

Contributions to Local Feature Extraction, Description and Matching in 2D Images

Iñigo Barandiaran Martirena

ibarandiaran@vicomtech.org

Department of Computer Science and Artificial Intelligence
University of the Basque Country



PhD Advisor:
Prof. Dr. Manuel Graña Romay

Vicomtech Internal Supervisor:
Dr. Marcos Nieto Doncel

San Sebastián, 12th July 2012



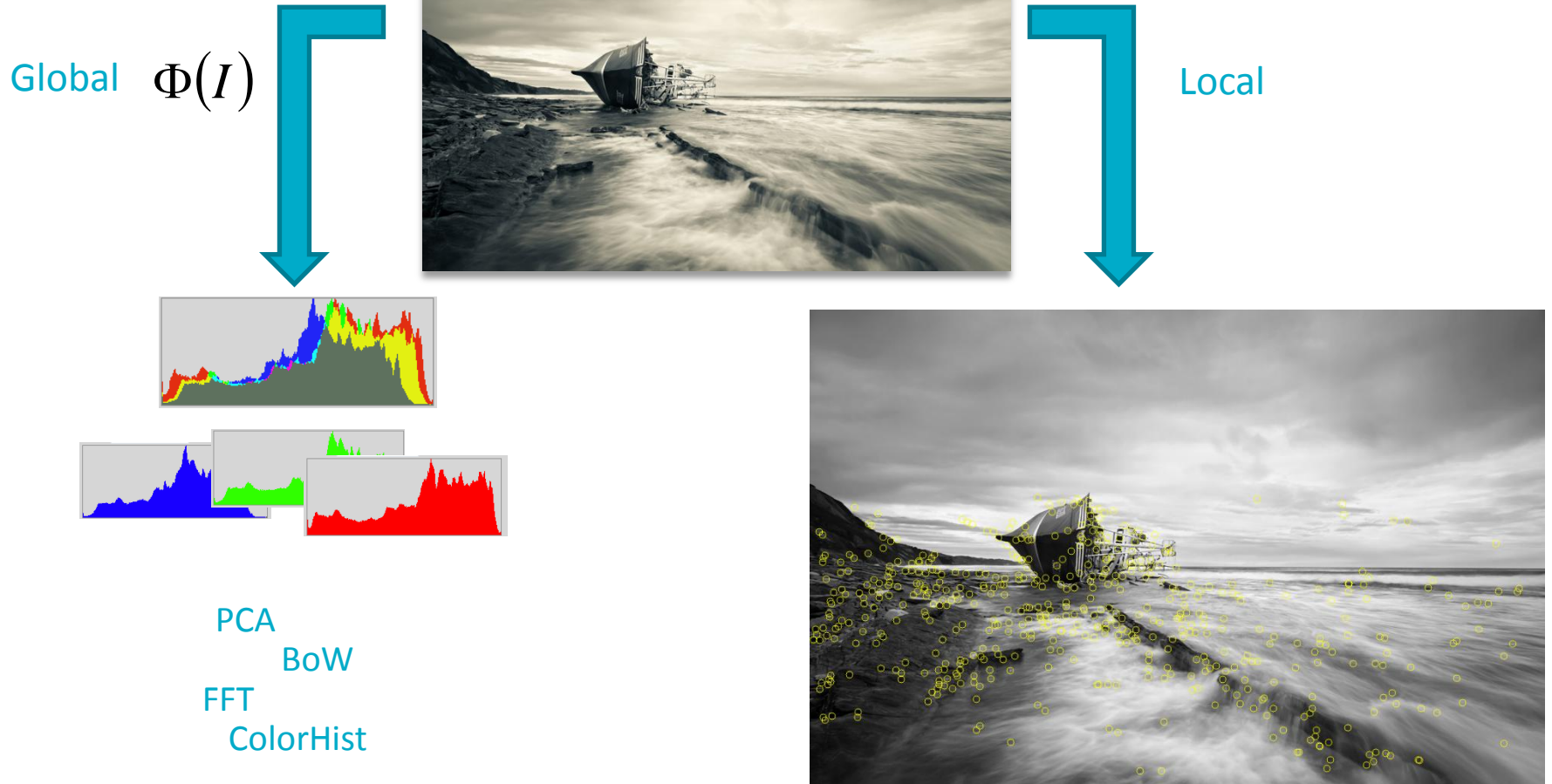
Summary

- 1. Introduction**
- 2. Data and Evaluation Framework**
- 3. Interest Point Extraction**
- 4. Feature Descriptors**
- 5. Feature Matching**
- 6. Conclusions**



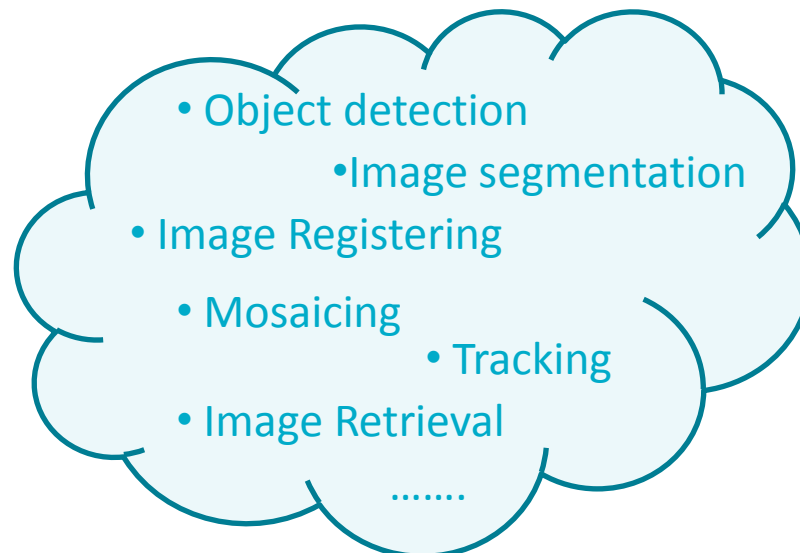
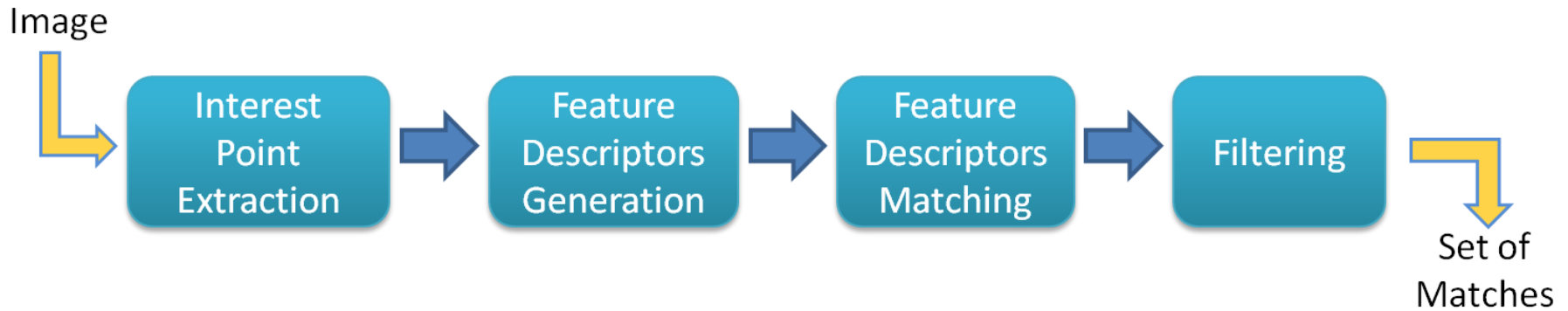
Global vs Local

1. Introduction
2. Data and Evaluation
3. Feature Extraction
4. Feature descriptors
5. Feature Matching
6. Conclusions



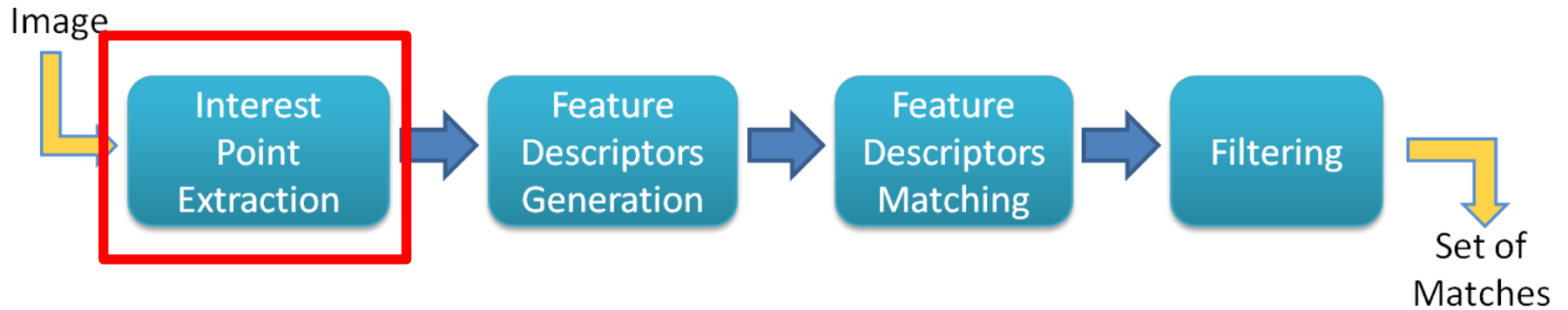
Processing Pipeline

1. Introduction
2. Data and Evaluation
3. Feature Extraction
4. Feature descriptors
5. Feature Matching
6. Conclusions



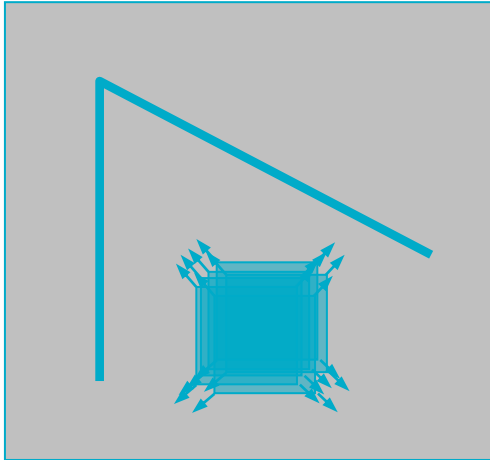
Processing Pipeline

1. Introduction
2. Data and Evaluation
3. Feature Extraction
4. Feature descriptors
5. Feature Matching
6. Conclusions

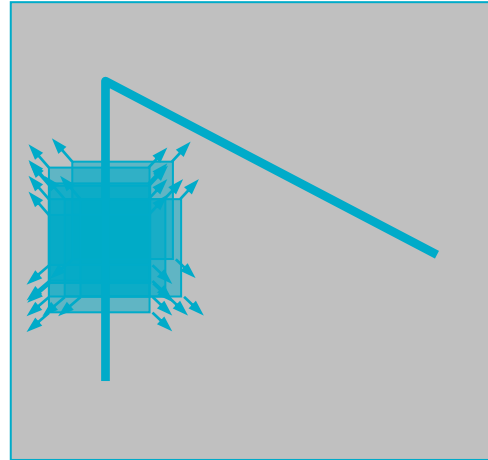


Interest Points

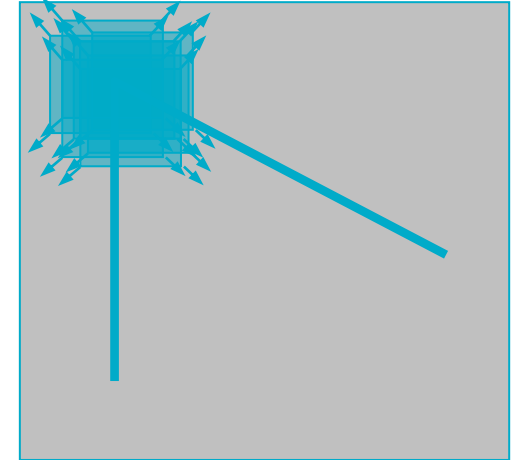
- Local Image structures around the interest point are rich in information content .



