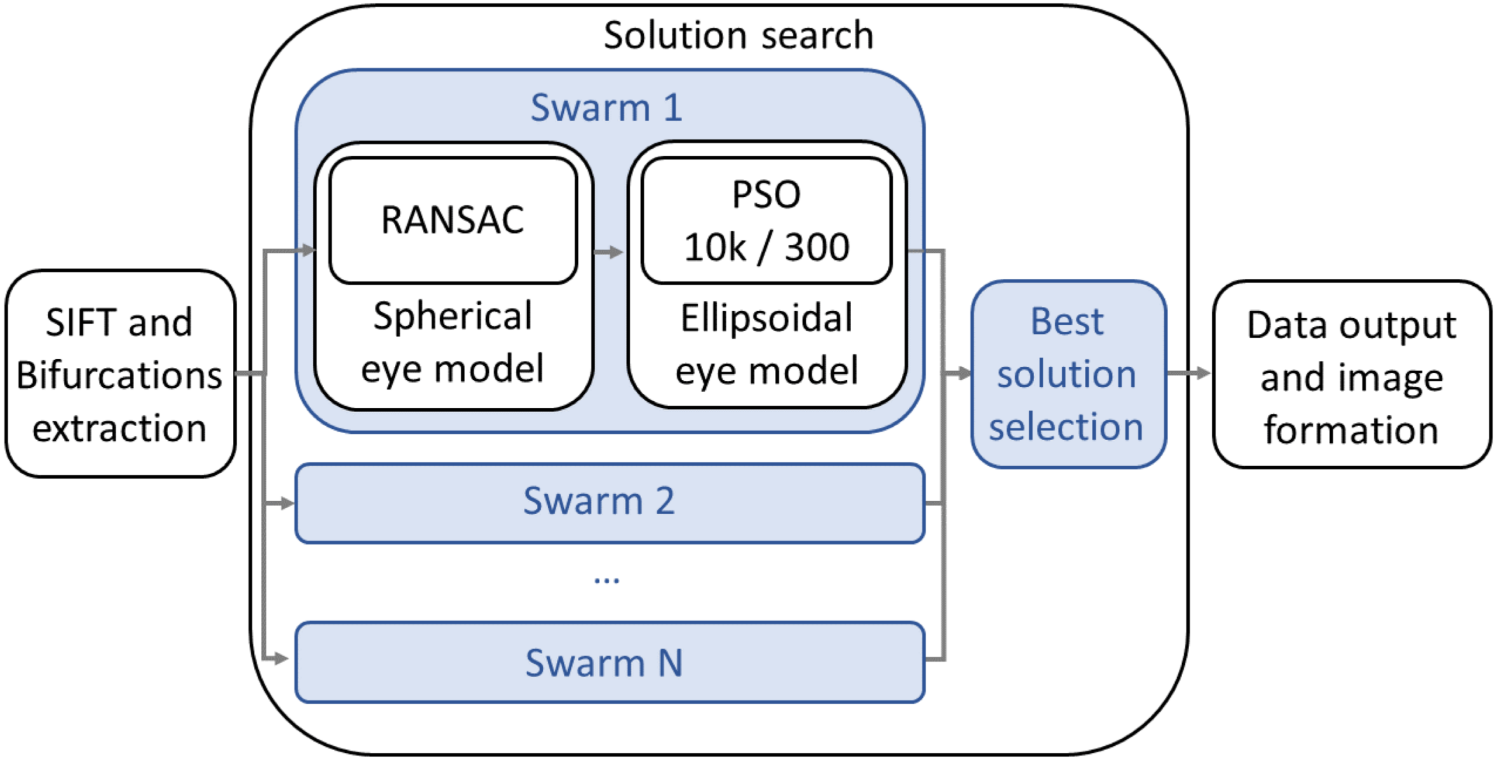
PSO budget



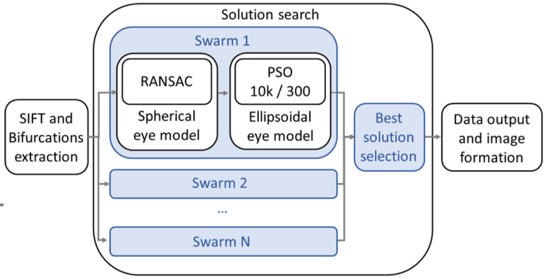


We aim for **high accuracy**, and we don't need to perform registration on real time. As such 300 generations and 10k particles per generation (3 million total) is our ideal trade-off between computational cost and accuracy

