# Towards a Unified Approach to Homography Estimation Using Image Features and Pixel Intensities

Lucas Nogueira[1], Ely Paiva[1] and Geraldo Silveira[2]

1 School of Mechanical Engineering - UNICAMP
2 Center for Information Technology Renato Archer
Campinas, Brazil

International Conference on Autonomic and Autonomous Systems, Sept 2020

# Presenter Info


Lucas Nogueira

- BSc in Computer Engineering @ University of Campinas in 2015
- MSc in Mechanical Engineering @ University of Campinas 2019
- Currently @ Carnegie Mellon University.
- Contact: `lcasanov@andrew.cmu.edu`
- Add me on Linkedin: `https://www.linkedin.com/in/lukscasanova`

# Outline

# Table of Contents

Feature-Based (FB):

- Feature **extraction** and **matching**
- Large convergence basin
- Lower precision
- The information space is the feature's **coordinates**

Intensity-Based (IB):

- NO intermediate steps. **Direct** estimation
- Smaller convergence basin
- Higher precision
- The information space is the pixel **intensities**

- We want the **precision** of IB methods and the **convergence** domain of FB methods
- The **Efficient Second-Order Minimization**[1] framework gives great results for IB methods, but its application to FB methods has been limited
- **Unify** minimization method, cost function, homography parametrization, Jacobians, and more

[1] Selim Benhimane and Ezio Malis. "Real-time image-based tracking of planes using efficient second-order minimization". In: *Proc. IEEE/RSJ IROS*. 2004, pp. 943–948.

# Contribution

Development of a vision-based estimation algorithm that:

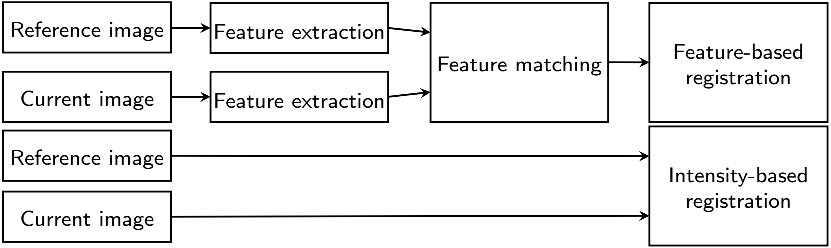
- Accurately and robustly estimates the homography matrix between two images
- Unifies the IB and FB approaches into a single optimization procedure
- Suitable to real-time settings (>30 Hz)
- Publicly available library and ROS package

# Table of Contents

- The goal is to optimally aligns two images of the same scene
- Existing methods are a combination of these components:
  1. information space
  2. transformation model
  3. similarity measure
  4. search strategy

Inputs to the Image Registration

- **Feature-based**: Geometric features coordinates
- **Intensity-based**: Pixel intensity values



- **Geometric** - changes due to motion:

$$\mathbf{p} \propto \mathbf{H}\,\mathbf{p}^* = \mathbf{w}(\mathbf{H}, \mathbf{p}^*)$$

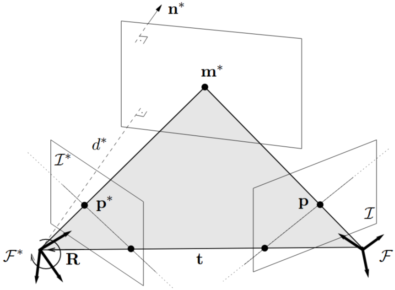
- **Photometric** - changes due to lighting:

$$\mathcal{I}'(\mathbf{p}) = \alpha\,\mathcal{I}(\mathbf{p}) + \beta$$

  - $\mathbf{SSD}(\mathcal{I}', \mathcal{I}^*) = \sum_i \left[\mathcal{I}'(\mathbf{p}_i) - \mathcal{I}^*(\mathbf{p}_i^*)\right]^2$
  - $\mathbf{ZNCC}(\mathcal{I}', \mathcal{I}^*) = \left\langle \frac{\mathcal{I}_v^* - \bar{\mathcal{I}}_v^*}{\left\|\mathcal{I}_v^* - \bar{\mathcal{I}}_v^*\right\|}, \frac{\mathcal{I}_v' - \bar{\mathcal{I}}_v'}{\left\|\mathcal{I}_v' - \bar{\mathcal{I}}_v'\right\|} \right\rangle$

