Artificial Eye Orientation Estimation with Camera-Based Measurements of Natural Visual Features

Adding a camera to obtain an unbiased real-time estimate for the three-dimensional orientation of a humanoid eye

Mariana Ribeiro dos Reis do Vale Martins

Advisor(s)/Supervisor(s)

Prof./Dr. Alexandre Bernardino
Prof./Dr. John Van Opstal
Prof./Dr. José Santos-Víctor

February 2019

Outline

1. Context, Motivation and Objective
2. Background and state of the art
3. Method and proposal
4. Preliminary work
5. Thesis development plan

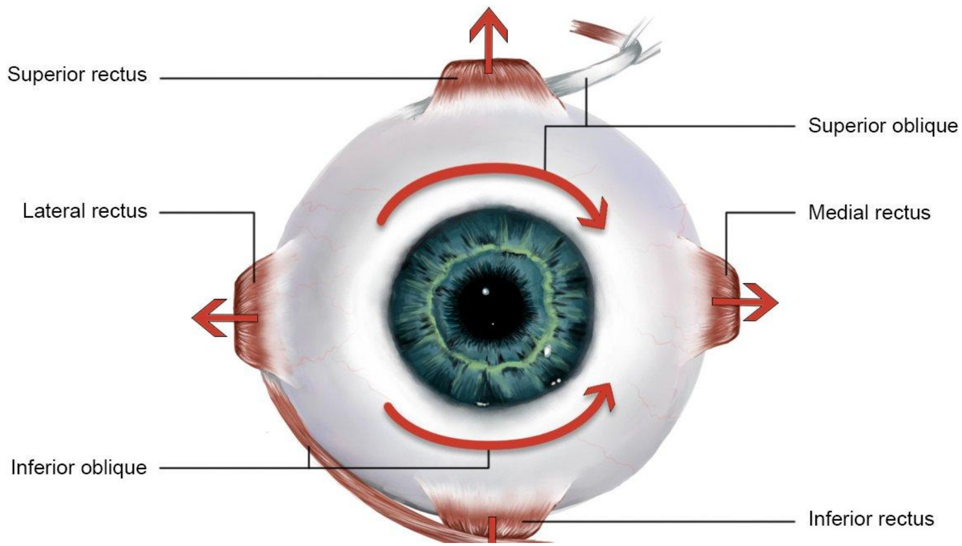
Outline

Context

How is the eye movement defined?

How are sensory-impaired systems different?

[1]

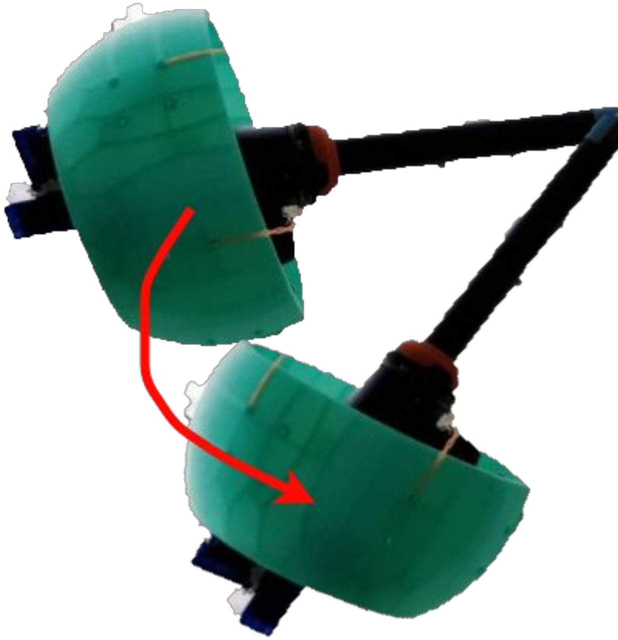


[2]

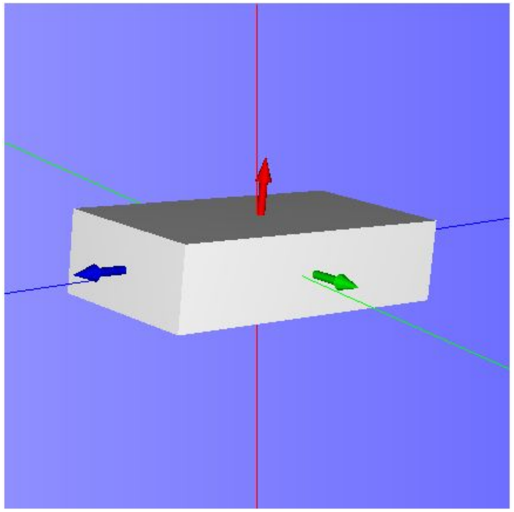
Apply concepts to an approximate model of the human eye