Multiple swarms



Multiple swarms



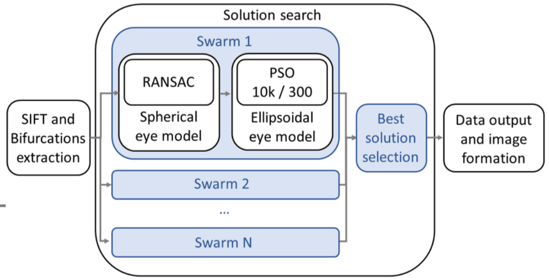
Both **RANSAC** and **PSO non-deterministic**

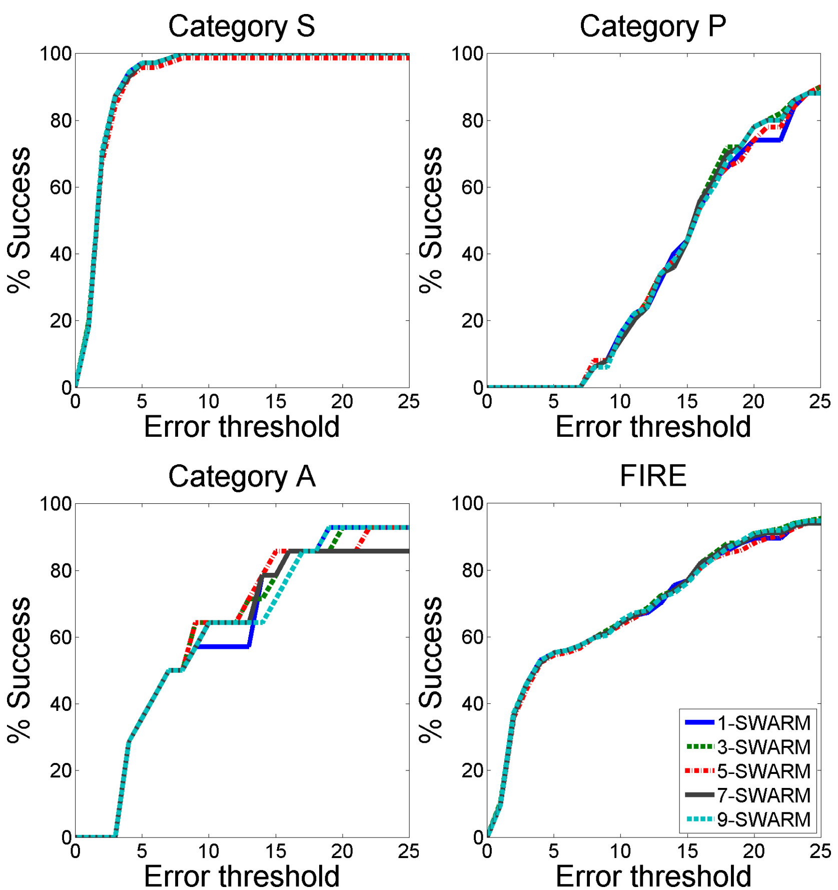
The solution search is executed **multiple parallel** times, denoted as **swarms**

The parameters of the **best overall** score in the objective function are **selected** as the solution.

This leads to an **increase** on the computational **cost**, but this solution offers **increased accuracy**, **robustness** and **reliability**.