“flat” region:
no change in all
directions

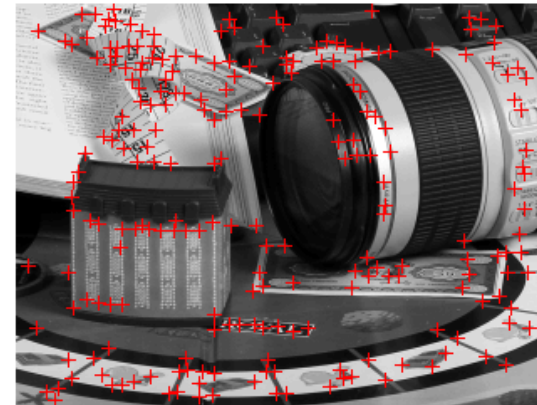


“edge”:
no change along the
edge direction



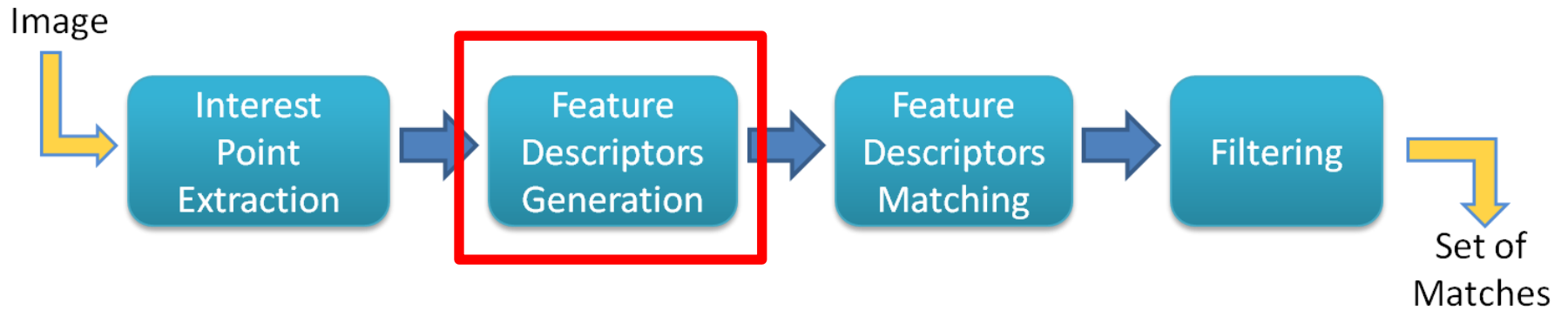
“corner”:
significant change in
all directions

Large local intensity changes (*Harris*)



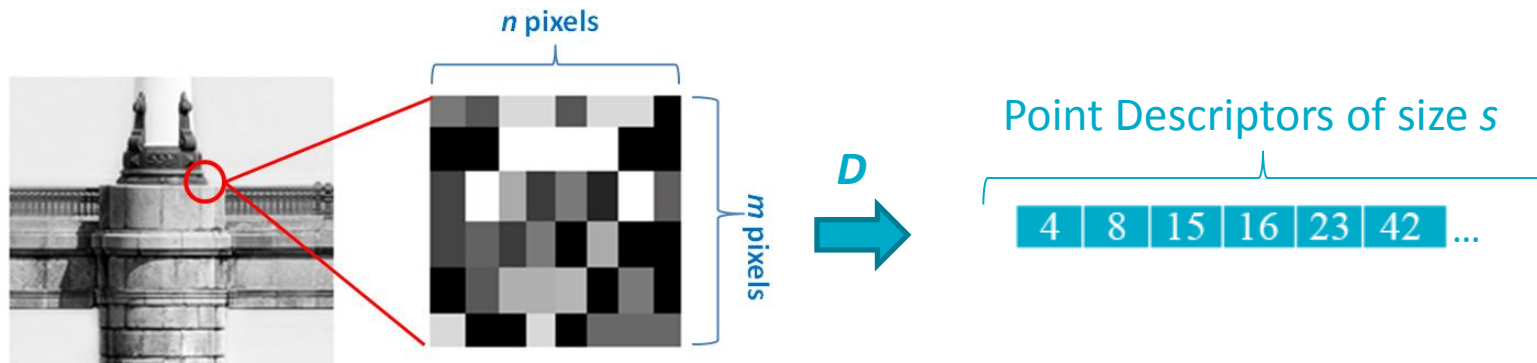
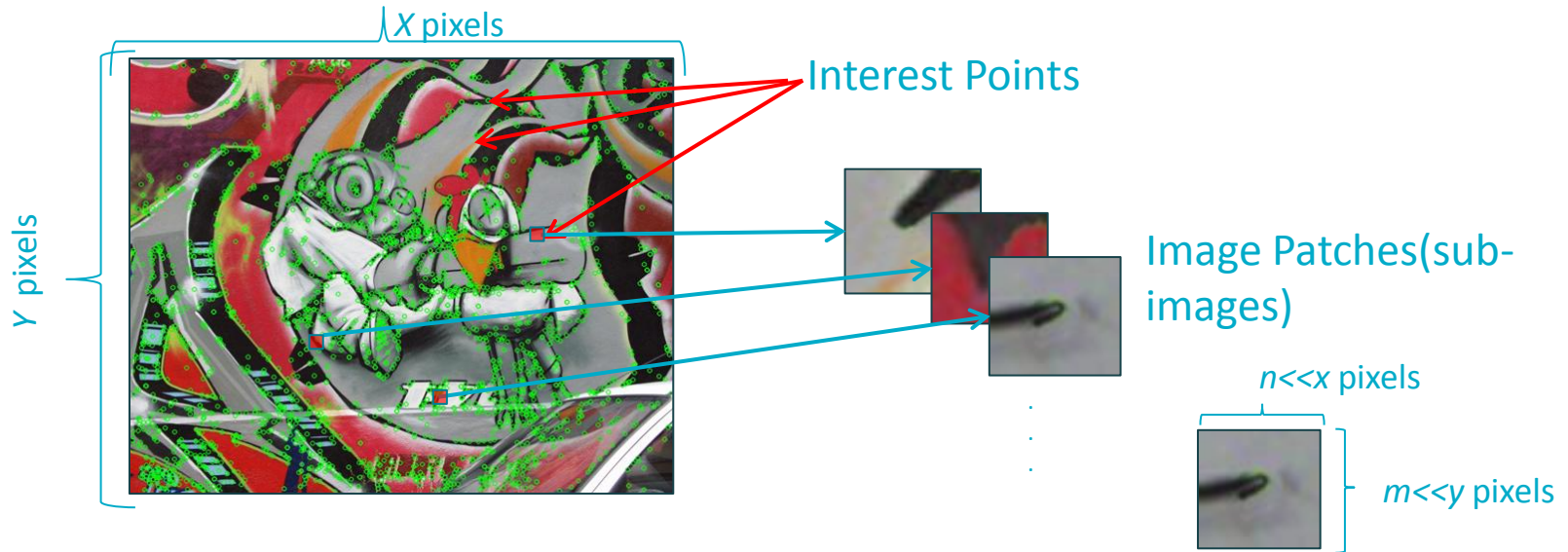
Processing Pipeline

1. Introduction
2. Data and Evaluation
3. Feature Extraction
4. Feature descriptors
5. Feature Matching
6. Conclusions



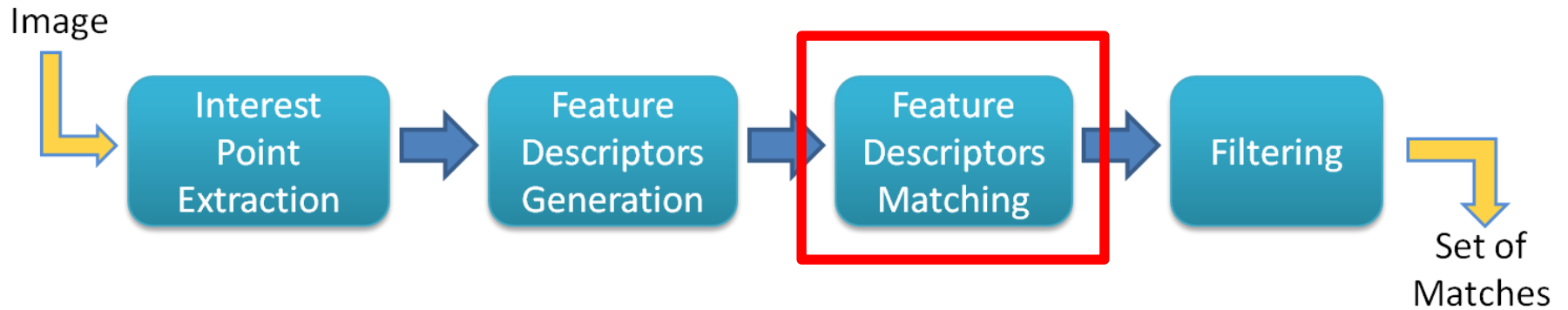
Feature Description

- Convert or represents image patches numerically
- Represent surrounding of interest point p_i univocally



Processing Pipeline

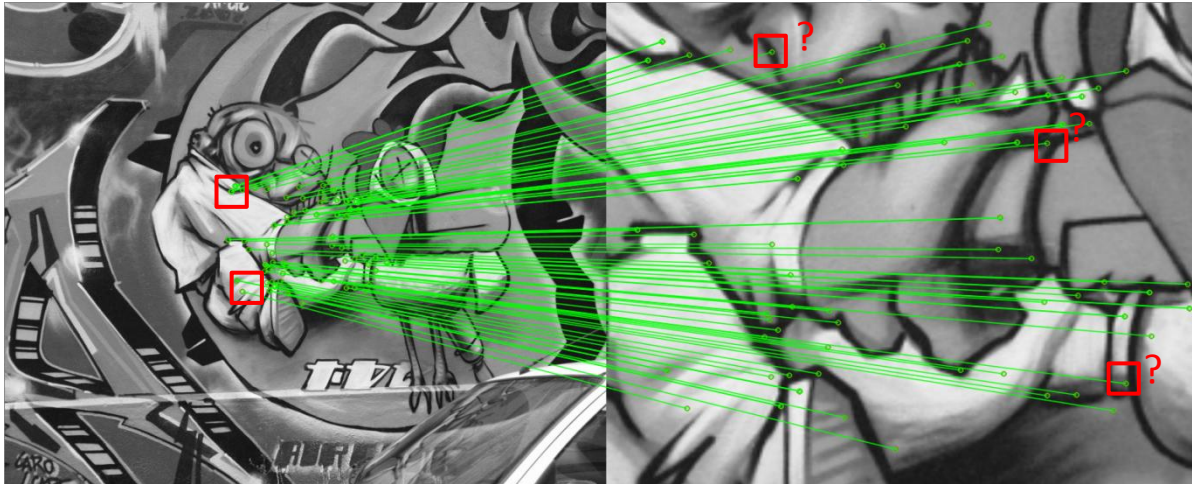
1. Introduction
2. Data and Evaluation
3. Feature Extraction
4. Feature descriptors
5. Feature Matching
6. Conclusions



Feature Matching

Image A

Image B

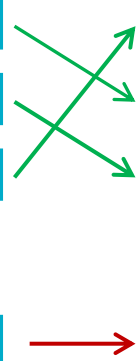


Descriptors A

Descriptors B

1	23	12	5	99
13	123	152	56	10
133	189	7	57	5
⋮				
4	8	15	16	23

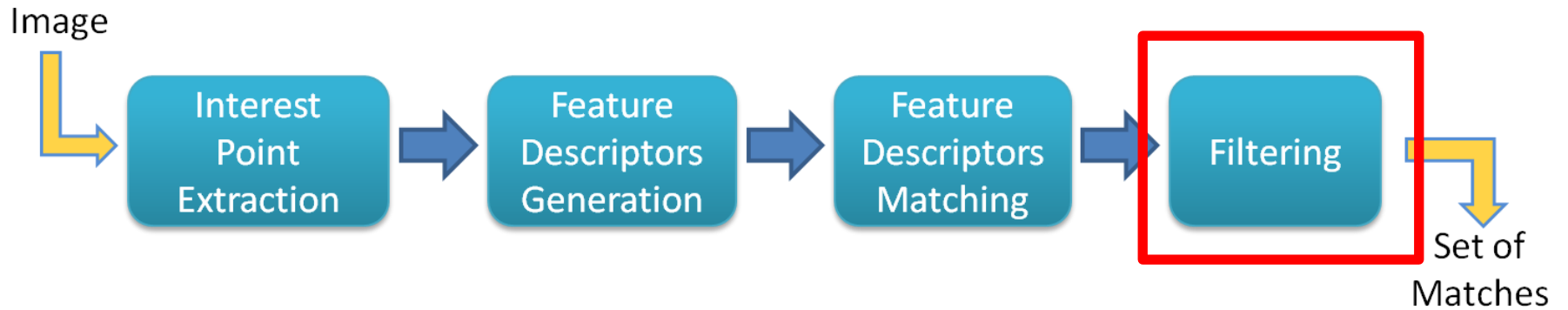
13	233	12	90	101
113	188	34	22	88
33	122	7	173	5
⋮				
42	6	99	65	107



- Brute – force (Euclidean, Hamming...)
- K- Nearest Neighbour
- Supervised Classification
- ...