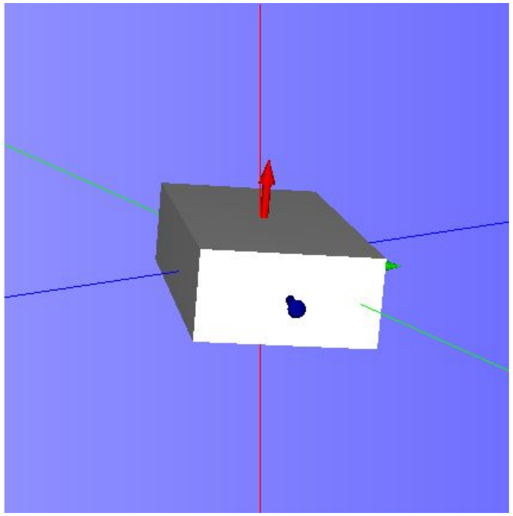
Context

[2]



Motivation





IMU sensor presents significant drift.

Eye’s orientation should be established with better than 0.5 deg resolution!

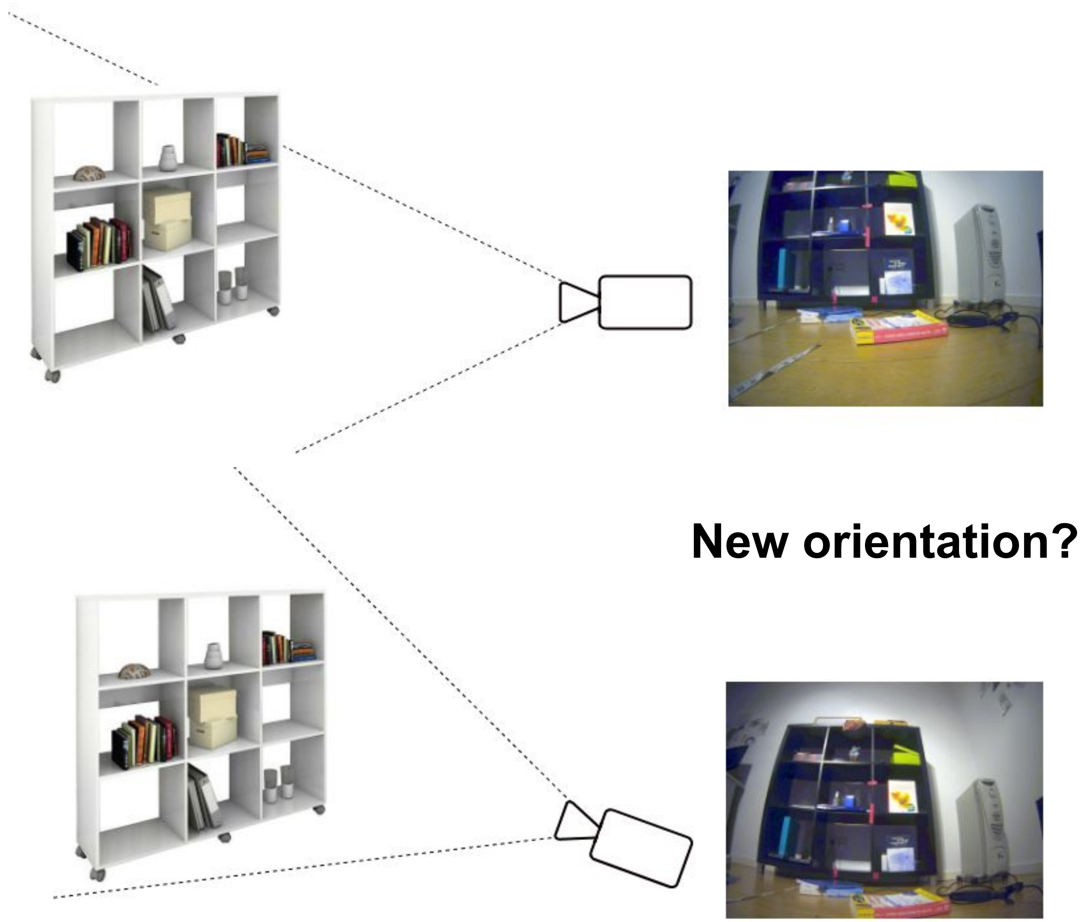
Objective

Estimate the eye's pose as best as possible using a camera with natural visual features.