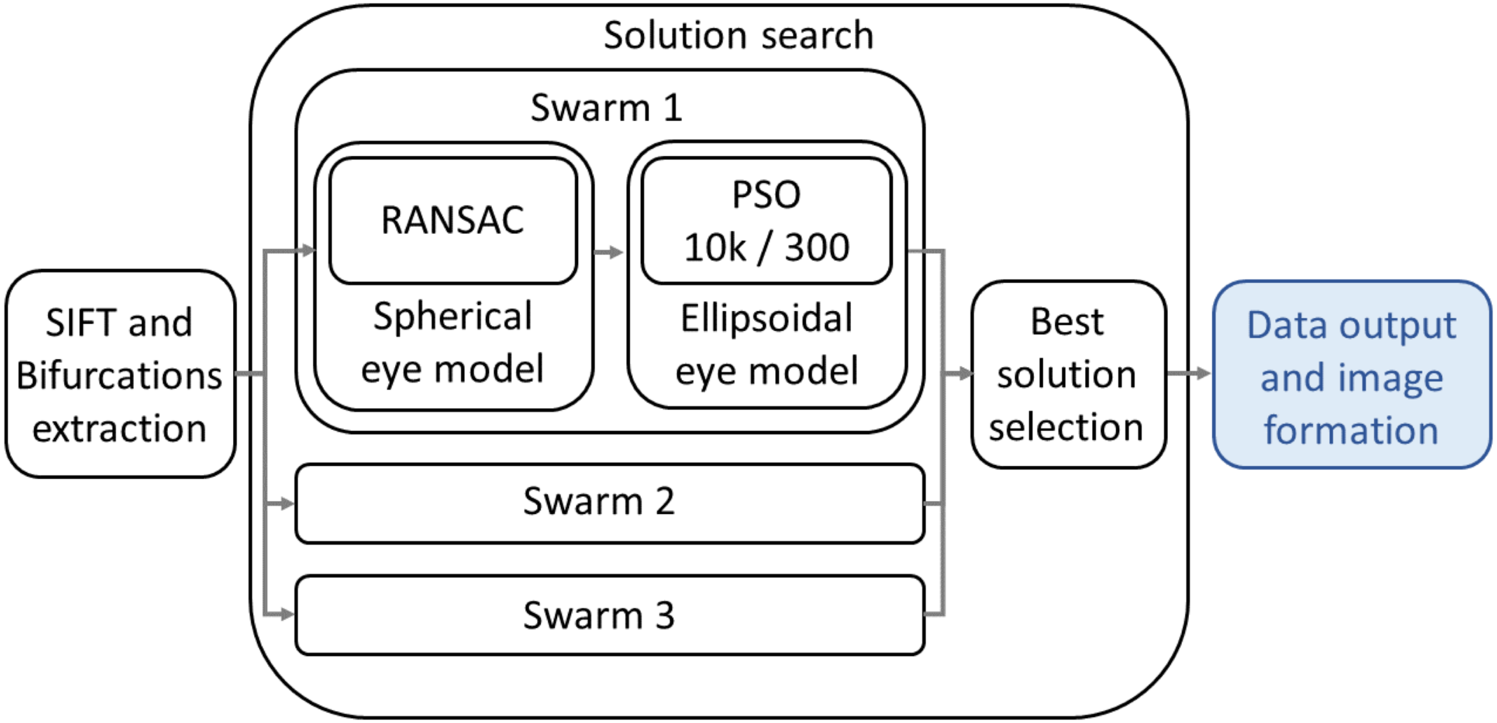
Multiple swarms





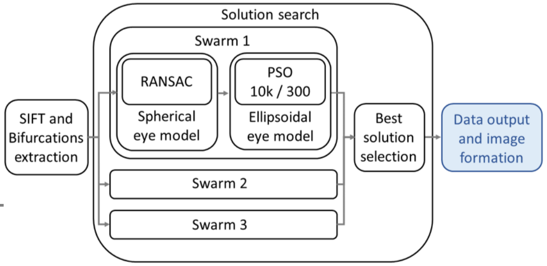
Swarms	S	P	A	FIRE
1	**0.945**	0.370	0.623	0.699
2	**0.945**	**0.383**	0.623	0.703
3	**0.945**	**0.383**	0.634	**0.705**
4	**0.945**	0.377	0.640	0.703
5	0.931	0.374	0.634	0.693
6	0.944	0.378	0.614	0.700
7	0.944	0.378	0.614	0.700
8	**0.945**	0.380	0.620	0.702
9	**0.945**	0.378	0.623	0.702
10	0.944	0.381	**0.643**	0.704

Data output



SIFT and Bifurcations extraction
Solution search
Swarm 1
RANSAC
Spherical eye model
PSO
10k / 300
Ellipsoidal eye model
Swarm 2
Swarm 3
Best solution selection
Data output and image formation