# Plane-induced Homography

The homography relates coordinates of the same scene point in two separate views of a planar object




$$\mathbf{p} \propto \mathbf{H}\,\mathbf{p}^*$$

# Search Strategy

- **ESM**:
  - The ESM increment is defined as:

$$\widetilde{\mathbf{x}}_n = -2(\mathbf{J}(\widehat{\mathbf{x}}_n) + \mathbf{J}(\mathbf{x}^*))^{+}\mathbf{f}(\widehat{\mathbf{x}}_n),$$

    where $\mathbf{J}(\mathbf{x})$ is the Jacobian of function $\mathbf{f}$
  - Note that ESM only requires calculating the first-order derivatives
- **Multiresolution Pyramid**:
  - Increase in the computational efficiency
  - Avoidance of spurious local minima
  - Larger domain of convergence

# Table of Contents

We want to find a vector $\mathbf{x}^* \in \mathbb{SL}(3) \times \mathbb{R}^2 = \{\mathbf{H}^*, \alpha^*, \beta^*\}$ such that:

$$\alpha^* \mathcal{I}(\mathbf{w}(\mathbf{H}^*, \mathbf{p}_i^*)) + \beta^* = \mathcal{I}^*(\mathbf{p}_i^*), \qquad \forall i = 1, \ldots, m \tag{1}$$

$$\mathbf{w}(\mathbf{H}^*, \mathbf{q}_j^*) = \mathbf{q}_j, \qquad \forall j = 1, \ldots, n \tag{2}$$

- 2 different information types: Coordinates AND Intensities
- Proposed cost function:

$$\min_{\mathbf{x}=\{\mathbf{H},\alpha,\beta\}} \quad \frac{1}{2}\left( w_{IB} \left\| \mathbf{y}_{IB}(\mathbf{x}) \right\|^2 + w_{FB} \left\| \mathbf{y}_{FB}(\mathbf{x}) \right\|^2 \right)$$

with $w_{IB} + w_{FB} = 1$

# Jacobian

The Jacobian derivation is simply the stacking of the IB and FB Jacobians

$$\mathbf{J}_{UN} = \begin{bmatrix} \frac{w_{IB}}{m}\mathbf{J}_{IB} \\ \frac{w_{FB}}{2n}\mathbf{J}_{FB} \end{bmatrix}$$

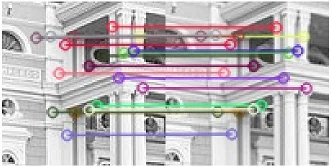
The choice of weights depends on the distance of feature correspondences

$$RMSD(\mathbf{y}_{FB}) = \sqrt{\frac{\sum_{j=1}^{n} d_j^2}{n}} = d_{FB}$$

$$w_{FB} = 1 - e^{-d_{FB}}$$


## Local Search

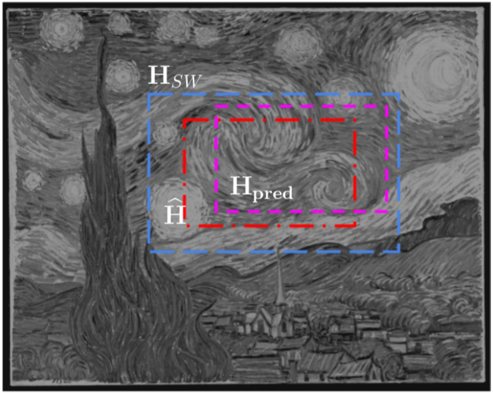
Features in current *template*



## Global Search

Features in current *image*

Idea: Generate multiple candidates solutions, evaluating them according to ZNCC metric, and choose the best one.



# Table of Contents

# Benchmark Tests Procedure

Randomly generate test homographies and evaluates if the estimation correctly recovers it