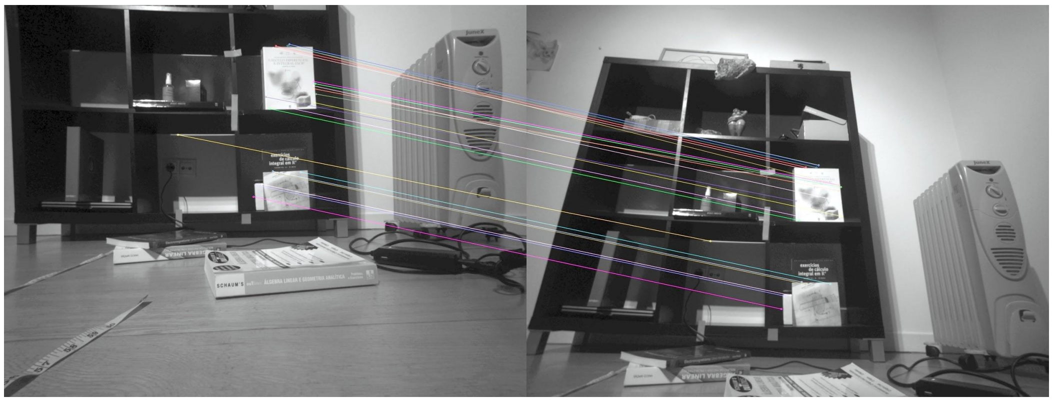
[3]