INSTITUTE OF COMPUTER SCIENCE



Data output

R, t, A, Q solution

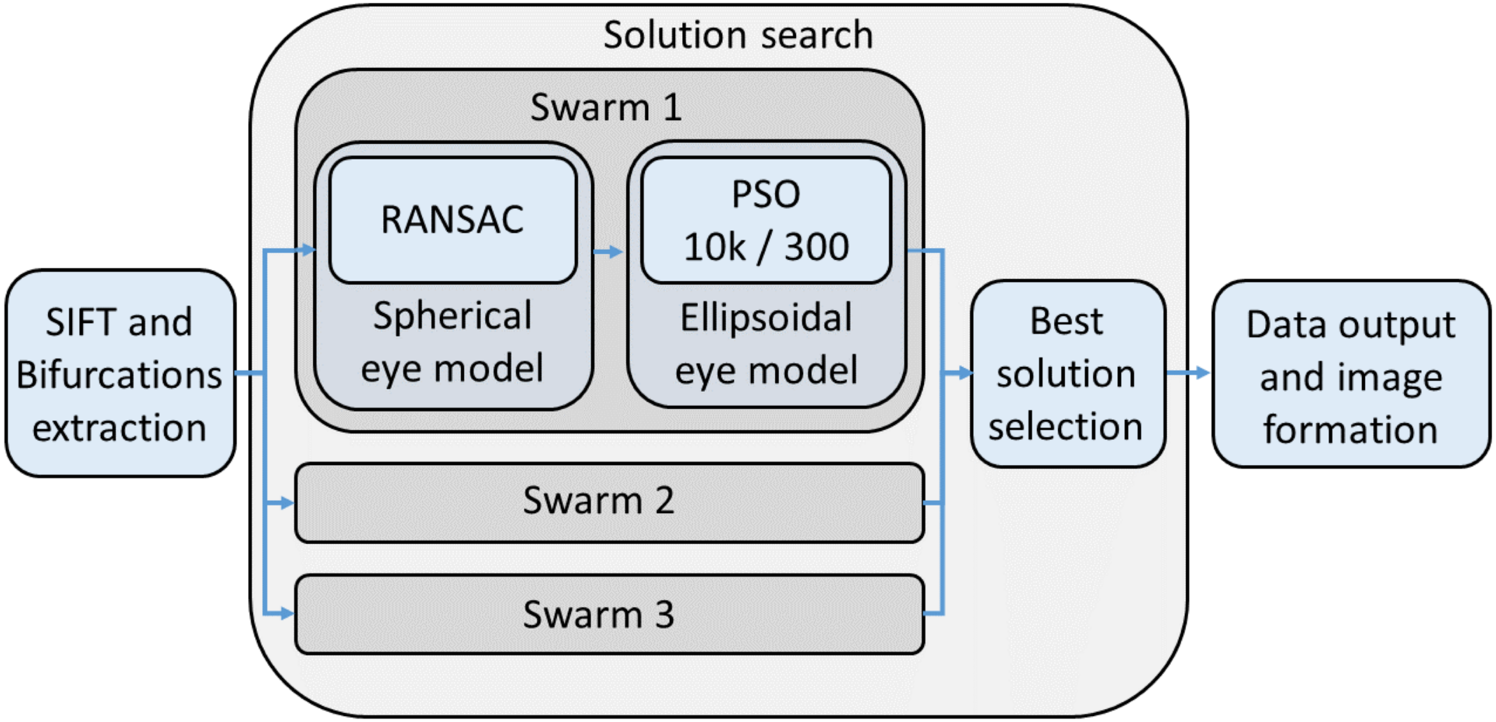
Warped 2D images

2D floating point pixel coordinates and color information for both images

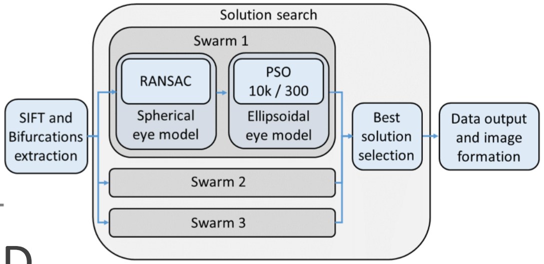
3D coordinates on the eye model and color information for both images

Control point transformations (if control points were provided)

Final configuration



Competing methods



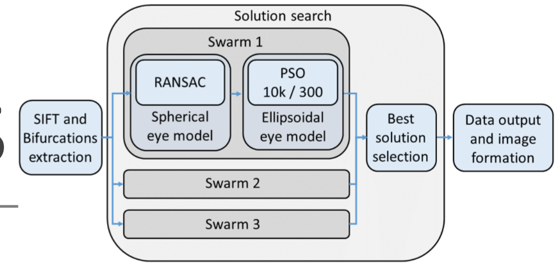
Method is compared to GDB-ICP and Harris-PIIFD