Processing Pipeline

1. Introduction
2. Data and Evaluation
3. Feature Extraction
4. Feature descriptors
5. Feature Matching
6. Conclusions

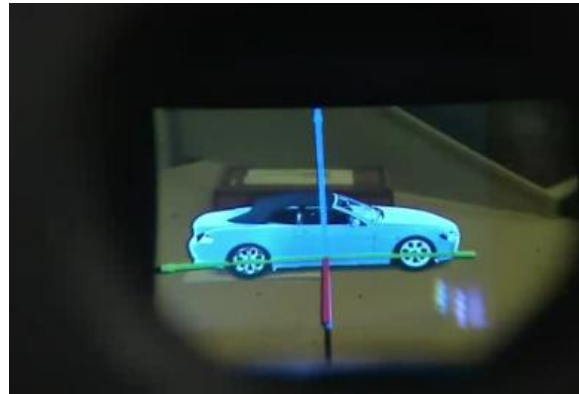


- Wrong Correspondences (Matching errors)
- Robust Estimation



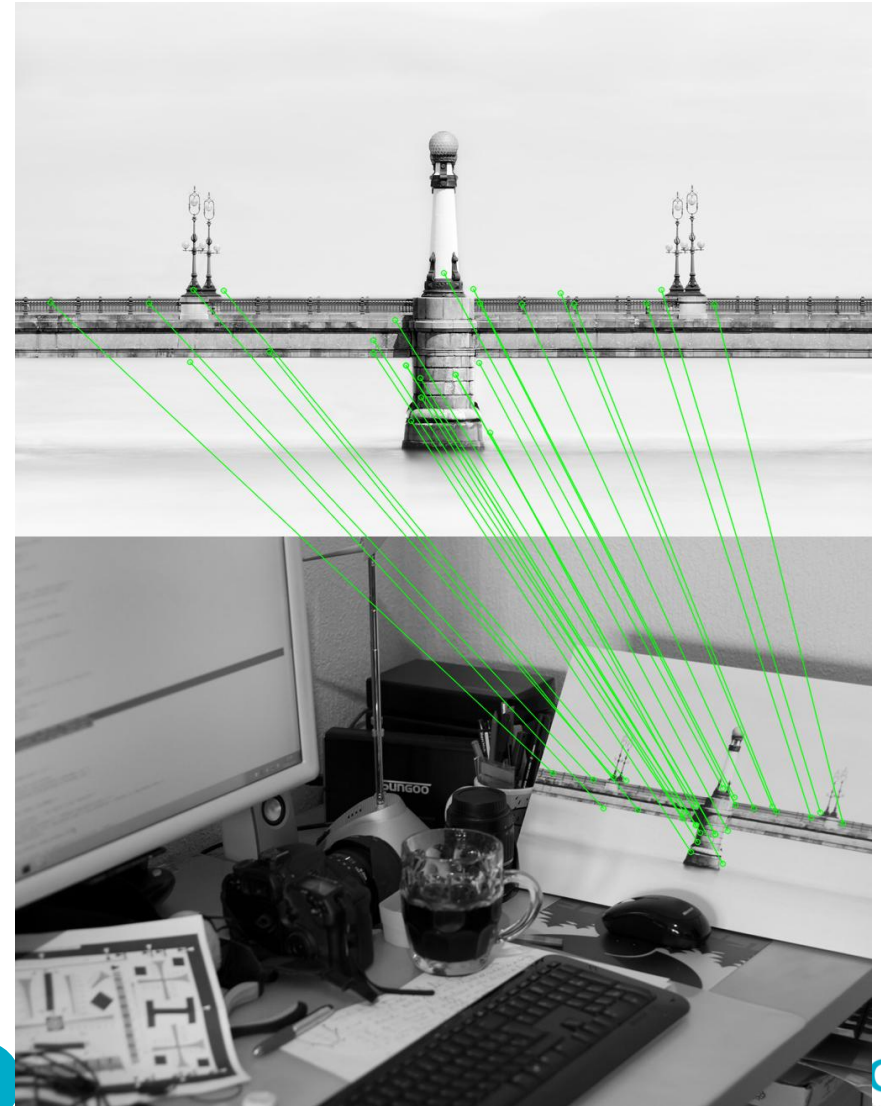
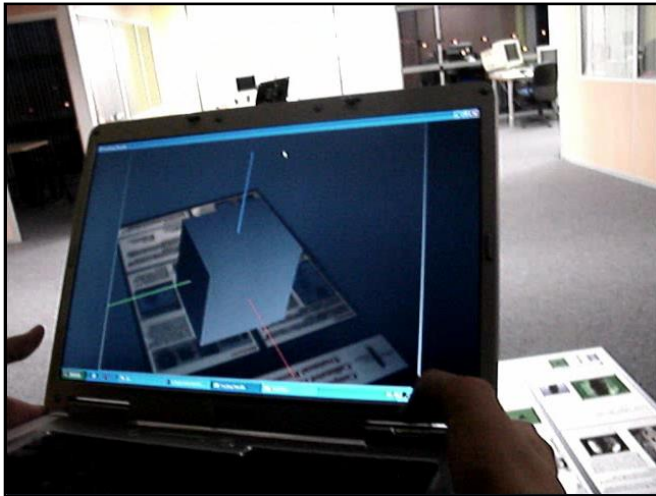
Application/Projects Samples

1. Introduction
2. Data and Evaluation
3. Feature Extraction
4. Feature descriptors
5. Feature Matching
6. Conclusions



Application/Projects Samples

1. Introduction
2. Data and Evaluation
3. Feature Extraction
4. Feature descriptors
5. Feature Matching
6. Conclusions



Application/Projects Samples

1. Introduction
2. Data and Evaluation
3. Feature Extraction
4. Feature description
5. Feature Matching
6. Conclusions



Dense Depthmap
extraction

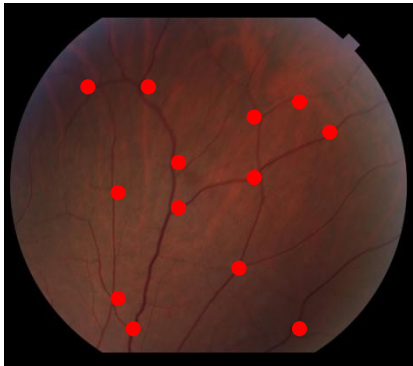
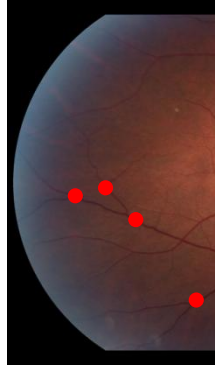
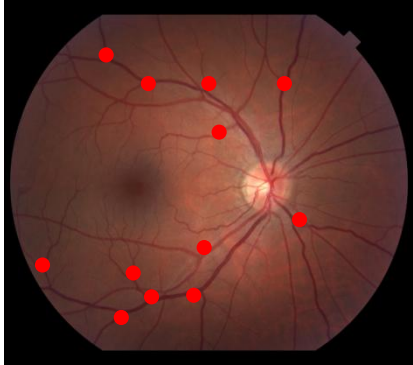


3D Reconstruction

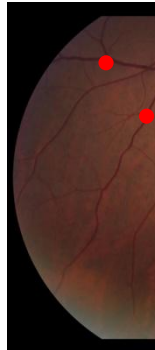


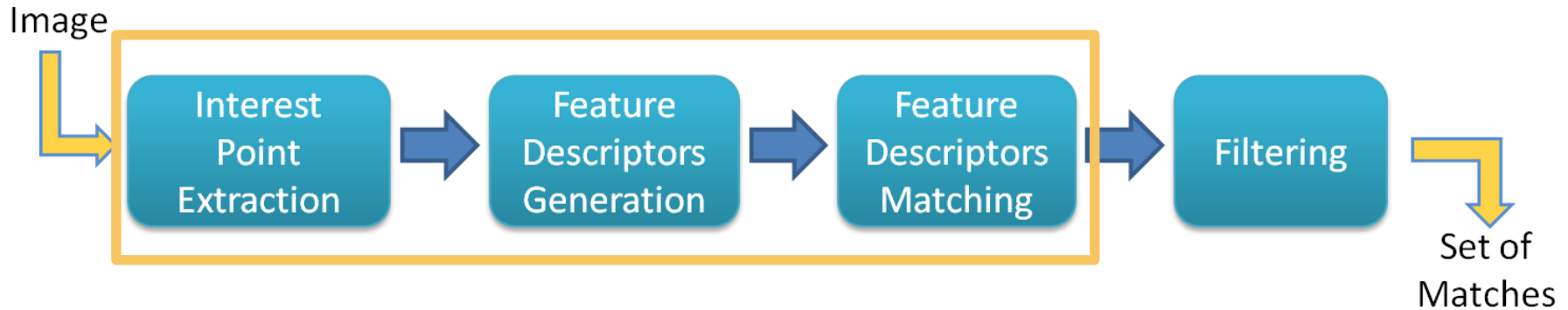
Application/Projects Samples

1. Introduction
2. Data and Evaluation
3. Feature Extraction
4. Feature descriptors
5. Feature Matching
6. Conclusions

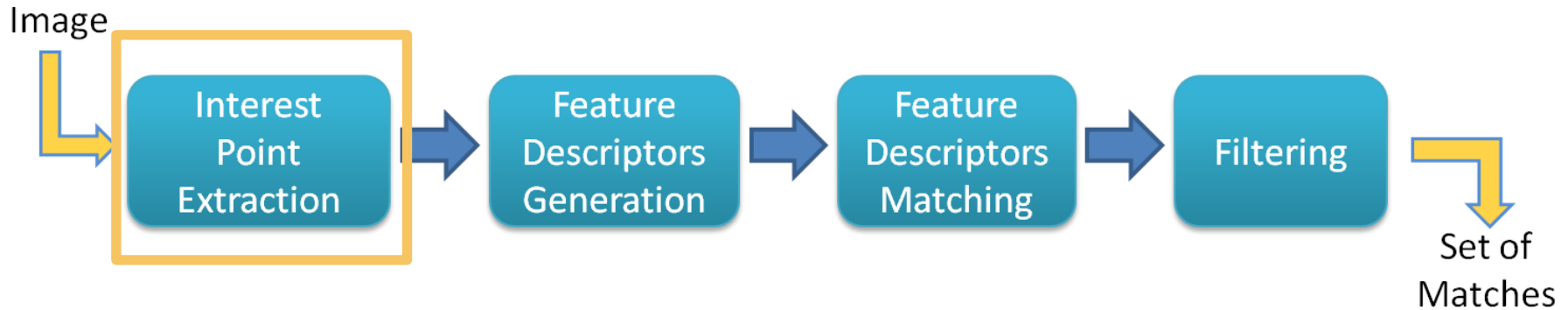


...



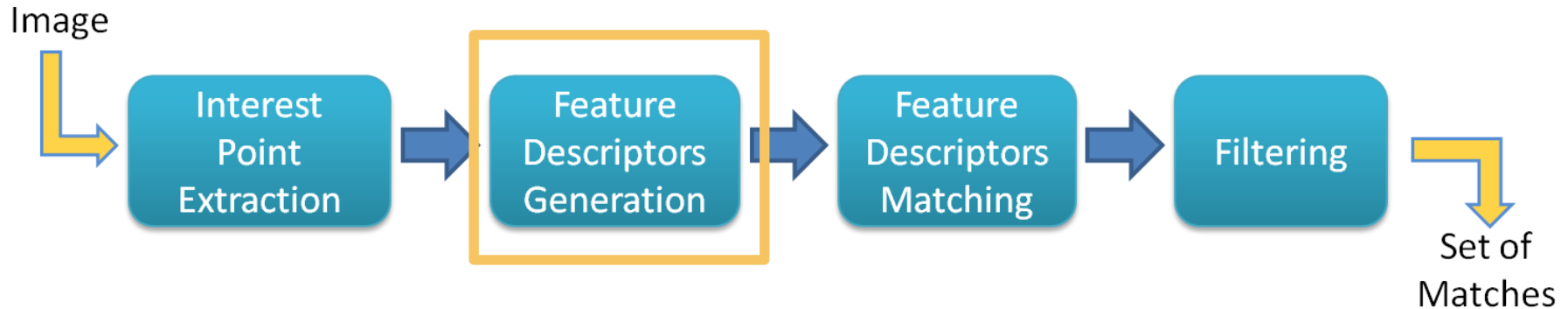


- **Data set and Evaluation Framework**
 - Normalized Images
 - Modular/Extensible design
 - Publicly available (www.vicomtech.tv/keypoints)

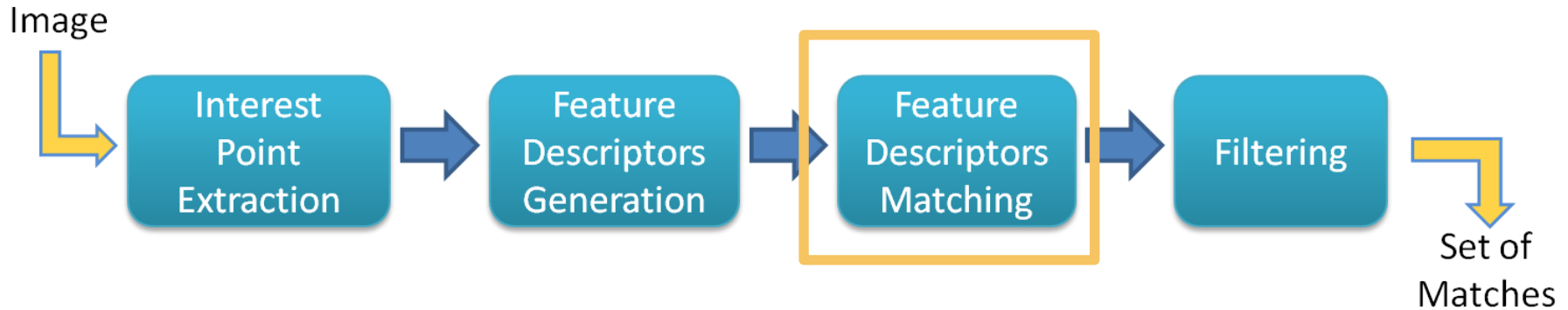


■ Interest Point Evaluation

- State-of-the-art Local Interest Point extraction
- Evaluation against geometric and photometric transformations

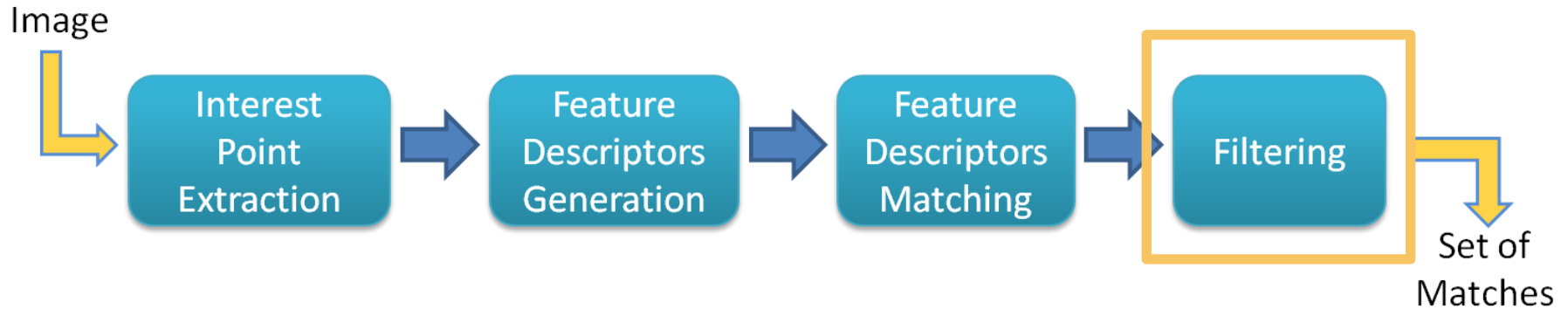


- **Proposal of DITEC Local Descriptor**
 - Local descriptor based on Trace Transform
- **Evaluation**
 - State-of-the art descriptors review
 - Evaluation against geometric and photometric transformations