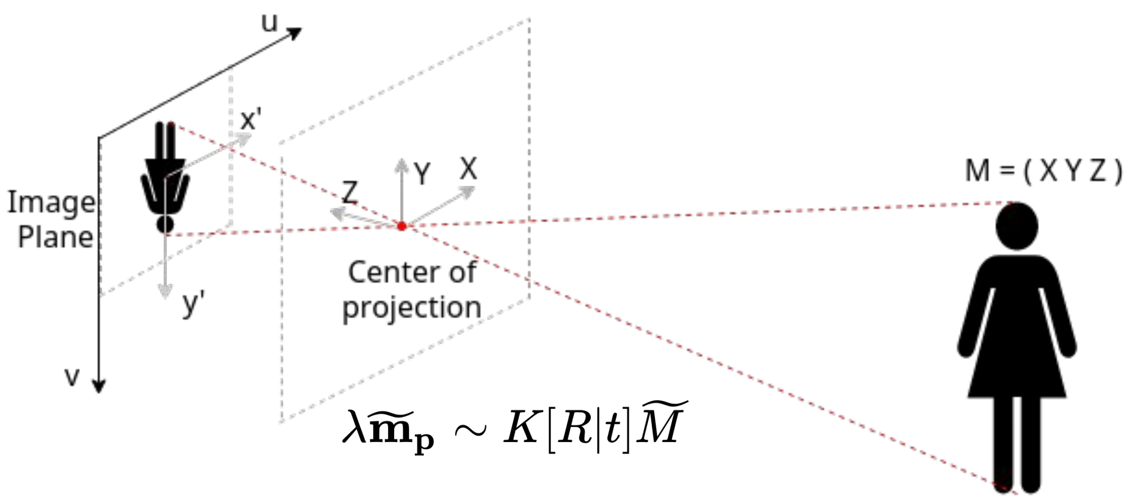
Outline

Background and state of the art



Background and state of the art





$$\lambda\widetilde{\mathbf{m}}_{\mathbf{p}} \sim K[R|t]\widetilde{M}$$

1) Detect features
2) Find matches
3) Determine camera orientation

Background and state of the art

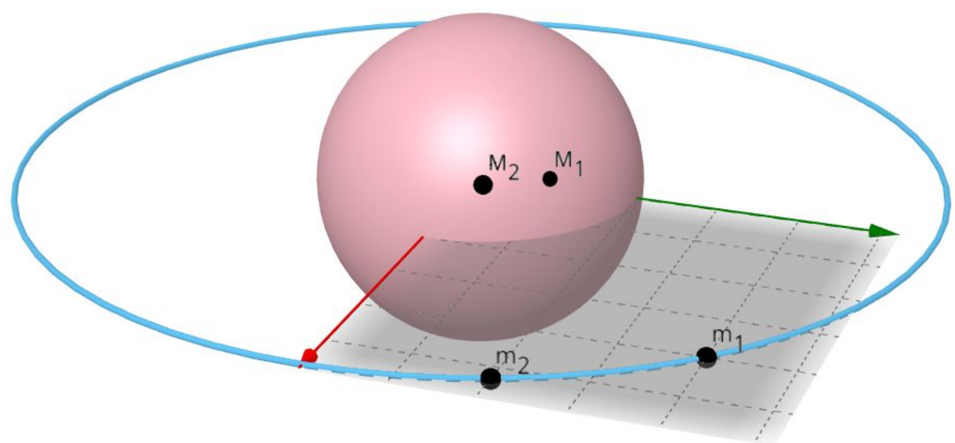
Accuracy will depend on ...

Feature Detection	Feature Matching	Calculate rotation
1. MSER 2. SIFT 3. SURF [5] [6] [7] [8]	1. FLANN 2. Pyramid Match Kernel [9][10]	1. Orthogonal Procrustes Problem 2. Epipolar Geometry [11][12][13]

Background and state of the art

1. Orthogonal Procrustes Problem

[6]



Camera has NO depth (λ) information

$$\|M_1 - RM_2\|^2$$ [5]

$$M_1 = \lambda \widetilde{\mathbf{m}}_\mathbf{1} \qquad \|(\lambda x, \lambda y, \lambda)\| = 1, \lambda = \frac{1}{\sqrt{x^2 + y^2 + 1}}$$

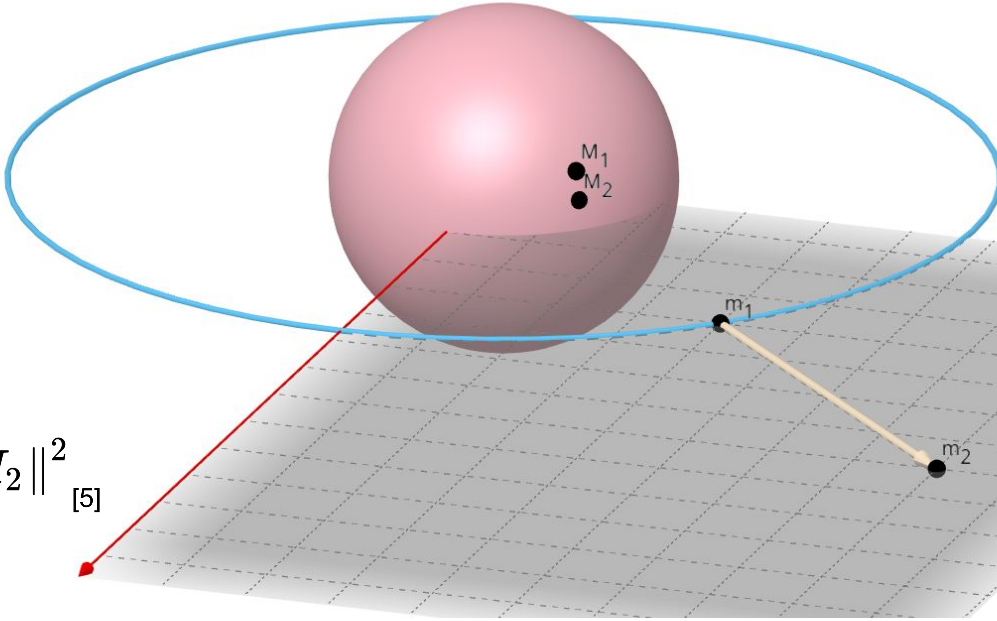
$$\widetilde{\mathbf{m}}_\mathbf{1} = [x\ y\ 1]^T$$