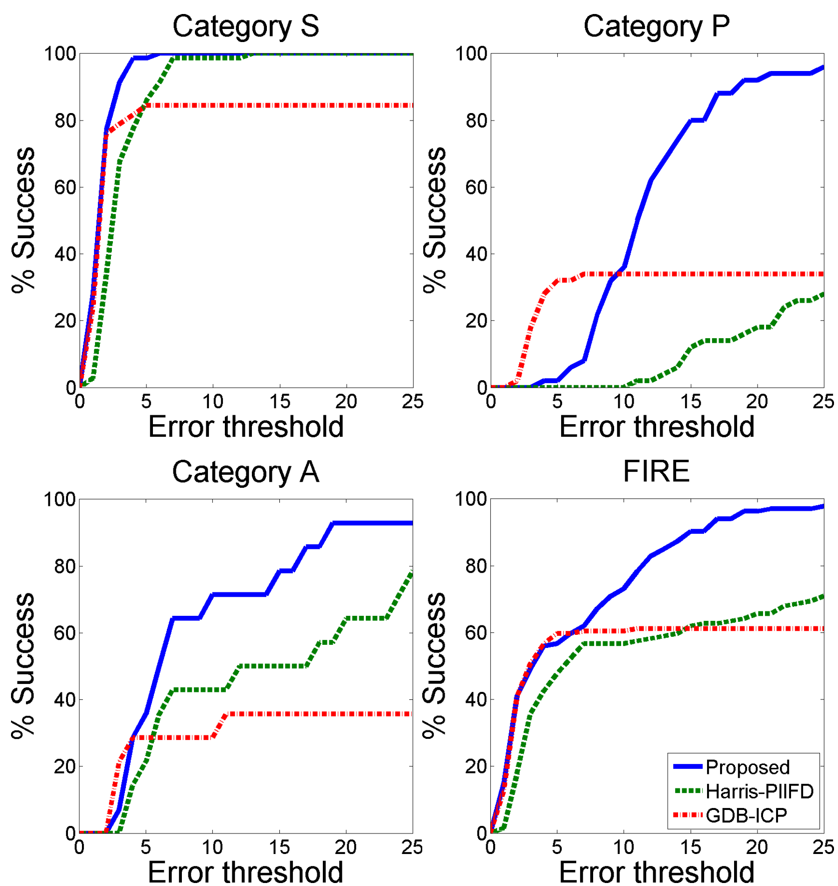
They are widely applied in the field.

Generalized Dual-Bootstrap Iterative Closest Point (GDB-ICP) [Yang 2007]: a local registration method in the spatial domain that performs quadratic registration of intra- and cross-modal retinal images. Face and corner points are used to iteratively register the image pair.

Harris-PIIFD [Chen 2010]: a local registration method in the spatial domain that performs quadratic registration of intra- and cross-modal retinal images. Corner points are selected and PIIFD are extracted. An adaptive transformation is used to register the image pairs.

Competing methods





Method	S	P	A	FIRE
Proposed	**0.958**	**0.541**	**0.660**	**0.773**
Harris-PIIFD	0.900	0.090	0.443	0.553
GDB-ICP	0.814	0.303	0.303	0.576