▪ Supervised classification descriptor Matching

- Random Forest Ensemble
- Tracking by Detection



- **Robust Homography Estimation**
 - Evaluation of Random Sampling approaches

1. Introduction

2. Data and Evaluation Framework

3. Interest Point Extraction

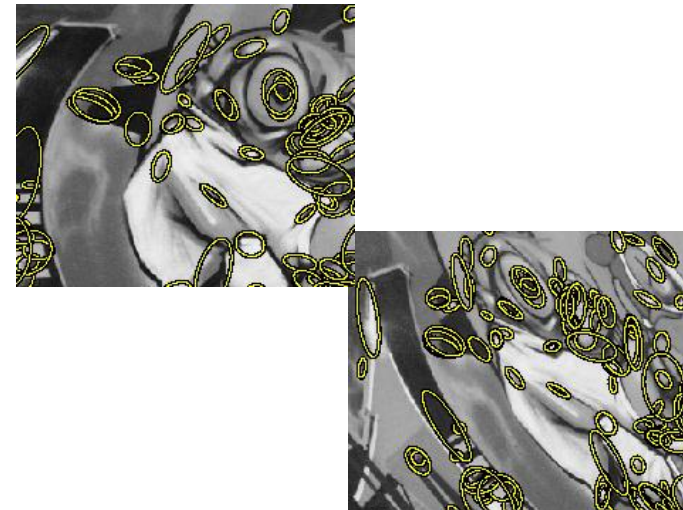
4. Feature Descriptors

5. Feature Matching

6. Conclusions

- Our Approach

- Based on Mikolajczyk framework(De facto standard)
- Evaluation of detectors, descriptors and matching strategies.
- Repeatability and matching accuracy measurements
- Inlier/Outlier detection (Ransac)
- OpenSource



- Ground Truth
 - Point/region Transfer via Known Homography H
 - Rectified Images



Evaluation Framework

- Measures/Performance

- Repeatability measurement

- Transfer error $d_{ij} = d(x_{jb}, Hx_{ia})^2 + d(x_{ia}, H^{-1}x_{jb})^2$

- Overlapping error $\epsilon_s \leq 1 - \left(\frac{R_{ia} \cap H^T R_{jb} H}{R_{ia} \cup H^T R_{jb} H} \right)$

- Matching Accuracy (descriptors)

- Computation Times.

Public Data Sets

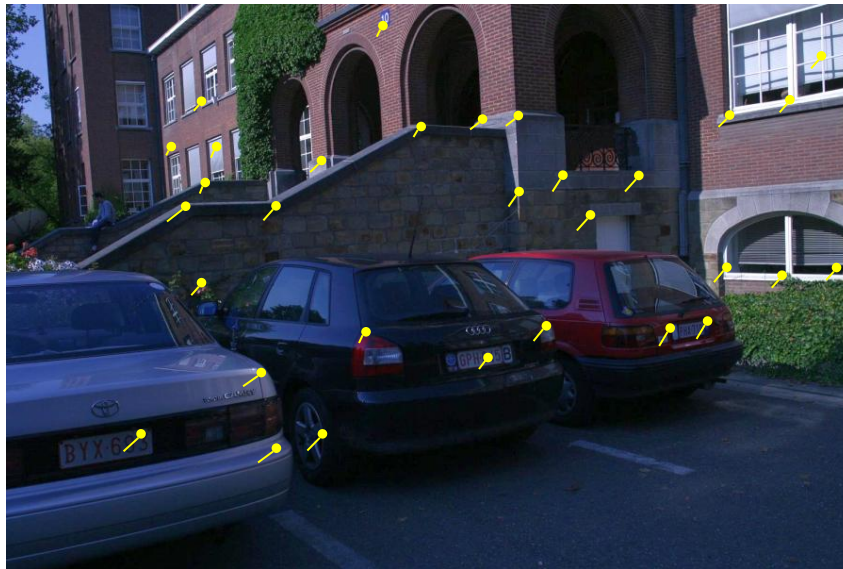
1. Introduction
2. Data and Evaluation
3. Feature Extraction
4. Feature descriptors
5. Feature Matching
6. Conclusions



<http://www.robots.ox.ac.uk/~vgg/research/affine/>

Public Data Sets

- Small Number of images.
- Geometric and Photometric transformations are not isolated.



Geometric Transformations

1. Introduction
2. Data and Evaluation
3. Feature Extraction
4. Feature descriptors
5. Feature Matching
6. Conclusions



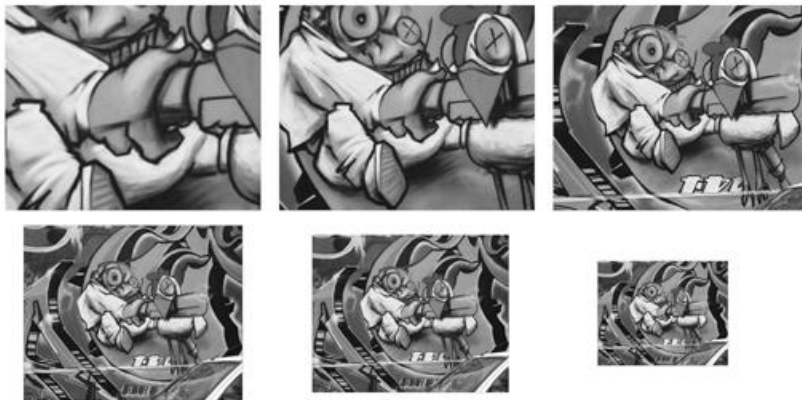
Geometric Transformations

1. Introduction
2. Data and Evaluation
3. Feature Extraction
4. Feature descriptors
5. Feature Matching
6. Conclusions

• Projective Transformation



• Scale Transformation

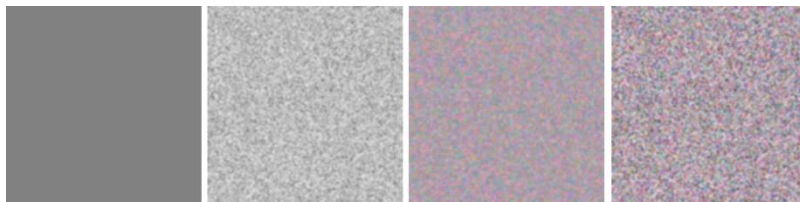


• Rotation Transformation



Photometric transformations

- Change amount/Intensity of light
- Fixed camera position/orientation
- Fixed acquisition parameters
- Noise (SNR)



Data Sets (real acquired images)

1. Introduction
2. Data and Evaluation
3. Feature Extraction
4. Feature descriptors
5. Feature Matching
6. Conclusions

- Light Intensity



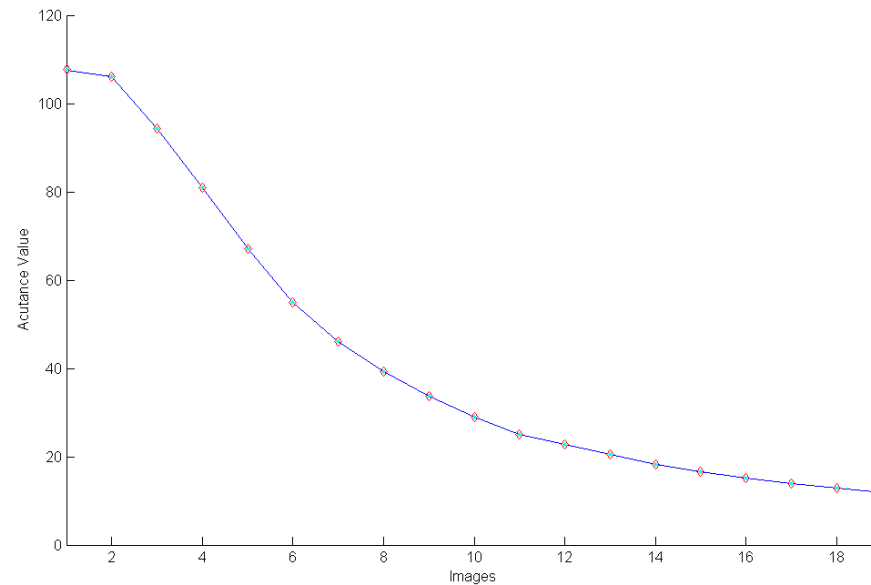
- Noise



Data Sets (real acquired images)

1. Introduction
2. Data and Evaluation
3. Feature Extraction
4. Feature descriptors
5. Feature Matching
6. Conclusions

– Focus



1. Introduction

2. Data and Evaluation Framework

3. Interest Point Extraction

4. Feature Descriptors

5. Feature Matching

6. Conclusions

Interest Point Extraction

Image A



Image B



No possible Match!

Detected Interest Points

- Repeatability/stability
- Efficiency

Contribution

- Review state-of-the-art in Interest Point Extraction
- Evaluation
 - Repeatability
 - Geometric transformations
 - Photometric transformations
 - Computation Time
 - Density

Evaluation

• Evaluation Methodology

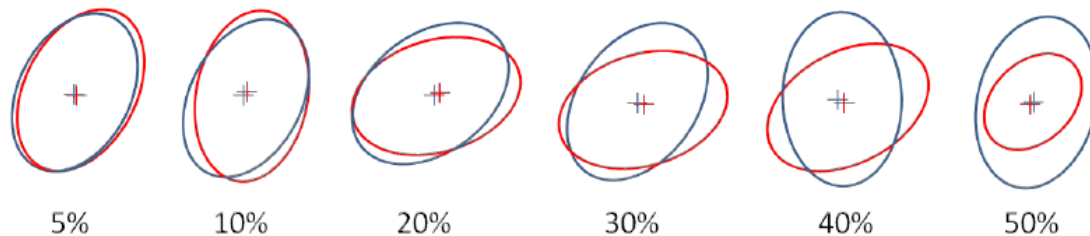
– Inputs

- A Set of normalized images $I = \{I_1, I_2, \dots, I_n\}$
- A set of Algorithms $A = \{A_1, A_2, \dots, A_w\}$
- A set of bijective functions $S_0 = \{f_{1,2}, f_{1,3}, \dots, f_{i,j}\}$ such that $f_{i,j} : I_i \rightarrow I_j$ representing the Ground Truth Data (Homographies)

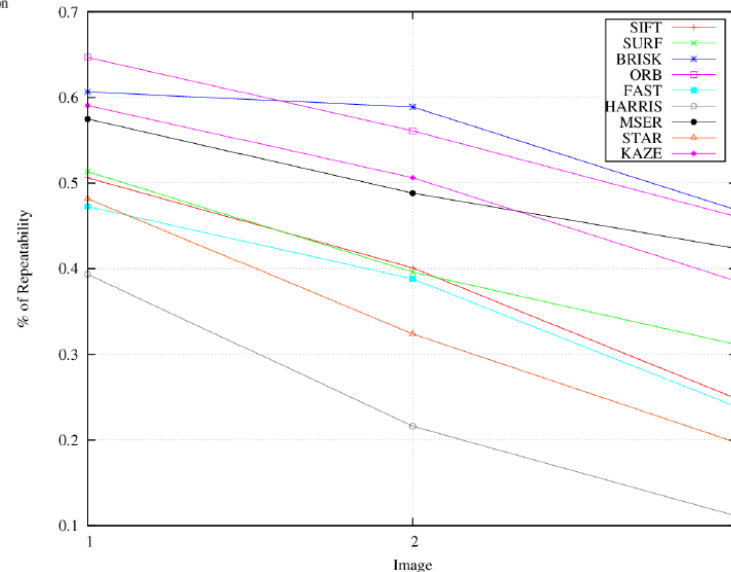
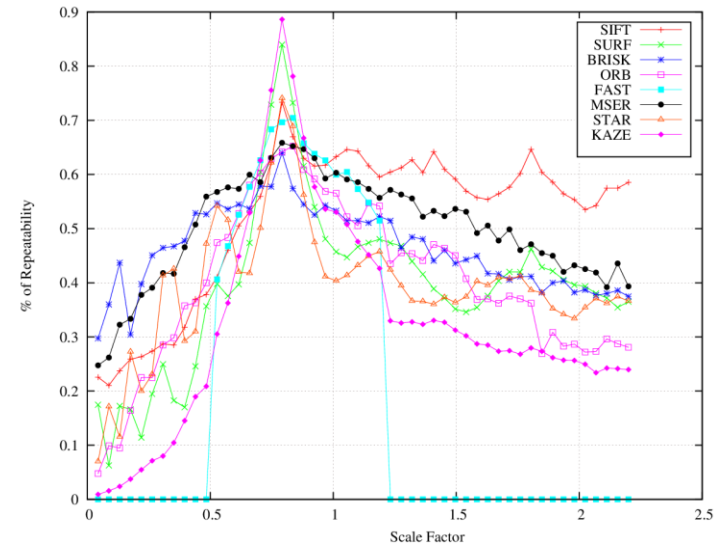
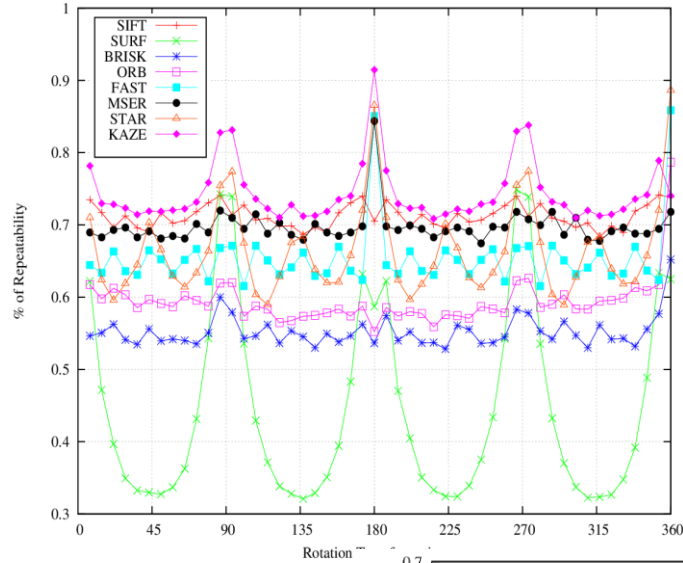
– OutPut