Background and state of the art

1. Orthogonal Procrustes Problem

$$\left\| M_1 - RM_2 \right\|^2$$ [5]

$$\left\| M_1 - RM_2 - \mathbf{1}\mathbf{t}^T \right\|$$

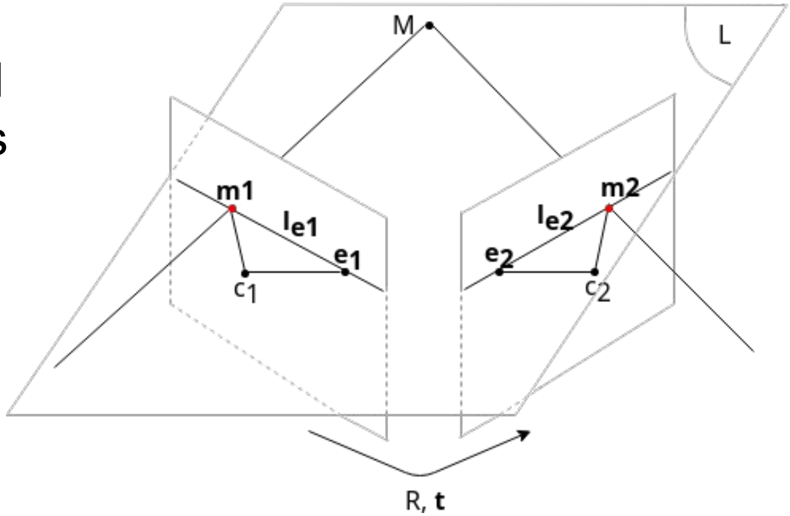


Background and state of the art

2. Epipolar Geometry

Min 7 needed point matches

Must have translation



$$E = [t]_\times R$$ [8]

Outline

1. Context, Motivation and Objective
2. Background and state of the art
3. **Method and proposal**
4. Preliminary work
5. Thesis development plan

Method and proposal

$$\mathbf{t} = (R - I)\mathbf{b}$$

Minimization of the Back projection Error

$$J(\mathbf{m_{1i}}, \mathbf{m_{2i}}, R, \mathbf{b}, Z_{1i}) = \sum_{i=1}^{N} \left[(u'_{1i} - u_{1i})^2 + (u'_{2i} - u_{2i})^2 + (v'_{1i} - v_{1i})^2 + (v'_{2i} - v_{2i})^2 \right]$$

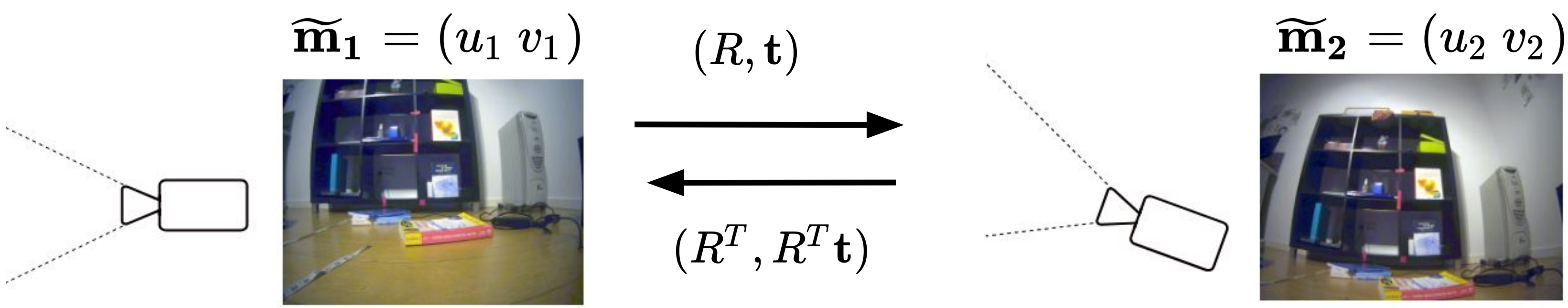
Z_{1i} is the depth for view 1

Method and proposal

Minimization of the Back projection Error

$$\boxed{J(\mathbf{m_{1i}}, \mathbf{m_{2i}}, R, \mathbf{b}, Z_{1i}) = \sum_{i=1}^{N} \left[(u'_{1i} - u_{1i})^2 + (u'_{2i} - u_{2i})^2 + (v'_{1i} - v_{1i})^2 + (v'_{2i} - v_{2i})^2 \right]}$$

Z_{1i} is the depth for view 1



$$\widetilde{\mathbf{m}}_\mathbf{1} = (u_1 \; v_1) \longrightarrow \boxed{(R, \mathbf{t})} \longrightarrow \widetilde{\mathbf{m}}'_\mathbf{2} = (u'_2 \; v'_2)$$

$$\widetilde{\mathbf{m}}_\mathbf{2} = (u_2 \; v_2) \longrightarrow \boxed{(R^T, R^T\mathbf{t})} \longrightarrow \widetilde{\mathbf{m}}'_\mathbf{1} = (u'_1 \; v'_1)$$