II – Contributions

Datasets

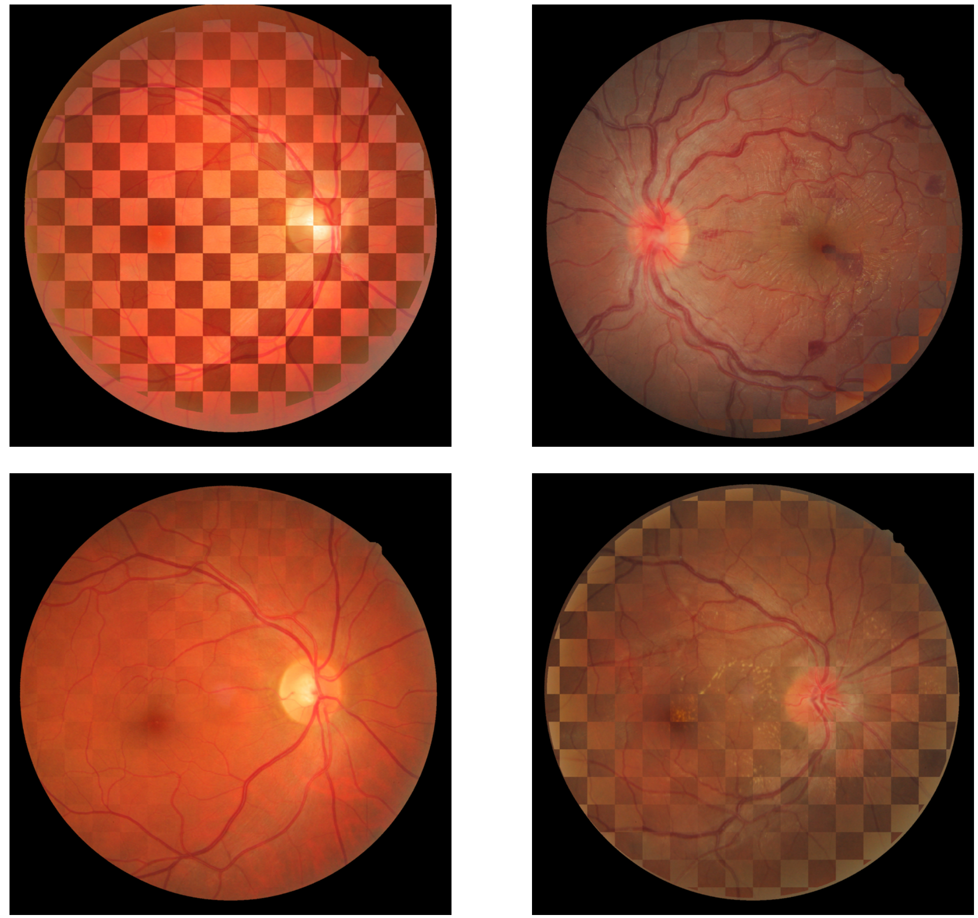
Registration framework and experiments

Registration applications

Longitudinal Studies – FIRE

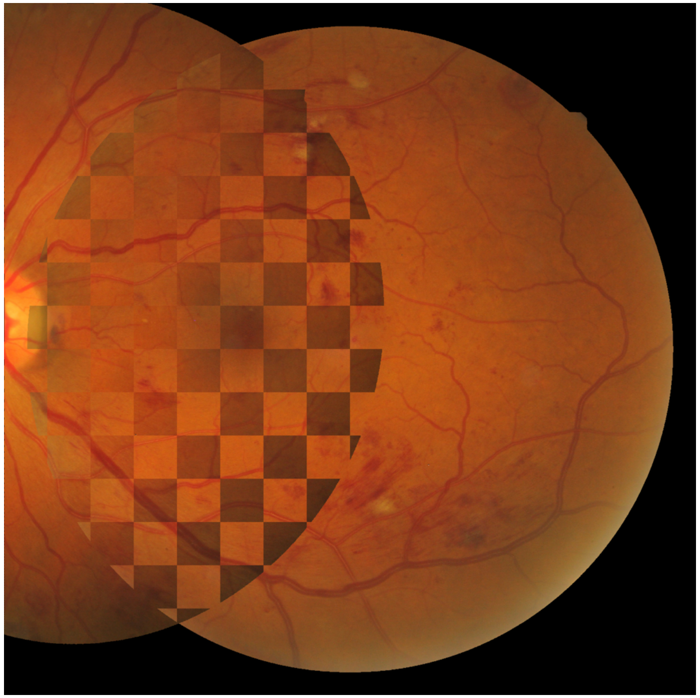
Analyze **differences** across **time**

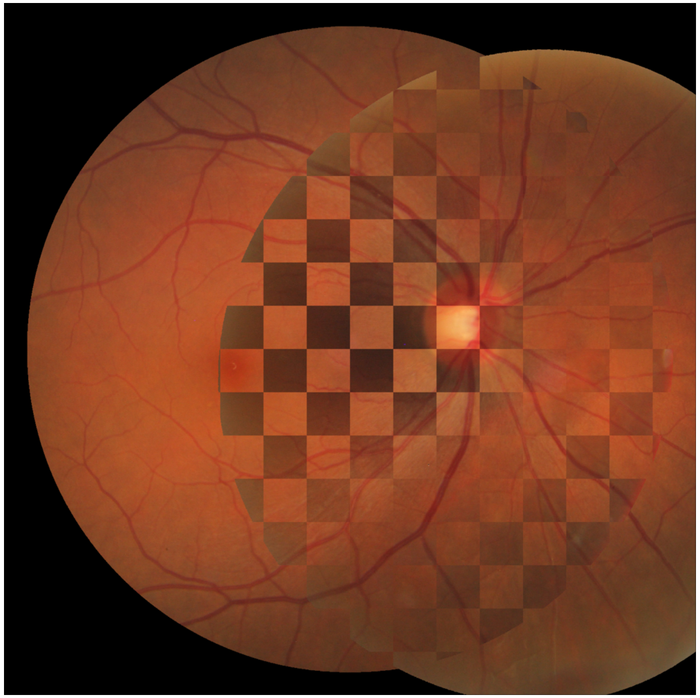
Study evolution of **diseases** such as hypertensive retinopathy