- A set of performance measures for each algorithm $A_w (1 \leq m \leq w)$ grading the quality of A_w against the ground truth data.

$$\text{repeatabilityScore} = \frac{\text{numberOfTrueMatches}}{\text{numberOfDetectedPoints}} \quad \text{TrueMatch} = \left\{ \begin{array}{ll} 1 & \text{if } (d_{ij} < t) \& (\epsilon_s < r) \\ 0 & \text{otherwise} \end{array} \right\}$$

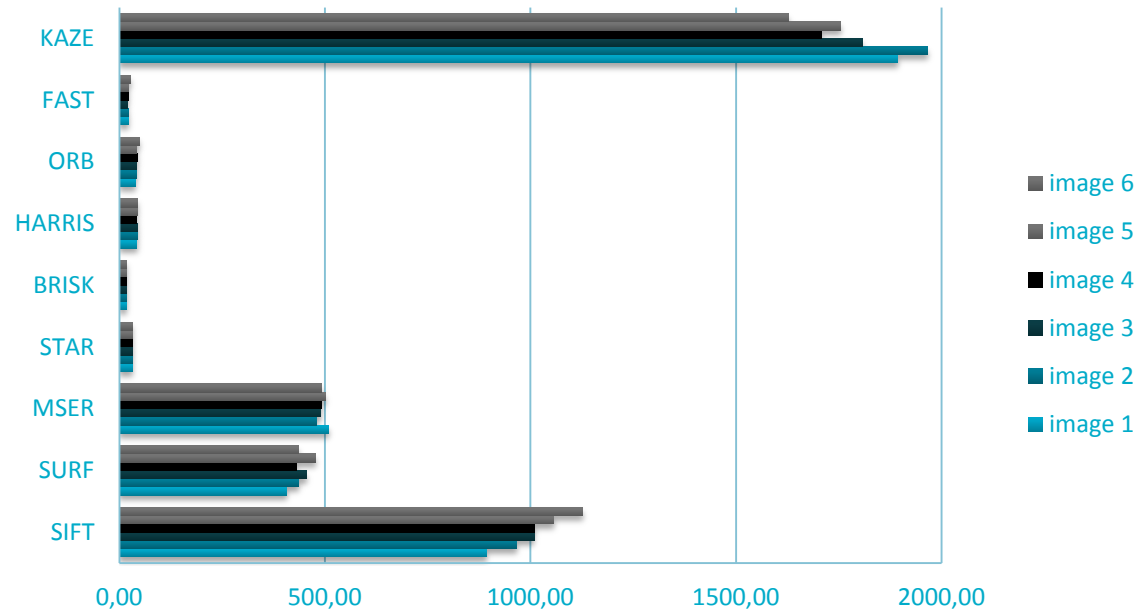


● Geometric Transformations



Evaluation

- Computation Time



Conclusions

- Detectors with scale estimation are more robust to affine and projective transformation.
- FAST/BRISK better for Real-Time processing.
- FAST Not robust to severe image transformation(scale, rotation or projective).
- None is projective invariant but MSER.
- MSER needs stable, big, homogeneous regions.
- SIFT and BRISK are very robust to rotation and scale transformation.
- SIFT shows great performance but is slow.
- BRISK show the best compromise between accuracy and computational cost



1. Introduction

2. Data and Evaluation Framework

3. Interest Point Extraction

4. Feature Descriptors

5. Feature Matching

6. Conclusions

Feature Description

Image A

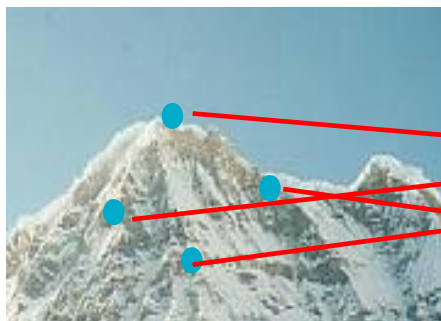


Image B



Descriptors Image A

10	12	0	6	34
----	----	---	---	----

.

.

Descriptors Image B

40	2	12	60	43	27
----	---	----	----	----	----

.

.

- Feature description
 - Retain local information.
 - Describe region univocally.
 - Robust to geometric or photometric transformations.

- New proposal for local feature description
 - Based on the Trace Transform
 - A generalization of RADON

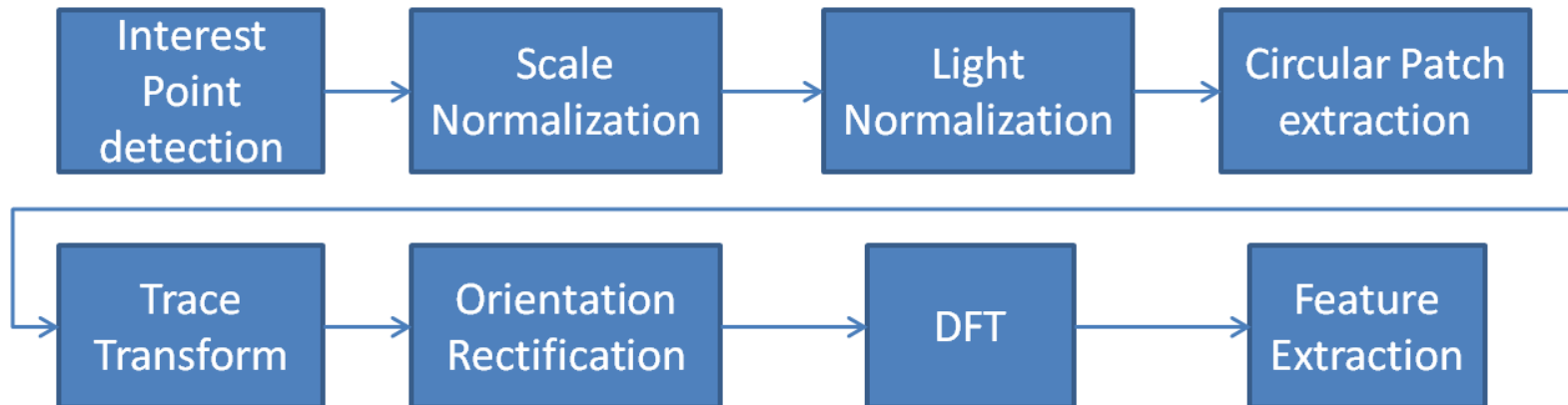
Radon Transform

$$R(\phi, \rho) = \iint f(x, y) \delta(x \cos \phi + y \sin \phi - \rho) dx dy$$

<i>IF1</i>	$\int \xi(t) dt$
<i>IF2</i>	$(\int \xi(t) ^q dt)^r$
<i>IF3</i>	$\int \xi(t)' dt$
<i>IF4</i>	$\int (t - (\int t \xi(t) dt / IF1))^2 \xi(t) dt$
<i>IF5</i>	$(IF4 / IF1)^{1/2}$
<i>IF6</i>	$\max(\xi(t))$
<i>IF7</i>	$IF6 - \min(\xi(t))$
<i>IF8</i>	Amplitude of 1st harmonic of $\xi(t)$
<i>IF9</i>	Amplitude of 2nd harmonic of $\xi(t)$
<i>IF10</i>	Amplitude of 3rd harmonic of $\xi(t)$
<i>IF11</i>	Amplitude of 4th harmonic of $\xi(t)$

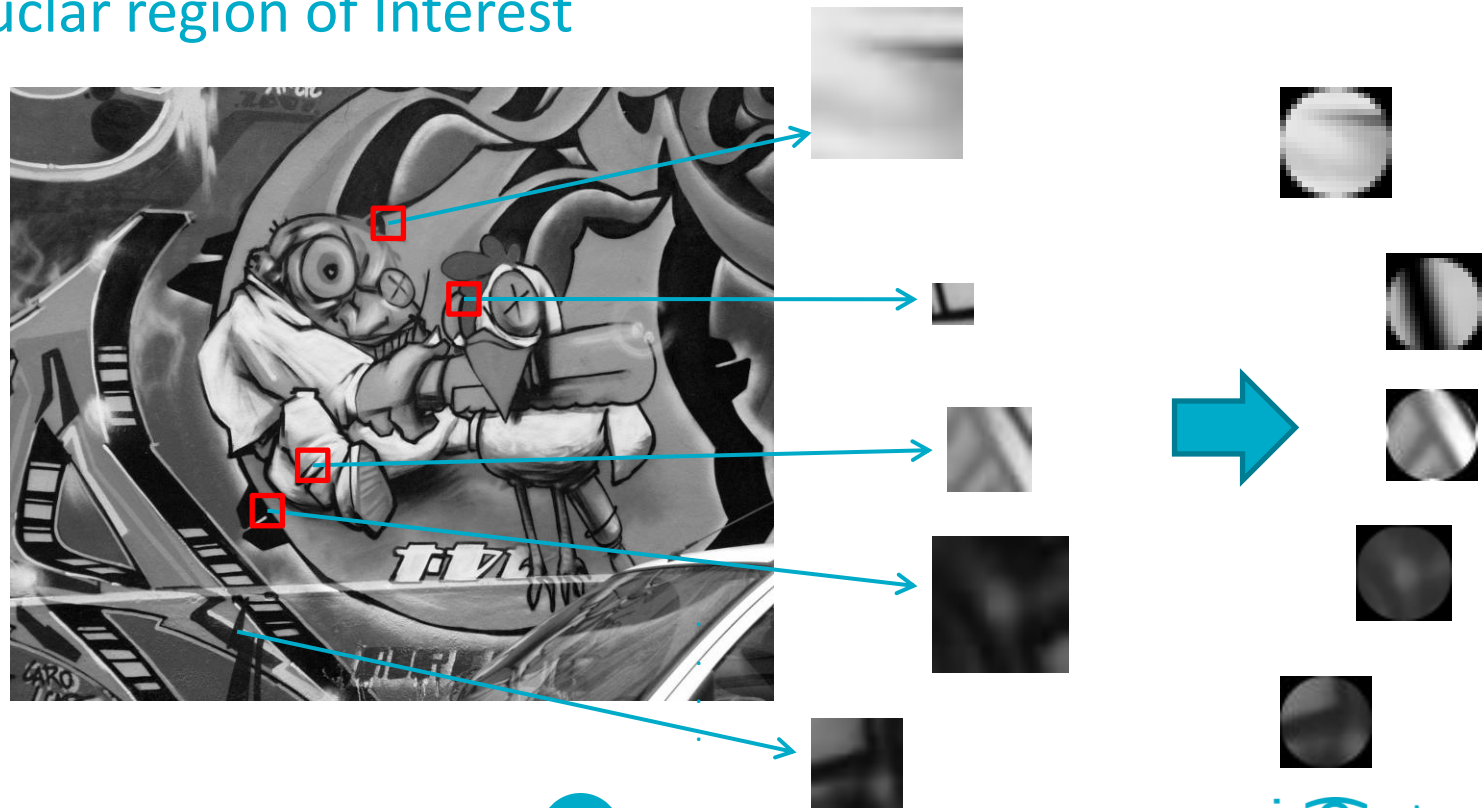
Local DITEC

- Proposed pipeline for DITEC descriptor
 - Interest point extraction is carried out by an external mechanism.
 - Interest Point detector with scale estimation.