Method and proposal

Minimization of the Back projection Error

$$\boxed{J(\mathbf{m_{1i}}, \mathbf{m_{2i}}, R, \mathbf{b}, Z_{1i}) = \sum_{i=1}^{N} \left[(u'_{1i} - u_{1i})^2 + (u'_{2i} - u_{2i})^2 + (v'_{1i} - v_{1i})^2 + (v'_{2i} - v_{2i})^2\right]}$$

Z_{1i} is the depth for view 1

$$Z_2 = Z_1 \mathbf{r_3}^T \widetilde{\mathbf{m}}_\mathbf{1} - \mathbf{t_3}$$

$$u'_1 = \frac{Z_2 \mathbf{r_1} \widetilde{\mathbf{m}}_\mathbf{2} - \mathbf{r_1}\mathbf{t}}{Z_1} \qquad u'_2 = \frac{Z_1 \mathbf{r_1}^T \widetilde{\mathbf{m}}_\mathbf{1} + \mathbf{t_1}}{Z_1 \mathbf{r_3}^T \widetilde{\mathbf{m}}_\mathbf{1} + \mathbf{t_3}}$$

$$v'_1 = \frac{Z_2 \mathbf{r_2} \widetilde{\mathbf{m}}_\mathbf{2} - \mathbf{r_2}\mathbf{t}}{Z_1} \qquad v'_2 = \frac{Z_1 \mathbf{r_2}^T \widetilde{\mathbf{m}}_\mathbf{1} + \mathbf{t_2}}{Z_1 \mathbf{r_3}^T \widetilde{\mathbf{m}}_\mathbf{1} + \mathbf{t_3}}$$

$$\widetilde{\mathbf{m}}_\mathbf{1} = (u_1 \; v_1) \longrightarrow \boxed{(R, \mathbf{t})} \longrightarrow \widetilde{\mathbf{m}}'_\mathbf{2} = (u'_2 \; v'_2)$$

$$\widetilde{\mathbf{m}}_\mathbf{2} = (u_2 \; v_2) \longrightarrow \boxed{(R^T, R^T\mathbf{t})} \longrightarrow \widetilde{\mathbf{m}}'_\mathbf{1} = (u'_1 \; v'_1)$$

Method and proposal

Compare Feature detection and matching methods

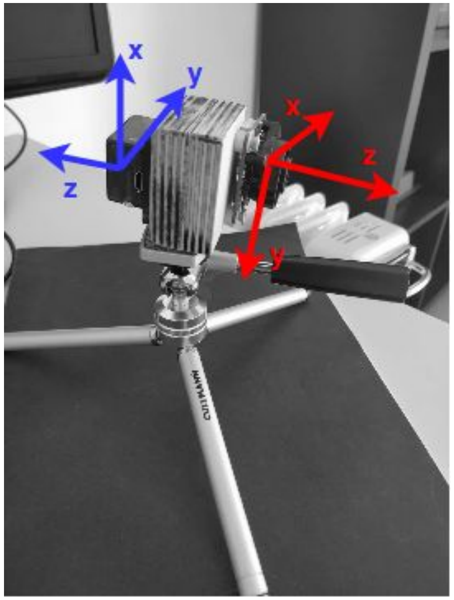
Compare Procrustes, Epipolar Geometry and our method

Study the benefits of camera over sensor, or mutual usage

Outline

Preliminary work

Comparison of rotations around each axis between the IMU sensor and the camera

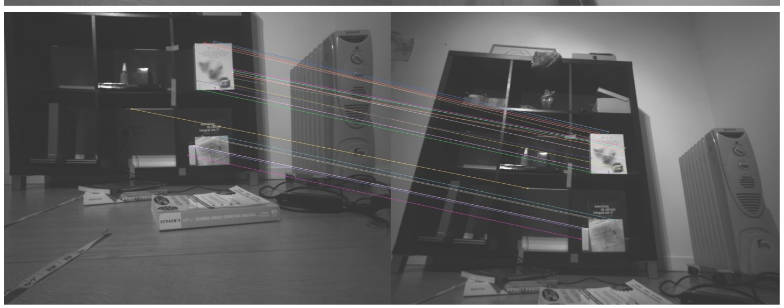


1) Undistort
2) SURF + FLANN + $\sqrt{\sum_{i=1}^{64} |a_i - b_i|^2}$
3) Orthogonal Procrustes Problem