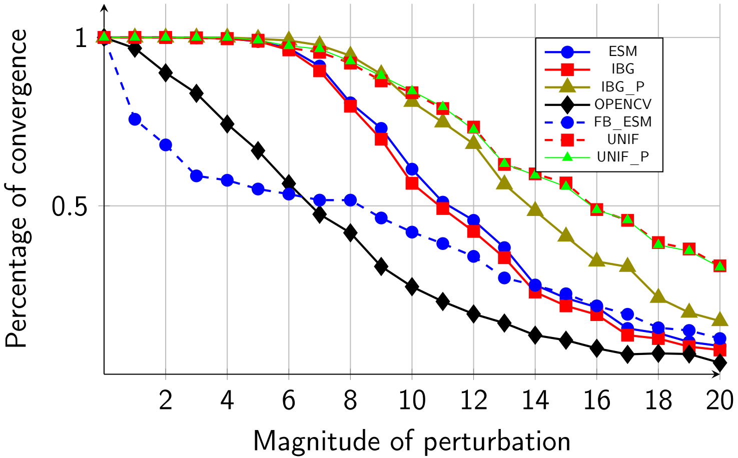
Reference Image

Reference Template

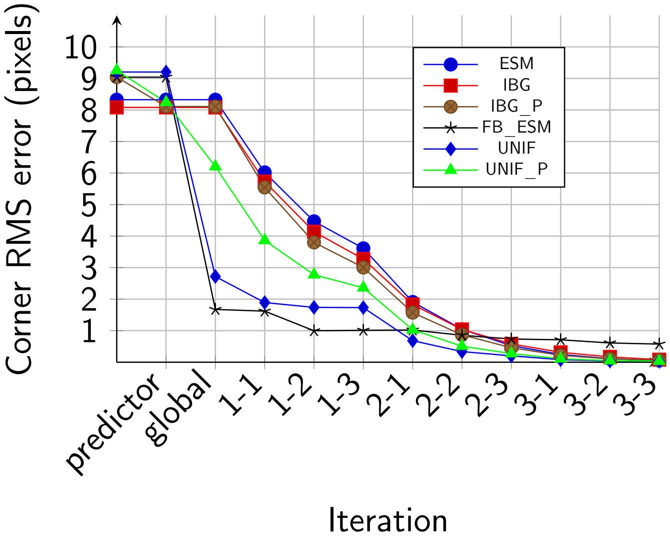
Transformed with $\sigma = 5$

Transformed with $\sigma = 10$

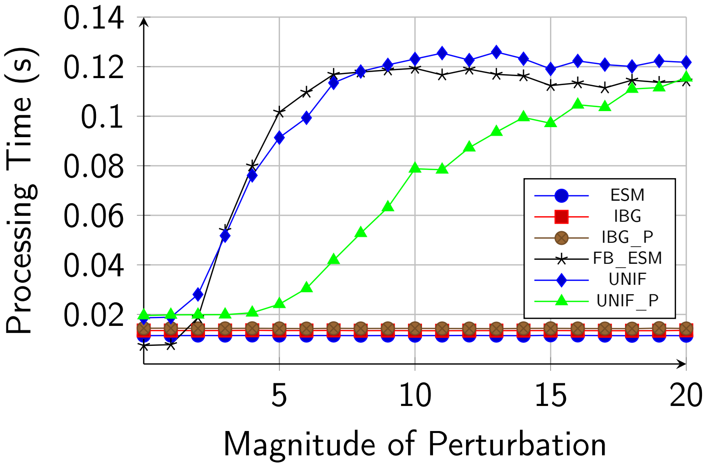
Convergence criteria: RMS of corner point error $< 1$ pixels


Percentage of convergence
1
0.5
2
4
6
8
10
12
14
16
18
20
Magnitude of perturbation
ESM
IBG
IBG_P
OPENCV
FB_ESM
UNIF
UNIF_P

# Rate of Convergence ($\sigma = 10$)



# Timing

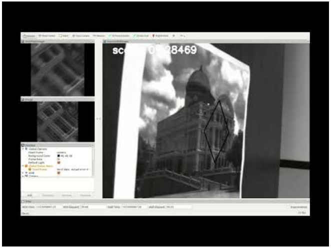


Processing Time (s)
0.14
0.12
0.1
0.08
0.06
0.04
0.02
5
10
15
20
Magnitude of Perturbation
ESM
IBG
IBG_P
FB_ESM
UNIF
UNIF_P

# Table of Contents

The proposed algorithm is available as a ready to use C++ library and as a ROS package, with its technical report[2] at:

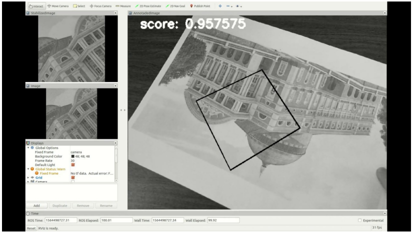
`https://github.com/visiotec/vtec_ros`



[2] Lucas Nogueira, Ely de Paiva, and Geraldo Silveira. *VTEC robust intensity-based homography optimization software*. Tech. rep. CTI-VTEC-TR-01-19. Brazil: CTI, 2019.

# Unified Visual Tracking

Using the proposed unified method, we built a visual tracker. An important improvement is the implicit ability to reinitialize itself.



# Table of Contents

# Conclusions And Future Work

- The unified method has the convergence of FB and the precision of IB methods
- Further tests can be conducted in the FB component: other feature detectors; choice of weights; outlier rejection
- Application to homography-based visual servoing
- Different camera sensor types: omnidirectional, color images
- Extension to nonplanar surfaces

Thank you!