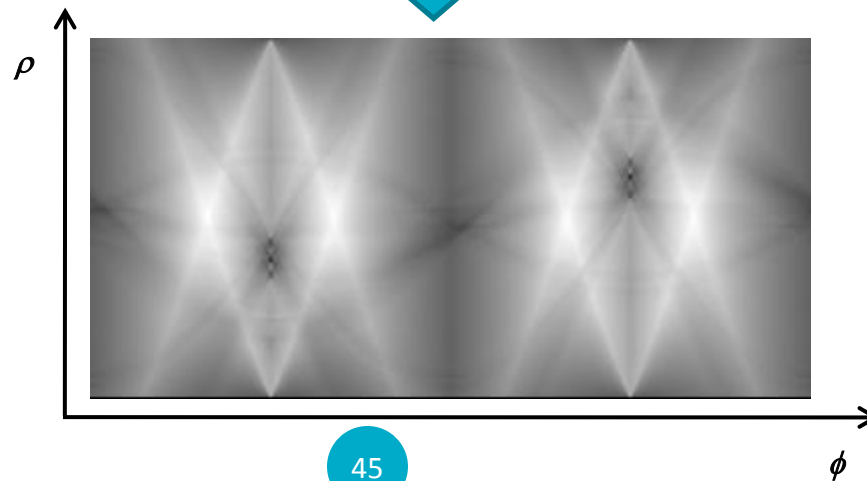
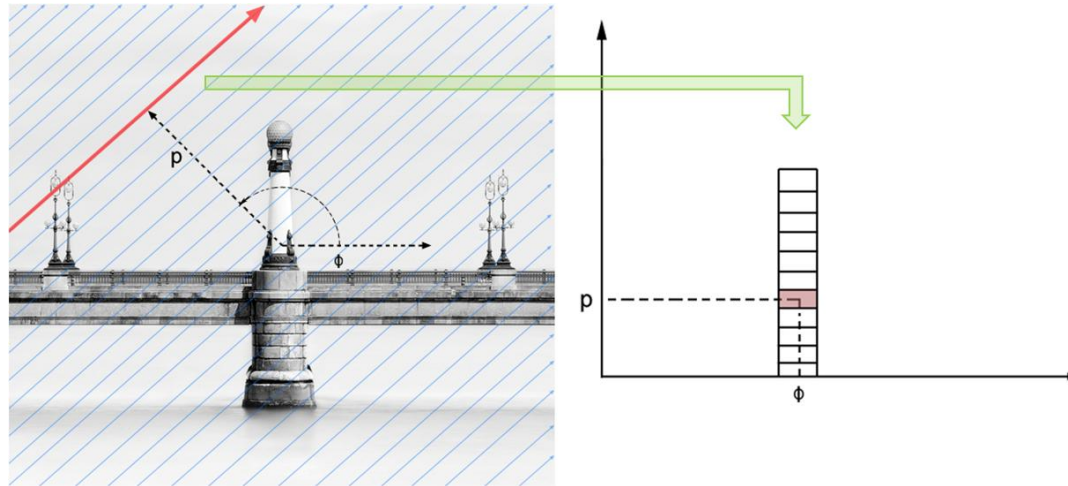


Local DITEC

- Patch Extraction
 - Scale Normalization
 - Histogram normalization
 - Circular region of Interest

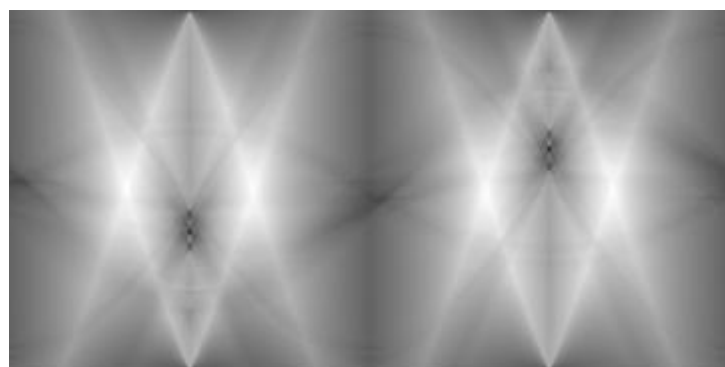


- Trace Transform

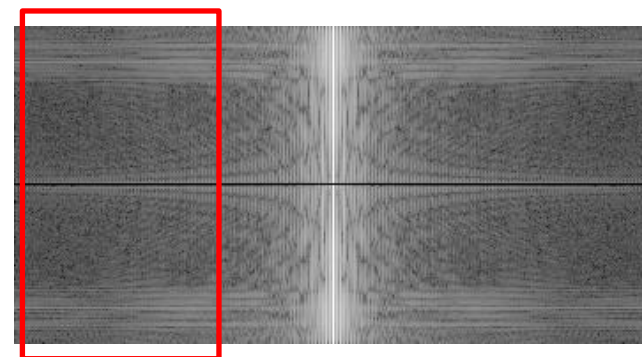


Local DITEC

- Discrete Fourier Transform



DFT ROWS



Modulus



DC Removal

4	8	15	16	23	42	...
---	---	----	----	----	----	-----



Normalization (*Lowe*)

0.1	0.1	0.92	0.18	0.27	0.55	...
-----	-----	------	------	------	------	-----

Parameter Analysis

- Local DITEC parameters
 - Trace transform computation (ϕ, ρ)
 - Size of local image patches
 - Descriptor Dimensionality
- Definition of descriptor parameters
 - Parameter sensitivity.
 - Define parameters ranges.
 - Influence of parameters in output (matching accuracy)

Evaluation

- Evaluation Methodology

- Matching accuracy
- Filter Correspondences through Known Homography
- 1-NearestNeighbour

$$Accuracy = \frac{TrueCorrectMatches}{Correspondences}$$

Image A



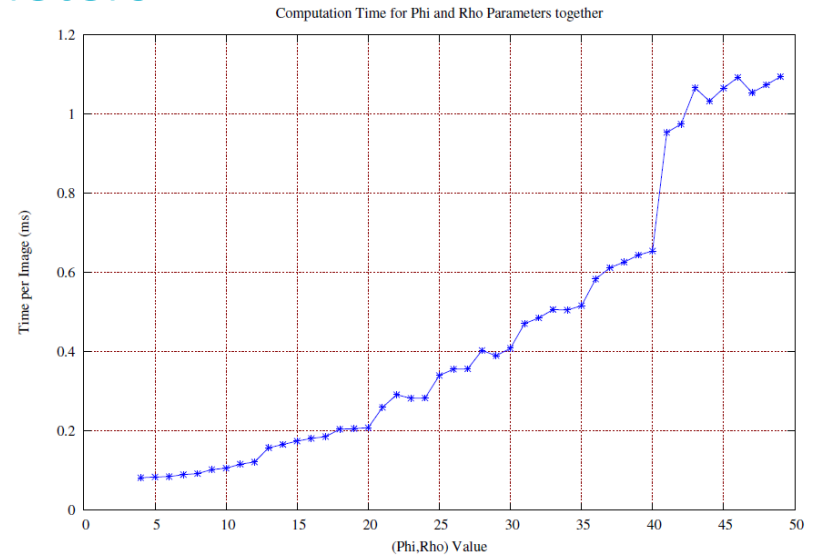
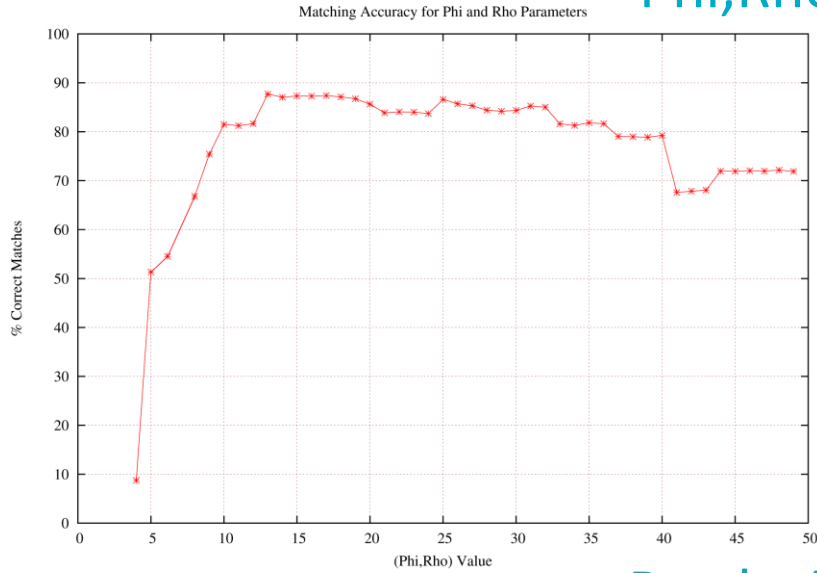
Image B



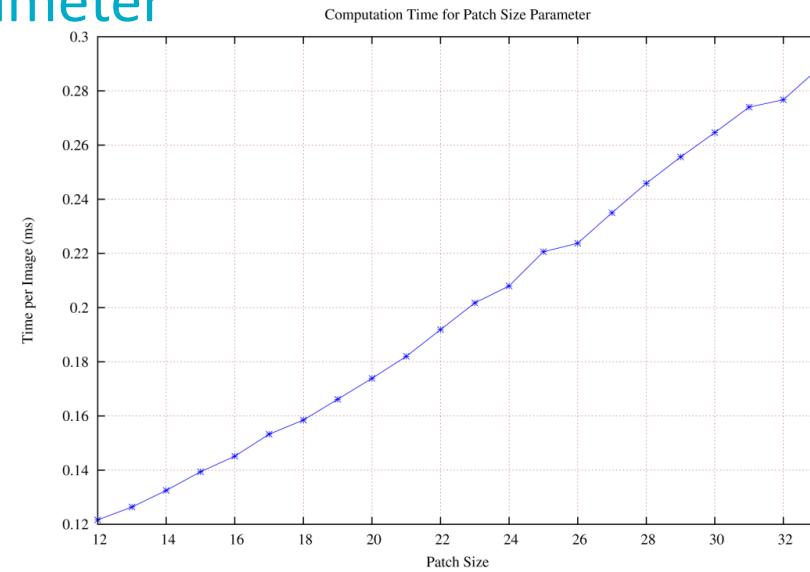
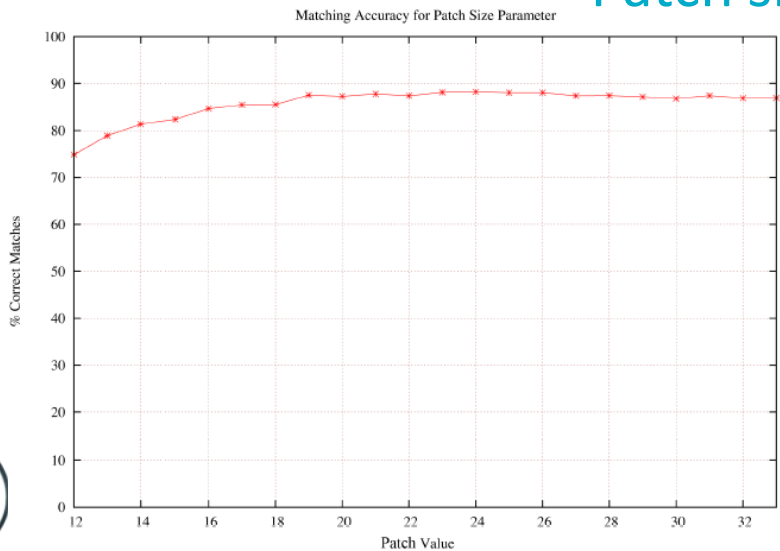
Parameter Analysis

1. Introduction
2. Data and Evaluation
3. Feature Extraction
4. Feature descriptors
5. Feature Matching
6. Conclusions

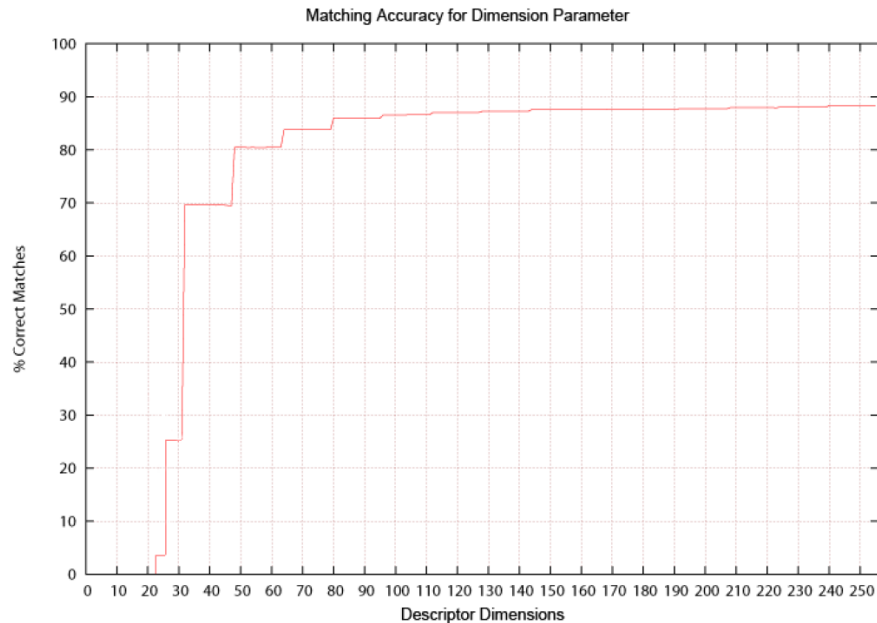
Phi,Rho Parameters



Patch size Parameter



Descriptor Dimensionality



- Sensitivity index

$$SI = \frac{D_{max} - D_{min}}{D_{max}}$$

<i>Parameter</i>	<i>SI</i>
Number of Samples in Phi	0.386
Number of Samples in Rho	0.331
Descriptor Dimensionality	0.714
Size of local image Patch	0.138