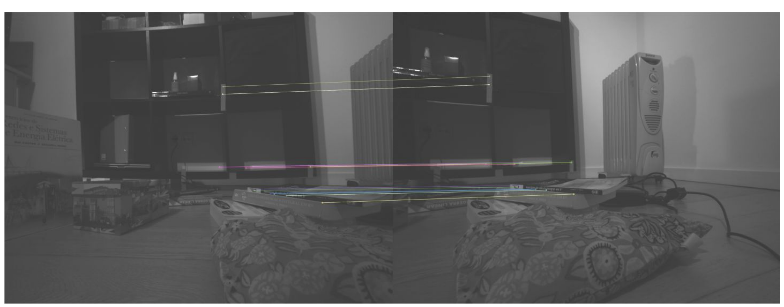
$$\|(\lambda x, \lambda y, \lambda)\| = 1, \lambda = \frac{1}{\sqrt{x^2 + y^2 + 1}}$$

Preliminary work

θ	θ_{Cam}	θ_{IMU}	$n_{matches}$
90^o	86.76^o	90.9^o	7



θ	θ_{Cam}	θ_{IMU}	$n_{matches}$
19^o	21.37^o	16.56^o	23
5^o	5.41^o	5.35^o	28

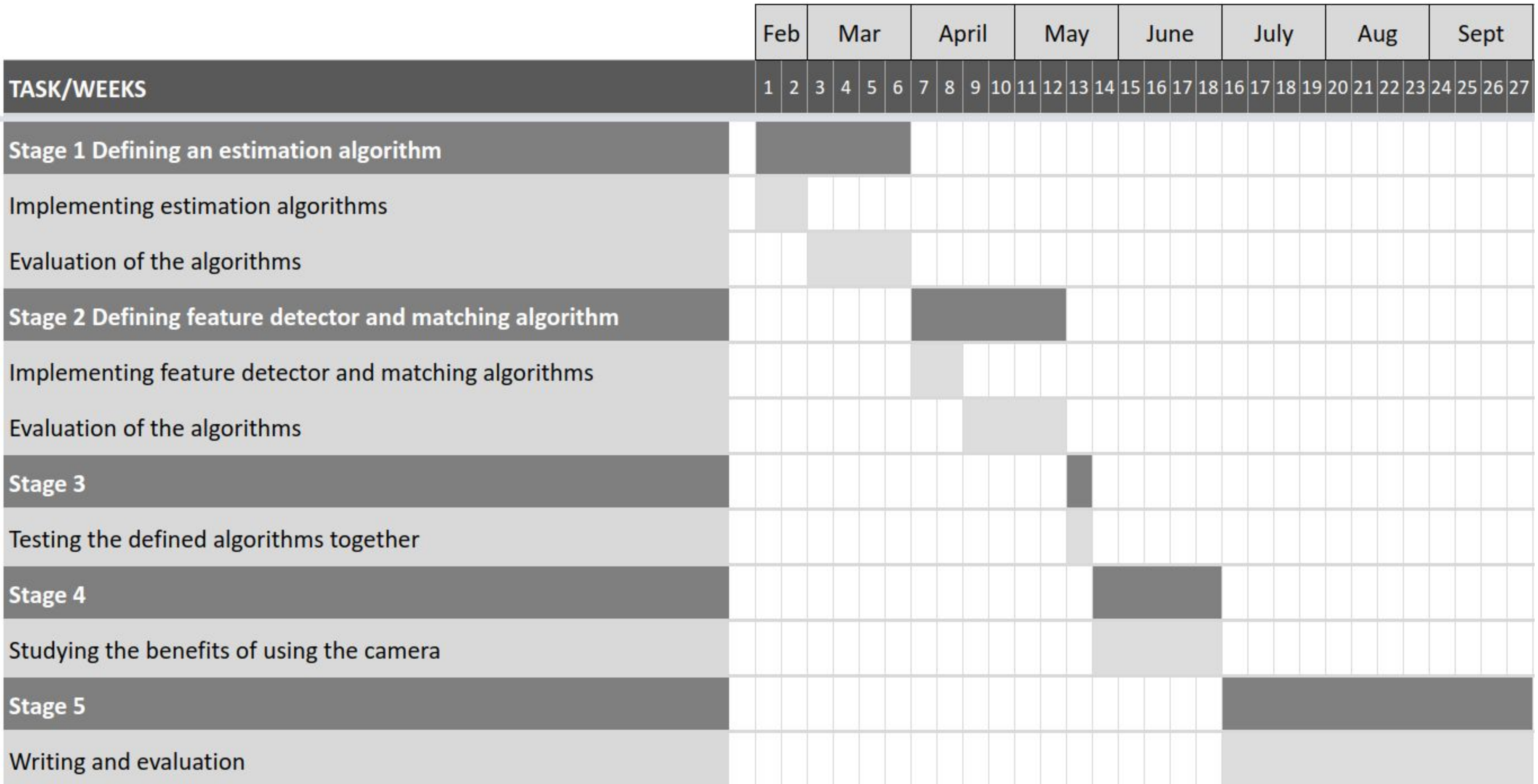


θ	θ_{Cam}	θ_{IMU}	$n_{matches}$
30^o	31.75^o	10.63^o	11
5^o	5.27^o	6.12^o	27

Outline

1. Context, Motivation and Objective
2. Background and state of the art
3. Method and proposal
4. Preliminary work
5. **Thesis development plan**

Thesis development plan



[1] J. van Opstal. http://www.mbfys.ru.nl/~johnvo/OrientWeb/orient_1.html#Abstract. Acessed on December 15th 2018.

[2] Eye Movement. https://step1.medbullets.com/neurology/113079/eye-movement. Accessed on January 17th 2019.

[3] M. Lucas, “Construction and Characterization of a Biomimetic Robotic Eye Model with Three Degrees of Rotational Freedom: A Testbed for Neural Control of Eye Movements,” 2017.

[4] IDS camera. https://en.ids-imaging.com/news-article/en_usb-3_1-le-18_1-mp.html. Accessed on January 27th 2019.

[5] E. Salahat and M. Qasaimeh, “Recent Advances in Features Extraction and Description Algorithms: A Comprehensive Survey,” 2017.