Mosaicing – FIRE

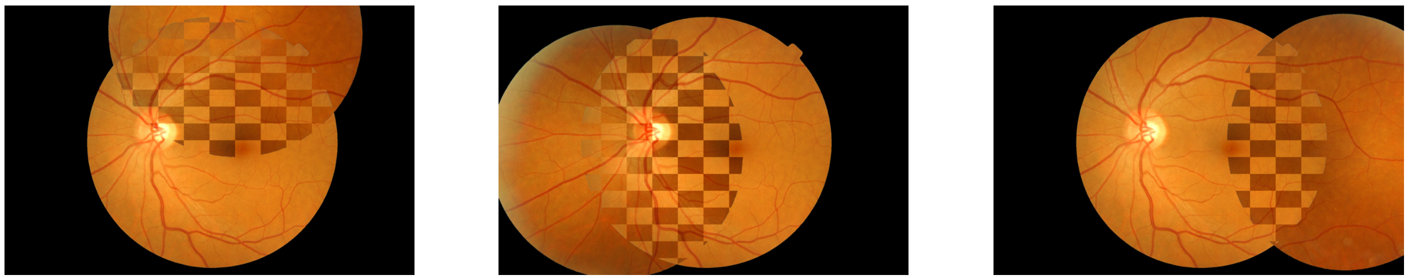
Display larger area of the retina in a **single image**



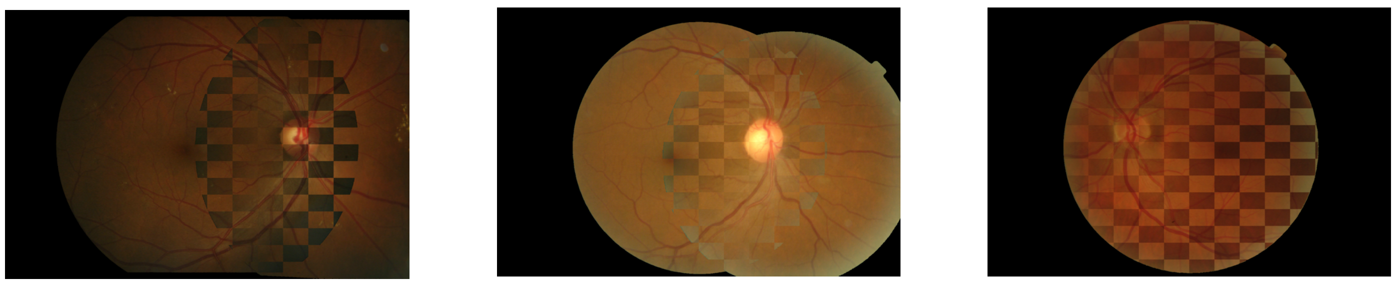


Mosaicing – other datasets

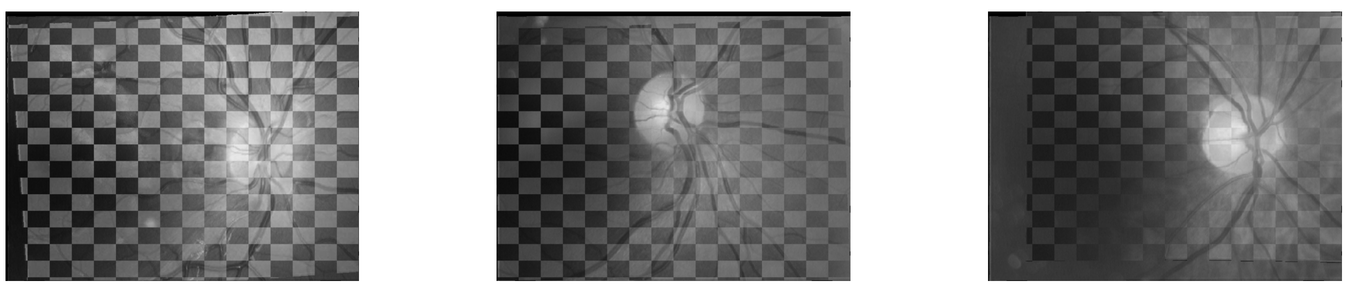
RODREP



e-ophtha



VARIA



Multi-frame Super Resolution

Combine **multiple images** of the **same scene,** acquired from slightly different viewpoints to create an image of **higher resolution** and **definition**.

The basis of SR methods is image registration. It enables the utilization of pixels from different images to be considered as additional samplings of the same function.

Not study of SR per se, but the suitability of the proposed method for it.

2 sets of 9 images, downscaled to 1/3. Super resolution used to generate image of the original size.