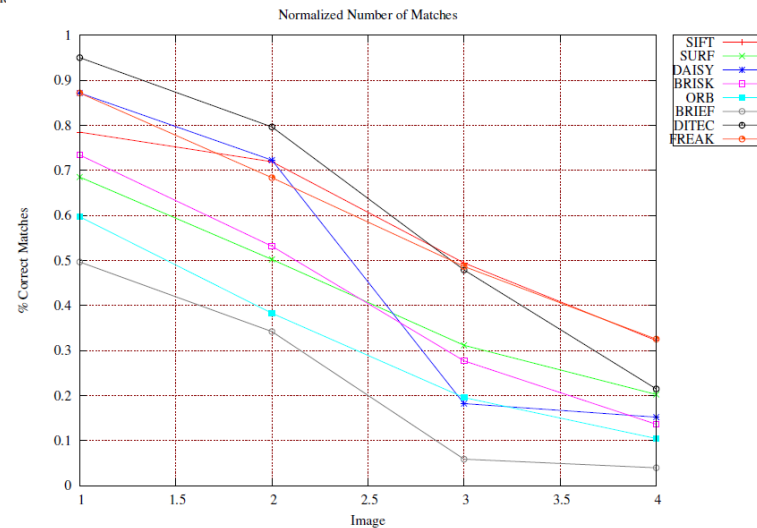
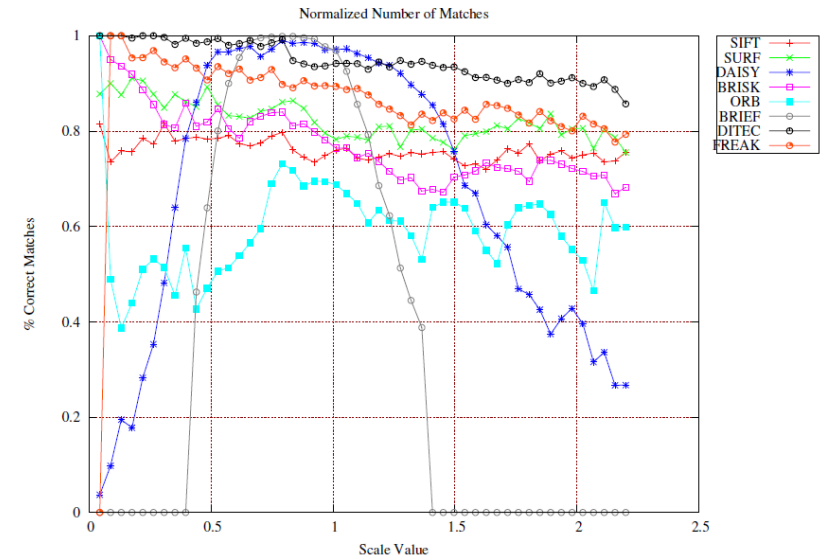
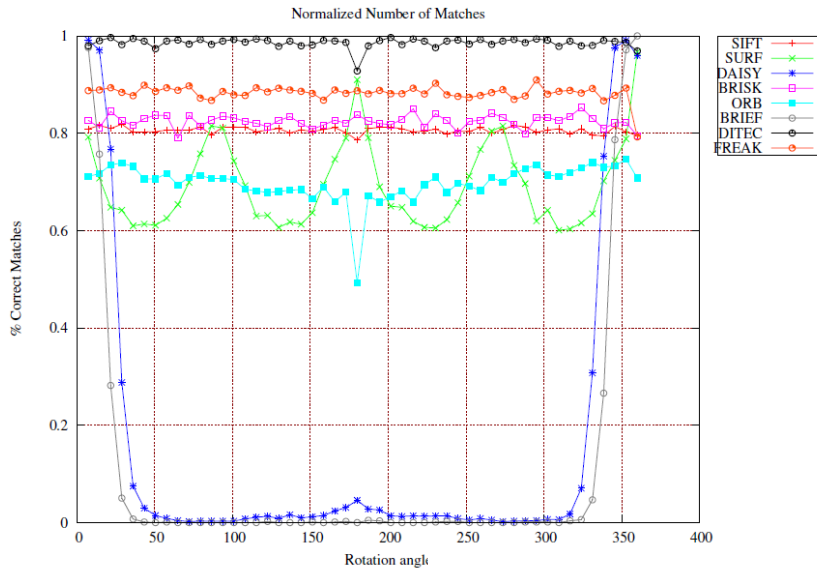


Conclusions

- Descriptor dimensionality is the most important parameter.
- Dimensionality between 100 and 140 shows best results.
- Phi and Rho can be set as a single parameter.
- Patch size can be set automatically (24px for a 1024px image)



● Geometric Transformations



Evaluation

- Computation Times
 - 10 runs of 1000 descriptors
 - 2Ghz single Core.

Descriptor	Extraction (ms)	Extraction+Matching (ms)
DAISY	0,86	0,89
SIFT	0,113	0,174
DITEC	0,112	0,173
SURF	0,06	0,11
FREAK	0,05	0,099
ORB	0,017	0,044
BRISK	0,015	0,069
BRIEF	0,008	0,036

Conclusions

- DITEC is robust against rotation transformation without using dominant orientation estimation.
- DITEC shows great discriminant capabilities.
- BRIEF shows poor performance regarding scale and rotation transformation.
- BRIEF is the less computational demanding approach.
- FREAK is the best trade-off between computational resources and performance.
- Binary descriptors are preferable.



Summary

1. Introduction

2. Data and Evaluation Framework

3. Interest Point Extraction

4. Feature Description

5. Feature Matching

6. Conclusions



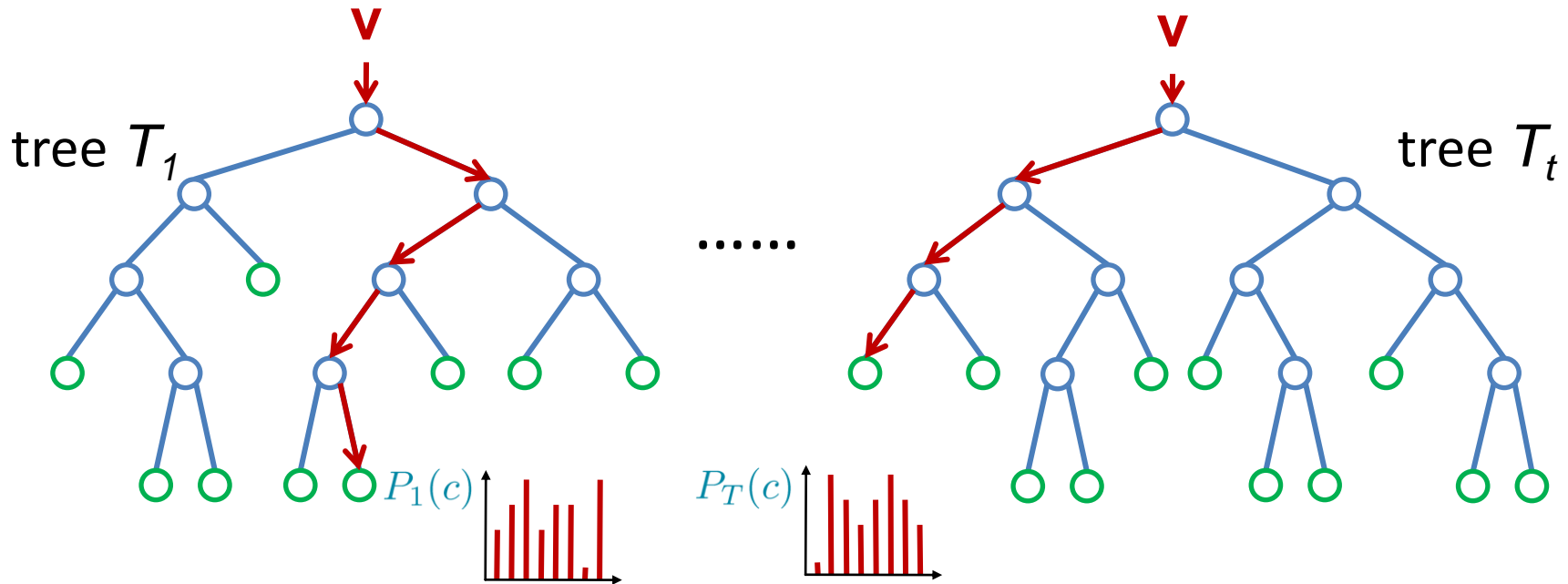
Feature Matching: Supervised Classification

- Matching Features as a classification problem
 - Supervised classification
 - No descriptors generation
 - Learning stage needed
 - Tracking by detection
- Random Forest
 - Multiclass support
 - Fast Training
 - Parallelizable



Random Forest

- Forest is an ensemble of several decision trees

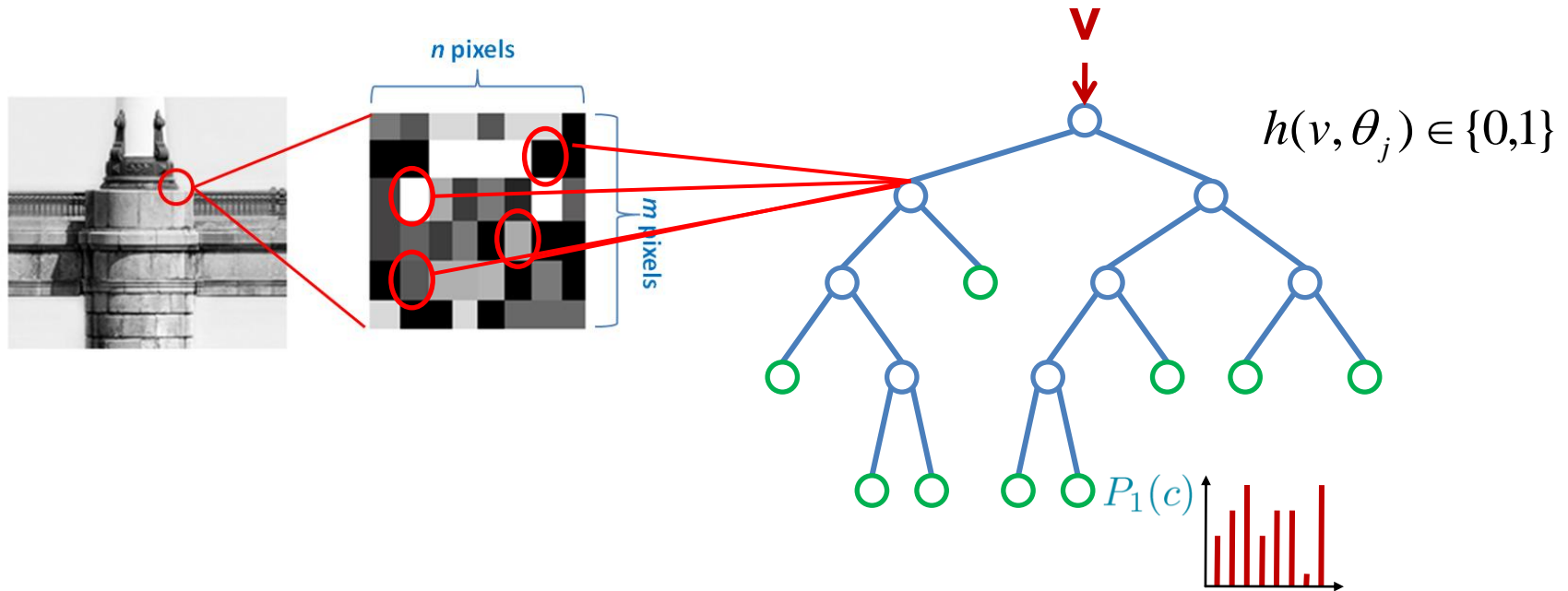


- Combining trees outputs (averaging class posteriors)

$$P(c | v) = \frac{1}{T} \sum_{t=1}^T P_t(c | v)$$

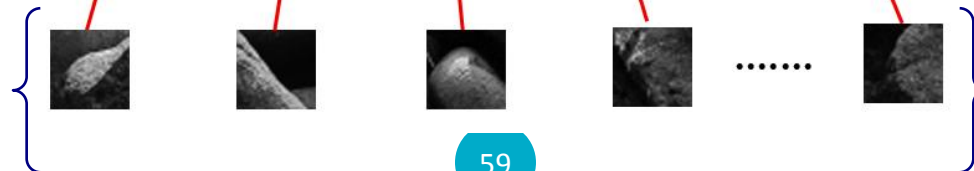
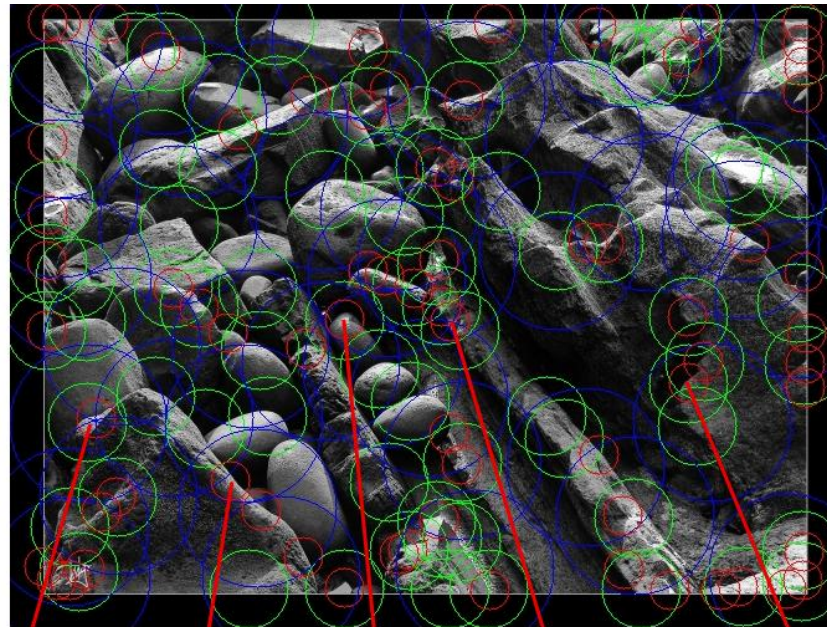
Supervised Feature Matching

- Descriptor is the patch by itself
 - High dimensional data ($m \times n$)
 - Exhaustive evaluation is not feasible (Split criteria)
 - Random feature selection
 - Simple test (*Weak Learners*)



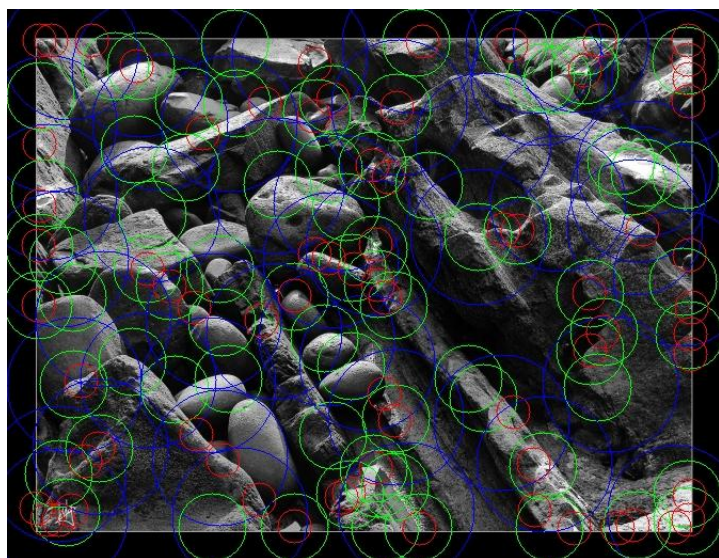
Supervised Feature Matching

- Building the training set.
 - Frontal view of the object to be detected.
 - Interest Point extraction.
 - Sub-images (*patches*) are extracted for each class.