[6] D. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints,” 2004.

[7] H. BayTinne, T. TuytelaarsLuc and L. Van Gool, “SURF: Speeded Up Robust Features,” 2006.

[8] D. Mistry and A. Banerjee, “Comparison of Feature Detection and Matching Approaches: SIFT and SURF,” 2017.

[9] M. Muja and D. Lowe, “Fast approximate nearest neighbors with automatic algorithm configuration,”, 2009.

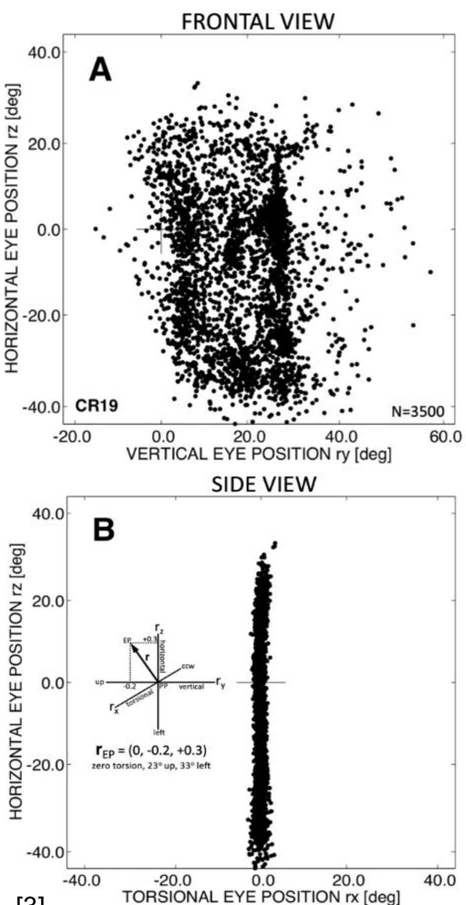
[10] K. Grauman and T. Darrell, ”The Pyramid Match Kernel: Efficient Learning with Sets of Features”, 2005

[11] J. Gower and G. Dijksterhuis, Procrustes Problems. Oxford Statistical Science Series, 2004. Chapter 3 and 4.

[12] Procrustes. http://www.mythweb.com/today/today07.html. Accessed on 16th January 2019.

[13] R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision. Cambridge University Press, 2 ed., 2003. Part II: Chapter 9.

Motivation



[3]

The eye does not achieve all possible 3D orientations.

But all can be reached starting from the “primary position” - normal to the plane - and then rotating about an axis that lies in the **Listing’s plane**.

The component of the rotation axis normal to this plane is the torsional component and is close to zero.

[4]

How is the “primary position” determined?