	SNR	SSIM	MSE
Proposed	**101.252**	**0.969**	**83.393**
GDB-ICP	84.892	0.816	186.658
Harris-PIIFD	-8.394	0.626	4322.3

Set 1

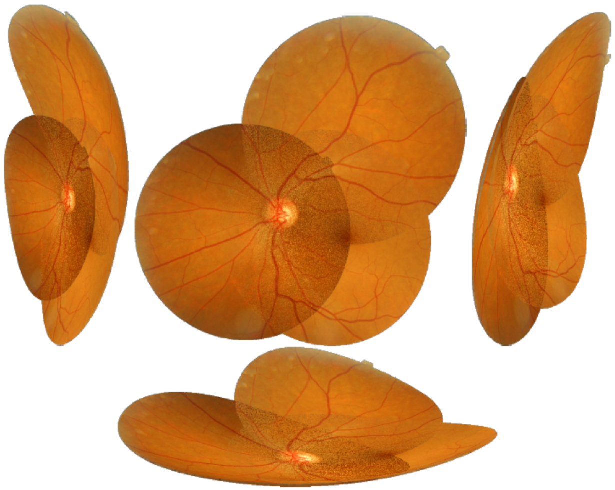
	SNR	SSIM	MSE
Proposed	**56.515**	**0.930**	**190.004**
GDB-ICP	55.256	0.929	199.528
Harris-PIIFD	39.307	0.589	415.985

Set 2

Eye shape estimation

Two sets of synthetic images.

- Fixed pose ellipsoidal model (9 DoF).
- Generated from the same poses.
- 45^o and 100^o FOV
- One experiment searches all 9 parameters. In the other, *c* is fixed to the actual value.
- Error is indicated as the average of the percentage of the ground truth values.



	45^o	100^o
8 DoF search	0.25% (30μm)	0.06% (7.2μm)
9 DoF search	6.54% (785μm)	0.52% (62μm)

Part III

Discussion

III – Discussion

Future work

Discussion

INSTITUTE OF COMPUTER SCIENCE

Future work

Cross-modal keypoints (PIIFD not accurate enough)

Global information extraction methods

More complex eye and camera models

- Currently using pinhole camera model and smooth surface ellipsoid

III – Discussion

Future work

Discussion

Discussion

A retinal image **registration method** is **proposed**.

- It **estimates** the relative **pose** of the **cameras** as well as the general **shape** and **pose** of the **eye**.
- This enables **2D registration** of the retinal images. Also attempts to **reconstruct** the **eye shape**.
- The method has been shown to **outperform competing** methods
- The **suitability** of the proposed registration framework for applications such as **longitudinal studies, mosaicing**, **super resolution** and **eye shape estimation** is explored.
- Executable made **publicly available**

Discussion

The framework allows for **generating synthetic data** such as **3D** eye models. **2D** retinal images can be generated from an existing image texture. Additionally it allows to **evaluate** 3D **shape estimation**.

FIRE dataset, with real images has been compiled and made **publicly available**.

- **Three categories** of retinal image pairs. Each category with the intention of covering a **different challenge** in retinal image registration.
- For each image pair, **ground truth** in the form of control points is provided.

Publications

Journal papers

- *To appear:* C. Hernandez-Matas, X. Zabulis, A. Triantafyllou, P. Anyfanti, S. Douma, A.A. Argyros, "*FIRE: Fundus Image Registration Dataset*", Journal for Modeling in Ophthalmology, 2017.
- C. Hernandez-Matas, X. Zabulis, A. Triantafyllou, P. Anyfanti, A.A. Argyros, "*Retinal Image Registration under the Assumption of a Spherical Eye*", Computerized Medical Imaging and Graphics, Volume 55, January 2017, Pages 95-105.

Conference papers

- *To appear*: C. Hernandez-Matas, X. Zabulis, A.A. Argyros, "*An Experimental Evaluation of the Accuracy of Keypoints-based Retinal Image Registration*", 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju Island, July 11-15, 2017
- C. Hernandez-Matas, X. Zabulis, A.A. Argyros, "*Retinal Image Registration Through Simultaneous Camera Pose and Eye Shape Estimation*", 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 3247-3251, Orlando, August 16-20, 2016
- C. Hernandez-Matas, X. Zabulis, A.A. Argyros, "*Retinal Image Registration Based on Keypoint Correspondences, Spherical Eye Modeling and Camera Pose Estimation*", 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 5650-5654, Milan, August 25-29, 2015.
- C. Hernandez-Matas, X. Zabulis, "*Super Resolution for Fundoscopy Based on 3D Image Registration*", 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 6332-6338, Chicago, August 26-30, 2014.

Downloads

Registration executable

http://www.ics.forth.gr/cvrl/rempe

FIRE dataset

http://www.ics.forth.gr/cvrl/fire

Special thanks









Panagiotis
Koutlemanis



Thank you for your attention!