Supervised Feature Matching

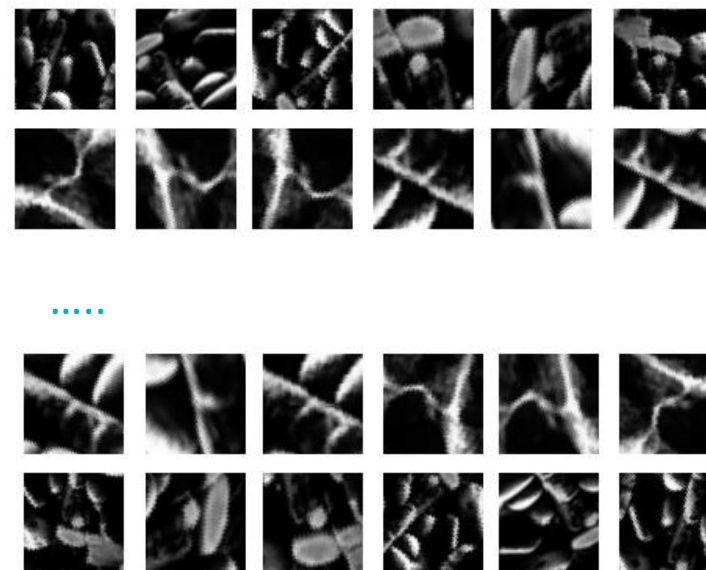
- Building the training set.
 - Generate Random Affine transformations.
 - Generate new examples of each Class.



Random
transformations

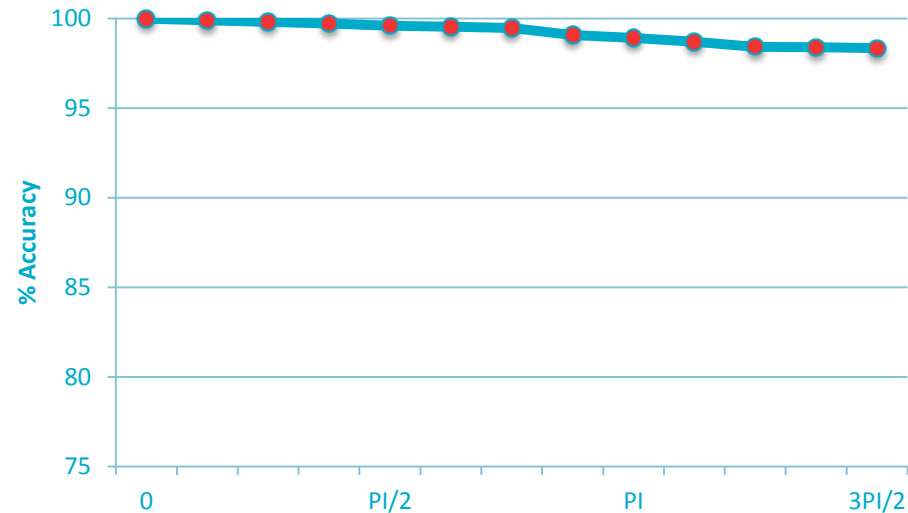


Training Set

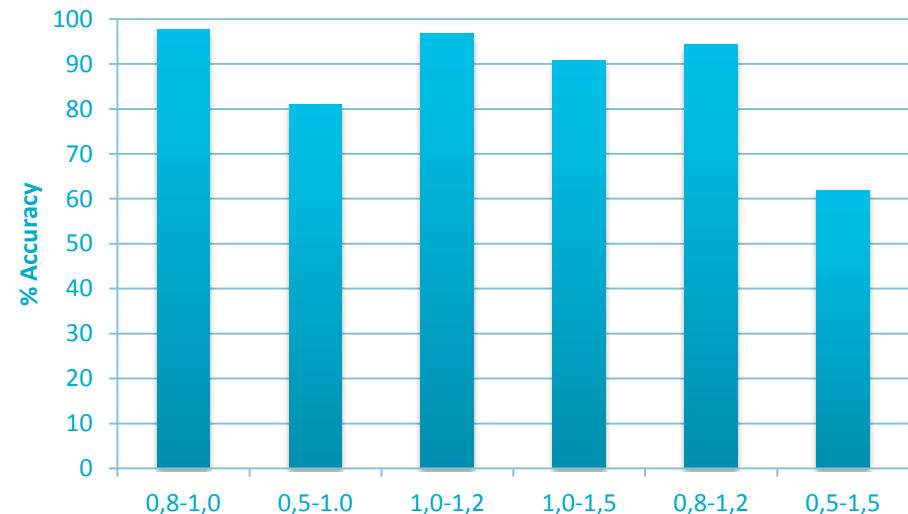


Supervised Feature Matching

- **Rotation Transformation**
 - 50 Trees, 15 Depth.
 - 225 Different Classes.
 - 400 Images per class.
 - No variation in scale
 - 10-fold cross validation



- **Scale Transformation**
 - 50 Trees, 15 Depth.
 - 225 Different Classes.
 - 400 Images per class.
 - Full rotation Range
 - 10-fold cross validation



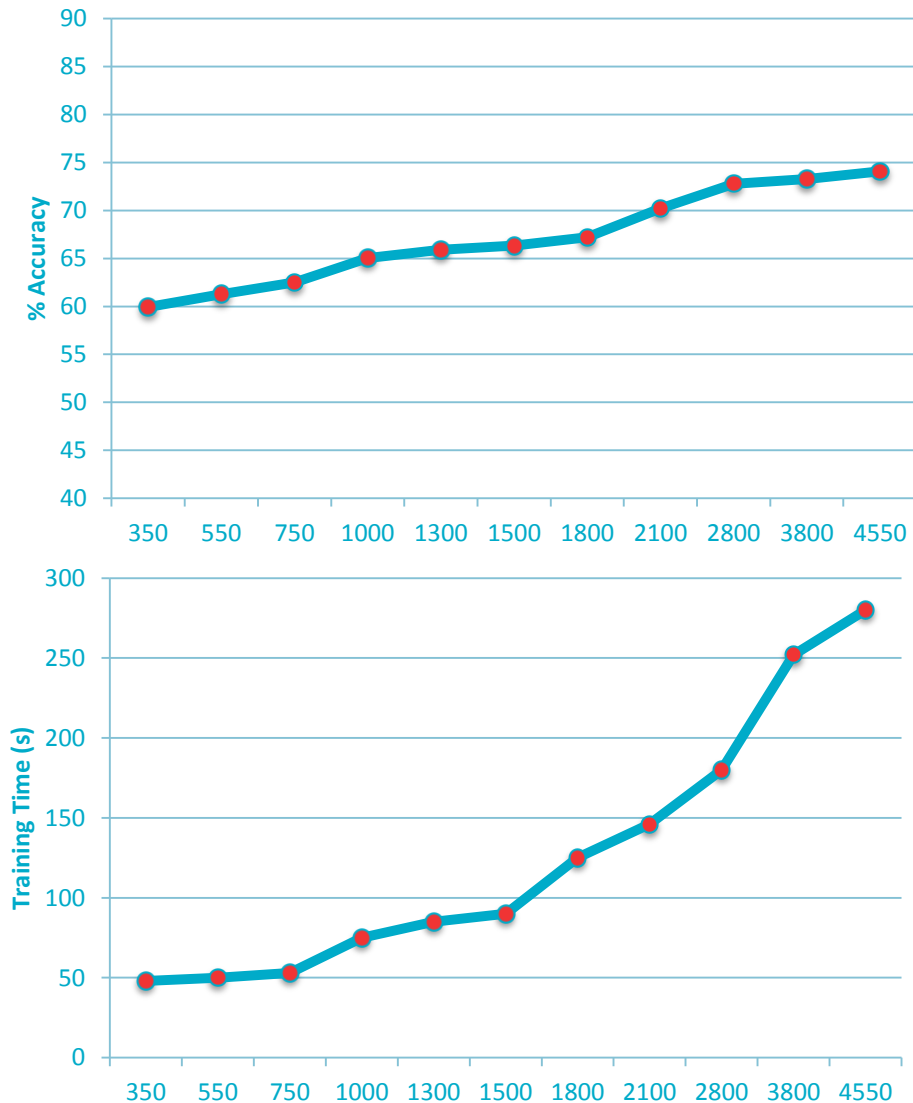
Supervised Feature Matching

- Size of the training Set

- 50 Trees, 15 Depth.
- 225 Different Classes.
- Full Rotation Range and [0.5-1.5] Scale Range.
- 10-fold cross validation

- Training Time

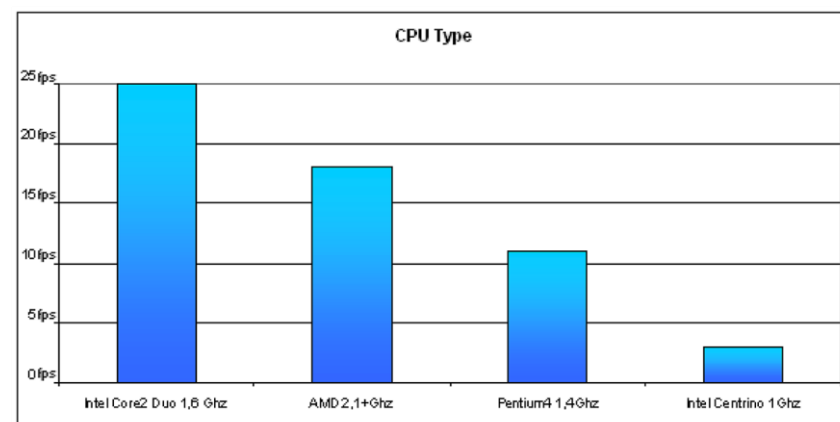
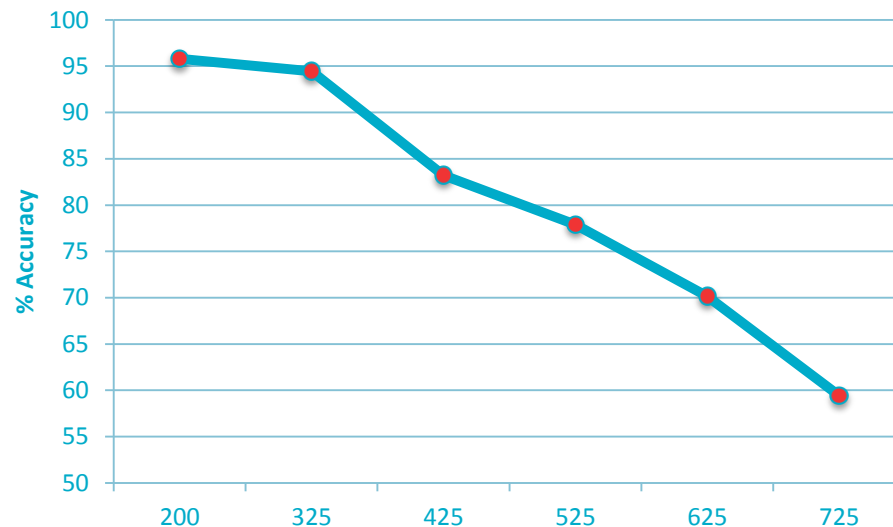
- 50 Trees, 15 Depth.
- Full Rotation Range and [0.5-1.5] Scale Range.
- 3000 Training images per class.
- 10-fold cross validation



Supervised Feature Matching

- **Number of different Classes.**
 - 50 Trees, 15 Depth.
 - Full Rotation Range and [0.8-1.2] Scale Range.
 - 3000 Training images per class.
 - 10-fold cross validation

- **Running Time**
 - 50 Trees, 15 Depth.
 - 225 Classes



Supervised Feature Matching

- The classifier is more sensitive to variations in scale.
- The classifier is robust against variations in object orientation.
- The higher the classification accuracy, the better the performance of the tracking.
- Limited number of interest points (classes).
- Need for off-line learning.
- Need memory resources for storing forest.
- Real-time performance.
- Highly paralelizable.



Summary

1. Introduction

2. Data and Evaluation Framework

3. Interest Point Extraction

4. Feature Description

5. Feature Matching

6. Conclusions



Conclusions and Results

- We have designed and implemented an evaluation Framework and a set of Ground truth Data.
- Regarding interest points
 - Review current state-of-the-art.
 - Evaluation of robustness and repeatability against photometric and geometric transformations, as well as computational costs.
- Regarding Feature descriptors
 - DITEC Descriptor.
 - Evaluation of state-of-the-art descriptors against photometric and geometric transformation.
- Regarding Supervised Matching
 - Good alternative to conventional point matching.
 - Limited number of points.
 - Limited applicability (known points, need of training phase).
- Several of these findings/approaches were applied in several research projects
 - Augmented reality.
 - Tracking .
 - Reconstruction.



Questions?

THANK YOU FOR
YOUR
ATTENTION

